Care-as-usual control groups across different settings in randomized trials on psychotherapy for adult depression: a meta-analysis

Pim Cuijpers[1] , Soledad Quero[2], Davide Papola[1,3], Ioana A. Cristea[4] and Eirini Karyotaki[1]

[1]Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health research institute, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands; [2]Department of Basic, Clinical Psychology, and Psychobiology, Universitat Jaume I, Castellón, Spain; [3]Department of Neuroscience, Biomedicine and Movement Science, Section of Psychiatry, WHO Collaborating Centre for Research and Training in Mental Health and Service Evaluation, University of Verona, Verona, Italy and [4]Department of Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy

Abstract

Background. Care-as-usual (CAU) is often used as a control condition in psychotherapy research, but it may vary considerably what that entails, ranging from no treatment, to routine treatment in primary care, general medical care, perinatal care, and specialized mental health care.

Methods. We conducted a meta-analysis of trials comparing psychotherapy for depression to CAU, with a focus on the different categories of CAU and countries where the studies were conducted. We used an existing database of randomized trials on psychotherapy for depression that is updated every year.

Results. A total of 140 studies with 15 419 patients were included. We found no significant differences in effects between categories of CAU (effect sizes ranging from $g = 0.43$ for CAU in primary care to $g = 0.73$ for no treatment), but heterogeneity was high in all CAU categories. After stratifying effects across specific countries (within CAU categories) we found that heterogeneity was considerably lower and there were several significant differences between countries. Overall, effects were larger in non-Western countries ($g = 0.84$ to 1.28) compared to those in Western countries ($g = 0.52$; p for difference = 0.002). Effects were smaller in studies with risk of bias ($p = 0.01$).

Conclusions. There are no significant differences between major categories of CAU when compared to psychotherapy conditions in randomized trials. However, effects of psychotherapy differ considerably across CAU conditions in specific countries. CAU therefore is a heterogeneous control condition in psychotherapy research.

Introduction

It is well-established that psychological therapies have significant effects on adult depression, and that is true for cognitive behavior therapy, interpersonal therapy, behavioral activation therapy, problem solving therapy, non-directive counseling, and psychodynamic therapies (Cuijpers, Karyotaki, de Wit, & Ebert, 2019b). Meta-analyses of direct and indirect comparisons between psychological therapies suggest that there are no significant differences between the effects of different types of therapies (Barth et al., 2013; Cuijpers, van Straten, Andersson, & van Oppen, 2008), that they can be delivered in several treatment formats (Cuijpers, Noma, Karyotaki, Cipriani, & Furukawa, 2019d), that short term effects are comparable to those of antidepressant medication (Cuijpers et al., 2020), but that long term effects may be better (Karyotaki et al., 2016), and that combined treatment is more effective than either psychotherapy or pharmacotherapy alone (Cuijpers et al., 2020).

The effect sizes that are found for psychological therapies, however, depend very much on the type of control group that is used to compare them with. In drug trials patients, clinicians, and other staff can be blinded for who gets the drug and who gets placebo. Because that is not possible in psychotherapy trials, researchers have to use other types of control groups when examining the effects of psychotherapies. Apart from comparisons of psychotherapy to other active treatments (such as another psychotherapy or a drug), several other types of control conditions are typically used in such randomized trials, including care-as-usual (CAU), waiting lists, pill placebo, and psychological placebo.

Each of these types of control groups has its own problems (Gold et al., 2017; Mohr et al., 2009). For example, in waiting list control groups improvement rates may be lower than natural recovery rates and these control groups may inflate the effect sizes of therapies (Cristea,

2017; Cuijpers, Karyotaki, Reijnders, & Ebert, 2019c; Furukawa et al., 2014; Mohr et al., 2009). Pill placebo can only be used when in the same trial active drugs are examined, otherwise participants know they get the placebo. Psychological placebos are also problematic, especially in depression, because non-directive counseling, which is typically used to control for factors that all therapies have in common, has been found to have considerable effects in depression (Cuijpers et al., 2012). It is not clearly established whether non-directive counseling is indeed less effective than other therapies.

CAU is one of the most credible control conditions in psychotherapy research, because it can indicate whether a new intervention has additional value above what is usually done in routine care. However, CAU also has its problems. One problem is the big heterogeneity across settings where the CAU is provided. CAU can be provided in specialized mental health care, where patients in the 'control' condition get specialized care from highly trained psychologists, psychiatrists, social workers, and often teams of such professionals. It can also be provided in primary care, where general practitioners usually treat patients. And although general practitioners are usually trained in handling mental disorders, their level of training cannot be compared to the specialized mental health clinicians. Another type of CAU is general medical care. Patients can be randomized to either a psychological treatment delivered in general medical settings or to the care they usually get in these settings, which typically does not mean care from clinicians specialized in delivering mental health care. No treatment can also be considered to be a specific type of CAU, because when someone is randomized to 'no treatment' this person can still find treatment elsewhere, outside the study (either with information about services provided in the trial, or without such information). Furthermore, CAU depends very much on the healthcare system of the country or the region where the study is conducted. In many high-income countries, CAU typically means that most patients have access to a range of treatment options, while in low and middle income (LAMI) countries, CAU often means no treatment at all, or only medication (Cuijpers, Karyotaki, Reijnders, Purgato, & Barbui, 2018).

Although randomized trials with CAU as the control group have been examined in many meta-analyses, the different types of CAU and the setting across countries has not yet been examined extensively. In one meta-analysis of psychotherapies for depression and anxiety (Watts, Turnell, Kladnitski, Newby, & Andrews, 2015), several subcategories of CAU were distinguished (primary care, specialized care). However, in this meta-analysis, the number of categories of CAU was limited, the number of studies was small (48 trials across depression and anxiety, while we could include 140 trials focused only on depression, see below), and country was not examined as a moderator. Furthermore, publication bias was not examined, nor were sensitivity analyses conducted (for example with studies with low risk of bias). This meta-analysis suggested that the effects of psychotherapy were smallest compared to CAU in primary care and largest when compared to CAU in which minimal treatment is given.

