Revisiting Child and Adolescent Depression Treatments

A Quantitative Critique of the Evidence Base

Argyris Stringaris

Charlotte Burman

Rudolf Uher

Dayna Bhudia

Despoina Miliou

Ioannis-Marios Rokas

Marinos Kyriakopoulos

Lucy Foulkes

Carmen Moreno

Samuele Cortese

Georgina Krebs

2024-05-15

# Abstract

**Question:** Psychotherapy and medication are the main treatments for adolescent depression, with guidelines recommending psychotherapy as the first-line treatment for most cases. Given the scarcity of head-to-head comparisons, recommendations are based on indirect comparisons of outcomes obtained in randomised controlled trials (RCTs) within each modality. Here we examine whether basic conditions required to draw valid inferences from these indirect comparisons are met. **Study selection and analysis:** Drawing upon studies included in two recent meta-analyses, and supplementing with an original systematic search, 88 RCTs of psychotherapy and medication for child and adolescent depression (mean age 4-18 years) were identified. Using meta-analytic approaches, we examined comparability of a) participant characteristics and b) trial characteristics in medication and psychotherapy trials. Lastly, we examined whether psychotherapy controls are well-matched to active conditions. **Findings:** Participants in medication RCTs had higher depression severity and were more likely to be male compared to psychotherapy RCTs. The pooled effect size of medication controls was very large (SMD=-1.9), and was medium for psychotherapy controls (SMD=-0.5). Medication RCTs also had significantly more trial sites. Within psychotherapy RCTs, control conditions were less intensive on average than active conditions (e.g. fewer contact hours). **Conclusions:** Medication and psychotherapy RCTs differ on both participant and methodological characteristics, thereby violating the conditions for valid comparison between modalities. Psychotherapy controls often involve little therapist contact and are easy-to-beat comparators. These findings cast doubt on the confidence with which psychotherapy is recommended over medication for youth depression, and highlight the pressing need to improve the evidence base.

# Key messages

Psychotherapy is recommended before medication for most cases of depression in children and adolescents, a recommendation that is based on indirect comparisons of outcomes from randomised controlled trials (RCTs) within each treatment modality. We examine the validity of these inferences by scrutinising the comparability of psychotherapy and medication RCTs. We find significant differences in sample characteristics (namely depression severity and sex composition) and trial design features, such that the effect sizes of medication controls (i.e. pill placebo) are much larger than those for psychotherapy controls, and that medication RCTs feature significantly more trial sites. We also examine the quality of controls used in psychotherapy RCTs and find that they are poorly matched to active intervention arms in ways such as human contact hours, and hence represent poor and easy to beat comparators. Our findings underscore the need for a higher quality evidence base upon which to base treatment guidelines and clinical decision making.

# Background

Should a child or an adolescent receive psychotherapy or medication for their depression, and what information should be used to guide decision-making?

For adolescent depression, there are limited data from head-to-head trials of medication and psychotherapy, and hence recommendations must be derived from indirect comparisons of treatment efficacy. The UK National Institute of Health and Care Excellence (NICE) guidelines for adolescent depression recommend that youth are first offered psychological therapy (specifically cognitive behaviour therapy (CBT) or interpersonal therapy (IPT)) over medication in most cases (NICE 2019). This is in keeping with two sources of evidence relating to child and adolescent depression: meta-analyses of medication randomised controlled trials (RCTs) that cast doubt on the efficacy of antidepressants, with the exception of fluoxetine (Cipriani et al. 2016); and meta-analyses of psychotherapy RCTs that conclude psychotherapy to be efficacious (Cuijpers et al. 2021). However, a recent network meta-analysis (NMA) (Zhou et al. 2020), an established method of comparing treatments using both direct and indirect (i.e. treatment A with treatment C, via studies that directly compare A with B and B with C) evidence, concluded that only fluoxetine alone and fluoxetine administered together with CBT were significantly more effective than medication control (i.e. pill placebo) or psychotherapy controls. The only head-to-head RCT comparing modalities found that fluoxetine, alone and in combination with CBT, was superior to pill placebo, though CBT alone was not (March et al. 2004). Given this confusing evidence base, how should patients, carers, clinicians and policy makers make treatment decisions?

In this paper, we examine whether the existing evidence for adolescent depression treatment can offer valid answers to this question. We provide a conceptual framework and test a series of hypotheses using data from existing trials. Two points are crucial to indirect comparisons of treatment modalities. First, whether the participants of trials in one modality are comparable to those in another modality. Second, whether key conditions of the trial, such as the effects of control conditions or the number of sites involved, are comparable.

Starting with the first point, comparison between different trials assumes that they sample from the same population. If not, the validity of any comparisons, including those conducted through NMA (which rests on the principal of transitivity, i.e. the requirement that the different sets of randomized trials are similar on average (Chaimani et al. 2023)) are questionable (see Supplement for further details).

The assumption that medication and psychotherapy trials sample from the same population may not be valid as patients and parents often have treatment preferences (Jaycox et al. 2006; Langer et al. 2021; McHugh et al. 2013), meaning that there is likely to be a self-selection bias in who participates in psychotherapy and medication trials. Moreover, treatment preferences correlate with clinically-relevant characteristics of the participants, including severity and sex. Some of these characteristics, such as severity, may moderate treatment response (Courtney et al. 2022; Lorenzo-Luaces, Rodriguez-Quintana, and Bailey 2020) and therefore, if they differ across psychotherapy and medication trials, they may confound comparisons.

In terms of the second point, differences in trial design may impact outcomes in a differential way between medication and psychotherapy trials (Del Giovane, Cortese, and Cipriani 2019). Most obviously, participants in psychotherapy trials are unblinded to treatment allocation, with the exception perhaps of trials that compare two equally plausible treatment arms (Calvo et al. 2014). This creates differential expectations which may favour the psychotherapy active condition and disadvantage the psychotherapy control; participants in the active condition are content to be receiving the “cutting edge” treatment whilst those in the control are dissatisfied for having missed out (i.e. “disappointment bias” (Relton et al. 2017). By contrast, in new antidepressant trials, patients (and raters) were largely unable to judge treatment allocation (Lin et al. 2022), suggesting that expectancy effects are well-matched across conditions. Since expectancy is substantially associated with treatment outcomes (Constantino et al. 2011), if expectancy differs between medication and psychotherapy trials, comparisons between them, including in NMA, become questionable.

Another difference in design is the number of trial sites. Previous research has shown that the number of sites in medication trials is positively related to the magnitude of placebo response (Bridge et al. 2009; Dechartres et al. 2011). This phenomenon may be due to lower quality of assessments in trials with more sites, with higher rates of classification errors, entailing higher rates of apparent spontaneous remission or regression to the mean.

An inter-related issue concerns the effect of control conditions. Often, in the public domain, psychotherapy and medication are compared to each other on the basis of their respective effect sizes. However, these effect sizes represent differences between the active intervention and the control condition for each modality. For these effect sizes to be comparable, placebo and psychotherapy controls ought to be equal in their effects. Otherwise, misleading conclusions could be drawn, e.g. two effect sizes of 40% would be considered equal, even if one arose from a difference of 100%-60% and another arose from a difference of 40%-0%.

Additionally, control conditions in RCTs should generate counterfactual conditions to the intervention (Guo and Fraser 2014): what would have been the outcome had an individual not received the intervention, with all else being equal. A pill placebo condition, where the appearance of the drug is faithfully emulated, is an effort for all else to be equal. In psychotherapy trials, control conditions may not be so well matched to the intervention (e.g. with respect to number of hours of therapist contact).

# Objective

In this study we examine RCTs of psychotherapy and medication for child and adolescent depression (mean age 4-18 years). We hypothesise that there are substantial methodological differences between psychotherapy and medication RCTs, making comparisons problematic, and we examine this in the following ways. First, we conduct meta-analyses to compare sample characteristics of medication and psychotherapy trials including: a) baseline severity of depression; b) percentage females; and c) mean age. Second, we examine trial characteristics including the efficacy of the control arms, using random effects meta-regression, and the number of trial sites. Third, we examine the quality of psychotherapy control conditions by scrutinising the extent to which they are matched to the active intervention in ways such as number and frequency of sessions, and hence whether they represent fair pairings from which to draw valid efficacy inferences.

# Study selection and analysis

The study protocol was registered on the Open Science Framework (OSF) and can be found [here](https://osf.io/bfmc6).

A more detailed description of our methods, including the formalisms on which analyses are based, can be found in the Supplement.

## Included studies

We included RCTs identified in a recent meta-analysis of psychotherapy versus control (Cuijpers et al. 2021) and an NMA examining the efficacy of antidepressants (Cipriani et al. 2016) for depression in children and adolescents. For the psychotherapy trials, we utilised open data from the previous meta-analysis (see [here](https://docs.metapsy.org/databases/depression-childadol-psyctr/)). For medication trials, we were unable to access the full dataset used in the NMA and hence extracted data from the included studies ourselves. We excluded three studies because they had no control arm. We were unable to locate and complete extraction for two RCTs (Almeida-Montes, 2005; Eli Lilly, 1986). Many studies did not report complete data; we contacted all corresponding authors to request missing data, though did not receive any responses.

For medication trials, we also conducted a systematic search for studies published after the final search date of Cipriani’s (2016) review up to the final search date of Cuijpers et al’s (2021) review to ensure we analysed an equivalently up-to-date database of medication trials. Two authors screened 450 titles and abstracts, and 38 full text records. Seven studies met inclusion criteria and one author completed data extraction for these papers (see [Figure 1](#fig-prisma) for PRISMA diagram).

## Statistical Analysis

### Sample characteristics

We conducted a series of random effects meta-analyses and tested subgroup differences (psychotherapy vs medication trials) in severity of depressive symptoms at baseline, sex and age. Meta-analyses were implemented using R’s meta package (version 7.0-0).

### Trial design

#### Measures of Effect

As the measure of effect of each individual study, we used the within-group Standardised Mean Difference (SMD) for the primary depression symptom scale used (selected according to the hierarchy defined in the Supplement).

Where individual studies did not report all data required to calculate the SMD, we imputed missing data according to the methods summarised in this Cochrane Handbook (Higgins, Li, and Deeks 2023).

For the purposes of meta-analysis, it is necessary to estimate a standard error of the SMD. This requires a correlation between the pre- and post-measures, a statistic typically not reported. To ensure that our results are not biased by misestimation, we simulated n = 1000 datasets for different values (0.45 to 0.9) of this correlation and used these datasets in the subsequent analyses (please refer to Supplement for full details).

