Cochrane Library

Cochrane Database of Systematic Reviews

Control interventions in randomised trials among people with mental health disorders (Review)

Faltinsen E, Todorovac A, Staxen Bruun L, Hróbjartsson A, Gluud C, Kongerslev MT, Simonsen E, Storebø OJ

Faltinsen E, Todorovac A, Staxen Bruun L, Hróbjartsson A, Gluud C, Kongerslev MT, Simonsen E, Storebø OJ.
Control interventions in randomised trials among people with mental health disorders.
Cochrane Database of Systematic Reviews 2022, Issue 4. Art. No.: MR000050.
DOI: 10.1002/14651858.MR000050.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Methodology Review]

Control interventions in randomised trials among people with mental health disorders

Erlend Faltinsen[1,2,3a], Adnan Todorovac[1b], Laura Staxen Bruun[1], Asbjørn Hróbjartsson[3,4], Christian Gluud[5,6], Mickey T Kongerslev[1,7], Erik Simonsen[1,8], Ole Jakob Storebø[1,9]

[1]Psychiatric Research Unit, Region Zealand Psychiatry, Slagelse, Denmark. [2]Centre for Evidence-Based Medicine Odense (CEBMO) and Cochrane Denmark, Department of Clinical Research, University of Southern Denmark, Odense, Denmark. [3]Open Patient data Exploratory Network (OPEN), Odense University Hospital, Odense, Denmark. [4]Centre for Evidence-Based Medicine Odense (CEBMO) and Cochrane Denmark, Department of Clinical Research, University of Southern Denmark, Odense, Denmark. [5]Copenhagen Trial Unit, Centre for Clinical Intervention Research, The Capital Region, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. [6]Department of Regional Health Research, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark. [7]Department of Psychology, University of Copenhagen, Copenhagen, Denmark. [8]Institute of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [9]Department of Psychology, University of Southern Denmark, Odense, Denmark

[a]Shared first authorship. [b]Shared first authorship

Contact: Erlend Faltinsen, egfaltinsen@health.sdu.dk.

Editorial group: Cochrane Methodology Review Group.
Publication status and date: New, published in Issue 4, 2022.

Citation: Faltinsen E, Todorovac A, Staxen Bruun L, Hróbjartsson A, Gluud C, Kongerslev MT, Simonsen E, Storebø OJ.Control interventions in randomised trials among people with mental health disorders. *Cochrane Database of Systematic Reviews* 2022, Issue 4. Art. No.: MR000050. DOI: 10.1002/14651858.MR000050.pub2.

ABSTRACT

Background

Control interventions in randomised trials provide a frame of reference for the experimental interventions and enable estimations of causality. In the case of randomised trials assessing patients with mental health disorders, many different control interventions are used, and the choice of control intervention may have considerable impact on the estimated effects of the treatments being evaluated.

Objectives

To assess the benefits and harms of typical control interventions in randomised trials with patients with mental health disorders. The difference in effects between control interventions translates directly to the impact a control group has on the estimated effect of an experimental intervention. We aimed primarily to assess the difference in effects between (i) wait-list versus no-treatment, (ii) usual care versus wait-list or no-treatment, and (iii) placebo interventions (all placebo interventions combined or psychological, pharmacological, and physical placebos individually) versus wait-list or no-treatment. Wait-list patients are offered the experimental intervention by the researchers after the trial has been finalised if it offers more benefits than harms, while no-treatment participants are not offered the experimental intervention by the researchers.

Search methods

In March 2018, we searched MEDLINE, PsycInfo, Embase, CENTRAL, and seven other databases and six trials registers.

Selection criteria

We included randomised trials assessing patients with a mental health disorder that compared wait-list, usual care, or placebo interventions with wait-list or no-treatment .

Data collection and analysis

Titles, abstracts, and full texts were reviewed for eligibility. Review authors independently extracted data and assessed risk of bias using Cochrane's risk of bias tool. GRADE was used to assess the quality of the evidence. We contacted researchers working in the field to ask for data from additional published and unpublished trials.

A pre-planned decision hierarchy was used to select one benefit and one harm outcome from each trial. For the assessment of benefits, we summarised continuous data as standardised mean differences (SMDs) and dichotomous data as risk ratios (RRs). We used risk differences (RDs) for the assessment of adverse events. We used random-effects models for all statistical analyses. We used subgroup analysis to explore potential causes for heterogeneity (e.g. type of placebo) and sensitivity analyses to explore the robustness of the primary analyses (e.g. fixed-effect model).

Main results

We included 96 randomised trials (4200 participants), ranging from 8 to 393 participants in each trial. 83 trials (3614 participants) provided usable data. The trials included 15 different mental health disorders, the most common being anxiety (25 trials), depression (16 trials), and sleep-wake disorders (11 trials).

All 96 trials were assessed as high risk of bias partly because of the inability to blind participants and personnel in trials with two control interventions. The quality of evidence was rated low to very low, mostly due to risk of bias, imprecision in estimates, and heterogeneity.

Only one trial compared wait-list versus no-treatment directly but the authors were not able to provide us with any usable data on the comparison.

Five trials compared usual care versus wait-list or no-treatment and found a SMD –0.33 (95% CI -0.83 to 0.16, I^2 = 86%, 523 participants) on benefits.

The difference between all placebo interventions combined versus wait-list or no-treatment was SMD −0.37 (95% CI −0.49 to −0.25, I^2 = 41%, 65 trials, 2446 participants) on benefits. There was evidence of some asymmetry in the funnel plot (Egger's test P value of 0.087). Almost all the trials were small. Subgroup analysis found a moderate effect in favour of psychological placebos SMD -0.49 (95% CI −0.64 to −0.30; I^2 = 53%, 39 trials, 1656 participants). The effect of pharmacological placebos versus wait-list or no-treatment on benefits was SMD -0.14 (95% CI −0.39 to 0.11, 9 trials, 279 participants) and the effect of physical placebos was SMD −0.21 (95% CI −0.35 to −0.08, I^2 = 0%, 17 trials, 896 participants). We found large variations in effect sizes in the psychological and pharmacological placebo comparisons. For specific mental health disorders, we found significant differences in favour of all placebos for sleep-wake disorders, major depressive disorder, and anxiety disorders, but the analyses were imprecise due to sparse data.

We found no significant differences in harms for any of the comparisons but the analyses suffered from sparse data.

When using a fixed-effect model in a sensitivity analysis on the comparison for usual care versus wait-list and no-treatment, the results were significant with an SMD of –0.46 (95 % CI –0.64 to –0.28). We reported an alternative risk of bias model where we excluded the blinding domains seeing how issues with blinding may be seen as part of the review investigation itself. However, this did not markedly change the overall risk of bias profile as most of the trials still included one or more unclear bias domains.

Authors' conclusions

We found marked variations in effects between placebo versus no-treatment and wait-list and between subtypes of placebo with the same comparisons. Almost all the trials were small with considerable methodological and clinical variability in factors such as mental health population, contents of the included control interventions, and outcome domains. All trials were assessed as high risk of bias and the evidence quality was low to very low.

When researchers decide to use placebos or usual care control interventions in trials with people with mental health disorders it will often lead to lower estimated effects of the experimental intervention than when using wait-list or no-treatment controls. The choice of a control intervention therefore has considerable impact on how effective a mental health treatment appears to be. Methodological guideline development is needed to reach a consensus on future standards for the design and reporting of control interventions in mental health intervention research.

PLAIN LANGUAGE SUMMARY

Control interventions in randomised trials for people with a mental health disorder

This systematic review assesses the effects of different control interventions in randomised trials including patients with a mental health disorder. In randomised trials, patients are assigned by chance to one of two or more groups – usually an experimental intervention and a control intervention. There are many types of control interventions in mental health intervention research. Some of the most common are different types of placebos that lack what is assumed to be the active component in the experimental intervention, and usual care, where patients receive the standard treatment for their mental health disorder in the area where they live. Two other types of control interventions are wait-list or no-treatment where patients receive no trial-related care during the study (although some patients may receive care outside

the studies). Wait-list patients are often offered the experimental intervention after the trial has been finalised if it is likely to provide more benefits than harms, while no-treatment participants are not offered the experimental intervention by the researchers.

We searched for randomised trials with patients with mental health disorders where wait-list, usual care, or placebo interventions were compared with either wait-list or no-treatment. We looked at differences between all the types of control interventions on beneficial effects and whether they caused any adverse effects. We included 96 trials with a total of 4200 participants. Only 83 trials (3614 participants) provided usable data. Fifteen different mental health disorders were included. We found that all the trials were at high risk of bias in how they had been conducted, which reduced the interpretability of our findings. However, the risk of bias was mostly due to lack of blinding in the placebo studies, which may be seen as an aspect of the review's methodological question rather than a flaw with the review itself. We found no clinically important differences for usual care or wait-list control interventions in the main analyses, however in our secondary analyses we found a clinically important favourable difference for usual care. In general, placebo control interventions tended to be favourable over no-treatment or wait-list control interventions across mental health disorders. We found no clinically important differences on adverse events.

This review suggests that different control interventions have a tendency to yield very different estimates for the effects of the experimental intervention and that the choice of control intervention has a large impact on how effective a mental health treatment appears to be. Control interventions in trials with patients with mental health disorders are often poorly reported upon, and guidelines are needed to inform researchers on how to properly design, report, and interpret these trials.

SUMMARY OF FINDINGS

Summary of findings 1. Summary of findings

Benefits and harms of wait-list compared with no-treatment for mental health disorders

Patient or population: patients with mental health disorders

Settings: inpatient and outpatient

Intervention: wait-list

Comparison: no-treatment

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Wait-list/no-treatment	Usual care				
Wait-list compared with no-treatment						Only one cluster-randomised trial compared a wait-list intervention to a no-treatment intervention was included (Howlin 2007). However, no usable data were provided in the full report, and the authors did engage in correspondence. Eighty-four elementary school children with a autism spectrum disorder were randomised to either, i) immediate treatment, ii) delayed treatment (wait-list), and iii) no-treatment. Conclusions were that Picture Exchange Communication System (PECS) training indicated modest effectiveness for children with autism spectrum disorder. In general there were no differences on across outcome measures between the wait-list and no-treatment intervention groups.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **RR:** Risk Ratio; **RD:** Risk Difference; **RCT:** Randomised clinical trial

GRADE Working Group grades of evidence
High quality: further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low quality: we are very uncertain about the estimate.

Summary of findings 2. Summary of findings

Benefits and harms of usual care compared with wait-list or no-treatment for mental health disorders

Patient or population: patients with mental health disorders

Settings: inpatient and outpatient

Intervention: usual care

Comparison: wait-list or no-treatment

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Wait-list/no-treatment	Usual care				
Usual care compared with wait-list/no-treatment (Variety of continuous outcome) (Post-treatment)		The mean score in the usual care group was **0.33 points lower** (0.83 lower to 0.16 higher)		523 (5 RCTs)	⊕⊖⊖⊖ **very low** [a,b,c]	TSA adjusted CI = -2.32 to 1.15 TSA RIS = 1536
Serious adverse events for all placebos						No data
Non-serious adverse events for all placebos						No data

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: confidence interval; **RR:** risk ratio; **RD:** risk difference; **RCT:** randomised clinical trial; **TSA:** Trial Sequential Analysis

GRADE Working Group grades of evidence
High quality: further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: we are very uncertain about the estimate.

[a] We downgraded the quality of evidence by one level due to risk of bias

[b] We downgraded the quality of evidence by one level due to inconsistency (in terms of either clinical and methodological heterogeneity)

[c] We downgraded the quality of evidence by one level due to imprecision (wide confidence intervals)

Summary of findings 3. Summary of findings

Benefits and harms of placebos compared with wait-list or no-treatment for mental health disorders

Patient or population: patients with mental health disorders

Settings: inpatient and outpatient

Intervention: all placebos combined, psychological, pharmacological and physical placebos

Comparison: wait-list or no-treatment

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Wait-list/no-treatment	Placebos				
All placebos compared with wait-list or no-treatment (Variety of continuous outcome) (Post-treatment)		The mean score in the placebo group was **0.37 points lower** (0.49 lower to 0.25 lower)		2446 (65 RCTs)	⊕⊕⊖⊖ **low** [a,b]	TSA adjusted CI = −1.85 to −0.84 TSA RIS = 397
Psychological placebos compared with wait-list or no-treatment (Variety of continuous outcome) (Post-treatment)		The mean score in the placebo group was **0.49 points lower** (0.66 lower to 0.31 lower)		1263 (38 RCTs)	⊕⊕⊖⊖ **low** [a,b]	TSA adjusted CI = −2.54 to −1.02 TSA RIS = 454
Pharmacological placebos compared with wait-list or no-treatment (Variety of continuous outcome) (Post-treatment)		The mean score in the placebo group was **0.14 points lower** (0.39 lower to 0.11 higher)		279 (9 RCTs)	⊕⊖⊖⊖ **very low** [a,b,c]	TSA adjusted CI = −9.43 to 6.15 TSA RIS = 229

Physical placebos compared with wait-list or no-treatment (Variety of continuous outcome) (Post-treatment)		The mean score in the placebo group was **0.21 points lower** (0.35 lower to 0.08 lower)		896 (17 RCTs)	⊕⊕⊖⊖ **low** [a,b]	TSA adjusted CI = −3.64 to −0.49 TSA RIS = 194
Serious adverse events for all placebos compared with wait-list or no treatment (Spontaneous reporting of dichotomous outcomes) (Post-treatment)	**43 per 1000**	**27 per 1000 (95% CI 32 fewer to 23 higher)**	**RD -0.00** (95% CI -0.03 to 0.03)	517 (11 RCTs)	⊕⊖⊖⊖ **very low a,b,c**	Not possible to calculate TSA on serious adverse events due to too little information use
Non-serious adverse events for all placebos compared with wait-list or no treatment (Spontaneous reporting of dichotomous outcomes) (Post-treatment)	**93 per 1000**	**96 per 1000** **(95% CI 2 fewer to 7 higher)**	**RD 0.03** (95% CI -0.02 to 0.08)	590 (14 trials)	⊕⊖⊖⊖ **very low a,b,c**	Not possible to calculate TSA on serious adverse events due to too little information use

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: Confidence interval; **RR:** Risk Ratio; **RD:** Risk Difference; **RCT:** Randomised clinical trial;**RIS:** required information size; **TSA:** trial sequential analysis

GRADE Working Group grades of evidence
High quality: further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: we are very uncertain about the estimate.

[a] We downgraded the quality of evidence by one level due to risk of bias

[b] We downgraded the quality of evidence by one level due to inconsistency (in terms of either clinical and methodological heterogeneity)

[c] We downgraded the quality of evidence by one level due to imprecision (wide confidence intervals)

BACKGROUND

Description of the methods being investigated

Control interventions in randomised trials provide a frame of reference for the experimental intervention and allow causal estimations of treatment efficacy and adverse events (Higgins 2019; Kazdin 2016; Sibbald 1998). This systematic review assesses the benefits and harms of different control interventions in randomised trials that include patients with a mental health disorder.

We included the following experimental interventions in the review (which are often described as control interventions — see Table 1): (a) wait-list, (b) usual care, (c) psychological placebo, (d) pharmacological placebo, and (e) physical placebo. We included the following control interventions: (a) wait-list and (b) no-treatment. We also planned to compare wait-list with no-treatment. We conducted analyses across all included patient populations and within specific mental health disorders. We made direct comparisons between the control interventions by including trials with more than one control arm (often three-armed randomised trials).

Wait-list participants are typically assessed before and after a given time period, and they receive the experimental intervention after the final research assessment if it provides more benefits than harms. *No-treatment* participants are also assessed on repeated occasions but are not promised the experimental intervention after the final assessment (Comer 2013). Furukawa and colleagues have proposed that wait-list participants could become motivated to remain in poor health in order to receive a desired therapy after the trial has ended, and that those receiving no-treatment might actively seek out other forms of care outside the trial during the trial period (Furukawa 2014). Wait-list participants could therefore be subject to so-called nocebo effects (i.e. negative effects from inert interventions) (Colloca 2020), but the evidence on this is preliminary (Greville-Harris 2015; Furukawa 2014). Wait-list and no-treatment comparators control for maturation, spontaneous improvement, regression to the mean, and observer-expectancy effects (Comer 2013 ; Kienle 1997). Careful monitoring of participants in wait-list and no-treatment interventions is important to ensure toleration of treatment delays and ethical compliance Comer 2013; Mohr 2009).

Usual care (sometimes also referred to as treatment as usual) is a control intervention that attempts to mirror the locally accepted treatment practices for a given mental health disorder. This control intervention may include both pharmacological and psychological treatments that are administered by relevant practitioners (Freedland 2011). The research teams are often not involved in the care of these patients. Usual care control groups are typically subject to large clinical and methodological heterogeneity, the practitioners receive little supervision, and the interventions often use a mixture of different theoretical approaches (Comer 2013 ; Kazdin 2015 ; Löfholm 2013). Despite these issues, usual care arguably reflects routine practice better than highly controlled psychiatric interventions (Kazdin 2015; Mohr 2014) and when delivered well this type of control intervention is useful for determining whether novel psychiatric treatments are favourable to current practices (Mohr 2009). Usual care is sometimes standardised (Bateman 2009; Chanen 2008), which may involve manualisation, optimising of treatment structure, and adherence procedures (e.g. through supervision) (Bateman 2017; Cristea 2017; Kongerslev 2015).

This Cochrane methodology review distinguishes between three types of placebos. First, *psychological placebos* are designed to target the shared components of psychological treatments, such as attending sessions, the therapeutic relationship and patient expectations (Frank 1991; Hróbjartsson 2012; Rosenzweig 1936). It is both methodologically and theoretically difficult to discriminate between psychological placebos and psychological treatments (Borkovec 2005; Hróbjartsson 2012; Locher 2018; Mohr 2014; Wampold 2010; Wampold 2016). However, psychological placebos can be methodologically useful for differentiating between the proposed active and non-active components in psychological treatments (Mohr 2009). Second, *pharmacological placebos* are inert substances in pill, liquid or other forms that do not contain the active ingredients of a given pharmacological treatment. Participants typically receive a pill containing starch, sugar, or lactose (Double 1993; Meissner 2011). The pharmacological placebo will need to match the active drug treatment (e.g. antidepressant medication) in size, form, colour, weight, smell, texture, solubility and taste, but not include any of the active components in the experimental intervention (Wager 2015). Third, *physical placebos* target the inert components of physical treatments (e.g. acupuncture, exercise regimens, or surgery). Here an example could be a staged electromagnetic stimulation procedure where the machine is not turned on or electrodes are attached to inactive sites (Sommer 2006).

Why it is important to do this review

The need to improve and develop treatments for mental health disorders is great (Holmes 2018; Karterud 2020; Leichsenring 2019; Weisz 2019). The type of comparator used in randomised trials with patients with mental health disorders may influence estimates of the effects of the experimental intervention, and it is important to know comparative benefits and harms of different types of comparator. However, these is a lack of consensus on how to design and report control interventions in randomised trials with these patients and evidence-based guidelines are needed (Erlen 2015; Freedland 2011; Gold 2018; Kube 2017; Lund 2014; Mohr 2009). One aim of this review was to provide an empirical basis for future methodological guideline development in this field (Hoffmann 2013; Tajika 2015).

Wait-list control and no-treatment interventions may yield different effects in favour of experimental treatments depending on how they are structured, designed, and delivered, and it is very important to describe such factors. Wait-list and no-treatment conditions are also some of the most commonly used control interventions in psychiatric research (Mohr 2014) but may induce unwanted adverse events in participants, for instance from waiting to receive a treatment that patients may critically need (Furukawa 2014). If participants allocated to wait-list and no-treatment interventions show significantly more adverse events than those allocated to other control interventions, the ethical concerns and risks of overestimating the effects of clinical interventions in randomised trials should also be investigated (Cunningham 2013; Furukawa 2014).

We need more evidence on the content and effects of usual care as a control intervention in randomised trials (Rosenberg 2014; Swanson 2014) given the lack of discussion on how to design

usual care conditions properly and how the use of usual care as a control condition may influence the reported effects (e.g. in favour of experimental treatments in study reports) (Kazdin 2015).

In a series of prior reviews, Hróbjartsson and Gøtzsche found, in general, no clinically important effects of psychological, pharmacological, and physical placebos versus wait-list and no-treatment interventions for various medical and psychiatric conditions (Hróbjartsson 2001; Hróbjartsson 2002; Hróbjartsson 2004; Hróbjartsson 2010). For example, the most recent update of their review (2010) included 44 trials with dichotomous outcomes and 158 trials with continuous outcomes, and they found moderate heterogeneity for both outcome domains (I^2 = 45% and 42%, respectively). For continuous outcomes, they also found large variation in effects between small and large trials (asymmetric funnel plots). Although the design of Hróbjartsson's and Gøtzsche's reviews is similar to the design of this review, their objective was to investigate the clinical relevancy of placebos, whereas this review is focused on methodological questions related to control interventions in randomised trials with patients with mental health disorders. It is, however, relevant to compare the two reviews methodologically. The present review is also interested in how placebo interventions may depend on factors such as type of mental health disorder, context of administration, information given to participants, and type of outcome measure (Charlesworth 2017; Fässler 2015; Holmes 2016; Hróbjartsson 2010; Howick 2019; Jensen 2017; Meissner 2011; O'Leary 1978; Rutherford 2014; Vase 2019Walach 2011; We 2012; Weimer 2015; Yeung 2017), which were also investigated by Hróbjartsson and Gøtzsche.

This review is based on our published protocol (Faltinsen 2019).

OBJECTIVES

Our objectives were to assess the comparative benefits and harms of different control interventions used in randomised trials with patients with mental health disorders. We specifically wanted to assess whether different control interventions yield different effect estimates compared with wait-list or no-treatment. We included the most common control interventions in mental health intervention research: wait-list, usual care and placebos and compared these with wait-list or no-treatment. We also wanted to compare wait-list with no-treatment interventions.

We compared the following interventions:

1. wait-list versus no-treatment interventions;
2. usual care versus wait-list or no-treatment interventions;
3. all placebos combined, psychological, pharmacological, and physical placebos versus wait-list or no-treatment interventions.