Another meta-analysis focused on differences between all types of control groups in psychotherapies for depression (Mohr et al., 2014). In this meta-analysis, no differences between subcategories of CAU were tested, nor was the country where the study was conducted included as a predictor of the outcome. Furthermore, only 34 trials with CAU control groups were included in this meta-analysis.

A third meta-analysis examining CAU control conditions in psychotherapy for depression did look at the type of CAU, as well as country where the study was conducted (Kolovos et al., 2017). However, this study only focused on change within the CAU conditions, which has been known to result in extremely large levels of heterogeneity (Cuijpers, Weitz, Cristea, & Twisk, 2017). Furthermore, only two types of CAU were distinguished (primary and specialized care) and the number of included trials was small (only 38). In this meta-analysis no significant difference was found between CAU in primary care and specialized care.

We decided to conduct a new meta-analysis of randomized trials examining the effects of psychotherapies for depression compared to different categories of CAU control groups, and to explore if these differ across countries. It is very well possible that CAU control groups differ depending on the setting where the trial is conducted, and depending on the organization of the health care system where the CAU is provided. This is important because these differences may be an important source of heterogeneity when assessing the effects of (psychological) treatments compared to CAU.

Methods

Identification and selection of studies

We used an existing database of studies on the psychological treatment of depression. This database has been described in detail elsewhere (Cuijpers, Karyotaki, & Ciharova, 2019a), and has been used in a series of earlier published meta-analyses (Cuijpers, 2017). The protocol for the current meta-analysis has been registered at the Open Science Foundation as part of the main meta-analytic project (https://osf.io/p8r52).

For the meta-analytic database we searched four major bibliographical databases (PubMed, PsycINFO, Embase, and the Cochrane Library) by combining terms (both index terms and text words) indicative of depression and psychotherapies, with filters for randomized controlled trials. The full search string for one database (PubMed) is given in online supplementary Appendix A. We also searched a number of bibliographical databases to identify trials in non-Western countries (Cuijpers et al., 2018), because the number of trials on psychological treatments in these countries is growing rapidly. Furthermore, we checked the references of earlier meta-analyses on psychological treatments of depression. The database is continuously updated and was developed through a comprehensive literature search (from 1966 to 1 January 2019). All records were screened by two independent researchers and all papers that could possibly meet inclusion criteria according to one of the researchers were retrieved as full-text. The decision to include or exclude a study in the database was also done by the two independent researchers, and disagreements were solved through discussion.

For the current meta-analysis, we included studies that were: (a) a randomized trial (b) in which a psychological treatment (c) for adults suffering from depression was (d) compared with a CAU control group. A diagnosis of depression could be established with a diagnostic interview or with a score above a cut-off on a self-report measure. Co-morbid mental or somatic disorders were not used as an exclusion criterion. Studies on inpatients were excluded, as were studies on children and adolescents. We also excluded maintenance studies, aimed at people who had already recovered or partly recovered after an earlier treatment.

We included the following types of CAU: (1) CAU in primary care, meaning that patients were recruited from primary care and receiving the usual care given in that context; (2) CAU in specialized mental health care; (3) CAU in perinatal care; (4) CAU in general medical care (in patients with comorbid general medical disorders); and (5) no treatment, meaning that they were not recruited from one specific setting, and that they did not receive any treatment in the context of the trial, but were allowed to seek treatment anywhere. In the case of no treatment, we did allow minimal support from the study, like sharing the results of the screening, advise to seek treatment elsewhere, information booklets, or one information session. Studies providing care as usual in other settings were excluded, for example Headstart (Beeber et al., 2010) or inmates (Eseadi, Obidoa, Ogbuabor, & Ikechukwu-Ilomuanya, 2018).

Quality assessment and data extraction

As in our previous meta-analyses using our database of randomized trials, we assessed the validity of included studies using four criteria of the 'Risk of bias' assessment tool, developed by the Cochrane Collaboration (Higgins, Altman, Gøtzsche, Jüni, Moher, Oxman, et al., 2011). This tool assesses possible sources of bias in randomized trials, including the adequate generation of allocation sequence; the concealment of allocation to conditions; the prevention of knowledge of the allocated intervention (masking of assessors); and dealing with incomplete outcome data (this was assessed as positive when intention-to-treat analyses were conducted, meaning that all randomized patients were included in the analyses). Assessment of the validity of the included studies was conducted by two independent researchers, and disagreements were solved through discussion.

We also coded participant characteristics (depressive disorder or scoring high on a self-rating scale; recruitment method; target group; proportion of women; mean age); characteristics of the psychotherapies (type; treatment format; number of sessions); and general characteristics of the studies (country where the study was conducted; year of publication).

We categorized the countries where the studies were conducted into low-, lower-middle, upper-middle, and high-income countries according to the definition of the World Bank (http://data.worldbank.org), for the year in which the study was published. We, arguably, considered Europe, North America, and Australia as Western countries and all other countries non-Western.

Type of treatment was defined according to the generic definitions of therapies given in Cuijpers et al. (2019b), and treatment format was coded as individual, group or guided self-help (including internet-based guided self-help; Cuijpers et al., 2019d).

