The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size

Cuijpers P, van Straten A, Bohlmeijer E, Hollon SD, Andersson G

CRD summary
The review concluded that while the effects of psychotherapy for adult depression were significant, they had been overestimated in meta-analytic studies. The authors’ conclusion reflected the results, but given that the classification of studies based on quality appeared to be arbitrary, some caution is warranted.

Authors’ objectives
To determine the impact of study quality on the effect size of psychotherapy for adult depression.

Searching
A database (Psychotherapy - randomised controlled and comparative trials) for studies on psychological treatment of depression in general that was developed from a previous search of PubMed, PsycINFO, EMBASE and Cochrane Central Register of Controlled Trials (CENTRAL) (1966 through to December 2007) was searched. Search terms were reported (see URL For Additional Data). Dissertation Abstracts International, primary studies from relevant reviews and references from included studies were checked. No language restrictions were applied.

Study selection
Randomised control trials (RCTs) were eligible for inclusion if they compared psychological treatments with wait list, care as usual, oral placebo or psychological placebo in adults (18 years or greater) with a depressive disorder or elevated level of depressive symptomatology. Psychological treatments were defined as interventions in which the core element was verbal communication between therapist and patient or bibliotherapy with some personal support (such as telephone or email) from a therapist. Comorbid medical or psychiatric disorders were allowed. Studies aimed at maintenance treatments and relapse prevention, studies of therapies in which the psychological intervention could not be assessed independently from other elements of the intervention (such as managed care interventions and disease management programmes) and studies in which the standardised effect size could not be calculated were excluded.

Participants in more than half the studies met diagnostic criteria for a mood disorder; another definition of depression (typically a high score on a self-report questionnaire) was used in the other studies. For high-quality studies only, interventions included cognitive-behavioural therapy (with or without depression recurrence prevention programme), psychodynamic therapy, counselling, behavioural activation, problem-solving therapy, self-management therapy and interpersonal psychotherapy. Most studies involved individual rather than group therapy. Diagnoses ranged from major depressive disorder (with or without atypical features), minor depressive disorder, adjusted disorder and dysthymia. Participants included women with postpartum depression, adults in general, male veterans with chronic combat-related post-traumatic stress disorder and depressive disorder, low-income, young minority women and older adults. Beck Depression Inventory (BDI) and Hamilton Rating Scale for Depression (HAMD) were the most frequently used rating scales, others included Edinburgh Postnatal Depression Scale, Montgomery-Asberg Depression Rating Scale, Geriatric Depression Rating Scale and Hopkins Symptom Checklist Depression Scale.

The authors did not state how papers were selected for the review.

Assessment of study quality
The quality of included studies was assessed using eight criteria: participants met diagnostic criteria for a depressive disorder; study referred to use of a treatment manual; trained therapist used; treatment integrity checked during study; intention-to-treat analysis; sufficient statistical power; randomisation conducted by independent party; and outcome assessors blinded. A study was considered to be of high quality if it met all eight criteria.

Two reviewers independently assessed study quality. Any disagreements were resolved by discussion.
Data extraction
Data required for calculation of effect sizes (standardised mean difference) were extracted for each study. Only instruments that explicitly measured depression were used in these calculations.

The authors did not state how many reviewers performed data extraction.

Methods of synthesis
Pooled mean effect sizes with 95% confidence intervals (CI) were calculated using a random-effects model. Numbers needed to treat (NNT) were calculated. Statistical heterogeneity was assessed using Q statistic and $I^2$ statistic. Subgroup analysis (mixed effects analysis) and multivariate meta-regression were conducted.

Results of the review
The review included 115 studies with 178 comparisons (n=8,140). Eleven studies (16 comparisons) were considered to be of high quality. Eight interventions examined cognitive-behavioural therapies. Thirteen studies used an individual treatment format (no intervention used a self-help format).

Overall mean effect size of all comparisons was 0.68 (95% CI 0.60 to 0.76, $I^2=70.3\%$). The mean effect size for high quality studies ($d=0.22$) was significantly smaller than in other studies ($d=0.74$, $p<0.001$). NNT in the high-quality studies was eight; NNT in lower quality studies was two. No evidence of heterogeneity was found in the high-quality studies; substantial heterogeneity was found in lower quality studies ($I^2=73.5\%$).

Limiting the sample to Beck Depression Inventory or Hamilton Rating Scale for Depression, or to a subset of the other studies using care-as-usual or non-specific controls similar to those found in the higher quality studies did not substantially alter the results.

The effect size was significantly lower in studies with intention to treat analysis, in studies with a sample size of 50 or more, studies in which randomisation was done by an independent party and studies in which assessors of outcomes did not know to which condition the respondents were assigned. Results of the meta-regression were reported.

Authors' conclusions
Although the effects of psychotherapy for adult depression were significant, they had been overestimated in meta-analytic studies even after controlling for the type of control condition used.

CRD commentary
The review question was defined in terms of participants, study design and intervention. Attempts were made to identify relevant evidence from multiple sources (which included unpublished studies) regardless of language, which minimised the likelihood of publication bias. With the exception of study quality, the authors did not report whether sufficient attempts to minimise errors and bias were made during study selection and data extraction. The validity of studies was assessed, but individual results were not reported. Some study characteristics were provided and the methods of synthesis appeared appropriate. Multiple comparisons were included and this was considered by the authors in their analysis. The authors acknowledged a number of limitations, these included that some quality criteria were not considered (such as attrition), that definition of high quality in others could be considered arbitrary and the small number of studies that met all the quality criteria.

The authors' conclusion reflected the results, but given that the classification of studies based on quality appeared to be arbitrary, some caution is warranted.

Implications of the review for practice and research
Practice: The authors did not state any implications for practice.

Research: The authors indicated that there was a need to develop more powerful interventions or enhance the power of interventions already available.
Funding
None reported.

Bibliographic details

PubMedID
19490745

DOI
10.1017/S0033291709006114

Original Paper URL
http://journals.cambridge.org/action/displayAbstract?aid=7018756

Additional Data URL
www.evidencebasedpsychotherapies.org

Indexing Status
Subject indexing assigned by NLM

MeSH
Adult; Cohort Effect; Depressive Disorder /diagnosis /therapy; Humans; Psychotherapy /methods; Randomized Controlled Trials as Topic; Severity of Illness Index

AccessionNumber
12010003413

Date bibliographic record published
04/08/2010

Date abstract record published
01/12/2010

Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.