METHODS

Criteria for considering studies for this review

Types of studies

Randomised trials comparing wait-list, usual care, or placebo interventions with either wait-list or no-treatment interventions were eligible. Parallel trials irrespective of language, publication year, and publication type were eligible. We included one cross-over trial, but only used data from the first phase of the trial as a regular parallel trial. We included one cluster-randomised trial. In case of articles published in languages other than English, we sought translation of the relevant sections. Unpublished studies where methods and results could be assessed in written form were eligible.

Types of data

All patients in each included trial were required to have a formal diagnosis of a mental health disorder according to the Diagnostic and Statistical Manual of Mental Disorders (DSM), First Edition (DSM-I; APA 1952), Second Edition (DSM-II; APA 1968), Third Edition (DSM-III; APA 1980), Third Edition Revised (DSM-III-R; APA 1987), Fourth Edition (DSM-IV; APA 1994), Fourth Edition Text Revision (DSM-IV-TR; APA 2000), and Fifth Edition (DSM-5; APA 2013), or according to the International Classification of Diseases and Related Health Problems (ICD), Sixth Edition (ICD-6; WHO 1949), Seventh Edition (ICD-7; WHO 1955), Eighth Edition (ICD-8; WHO 1967), Ninth Edition (ICD-9; WHO 1975), 10th Edition (ICD-10; WHO 1993), or 11th Edition (ICD-11; WHO 2018). In some instances, the diagnostic classification system was not mentioned in the full report, but the participants fulfilled all symptoms to receive a diagnosis of a mental health disorder or they were formally diagnosed by a mental health professional. For trials published before the introduction of DSM or ICD criteria in 1949, participants were eligible if they had received a formal diagnostic assessment of a mental health disorder by a health professional.

We categorised the different mental health disorders according to the current nomenclature in the DSM-5 (APA 2013). If all participants in a trial had a mental health disorder, but not the same one, we included the trial in all the analyses except those on specific mental health disorders (see Types of outcome measures). We included participants with or without comorbid conditions. Eligible participants were included irrespective of location, setting, and other demographic variables (including age).

Types of methods

Experimental interventions

We defined wait-list, usual care, and placebo interventions as any interventions that were clearly labelled or reflected the properties of wait-list, usual care, or placebo interventions, according to the criteria below (and in Table 1). We anticipated that most of the included interventions would be control interventions in three-group randomised trials. The properties of the interventions deemed experimental for this methodology review were defined as the following (based on the work by Hróbjartsson 2010 ; Comer 2013; and Kazdin 2016.

1. Wait-list: an intervention where participants are assessed on one or more occasions, and are promised the 'active' intervention after the trial has ended.
2. Usual care: an intervention that reflects locally accepted treatment practices for a given mental health disorder. It is provided either by private or public practitioners and may involve pharmacological, psychological treatment or both.
3. Psychological placebo: an intervention that targets the non-specific or shared components of psychological treatments, such as treatment exposure and human interaction variables, attending sessions, and patient expectations.

4. Pharmacological placebo: an intervention that includes an inert substance, typically in the form of a pill or liquid, which does not contain the active ingredients of a given medication.
5. Physical placebo: an intervention that includes the inert components of a physical treatment (such as acupuncture, exercise regimens, surgery, or electromagnetic stimulation).

Comparator interventions

We included two comparators: wait-list and no-treatment (see Table 1). When wait-list was the experimental intervention, we only compared it with no-treatment interventions. We defined these comparator interventions as any interventions that were clearly labelled as, or reflected the properties of wait-list and no-treatment interventions. The properties of no-treatment interventions were defined as the following (based on the work by Comer 2013).

1. No-treatment: an intervention where participants are assessed on repeated occasions without receiving the experimental intervention. Unlike wait-list interventions, no-treatment participants are not promised the experimental treatment after trial completion.

Description of main comparisons

We conducted the comparisons on placebo and usual care interventions in the following order.

1. We first pooled wait-list and no-treatment interventions when compared with placebo and usual care interventions.
2. We then conducted subgroup analyses (see Subgroup analysis and investigation of heterogeneity) between wait-list and no-treatment interventions for all these pooled comparisons. If there were significant differences or substantial heterogeneity between the wait-list and no-treatment interventions for a given comparison, we conducted separate main analyses for the two comparison interventions. We expressed low confidence in these analyses if they had insufficient statistical power.

Types of outcome measures

Primary outcomes

1. Outcomes measuring the efficacy of wait-list, usual care, and placebo interventions versus wait-list or no-treatment interventions for all mental health disorders combined.
2. Serious adverse events in wait-list, usual care, and placebo interventions versus wait-list or no-treatment interventions for all mental health disorders combined and for specific mental health disorders.

Secondary outcomes

1. Outcomes measuring the efficacy of wait-list, usual care, and placebo interventions versus wait-list or no-treatment interventions for specific mental health disorders.
2. Non-serious adverse events in wait-list, usual care, and placebo interventions versus wait-list or no-treatment interventions for all mental health disorders and for specific mental health disorders.

Description of outcomes

We conducted analyses across all included mental health disorders and within specific disorders. We grouped the specific disorders according to the classification in the DSM-5 (APA 2013). We only calculated the efficacy for specific mental health disorders that had been included in at least three included trials. This was a pragmatic threshold inspired by Hróbjartsson and Gøtzsche (Hróbjartsson 2010) to reduce spurious positive and negative findings in single trials.

For the outcomes measuring efficacy, we selected one outcome from each trial report. We conducted separate analyses on dichotomous and continuous outcomes (see Measures of the effect of the methods). We used the following decision hierarchy to select the outcomes measuring effect.

1. We first included the outcome indicated as the primary outcome in the trial report (e.g. the one used for the sample size calculation). We preferred data from end of treatment over follow-up data. This choice was inspired by Hróbjartsson 2010.
2. If the trial did not differentiate between primary and secondary outcomes or if more than one primary outcome was stated, we preferred continuous over dichotomous outcomes.
3. If there were multiple continuous outcomes, we preferred observer-reported over patient-reported outcomes, and blinded over non-blinded outcomes.
4. If trials reported several observer-reported outcomes, we included the outcomes that best captured the core symptoms of the mental health population being treated. Here, we preferred global scores over sub-scores.
5. We then identified the outcome measure with the best psychometric properties (e.g. validity and reliability).
6. If still undecided, we randomly selected the outcome measure to use.

Serious adverse events were defined as any event that lead to death (e.g. suicide), is life-threatening (e.g. suicidality), required in-patient hospitalisation (e.g. self-harm), prolonged hospitalisation, resulted in persistent or significant disability, or was any other important event that jeopardised the patient's life or required intervention for prevention (ICH 2005). All other adverse events were considered non-serious adverse events (ICH 2005). We conducted separate analyses for specific serious adverse events (e.g. suicide and self-harm). We combined all non-serious adverse events into a single estimate.

We extracted adverse events from studies as measured by standardised psychometric rating scales, such as laboratory values, or spontaneous reporting. We also located adverse events as described in the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) (ICH 2005). Most of the adverse events data from the reports were spontaneously reported. Adverse events in randomised trials generally (Allen 2018) and for psychiatric treatments in particular, can be difficult to detect, and valid instruments to detect them are lacking (Lilienfeld 2007; Linden 2014; Pagsberg 2017; Storebø 2018). However, strategic searches for adverse events using standardised questionnaires are becoming more common (Pagsberg 2017; Storebø 2018). We corresponded with trial authors if they did not report data on adverse events.

Search methods for identification of studies

Electronic searches

We searched the electronic databases and trial registries listed below (guided by Bramer 2017) using the search strategies shown in Appendix 1 . The strategy for MEDLINE was used as a template for the other databases and trial registries, with modified syntax and controlled terms as necessary.

Bibliographic databases (April 2018)

1. MEDLINE Ovid (1946 to current) (see Appendix 1 for search strategy)
2. PsycINFO Ovid (1806 to current)
3. Embase Ovid (1974 to current)
4. Cochrane Central Register of Controlled Trials (CENTRAL; current issue), in The Cochrane Library.
5. Allied and Complementary Medicine Database (AMED; 1900 to current)
6. Web of Science Core Collection (1900 to current)
7. ProQuest Dissertations and Theses A&I (1743 to current)
8. Sociological Abstracts ProQuest (1952 to current)
9. Google Scholar (https://scholar.google.no/)
10. BIOSIS Previews/Thomson Reuters (969 to current)
11. Open Grey (1997 to current)

Clinical trial registries (March 2019)

1. Australian New Zealand Clinical Trials Registry (ANZCTR; www.anzctr.org.au/BasicSearch.aspx).
2. Clinical Trials (clinicaltrials.gov).
3. EU Clinical Trials Register (www.clinicaltrialsregister.eu/ctr-search/search).
4. ISRCTN Registry (www.isrctn.com).
5. UK Clinical Trials Gateway (www.ukctg.nihr.ac.uk/#popoverSearchDivId).
6. World Health Organization International Clinical Trials Registry Platform (WHO ICTRP; http://apps.who.int/trialsearch/)

Searching other resources

We searched other resources at the end of the screening process. We surveyed relevant journals such as ACTA Psychiatrica Scandinavica, the American Journal of Psychiatry, Biological Psychiatry, the British Journal of Psychiatry, the BMJ, the International Journal of Clinical Psychopharmacology, JAMA Psychiatry, Journal of the American Academy of Child and Adolescent Psychiatry, Journal of Clinical Psychiatry, Journal of Clinical Psychopharmacology, Journal of Psychopharmacology, Lancet Psychiatry, Psychopharmacology, Psychotherapy Research and the Scandinavian Journal of Child and Adolescent Psychiatry and Psychology. We also reviewed abstracts of key psychiatric conferences, given the large proportion of conference abstracts that do not go on to full publication (Scherer 2018) and asked for relevant unpublished studies from experts in the field. We also checked the references in relevant literature.

Data collection and analysis

We conducted this review according to the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2019), and performed analyses using the latest version of RevMan (Review Manager 5) .

Selection of studies

Because we expected to retrieve large numbers of records from the electronic literature search, titles and abstracts were screened only once (divided equally between review authors EF and AT). For quality assurance, an additional review author (OJS) screened a random sample of the retrieved records to check whether there were differences in the included and excluded records between screeners. Three review authors (EF, AT and LB) independently screened the full-text reports for studies judged to be potentially eligible. They discussed any disagreements, and an arbiter (OJS) made the final decision if agreement was not reached. Full-text reports were obtained and assessed for inclusion based on the eligibility criteria (see Criteria for considering studies for this review). Randomised trials in this general topic area that do not fulfil the inclusion criteria are listed as excluded studies. We used EPPI Reviewer 4, an online software application for systematic review development, for screening of abstracts and full-text reports (Thomas 2010). We included a PRISMA flow diagram to show the flow of included and excludes studies in the full review (Moher 2009).

Data extraction and management

Three review authors (EF, AT and LB) independently extracted data from the included studies. We resolved disagreements by discussion or using an arbiter (OJS), if necessary. Two review authors (EF and AT) entered data into Review Manager 5. We requested missing information by contacting relevant authors (Young 2011). We developed a data extraction form to facilitate standardisation of the data extraction process. The form included the following items: methods (e.g. trial design, setting, and country), types of participants (e.g. baseline demographics, inclusion and exclusion criteria), description of experimental and comparator interventions and their components (e.g. duration and intensity), outcome measures, and risk of bias assessment (see Appendix 2).

Assessment of risk of bias in included studies

Three review authors (EF, AT and LB) assessed the risk of bias using Cochrane's risk of bias tool (RoB) 1.0 (Higgins 2011). There is an updated version of this tool (Eldridge 2016 ; Higgins 2017), but because it was still at the pilot stage when we rated risk of bias, we used the original version. For each included study, the data extractors independently categorised the risk of bias domains listed in Appendix 3 as being low, unclear (uncertain), or high risk of bias, according to the guidelines in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). Potential disagreements were resolved by discussion or using an arbiter (OJS), if necessary.

We defined trials at 'low risk of bias' as having low risk of bias on all domains. We defined trials with one or more unclear risk of bias domain as trials at 'high risk of bias'. We evaluated the influence of risk of bias on our results (see Sensitivity analysis) due to the risk of overestimating beneficial intervention effects and underestimating adverse events in randomised trials with unclear or inadequate

methodological quality (Kjaergard 2001; Lundh 2017; Moher 1998; Savović 2012; Savovic 2018; Schulz 1995; Wood 2008; Savovic 2018).

At the protocol stage (Faltinsen 2019), we decided to include all the domains in RoB 1.0 including the blinding domains when assessing risk of bias in the included studies. However, we recognise that the blinding domains are the subject of the investigation in this review in the placebo comparisons, (i.e. one goal of our review is to assess differences in blinding between placebos and no-treatment or wait-list seeing) and we therefore decided to report two solutions to the bias assessment post hoc for placebo interventions: one bias assessment including the blinding domains and one without.

We assessed conflicts of interest in the included studies as a separate bias category outside of Cochrane's risk of bias tool. We assessed both financial and non-financial conflicts of interest. Conflicts of interest were defined as situations in which professional judgments or actions regarding a primary interest are unduly influenced by a secondary interest (Institute of Medicine 2009). Examples of financial conflicts could be when a study's authors had received payment from a company manufacturing one of the study interventions. A non-financial conflict of interest (often termed affiliation bias in psychotherapy research) could be if a study's authors had developed the treatment manual for the intervention being evaluated (Munder 2013).

Dichotomous data

We summarised dichotomous data as risk ratios (RR) for outcomes for efficacy and risk differences (RD) for adverse events. We used 95% confidence intervals (CIs) for both, and Trial Sequential Analysis (TSA)-adjusted CIs if possible (see Subgroup analysis and investigation of heterogeneity).

Continuous data

For continuous data, we estimated standardised mean differences (SMD). We used SMD because we anticipated variation in the types of outcome measures. We calculated SMDs using scores from the end of intervention. We considered a statistical significant SMD effect size of: 0.15 or less to have no clinically meaningful effect; 0.15 to 0.40 to have a clinical meaningful but small effect; 0.40 to 0.75 to have a moderate effect; and greater than 0.75 to have a large treatment effect (Cohen1988). When the trials only reported change data, we pooled these with scores from the end of intervention (da Costa 2013). We explored whether inclusion of change data affected the outcomes by performing a sensitivity analysis (see Sensitivity analysis). If the direction of a given scale was opposite to that of most other scales, we multiplied the corresponding mean values by -1.00 to ensure adjusted values. If the trials did not report means and standard deviations (SDs), but reported other values such as t-tests and P values, we attempted to transform these into means and SDs.

We used data from means and SDs in intention-to-treat (ITT) analyses as well as replacing missing values when available. We otherwise conducted the analyses based on the available data. We performed all calculations using RevMan software (Review Manager 5).

We summarised the outcomes measuring adverse events from count data (e.g. spontaneous reporting) as RD (see Subgroup analysis and investigation of heterogeneity).

Unit of analysis issues

We only included the first phase of cross-over trials. We calculated study estimates on the basis of post-treatment group results. If trials were cluster-randomised we planned to appropriately control for cluster effects (robust standard errors or hierarchical linear models). If the necessary information was unclear or not available in the trial reports, we attempted to contact the original authors for further information. We used sensitivity analyses to assess the potential biases of inadequately controlled cluster-randomised trials (Donner 2002) (see Sensitivity analysis).

Dealing with missing data

We contacted trial authors for relevant missing data on our primary and secondary outcomes (Young 2011). However, we did not contact authors of trials published before 1990 because of a lack of reliable contact information and the probability that these data would not have been preserved. If authors did not respond after two attempts to contact them, we stopped communications. If we were not able to obtain missing data, we used the available data (incomplete data) in our analyses. If data were not reported in a usable way, we consulted a statistician to explore its transformation. For a description of each trial with missing data see Table 2 .

Assessment of heterogeneity

We expected to find evidence of substantial heterogeneity. We created subgroups based on study characteristics such as different control intervention, study duration, participants etc. (see Subgroup analysis and investigation of heterogeneity). We evaluated methodological heterogeneity by comparing trial designs. Assessment of statistical heterogeneity was carried out for comparisons by visual inspection of the graphs and the I^2 statistic (Higgins 2003). I^2 values between 0% and 40% indicated little heterogeneity; between 30% and 60% indicated moderate heterogeneity: between 50% and 90% indicated substantial heterogeneity; and between 75% and 100% indicated considerable heterogeneity (Higgins 2019). We also assessed statistical heterogeneity by Chi^2 tests ($P < 0.10$) and tau^2, an estimate of between-study variability.

Assessment of reporting biases

Funnel plots were provided for comparisons that had a sufficient number of included trials. Asymmetry in the funnel plot could be due to publication bias or other reasons for heterogeneity between small and large trials (Higgins 2019). Egger's statistical test was performed for primary outcomes included in the Summary of findings 3 to test for small-study effects (Egger 1997). A visual inspection of funnel plots and Egger's statistical test was not applied if there were fewer than 10 trials in the meta-analysis, in keeping with the recommendations in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2019).

Data synthesis

We performed statistical analyses according to the recommendations in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2019). We applied the inverse variance method to give estimates from trials with less variance (mostly, larger studies) more weight. We used the random-effects model for meta-analysis because some clinical heterogeneity was expected to be present in most cases. We tested whether a fixed-

effect model provided different effect estimates in a sensitivity analysis (see Sensitivity analysis). If pooling of data seemed feasible, we combined the included study effects and calculated the associated 95% CIs.

Subgroup analyses (pre-specified)

1. Type of active interventions: i) psychological intervention, ii) pharmacological intervention, iii) physical intervention, or iv) other or combination of interventions.
2. Overall risk of bias: i) high risk of bias compared with ii) low risk of bias.
3. Type of outcome domain: i) blinded observer-reported, ii) non-blinded observer-reported, or iii) patient-reported.
4. Type of comparator intervention: i) wait-list or ii) no-treatment.
5. Awareness of placebo intervention: i) participants were aware that they might receive a placebo or ii) participants were not aware of this.
6. Trial objective: i) a trial's objective was clearly to assess the effects of placebo, usual care, or wait-list interventions, or ii) no such objectives were stated.
7. Mean age of participants: i) < 18 years, ii) 18 to 50 years, or iii) > 50 years.
8. Duration of intervention: i) three months or above or ii) below three months.
9. Type of usual care: i) pharmacological, ii) psychological, iii) physical, or iv) other.
10. Standardised usual care: i) the usual care intervention was intentionally standardised or manualised or ii) no standardisation or manualisation.
11. Mode of psychological treatment in usual care and psychological placebo: i) individual psychological treatment or ii) group psychological treatment.

Subgroup analyses (post hoc)

1. Mental health diagnoses: i) formal diagnosis according to DSM/ICD, ii) fulfil symptoms of disorder ICD/DSM while not stating classifications systems, or iii) population is classified as having a mental disorder, but full diagnostic criteria not reported.
2. Type of psychological placebo: i) interaction placebo, ii) educational placebo, or iii) exposure placebo.
3. Type of physical placebo: i) acupuncture or acupressure placebo, ii) exercise and relaxation placebo, iii) technical device placebo, or iv) electromagnetic stimulation placebo.
4. Conflicts of interest: i) risk of non-financial and financial conflicts of interest, or ii) no risk of conflicts of interest (Leichsenring 2019).
5. Imputed data: i) analyses with available outcome data or ii) analyses following the ITT principle.

Diversity-adjusted required information size (RIS) and Trial Sequential Analysis (TSA)

Trial Sequential Analysis (TSA) is a methodology that combines a required information size (RIS) calculation for meta-analyses with a threshold for statistical significance (Brok 2009 ; Thorlund 2009; Wetterslev 2008; Wetterslev 2009; Wetterslev 2017). The TSA enables quantification of the statistical reliability of the data in cumulative meta-analysis, and adjusted P values for sparse data and for repetitive testing on accumulating data (Brok 2008; Brok 2009; Thorlund 2009; Wetterslev 2008; Wetterslev 2017). Similar to an a priori sample size estimation in a single randomised trial, a meta-analysis should include a RIS at least as large as the sample size of an adequately powered single trial to control the risks of random error. The TSA program can calculate the RIS in a meta-analysis and provide an alpha-spending boundary to adjust the significance level for sparse data and repetitive testing (Copenhagen Trial Unit 2018; Wetterslev 2008; Wetterslev 2017). This enables one to control for the risk of random error.

Multiple analyses of accumulating data when new trials emerge lead to repeated significance testing and introduces multiplicity issues. Therefore, the use of a conventional naïve P value exacerbates the risk of random errors (Berkey 1996; Thorlund 2011; Wetterslev 2017). By analysing meta-analyses that do not reach the RIS with trial sequential alpha-spending monitoring boundaries (analogous to interim monitoring boundaries in a single trial), this can be controlled for (Wetterslev 2008; Wetterslev 2017).

We calculated a RIS on the outcomes reported in the summary of findings tables in this review (i.e. the major findings of the review). If the TSA does not find significant results (no crossing of the alpha-spending boundary and no crossing of the conventional boundary of P = 0.05) before the RIS has been reached, several conclusions may be inferred. We will either conclude that more trials are needed to reject or accept an intervention effect used for the calculation of the required sample size, or reject the anticipated effect, if the cumulative Z-curve enters the futility area. We used an assumption that the minimal relevant clinical difference (MIREDIF) was approximately ½ SD on the used scale, which can be used as a MIREDIF (Norman 2003).

We calculated the diversity-adjusted required information size (DARIS; that is the number of participants required to detect or reject a specific intervention effect in a meta-analysis), and performed TSAs for the primary outcomes reporting continuous data at the end of treatment, based on the following a priori assumptions:

1. the SD of the primary outcomes;
2. an anticipated MIREDIF as a ½ SD on the used scale;
3. a maximum type I error of 3.3% (due to two primary outcomes; Jakobsen 2014);
4. a maximum type II error of 10% (minimum 90% power; Castellini 2018); and
5. the diversity observed in the meta-analysis.

For the outcomes 'total serious adverse events' (dichotomous data), we calculated the diversity-adjusted required information size (DARIS; i.e. number of participants in the

1. proportion of participants in the control group with serious adverse events;
2. relative risk reduction of 25%;
3. type I error of 3.3%;
4. type II error of 10%;
5. observed diversity of the meta-analysis; and
6. we included trials with zero events by substituting 0.5 for zero (Thorlund 2011).