Outcome measures

For each comparison between a psychotherapy and a CAU condition, the effect size indicating the difference between the two groups at the post-test was calculated (Hedges' g) (Hedges & Olkin, 1985). Effect sizes of 0.8 can be assumed to be large, while effect sizes of 0.5 are moderate, and effect sizes of 0.2 are small (Cohen, 1988). Effect sizes were calculated by subtracting (at post-test) the average score of the psychotherapy group from the average score of the control group, and dividing the result by the pooled standard deviation. Because some studies had relatively small sample sizes we corrected the effect size for small sample bias (Hedges & Olkin, 1985). If means and standard deviations were not reported, we converted dichotomous outcomes into effect sizes using the methods described by Borenstein and colleagues (Borenstein, Hedges, Higgins, & Rothstein, 2009). If dichotomous outcomes were not available either, we used other statistics (such as t value or p value) to calculate the effect size.

In order to calculate effect sizes, we used all measures examining depressive symptoms, such as the Beck Depression Inventory/BDI (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961); the BDI-II (Beck, Steer, & Brown, 1996); or the Hamilton Rating Scale for Depression/HAMD-17 (Hamilton, 1960). When more than one depression measure was used in a study, we pooled the outcomes within a study before pooling across studies, using the procedures described by Borenstein et al. (2009). All effect sizes were calculated in Comprehensive Meta-analysis (version 3.3.070).

Apart from the effect sizes, we also calculated the acceptability of the interventions as study drop-out for any reason. We calculated the relative risk (RR) of dropping out as the proportion of study drop-outs in the experimental group divided by the proportion of drop-outs in the comparison group. Drop-out in each arm was calculated as the number of people randomized minus the number of people for whom data were available at the post-test. The RRs for acceptability were calculated in R (see below).

Meta-analyses

To calculate pooled mean effect sizes, we used the 'meta' and 'metafor' packages in R, and ran all analyses in R studio (version 1.1.463 for Mac). Because we expected considerable heterogeneity among the studies, we employed a random effects pooling model in all analyses.

Numbers-needed-to-be-treated (NNT) were calculated using the formulae provided by Furukawa (1999), in which the control group's event rate was set at a conservative 19% (based on the pooled response rate of 50% reduction of symptoms across trials in psychotherapy for depression). As a test of homogeneity of effect sizes, we calculated the I^2-statistic and its 95% confidence interval, which is an indicator of heterogeneity in percentages. A value of 0% indicates no observed heterogeneity, and larger values indicate increasing heterogeneity, with 25% as low, 50% as moderate, and 75% as high heterogeneity (Higgins, Thompson, Deeks, & Altman, 2003).

We tested for publication bias by inspecting the funnel plot on primary outcome measures and by Duval and Tweedie's trim and fill procedure (Duval & Tweedie, 2000), which yields an estimate of the effect size after correction for the funnel plot asymmetry. We also conducted Egger's test of the intercept to quantify the bias captured by the funnel plot and to test whether it was significant. Studies with very large effect sizes ($g > 1.5$) were considered to be outliers and were excluded in sensitivity analyses.

In order to examine potential differences between the CAU categories as well as the effect sizes across different countries, we first conducted subgroup analyses. These analyses were conducted according to the mixed effects model as implemented in the Metafor package, in which effect sizes within subgroups are pooled according to the random effects model and the difference between subgroups according to a fixed effects model. We also conducted multivariate meta-regression analyses with the effect size as the dependent variable, and as predictors we used the five CAU categories (no treatment as reference category), the different countries, and the other major characteristics of the studies
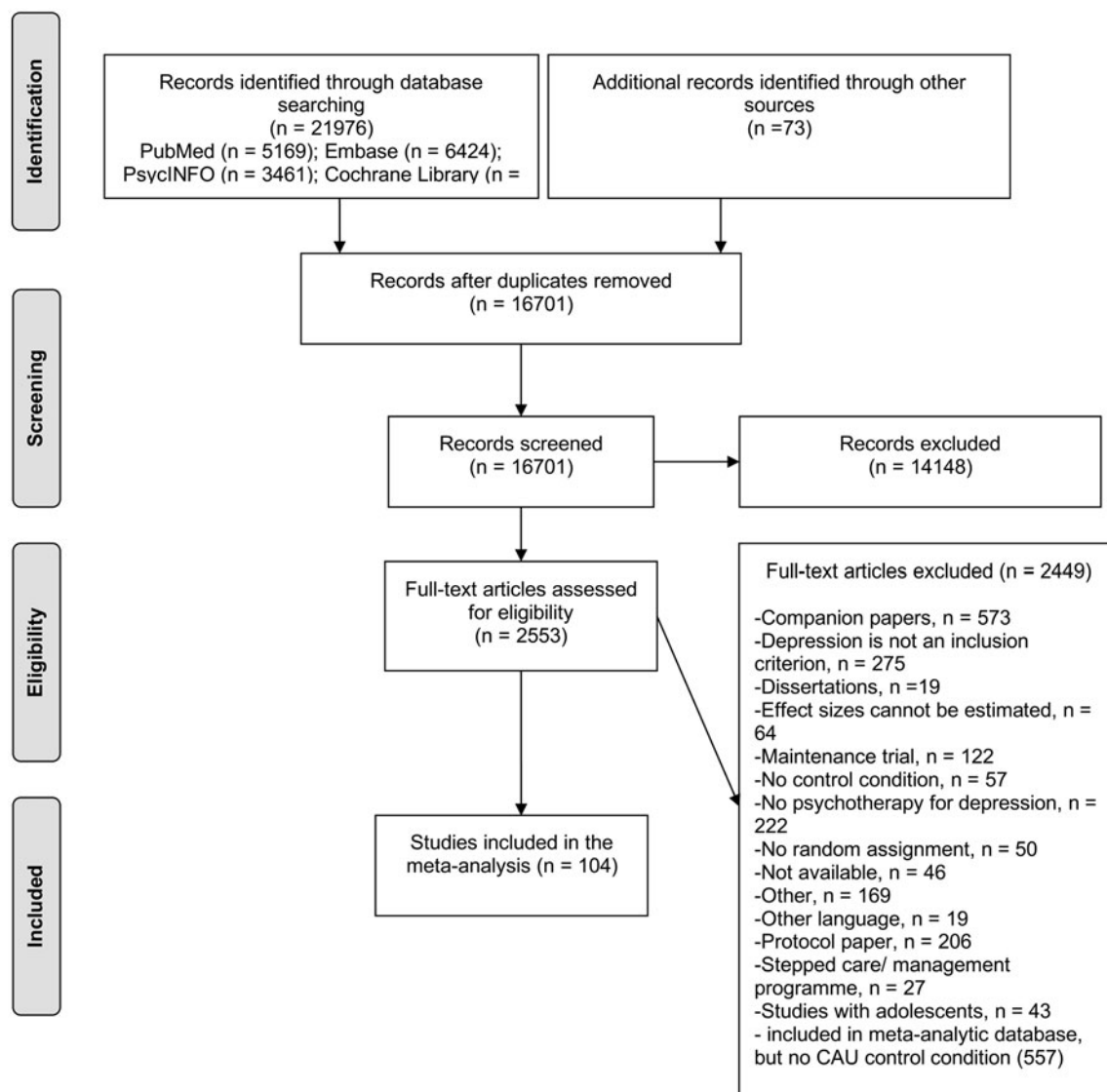
Fig. 1. Flowchart on the selection of studies.