#### Random Effects Metaregression

We estimated the pooled standardised mean difference for each arm by using a random effects meta-analysis implemented in R’s metafor package (version 4.4-0).

We present the SMDs of each of the four treatment arms (medication control, medication active, psychotherapy control, psychotherapy active) under investigation. The SMDs are the means across the 1000 simulated datasets.

#### Number of sites

We also conducted a t-test to compare mean number of trial sites between psychotherapy and medication trials.

### Sensitivity Analyses

We conducted a series of sensitivity analyses. For each of the meta-analyses we excluded studies that 1) used waitlist as their control and 2) recruited participants with subclinical levels of depression. Next, we conducted two analyses where we included only trials that used the Children’s Depression Rating Scale, Revised (CDRS-R) or the Hamilton Depression Rating Scale (HAM-D) as outcome instruments.

Further, we tested whether the simulated values for the standard error had a substantial influence on the estimation of the differences between the medication and psychotherapy control conditions. To inspect whether this is the case, we plotted the z-value of the difference between the two coefficients against the number of simulations. We make inference on the stability of the difference, by counting the proportion of times that the z-value is above the critical value of z = 1.645 corresponding to an alpha = 0.05.

Finally, we examined whether differential regression to the mean may account for differences in effect for psychotherapy and medication trials.

### Comparing the control and active arms of psychotherapy trials

We ran t-tests to compare the active and control arms of psychotherapy trials on key variables of interest regarding the intensity of the interventions: the number, duration and intensity of sessions, and the total cumulative hours and duration of the intervention.

# Findings

## Included studies

The data for the studies included in this meta-analysis are summarised in [Table 4](#tbl-all-trials) and are also available as a csv dataframe on [<https://github.com/transatlantic-comppsych/apples_oranges>]. Please see [Figure 1](#fig-prisma) for a summary of the systematic review and the sources of included RCTs.

In total, there were 88 RCTs which included 48 active arms and 36 control arms of medication trials; and 61 active arms and 58 control arms from psychotherapy RCTs. Note that the number of active and control arms does not exactly match because some studies feature more than one control or active arm.

Placebo pill was the control condition for all medication trials. In psychotherapy trials, the control arm included 14 waitlist, 25 treatment-as-usual (TAU), and 19 other control conditions (e.g. relaxation training, attention control, counselling).

|  |
| --- |
| Figure 1: PRISMA chart summarising sources of included studies |

## Sample characteristics at baseline in medication and psychotherapy trials

[Table 1](#tbl-baseline_results) summarises the results from each of the meta-analyses examining sample characteristics at baseline. The summary statistics are provided for each subgroup (i.e. for medication and psychotherapy RCTs) and the p-value derives from the test for subgroup differences.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1: Sample characteristics at baseline across medication and psychotherapy studies: Results for overall sample and sensitivity analyses   | **Subgroup** | **K** | **Mean** | **SE** | **Lower CI** | **Upper CI** | **T2** | **p-value** | | --- | --- | --- | --- | --- | --- | --- | --- | | **Baseline Severity of Depressive Symptoms\*** | | | | | | | | | **Overall** |  |  |  |  |  |  | **0.009** | | Psychotherapy | 46 | 0.36 | 0.02 | 0.32 | 0.4 | 0.02 |  | | Medication | 31 | 0.42 | 0.01 | 0.39 | 0.44 | 0 |  | | **Excluding subclinical** |  |  |  |  |  |  | **0.119** | | Psychotherapy | 38 | 0.38 | 0.02 | 0.34 | 0.42 | 0.01 |  | | Medication | 31 | 0.42 | 0.01 | 0.39 | 0.44 | 0 |  | | **Excluding waitlist** |  |  |  |  |  |  | **0.024** | | Psychotherapy | 38 | 0.36 | 0.02 | 0.31 | 0.41 | 0.02 |  | | Medication | 31 | 0.42 | 0.01 | 0.39 | 0.44 | 0 |  | | **Percent Female** | | | | | | | | | **Overall** |  |  |  |  |  |  | **0.029** | | Psychotherapy | 45 | 61.05 | 2.43 | 56.16 | 65.94 | 264.99 |  | | Medication | 28 | 53.72 | 2.33 | 48.94 | 58.51 | 152.15 |  | | **Excluding subclinical** |  |  |  |  |  |  | **0.035** | | Psychotherapy | 38 | 61.4 | 2.8 | 55.72 | 67.07 | 297.93 |  | | Medication | 28 | 53.72 | 2.33 | 48.94 | 58.51 | 152.15 |  | | **Excluding waitlist** |  |  |  |  |  |  | **0.044** | | Psychotherapy | 37 | 61.01 | 2.77 | 55.39 | 66.63 | 284.33 |  | | Medication | 28 | 53.72 | 2.33 | 48.94 | 58.51 | 152.15 |  | | **Age** | | | | | | | | | **Overall** |  |  |  |  |  |  | **0.287** | | Psychotherapy | 50 | 14.23 | 0.35 | 13.54 | 14.93 | 5.97 |  | | Medication | 28 | 13.69 | 0.37 | 12.95 | 14.44 | 3.7 |  | | **Excluding subclinical** |  |  |  |  |  |  | **0.336** | | Psychotherapy | 41 | 14.21 | 0.39 | 13.42 | 15.01 | 6.33 |  | | Medication | 28 | 13.69 | 0.37 | 12.95 | 14.44 | 3.7 |  | | **Excluding waitlist** |  |  |  |  |  |  | **0.334** | | Psychotherapy | 42 | 14.21 | 0.38 | 13.43 | 14.98 | 6.15 |  | | Medication | 28 | 13.69 | 0.37 | 12.95 | 14.44 | 3.7 |  | | \*These are baseline depression scores transformed to reflect percentage of a scale range (see Supplement for detailed description).To take an example, the CDRS gives a possible total score from 17 to 113 (i.e. range of 96). Mean severity was 0.36 for psychotherapy studies and 0.42 for medication studies, which would translate to 51.56 (17 + 0.36 x 96) and 57.32 (17 + 0.42 x 96), respectively, as equivalent scores on the CDRS. | | | | | | | | |

### Baseline severity

On average, severity of depressive symptoms at baseline was significantly higher in medication trials compared to psychotherapy trials (see [Table 1](#tbl-baseline_results)). When excluding RCTs that used waitlist as their control, baseline severity remained significantly higher in medication trials compared to psychotherapy trials. This difference did not reach statistical significance when excluding studies that recruited samples with sub-clinical depression.

To ensure that this was not an artefact of variable transformation, we also compared means at baseline in the two instruments, CDRS-R and HAM-D, on which there was a sufficient number of studies to metanalyse. As can be seen in [Table 5](#tbl-hamd-baseline) and [Table 6](#tbl-cdrs-baseline) (see Supplementary Materials), the number of studies is much smaller, but the pattern of differences is the same for the HAM-D and the CDRS, though it does not reach statistical significance for the latter.

### Sex

For this analysis, we excluded the two psychotherapy trials which included entirely female samples (Moeini, 2019; Shomaker, 2016). As can be seen in [Table 1](#tbl-baseline_results), psychotherapy trials featured a significantly higher percentage of females when compared to medication trials. On average, samples were 61.05% (*SE* = 2.43) female across psychotherapy trials and 53.72% (*SE* = 2.33) female across medication trials. Excluding sub-clinical and waitlist control studies yielded similar results.

### Age

As can be seen in [Table 1](#tbl-baseline_results), mean age was 14.23 (*SE* = 0.35) across psychotherapy trials and 13.69 (*SE* = 0.37) across medication trials, with no significant between group differences. There were no significant differences in mean age between modalities on further sensitivity analyses.

## Trial design

### Standardised mean differences of control conditions in psychotherapy and medication studies

We applied metaregression to obtain the SMDs and confidence intervals of each of the four study arms. As can be seen in [Figure 2](#fig-plot-means-all) there were substantial differences between the four arms of the meta-analysis with striking differences between the medication and the psychotherapy control arms. In particular, pill placebo had an SMD = -1.9 (95% CI: -2.1 to -1.7) whereas psychotherapy controls had an SMD = -0.5 (95% CI: -0.75 to -0.25 ).

|  |
| --- |
| Figure 2: Meta-analytic estimates of within-group changes: all studies |

In [Table 2](#tbl-results-overall), we present the regression that tests our hypothesis about differences between medication and psychotherapy controls. Here, medication control is the reference category to which all others are compared. The strongest difference between arms, as judged by the z-value, is between the psychotherapy and medication controls with a z-value 10.94 (p<0.0001).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 2: Results from metaregression with overall sample   | **Condition** | **Coefficient** | **SE** | **z value** | **Lower CI** | **Upper CI** | | --- | --- | --- | --- | --- | --- | | Medication Control | -1.90 | 0.10 |  | -2.10 | -1.70 | | Medication Active | -0.24 | 0.13 | -1.79 | -0.50 | 0.02 | | Psychotherapy Control | 1.40 | 0.13 | 10.94 | 1.14 | 1.65 | | Psychotherapy Active | 0.75 | 0.13 | 5.82 | 0.49 | 1.00 | | T2 = 0.2; *I*2 = 92.5; *K* = 183.0; *R*2 = 64.9. | | | | | | |

##### Sensitivity analyses

We conducted a series of sensitivity analyses (all figures can be found in the Supplementary Materials). First, we analysed the data after excluding waitlist control studies, which yielded a pattern of results is very similar to that of the overall analyses (see [Figure 3](#fig-plot-means-no-wl)). Second, we analysed the data after excluding sub-clinical studies, which again yielded a similar pattern of results (see [Figure 4](#fig-plot-means-clin)). Third, we examined the data including only those studies that used the CDRS (see [Figure 5](#fig-plot-means-cdrs)) or the HAM-D (see [Figure 6](#fig-plot-means-hamd)). Medication control and psychotherapy control conditions remained significantly different, however the small number of studies using each of these instruments resulted in less precise meta-analytic estimates of the SMDs. Finally, we showed different values for the pre-post measure correlation had minimal effect on the estimated outcomes (see [Figure 7](#fig-stab-sims)).