It was not possible to calculate TSA on 'total serious adverse events' and 'non-serious adverse events' due to a lack of information.

Summary of findings tables

We used the GRADE approach to construct three summary of findings tables to document primary review outcomes. GRADE evaluates the quality of a body of evidence based on the confidence that an effect estimate or association reflects the item being assessed. These considerations were based on within-trial risk of bias, directness of evidence, heterogeneity of data, precision of effect estimates and risk of publication bias (Andrews 2013a; Andrews 2013b; Balshem 2018; Brunetti 2013; GRADE Working Group 2004; Guyatt 2011a; Guyatt 2011b; Guyatt 2011c; Guyatt 2011d; Guyatt 2011e; Guyatt 2011f ; Guyatt 2011g; Guyatt 2011h; Guyatt 2013a; Guyatt 2013b; Mustafa 2013). When possible, we used the SMD or the RR for the summary of findings table. We used the TSA as the rating for imprecision (Jakobsen 2014).

Sensitivity analysis

Trials contributing to statistical heterogeneity ('outliers') were removed to evaluate their impact on the overall pooled effect estimate. We removed outliers one by one and assessed the impact on the overall outcome.

We conducted sensitivity analyses to determine whether findings were sensitive to the following decisions made during the review process.

1. Analytical technique (e.g. fixed-effect compared with random-effects models)
2. Combination of data in continuous outcomes (end of intervention or compared with change scores)
3. Trial Sequential Analysis (TSA) as a sensitivity analysis for the imprecision rated with GRADE (Castellini 2018)
4. Including wait-list interventions described as no-interventions
5. Including no-interventions described as wait-list interventions

More information on the sensitivity analyses that we were not able to conduct is given in Differences between protocol and review .

RESULTS

Description of studies

See Characteristics of included studies , Characteristics of excluded studies , Characteristics of studies awaiting classification , and Characteristics of ongoing studies .

Results of the search

All electronic databases and search periods are listed in the Methods section (see Electronic searches). The search was conducted in April 2018. The search strategy was comprehensive (see Appendix 1 and generated 64,529 records, but only 58,943 records could be exported from the databases and imported to Endnote. We consulted our research librarian and identified some records that did not include any title, abstract, or keywords, and they could therefore not be retrieved. Another 1034 records of clinical trials across the remaining trial registries could not be imported and had to be manually screened. Ten records were identified from references in other reviews.

Figure 1 shows our PRISMA flowchart. After duplicate check, 13,134 studies were excluded. We used an EPPI Reviewer 4 text mining software filter to identify reports that with 97% certainty was a systematic review or a randomised trial. This filter was used in three phases. In total, 10,167 reports were excluded because they did not fulfil the criteria for being a randomised trial, leaving 35,642 reports for abstract screening in EPPI reviewer 4, and 1034 clinical trials. Following screening of titles and abstracts, 1243 records were identified for full-text screening. Six trials are still awaiting classification due to difficulties locating the trial reports (see Studies awaiting classification), while four trials are ongoing (see Ongoing studies). In total, 96 randomised trials described in 122 reports were eligible for the full review (see Figure 1 for a more detailed description).

Figure 1. Study flow diagram. * Powers 2008 was included in both psychological and pharmacological placebo " Brill 1964 was included in both pharmacological placebo and usual care < Klerman 1974 was divided into two

different trials -: Peck 1974 was included only in all placebos analyses due to that the placebo group was a mix of psychological and pharmacological placebo

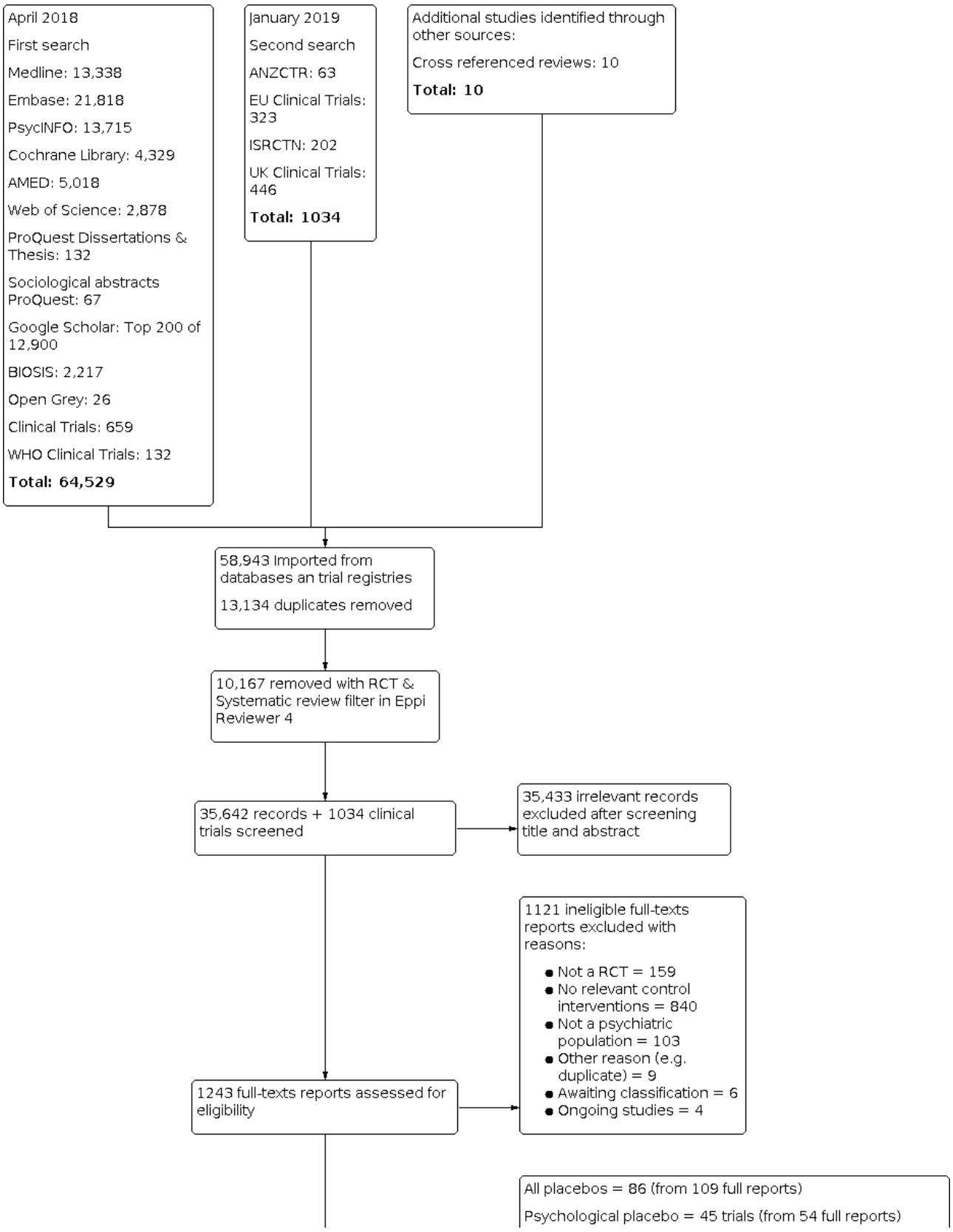
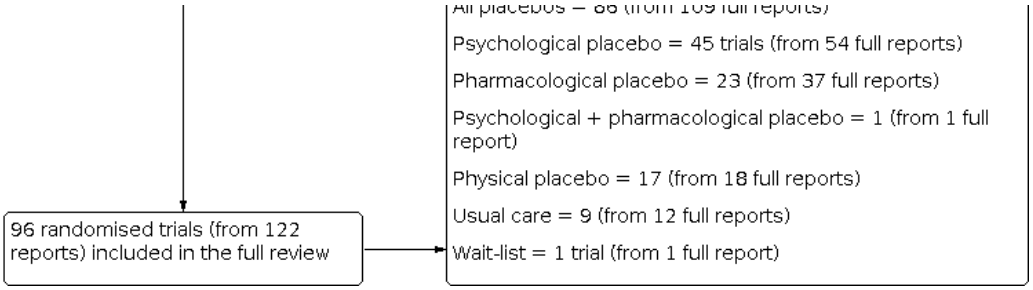


Figure 1. (Continued)



Author correspondence

We contacted authors from 35 trials with unclear or missing data and requested the necessary data but only 16 responded (Table 3) The other trials did not provide contact information or were below the threshold for contact.

Included studies

Here we summarise the key characteristics of the 96 included trials. Further detail can also be found in Characteristics of included studies.

Design

We included 96 trials (94 parallel-group trials, one cluster-randomised trial, and one cross-over trial). The only cross-over trial did not provide any usable data (Sibilio 1957). Only one trial compared wait-list with no-treatment (Howlin 2007). This was the only cluster-randomised trial identified. Nine trials compared usual care versus either wait-list or no-treatment (Brill 1964b; Crisp 1991; Glogowska 2000; Matson 1980; Milby 1980; Rapee 2006; Rapee 2007; Robin 1976; Teri 1997). We included 45 trials on psychological placebos, 23 trials on pharmacological placebos and 17 trials on physical placebos.

One parallel-group trial compared a pooled group of psychological and pharmacological placebo with no-treatment (Peck 1976). One trial included three control groups (wait-list, usual care and pharmacological placebo) and was split into two trials (Brill 1964a; Brill 1964b).

Settings

Seventy-four trials were conducted in outpatient settings and 20 trials were conducted in inpatient settings. Two trials were conducted combining inpatient and outpatient settings (Table 3).

Sample sizes

There was considerable variation in sample sizes between the trials. The total number of participants ranged from eight participants (Kilmann 1987; Peck 1976) to 393 participants (Proudfoot 2013). Only five trials included more than 100 participants (Table 3).

Participants

The 96 trials included a total of 4200 participants but 586 participants could not be included due to missing data. The mean age ranged from 2.9 years (Glogowska 2000) to 86.5 years (Kwan 2017). Nineteen trials only included females, and 14 trials only included males. Seven trials did not state the sex of the participants (Table 3). All remaining trials included both sexes.

Diagnostic criteria

Participants were diagnosed as having a formal mental health disorder according to DSM-II (two trials), DSM-III (four trials), DSM-III-R (eight trials), DSM-IV (15 trials), DSM-IV-TR (three trials) and ICD 9th edition (one trial) (Table 3). The most commonly used assessment instrument was Structured Clinical Interview for DSM (SCID; Spitzer 1989), which was used by eight trials. Twenty-four trials fulfilled the symptoms of a mental health disorder from the available diagnostic classifications system at the time of the trial, but did not report a classification system. Thirty-five trials reported a population classified as having a mental health disorder, but full diagnostic criteria were not reported (for more information see Table 3).

Diagnoses

The 96 trials included participants with the following 15 diagnoses: different forms of anxiety disorders (such as specific anxiety, social anxiety, or panic disorder, 25 trials); depression (16 trials); sleep-wake disorders (11 trials); substance use disorders of different kind such as cocaine and alcohol dependency (eight trials); other unspecified disorders mentioned as 'psychiatric patients' only (eight trials); neurodegenerative diseases (six trials); schizophrenia (five trials); attention deficit hyperactivity disorder (ADHD) or attention deficit disorder (ADD) (five trials); post traumatic stress disorder (PTSD) (four trials); learning disability (three trials); intellectual disability (two trials); and one trial each for anorexia, autism, bulimia, encopresis, and erectile dysfunction (Table 3).

Experimental interventions

In the original trials, the placebo groups were all control interventions. We turned these control interventions into our experimental interventions in this systematic review.

Types of interventions

Only one trial (cluster-randomised) compared a wait-list intervention versus a no-treatment intervention.

Three trials included usual care as a standard treatment, three trials included it as a form of outpatient psychotherapy, one trial as community-based therapy, one trial as typical care control, and one trial did not specify its format (Robin 1976).

The 44 trials with psychological placebo included seven different labels for the psychological placebos. Thirteen trials used the

term attention placebo control, 11 trials used non-specific placebo counselling or treatment; seven trials used a quasi-desensitisation placebo; four trials used a non-specific educational placebo; two trials used a form of active treatment such as present-centred therapy (Foa 2018), or emotion-focused supportive therapy (Ehlers 2014). The other trials used different variations of placebo definitions, such as credible placebo or imagery relief placebo (Table 3).

Twenty trials used a psychological placebo with an interactive component; 16 trials included a psychological placebo with an exposure component; nine trials had a psychoeducational character (Table 3); and one trial combined psychological and pharmacological placebo (Peck 1976).

Sixteen trials provided pharmacological placebos in pill form, one trial used implants, and one trial used injection (Table 3).

Four trials used a pharmacological placebo with psychological treatment as an add-on treatment. Six trials provided a physical placebo as a technical device, five trials as either acupuncture or acupressure, three trials as exercise and relaxation, and two trials as electromagnetic stimulation (Table 3).

Format of interventions

Twenty-five trials administered psychological placebos individually, whereas 18 trials administered them in groups. Three trials used a combination of individual and group administration of psychological placebos (Table 3). All pharmacological and physical treatments were provided on an individual basis, except for one trial that combined a pharmacological treatment with group psychological treatment (Crouch 1988).

Duration of interventions

Seventy-four trials had a duration of less than three months, while 21 had a duration of three months or more (Table 3). One trial did report the duration of the interventions (Hippman 2016). Where reported, the duration of treatment ranged from a single session (Etringer 1982; Karst 2007; Powers 2004; Powers 2008a; Powers 2008b; Wilson 1980; Wolitzky 2009) to two years of treatment (McLachlan 1991).

Control comparators

Fifty-six trials included a no-treatment control and 39 trials used a wait-list control intervention. Ten trials labelled their comparator as a wait-list intervention, but their description and definition of it led us to classify it as a no-treatment control intervention. Four trials labelled their comparator as no-treatment, but their description and definition led us to classify it as a wait-list control intervention, and four trials received an add-on psychotherapeutic treatment to the wait-list group. One trial labelled their wait-list as a 'minimal contact group', three trials labelled their wait-list as a 'delayed treatment group', and one trial received an add-on drug treatment (Table 3).

Concomitant treatment

Twenty trials did not allow concomitant psychotherapy to the placebo, treatment as usual, no-intervention, or wait-list groups, and 18 trials allowed the participant to receive a concomitant psychotherapy to the placebo, treatment as usual, no-intervention, or wait-list groups, (Table 3). The remaining trials did not report any information about concomitant treatments.

Twenty-nine trials allowed the participants to receive a concomitant pharmacotherapy to the placebo, treatment as usual, no-intervention or wait-list groups, while 19 trials did not allow any sort of concomitant pharmacotherapy (Table 3). The other trials did not report any information about concomitant treatments.

Outcomes

Benefits

We followed our hierarchy for selecting outcomes measuring potential benefits (see Types of outcome measures). For more information for the individual trials, see Characteristics of included studies . We included 59 different outcomes for the placebo analyses (see Characteristics of included studies). The most common outcomes were Behavioral Avoidance Test (BAT) in eight trials, Daily Sleep Questionnaire (DSQ) in six trials, and Beck Depression Inventory (BDI) in four trials (Table 3). All outcomes included in the usual care analysis were different. The outcome in the cluster-randomised trial of wait-list versus no-treatment included an outcome that was not used in any of the other analyses.

Adverse events

Only 11 trials reported serious adverse events, and only 14 trials reported non-serious adverse events (Table 3). This was reported in the following ways: one trial used a complaint list (Ayen 2004), another trial used an assessment with clinician-administered posttraumatic stress disorder (PTSD) Scale for DSM-5 (CAPS-5), a third trial used a disulfiram-ethanol reaction (DER), and the remaining 11 trials reported adverse events as a spontaneous reporting.

Excluded studies

In total, we excluded 1121 full-text reports. Of the excluded full-text reports, 159 were not a randomised trial, 840 did not compare a placebo or usual care control intervention versus either wait-list or no-treatment intervention. One hundred and three studies were excluded because the participants did not belong to a psychiatric population. Lastly, nine duplicates were identified in the full-text screening and excluded (Figure 1). Thirty-three excluded studies were close enough to the inclusion criteria to be listed in Characteristics of excluded studies.

Studies awaiting classification

Six full reports are awaiting classification due to difficulties in retrieving them (Studies awaiting classification). We were not able to locate the full text for these trials. Three were reported as an abstract (Bommert 1978; McLachlan 1993; Trianes Torres 1991), and three were only reported as a title (Brandes 2010; Newton-Cross 2017; Schwarzler 1999). We tried to contact the authors of the most recent studies (Brandes 2010 ; Newton-Cross 2017), but did not receive any response. After two attempts, we terminated our correspondence.

Ongoing studies

We identified four ongoing studies that assessed different type of placebos or usual care versus wait-list or no-treatment (Heitman 2017; ISRCTN21392756; ISRCTN35717198; NCT00044629) (Characteristics of ongoing studies).

Risk of bias in included studies

Figure 2 and Figure 3 show our assessment of the risk of bias for each included study (see also Characteristics of included studies).

Figure 2. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.

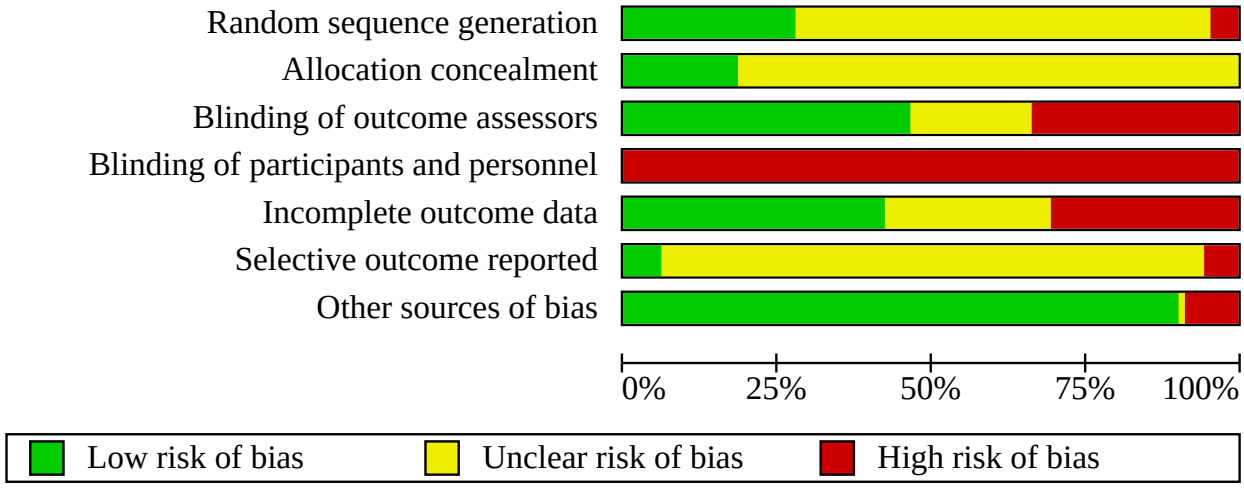


Figure 3. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.