as predictors. The method by Hartung and Knapp was used to adjust test statistics and confidence intervals.

The RRs indicating the acceptability of the interventions were pooled across studies, with the Hartung and Knapp to adjust test statistics and confidence intervals, and a value of 0.1 added for studies with a zero cell count.

Results

Selection and inclusion of studies

After examining a total of 21 976 abstracts (16 701 after removal of duplicates), we retrieved 2553 full-text papers for further consideration. We excluded 2449 of the retrieved papers. The PRISMA flowchart describing the inclusion process, including the reasons for exclusion, is presented in Fig. 1. A total of 140 randomized controlled trials (with 158 comparisons between a psychotherapy and a control group) met inclusion criteria for this meta-analysis. These studies included a total of 15 419 patients (8056 in the psychotherapy conditions and 7363 in the control

conditions). Selected characteristics of the included studies are given in online supplementary Appendix B and the references are given in online supplementary Appendix C.

Characteristics of included studies

Of the 140 studies, CAU was delivered in primary care in 34 studies (24.3%), 26 in specialized mental health care (18.6%), 36 in general medical care (25.7%), 20 in perinatal care (14.3%), and in 24 studies (17.1%) CAU was categorized as no treatment. Most studies (76.4%) were conducted in Western, high-income countries (36 in the US, 27 in the UK, 32 in Europe, 11 in Australia, and 1 in Canada), 7 were conducted in non-Western high-income countries, 17 in upper-middle, 8 in lower-middle income countries, and 1 in a low-income country.

A total of 23 studies (16.4%) were conducted between 1981 and 2005, 26 between 2006 and 2010 (18.6%), 59 between 2011 and 2015 (42.1%), and 32 between 2016 and 2018 (22.9%).

In 93 studies (58.9%) patients met criteria for a depressive disorder according to a diagnostic interview, while in the

remaining 65 studies (41.1%) they scored above a cut-off on a self-report measure. Of the 140 studies, 111 (79.3%) were aimed at adults in general, while 29 (20.7%) were aimed at older adults.

In 88 of the 158 comparisons between a treatment and a CAU condition, cognitive behavior therapy (CBT) was used as the intervention (55.7%), 17 used interpersonal psychotherapy (10.8%), 13 non-directive counseling (8.2%), 8 psychodynamic therapy (5.1%), 8 behavioral activation therapy (5.1%), and the remaining 24 comparisons used another therapy (17.1%). In 80 comparisons an individual treatment format was used (50.6%), 43 used a group format (27.2%), while the remaining comparisons used another format (guided self-help, telephone, mixed; 22.2%). A total of 49 comparisons (31.0%) had less than 6 sessions, 70 had 7 to 11 sessions (44.3%), and 39 comparisons had more than 12 sessions (24.7%).

Risk of bias

The risk of bias in most studies was considerable. A total of 81 of the 140 studies reported an adequate sequence generation (57.9%). Seventy-eight studies reported properly concealed allocation (55.7%). Forty-one studies (29.3%) reported using blinded outcome assessors, and 92 (65.7%) used only self-report outcomes. In 92 (65.7%) studies intent-to-treat analyses were conducted. Only 48 studies (34.3%) met all quality criteria, 68 (48.6%) met two or three of the criteria, and the 24 remaining studies (17.1%) met no or only one criterion.

Overall effects of psychotherapies compared with CAU

The overall effect size of all 158 comparisons between psychotherapy and CAU, across all CAU categories, was $g = 0.61$ (95% CI 0.52–0.70) with high heterogeneity ($I^2 = 78\%$; 95% CI 74–81). This effect size corresponds with an NNT of 4.89. The results are presented in Table 1.

In 15 studies more than one psychotherapy arm was included (13 had two psychotherapy arms, and 2 had three arms). Because these effect sizes were not independent of each other, they may have artificially reduced heterogeneity and influenced the effect sizes. We conducted two sensitivity analyses to examine this. In the first analysis we only included the largest effect size from each study, and in the second only the smallest effect size. As can be seen in Table 1, the effect sizes and levels of heterogeneity were comparable to those in the main analyses. We excluded outliers with an effect size of $g = 1.5$ or higher, and found that the effect size was somewhat smaller ($g = 0.48$; 95% CI 0.42–0.54; NNT = 6.44), but heterogeneity was still moderate to large ($I^2 = 65\%$; 95% CI 58–71). When we limited the analyses to the 55 comparisons with low risk of bias, the effect size was moderate ($g = 0.51$; 95% CI 0.35–0.67; NNT = 6.01) with high heterogeneity ($I^2 = 73\%$; 95% CI 64–79).