##### Addressing Regression to the Mean

We addressed potential regression to the mean by residualising SMDs by the percentage adjusted baseline scores in one regression model (see [Table 7](#tbl-adj-smds) in the Supplementary Materials for the adjusted SMDs with their confidence intervals). The difference between the medication control and psychotherapy control arms remained similar to the overall analyses.

### Number of trial sites

Average number of trial sites was significantly higher in medication trials (*M* = 35.96, *SD* =25.16) compared to psychotherapy studies (*M* =3.07, *SD* =3.16)(*t* (27.53) = 6.89, *p* =< 0.001). Of those studies for which we had data on number of sites, 26 of 28 (93%) medication trials were multisite, compared to 24 of 45 (54%) psychotherapy studies.

## Comparing the nature and intensity of control conditions in psychotherapy trials

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 3: Comparing the intensity of the intervention between active and control arms of psychotherapy studies   | **Group** | **N** | **Mean** | **SD** | **Cohen's d** | **Upper CI** | **Lower CI** | **t** | **df** | **p-value** | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Number of sessions** | | | | | | | | | | | Active | 63 | 13.19 | 11.40 | 0.78 | 0.36 | 1.20 | 4.37 | 98.63 | < 0.001 | | Control | 39 | 5.56 | 6.19 |  |  |  |  |  |  | | **Intensity (sessions per week)** | | | | | | | | | | | Active | 57 | 1.32 | 0.72 | 1.06 | 0.60 | 1.51 | 4.99 | 75.49 | < 0.001 | | Control | 35 | 0.57 | 0.68 |  |  |  |  |  |  | | **Session length (mins)** | | | | | | | | | | | Active | 55 | 64.63 | 31.32 | 1.11 | 0.65 | 1.57 | 4.99 | 66.42 | < 0.001 | | Control | 35 | 28.24 | 35.12 |  |  |  |  |  |  | | **Total intervention hours** | | | | | | | | | | | Active | 55 | 13.96 | 10.18 | 0.97 | 0.51 | 1.42 | 4.74 | 84.99 | < 0.001 | | Control | 35 | 4.97 | 7.77 |  |  |  |  |  |  | |

As seen in [Table 3](#tbl-intensity-of-int), active conditions featured significantly more sessions when compared to control conditions. Sessions in active conditions were longer and more frequent, resulting in significantly more intervention hours overall. Notably, many control conditions were very poorly described and their intensity could not be quantified, resulting in missing data for a number of trials. We performed a sensitivity analysis where we excluded trials that used a WL control condition; with the exception of number of sessions, differences between active and control arms were no longer statistically significant though remained substantial, with between-group Cohen’s ds ranging from 0.41-0.54 (please see [Table 8](#tbl-intensity-of-int-no-wl) in the Supplement).

# Conclusions and Clinical Implications

This paper sought to address the question of whether psychotherapy and medication can be meaningfully compared on the basis of the existing evidence. Specifically, our paper answered two questions that are fundamental to any attempts at comparing the two modalities. First, whether the participants of trials in one modality are comparable to those in another modality. Second, whether key conditions of the trial, such as the effects of control conditions or the number of sites involved, are comparable.

Starting with the first question, we found that participants in medication trials are comparable on age but more likely to be male and have more severe depression compared to those in psychotherapy trials. These are indicators that different people enter medication and psychotherapy trials, thereby violating basic assumptions of comparability between trials for each modality.

Severity is particularly important as it may moderate treatment response, with some evidence suggesting that those with higher baseline scores respond more to antidepressant therapy (Stone et al. 2022) or that their response to pill placebo is lower (Bridge et al. 2009). Other studies argue against severity as a treatment moderator (Tröger et al. 2024; Weitz et al. 2015), however, these are within modality (e.g. within CBT studies), i.e. within those people who have chosen to be in the particular trial and modality. Moreover, there is evidence that severity may represent different subtypes in terms of course of depression and its outcomes in real life settings (Lamers et al. 2016; Simmonds-Buckley, Catarino, and Delgadillo 2021).

We then turned to our second basic question, about whether key conditions of the trial design were comparable between modalities. We found a series of critical differences between the two modalities.

First, medication trials are vastly more likely to be multi-site than their psychotherapy counterparts; in the current study, 93% of medication RCTs were multisite compared to 54% of psychotherapy RCTs. Multisite trials are associated with higher pill placebo response (Bridge et al. 2009), and are less common in non-industry trials which are also more likely to show lower pill placebo efficacy. Also, in single-site trials, principal investigators are often intellectually invested in the treatment (in psychotherapy these are often treatments developed or modified by the PI); this is in stark contrast to the incentive structure in multi-site trials where the number of recruited participants is the primary unit of reimbursement.

Second, psychotherapy controls have moderate effect size (-0.5) whereas pill-placebo has very large effect size (-1.9). Our analysis could be critiqued as it relies on comparing within arm symptom change per trial, therefore effectively breaking randomization. This criticism would apply if our aim were to draw inferences about the efficacy of each arm — in which case preserving the randomization (in order to balance confounders) is critical. However, we note that our findings are largely in keeping with those of the NMA, which is designed to preserve the randomization structure. More importantly, we do not make the claim with our analyses that these differences are genuinely due to efficacy differences; they may well be because people who attend psychotherapy and medication trials are different and therefore respond differently. In either case (difference in efficacy vs difference in trial participant profile), the vast disparity in the response to control conditions is reason for major concern about our ability to draw inferences concerning the comparison of the two modalities. This is particularly problematic as clinicians as well as policy makers often resort to between-group effect sizes to summarise findings. Our results make obvious that comparing treatments based on the between-group effect sizes of each modality is misleading.

Moving beyond comparison between the two modalities, we examined whether psychotherapy controls are reasonable counterfactuals to receiving the treatment. An obvious disadvantage of psychotherapy trials is that they are typically unblinded (and hard to blind). Yet, our results show that psychotherapy trials are unlikely to fulfill some other basic conditions of the “all else is equal” assumption of a counterfactual condition. In order to test that a psychological treatment is effective per se (e.g. because of the specific techniques) rather than because of generic effects (e.g. pleasant human contact), aspects such as therapist contact time should be matched. Many (24%) psychotherapy RCTs use waitlist controls, which by definition do not match for hours of therapist contact, and are often associated with disappointment bias. TAU and other psychotherapy control conditions varied drastically; 9 RCTs used controls that exactly matched the active arm in total number of contact hours, though several studies used bibliotherapy (e.g. an educational brochure) or online-only control conditions which did not involve any direct therapist contact. Importantly, controls were often very poorly described, resulting in difficulties quantifying their intensity and hence evaluating their adequacy as counterfactual conditions. Overall, there is a lack of rigor in matching control to active treatment conditions in psychotherapy RCTs, with the latter typically featuring considerably more contact hours, which may artificially inflate estimates of the efficacy of psychological treatments.

Given all of the above, the certainty with which guidelines recommend psychotherapy over medication for adolescent depression is surprising. Indeed, we believe that our findings have several profound implications for patients, their families, clinicians and policy makers.

First, the grounds for a comparison between medication and psychotherapy should be seen as shaky, rather than offering confidence, and there is an urgent need to revisit guidelines and public information in light of the limitations.

Second, the low quality of control conditions of psychotherapy trials for adolescent depression should prompt a rethinking of what needs to be done to create fair comparators. Indeed, investment should be directed into providing rigorous evidence that establishes depression psychotherapies as more efficacious than fair controls. There are examples of psychotherapy RCTs where such rigor has been applied (e.g. Bolton, 2007; Liddle, 1990; Rohde, 2004) in matching active and control arms on variables such as therapist time and attention, provision of homework, and small group interaction.

Third, our findings make clear the inherent difficulties of comparing psychotherapy with medication trials (Del Giovane, Cortese, and Cipriani 2019). The first obstacle is the comparability of the populations taking part. Head-to-head comparisons of psychotherapy with medication (as has been done in (March et al. 2004)) are more favourable in this regard, yet even so these trials might sample the population of those who are indifferent to which treatment they would receive. And even in such a design, difficulties with blinding of the psychotherapy control would have to be overcome to draw valid inferences.

In summary, the current findings give cause for consternation about the state of the evidence for treatments of youth depression. Our data question the state of knowledge about the efficacy of psychotherapies and, in particular, the extent to which giving them primacy in the treatment of depression is justified and beneficial for young people. Returning to our motivating question, the stakeholders, including patients and clinicians, deserve better evidence on which to base their choices.

# Supplementary Materials

## Systematic review search terms and PRISMA chart

We conducted a systematic search for medication studies published from 31 May 2015 up to 1 Jan 2021 (i.e. after the final search date of Cipriani’s (2016) review up to the final search date of Cuijpers et al’s (2021) review). We searched PubMed, the Cochrane Central Register of Controlled Trials, Embase, Web of Science, CINAHL, PsycINFO and LiLACS for randomised controlled trials (RCTs) comparing any antidepressant with placebo in the treatment of children and adolescents with a primary diagnosis of major depressive disorder. We used the same search terms as Cipriani (2016) with one additional search term to include only placebo-controlled trials (see below). We additionally applied filters to specify our date range, and to exclude reviews and non-human studies. We also searched clinical trial registers for published and unpublished studies however all RCTs meeting inclusion criteria had already been identified from the database search outlined above. Please see [Figure 1](#fig-prisma) for the PRISMA flow diagram.

We used Covidence, an online software tool, to manage our systematic review. Our search produced 538 studies, 88 of which were duplicates and subsequently removed. Two authors screened 450 titles and abstracts, and 38 full text records. Seven studies met inclusion criteria and data extraction was completed for these papers.

**Search terms**

Explicit search strategy: title/abstract = (depress\* or dysthymi\* or “mood disorder\*” or “affective disorder\*“) AND (adolesc\* or child\* or boy\* or girl\* or juvenil\* or minors or paediatri\* or pediatri\* or pubescen\* or school\* or student\* or teen\* or young or youth\*) AND (selective serotonin reuptake inhibitor or SSRI or citalopram or fluoxetine or paroxetine or sertraline or escitalopram or fluvoxamine or serotonin norepinephrine reuptake inhibitor\* or SNRI or venlafaxine or duloxetine or milnacipran or reboxetine or bupropion or noradrenergic and specific serotonergic antidepressants or NaSSA or mirtazapine or TCA or tricyclic or amersergide or amineptine or amitriptyline or amoxapine or butriptyline or chlorpoxiten or clomipramine or clorimipramine or demexiptiline or desipramine or dibenzipin or dothiepin or doxepin or imipramine or lofepramine or melitracen or metapramine or nortriptyline or noxiptiline or opipramol or protriptyline or quinupramine or tianeptine or trimipramine) AND (placebo)

**References for included trials**

References for trials included in the existing meta-analyses drawn upon for this study can be found [here](https://docs.metapsy.org/databases/depression-childadol-psyctr/) for psychotherapy RCTs and [here](https://ora.ox.ac.uk/objects/uuid:e0b5ae23-d562-4348-94b8-84f70b7812c5) for medication RCTs. Below are references for the seven additional RCTs identified in the original systematic review conducted for the current study.