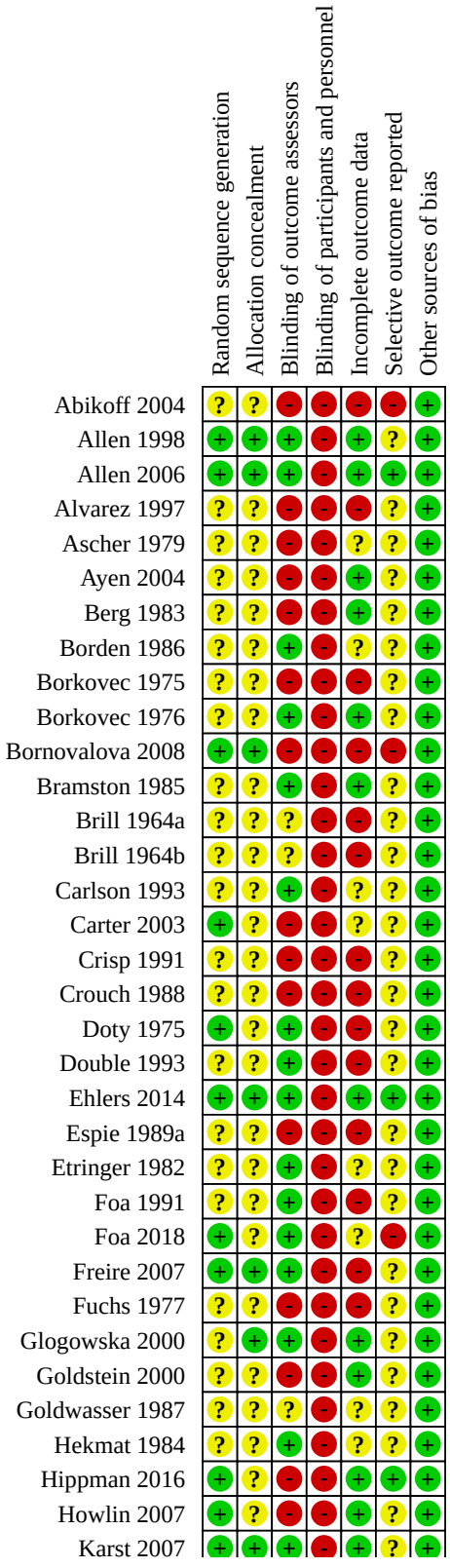


Figure 3. (Continued)

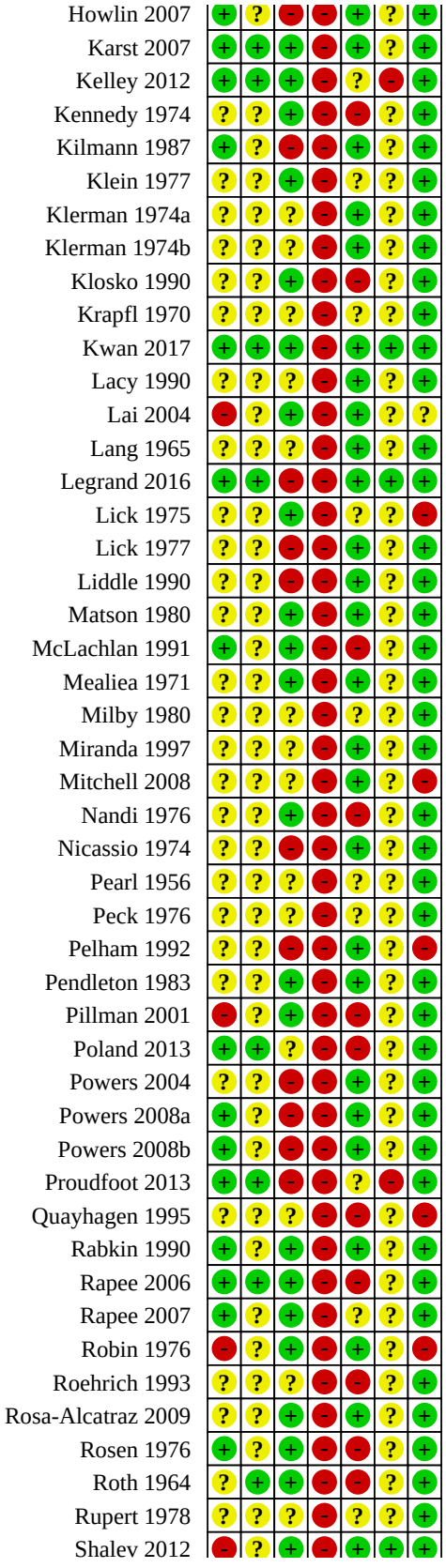
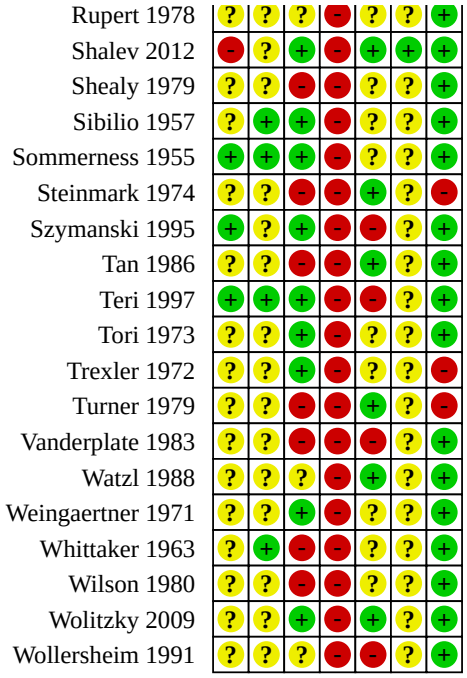


Figure 3. (Continued)



We judged all trials to be at high risk of bias overall. All trials were rated at high risk of bias on blinding of participants and personnel because of the difficulties with blinding a trial with a no-treatment or wait-list comparator. However, the remaining risk of bias domains also had a large proportion of unclear risk of biases. We used all eligible trials in the meta-analysis, as recommended in the Cochrane Handbook for Systematic Reviews of Interventions when all trials are assigned the same risk of bias (Higgins 2011; Higgins 2019). We incorporated our risk of bias assessment when considering the quality of the evidence using the GRADE approach (Higgins 2011). Below is a breakdown of how the included trials scored on each risk of bias domain.

Allocation

Random sequence generation

Evidence suggests that trials which lack sufficient reporting of randomisation processes are more likely to present larger effect estimates for beneficial outcomes (Chalmers 1983; Schulz 1995; Kjaergard 2001; Savović 2012; Savovic 2018; Wood 2008). Trials were regarded as low risk of bias if they provided detailed description of their randomisation process such as stratification methods rather than just stating that it was randomised. Twenty-seven trials provided sufficient information on how the randomisation sequence had been generated, and were rated at low risk of bias. The random sequence generation was rated at high risk of bias in four trials (Table 4). Examples included patients being randomised according to when they entered the treatment program (Poland 2013) or allowing participants to decline up to two treatments arms (Shalev 2012). The remaining 65 trials did not provide sufficient information on how the treatment allocation had been conducted and were assessed as unclear risk of bias.

Allocation concealment

We classified trials as low risk of bias if the allocation was conducted off-site (centralised) by computer software or by an independent research coordinator not involved in delivering the therapy. Eighteen trials provided information about how the allocation was concealed and were therefore rated at low risk of bias. The other 78 trials did not provide any information regarding allocation concealment and were assessed as unclear risk of bias (Table 4)

Blinding

Blinding of outcome assessors

Forty-five trials that reported that outcomes assessors were kept blind to treatment allocation were rated as low risk of bias. Thirty-two trials were rated as high risk of bias due to reporting of inadequate blinding of outcomes assessors. The remaining 19 trials were assessed as unclear risk of bias due to a lack of sufficient information (Table 4).

Blinding of participants and personnel

We judged all trials as high risk of bias in this domain because the participants would be aware of whether they received treatment or not (e.g. allocated to either placebo or wait-list).

Incomplete outcome data

Incomplete outcome data

Forty-one trials were assessed as low risk of bias, due to the use of appropriate methods for handling missing data, such as intention-to-treat (ITT) analyses. Twenty-nine trials either reported data on completers only or did not address the missing data and were considered as high risk of bias. The other trials did not report

adequately information regarding missing data, and were therefore assessed as unclear risk of bias (Table 4).

Selective reporting

Most trials (n = 85) did not have a published protocol prior to initiation or did not provide sufficient information in the report to judge reporting bias and were considered as unclear risk of reporting bias. Six trials had a prior published protocol that provided sufficient information about all outcomes and did not exclude any of these in the full report. These trials were rated as low risk of biased reporting. Five trials published a protocol before the start of the trial, but we found discrepancies such as missing outcomes or an addition of outcomes and we rated them as high risk of bias.

Other potential sources of bias

Eighty-seven trials were rated as low risk of other biases, eight trials were rated as high risk of bias in other potential sources of bias. These included four trials with researchers or authors who provided the treatment, two trials with attention bias or differences in duration of treatment, one trial with potential carry-over effects, one trial that exceeded the passivity of placebo, and one trial with a time bias or assessment at different point for the groups. One trial was rated as unclear risk of others bias because of confounding differences between groups on the number of medical diagnoses other than dementia among participants (Table 4).

Conflicts of interest

We assessed six trials to be at risk of bias because of conflicts of interest. This could be a non-financial affiliation bias, for instance if one of the investigators had developed a treatment evaluated in the trial, or bias from trials funded by a company manufacturing one of the interventions. We included subgroup analyses to test the difference between trials judged to be at risk of affiliation bias and those judged not to be at this risk (Analysis 16.17).

Effect of methods

Effects of interventions for all mental health disorders

Here, we present the results for each of the primary and secondary outcomes for the 19 comparisons. Seventy-one trials reported data as continuous and 11 trials reported dichotomous data, whereas 13 trials did not report usable data. It was only possible to generate missing data for nine trials. For more information see Table 2 .

Wait-list versus no-treatment

We included a single cluster-randomised trial that compared a wait-list intervention versus a no-treatment intervention (Howlin 2007). However, no usable data were provided in the full report and the trial authors did not respond to our request for additional data. In this trial, 84 elementary school children with autism spectrum disorder were randomised to either immediate treatment, delayed treatment (wait-list) or no-treatment. The study's conclusion was that Picture Exchange Communication System (PECS) training showed modest effectiveness for these children. In general, there were no differences across outcome measures between the wait-list and the no-treatment intervention groups.

Outcomes measuring benefits for usual care versus wait-list or no-treatment

Nine trials compared usual care versus wait-list or no-treatment. Two of these did not report usable data (Table 5). Five trials reported continuous data and two trials reported dichotomous data.

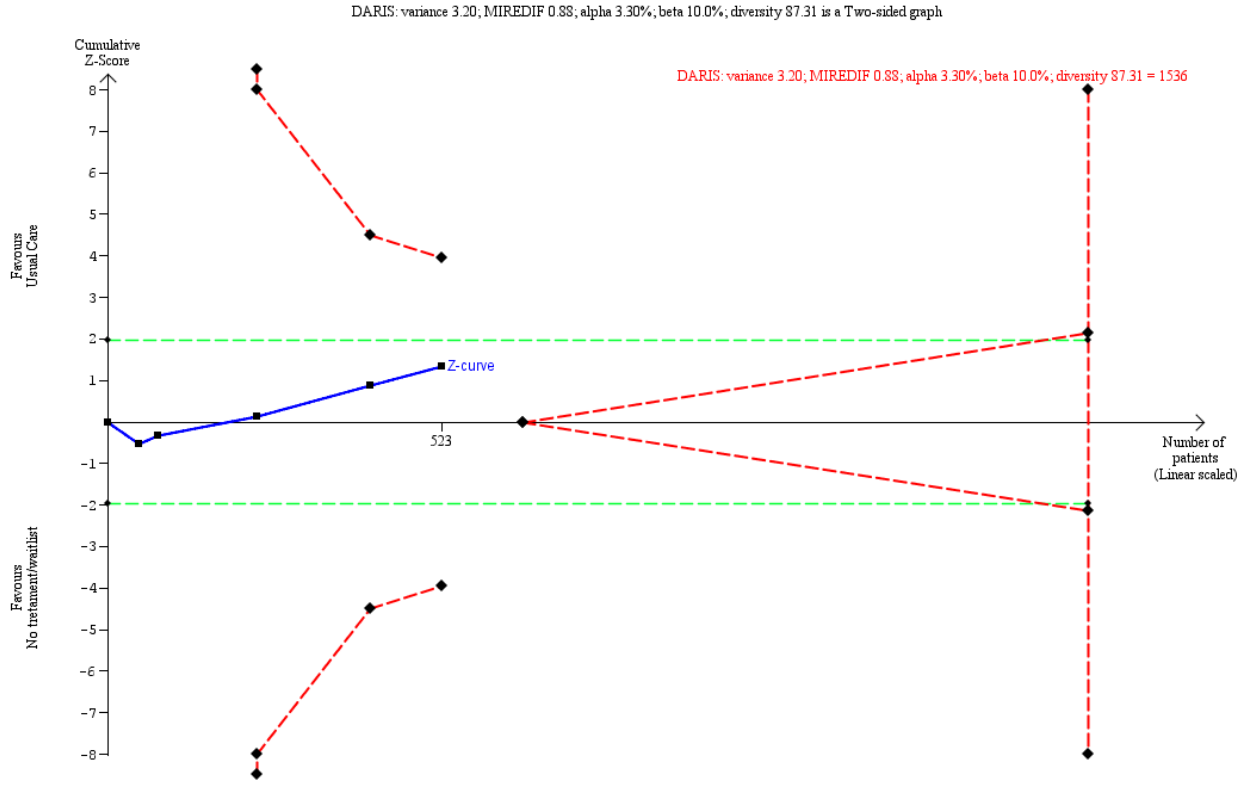
Usual care versus wait-list or no-treatment (continuous data)

No differences were found for beneficial effects comparing usual care versus wait-list or no-treatment when using a random-effects model (standardised mean difference (SMD) −0.33, 95% confidence interval (CI) −0.83 to 0.16; 5 trials, 523 participants; P = 0.13; I^2 = 86% ; very low-quality evidence; Analysis 2.1). No differences were found between the subgroups: usual care versus wait-list (SMD −0.53, 95% CI −1.17 to 0.10; 3 trials, 443 participants; P = 0.10; I^2 = 91%), and usual care versus no-treatment (SMD 0.08, 95% CI −0.38 to 0.53; 2 trials, 80 participants; P = 0.74; I^2 = 0%). Test for subgroup difference: Chi^2 = 2.33, df = 1, P = 0.13, I^2 = 57.1%.

When using a fixed-effect model, usual care had a beneficial effect compared with wait-list or no-treatment (SMD −0.45, 95% CI −0.62 to −0.27; 5 trials, 523 participants; P = 0.01; I^2 = 83.3%; Analysis 25.1), and there were differences between the subgroups: usual care versus wait-list (SMD −0.54, 95% CI −0.73 to −0.35; 3 trials, 443 participants; P < 0.00001; I^2 = 91%) and usual care versus no-treatment (SMD 0.08, 95% CI −0.38 to 0.53; 2 trials, 80 participants; P = 0.74; I^2 = 0%). Test for subgroup differences: Chi^2 = 5.99, df = 1; P = 0.01; I^2 = 83.3%.

The TSA showed the cumulated Z curve enters the futility area, and therefore the anticipated intervention effect can be rejected (TSA-adjusted confidence interval −2.32 to 1.15) (see Figure 4). Inspection of the funnel plot and Egger's test were not possible due to insufficient data.

Figure 4. When we compared usual care with wait-list and no-treatment, we performed trial sequential analysis (TSA) on the primary outcome. The analysis shows that the required information size was not reached. See Figure 4 above. MIREDIF: Minimum relevant difference



Usual care versus wait-list or no-treatment (dichotomous data)

No differences were found for beneficial effects comparing usual care versus wait-list or no-treatment (RR 0.93, 95% CI 0.34 to 2.57; 2 trials, 260 participants; P = 0.89; I^2 = 79%; very low-quality evidence; Analysis 2.2). Tests for subgroup differences were not done because both these trials were versus wait-list.

Serious adverse events of usual care versus wait-list or no-treatment

None of the trials in this comparison reported data on serious adverse events.

Non-serious adverse events of usual care versus wait-list or no-treatment

None of the trials in this comparison reported data on non-serious adverse events.

Outcomes measuring benefits for all placebos versus wait-list or no-treatment

86 trials compared all placebos versus wait-list or no-treatment. 12 of these did not report usable data (Table 5).

All placebos versus wait-list or no-treatment (continuous data)

All placebo interventions showed beneficial effect compared with wait-list or no-treatment (SMD −0.37, 95% CI −0.49 to −0.25; 65 trials, 2446 participants; P < 0.00001, I^2 = 41% ; low-quality evidence; Analysis 6.1) . Differences were identified between subgroups: all placebos versus wait-list (SMD −0.55, 95% CI −0.76 to −0.35; 31 trials, 1410 participants; P < 0.00001; I^2 = 62%), and all placebos versus no-treatment (SMD −0.18, 95% CI −0.30 to −0.05; 34 trials, 1036 participants; P = 0.005; I^2 = 0%). Test for subgroup differences: Chi ^{2} = 9.63, df = 1 ; P = 0.002; I^2 = 89.6%).

The Trial Sequential Analysis (TSA) showed that the required information size (RIS) was reached (n = 397) and that there was no risk of type 1 error (TSA adjusted confidence interval −1.85 to −0.84) (see Figure 5 in Appendix 4). Inspection of the funnel plot (see Figure 6 in Appendix 4) suggested a small potential bias (asymmetry), but we found no evidence of possible publication bias: Egger's regression intercept (bias) −0.699 (two tailed, P = 0.087).

All placebos versus wait-list or no-treatment (dichotomous data)

We found no differences for beneficial effect comparing all placebos versus wait-list or no-treatment (risk ratio (RR) 1.05, 95% CI 0.74 to 1.48; 9 trials, 385 participants; P = 0.79; I^2 = 58%; very low-quality evidence; Analysis 6.2) . Tests for subgroup differences were not done because all these trials were versus no-treatment.

Serious adverse events of all placebos versus wait-list or no-treatment

Eleven trials compared versus placebos with wait-list or no-treatment and reported serious adverse events (Table 5).

All placebos versus wait-list or no-treatment (dichotomous data)

We found no differences for serious adverse events comparing all placebos versus wait-list or no-treatment (risk difference (RD) −0.00, 95% CI −0.03 to 0.03; 11 trials, 517 participants; P = 0.89; I^2 = 0%; very low-quality evidence; Analysis 7.1)

It was not possible to construct a TSA-figure on serious adverse events due to insufficient data. Inspection of the funnel plot (Figure 7 in Appendix 4) suggested no potential bias (asymmetry). We found no evidence of possible publication bias: Egger's regression intercept (bias) −1.192 (two tailed, P = 0.408).

Psychological placebos versus wait-list (dichotomous data)

We found no differences for serious adverse events comparing psychological placebos versus wait-list (RD −0.01, 95% CI −0.07 to 0.04; 2 trials, 207 participants; P = 0.68; I^2 = 0%; very low-quality evidence; Analysis 7.2)

Pharmacological placebos versus no-treatment (dichotomous data)

We found no differences for serious adverse events comparing pharmacological placebos versus no-treatment (RD 0.01, 95% CI −0.08 to 0.09; 4 trials, 125 participants; P = 0.89; I^2 = 0%; very low-quality evidence; Analysis 7.3)

Physical placebos versus wait-list or no-treatment (dichotomous data)

We found no differences for serious adverse events comparing physical placebos versus wait-list or no-treatment (RD 0.00, 95% CI −0.04 to 0.04; 5 trials, 185 participants; P = 1.00; I^2 = 0%; very low-quality evidence; Analysis 7.4)

Non-serious adverse events of all placebos versus wait-list or no-treatment

Fourteen trials compared all placebos versus wait-list or no-treatment and reported non-serious adverse events (Table 5).

All placebos compared with wait-list or no-treatment (dichotomous data)

We found no differences for non-serious adverse events comparing all placebos versus wait-list or no-treatment (RD 0.03, 95% CI −0.02 to 0.08; 14 trials, 590 participants; P = 0.27; I^2 = 33% ; very low-quality evidence; Analysis 8.1).

Psychological placebos versus wait-list or no-treatment (dichotomous data)

We found no differences for non-serious adverse events comparing psychological placebos versus wait-list or no-treatment (RD 0.01, 95% CI −0.18 to 0.19; 5 trials, 280 participants; P = 0.96; I^2 = 66%; very low-quality evidence; Analysis 8.2).

Pharmacological placebos versus wait-list or no-treatment (dichotomous data)

We found no differences for non-serious adverse events comparing pharmacological placebos versus no-treatment (RD 0.08, 95% CI −0.04 to 0.21; 4 trials, 125 participants; P = 0.18; I^2 = 46%; very low-quality evidence; Analysis 8.3).

Physical placebos versus wait-list or no-treatment (dichotomous data)

We found no differences for non-serious adverse events comparing physical placebos versus wait-list or no-treatment (RD 0.00, 95% CI −0.04 to 0.04; 5 trials, 185 participants; P = 1.00; I^2 = 0%; very low-quality evidence; Analysis 8.4).

Outcomes measuring benefits for psychological placebos versus wait-list or no-treatment

Forty-four trials compared psychological placebos versus wait-list or no-treatment. Five trials did not report usable data.(Table 5). One trial reported dichotomous data and 39 trials reported continuous data.

Psychological placebos versus wait-list or no-treatment (continuous data)

Psychological placebos showed a beneficial effect compared with wait-list or no-treatment interventions (SMD −0.49, 95% CI −0.64 to −0.30; 38 trials, 1656 participants; P < 0.00001; I^2 = 56%; low-quality evidence; Analysis 9.1). Differences were identified between subgroups: psychological placebos versus wait-list (SMD −0.66, 95% CI −0.92 to −0.40; 23 trials, 721 participants; P < 0.00001; I^2 = 41%), and psychological placebos versus no-treatment (SMD −0.21, 95% CI −0.38 to −0.04; 15 trials, 542 participants; P = 0.02; I^2 = 0%). Test for subgroup difference: Chi^2 = 8.03, df = 1, P = 0.005; I^2 = 87.5%.

The TSA showed that the RIS was reached (n = 454), and that there was no risk of type 1 error (TSA-adjusted confidence interval −2.54 to −1.02) (Figure 8 in Appendix 4). Inspection of the funnel plot (Figure 9 in Appendix 4) suggested no potential bias (asymmetry), and we found no evidence of possible publication bias: Egger's regression intercept (bias) −0.915 (two tailed, P = 0.259).

Psychological placebos versus wait-list or no-treatment (dichotomous data)

No data were applicable; Analysis 9.2 .

Outcomes measuring benefits for pharmacological placebos versus wait-list or no-treatment

Twenty-three trials compared pharmacological placebos versus wait-list or no-treatment. Six trials did not report usable data (Table 5). Nine trials reported continuous data and eight trials reported dichotomous data.

Pharmacological placebos versus wait-list or no-treatment (continuous data)

We found no differences for beneficial effect comparing pharmacological placebos versus wait-list or no-treatment (SMD −0.14, 95% CI −0.39 to 0.11; 9 trials, 279 participants; P = 0.28; I^2 = 0%; very low-quality evidence; Analysis 10.1) No differences were identified between subgroups: pharmacological placebos versus wait-list (SMD −0.51, 95% CI −1.41 to 0.38; 1 trial, 20 participants; P = 0.26; I^2 = not applicable), and pharmacological placebos versus no-treatment (SMD −0.11, 95% CI −0.37 to 0.16; 8 trials, 259

participants; P = 0.43; I^2 = 0%). Test for subgroup differences: Chi ^{2} = 0.73, df = 1, P = 0.39; I^2 = 0%.

The TSA shows that the cumulated Z curve enters the futility area, and therefore the anticipated intervention effect can be rejected (TSA adjusted confidence interval −9.43 to 6.15) (Figure 10 in Appendix 4). Inspection of the funnel plot was not possible due to insufficient data. We found no evidence of possible publication bias: Egger's regression intercept (bias) −1.192 (two tailed, P = 0.408).

Pharmacological placebos versus wait-list or no-treatment (dichotomous data)

We found no differences for beneficial effect comparing pharmacological placebos versus wait-list or no-treatment (RR 1.05, 95% CI 0.74 to 1.48; 8 trials, 366 participants; P = 0.79; I^2 = 58%; very low-quality evidence; Analysis 10.2). Test for subgroup differences: not applicable.

Outcomes measuring benefits for physical placebos versus wait-list or no-treatment

17 trials compared physical placebos versus wait-list or no-treatment. One of these did not report usable data (Table 5 in Appendix 4).

Physical placebos compared with wait-list or no-treatment for continuous data

Physical placebos had a beneficial effect compared with wait-list or no-treatment (SMD −0.21, 95% CI −0.35 to −0.08; 17 trials, 896 participants; P = 0.002; I^2 = 0%; low-quality evidence; Analysis 11.1). No differences were found between subgroups: physical placebos versus wait-list (SMD −0.30, 95% CI −0.54 to −0.06; 7 trials, 669 participants; P = 0.02; I^2 = 37%), and physical placebos versus no-treatment (SMD −0.15, 95% CI −0.42 to 0.11; 10 trials, 227 participants; P = 0.26; I^2 = 0%). Test for subgroup differences: Chi ^{2} = 0.64, df = 1, P = 0.42, I^2 = 0%.