We found strong indications for publication bias. Egger's test of the asymmetry of the funnel plot was highly significant ($p < 0.001$) and Duval and Tweedie's trim and fill procedure resulted in an adjusted effect size of $g = 0.44$ (95% CI 0.32–0.55; NNT = 7.11; $I^2 = 84\%$, 95% CI 82–86), with 25 imputed studies.

The effects of psychotherapies in different CAU categories and across countries

We found no significant differences between the different CAU in primary care, general medical care, specialized mental health care, perinatal care, and no treatment ($p = 0.21$; Table 1). Heterogeneity was high in all CAU subcategories (I^2: 72 to 80%). We did find a significant difference between CAU (all subcategories together) in Western high-income countries, non-Western high-income countries, and low- and middle-income countries ($p = 0.002$). The effects in Western countries were significantly smaller than in other countries. Heterogeneity was again high in all subgroups (I^2: 70 to 92%).

Within each CAU category we selected specific countries, if there were at least three comparisons between psychotherapy and CAU. As can be seen in Table 1, We found that the effect sizes significantly differed between countries for no treatment (Netherlands, Spain, UK, US; $p < 0.001$), for primary care (Netherlands, UK, US; $p < 0.001$), for general medical care (Germany, UK, US; $p = 0.03$), for specialized mental health care (Netherlands, US; $p < 0.001$), but not for perinatal care (Australia, China, UK, US; $p = 0.84$). Heterogeneity was low to moderate in most subgroups (except in primary care in the US, and across three countries in perinatal care, where it was still high).

Subgroup and meta-regression analyses

In order to examine whether potential differential effects of therapies compared to CAU were caused by other characteristics of the participants, interventions, and studies we first ran a series of subgroup analyses to examine differences in effects between subgroups of studies. The results are presented in Table 2. As can be seen, none of the examined subgroups differed significantly from each other, except for studies with higher risk of bias showing larger effect sizes than those with low risk of bias ($p = 0.01$).

We ran two separate multivariate meta-regression models. In the first model, we included the CAU categories (no treatment as reference category), the income level of the country (Western high-income countries as reference category), and the other characteristics of the participants, therapies, and studies as predictors. In the second model, we used the same predictors, except that we included the specific countries as predictor instead of the income level (only countries with 10 or more comparisons, and the US as the reference category). The results are presented in Table 3. As can be seen, we found only one significant predictor in the first model. The effect sizes found in high-income non-Western countries were larger than in other countries. None of the other predictors were significant. In the second model, none of the predictors was significant.

Acceptability

We found no significant difference between therapies and CAU for acceptability (RR = 1.07; 95% CI 0.96–1.20; $I^2 = 44$; 95% CI 31–54; Table 4). We also found no significant difference across CAU categories ($p = 0.33$), or between the income levels of the countries ($p = 0.45$). We did find that a significant effect for acceptability of usual general medical care (RR = 1.25; 95% CI 1.01–1.55; $I^2 = 17$; 95% CI 0–46), indicating that acceptability of usual general medical care was higher than the acceptability of the psychotherapy conditions.

Table 1. Effects of psychotherapies compared with CAU control groups, across categories of CAU, and across income level of the country where the study was conducted: Hedges' g^a

	N_{st}	g	95% CI	I^2	95% CI	p^b	NNT
All comparisons	158	0.61	0.52–0.70	78	74–81		4.89
One ES per study (highest)	140	0.62	0.53–0.72	79	76–82		4.80
One ES per study (lowest)	140	0.60	0.50–0.70	79	76–82		4.99
Outliers excluded ($g \geqslant 1.5$)	144	0.48	0.42–0.54	65	58–71		6.44
Adjusted for publication bias[c]	183	0.44	0.32–0.55	84	82–86		7.11
Only low risk of bias	55	0.51	0.35–0.67	73	64–79		6.01
CAU categories						0.21	
No treatment	29	0.73	0.51–0.94	80	71–86		3.98
Primary care	37	0.46	0.31–0.60	72	61–80		6.76
General medical care	38	0.58	0.42–0.74	74	64–81		5.18
Perinatal care	25	0.69	0.41–0.96	79	70–86		4.25
Specialized mental health	29	0.68	0.36–0.99	80	72–86		4.32
Income level countries							
High income, Western	123	0.52	0.42–0.61	70	64–75	0.002	5.88
High income, non-Western	7	1.28	0.02–2.54	92	87–96		1.58
Upper middle	19	0.84	0.58–1.09	82	73–88		3.39
Low/lower middle	9	0.85	0.62–1.08	76	53–87		3.35
No treatment							
Netherlands	3	0.42	−0.01–0.85	25	0–92	<0.001	7.49
Spain	4	1.62	0.74–2.50	56	0–86		1.34
UK	4	0.38	−0.18–0.95	0	0–84		8.38
US	8	0.62	0.18–1.05	57	6–81		4.80
Primary care						<0.001	–
Netherlands	3	0.05	−0.22–0.32	0	0–69		72.13
UK	16	0.39	0.12–0.54	19	0–55		8.14
US	5	0.95	0.16–1.74	87	73–94		2.95
General medical care						0.03	–
Germany	4	0.19	−0.17–0.55	57	0–86		17.93
UK	5	0.30	−0.24–0.84	41	0–78		10.91
US	16	0.55	0.37–0.73	44	0–69		5.51
Perinatal care						0.84	
Australia	8	0.57	0.37–0.77	0	0–27		5.29
China	4	0.65	0.12–1.17	92	83–96		4.55
UK	6	0.97	−0.39–2.32	90	81–95		2.88
US	3	0.39	−1.64–2.43	87	62–95		8.14
Specialized mental health						<0.001	–
Netherlands	3	0.80	0.11–1.50	39	0–81		3.58
US	7	0.21	0.08–0.33	0	0–3		16.10

[a]According to the random effects model.
[b]p value indicates the significance of the difference between subgroups.
[c]Egger's test was significant ($p < 0.001$) and the number of imputed studies using Duval and Tweedie's trim and fill procedure was 25.