*Publications*

Atkinson, S., Lubaczewski, S., Ramaker, S., England, R. D., Wajsbrot, D. B., Abbas, R., & Findling, R. L. (2018). Desvenlafaxine versus placebo in the treatment of children and adolescents with major depressive disorder. *Journal of Child and Adolescent Psychopharmacology, 28(1),* 55-65. DOI: 10.1089/cap.2017.0099

Durgam, S., Chen, C. Z., Migliore, R., Prakash, C., Edwards, J., & Findling, R. L. (2018). A phase 3, double-blind, randomized, placebo-controlled study of vilazodone in adolescents with major depressive disorder. *Pediatric Drugs, 20(4),* 353-363. DOI: 10.1007/s40272-018-0290-4

Findling, R. L., McCusker, E., & Strawn, J. R. (2020). A randomized, double-blind, placebo-controlled trial of vilazodone in children and adolescents with major depressive disorder with twenty-six-week open-label follow-up. *Journal of Child and Adolescent Psychopharmacology, 30(6),* 355-365. DOI: 10.1089/cap.2019.0176

Le Noury, J., Nardo, J. M., Healy, D., Jureidini, J., Raven, M., Tufanaru, C., & Abi-Jaoude, E. (2015). Restoring Study 329: efficacy and harms of paroxetine and imipramine in treatment of major depression in adolescence. *BMJ, 351,*, h4320. DOI: 10.1136/bmj.h4320

Weihs, K. L., Murphy, W., Abbas, R., Chiles, D., England, R. D., Ramaker, S., & Wajsbrot, D. B. (2018). Desvenlafaxine versus placebo in a fluoxetine-referenced study of children and adolescents with major depressive disorder. Journal of Child and Adolescent Psychopharmacology, 28(1), 36-46. DOI: 10.1089/cap.2017.0100

*Unpublished clinical trials*

Active Reference (Fluoxetine) Fixed-dose Study of Vortioxetine in Paediatric Patients Aged 12 to 17 Years With Major Depressive Disorder (MDD). ClinicalTrials.gov Identifier: NCT02709746. Accessed 17 Jan 2024.

Safety and Efficacy of Levomilnacipran ER in Adolescent Participants With Major Depressive Disorder. ClinicalTrials.gov Identifier: NCT02431806. Accessed 17 Jan 2024.

**Hierarchy of depression symptom severity measurement scales**

Where multiple depression rating scales were used, we selected the best available measure according to the following hierarchy used in Cipriani et al. (2016).

1. Children’s Depression Rating Scale (CDRS)
2. Hamilton Depression Rating Scale (HAMD)
3. Montgomery Asberg Depression Rating Scale (MADRS)
4. Beck Depression Inventory (BDI)
5. Children’s Depression Inventory (CDI)
6. Schedule for Affective Disorders and Schizophrenia for School-Aged Children (K-SADS)
7. Mood and Feeling Questionnaire (MFQ)
8. Reynolds Adolescent Depression Scale (RADS)
9. Bellevue Index of Depression (BID)
10. Child Depression Scale (CDS)
11. Centre for Epidemiological Studies Depression Scale (CES-D)
12. Child Assessment Schedule (CAS)
13. Child Behaviour Checklist-Depression (CBCL-D)

## Full description of methods (including formalisms)

## Included studies

We drew upon RCTs included in two recent comprehensive meta-analyses with open data available for each medication and psychotherapy, and supplemented them with an updated systematic review. Please refer to these original meta-analyses for a detailed description of their search strategy and study selection criteria. Psychotherapy studies were drawn from a systematic review and meta-analysis of randomised trials comparing psychotherapy for youth depression against control conditions (Cuijpers et al. 2021) (dataset available [here](https://docs.metapsy.org/databases/depression-childadol-psyctr/)). Whilst Cuijpers et al. (2021) excluded studies for which the primary outcome variable could not be calculated due to missing data, we included these studies and performed the imputations outlined below. We also included studies which had data available for other variables in interest, including number of sites or baseline demographics; hence we have more psychotherapy studies included in this review compared to the original meta-analysis. Whilst the online database is regularly updated, we chose to exclude studies published after the final date of Cuijpers et al.’s (2021) literature search.

Medication studies were drawn from a network meta-analysis examining the efficacy and tolerability of antidepressants and placebo for major depressive disorder in children and adolescents (Cipriani et al. 2016). A dataset was made available online though did not include means or standard deviations at baseline or post-test. We were unable to access the full dataset used in this meta-analysis, and hence completed extraction from the included studies ourselves. We excluded three studies because they had no control arm. We were unable to locate and therefore complete extraction for two RCTs (Almeida-Montes, 2005; Eli Lilly, 1986). Many studies did not report complete data; we contacted all corresponding authors to request missing data, though did not receive any responses.

We conducted a systematic search for medication studies published after the final search date of Cipriani’s (2016) review up to the final search date of Cuijpers et al’s (2021) review to ensure we analysed an equivalently up-to-date database of medication trials. Please see the Supplementary Materials for further details. Our search produced 538 studies, 88 of which were duplicates and subsequently removed. Two authors screened 450 titles and abstracts, and 38 full text records. Seven studies met inclusion criteria and data extraction was completed for these papers.

## Statistical Analysis

### Sample characteristics

We conducted a series of random effects meta-analyses and tested for subgroup differences between psychotherapy and medication trials in sample characteristics including sex, age, and severity of depressive symptoms at baseline. Meta-analyses were implemented using R’s Meta package.

In order to compare depression severity across the variety of instruments the studies used, we performed a min-max normalisation to turn each study arm mean score at baseline into a percentage using the following formalism:

where,

is the mean score for each study arm on the primary outcome questionnaire, and and are the minimum and maximum possible values of the scale in question, respectively. The standard deviation is calculated thus:

where is the original standard deviation of the mean at baseline.

### Trial design

#### Measures of Effect

As the measure of effect of each individual study, we used the within-group Standardised Mean Difference (SMD), which we defined following Cumming (2013) as:

where, and refer to the means of the main outcome score at the end and beginning of the intervention respectively and and to the respective standard deviations. Where individual studies did not report all data required to calculate the SMD, we imputed missing data according to the methods summarised in this Cochrane Handbook (Higgins, Li, and Deeks 2023), in the following order. If a study reported the standard error of the mean, the SD was obtained simply by multiplying the SE by the square root of the sample size. For conditions where the SD was missing at one time point, the baseline SD was substituted by the post-test SD, and vice versa. If the SD was not available at either time point, missing values were replaced by the mean of the SDs available for comparable cases (defined as same trial type (psy or med), same instrument, same timepoint (pre or post), and same arm (control or active)). Where there were missing means at either baseline or post-test, missing values were calculated using mean change scores, preferring the change scores reported in the paper itself, though where this was unavailable, using the change scores reported in the dataset from Cipriani’s meta-analysis (for medication studies only).

For the purposes of meta-analysis, it is necessary to estimate a standard error of the SMD. This is calculated according to:

where refers to the study sample size and refers to the correlation between the outcome score obtained at baseline and at the end point. This correlation is typically not reported in studies and is often imputed using previously reported correlations for the instruments used. However, this practice has given rise to concerns about misestimation. Whilst such misestimation is possible, there is no reason to expect that it would be systematic, i.e. bias estimation of the effects for the control group of medication compared to those of psychotherapy. Still, to alleviate such concerns we have used a simulations.

In particular, we simulated one thousand truncated distribution of standard errors with the following general characteristics:

for which we chose the mean to be , the standard deviation to be , and the upper and lower bounds to be and , respectively. We then used these simulated datasets in the subsequent meta-analyses.

#### Random Effects Metaregression

We estimated the pooled standardised mean difference for each arm by using a random effects meta-analysis implemented in R’s metafor package. The main underlying assumption of random effects meta-analysis is that each study’s true effect size is affected not only by sampling error , but also by which represents heterogeneity between studies, allowing each study’s estimate to vary along a distribution of effects, and the distribution of true effect sizes termed . Therefore, we can estimate a two stage model with:

where is the estimated effect size for study i, has a normal distribution with as its true mean effect and sampling error . Whereas is a study-specific instantiation of the distribution of effect sizes, with representing heterogeneity.

This then gives rise to:

where,

describes the deviation of each study from the mean of the distribution, and,

describes the sampling error.

We can then specify the following model to obtain the means of each arm of the trials as follows:

where to obtain the mean of each level is the sum of , the intercept for the reference category of medication control, with the coefficient of each level, e.g. for level 3, the psychotherapy controls. The confidence intervals of the means are constructed in the standard way using the standard errors of the mean. Similarly, each coefficient represents the contrast between the reference category and each level, for an example and of main interest to us represents the contrast between psychotherapy and medication control arms. Inference on the contrasts is done as follows:

We used maximum likelihood (ML) to estimate model and applied Hartung-Knapp adjustment to reduce the chance of false positives (IntHout, Ioannidis, and Borm 2014).

We present the SMDs of each of the four treatment arms (medication control, medication active, psychotherapy control, psychotherapy active) under investigation. The SMDs are the means across the 1000 simulated datasets.

#### Number of sites

We also conducted a t-test to compare mean number of trial sites between psychotherapy and medication trials.

### Sensitivity Analyses

We conducted a series of sensitivity analyses. For each of the meta-analyses we excluded studies that 1) used waitlist as their control and 2) recruited participants with subclinical levels of depression. Next, we conducted two analyses where we included only trials that used the Children’s Depression Rating Scale, Revised (CDRS-R) or the Hamilton Depression Rating Scale (HAM-D) as outcome instruments.

Further, we tested whether the simulated values for the standard error had a substantial influence on the estimation of the differences between the medication and psychotherapy control conditions. To inspect whether this is the case, we plotted the z-value of the difference between the two coefficients against the number of simulations. We make inference on the stability of the difference, by counting the proportion of times that the z-value is above the critical value of z = 1.645 corresponding to an alpha = 0.05.