The TSA showed that the RIS was reached (n = 194), and that there was no risk of type 1 error (TSA-adjusted confidence interval −3.64 to −0.49) (Figure 11 in Appendix 4). Inspection of the funnel plot (Figure 12 in Appendix 4) suggested no potential bias (asymmetry), and we found no evidence of possible publication bias: Egger's regression intercept (bias) −0.078 (two tailed, P = 0.860).

Subgroup analyses

We found significant subgroup differences between using wait-list or no-treatment as comparators in the analyses on all placebos and psychological placebos (Appendix 5). We found larger differences in favour of all placebos or psychological placebos when comparing them with wait-list rather than no-treatment.

For specific mental health disorders: on all placebos versus wait-list or no-treatment, we found differences in favour of placebos for sleep-wake disorders (SMD −0.34, 95% CI −0.60 to −0.07, I^2 = 0%), depression (SMD −0.42, 95% CI −0.78 to −0.05, I^2 = 51%), post-traumatic stress disorder (SMD −0.54, 95% CI −1.06 to −0.02, I^2 = 74%), and anxiety disorders (SMD −0.57, 95% CI −0.93 to −0.21, I^2 = 66%) . However, sparse data in these analyses made the results imprecise. Psychological placebos showed a beneficial effect for patients with sleep-wake disorders (SMD −0.44, 95% CI −0.76 to −0.12, I^2 = 0%) and for patients with post-traumatic stress disorder (SMD −0.75, 95% CI −1.23 to −0.27, I^2 = 55%) versus wait-list or no-treatment.

In the other subgroup analyses, we only found significant differences in the analyses comparing non-blinded observer-reported outcomes with blinded observer-rated and patient-reported outcomes. We also conducted a post-hoc subgroup analysis on types of psychological placebos: those with an interactional component (e.g. talking to a counsellor in a non-directive manner), psychological placebos with a psychoeducational component and those with an exposure element. For this subgroup analysis, we found that interactional placebos yielded significantly higher effects than the other two types of psychological placebo.

There were no significant differences in the other subgroup analyses (see Appendix 5 for all the estimates of the subgroup analyses).

Sensitivity analyses

Due to a lack of sufficient data, it was not possible to conduct some of our predefined sensitivity analyses (see Table 6). We used both the fixed-effect and the random-effects models in all meta-analyses. Statistical significance did not change when we applied a fixed-effect model to analyses regarding all placebos, psychological placebos, pharmacological placebos and physical placebos. However, the statistical significance did change for usual care versus no-treatment or wait-list (Analysis 25.1). We therefore report the results of the random-effects model for placebo interventions versus wait-list or no treatment, and both the random-effects and fixed-effect models for usual care versus wait-list or no treatment.

We also tested if different type of data collection (e.g. measures of adverse events) impacted our results and found no differences. For another sensitivity analysis, for outcomes at the end of intervention, we removed change scores to see if it affected the results. No statistical significant differences were detected.

We performed Trial Sequential Analysis (TSA) on all relevant primary outcomes included in Summary of findings 2 and Summary of findings 3 . The required information size (RIS) was reached for all placebos, psychological, and physical placebos compared with wait-list and no-treatment. It was not reached for pharmacological placebos, where the cumulated Z curve entered the futility area, and therefore the anticipated intervention effect could be rejected. The RIS was also not reached for usual care compared with wait-list or no-treatment. It was not possible to calculate TSA on serious adverse events and non-serious adverse events because of insufficient information. We also tested whether removing the trials named by trial authors as wait-list or no-treatment but fitting our criteria for no-treatment or wait-list respectively would change the subgroup analysis between these two groups and found no significant differences.

Summary of findings tables

We did not assess the quality of evidence and report effect estimates in Summary of findings 1 because it only included one study. In the second table, on usual care versus no-treatment or wait-list, we rated the quality of evidence as low using GRADE (Summary of findings 2). In the third table on placebo interventions,

we included six comparisons and the quality of evidence was rated low to very low (Summary of findings 3).

DISCUSSION

Summary of main results

This review includes 96 randomised trials, out of which 83 trials provided usable data (3614/4200 participants or 86%). The trials included 15 different mental health disorders. We only found one trial that compared wait-list versus no-treatment directly and the authors were not able to provide usable data for this comparison. The comparison on usual care versus wait-list or no-treatment was not significant with an standardised mean difference (SMD) of -0.33 (95% confidence interval (CI) -0.83 to 0.16, I^2 = 86%, 5 trials, 523 participants), although a sensitivity analysis showed significant differences when using a fixed-effect instead of a random-effect model with an SMD of -0.46 (95% CI -0.64 to -0.28). We found significant differences between all placebo interventions combined versus wait-list and no-treatment with an SMD of -0.37 (95% CI -0.49 to -0.25, I^2= 41%, 65 trials, 2446 participants), but there was evidence of some asymmetry in the funnel plot and almost all the trials were small. We found a moderate effect in favour of psychological placebos (SMD -0.49, 95% CI -0.64 to -0.30, I^2 = 53%, 39 trials, 1656 participants) and small effects in favour of pharmacological placebos (SMD -0.14, 95% CI -0.39 to 0.11, I^2 = 0%, 9 trials, 279 participants) and physical placebos (SMD -0.21, 95% CI -0.35 to -0.08, I^2 = 0%, 17 trials, 896 participants). There were significant differences in favour of all placebos in the comparisons on specific mental disorders, but the analyses suffered from sparse data. No differences were found on harms in any of the analyses.

The present systematic review has many strengths. We developed a protocol for this review (Faltinsen 2019) in accordance with the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2019). We conducted extensive searches in relevant databases, with no restrictions to language, publication year, or publication type. Two independent review authors selected trials, extracted data, assessed the risk of bias, and graded the quality of the evidence. Disagreements were resolved by discussion. We used Trial Sequential Analysis (TSA) to estimate the required Information Size (RIS) needed to either accept or reject a certain intervention effect. Another strength of the review is the large number of included trials and the fact that we could make direct comparisons between the different control interventions. This enabled a comprehensive assessment of the effect of the included controls.

The results of this review are affected by the statistical heterogeneity in the analyses, which may be due in part to methodological and clinical heterogeneity in variables such as the included mental health populations, outcome domains, and the contents of the control interventions. To investigate some of the heterogeneity stemming from the pooled mental health populations in the first primary outcome, we conducted comparisons on specific mental health disorders. However, out of the 15 different identified diagnoses across the included trials, we were only able to run comparisons on seven mental health disorders, and these analyses were limited by sparse data, which made the results imprecise. The majority of the review's meta-analyses are therefore conducted across all included mental disorders.

Another example of methodological heterogeneity is the large variability in the contents of the included psychological placebos, which is a much-discussed issue in the literature on mental health control interventions (Comer 2013; Kazdin 2015; Borkovec 2005; Hróbjartsson 2012). It is in part difficult to properly design a psychological placebo because of the issue of targeting hypothesised specific factors of complex psychological treatments Borkovec 2005; Hróbjartsson 2012; Mohr 2014). Out of the three types of placebo interventions, psychological placebos showed the largest reported difference compared with wait-list and no-treatment for all included mental health populations (SMD -0.49, 95% CI -0.65 to -0.30). In an attempt to further investigate the methodological heterogeneity within this control intervention, we conducted a post-hoc subgroup analysis on psychological placebos and divided them into three separate groups: interactional, psychological placebos with a psychoeducational component, and those with an exposure element (Analysis 17.8). The term interactional placebo referred in this case to psychological placebos that control for human interaction variables in treatments such as psychotherapy. We found that interactional placebos yielded significantly higher effects than the other two groups, which may indicate that control interventions that involve a human interaction element yield higher effect sizes compared to no-treatment or wait-list. This is in line with previous frameworks on the shared factors of psychological treatments (Wampold 2010; Hafliðadóttir 2021), although it is not possible to draw strong conclusions on this matter with this exploratory subgroup analysis. Rather, it points to the fact that the psychological placebos as an intervention are methodologically heterogeneous in their contents.

Overall completeness and applicability of evidence

We were able to include data from 83 of 96 trials or 3614 participants out of 4200 in total (86% of the total participant pool). In order to include a sufficient number of trials and give a global estimate of efficacy, we combined different outcomes across the included trials. More specifically, one outcome was chosen from each trial based on our predetermined outcome hierarchy. We included 66 different outcomes, which is a source of methodological heterogeneity (Higgins 2019). When we looked at specific mental health disorders, the outcome measures were often more similar. For instance, for depression, the outcome was often a depression inventory to rate symptoms (see Table 3).

Quality of the evidence

We assessed all included trials as high risk of bias, partly due to lack of blinding of participants and personnel. It is not possible to maintain blinding when comparing a control intervention where participants receive some form of treatment with a control intervention where no treatment is provided (wait-list or no-treatment). This is because the participants and often the personnel will know what treatment is provided to whom, which makes the results prone to bias and systematic errors (Higgins 2019). However, the trials also suffered from other forms of risk of bias. In fact, only three trials (Allen 2006; Ehlers 2014; Kwan 2017) would have been rated at low risk of bias if the blinding of the participants and personnel rating were excluded from the risk of bias assessment. Blinding of participants and personnel is a persistent issue in randomised trials with psychosocial interventions (Guidi 2018; Juul 2020) and there is a need to consider how to address the specific methodological challenges relevant to

these types of trials (Guidi 2018; Munder 2018). We also conducted a post-hoc subgroup analysis and found no differences between the trials at low, unclear and high risk of bias, when we removed the blinding of participants and personnel from the assessment (Analysis 16.18 , test for subgroup differences P = 0.26, I^2 = 26%). However, this analysis was exploratory and many of the trials provided insufficient information on the bias domains, which makes it difficult to judge the overall true extent of systematic bias in the included studies. The reader should take these factors into consideration when evaluating the review's risk of bias profile and its impact on the quality of the evidence.

We intended for this review to provide support for an empirical and methodological reflection of the benefits and harms of control interventions in mental health intervention research. Realising that the subject matter is complex, one may look at the low to very low quality of evidence as a reflection of the state of control intervention design rather than a criticism of the interpretability of our review. Although the included trials were rated as low to very low quality of evidence, this is a reflection of the heterogeneous objectives of this review and to some extent the fact that blinding of participants is not possible with the included designs. As a consequence, we chose to report an alternative risk of bias profile where we excluded the blinding domains (see Figure 5; Figure 6) in addition to the conventional risk of bias assessment. However, this did not change the overall quality of evidence.

Figure 5. Alternative Risk of bias graph

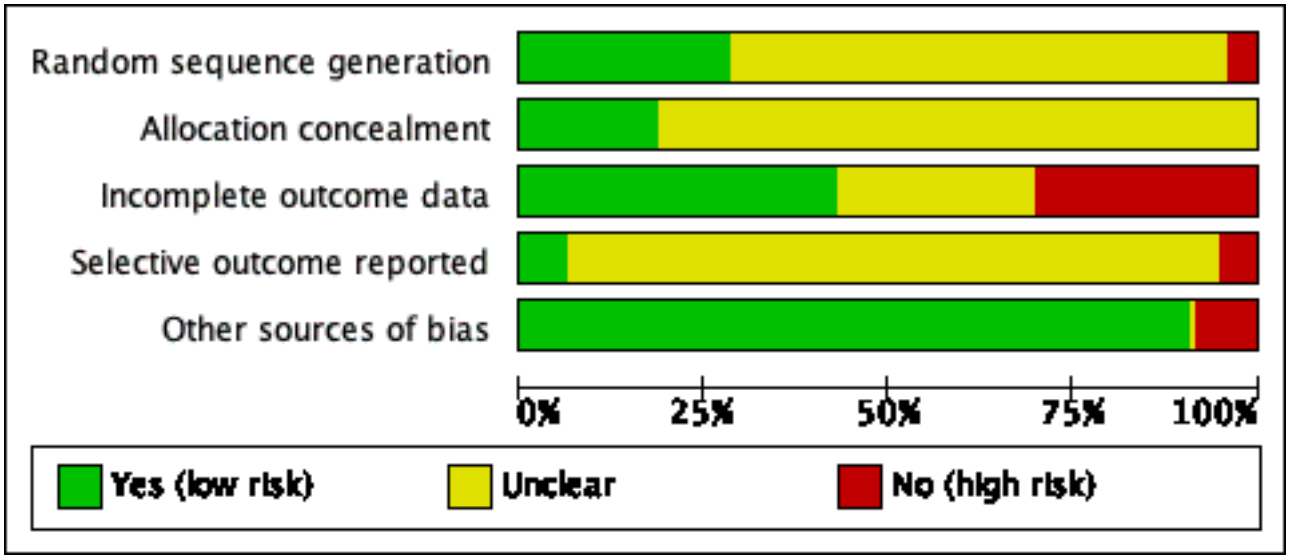


Figure 6. Alternative Risk of bias summary

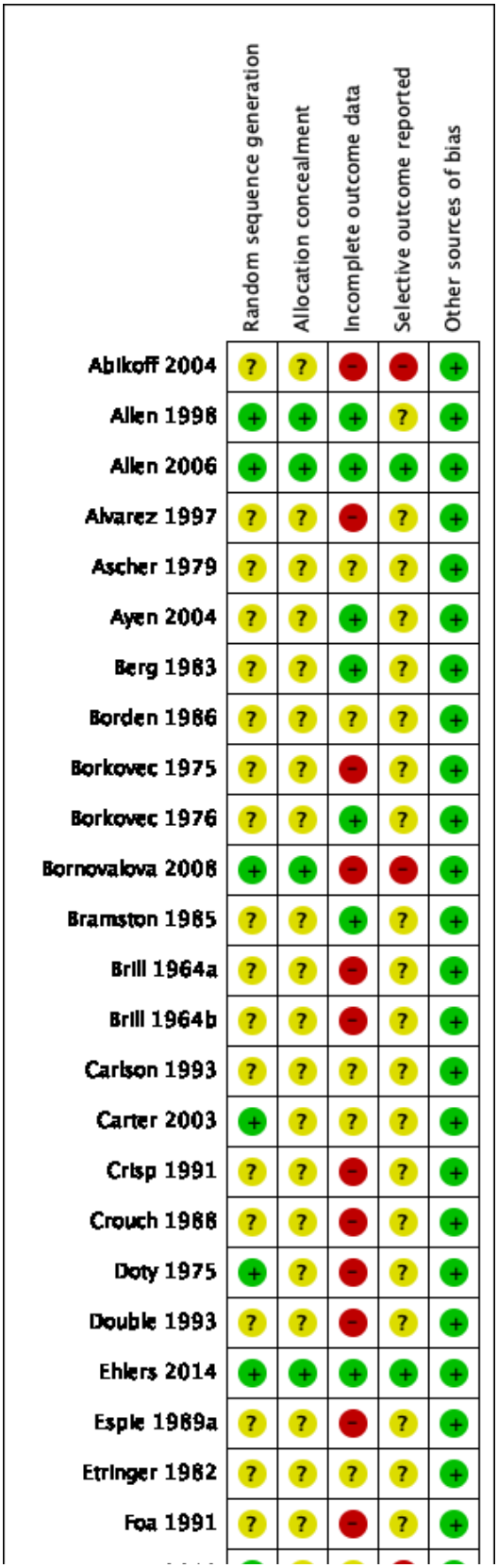


Figure 6. (Continued)

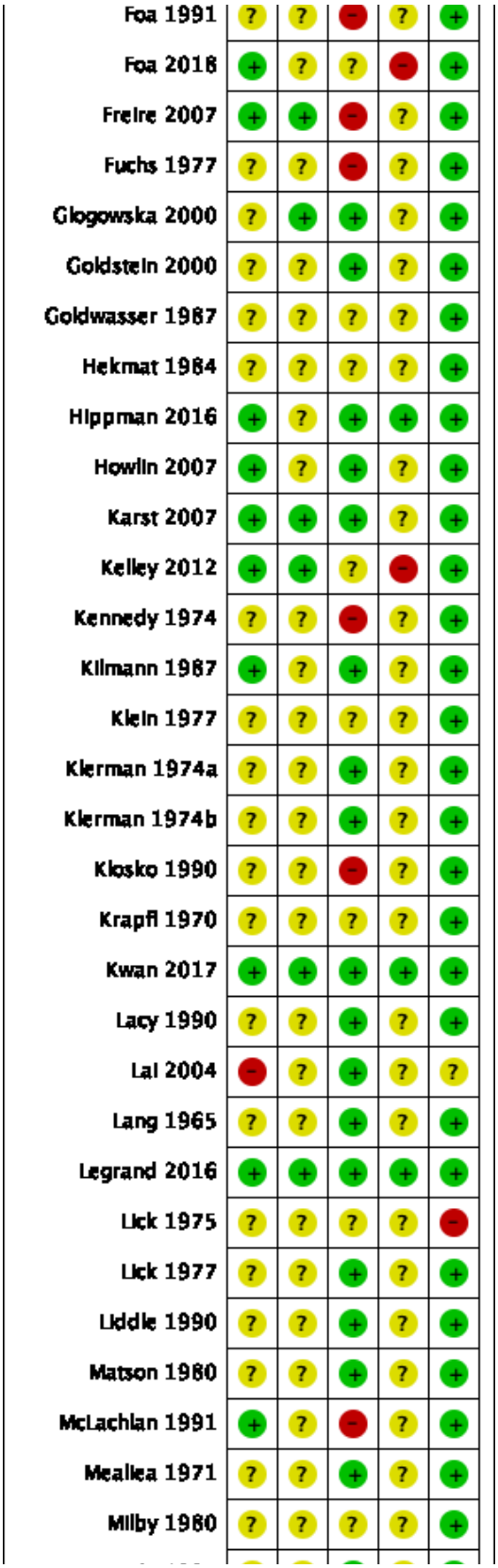


Figure 6. (Continued)

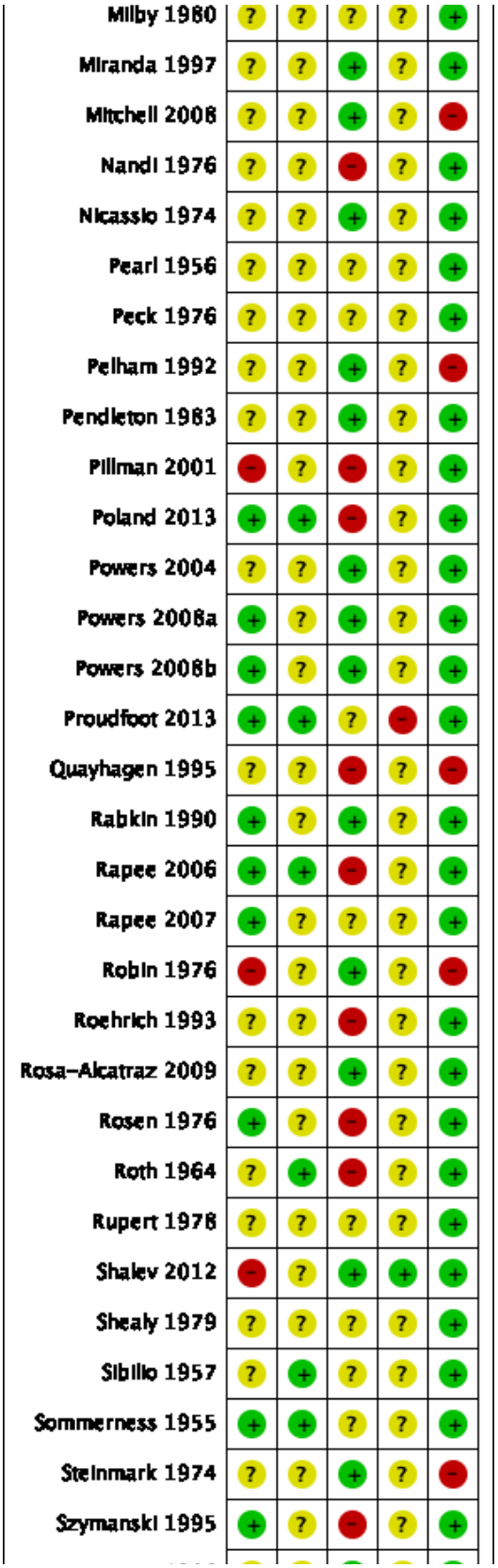
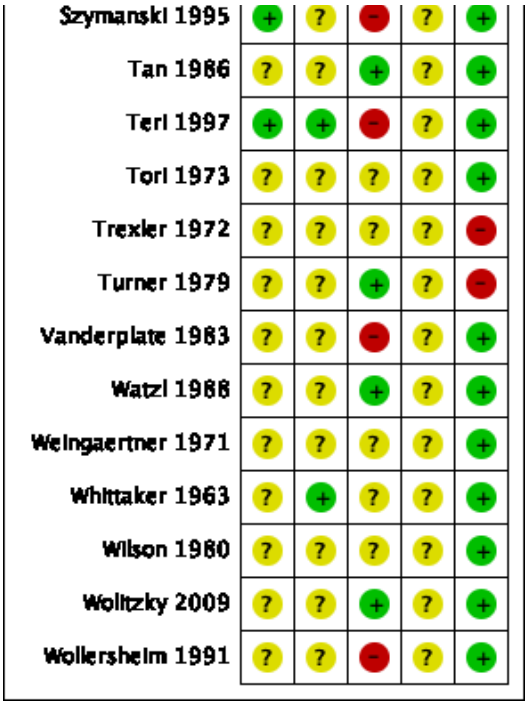


Figure 6. (Continued)



Many trials had small sample sizes which led to imprecise estimates. The funnel plot for the comparison on all placebos was somewhat asymmetrical, which may reflect poor methodological quality, true heterogeneity, or selection bias (Higgins 2019). Upon visual inspection, the funnel plot indicated that some data points might be missing on the lower right corner of the plot, and this could have important implications for the interpretability of the data. The results in this comparison may be sensitive to trial size and should be interpreted with caution.