We also examined acceptability of therapies compared to the CAU categories across specific countries (we examined the same countries for which we also examined the effect sizes). The results are given in Table 4. We found no significant differences between countries for primary, general medical, perinatal, and specialized mental health care. But we did find significant

Table 2. Effects of psychotherapies compared with CAU control groups, across categories of CAU, and across income level of the country where the study was conducted: Hedges' g^a

	N_{st}	g	95% CI	I^2	95% CI	p^b	NNT
Depression							
Diagnosis	93	0.61	0.49–0.74	76	70–80	0.97	4.89
Cut-off	65	0.61	0.47–0.75	81	76–85		4.89
Type of therapy							
CBT	88	0.66	0.52–0.80	79	74–83	0.07	4.47
IPT	17	0.45	0.21–0.69	79	66–86		6.93
Supportive	13	0.43	0.28–0.58	0	0–37		7.29
Other	40	0.63	0.46–0.79	82	76–86		4.71
Format							
Individual	80	0.51	0.40–0.62	60	49–69	0.12	6.01
Group	43	0.80	0.58–1.02	84	79–88		3.58
GSH	11	0.47	0.21–0.73	76	57–87		6.59
Telephone	11	0.56	0.19–0.94	79	64–88		5.40
Other	13	0.75	0.38–1.11	90	85–94		3.86
Number of sessions							
6 or less	49	0.57	0.42–0.72	76	68–81	0.82	5.29
7 to 11	70	0.63	0.49–0.76	73	66–79		4.71
12 or more	39	0.63	0.40–0.86	85	80–88		4.71
Risk of Bias							
0–1 (high)	26	0.90	0.68–1.12	64	46–76	*0.01*	3.14
2–3	77	0.60	0.47–0.73	81	76–84		4.99
4 (low)	55	0.51	0.35–0.67	73	64–79		6.01
Year							
<2000	16	0.55	0.35–0.75	43	0–69	0.77	5.51
2001–2005	15	0.48	0.26–0.70	77	62–86		6.44
2006–2010	28	0.59	0.41–0.77	63	44–75		5.08
2011–2015	64	0.66	0.48–0.83	84	80–87		4.47
2016–2018	35	0.61	0.40–0.82	79	71–85		4.89

[a]According to the random effects model.
[b]The *p*-value in this column refers to the difference between subgroups.

differences between countries for no treatment ($p < 0.001$). However, the number of studies within each country was small, so these results should be interpreted with caution.

For specific countries we found that acceptability of CAU was significantly higher in Spain for no treatment, lower in the UK for no treatment, and higher in the US for general medical care.

Discussion

The goal of this meta-analysis was to examine whether the effects of psychotherapies for depression varied across different categories of CAU and to explore whether these categories differed across countries. We did not find that the effects of psychotherapies differed significantly across categories of CAU, including CAU in primary care, perinatal care, general medical care, specialized mental health care, and no treatment. This is surprising, because the different CAU categories differ so strongly from each other, ranging from no treatment to regular specialized mental health care. However, heterogeneity was very high in each of the CAU categories indicating considerable variations between effect sizes within each category.

We also found that when we examined differences between countries within each of the CAU categories, heterogeneity was considerably lower and there were several significant differences between countries. This suggests that effects of psychotherapy could differ considerably across countries because of differences between CAU. We consider the latter as a reasonable conclusion given that primary care in the US is not the same as primary care in the UK or in LAMI countries. This finding may be a partial explanation for the absence of differences between the CAU categories and the high heterogeneity within each of these categories. This implies that the CAU conditions that were included in

Table 3. Standardized regression coefficients of effect sizes across different CAU categories: multivariate metaregression analyses

	Model 1: income level			Model 2: countries		
	Coeff.	S.E.	p	Coeff.	S.E.	p
CAU category						
No treatment	Ref.			Ref.		
GP	−0.18	0.16	0.26	−0.15	0.16	0.36
General medical	−0.06	0.15	0.70	−0.05	0.16	0.75
PPD	−0.04	0.17	0.82	−0.03	0.18	0.88
Specialized MH	−0.09	0.17	0.59	−0.07	0.17	0.68
Western (y/n)						
High income, Western	Ref.					
High income, non-Western	0.57	0.23	*0.01*			
Upper middle	0.23	0.16	0.15			
Low/lower middle	0.34	0.20	0.09			
Country						
US				Ref.		
UK				0.02	0.17	0.91
Australia				0.08	0.21	0.72
Netherlands				−0.14	0.21	0.49
Other country				0.13	0.13	0.34
Diagnosis (y/n)	0.11	0.10	0.27	0.10	0.11	0.36
Type						
CBT	Ref.			Ref.		
IPT	−0.24	0.16	0.13	−0.24	0.17	0.16
Supportive	−0.14	0.19	0.47	−0.14	0.20	0.48
Other	0.04	0.12	0.74	0.05	0.12	0.71
Format						
Individual	Ref.			Ref.		
Group	0.11	0.13	0.41	0.15	0.13	0.25
Other/mixed	0.05	0.12	0.67	0.04	0.13	0.75
N sessions	−0.00	0.01	0.66	−0.00	0.01	0.69
Year	−0.00	0.01	0.76	−0.00	0.01	0.98
Risk of bias	−0.08	0.05	0.13	−0.09	0.05	0.09
Intercept	6.25	17.84	0.73	1.38	19.05	0.94

this meta-analysis are not necessarily comparable across the different categories. Although the effects were comparable across categories, heterogeneity within each category was high and was only acceptably low when we looked at CAU within countries and within categories.