Finally, we examined whether differential regression to the mean may account for differences in effect for psychotherapy and medication trials.

### Comparing the control and active arms of psychotherapy trials

We ran t-tests to compare the active and control arms of psychotherapy trials on key variables of interest regarding the intensity of the interventions. We extracted data pertaining to the number, duration and intensity of sessions, and the total cumulative hours and duration of the intervention. Where a range was provided, the maximum was encoded (e.g. if a paper reported that an intervention involved 8-10 sessions lasting 50-60 minutes, we encoded the number and duration of sessions as 10 and 60, respectively). If sessions varied in frequency across an intervention, we calculated an average by dividing total number of sessions by length of intervention period. Similarly, if the length of sessions varied across the course of the intervention, we calculated a weighted average. Phone call, web-chat and online sessions were encoded as sessions, however guided self-help components were not.

## Summary of all included trials

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 4: Summary of included RCTs   | **Study** | **Arm** | **Description** | **N** | **Instrument** | **Baseline M** | **Baseline SD** | **Post M** | **Post SD** | **Cohen's d** | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Medication** | | | | | | | | | | | Atkinson, 2014 | Active | Duloxetine | 117 | cdrs | 59.20 | 10.50 | 35.00 | 10.50 | -2.30 | | Active | Fluoxetine | 117 | cdrs | 58.80 | 10.60 | 35.60 | 10.60 | -2.19 | | Control | Placebo | 103 | cdrs | 60.20 | 11.70 | 35.00 | 11.70 | -2.15 | | Atkinson, 2018 | Active | Desvenlafaxine (high dose) | 121 | cdrs | 58.45 | 9.45 | 34.05 | 9.45 | -2.58 | | Active | Desvenlafaxine (low dose) | 122 | cdrs | 58.52 | 9.18 | 34.82 | 9.18 | -2.58 | | Control | Placebo | 119 | cdrs | 57.28 | 8.94 | 34.38 | 8.94 | -2.56 | | Berard, 2006 | Active | Paroxetine | 182 | madrs | 25.90 | 6.75 | 12.30 | 6.75 | -2.01 | | Control | Placebo | 93 | madrs | 25.90 | 5.79 | 13.10 | 5.79 | -2.21 | | Bristol-Myers Squibb, 2002a | Active | Nefazodone |  | cdrs |  | 9.33 |  | 10.83 |  | | Control | Placebo |  | cdrs |  | 9.30 |  | 10.87 |  | | Bristol-Myers Squibb, 2002b | Active | Nefazodone | 90 | cdrs | 61.20 | 9.33 | 38.00 | 10.83 | -2.30 | | Control | Placebo | 93 | cdrs | 58.30 | 9.30 | 36.70 | 10.87 | -2.14 | | Durgam, 2018 | Active | Vilazodone (15mg) | 175 | cdrs | 57.80 | 8.70 | 33.80 | 12.00 | -2.32 | | Active | Vilazodone (30mg) | 180 | cdrs | 56.80 | 8.50 | 32.50 | 11.50 | -2.43 | | Control | Placebo | 171 | cdrs | 57.50 | 8.60 | 34.00 | 12.90 | -2.19 | | Emslie, 1997 | Active | Fluoxetine | 48 | cdrs | 58.50 | 10.50 | 38.40 | 14.80 | -1.59 | | Control | Placebo | 48 | cdrs | 57.60 | 10.40 | 47.10 | 17.00 | -0.77 | | Emslie, 2002a | Active | Fluoxetine | 109 | cdrs | 57.10 | 9.90 | 35.10 | 13.50 | -1.88 | | Control | Placebo | 110 | cdrs | 55.10 | 11.80 | 40.20 | 13.50 | -1.18 | | Emslie, 2002b | Active | Nefazodone | 99 | cdrs |  | 9.33 |  | 10.83 |  | | Control | Placebo | 96 | cdrs |  | 9.30 |  | 10.87 |  | | Emslie, 2006 | Active | Paroxetine | 101 | cdrs | 60.70 | 9.37 | 38.12 | 9.37 | -2.41 | | Control | Placebo | 102 | cdrs | 62.60 | 8.96 | 39.22 | 8.96 | -2.61 | | Emslie, 2007a | Active | Venlafaxine | 80 | cdrs | 54.40 | 13.19 | 36.30 | 13.19 | -1.37 | | Control | Placebo | 85 | cdrs | 54.40 | 12.82 | 38.30 | 12.82 | -1.26 | | Emslie, 2007b | Active | Venlafaxine | 102 | cdrs | 57.30 | 13.06 | 32.70 | 13.06 | -1.88 | | Control | Placebo | 94 | cdrs | 57.30 | 13.43 | 34.50 | 13.43 | -1.70 | | Emslie, 2009 | Active | Escitalopram | 154 | cdrs | 57.60 | 8.19 | 35.50 | 8.19 | -2.70 | | Control | Placebo | 157 | cdrs | 56.00 | 8.27 | 37.20 | 8.27 | -2.27 | | Emslie, 2014 | Active | Duloxetine | 108 | cdrs | 59.30 | 10.90 | 35.00 | 10.90 | -2.23 | | Active | Fluoxetine | 117 | cdrs | 57.90 | 10.10 | 36.40 | 10.10 | -2.13 | | Control | Placebo | 122 | cdrs | 58.20 | 9.40 | 37.40 | 9.40 | -2.21 | | Findling, 2009 | Active | Fluoxetine | 18 | cdrs | 53.00 | 9.84 | 34.60 | 13.66 | -1.57 | | Control | Placebo | 16 | cdrs | 53.94 | 9.84 | 31.31 | 13.68 | -1.92 | | Findling, 2020 | Active | Fluoxetine | 97 | cdrs | 58.10 | 8.30 | 34.70 | 12.80 | -2.22 | | Active | Vilazodone | 186 | cdrs | 58.50 | 9.40 | 37.20 | 14.20 | -1.81 | | Control | Placebo | 182 | cdrs | 57.70 | 9.20 | 37.80 | 13.70 | -1.74 | | Forest, 2020 | Active | Fluoxetine | 134 | cdrs | 61.50 | 9.33 | 37.13 | 10.83 | -2.42 | | Active | Levomilnacipran (40mg) | 134 | cdrs | 61.80 | 9.33 | 38.52 | 10.83 | -2.31 | | Active | Levomilnacipran (80mg) | 138 | cdrs | 59.40 | 9.33 | 36.76 | 10.83 | -2.25 | | Control | Placebo | 140 | cdrs | 61.10 | 9.30 | 38.20 | 10.87 | -2.27 | | Geller, 1990 | Active | Nortriptyline | 12 | cdrs | 51.30 | 4.40 | 34.70 | 7.80 | -2.72 | | Control | Placebo | 19 | cdrs | 52.40 | 3.70 | 37.80 | 9.10 | -2.28 | | Geller, 1992 | Active | Nortriptyline | 26 | cdrs | 49.90 | 4.20 | 32.90 | 11.40 | -2.18 | | Control | Placebo | 24 | cdrs | 49.60 | 4.60 | 32.00 | 9.80 | -2.44 | | Hughes, 1990 | Active | Imipramine |  | cdrs |  | 9.33 |  | 10.83 |  | | Control | Placebo |  | cdrs |  | 9.30 |  | 10.87 |  | | Keller, 2001 | Active | Imipramine | 95 | hamd | 18.11 | 4.19 | 9.20 | 7.85 | -1.48 | | Active | Paroxetine | 93 | hamd | 18.98 | 4.15 | 8.24 | 7.68 | -1.82 | | Control | Placebo | 87 | hamd | 18.97 | 4.10 | 9.88 | 7.74 | -1.54 | | Klein, 1998 | Active | Desipramine | 23 | hamd | 21.44 | 3.70 | 10.23 | 2.10 | -3.87 | | Control | Placebo | 22 | hamd | 21.33 | 5.20 | 14.61 | 2.10 | -1.84 | | Kutcher, 1994 | Active | Desipramine | 30 | hamd | 22.63 | 5.17 | 12.68 | 8.68 | -1.44 | | Control | Placebo | 30 | hamd | 23.77 | 5.31 | 13.42 | 8.43 | -1.51 | | Kye, 1996 | Active | Amitriptyline | 18 | hamd | 12.00 | 4.50 | 8.00 | 4.90 | -0.85 | | Control | Placebo | 13 | hamd | 13.20 | 4.10 | 8.80 | 4.50 | -1.02 | | Le Noury, 2015 | Active | Imipramine | 95 | hamd | 18.10 | 4.19 | 9.10 | 4.19 | -2.15 | | Active | Paroxetine | 93 | hamd | 18.90 | 4.24 | 8.20 | 4.24 | -2.52 | | Control | Placebo | 87 | hamd | 19.00 | 4.10 | 9.90 | 4.10 | -2.22 | | Lundbeck, 2020 | Active | Fluoxetine | 150 | cdrs | 61.80 | 8.90 | 39.80 | 8.90 | -2.47 | | Active | Vortioxetine (10mg) | 145 | cdrs | 61.20 | 9.40 | 44.10 | 9.40 | -1.82 | | Active | Vortioxetine (20mg) | 159 | cdrs | 62.50 | 9.80 | 43.60 | 9.80 | -1.93 | | Control | Placebo | 153 | cdrs | 60.60 | 9.10 | 42.40 | 9.10 | -2.00 | | March, 2004 | Active | Fluoxetine | 109 | cdrs | 58.94 | 4.00 | 36.30 | 8.18 | -3.72 | | Control | Placebo | 112 | cdrs | 61.18 | 4.27 | 41.77 | 7.99 | -3.17 | | Organon, 2002a | Active | Mirtazapine | 82 | cdrs | 50.93 | 9.33 | 35.08 | 10.83 | -1.57 | | Control | Placebo | 44 | cdrs | 51.93 | 9.30 | 37.24 | 10.87 | -1.46 | | Organon, 2002b | Active | Mirtazapine | 88 | cdrs | 48.87 | 9.33 | 35.39 | 10.83 | -1.34 | | Control | Placebo | 45 | cdrs | 47.57 | 9.30 | 38.76 | 10.87 | -0.87 | | Paxil (GlaxoSmithKline), 2009 | Active | Paroxetine | 29 | cdrs |  | 9.33 |  | 10.83 |  | | Control | Placebo | 27 | cdrs |  | 9.30 |  | 10.87 |  | | Puig-Antich, 1987 | Active | Imipramine | 20 | ksads | 3.10 | 0.43 | 1.90 | 0.68 | -2.16 | | Control | Placebo | 22 | ksads | 3.00 | 0.66 | 1.90 | 0.86 | -1.45 | | Von Knorring, 2006 | Active | Citalopram | 124 | madrs | 30.00 | 5.50 | 17.91 | 5.50 | -2.20 | | Control | Placebo | 120 | madrs | 30.00 | 5.50 | 18.09 | 5.50 | -2.17 | | Wagner, 2003 | Active | Sertraline | 189 | cdrs | 64.30 | 11.00 | 41.46 | 11.00 | -2.08 | | Control | Placebo | 187 | cdrs | 64.60 | 11.00 | 44.41 | 11.00 | -1.84 | | Wagner, 2004 | Active | Citalopram | 89 | cdrs | 58.80 | 10.90 | 37.10 | 10.90 | -1.99 | | Control | Placebo | 85 | cdrs | 57.80 | 11.10 | 41.30 | 11.10 | -1.49 | | Wagner, 2006 | Active | Escitalopram | 131 | cdrs | 54.50 | 9.33 | 32.60 | 10.83 | -2.17 | | Control | Placebo | 133 | cdrs | 56.60 | 9.30 | 36.40 | 10.87 | -2.00 | | Weihs, 2018 | Active | Desvenlafaxine | 115 | cdrs | 56.34 | 9.59 | 33.74 | 9.59 | -2.36 | | Active | Fluoxetine | 112 | cdrs | 56.26 | 8.34 | 31.46 | 8.34 | -2.97 | | Control | Placebo | 112 | cdrs | 57.06 | 8.91 | 33.96 | 8.91 | -2.59 | | **Psychotherapy** | | | | | | | | | | | Ackerson, 1998 | Active | cbt | 12 | hamd | 19.90 | 5.50 | 8.80 | 5.30 | -2.06 | | Control | wl | 10 | hamd | 21.00 | 5.00 | 20.50 | 3.40 | -0.12 | | Arnarson, 2009 | Active | cbt |  |  |  |  |  |  |  | | Control | cau |  |  |  |  |  |  |  | | Asarnow, 2002 | Active | cbt |  | cdi |  | 7.10 | 9.54 | 7.10 |  | | Control | wl |  | cdi |  | 10.56 | 12.33 | 10.56 |  | | Bolton, 2007 | Active | ipt | 105 | apai | 43.50 | 10.10 | 27.80 | 17.20 | -1.15 | | Control | other ctr | 105 | apai | 44.20 | 11.20 | 40.60 | 15.70 | -0.27 | | Control | wl | 104 | apai | 44.20 | 10.80 | 37.30 | 15.90 | -0.52 | | Charkhandeh, 2016 | Active | cbt | 65 | cdi | 29.46 | 5.47 | 19.94 | 5.59 | -1.72 | | Control | other ctr | 63 | cdi | 30.00 | 5.37 | 26.33 | 5.88 | -0.65 | | Control | wl | 60 | cdi | 30.35 | 5.45 | 30.38 | 4.66 | 0.01 | | Clarke, 1995 | Active | cbt | 55 | hamd | 3.55 | 3.20 | 1.87 | 2.50 | -0.59 | | Control | cau | 70 | hamd | 3.86 | 3.10 | 2.91 | 4.30 | -0.26 | | Clarke, 1999 | Active | cbt | 32 | hamd | 15.10 | 6.00 | 6.70 | 7.10 | -1.28 | | Active | cbt | 37 | hamd | 13.00 | 5.30 | 4.60 | 4.80 | -1.66 | | Control | wl | 27 | hamd | 14.50 | 5.90 | 7.70 | 7.00 | -1.05 | | Clarke, 2001 | Active | cbt | 45 | hamd | 3.20 | 3.40 | 1.80 | 2.10 | -0.51 | | Control | cau | 49 | hamd | 3.10 | 3.20 | 2.90 | 4.60 | -0.05 | | Clarke, 2002 | Active | cbt | 41 | hamd | 12.00 | 5.30 | 5.50 | 5.20 | -1.24 | | Control | cau | 47 | hamd | 11.40 | 5.00 | 6.00 | 5.10 | -1.07 | | De Cuyper, 2004 | Active | cbt | 9 | cdi | 12.67 | 6.00 | 10.11 | 6.03 | -0.43 | | Control | wl | 11 | cdi | 15.27 | 4.54 | 11.73 | 5.66 | -0.69 | | De Jonge-Heesen, 2020 | Active | cbt | 66 | cdi | 16.18 | 4.92 | 13.32 | 7.07 | -0.48 | | Control | other ctr | 64 | cdi | 15.68 | 7.08 | 14.71 | 9.06 | -0.12 | | Diamond, 2002 | Active | other psy | 16 | bdi | 23.80 | 7.40 | 11.80 | 8.80 | -1.48 | | Control | wl | 16 | bdi | 28.00 | 7.10 | 18.50 | 11.10 | -1.04 | | Diamond, 2010 | Active | other psy | 35 | bdi | 33.00 | 9.66 | 12.60 | 13.88 | -1.73 | | Control | cau | 31 | bdi | 33.00 | 9.37 | 18.50 | 15.91 | -1.15 | | Esposito-Smythers, 2019 | Active | cbt | 74 | cdi | 27.00 | 8.70 | 13.20 | 9.30 | -1.53 | | Control | cau | 73 | cdi | 26.30 | 9.80 | 13.80 | 9.20 | -1.32 | | Fristad, 2019 | Active | cbt | 19 | cdrs | 42.00 | 9.00 | 30.00 | 9.00 | -1.33 | | Control | other ctr | 18 | cdrs | 44.00 | 13.00 | 31.00 | 11.00 | -1.08 | | Gillham, 2006 | Active | cbt | 147 | cdi | 13.19 | 7.81 | 11.55 | 8.21 | -0.20 | | Control | cau | 124 | cdi | 12.57 | 7.17 | 11.37 | 7.90 | -0.16 | | Idsoe, 2019 | Active | cbt | 133 | ces\_d | 33.08 | 9.97 | 26.85 | 11.82 | -0.57 | | Control | cau | 95 | ces\_d | 32.01 | 9.75 | 29.55 | 10.77 | -0.24 | | Israel, 2013 | Active | other psy | 11 | hamd | 20.60 | 4.60 | 12.50 | 7.20 | -1.37 | | Control | cau | 9 | hamd | 19.70 | 5.50 | 19.40 | 5.20 | -0.06 | | Kahn, 1990 | Active | cbt | 17 | cdi | 31.11 | 9.58 | 7.29 | 66.03 | -0.63 | | Active | other psy | 17 | cdi | 27.18 | 7.84 | 13.58 | 7.38 | -1.79 | | Control | other ctr | 17 | cdi | 26.94 | 10.83 | 12.88 | 10.71 | -1.31 | | Control | wl | 17 | cdi | 28.06 | 9.75 | 26.94 | 15.41 | -0.09 | | Lewinsohn, 1990 | Active | cbt | 19 | bdi | 21.26 | 11.35 | 6.47 | 8.53 | -1.49 | | Active | cbt | 21 | bdi | 21.67 | 11.34 | 10.00 | 11.91 | -1.00 | | Control | wl | 19 | bdi | 23.84 | 11.43 | 20.47 | 10.28 | -0.31 | | Liddle, 1990 | Active | cbt | 11 | cdi | 21.00 | 4.45 | 14.45 | 6.74 | -1.17 | | Active | cbt | 11 | cdi | 21.00 | 4.45 | 14.45 | 6.74 | -1.17 | | Control | other ctr | 10 | cdi | 22.30 | 4.24 | 19.30 | 6.93 | -0.54 | | Control | wl | 10 | cdi | 20.70 | 3.34 | 16.90 | 6.79 | -0.75 | | Listug-Lunde, 2013 | Active | cbt | 8 | cdi | 21.00 | 5.29 | 14.38 | 9.93 | -0.87 | | Control | cau | 8 | cdi | 20.37 | 4.10 | 13.25 | 9.87 | -1.02 | | Luby, 2012 | Active | other psy | 25 | pfc | 42.80 | 5.80 | 30.10 | 11.30 | -1.49 | | Control | other ctr | 18 | pfc | 39.80 | 10.30 | 33.70 | 10.60 | -0.58 | | Makover, 2019 | Active | other psy |  | smfq |  |  |  |  |  | | Control | cau |  | smfq |  |  |  |  |  | | March, 2004 | Active | cbt | 111 | cdrs | 59.64 | 4.52 | 42.06 | 9.18 | -2.57 | | Martinovic, 2006 | Active | cbt | 16 | hamd | 5.90 | 0.80 | 3.30 | 1.29 | -2.49 | | Control | cau | 16 | hamd | 5.70 | 0.70 | 5.80 | 1.98 | 0.07 | | Moeini, 2019 | Active | cbt | 64 | ces\_d | 24.60 | 11.70 | 18.50 | 14.00 | -0.47 | | Control | cau | 64 | ces\_d | 22.30 | 11.80 | 21.40 | 15.60 | -0.07 | | Mufson, 1999 | Active | ipt | 24 | hamd | 19.20 | 7.50 | 6.30 | 7.70 | -1.70 | | Control | other ctr | 24 | hamd | 18.70 | 8.60 | 11.80 | 8.90 | -0.79 | | Mufson, 2004 | Active | ipt | 34 | hamd | 18.90 | 5.90 | 8.70 | 8.00 | -1.47 | | Control | cau | 29 | hamd | 18.30 | 5.00 | 12.80 | 8.40 | -0.82 | | Reynolds, 1986 | Active | cbt | 9 | bdi | 21.11 | 7.75 | 6.36 | 3.15 | -2.71 | | Control | other ctr | 11 | bdi | 17.09 | 6.36 | 5.77 | 4.00 | -2.19 | | Control | wl | 10 | bdi | 16.90 | 5.48 | 18.31 | 9.82 | 0.18 | | Rohde, 2004 | Active | cbt | 45 | hamd | 14.20 | 5.20 | 6.00 | 6.30 | -1.43 | | Control | other ctr | 48 | hamd | 6.00 | 6.30 | 8.30 | 5.40 | 0.39 | | Rohde, 2014a | Active | cbt | 126 | ksads | 1.37 | 0.35 | 1.40 | 0.32 | 0.09 | | Control | other ctr | 124 | ksads | 1.38 | 0.36 | 1.50 | 0.41 | 0.31 | | Rossello, 1999 | Active | cbt | 25 | cdi | 20.12 | 6.95 | 13.28 | 7.61 | -0.94 | | Active | ipt | 23 | cdi | 21.21 | 7.53 | 10.79 | 6.51 | -1.48 | | Control | wl | 23 | cdi | 20.13 | 5.99 | 15.83 | 6.83 | -0.67 | | Sanford, 2006 | Active | other psy | 16 | rads |  |  |  |  |  | | Control | cau | 15 | rads |  |  |  |  |  | | Santomauro, 2016 | Active | cbt | 11 | dass | 24.20 | 8.97 | 17.20 | 8.95 | -0.78 | | Control | wl | 12 | dass | 21.60 | 11.03 | 23.93 | 11.58 | 0.21 | | Shomaker, 2016 | Active | cbt | 61 | ces\_d | 25.30 | 7.30 | 13.40 | 7.30 | -1.63 | | Control | other ctr | 58 | ces\_d | 24.50 | 7.50 | 13.80 | 7.50 | -1.43 | | Srivastava, 2020 | Active | cbt | 11 | cdrs | 67.80 | 3.10 | 33.70 | 11.20 | -4.77 | | Control | cau | 10 | cdrs | 66.90 | 3.40 | 45.30 | 9.70 | -3.30 | | Stallard, 2012 | Active | cbt | 392 | smfq | 10.64 | 4.91 | 8.22 | 6.45 | -0.43 | | Control | other ctr | 374 | smfq | 10.60 | 4.67 | 8.50 | 5.88 | -0.40 | | Control | cau | 298 | smfq | 10.56 | 4.93 | 6.81 | 5.70 | -0.71 | | Stark, 1987 | Active | pst | 10 | cdrs | 33.50 | 10.27 | 24.16 | 6.01 | -1.15 | | Active | other psy | 9 | cdrs | 37.22 | 8.36 | 22.91 | 4.36 | -2.25 | | Control | wl | 9 | cdrs | 30.33 | 6.28 | 28.15 | 6.21 | -0.35 | | Stice, 2008 | Active | cbt | 80 | bdi | 18.20 | 7.53 | 14.25 | 8.98 | -0.48 | | Active | cbt | 89 | bdi | 20.03 | 10.35 | 10.77 | 9.04 | -0.96 | | Active | sup | 88 | bdi | 20.27 | 9.83 | 14.67 | 10.62 | -0.55 | | Control | cau | 84 | bdi | 19.60 | 9.23 | 16.71 | 9.74 | -0.30 | | Stikkelbroek, 2020 | Active | cbt |  | ksads |  |  |  |  |  | | Control | cau |  | ksads |  |  |  |  |  | | Szigethy, 2007 | Active | cbt | 22 | cdi | 25.70 | 10.80 | 10.70 | 8.00 | -1.60 | | Control | cau | 19 | cdi | 21.80 | 8.10 | 16.70 | 11.10 | -0.53 | | Tang, 2009 | Active | ipt | 35 | bdi | 32.66 | 10.06 | 19.97 | 14.68 | -1.03 | | Control | cau | 38 | bdi | 32.32 | 8.70 | 31.58 | 12.01 | -0.07 | | Topooco, 2018 | Active | cbt | 33 | bdi | 33.10 | 9.40 | 19.90 | 7.20 | -1.59 | | Control | other ctr | 37 | bdi | 32.30 | 10.20 | 25.20 | 7.80 | -0.79 | | Topooco, 2019 | Active | cbt | 35 | bdi | 31.60 | 10.00 | 16.00 | 11.30 | -1.46 | | Control | other ctr | 35 | bdi | 28.80 | 7.90 | 24.80 | 10.40 | -0.44 | | Vostanis, 1996 | Active | cbt | 29 | mfq | 33.40 | 12.20 | 17.60 | 5.20 | -1.82 | | Control | other ctr | 28 | mfq | 28.60 | 14.40 | 18.40 | 15.80 | -0.68 | | Weisz, 1997 | Active | cbt | 16 | cdrs | 45.25 | 16.01 | 33.19 | 10.86 | -0.90 | | Control | cau | 32 | cdrs | 38.38 | 11.15 | 34.94 | 10.93 | -0.31 | | Weisz, 2009 | Active | cbt | 32 | cdi | 10.88 | 7.91 | 8.00 | 6.32 | -0.40 | | Control | cau | 25 | cdi | 11.29 | 7.93 | 8.47 | 8.44 | -0.34 | | Wood, 1996 | Active | cbt | 24 | mfq | 25.80 | 9.80 | 14.90 | 10.40 | -1.08 | | Control | other ctr | 24 | mfq | 28.70 | 11.60 | 19.80 | 13.30 | -0.71 | | Young, 2006 | Active | ipt | 27 | ces\_d | 24.10 | 6.90 | 6.40 | 4.80 | -3.03 | | Control | other ctr | 14 | ces\_d | 27.40 | 6.60 | 17.40 | 10.50 | -1.17 | | Young, 2010 | Active | ipt |  | cdrs |  |  |  |  |  | | Control | cau |  | cdrs |  |  |  |  |  | | Young, 2016 | Active | ipt | 95 | ces\_d | 15.51 | 8.52 | 11.12 | 8.57 | -0.51 | | Control | other ctr | 91 | ces\_d | 15.07 | 8.65 | 12.62 | 9.28 | -0.27 | | Yu, 2002 | Active | cbt | 104 | cdi | 17.44 | 9.47 | 13.64 | 9.01 | -0.41 | | Control | cau | 116 | cdi | 16.72 | 9.29 | 16.02 | 10.16 | -0.07 | |