Almost all the included trials had three arms (interventions) with one experimental arm. Potential conflicts of interest when the researches had vested interests in the experimental intervention, financially or non-financially, may have produced bias and threatened the validity of the results (Lundh 2017; Boutron 2021). We conducted a post-hoc analysis to test conflicts of interest, but found no significant differences, but the data in the subgroup analysis were sparse. We should therefore not rule out the possibility of conflicts of interest towards the experimental intervention in some of the trials having impacted effect estimates. We graded all the overall results as either very low or low quality of evidence according to GRADE based on risk of bias, inconsistency of the evidence, and imprecision (Guyatt 2011a).

There are also many potential issues with reporting in the included trials. For instance, usual care, wait-list and no-treatment can vary in their contents and trial authors do not always specify how their control interventions were designed (Cuijpers 2021; Watts 2015), which makes it hard to determine how much overlap there is between the controls in the review. This should not, however, be an argument against the use of controls such as psychological placebos in trials with mental health populations because they may be methodologically useful for differentiating the active and non-active factors in psychological treatments (Mohr 2009). Instead, controls such as psychological placebos should arguably be designed to control for everything but the hypothesised mechanism of causality in a psychological treatment (Hróbjartsson 2012; Locher 2018). Another potential issue is with the diagnostic classification systems used in the included studies, as some studies used older versions of the Diagnostic and Statistical Manual of Mental Disorders (DSM), which are not necessarily comparable to current versions.

Potential biases in the review process

Because of the large amount of records to screen, we chose to single-screen records. This may have affected the selection of results and produced bias. It is very challenging to locate accurately all relevant records with such as broad search strategy (two control interventions in a three-arm randomised trial with any mental health disorder), and we may have missed relevant trials. However, for the placebo interventions, we did include 50 more trials for mental health disorders than Hróbjartsson 2010 reported in an earlier review on placebos for all medical conditions.

During the inclusion phase, we identified trials where patients fulfilled the symptoms of a mental health disorder according to International Classification of Diseases (ICD) or DSM, but where the trial did not explicitly state what classification system was used. In some instances we also included trials where the population was classified as having a mental health disorder but the full diagnostic symptoms were not reported. We chose to include these trials in our analyses, and included a post-hoc subgroup analysis to investigate potential differences.

There were some more minor changes from the protocol to the final report and these are all listed in Differences between protocol and review and Table 6.

Agreements and disagreements with other studies or reviews

The earlier reviews by Hróbjartsson and Gøtzche (Hróbjartsson 2001; Hróbjartsson 2002; Hróbjartsson 2004; Hróbjartsson 2010) found that the efficacy of their included placebo interventions yielded, on average, a small to moderate effect and that placebos may influence patient-reported outcomes. Our analysis on all placebo types combined for continuous outcomes yielded an SMD of -0.37 (95% CI -0.49 to -0.25), which is not considerably different from the results found in Hróbjartsson and Gøtzsche's work (SMD -0.23, 95% CI -0.28 to -0.17) for all placebos combined on continuous outcomes. Both these effect estimates are small to moderate. They also found moderate statistical heterogeneity present in both the binary and continuous outcomes and funnel plot asymmetry in the analyses on continuous outcomes, which is similar to the findings in the present review.

We found some indication of funnel plot asymmetry in the comparison on all placebos, moderate heterogeneity in most analyses and high variability between effect sizes. In this review, psychological and physical placebos showed larger differences compared with no-treatment or wait-list than pharmacological placebos, which is also similar to what was found in Hróbjartsson 2010 . It could be true that placebo interventions yield larger differences compared with wait-list or no-treatment in randomised trials with mental health populations, perhaps because they involve more subjective outcome measures or that mental health disorders are more prone to be affected by placebo administration (Weimer 2015). Our subgroup analyses for all placebos combined indicated that blinded and non-blinded observer-reported outcomes provided a higher placebo effect size compared with patient-reported outcomes. Hróbjartsson 2010 found the opposite: a higher placebo effect for patient-reported outcomes over observer-reported. Here it should be noted that the observer-reported outcomes in Hróbjartsson 2010 often measured a somatic variable, whereas observer-reported outcomes in the present review were often psychometric instruments rated by an observer. Thus, the conflicting results from the reviews may not be directly comparable.

Placebo and usual care control groups compared with wait-list controls were found to yield higher effect estimates than compared with no-treatment controls. The findings were only significant for all placebos combined and for psychological placebos. Our review found similarities to the work of Mohr and colleagues (Mohr 2014). In their meta-analysis focusing on studies on depression, they found significant differences in effect sizes generated across different control interventions. Another recent network meta-analysis that assessed control intervention's influence on effect estimates of active psychotherapies for depression found weaker effect estimates for wait-list and no-treatment than psychological and pharmacological placebos (Michopoulos 2021). Our review supports these findings and the importance of considering the type of control intervention in a randomised trial with mental health populations because it can drastically influence reported effect estimates. Cuijpers and Cristea has also proposed that to ensure a higher effect estimate of the active treatment in a randomised trial, a wait-list intervention should be preferred (Cuijpers 2016). Wait-list control interventions might bias the true effect of different active treatments and therefore potentially produce a skewed view of the effect of those treatments, but we are not able to make conclusions about this based on the results in this review.

Little research has been done on the harms of psychological treatments (Lilienfeld 2007; Linden 2014; Pagsberg 2017; Storebø 2018). We did not find indications of wait-list or no-treatment interventions being more harmful than any placebos. In the case of usual care, no trials reported or mentioned adverse events, and we are only able to give anecdotal evidence on adverse events here. Usual care is a highly heterogeneous control intervention and very few of our included trials accurately reported its contents because the researchers are often unaware of this themselves. Furukawa and colleagues previously speculated that wait-list interventions could lead to negative effects in patients from waiting for an experimental treatment after the study period (Furukawa 2014), but our review was not able to confirm this because of sparse adverse event data. It should be a priority to identify whether wait-list interventions might produce unfavourable harms in randomised trials with participants with mental health diagnoses in future research.

It has previously been argued that a decision framework should be put in place for how to properly choose a control intervention in trials with patients with mental health disorders, and that such a framework should take into account factors like trial phase, participation risk and available levels of resources (Gold 2018). It is evident from this review that the effect sizes in trials with patients with mental health disorders may vary widely depending on what control intervention is used, and it seems reasonable, therefore, to demand methodological standards for when it is appropriate to use a particular control in a trial. It may for instance be recommended that wait-list or no-treatment controls should only be used in the early stages of testing a new behavioural treatment, seeing that they often produce high effect sizes in favour of experimental interventions, which may give a misguided impression of the intervention's effectiveness. Overall, it seems important that a control intervention should be properly designed and tailored to the specific objectives under investigation, and that there should be some agreement among researchers on when a type of control design is appropriate in a mental health trial and when it is not (Mohr 2009; Mohr 2014).

AUTHORS' CONCLUSIONS

Implication for systematic reviews and evaluations of healthcare

The choice of a control intervention in randomised trials with patients with mental health disorders has a considerable impact on the reported estimate of benefits in published reports. When psychiatric interventions are compared with some kind of placebo intervention, the beneficial effect of the psychiatric intervention is lower than when compared with other control interventions. The difference in effect size for the experimental intervention might be approximately a standardised mean difference (SMD) value from 0.3 to 0.4 lower when using a placebo control intervention compared with wait-list or no intervention.

Mental health systematic reviews and evaluations of healthcare should put equal emphasis on the reporting of the contents of control interventions, because they may have the same influence on effect size estimates as experimental interventions. This may be especially true for reviews dealing with psychosocial interventions,

where the contents of control interventions such as usual care or psychological placebos are often underreported and unclear. People using reviews and evaluations should therefore be aware of how the choice of a control intervention (such as wait-list controls or placebo controls) impacts on the reported effect estimates, both unfavourably and favourably. The evidence in the included trials in this review were rated as low to very low quality, only partly because of the inability to blind participants in randomised trials with a no-treatment or wait-list comparator. The issue of blinding may, however, be viewed as a methodological issue with the type of control design in mental health intervention research and not a flaw in the interpretability of this review itself.

The choice of a control intervention in a randomised trial in patients with mental health disorders has a considerable impact on the reported estimate of benefits in the published reports. Placebos or usual care tend to increase the differences compared with wait-list or no-treatment. Methodological guidelines need to be developed to reach a consensus on future standards for the design and reporting of control interventions in the field of mental health.

Implication for methodological research

Currently, descriptions of both the experimental and control interventions are often poorly reported (Hoffmann 2013) and they need to be more adequately described (Guidi 2018). Methodological guidelines on how to properly report and design control interventions in randomised trials in mental health research are needed to advance the evidence base in the field. An adequate and systematised description of interventions would provide a platform for researchers to build on findings about control intervention design or replicate results (Hoffmann 2013; Tajika 2015). Control interventions should be developed to answer the specific research question at hand and should be chosen based on available resources, ethical concerns and the phase of research for a particular treatment (Gold 2018). Future research should support the development of a methodological guideline on how to properly design and report control interventions in randomised trials with patients with mental health disorders, to ensure the validity and reliability of future mental health trials.

ACKNOWLEDGEMENTS

We thank the Cochrane Methodology Review Group's editorial team and peer reviewers. We thank Richard Kirubakaran for helping with the analysis regarding Egger's test. We also thank research librarian Trine Kæstel for her valuable guidance on the search strategy.

REFERENCES

References to studies included in this review

Abikoff 2004 *{published data only}*

* Abikoff H, Hechtman L, Klein RG.Social functioning in children with ADHD treated with long-term methylphenidate and multimodal psychosocial treatment. *Journal of the American Academy of Child and Adolescent Psychiatry* 2004;**43**(7):820-9.

Abikoff H.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to A Todorovac 23rd of August 2019.

Klein RG, Abikoff H, Hechtman L, Weiss G.Design and rationale of controlled study of long-term methylphenidate and multimodal psychosocial treatment in children with ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry* 2004;**43**(7):792-801.

Allen 1998 *{published data only}*

Allen JB.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to A Todorovac 28th of August 2019.

Allen JJ, Schnyer RN, Hitt SK.The efficacy of acupuncture in the treatment of major depression in women. *Psychological Science* 1998;**9**(5):397-401.

Allen 2006 *{published data only}*

Allen JB.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 28th of August 2019.

Allen JJ, Schnyer RN, Chambers AS, Hitt SK, Moreno FA, Manber R.Acupuncture for depression: a randomized controlled trial. *Journal of Clinical Psychiatry* 2006;**67**(11):1665-73.

Alvarez 1997 *{published data only}*

Alvarez JR.The effects of anger management training on depression, anger, self-concept, and treatment responsibility in a sample of drug-dependent individuals in a therapeutic community. *Dissertation Abstracts International: Section B: the Sciences and Engineering* 1997;**58**(4-B):2108.

Ascher 1979 *{published data only}*

Ascher LM, Turner R.Paradoxical intention and insomnia: an experimental investigation. *Behaviour Research and Therapy* 1979;**17**(4):408-11.

Ayen 2004 *{published data only}*

Ayen I, Hautzinger M.Cognitive behavior therapy for depression in menopausal women. A controlled, randomized treatment study [Kognitive Verhaltenstherapie bei Depressionen im Klimakterium: Eine kontrollierte, randomiserte Interventionsstudie]. *Zeitschrift für Klinische Psychologie und Psychotherapie* 2004;**33**(4):290-9.

Ayen I.Inclusion of study in Cochrane Systematic Review (personal communication). Email to EG Faltinsen 9th of September 2019.

Berg 1983 *{published data only}*

Berg I, Forsythe I, Holt P, Watts J.A controlled trial of 'Senokot' in faecal soiling treated by behavioural methods. *Journal of Child Psychology and Psychiatry* 1983;**24**(4):543-9.

Borden 1986 *{published data only}*

* Borden, KA.Attributional Outcomes: the Subtle Messages of Treatments for Attention Deficit Disorder. Dissertation, University of Illinois at Urbana-Champaign 1986.

Borden KA, Brown RT.Attributional outcomes: The subtle messages of treatments for attention deficit disorder. *Cognitive Therapy and Research* 1989;**13**(2):147-60.

Borkovec 1975 *{published data only}*

Borkovec TD, Kaloupek DG, Slama KM.The facilitative effect of muscle tension-release in the relaxation treatment of sleep disturbance. *Behavior Therapy* 1975;**6**(3):301-9.

Borkovec 1976 *{published data only}*

Borkovec TD, Weerts TC.Effects of progressive relaxation on sleep disturbance: an electroencephalographic evaluation. *Psychosomatic Medicine* 1976;**38**(3):173-80.

Bornovalova 2008 *{published data only}*

Bornovalova M.Inclusion of study in Cochrane Systematic Review (personal communication). Email to A Todorovac. 21st of August 2019.

Bornovalova MA.Distress tolerance treatment for inner-city drug users: a preliminary trial. Doctor's Dissertation, University of Maryland 2008.

Bramston 1985 *{published data only}*

Bramston P, Spence SH.Behavioural versus cognitive social-skills training with intellectually-handicapped adults. *Behaviour Research and Therapy* 1985;**23**(3):239-46.

Brill 1964a *{published data only}*

* Brill NQ, Koegler RR, Epstein LJ, Forgy EW.Controlled study of psychiatric outpatient treatment. *Archives of General Psychiatry* 1964;**10**:581-95.

Brill NQ.Results of psychotherapy. *California Medicine* 1966;**104**(4):249-53.

Brill 1964b *{published data only}*

Brill NQ, Koegler RR, Epstein LJ, Forgy EW.Controlled study of psychiatric outpatient treatment. *Archives of General Psychiatry* 1964;**10**:581-95.

Carlson 1993 *{published data only}*

Carlson CL, Pelham WE, Milich R, Hoza B.ADHD boys' performance and attributions following success and failure: drug effects and individual differences. *Cognitive Therapy and Research* 1993;**17**:269-87.

Carter 2003 *{published data only}*

Carter JC, Olmsted MP, Kaplan AS, McCabe RE, Mills JS, Aime A.Self-help for bulimia nervosa: a randomized controlled trial. *American Journal of Psychiatry* 2003;**160**(5):973-8.

Crisp 1991 *{published data only}*

* Crisp AH, Norton K, Gowers S, Halek C, Bowyer C, Yeldham D, et al.A controlled study of the effect of therapies aimed at adolescent and family psychopathology in Anorexia Nervosa. *British Journal of Psychiatry: the Journal of Mental Science* 1991;**159**:325-33.

Gowers S, Norton K, Halek C, Crisp AH.Outcome of outpatient psychotherapy in a random allocation treatment study of anorexia nervosa. *International Journal of Eating Disorders* 1994;**15**(2):165-77.

Gowers S, Norton K, Yelsham D.The St George's prospective treatment study of anorexia nervosa: a discussion of methodological problems. *International Journal of Eating Disorders* 1989;**8**:445-54.

Crouch 1988 *{published data only}*

* Crouch G, Robson M, Hallstrom C.Benzodiazepine dependent patients and their psychological treatment. *Progress in Neuro-psychopharmacology & Biological Psychiatry* 1988;**12**(4):503-10.

Hallström C, Crouch G, Robsen M, Shine P.The treatment of tranquillizer dependence by propranolol. *Postgraduate Medical Journal* 1988;**64**(Suppl 2):40-4.

Doty 1975 *{published data only}*

Doty, DW.Role playing and incentives in the modification of the social interaction of chronic psychiatric patients. *Journal of Consulting and Clinical Psychology* 1975;**43**(5):676-82.

Double 1993 *{published data only}*

Double DB, Warren GC, Evans M, Rowlands MP.Efficacy of maintenance use of anticholinergic agents. *Acta Psychiatrica Scandinavica* 1993;**88**:381-4.

Ehlers 2014 *{published data only}*

Ehlers A, Hackmann A, Grey N, Wild J, Liness S, Albert I, et al.A randomized controlled trial of 7-day intensive and standard weekly cognitive therapy for PTSD and emotion-focused supportive therapy. *American Journal of Psychiatry* 2014;**171**(3):294-304.

Espie 1989a *{published data only}*

Espie CA, Brooks DN, Lindsay WR.An evaluation of tailored psychological treatment of insomnia. *Journal of Behavior Therapy and Experimental Psychiatry* 1989;**20**(2):143-53.

* Espie CA, Lindsay WR, Brooks DN, Hood EM, Turvey T.A controlled comparative investigation of psychological treatments for chronic sleep-onset insomnia. *Behaviour Research and Therapy* 1989;**27**(1):79-88.

Etringer 1982 *{published data only}*

Etringer BD, Cash TF, Rimm DC.Behavioral, affective, and cognitive effects of participant modeling and an equally credible placebo. *Behavior Therapy* 1982;**13**(4):476-85.

Foa 1991 *{published data only}*

Foa EB, Rothbaum BO, Riggs DS, Murdock TB.Treatment of posttraumatic stress disorder in rape victims: a comparison between cognitive-behavioral procedures and counselling. *Journal of Consulting and Clinical Psychology* 1991;**59**(5):715-23.

Foa 2018 *{published data only}*

Foa EB, McLean CP, Zang Y, Rosenfield D, Yadin E, Yarvis JS, et al.Effect of prolonged exposure therapy delivered over 2 weeks vs 8 weeks vs present-centered therapy on PTSD symptom severity in military personnel: a randomized clinical trial. *JAMA* 2018;**319**(4):354-64.

Freire 2007 *{published data only}*

Freire AO, Sugai GC, Chrispin FS, Togeiro SM, Yamamura Y, Mello LE, et al.Treatment of moderate obstructive sleep apnea syndrome with acupuncture: a randomised, placebo-controlled pilot trial. *Sleep Medicine* 2007;**8**(1):43-50.

Fuchs 1977 *{published data only}*

Fuchs CZ, Rehn LP.A self-control behavior therapy program for depression. *Journal of Consulting and Clinical Psychology* 1977;**45**(2):206-15.

Glogowska 2000 *{published data only}*

Glogowska M, Roulstone S, Enderby P, Peters TJ.Randomised controlled trial of community based speech and language therapy in preschool children. *BMJ* 2000;**321**(7266):923-6.

Goldstein 2000 *{published data only}*

Goldstein AJ, Beurs E, Chambless Dl, Wilson KA.EMDR for panic disorder with agoraphobia: comparison with waiting list and credible attention-placebo control conditions. *Journal of Consulting and Clinical Psychology* 2000;**68**(6):947-56.

Goldwasser 1987 *{published data only}*

Goldwasser AN, Auerbach SM, Harkins SW.Cognitive, affective, and behavioural effects of reminiscence group therapy on demented elderly. *Aging and Human Development* 1987;**25**(3):209-22.

Hekmat 1984 *{published data only}*

Hekmat H, Lubitz R, Deal R.Semantic desensitization: a paradigmatic intervention approach to anxiety disorders. *Journal of Clinical Psychology* 1984;**40**(2):463-6.

Hippman 2016 *{published data only}*

Hippman C, Ringrose A, Inglis A, Cheek J, Albert AY, Remick R, et al.A pilot randomized clinical trial evaluating the impact of genetic counselling for serious mental illnesses. *Journal of Clinical Psychiatry* 2016;**77**(2):e190-8.

Howlin 2007 *{published data only}*

Howlin P, Gordon KR, Pasco G, Wade A, Charman T.The effectiveness of picture exchange communication system (PECS) training for teachers of children with autism: a pragmatic,group randomised controlled trial. *Journal of Child Psychology and Psychiatry* 2007;**48**(5):473-81.

Karst 2007 *{published data only}*

Karst M, Winterhalter M, Münte S, Francki B, Hondronikos A, Eckardt A, et al.Auricular acupuncture for dental anxiety: a randomized controlled trial. *Anesthesia and Analgesia* 2007;**104**(2):295-300.

Karst M.Inclusion of study in Cochrane Systematic Review (personal communication). Email to: EG Faltinsen 5th of September 2019.

Kelley 2012 *{published data only}*

Kelley J.Inclusion of study in Cochrane Systematic Review (personal communication). Email to: A Todorovac 12th of August 2019.

Kelley JM, Kaptchuk TJ, Cusin C, Lipkin S, Fava M.Open-label placebo for major depressive disorder: a pilot randomized controlled trial. *Psychotherapy and Psychosomatics* 2012;**81**(5):312-4.

Kennedy 1974 *{published data only}*

Kennedy TD, Kimura HK.Transfer, behavioral improvement, and anxiety reduction in systematic desensitization. *Journal of Consulting and Clinical Psychology* 1974;**42**(5):720-8.

Kilmann 1987 *{published data only}*

* Kilmann PR, Milan RJ, Boland JP, Nankin HR, Davidson E, West MO, et al.Group treatment of secondary erectile dysfunction. *Journal of Sex & Marital Therapy* 1987;**13**(3):168-82.

Kilmann PR, Milan RJ Jr, Boland JP, Davidson E, West MO, Sablis RF, et al.The sexual interactions of men with secondary erectile dysfunction. *Journal of Sex Education and Therapy* 1988;**14**(1):46-9.

Kilmann PR, Milan RJ Jr, Boland JP, Nankin HR, Sablis RF, West MO, et al.The sexual interactions of men with secondary erectile dysfunction. Paper presented at the annual meeting of the American Psychological Association in Los Angeles 1985;**1**:46-9.

Klein 1977 *{published data only}*

Klein SA, Deffenbacher JL.Relaxation and exercise for hyperactive impulsive children. *Perceptual and Motor Skills* 1977;**45**:1159-62.

Klerman 1974a *{published data only}*

* Klerman GL, Dimascio A, Weismann M, Prusoff B, Paykel ES.Treatment of depression by drugs and psychotherapy. *American Journal of Psychiatry* 1974;**131**(2):186-91.

Paykel ES, Dimascio A, Haskell D, Prusoff BA.Effects of maintenance amitriptyline and psychotherapy on symptoms of depression. *Psychological Medicine* 1975;**5**(1):67-77.

Paykel ES, DiMascio A, Klerman GL, Prusoff BA.Maintenance therapy of depression. *Pharmakopsychiatrie, Neuro-Psychopharmakologie* 1976;**9**:127-39.

Paykel ES, Weissman MM, Prusoff BA.Social maladjustment and severity of depression. *Comprehensive Psychiatry* 1978;**19**(2):121-8.