Overall, we found that the effects of psychotherapies were larger in non-Western countries than in Western countries. That is in line with a previous meta-analysis comparing the effect sizes of psychotherapy for depression in Western and non-Western countries (Cuijpers et al., 2018). There are several explanations for this finding. It is possible that these therapies simply work better in (some) non-Western countries, but it is not clear why that would be the case. Another explanation could be that CAU

in these countries means that patients get no treatment at all, while in Western countries CAU implies that patients have access to several treatments. A third explanation could be that the quality of the studies conducted in non-Western countries was not optimal.

The findings of this meta-analysis have several implications for future research. As indicated in the Introduction, each type of control condition in psychotherapy research has its own problems (Gold et al., 2017; Mohr et al., 2009). In the current study we could confirm that CAU is a heterogeneous control condition and when the effects of therapies are compared to CAU high levels of heterogeneity should be expected. It also implies that effects of therapy found in one country may not be comparable

Table 4. Acceptability of psychotherapies compared with CAU control groups, across categories of CAU, and across income level of the country where the study was conducted: relative risks[a]

	N_{st}	RR	95% CI	I^2	95% CI	p[b]
All comparisons	135	1.07	0.96–1.20	44	31–54	
One ES per study (highest)	119	1.08	0.95–1.22	46	34–57	
One ES per study (lowest)	119	1.10	0.98–1.25	43	28–54	
Adjusted for publication bias[c]	141	1.06	0.95–1.20	43	30–53	
Only low risk of bias	50	1.10	0.93–1.30	26	0–48	
CAU categories						0.33
No treatment	22	1.06	0.81–1.38	32	0–60	
Primary care	35	1.03	0.85–1.26	40	9–60	
General medical care	34	*1.25*	1.01–1.55	17	0–46	
Perinatal care	19	1.21	0.74–1.97	59	32–75	
Specialized mental health	25	0.90	0.69–1.16	28	0–56	
Income level countries						0.45
High income, Western	106	1.06	0.94–1.20	49	36–59	
High income, non-Western	7	1.08	0.50–2.34	0	0–51	
Upper middle	14	1.50	0.86–2.61	35	0–66	
Low/lower middle	8	0.88	0.55–1.41	17	0–60	
No treatment						*<0.001*
Netherlands	2	2.31	0.23–23.58	0	[a]	
Spain	2	*2.60*	1.24–5.45	0	[a]	
UK	4	*0.64*	0.41–1.00	0	0–60	
US	5	0.86	0.57–1.32	0	0–67	
Primary care						0.37
Netherlands	3	1.52	0.37–6.14	44	0–83	
UK	16	0.93	0.72–1.21	20	0–56	
US	4	1.00	0.43–2.29	48	0–83	
General medical care						0.51
Germany	3	1.17	0.31–4.42	86	58–95	
UK	5	1.09	0.38–3.12	7	0–81	
US	16	*1.56*	1.16–2.12	0	0–22	
Perinatal care						0.18
Australia	8	0.85	0.49–1.49	0	0–64	
China	3	1.10	0.05–23.29	78	28–93	
UK	4	1.74	0.73–4.11	0	0–78	
US	2	4.45	0–>1000	82	[a]	
Specialized mental health						0.65
Netherlands	3	0.49	0.05–5.15	77	24–93	
US	6	0.63	0.39–1.01	0	0–64	

[a]According to the random effects model.
[b]*p* value indicates the significance of the difference between subgroups.
[c]Egger's test was not significant (*p* < 0.84) and the number of imputed studies using Duval and Tweedie's trim and fill procedure was 6.

to those found in other countries, and meta-analyses should preferable be within one country and one setting. That requires, however, large numbers of studies and considerable resources. Both researchers and clinicians should be aware of these differences and should interpret the findings obtained in another country with caution when translating them to their own country. To our knowledge this is the first meta-analysis examining the effects of psychotherapy for depression in which a comprehensive

operationalization of CAU is conducted and differences across countries are examined. However, the results of this study have to be considered with caution because of several limitations. One important limitation is that although the total number of studies was large, the number per CAU category was limited, and certainly the number of studies from separate countries was small. Therefore, the confidence intervals around the effect sizes and levels of heterogeneity were broad, resulting in considerable uncertainty of the findings. Furthermore, the majority of studies had at least some risk of bias and the number of studies with low risk of bias across all domains was small. Another limitation is that we assumed that the effects of therapies were comparable across different types of therapies, treatment formats, and the number of sessions. Although this is what previous research suggests to be true (Barth et al., 2013; Cuijpers et al., 2019a, 2019b, 2019c), this is still an assumption that may not or only partially be true. We did find, however, that heterogeneity was relatively low within CAU categories in specific countries, suggesting that the effects of the therapies are indeed comparable, but may differ because the CAU differs across categories and across countries.

Another important problem of this meta-analysis is the strong influence of publication bias across the set of included studies. Although we used only statistical methods to assess publication bias, it has been validated in direct research that publication bias is indeed a real problem with considerable impact on the overall effect size of psychotherapies for depression (Driessen, Hollon, Bockting, Cuijpers, & Turner, 2015). In the current study we did not have the means to further explore the impact of publication bias, but this is certainly a limitation that should be taken into consideration when interpreting the findings.

Despite these limitations we can conclude that overall there are no significant differences between the effects of psychotherapies across main categories of CAU, but there are significant differences across categories between specific countries. We also could confirm that psychotherapies are more effective compared to CAU in non-Western countries than in Western countries.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0033291719003581

References

Barth, J., Munder, T., Gerger, H., Nuesch, E., Trelle, S., Znoj, H., … Cuijpers, P. (2013). Comparative efficacy of seven psychotherapeutic interventions for depressed patients: A network meta-analysis. *Plos Medicine*, 10, e1001454.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *BDI-II. Beck depression inventory second edition*. Manual. San Antonio: Psychological Corporation.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561–571.