## Metaregression sensitivity analyses

|  |
| --- |
| Figure 3: Meta-analytic estimates of within-group changes: waitlist studies excluded |

|  |
| --- |
| Figure 4: Meta-analytic estimates of within-group changes: subclinical studies excluded |

|  |
| --- |
| Figure 5: Meta-analytic estimates of within-group changes: CDRS studies only |

|  |
| --- |
| Figure 6: Meta-analytic estimates of within-group changes: HAM-D studies only |

#### Effect of standard errors of the SMDs

It could be argued that the choice of standard errors of the changes for the calculation of the confidence intervals could have affect the results in one or the other direction. To address such concerns we have simulated 1000 different datasets with SMDs coming from a broad distribution. If standard error distributions were influential, this should show up as substantial variability across simulations. We test this idea in the [Figure 7](#fig-stab-sims) which displays across the 1000 simulations the z-value of the contrast between medication and psychotherapy control arms (the mean of which we presented in [Table 2](#tbl-results-overall)). As can be seen, the variability in the z-score is minimal and consistently far away from the threshold for significance, i.e. the value of z = 1.645.

|  |
| --- |
| Figure 7: Stability of the Statistic of the Difference between Medication and Psychotherapy Control Arms |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 5: HAM-D Baseline Scores across Psychotherapy and Medication RCTs   | **Subgroup** | **K** | **Mean** | **SE** | **Lower CI** | **Upper CI** | **T2** | **p-value** | | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  | 0.025 | | Psychotherapy | 10 | 12.6 | 2.16 | 7.72 | 17.48 | 46.04 |  | | Medication | 5 | 18.89 | 1.8 | 13.9 | 23.89 | 15.65 |  | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 6: CDRS Baseline Scores across Psychotherapy and Medication RCTs   | **Subgroup** | **K** | **Mean** | **SE** | **Lower CI** | **Upper CI** | **T2** | **p-value** | | --- | --- | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  |  | 0.15 | | Psychotherapy | 4 | 46.23 | 7.36 | 22.81 | 69.65 | 215.43 |  | | Medication | 23 | 56.89 | 0.82 | 55.19 | 58.59 | 14.94 |  | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 7: SMDs adjusted by baseline scores   | **Condition** | **Coefficient** | **SE** | **Lower CI** | **Upper CI** | | --- | --- | --- | --- | --- | | Medication Control | -0.41 | 0.10 | -0.61 | -0.21 | | Medication Active | -0.64 | 0.13 | -0.91 | -0.38 | | Psychotherapy Control | 0.76 | 0.13 | 0.50 | 1.02 | | Psychotherapy Active | 0.11 | 0.13 | -0.15 | 0.37 |  | **Condition** | **Coefficient** | **SE** | **z value** | **Lower CI** | **Upper CI** | | --- | --- | --- | --- | --- | --- | | Medication Control | -0.41 | 0.10 | -3.97 | -0.61 | -0.20 | | Medication Active | -0.23 | 0.13 | -1.74 | -0.50 | 0.03 | | Psychotherapy Control | 1.17 | 0.13 | 8.78 | 0.91 | 1.43 | | Psychotherapy Active | 0.52 | 0.13 | 3.94 | 0.26 | 0.77 | | T2 = 0.3; *I*2 = 95.9; *K* = 183.0; *R*2 = 54.1. | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 8: Comparing the intensity of the intervention between active and control arms of psychotherapy studies: waitlist studies excluded   | **Group** | **N** | **Mean** | **SD** | **Cohen's d** | **Upper CI** | **Lower CI** | **t** | **df** | **p-value** | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Number of sessions** | | | | | | | | | | | Active | 47 | 13.53 | 12.83 | 0.41 | -0.10 | 0.93 | 2.04 | 67.86 | 0.045 | | Control | 23 | 8.96 | 5.90 |  |  |  |  |  |  | | **Intensity (sessions per week)** | | | | | | | | | | | Active | 41 | 1.23 | 0.65 | 0.51 | -0.05 | 1.07 | 1.78 | 32.53 | 0.084 | | Control | 19 | 0.89 | 0.71 |  |  |  |  |  |  | | **Session length (mins)** | | | | | | | | | | | Active | 40 | 61.12 | 29.53 | 0.54 | -0.01 | 1.10 | 1.88 | 33.37 | 0.069 | | Control | 20 | 44.17 | 34.40 |  |  |  |  |  |  | | **Total intervention hours** | | | | | | | | | | | Active | 40 | 12.40 | 8.34 | 0.46 | -0.10 | 1.01 | 1.64 | 36.53 | 0.111 | | Control | 20 | 8.54 | 8.74 |  |  |  |  |  |  | |

# References

Bridge, Jeffrey A., Boris Birmaher, Satish Iyengar, Rémy P. Barbe, and David A. Brent. 2009. “Placebo Response in Randomized Controlled Trials of Antidepressants for Pediatric Major Depressive Disorder.” *American Journal of Psychiatry* 166 (1): 42–49. <https://doi.org/10.1176/appi.ajp.2008.08020247>.