Weissman MM, Klerman GL, Paykel ES, Prusoff B.Treatment effects on the social adjustment of depressed patients. *Archives of General Psychiatry* 1974;**30**:771-8.

Klerman 1974b *{published data only}*

Klerman GL, Dimascio A, Weismann M, Prusoff B, Paykel ES.Treatment of depression by drugs and psychotherapy. *American Journal of Psychiatry* 1974;**131**(2):186-91.

Klosko 1990 *{published data only}*

Klosko JS, Barlow DH, Tassinari R, Cerny JA.A comparison of alprazolam and behavior therapy in treatment of panic disorder: correction. *Journal of Psychotherapy Practice and Research* 1994;**3**(2):163-79.

* Klosko JS, Barlow DH, Tassinari R, Cerny JA.A comparison of alprazolam and behavior therapy in treatment of panic disorder. *Journal of Consulting and Clinical Psychology* 1990;**58**(1):77-84.

Krapfl 1970 *{published data only}*

Krapfl JE, Nawas MM.Differential ordering of stimulus presentation in systematic desensitization. *Journal of Abnormal Psychology* 1970;**75**(3):333-7.

Kwan 2017 *{published data only}*

Kwan R.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 26th of August 2019.

Kwan RY, Leung MC, Lai CK.Acupressure for agitation in nursing home residents with dementia: study protocol for a randomized controlled trial. *Trials* 2014;**15**(410):1-7.

* Kwan RY, Leung MCP, Lai CK.A randomized controlled trial examining the effect of acupressure on agitation and salivary cortisol in nursing home residents with dementia. *Dementia and Geriatric Cognitive Disorders* 2017;**44**(1-2):92-104.

Lacy 1990 *{published data only}*

Lacy MK.The effect of short term autogenic training on the stress reactivity of hospitalized emotionally disturbed adolescents. ProQuest Dissertations Publishing, Oklahoma State University 1996.

Lai 2004 *{published data only}*

Lai CK, Chi I, Kayser-Jones J.A randomized controlled trial of a specific reminiscence approach to promote the well-being of nursing home residents with dementia. *International Psychogeriatrics* 2004;**16**(1):33-49.

Lai K.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 6th of September 2019.

Lang 1965 *{published data only}*

Lang OJ, Lazovik AD.Experimental desensitization of a phobia. *Journal of Abnormal and Social Psychology* 1963;**66**:519-25.

* Lang PJ, Lazovik AD, Reynolds DJ.Desensitization, suggestibility, and pseudotherapy. *Journal of Abnormal Psychology* 1965;**70**(6):395-402.

Legrand 2016 *{published data only}*

Legrand F.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 14th of August 2019.

Legrand FD, Neff EM.Efficacy of exercise as an adjunct treatment for clinically depressed inpatients during the initial stages of antidepressant pharmacotherapy: an open randomized controlled trial. *Journal of Affective Disorders* 2016;**191**:139-44.

Lick 1975 *{published data only}*

Lick J.Expectancy, false galvanic skin response feedback, and systematic desensitization in the modification of phobic behavior. *Journal of Consulting and Clinical Psychology* 1975;**43**(4):557-67.

Lick 1977 *{published data only}*

Lick JR, Heffler D.Relaxation training and attention placebo in the treatment of severe insomnia. *Journal of Consulting and Clinical Psychology* 1977;**45**(2):153-61.

Liddle 1990 *{published data only}*

Liddle B, Spence SH.Cognitive-behaviour therapy with depressed primary school children: A cautionary note. *Behavioural Psychotherapy* 1990;**18**(2):85-102.

Matson 1980 *{published data only}*

Matson JL, Marchetti A, Adkins JA.Comparison of operant- and independence-training procedures for mentally retarded adults. *American Journal of Mental Deficiency* 1980;**84**(5):487-94.

McLachlan 1991 *{published data only}*

* McLachlan DR, Dalton AJ, Kruck TP, Bell MY, Smith WL, Kalow W, et al.Intramuscular desferrioxamine in patients with Alzheimer's disease. *Lancet* 1991;**337**(8753):1304-8.

McLachlan DRC, Smith WL, Kruck TP.Desferrioxamine and Alzheimer's disease: video home behavior assessment of clinical course and measures of brain aluminum. *Therapeutic Drug Monitoring* 1993;**15**(6):602-7.

Mealiea 1971 *{published data only}*

Mealiea WL, Nawas MM.The comparative effectiveness of systematic desensitization and implosive therapy in the treatment of snake phobia. *Journal of Behavior Therapy and Experimental Psychiatry* 1971;**2**(2):85-94.

Milby 1980 *{published data only}*

Milby JB, Clarke C, Toro C, Thornton S, Rickert D.Effectiveness of urine surveillance as an adjunct to outpatient psychotherapy for drug abusers. *International Journal of the Addictions* 1980;**15**(7):993-1001.

Miranda 1997 *{published data only}*

Miranda A, Arlandis P, Soriano M.Instruction in strategies and attributional training: Effects on the problem-solving and self-concept of students' with learning disabilities [Instrucción en estrategias y entrenamiento atribucional: efectos sobre la resolución de problemas y el autoconcepto de los estudiantes con dificultades en el aprendizaje]. *Infancia y Aprendizaje* 1997;**20**(80):37-52.

Mitchell 2008 *{published data only}*

Mitchell C, Mansfield D, Rautenbach S.Coloured filters and reading accuracy, comprehension and rate: a placebo-controlled study. *Perceptual and Motor Skills* 2008;**106**(2):517-32.

Nandi 1976 *{published data only}*

Nandi DN, Ajmany S, Ganguli H, Banerjee G, Boral GC, Ghosh A, et al.A clinical evaluation of depressives found in a rural survey in India. *British Journal of Psychiatry* 1976;**128**:523-7.

Nicassio 1974 *{published data only}*

Nicassio P, Bootzin R.A comparison of progressive relaxation and autogenic training as treatments for insomnia. *Journal of Abnormal Psychology* 1974;**83**(3):253-60.

Pearl 1956 *{published data only}*

Pearl D, Vander Kamp H, Olsen AL, Greenberg PD, Armitage SG.The effects of reserpine on schizophrenic patients. *American Journal of Psychiatry* 1956;**112**:198-204.

Peck 1976 *{published data only}*

Peck CL.Desensitization for the treatment of fear in the high level adult retardate. *Behaviour Research and Therapy* 1977;**15**(2):137-48.

Pelham 1992 *{published data only}*

Milich R, Carlson CL, Pelham WE Jr, Licht BG.Effects of methylphenidate on the persistence of ADHD boys following failure experiences. *Journal of Abnormal Child Psychology* 1991;**19**(5):519-36.

* Pelham WE, Murphy DA, Vannatta K, Milich R, Licht BG, Gnagy EM, et al.Methylphenidate and attributions in boys with attention-deficit hyperactivity disorder. *Journal of Consulting and Clinical Psychology* 1992;**60**(2):282-92.

Pendleton 1983 *{published data only}*

Pedleton MG, Higgins RL.A comparison of negative practice and systematic desensitization in the treatment of acrophobia. *Journal of Behavior Therapy and Experimental Psychiatry* 1983;**14**(4):317-23.

Pillman 2001 *{published data only}*

Peterson MA, Patterson B, Pillman BM, Battista MA.Cognitive recovery following alcohol detoxification: a computerised remediation study. *Neuropsychological Rehabilitation* 2002;**12**(1):63-74.

* Pillman BM.Computerized cognitive remediation with recently detoxified alcoholics. Dissertation, California School of Professional Psychology, Fresno Campus 2000.

Poland 2013 *{published data only}*

Poland RE, Gertsik L, Favreau JT, Smith SI, Mirocha JM, Rao U, et al.Open-label randomized parallel-group controlled clinical trial of massage for treatment of depression in HIV-infected subjects. *Journal of Alternative and Complementary Medicine* 2013;**19**(4):334-40.

Powers 2004 *{published data only}*

Powers MB, Smits JA, Telch MJ.Disentangling the effects of safety-behavior utilization and safety-behavior availability during exposure-based treatment: a placebo-controlled trial. *Journal of Consulting and Clinical Psychology* 2004;**72**(3):448-54.

Powers 2008a *{published data only}*

Powers MB, Smiths JA, Whitley D, Bystritsky A, Telch MJ.The effect of attributional processes concerning medication taking on return of fear. *Journal of Consulting and Clinical Psychology* 2008;**76**(3):478-90.

Powers 2008b *{published data only}*

Powers MB, Smiths JA, Whitley D, Bystritsky A, Telch MJ.The effect of attributional processes concerning medication taking on return of fear. *Journal of Consulting and Clinical Psychology* 2008;**76**(3):478-90.

Proudfoot 2013 *{published data only}*

Proudfoot J, Clarke J, Birch MR, Whitton AE, Parker G, Manicavasagar V, et al.Impact of a mobile phone and web program on symptom and functional outcomes for people with mild- to-moderate depression, anxiety and stress: a randomized controlled trial. *BMC Psychiatry* 2013;**13**(312):1-12.

Proudfoot J.Re: Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 26th of August 2019.

Quayhagen 1995 *{published data only}*

Quayhagen MP, Quayhagen M, Corbeil RR, Roth PA, Rodgers JA.A dyadic remediation program for care recipients with dementia. *Nursing Research* 1995;**44**(3):153-9.

Rabkin 1990 *{published data only}*

Liebowitz M, Quitkin FM, Stewart JW, McGrath PJ, Harrison WM, Markowitz JS, et al.Antidepressant efficacy in atypical depression. *Archives of General Psychiatry* 1988;**45**(2):129-37.

Quitkin FM, McGrath PJ, Stewart JW, Harrison W, Wager SG, Nunes E, et al.Phenelzine and imipramine in mood reactive depressives: further delineation of the syndrome of atypical depression. *Archives of General Psychiatry* 1989;**46**(9):787-93.

* Rabkin JG, McGrath PJ, Quitkin FM, Tricamo E, Stewart JW, Klein DF.Effects of pill-giving on maintenance of placebo response in patients with chronic mild depression. *American Journal of Psychiatry* 1990;**147**(12):1622-6.

Rapee 2006 *{published data only}*

Rapee R.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 29th of August 2019.

Rapee RM, Abbott MJ, Lyneham HJ.Bibliotherapy for children With anxiety disorders using written materials for parents: a randomized controlled trial. *Journal of Consulting and Clinical Psychology* 2006;**74**(3):436-44.

Rapee 2007 *{published data only}*

Rapee R.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 29th of August 2019.

Rapee RM, Abbott MJ, Baillie AJ, Gaston JE.Treatment of social phobia through pure self-help and therapist-augmented self-help. *British Journal of Psychiatry* 2007;**191**:246-52.

Robin 1976 *{published data only}*

Robin A.Rationing out-patients: a defence of the waiting list. *British Journal of Psychiatry* 1976;**129**:138-41.

Roehrich 1993 *{published data only}*

Roehrich L, Goldman MS.Experience-dependent neuropsychological recovery and the treatment of alcoholism. *Journal of Consulting and Clinical Psychology* 1993;**61**(5):812-21.

Rosa-Alcatraz 2009 *{published data only}*

Olivares J.Program IAFS. Protocol for the treatment of social phobia in adolescents [Programa IAFS. Protocolo para el tratamiento de la fobia social enadolescentes]. Madrid: Pirámide 2005;**1**.

* Rosa-Alcázar AI, Olivares-Olivares PJ, Olivares J.The role of non-specific effects in the psychological treatment of adolescents with social phobia [Papel de los efectos inespecíficos en el tratamiento psicológico de adolescentes con fobia social]. *Anuario de Psicologia* 2009;**40**(1):43-61.

Rosen 1976 *{published data only}*

Rosen GM, Glasgow RE, Barrera M.A controlled study to assess the clinical efficacy of totally self-administrated systematic desensitization. *Journal of Consulting and Clinical Psychology* 1976;**44**(2):208-17.

Roth 1964 *{published data only}*

Lorr M, McNair DM, Weinstein GJ.Early effects of chlordiazepoxide (Librium) used with psychotherapy. *Journal of Psychiatric Research* 1963;**1**:257-70.

* Roth I, Rhudick PJ, Shaskan DA, Slobin MS, Wilkinson AE, Young HH.Long-term effects on psychotherapy of initial treatment conditions. *Journal of Psychiatric Research* 1964;**2**:283-97.

Rupert 1978 *{published data only}*

Rupert PA, Holmes DS.Effects of multiple sessions of true and placebo heart rate biofeedback training on the heart rates and anxiety levels of anxious patients during and following treatment. *Psychophysiology* 1978;**15**(6):582-90.

Shalev 2012 *{published data only}*

Shalev AY, Ankri Y, Isreali-Shalev Y, Peleg T, Adessky R, Freedman S.Prevention of posttraumatic stress disorder by early treatment. *Archives of General Psychiatry* 2012;**69**(2):166-76.

Shealy 1979 *{published data only}*

Shealy RC.The effectiveness of various treatment techniques on different degrees and durations of sleep-onset insomnia. *Behaviour Research and Therapy* 1979;**17**(6):541-6.

Sibilio 1957 *{published data only}*

Sibilio JP, Andrew G, Dart D, Moore KB, Stehman A.Treatment of chronic schizophrenia with promazine hydrochloride. *AMA Archives of Neurology and Psychiatry* 1957;**78**(4):419-24.

Sommerness 1955 *{published data only}*

Sommerness MD, Lucero RJ, Hamlon JS, Erickson JL, Matthews R.A controlled study of reserpine on chronically disturbed patients. *JAMA Archives of Neurology and Psychiatry* 1955;**74**(3):316-9.

Steinmark 1974 *{published data only}*

Steinmark SW, Borkovec TD.Active and placebo treatment effects on moderate insomnia under counterdemand and positive demand instructions. *Journal of Abnormal Psychology* 1974;**83**(2):157-63.

Szymanski 1995 *{published data only}*

Szymanski J, O'Donohue W.The potential role of state-dependent learning in cognitive therapy with spider phobics. *Journal of Rational-Emotive & Cognitive-Behavior Therapy* 1995;**13**(2):131-50.

Szymanski J.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 20th of August 2019.

Tan 1986 *{published data only}*

Tan SY, Bruni J.Cognitive-behavior therapy with adult patients with epilepsy: a controlled outcome study. *Epilepsia* 1986;**27**(3):225-33.

Teri 1997 *{published data only}*

Teri L, Logsdon RG, Uomoto J, McCurry SM.Behavioral treatment of depression in dementia patients: a controlled clinical trial. *Journals of Gerontology. Series B, Psychological Sciences and Social Sciences* 1997;**52**(4):159-66.

Teri L.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 26th of August 2019.

Tori 1973 *{published data only}*

Tori C, Worell L.Reduction of human avoidant behavior. A comparison of counterconditioning, expectancy, and cognitive information approaches. *Clinical Psychology* 1973;**41**(2):269-78.

Trexler 1972 *{published data only}*

Trexler, L.Rational-emotive therapy, placebo, and no-treatment effects on public-speaking anxiety. Doctoral dissertation, Temple University AnnArbor, Mich.: University Microfilms 1971;**71-31**(063).

* Trexler LD, Karst TO.Rational-emotive therapy, placebo, and no-treatment effects on public-speaking anxiety. *Journal of Abnormal Psyhology* 1972;**79**(1):60-7.

Turner 1979 *{published data only}*

Ascher ML, Turner RM.A comparison of two methods for the administration of paradoxical intention. *Behaviour Research and Therapy* 1980;**18**(2):121-6.

* Turner RM, Ascher LM.Controlled comparison of progressive relaxation, stimulus control, and paradoxical intention therapies for insomnia. *Journal of Consulting and Clinical Psychology* 1979;**47**(3):500-8.

Vanderplate 1983 *{published data only}*

Vanderplate C.Electromyograph biofeedback Treatment of sleep onset Insomnia: a controlled Investigation. Dissertation, University of Southern Mississippi 1978.

Watzl 1988 *{published data only}*

Watzl H, Rist F, Olbrich R, Cohen R.Placebo injections and surveillance of alcohol intake during inpatient treatment of female alcoholics. *Addictive Behaviors* 1988;**13**(2):197-200.

Weingaertner 1971 *{published data only}*

Weingaertner AH.Self-administered aversive stimulation with hallucinating hospitalized schizophrenics. *Journal of Consulting and Clinical Psychology* 1971;**36**(3):422-29.

Whittaker 1963 *{published data only}*

Whittaker CB, Hoy RM.Withdrawal of perphenazine in chronic schizophrenia. *British Journal of Psychiatry* 1963;**109**:422-7.

Wilson 1980 *{published data only}*

Wilson A, Davidson WJ, Blanchard R, White J.Disulfiram implantation. A placebo-controlled trial with two-year follow-up. *Journal of Studies on Alcohol* 1978;**39**(5):809-19.

* Wilson A, Davidson WJ, Blanchard R.Disulfiram implantation: a trial using placebo implants and two types of controls. *Journal of Studies on Alcohol* 1980;**41**(5):429-36.

Wolitzky 2009 *{published data only}*

Wolitzky KB, Telch MJ.Augmenting in vivo exposure with fear antagonistic actions: a preliminary test. *Behavior Therapy* 2009;**40**(1):57-71.

Wollersheim 1991 *{published data only}*

Wollersheim JP, Wilson GL.Group treatment of unipolar depression: a comparison of coping, supportive, bibliotherapy, and delayed treatment groups. *Professional Psychology: Research and Practice* 1991;**22**(6):496-502.

References to studies excluded from this review

Adelman 1990 *{published data only}*

Adelman HS, MacDonald VM, Nelson P, Smith DC, Taylor L.Motivational readiness and the participation of children with learning and behavior problems in psychoeducational decision making. *Journal of Learning Disabilities* 1990;**23**(3):171-6.

Altenhofer 2007 *{published data only}*

Altenhöfer A, Schulz W, Schwab R, Eckert J.Psychotherapy of adjustment disorders. Is psychotherapy if limited to 12 sessions sufficiently effective? [Psychotherapie von Anpassungsstörungen. Ist eine auf 12 Sitzungen begrenzte Gesprächspsychotherapie ausreichend wirksam?]. *Psychotherapeut* 2007;**52**:24-34.

Andreasson 1990 *{published data only}*

Andréasson S, Parmander M, Allebeck P.A pilot study of treatment of alcohol abuse. Marked improvement of alcohol drinking habits of both treated and untreated persons. *Läkartidningen* 1990;**87**(9):627-30.

Banasiak 2005 *{published data only}*

Banasiak SJ, Paxton SJ, Hay P.Guided self-help for bulimia nervosa in primary care: a randomized controlled trial. *Psychological Medicine* 2005;**35**(9):1283-94.

Bass 2016 *{published data only}*

Bass J, Murray SM, Mohammed TA, Bunn M, Gorman W, Ahmed AM, et al.A randomized controlled trial of a trauma-informed support, skills, and psychoeducation intervention for survivors of torture and related trauma in Kurdistan, Northern Iraq. *Global Health, Science and Practice* 2016;**4**(3):452-66.

de Jongh 1995 *{published data only}*

de Jongh A, Muris P, ter Horst G, van Zuuren F, Schoenmakers N, Makkes P.One-session cognitive treatment of dental phobia: preparing dental phobics for treatment by restructuring negative cognition. *Behaviour Research and Therapy* 1995;**33**(8):947-54.

Dolhanty 2005 *{published data only}*

Dolhanty J.Symptom control and self-efficacy with electronic self-monitoring of bulimia nervosa: a randomized controlled trial. Dissertation, York University, North York, Ontario 2005.

Drozdek 2010 *{published data only}*

Drožđek B, Bolwerk N.Evaluation of group therapy With traumatized asylum seekers and refugees—the Den Bosch model. *Traumatology* 2010;**16**(4):117-27.

Endicott 1964 *{published data only}*

Endicott NA, Endicott J.Prediction of improvement in treated and untreated patients using the Rorschaf Prognostic Rating Scale. *Journal of Consulting Psychology* 1964;**28**:342-8.

Ertl 2011 *{published data only}*

Ertl V, Pfeiffer A, Schauer E, Elbert T, Neuner F.Community-implemented trauma therapy for former child soldiers in northern Uganda. *JAMA* 2011;**306**(5):503-12.

Gitlin 2013 *{published data only}*

Gitlin LN, Harris LF, McCoy M, Chernett NL, Jutkowitz E, Pizzi LT, et al.A community-integrated home based depression intervention for older African Americans: [corrected] Description of the Beat the Blues randomized trial and intervention costs. *BMC Geriatrics* 2012;**10**(4):1-12.

* Gitlin LN, Harris LF, McCoy MC, Chernett NL, Pizzi LT, Jutkowitz E, et al.A home-based intervention to reduce depressive symptoms and improve quality of life in older African Americans: a randomized trial. *Annals of Internal Medicine* 2013;**159**(4):243-52.

Pizzi LT, Jutkowitz E, Frick KD, Suh D, Prioli KM, Gitlin LN.Cost-effectiveness of a community-integrated home based depression intervention in older African Americans. *Journal of the American Geriatrics Society* 2014;**62**(12):2288-95.

Gournay 1994 *{published data only}*

Gournay K, Brooking J.Community psychiatric nurses in primary health care. *British Journal of Psychiatry* 1994;**162**(2):231-8.

Haertzen 1964 *{published data only}*

Haertzen CA.On the addiction research center inventory scores of former addicts receiving LSD and untreated schizophrenics. *Psychological Reports* 1964;**14**(1):483-8.

Hartelius 1968 *{published data only}*

Hartelius H.Objective trial of sedatives: comparison between Noludar, Mogadon, placebo, and "no tablet at all" [Objektiv sömnmedelsprövning: Jämförelse mellan Noludar, Mogadon, placebo och ingen tablett alls]. *Nordisk Medicin* 1968;**79**(25):811-3.

Hinsberger 2017 *{published data only}*

Hinsberger M, Holtzhausen L, Sommer J, Kaminer D, Elbert T, Seedat S, et al.Feasibility and effectiveness of narrative exposure therapy and cognitive behavioral therapy in a context of ongoing violence in South Africa. *Psychological Trauma: Theory, Research, Practice and Policy* 2017;**9**(3):282-91.