Beeber, L. S., Holditch-Davis, D., Perreira, K., Schwartz, T. A., Lewis, V., Blanchard, H., & Goldman, B. D. (2010). Short-term in-home intervention reduces depressive symptoms in Early Head Start Latina mothers of infants and toddlers. *Research on Nursing and Health*, 33, 60–76.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale: Erlbaum.

Cristea, I. A. (2017). The waiting list is an inadequate benchmark for estimating the effectiveness of psychotherapy for depression. *Epidemiology and Psychiatric Science*, 28, 278–279.

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

Cuijpers, P., Driessen, E., Hollon, S. D., van Oppen, P., Barth, J., & Andersson, G. (2012). The efficacy of non-directive supportive therapy for adult depression: A meta-analysis. *Clinical Psychology Review*, 32, 280–291.

Cuijpers, P., Karyotaki, E., & Ciharova, M. (2019a). A meta-analytic database of randomised trials on psychotherapies for depression. *Open Science Foundation*. https://doi.org/10.17605/OSF.IO/825C6.

Cuijpers, P., Karyotaki, E., de Wit, L., & Ebert, D. D. (2019b). The effects of fifteen evidence-supported therapies for adult depression: A meta-analytic review. *Psychotherapy Research*. epub ahead of print.

Cuijpers, P., Karyotaki, E., Reijnders, M., & Ebert, D. D. (2019c). Was Eysenck right after all? A reassessment of the effects of psychotherapy for adult depression. *Epidemiology and Psychiatric Sciences*, 28, 21–30.

Cuijpers, P., Karyotaki, E., Reijnders, M., Purgato, M., & Barbui, C. (2018). Psychotherapies for depression in low- and middle-income countries: A meta-analysis. *World Psychiatry*, 17, 90–101.

Cuijpers, P., Noma, H., Karyotaki, E., Cipriani, A., & Furukawa, T. (2019d). Individual, group, telephone, self-help and internet-based cognitive behavior therapy for adult depression: A network meta-analysis of delivery methods. *JAMA Psychiatry*, 76, 700–707.

Cuijpers, P., Noma, H., Karyotaki, E., Vinkers, C. H., Cipriani, A., & Furukawa, T. A. (2020). A network meta-analysis of the effects of psychotherapies, pharmacotherapies and their combination in the treatment of adult depression. *World Psychiatry*, in press.

Cuijpers, P., van Straten, A., Andersson, G., & van Oppen, P. (2008). Psychotherapy for depression in adults: A meta-analysis of comparative outcome studies. *Journal of Consulting and Clinical Psychology*, 76, 909–922.

Cuijpers, P., Weitz, E., Cristea, I., & Twisk, J. (2017). Pre-post effect sizes should be avoided in meta-analyses. *Epidemiology and Psychiatric Sciences*, 26, 364–368.

Driessen, E., Hollon, S. D., Bockting, C. L. H., Cuijpers, P., & Turner, E. H. (2015). Does publication bias inflate the apparent efficacy of psychological treatment for major depressive disorder? A systematic review and meta-analysis of us national institutes of health-funded trials. *Plos One*, 10, e0137864.

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, 455–463.

Eseadi, C., Obidoa, M. A., Ogbuabor, S. E., & Ikechukwu-Ilomuanya, A. B. (2018). Effects of group-focused cognitive-behavioral coaching program on depressive symptoms in a sample of inmates in a Nigerian prison. *International Journal of Offender Therapy and Comparative Criminology*, 62, 1589–1602.

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

Furukawa, T. A., Noma, H., Caldwell, D. M., Honyashiki, M., Shinohara, K., Imai, H., … Churchill, R. (2014). Waiting list may be a nocebo condition in psychotherapy trials: A contribution from network meta-analysis. *Acta Psychiatrica Scandinavica*, 130, 181–192.

Gold, S. M., Enck, P., Hasselmann, H., Friede, T., Hegerl, U., Mohr, D. C., & Otte, C. (2017). Control conditions for randomised trials of behavioural interventions in psychiatry: A decision framework. *The Lancet. Psychiatry*, 4, 725–732.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry*, 23, 56–62.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic 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. *British Medical Journal*, 343, d5928.

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

Karyotaki, E., Smit, Y., Holdt Henningsen, K., Huibers, M. J. H., Robayse, J., de Beurs, D., & Cuijpers, P. (2016). Combining pharmacotherapy and psychotherapy or monotherapy for major depression? A meta-

analysis on the long-term effects. *Journal of Affective Disorders*, *194*, 144–152.

Kolovos, S., van Tulder, M. W., Cuijpers, P., Prigent, A., Chevreul, K., Riper, H., & Bosmans, J. E. (2017). The effect of treatment as usual on major depressive disorder: A meta-analysis. *Journal of Affective Disorders*, *210*, 72–81.

Mohr, D. C., Ho, J., Hart, S. L., Baron, K. G., Berendsen, M., Beckner, V., … Duffecy, J. (2014). Control condition design and implementation features in controlled trials: A meta-analysis of trials evaluating psychotherapy

for depression. *Translational Behavioral Medicine: Practice, Policy and Research*, *4*, 407–423.

Mohr, D. C., Spring, B., Freedland, K. E., Beckner, V., Arean, P., Hollon, S. D., … Kaplan, R. (2009). The selection and design of control conditions for randomized controlled trials of psychological interventions. *Psychotherapy and Psychosomatics*, *78*, 275–284.

Watts, S. E., Turnell, A., Kladnitski, N., Newby, J. M., & Andrews, G. (2015). Treatment-as-usual (TAU) is anything but usual: A meta-analysis of CBT versus TAU for anxiety and depression. *Journal of Affective Disorders*, *175*, 152–167.