Calvo, Ana, Miguel Moreno, Ana Ruiz-Sancho, Marta Rapado-Castro, Carmen Moreno, Teresa Sánchez-Gutiérrez, Celso Arango, and María Mayoral. 2014. “Intervention for Adolescents With Early-Onset Psychosis and Their Families: A Randomized Controlled Trial.” *Journal of the American Academy of Child & Adolescent Psychiatry* 53 (6): 688–96. <https://doi.org/10.1016/j.jaac.2014.04.004>.

Chaimani, A, DM Caldwell, T Li, Julian PT Higgins, and G Salanti. 2023. “Chapter 11: Undertaking Network Meta-Analyses.” In *Cochrane Handbook for Systematic Reviews of Interventions Version 6.4*, 6.4 ed. Cochrane.

Cipriani, Andrea, Xinyu Zhou, Cinzia Del Giovane, Sarah E Hetrick, Bin Qin, Craig Whittington, David Coghill, et al. 2016. “Comparative Efficacy and Tolerability of Antidepressants for Major Depressive Disorder in Children and Adolescents: A Network Meta-Analysis.” *The Lancet* 388 (10047): 881–90. <https://doi.org/10.1016/S0140-6736(16)30385-3>.

Constantino, Michael J., Diane B. Arnkoff, Carol R. Glass, Rebecca M. Ametrano, and JuliAnna Z. Smith. 2011. “Expectations.” *Journal of Clinical Psychology* 67 (2): 184–92. <https://doi.org/10.1002/jclp.20754>.

Courtney, Darren B., Priya Watson, Karolin R. Krause, Benjamin W. C. Chan, Kathryn Bennett, Meredith Gunlicks-Stoessel, Terri Rodak, Kirsten Neprily, Tabitha Zentner, and Peter Szatmari. 2022. “Predictors, Moderators, and Mediators Associated With Treatment Outcome in Randomized Clinical Trials Among Adolescents With Depression: A Scoping Review.” *JAMA Network Open* 5 (2): e2146331. <https://doi.org/10.1001/jamanetworkopen.2021.46331>.

Cuijpers, Pim, Eirini Karyotaki, Marketa Ciharova, Clara Miguel, Hisashi Noma, Yvonne Stikkelbroek, John R. Weisz, and Toshi A. Furukawa. 2021. “The Effects of Psychological Treatments of Depression in Children and Adolescents on Response, Reliable Change, and Deterioration: A Systematic Review and Meta-Analysis.” *European Child & Adolescent Psychiatry*, October. <https://doi.org/10.1007/s00787-021-01884-6>.

Cumming, Geoff. 2013. *Understanding The New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. 1st ed. Routledge. <https://doi.org/10.4324/9780203807002>.

Dechartres, Agnes, Isabelle Boutron, Ludovic Trinquart, Pierre Charles, and Philippe Ravaud. 2011. “Single-Center Trials Show Larger Treatment Effects Than Multicenter Trials: Evidence From a Meta-epidemiologic Study.” *Annals of Internal Medicine* 155 (1): 39. <https://doi.org/10.7326/0003-4819-155-1-201107050-00006>.

Del Giovane, Cinzia, Samuele Cortese, and Andrea Cipriani. 2019. “Combining Pharmacological and Nonpharmacological Interventions in Network Meta-analysis in Psychiatry.” *JAMA Psychiatry* 76 (8): 867. <https://doi.org/10.1001/jamapsychiatry.2019.0574>.

Guo, Shenyang, and Mark W. Fraser. 2014. “Chapter 2: Counterfactual Framework and Assumptions.” In *Propensity Score Analysis: Statistical Methods and Applications*. SAGE Publications.

Higgins, Julian PT, Tianjing Li, and Jonathan J Deeks. 2023. “Chapter 6: Choosing Effect Measures and Computing Estimates of Effect.” In *Cochrane Handbook for Systematic Reviews of Interventions Version 6.4*, 6.4 ed. Cochrane.

IntHout, Joanna, John Pa Ioannidis, and George F Borm. 2014. “The Hartung-Knapp-Sidik-Jonkman Method for Random Effects Meta-Analysis Is Straightforward and Considerably Outperforms the Standard DerSimonian-Laird Method.” *BMC Medical Research Methodology* 14 (1): 25. <https://doi.org/10.1186/1471-2288-14-25>.

Jaycox, Lisa H., Joan Rosenbaum Asarnow, Cathy D. Sherbourne, Margaret M. Rea, Anne P. LaBorde, and Kenneth B. Wells. 2006. “Adolescent Primary Care Patients’ Preferences for Depression Treatment.” *Administration and Policy in Mental Health and Mental Health Services Research* 33 (2): 198–207. <https://doi.org/10.1007/s10488-006-0033-7>.

Lakens, Daniel. 2013. “Calculating and Reporting Effect Sizes to Facilitate Cumulative Science: A Practical Primer for t-Tests and ANOVAs.” *Frontiers in Psychology* 4.

Lamers, F., A. T. F. Beekman, A. M. Van Hemert, R. A. Schoevers, and B. W. J. H. Penninx. 2016. “Six-Year Longitudinal Course and Outcomes of Subtypes of Depression.” *British Journal of Psychiatry* 208 (1): 62–68. <https://doi.org/10.1192/bjp.bp.114.153098>.

Langer, David A., Tessa K. Kritikos, Joan R. Asarnow, and Martha C. Tompson. 2021. “Parent and Youth Preferences in the Treatment of Youth Depression.” *Child Psychiatry & Human Development* 52 (2): 236–47. <https://doi.org/10.1007/s10578-020-01006-4>.

Lin, Yi-Hsuan, Ethan Sahker, Kiyomi Shinohara, Noboru Horinouchi, Masami Ito, Madoka Lelliott, Andrea Cipriani, Anneka Tomlinson, Christopher Baethge, and Toshi A. Furukawa. 2022. “Assessment of Blinding in Randomized Controlled Trials of Antidepressants for Depressive Disorders 2000–2020: A Systematic Review and Meta-Analysis.” *eClinicalMedicine* 50 (August): 101505. <https://doi.org/10.1016/j.eclinm.2022.101505>.

Lorenzo-Luaces, Lorenzo, Natalie Rodriguez-Quintana, and Allen J. Bailey. 2020. “Double Trouble: Do Symptom Severity and Duration Interact to Predicting Treatment Outcomes in Adolescent Depression?” *Behaviour Research and Therapy* 131 (August): 103637. <https://doi.org/10.1016/j.brat.2020.103637>.

March, John, Susan Silva, Stephen Petrycki, John Curry, Karen Wells, John Fairbank, Barbara Burns, et al. 2004. “Fluoxetine, Cognitive-Behavioral Therapy, and Their Combination for Adolescents With Depression: Treatment for Adolescents With Depression Study (TADS) Randomized Controlled Trial.” *JAMA* 292 (7): 807. <https://doi.org/10.1001/jama.292.7.807>.

McHugh, R. Kathryn, Sarah W. Whitton, Andrew D. Peckham, Jeffrey A. Welge, and Michael W. Otto. 2013. “Patient Preference for Psychological Vs Pharmacologic Treatment of Psychiatric Disorders: A Meta-Analytic Review.” *The Journal of Clinical Psychiatry* 74 (06): 595–602. <https://doi.org/10.4088/JCP.12r07757>.

NICE. 2019. “Depression in Children and Young People: Identification and Management.”

Relton, Clare, Maarten Burbach, Clive Collett, James Flory, Sophie Gerlich, Soren Holm, Amanda Hunn, et al. 2017. “The Ethics of ‘Trials Within Cohorts’ (TwiCs): 2nd International Symposium: London, UK. 7-8 November 2016.” *Trials* 18 (S2): 244, s13063-017-1961-0. <https://doi.org/10.1186/s13063-017-1961-0>.

Simmonds-Buckley, Melanie, Ana Catarino, and Jaime Delgadillo. 2021. “Depression Subtypes and Their Response to Cognitive Behavioral Therapy: A Latent Transition Analysis.” *Depression and Anxiety* 38 (9): 907–16. <https://doi.org/10.1002/da.23161>.

Stone, Marc B, Zimri S Yaseen, Brian J Miller, Kyle Richardville, Shamir N Kalaria, and Irving Kirsch. 2022. “Response to Acute Monotherapy for Major Depressive Disorder in Randomized, Placebo Controlled Trials Submitted to the US Food and Drug Administration: Individual Participant Data Analysis.” *BMJ*, August, e067606. <https://doi.org/10.1136/bmj-2021-067606>.

Tröger, Anna, Clara Miguel, Marketa Ciharova, Nino De Ponti, Güldehan Durman, Pim Cuijpers, and Eirini Karyotaki. 2024. “Baseline Depression Severity as Moderator on Depression Outcomes in Psychotherapy and Pharmacotherapy.” *Journal of Affective Disorders* 344 (January): 86–99. <https://doi.org/10.1016/j.jad.2023.10.047>.

Weitz, Erica S., Steven D. Hollon, Jos Twisk, Annemieke Van Straten, Marcus J. H. Huibers, Daniel David, Robert J. DeRubeis, et al. 2015. “Baseline Depression Severity as Moderator of Depression Outcomes Between Cognitive Behavioral Therapy Vs Pharmacotherapy: An Individual Patient Data Meta-analysis.” *JAMA Psychiatry* 72 (11): 1102. <https://doi.org/10.1001/jamapsychiatry.2015.1516>.

Zhou, Xinyu, Teng Teng, Yuqing Zhang, Cinzia Del Giovane, Toshi A Furukawa, John R Weisz, Xuemei Li, et al. 2020. “Comparative Efficacy and Acceptability of Antidepressants, Psychotherapies, and Their Combination for Acute Treatment of Children and Adolescents with Depressive Disorder: A Systematic Review and Network Meta-Analysis.” *The Lancet Psychiatry* 7 (7): 581–601. <https://doi.org/10.1016/S2215-0366(20)30137-1>.