Keltner 1976 *{published data only}*

Keltner L, Marshall WL.Attribution and subject control factors in experimental desensitization. *Behavior Therapy* 1976;**7**(5):626-33.

Laessle 1988 *{published data only}*

Laessle RG, Waadt S, Schindler L, Schweiger U, Pirke M.Structured ambulatory group therapy in bulimia [Strukturierte ambulante Gruppentherapie bei Bulimie - Vorläufige Ergebnisse einer kontrollierten Therapiestudie]. *PPmP Psychotherapie, Psychosomatik und Medizinische Psychologie* 1988;**38**:324-7.

Litrownik 1978 *{published data only}*

Litrownik AJ, Freitas JL, Franzini LR.Self-regulation in mentally retarded children: assessment and training of self-monitoring skills. *American Journal of Mental Deficiency* 1978;**82**(5):499-506.

Mak 2008 *{published data only}*

Mak MK, Hui-Chan CW.Cued task-specific training is better than exercise in improving sit-to-stand in patients with Parkinson's Disease: a randomized controlled trial. *Movement Disorders : Official Journal of the Movement Disorder Society* 2007;**23**(4):501-9.

McCall 2005 *{published data only}*

Mccall WV, Perlis ML, Tu X, Groman AE, Krystal A, Walsh JK.A comparison of placebo and no-treatment during a hypnotic clinical trial. *International Journal of Clinical Pharmacology and Therapeutics* 2005;**43**(8):355-9.

McNeil 1991 *{published data only}*

McNeil JK, LeBlanc EM, Joyner M.The effect of exercise on depressive symptoms in the moderately depressed elderly. *Psychology and Aging* 1991;**6**(3):487-8.

Otteson 1979 *{published data only}*

Otteson JP.Curative caring: the use of buddy groups with chronic schizophrenics. *Journal of Consulting and Clinical Psychology* 1979;**47**(3):649-51.

Propst 1992 *{published data only}*

Propst LR, Ostrom R, Watkins P, Dean T, Mashburn D.Comparative efficacy of religious and nonreligious cognitive-behavioral therapy for the treatment of clinical depression in religious individuals. *Journal of Consulting and Clinical Psychology* 1992;**60**(1):94-103.

Sackley 2018 *{published data only}*

Sackley CM, Smith CH, Rick CE, Brady MC, Ives N, Patel S, et al.Lee Silverman Voice treatment versus standard speech and language therapy versus control in Parkinson's disease: a pilot randomized controlled trial (PD COMM pilot). *Pilot and Feasibility Studies* 2018;**4**(30):1-10.

Salt 2002 *{published data only}*

Salt J, Shemilt J, Sellars V, Boyd S, Coulson T, McCool S.The Scottish Centre for autism preschool treatment programme: II. The results of a controlled treatment outcome study. *International Journal of Research and Practice* 2002;**6**(1):33-46.

Schultz 2002 *{published data only}*

Schultz DE.The effectiveness of stress management on anxiety and depression. Doctor's Dissertation, California School of Professional Psychology Fresno Campus 2002.

Sheikh 1986 *{published data only}*

Sheikh JI, Hill RD, Yesavage JA.Long-term efficacy of cognitive training for age-associated memory impairment: a six-month follow-up study. *Developmental Neuropsychology* 1986;**2**(4):413-21.

Steiner 2011 *{published data only}*

Steiner NJ, Sheldrick RC, Gotthelf D, Perrin EC.Computer-based attention training in the schools for children with Attention-Deficit/Hyperactivity Disorder: a preliminary trial. *Clinical Pediatrics* 2011;**50**(7):615-22.

Timko 1995 *{published data only}*

Timko C, Moos RH, Finney JW, Moos BS, Kaplowitz MS.Long-term treatment careers and outcomes of previously untreated alcoholics. *Journal of Studies on Alcohol* 1995;**60**(4):437-47.

Tomporowski 1984 *{published data only}*

Tomporowski PD, Ellis NR.Effects of exercise on the physical fitness, intelligence, and adaptive behavior of institutionalized mentally retarded adults. *Applied Research in Mental Retardation* 1984;**5**(3):329-37.

Torkan 2014 *{published data only}*

Torkan H, Blackwell SE, Holmes EA, Kalantari M, Neshat-Doost HT, Maroufi M, et al.Positive imagery cognitive bias modification in treatment-seeking patients with major depression in Iran: a pilot study. *Cognitive Therapy and Research* 2014;**38**:132-45.

Weber 1975 *{published data only}*

Weber BA, Sulzbacher SI.Use of CNS stimulant medication in averaged electroencephalic audiometry with children with MBD. *Journal of Learning Disabilities* 1975;**8**(5):44-7.

Wilson 2002 *{published data only}*

Wilson PH, Thomas PR, Maruff P.Motor imagery training ameliorates motor clumsiness in children. *Journal of Child Neurology* 2002;**17**(7):491-8.

References to studies awaiting assessment

Bommert 1978 *{published data only}*

Brandes 2010 *{published data only}*

McLachlan 1993 *{published data only}*

Newton-Cross 2017 *{published data only}*

Schwarzler 1999 *{published data only}*

Trianes Torres 1991 *{published data only}*

References to ongoing studies

Heitman 2017 *{published data only}*

Internet-based attentional bias modification training as add-on to regular treatment in alcohol and cannabis dependent outpatients: a study protocol of a randomised control trial. Ongoing study. 18/05/2015. Contact author for more information.

ISRCTN21392756 *{published data only}*

Standardised stress management mental health training: does it have a beneficial effect?. Ongoing study. 06/09/2010. Contact author for more information.

ISRCTN35717198 *{published data only}*

Cognitive training with non-invasive brain stimulation to treat binge eating disorder. Ongoing study. 03/05/2018. Contact author for more information.

NCT00044629 *{published data only}*

Combined behavioral/pharmacologicaltTherapy for insomnia. Ongoing study. 05/09/2002. Contact author for more information.

Additional references

Abikoff 2004 (pers comm)

Abikoff H.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 23rd of August 2019.

Allen 1998 (pers comm)

Allen JB.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 28th of August 2019.

Allen 2006 (pers comm)

Allen JB.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 28th of August 2019.

Allen 2018

Allen EN, Chandler CI, Mandimika N, Leisegang C, Barnes K.Eliciting adverse effects data from participants in clinical trials. *Cochrane Database of Systematic Reviews* 2018, Issue 1. Art. No: MR000039. [DOI: 10.1002/14651858.MR000039.pub2]

Andrews 2013a

Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y, et al.GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *Journal of Clinical Epidemiology* 2013;**66**(7):719-25.

Andrews 2013b

Andrews JC, Schünemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, et al.GRADE guidelines: 15. Going from evidence to recommendation - determinants of a recommendation's direction and strength. *Journal of Clinical Epidemiology* 2013;**66**(7):726-35.

APA 1952

American Psychiatric Association.Diagnostic and Statistical Manual of Mental Disorders. 1st edition. Washington (DC): American Psychiatric Association, 1952.

APA 1968

American Psychiatric Association.Diagnostic and Statistical Manual of Mental Disorders. 2nd edition. Washington (DC): American Psychiatric Association, 1968.

APA 1980

American Psychiatric Association.Diagnostic and Statistical Manual of Mental Disorders. 3rd edition. Washington (DC): American Psychiatric Association, 1980.

APA 1987

American Psychiatric Association.Diagnostic and Statistical Manual of Mental Disorders. 3rd revised edition. Washington (DC): American Psychiatric Association, 1987.

APA 1994

American Psychiatric Association.Diagnostic and Statistical Manual of Mental Disorders. 4th edition. Washington (DC): American Psychiatric Association, 1994.

APA 2000

American Psychiatric Association.Diagnostic and Statistical Manual of Mental Disorders. 4th text revision edition. Washington (DC): American Psychiatric Association, 2000.

APA 2013

American Psychiatric Association.Diagnostic and Statistical Manual of Mental Disorders. 5th edition. Washington (DC): American Psychiatric Association, 2013. [https://www.psychiatry.org/psychiatrists/practice/dsm]

Ayen 2004 (pers comm)

Ayen I.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: EG Faltinsen 9th of September 2019.

Balshem 2018

Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al.GRADE guidelines: 3. Rating the quality of evidence. *Journal of Clinical Epidemiology* 2018;**64**(4):401-6.

Bateman 2009

Bateman A, Fonagy P.Randomized controlled trial of outpatient mentalization-based treatment versus structured clinical management for borderline personality disorder. *American Journal of Psychiatry* 2009;**166**(12):1355-64.

Bateman 2017

Bateman A, Campbell C, Luyten P, Fonagy P.A mentalization-based approach to common factors in the treatment of borderline personality disorder. *Current Opinion in Psychology* 2017;**21**:44-9.

Berkey 1996

Berkey CS, Mosteller F, Lau J, Antman EM.Uncertainty of the time of first significance in random effects cumulative meta-analysis. *Controlled Clinical Trials* 1996;**17**(5):357-71.

Borduin 2009

Borduin CM, Schaeffer CM, Heiblum N.A randomized clinical trial of multisystemic therapy with juvenile sexual offenders: effects on youth social ecology and criminal activity. *Journal of Consulting and Clinical Psychology* 2009;**77**(1):26-37.

Borkovec 2005

Borkovec TD, Sibrava NJ.Problems with the use of placebo conditions in psychotherapy research, suggested alternatives, and some strategies for the pursuit of the placebo phenomenon. *Journal of Clinical Psychology* 2005;**61**(7):805-18.

Bornovalova 2008 (pers comm)

Bornovalova M.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 21st of August 2019.

Boutron 2021

Boutron I, Page MJ, Higgins JP, Altman DG, Lundh A, Hróbjartsson A.Chapter 7: Considering bias and conflicts of interest among the included studies. In: HigginsJP, ThomasJ, ChandlerJ, CumpstonM, LiT, PageMJ, et al, editors(s). Cochrane Handbook for Systematic Reviews of Interventions version 6.2. Cochrane, 2021.

Bramer 2017

Bramer WM, Rethlefsen ML, Kleijnen J, Franco OH.Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Systematic Reviews* 2017;**6**:245.

Brok 2008

Brok J, Thorlund K, Gluud C, Wetterslev J.Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. *Journal of Clinical Epidemiology* 2008;**61**(8):763-9.

Brok 2009

Brok J, Thorlund K, Wetterslev J, Gluud C.Apparently conclusive meta-analyses may be inconclusive -trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses. *International Journal of Epidemiology* 2009;**38**(1):287-98.

Brunetti 2013

Brunetti M, Shemilt I, Pregno S, Vale L, Oxman AD, Lord J, et al.GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. *Journal of Clinical Epidemiology* 2013;**66**(2):140-50.

Castellini 2018

Castellini G, Bruschettini M, Gianola S, Gluud C, Moja L.Assessing imprecision in Cochrane systematic reviews: a comparison of GRADE and Trial Sequential Analysis. *Systematic Reviews* 2018;**7**(1):110.

Chalmers 1983

Chalmers TC, Celano P, Sacks HS, Smith H Jr.Bias in treatment assignment in controlled clinical trials. *New England Journal of Medicine* 1983;**309**(22):1358-61.

Chanen 2008

Chanen AM, Jackson HJ, McCutcheon LK, Jovev M, Dudgeon P, Hok PY, et al.Early intervention for adolescents with borderline personality disorder using cognitive analytic therapy: randomised controlled trial. *British Journal of Psychiatry* 2008;**193**(6):477-84.

Charlesworth 2017

Charlesworth JE, Petkovic G, Kelley JM, Hunter M, Onakpoya I, Roberts N, et al.Effects of placebos without deception compared with no treatment: a systematic review and meta-analysis. *Journal of Evidence Based Medicine* 2017;**10**(2):97-107.

Cohen1988

Cohen J.Statistical Power Analysis for the Behavioral Sciences. 2nd edition.. Hillsdale (NJ): Lawrence Erlbaum Associates, Publishers, 1988.

Colloca 2020

Colloca L, Barsky AJ.Placebo and nocebo effects. *New England Journal of Medicine* 2020;**382**(6):554-61.

Comer 2013

Comer JS, Kendall PC.Methodology, design, and evaluation in psychotherapy research. In: LambertMJ, editors(s). Bergin and Garfields's Handbook of Psychotherapy and Behavior Change. 6th edition. New York: Wiley, 2013:21-48.

Copenhagen Trial Unit 2018

TSA - Trial sequential analysis. Copenhagen Trial Unit. [www.ctu.dk/tsa (accessed April 23 2018)]

Cristea 2017

Cristea IA, Gentili C, Cotet CD, Palomba D, Barbui C, Cuijpers P.Efficacy of psychotherapies for borderline personality disorder: A systematic review and meta-analysis. *JAMA Psychiatry* 2017;**74**(4):319-28.

Cuijpers 2016

Cuijpers P, Cristea IA.How to prove that your therapy is effective, even when it is not: a guideline. *Epidemiology and Psychiatric Sciences* 2016;**25**(5):428-35.

Cuijpers 2021

Cuijpers P, Quero S, Papola D, Cristea IA, Karyotaki E.Care-as-usual control groups across different settings in randomized trials on psychotherapy for adult depression: a meta-analysis. *Psychological Medicine* 2021;**51**(4):634–44.

Cunningham 2013

Cunningham JA, Kypri K, McCambridge J.Exploratory randomized controlled trial evaluating the impact of a waiting list control design. *BMC Medical Research Methodology* 2013;**13**(1):150.

da Costa 2013

da Costa BR, Nüesch E, Rutjes AW, Johnston BC, Reichenbach S, Trelle S, et al.Combining follow-up and change data is valid in meta-analyses of continuous outcomes: a meta-epidemiological study. *Journal of Clinical Epidemiology* 2013;**66**(8):847-55.

Donner 2002

Donner A, Klar N.Issues in the meta-analysis of cluster randomized trials. *Statistics in Medicine* 2002;**21**(19):2971-80.

Egger 1997

Egger M, Smith GD, Phillips AN.Meta-analysis: principles and procedures. *BMJ* 1997;**315**(7121):1533-7.

Eldridge 2016

Eldrige S, Campbell M, Campbell M, Dahota A, Giraudeau B, Higgins J, et al.Revised Cochrane risk of bias tool for

randomized trials (RoB 2.0) Additional considerations for cluster-randomized trials. University of Bristol 2016.

Erlen 2015

Erlen JA, Tamres LK, Reynolds N, Golin CE, Rosen MI, Remien RH, et al.Assessing usual care in clinical trials. Western Journal of Nursing Research 2015;**37**(3):288-98.

Espie 1989b

Espie CA, Brooks DN, Lindsay WR.An evaluation of tailored psychological treatment of insomnia. *Journal of Behavior Therapy and Experimental Psychiatry* 1989;**20**(2):143-53.

Fässler 2015

Fässler M, Meissner K, Kleijnen J, Hróbjartsson A, Linde K.A systematic review found no consistent difference in effect between more and less intensive placebo interventions. *Journal of Clinical Epidemiology* 2015;**68**(4):442-51.

Frank 1991

Frank JD, Frank JB.Persuasion and Healing: a Comparative Study of Psychotherapy. 3rd edition. Baltimore: Johns Hopkins University Press, 1991.

Freedland 2011

Freedland KE, Mohr DC, Davidson KW, Schwartz JE.Usual and unusual care: existing practice control groups in randomized controlled trials of behavioral interventions. *Psychosomatic Medicine* 2011;**73**(4):323-35.

Furukawa 2014

Furukawa TA, Noma H, Caldwell DM, Honyashiki M, Shinohara K, Imai H, et al.Waiting list may be a nocebo condition in psychotherapy trials: a contribution from network meta-analysis. *Acta Psychiatrica Scandinavica* 2014;**130**(3):181-92.

Gold 2018

Gold SM, Enck P, Hasselmann H, Friede T, Hegerl U, Mohr DC, et al.Control conditions for randomised trials of behavioural interventions in psychiatry: a decision framework. *Lancet Psychiatry* 2018;**4**(9):725-32.

GRADE Working Group 2004

GRADE Working Group.Grading quality of evidence and strength of recommendations. *BMJ* 2004;**328**(7454):1490.

Greville-Harris 2015

Greville-Harris M, Dieppe P.Bad is more powerful than good: the nocebo response in medical consultations. *American Journal of Medicine* 2015;**128**(2):126-9.

Guidi 2018

Guidi J, Brakemeier EL, Bockting CL, Cosci F, Cuijpers P, Jarrett RB, et al.Methodological recommendations for trials of psychological interventions. *Psychotherapy and Psychosomatics* 2019;**87**(5):276-84.

Guyatt 2011a

Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al.GRADE guidelines: 1. Introduction - GRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology* 2011;**64**(4):383-94.

Guyatt 2011b

Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al.GRADE guidelines: 2. Framing the question and deciding on important outcomes. *Journal of Clinical Epidemiology* 2011;**64**(4):395-400.

Guyatt 2011c

Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al.GRADE guidelines: 4. Rating the quality of evidence - study limitations (risk of bias). *Journal of Clinical Epidemiology* 2011;**64**(4):407-15.

Guyatt 2011d

Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, et al.GRADE guidelines: 5. Rating the quality of evidence - publication bias. *Journal of Clinical Epidemiology* 2011;**64**(12):1277-82.

Guyatt 2011e

Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al.GRADE guidelines 6. Rating the quality of evidence - imprecision. *Journal of Clinical Epidemiology* 2011;**64**(12):1283-93.

Guyatt 2011f

Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand, et al.GRADE guidelines: 7. Rating the quality of evidence - inconsistency. *Journal of Clinical Epidemiology* 2011;**64**(12):1294-302.

Guyatt 2011g

Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al.GRADE guidelines: 8. Rating the quality of evidence - indirectness. *Journal of Clinical Epidemiology* 2011;**64**(12):1303-10.

Guyatt 2011h

Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, et al.GRADE guidelines: 9. Rating up the quality of evidence. *Journal of Clinical Epidemiology* 2011;**64**(12):1311-6.

Guyatt 2013a

Guyatt G, Oxman AD, Sultan S, Brozek J, Glasziou P, Alonso-Coello P, et al.GRADE guidelines: 11. Making an overall rating of confidence in effect estimates for a single outcome and for all outcomes. *Journal of Clinical Epidemiology* 2013;**66**(2):151-7.

Guyatt 2013b

Guyatt GH, Oxman AD, Santesso N, Helfand M, Vist G, Kunz R, et al.GRADE guidelines: 12. Preparing Summary of Findings tables - binary outcomes. *Journal of Clinical Epidemiology* 2013;**66**(2):158-72.

Hafliðadóttir 2021

Hafliðadóttir SH, Juhl CB, Nielsen SM, Henriksen M, Harris IA, Bliddal H, et al.Placebo response and effect in randomized clinical trials: meta-research with focus on contextual effects. *Trials* 2021;**22**(1):493.

Heitmann 2017 (pers comm)

Heitmann J.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 5th of February 2019.

Higgins 2003

Higgins JP, Thompson SG, Deeks JJ, Altman DG.Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**(7414):557-60.

Higgins 2011

Higgin JP, Green S, editors.Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration, 2011. [Available from handbook.cochrane.org.]

Higgins 2017

Higgins JP, Altman DG, Sterne JA.Chapter 8: Assessing risk of bias in included studies. In: Higgins JP, Green S, Editors. Cochrane Handbook for Systematic Reviews of Interventions version 5.1.0 (updated June 2017). The Cochrane Collaboration, 2017. Available from www.training.cochrane.org/handbook.

Higgins 2019

Higgin JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al (editors).Cochrane Handbook for Systematic Reviews of Interventions Version 6.0 (updated July 2019). Cochrane, 2019.. Available from www.training.cochrane.org/handbook.

Hoffmann 2013

Hoffmann TC, Erueti C, Glasziou PP.Poor description of non-pharmacological interventions: analysis of consecutive sample of randomised trials. *BMJ* 2013;**347**(7924):e1-10.

Holmes 2016

Holmes RD, Tiwari AK, Kennedy JL.Mechanisms of the placebo effect in pain and psychiatric disorders. *Pharmacogenomics Journal* 2016;**16**(6):491.

Holmes 2018

Holmes EA, Ghaderi A, Harmer CJ, Ramchandani PG, Cuijpers P, Morrison AP, et al.The Lancet Psychiatry Commission on psychological treatments research in tomorrow's science. *Lancet Psychiatry* 2018;**5**(3):237-86.

Howick 2019

Howick J, Webster R, Kirby N, Hood K.Rapid overview of systematic reviews of nocebo effects reported by patients taking placebos in clinical trials. *Trials* 2018;**19**(1):674.

Hróbjartsson 2001

Hróbjartsson A, Gøtzsche PC.Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. *New England Journal of Medicine* 2001;**344**:1594-602.

Hróbjartsson 2002

Hróbjartsson A, Götzsche PC.Placebo treatment versus no treatment. *Cochrane Database of Systematic Reviews* 2002, Issue 4. Art. No: CD003974. [DOI: 10.1002/14651858.CD003974]

Hróbjartsson 2004

Hróbjartsson A, Götzsche PC.Placebo interventions for all clinical conditions. *Cochrane Database of Systematic Reviews* 2004, Issue 2. Art. No: CD003974. [DOI: 10.1002/14651858.CD003974.pub2]

Hróbjartsson 2010

Hróbjartsson A, Götzsche PC.Placebo interventions for all clinical conditions. *Cochrane Database of Systematic Reviews* 2010, Issue 1. Art. No: CD003974. [DOI: 10.1002/14651858.CD003974.pub3]

Hróbjartsson 2012

Hróbjartsson A, Miller FG.Placebo in nonpharmacological randomized trials. In: BoutronI, RavaudP, MoherD, editors(s). Randomized Clincial Trials of Nonpharmacological Treatments. Taylor & Francis Group, LLC, 2012:11-26.

ICH 2005

ICH Expert Working Group.International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use. ICH harmonised tripartite guideline. Guideline for good clinical practice E6(R1). www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R1_Guideline.pdf (accessed 24 April 2018).

Institute of Medicine 2009

Lo B, Field MJ.Institute of Medicine (US) Committee on Conflict of Interest in Medical Research, Education, and Practice. Conflict of Interest in Medical Research, Education, and Practice. Washington DC: National Academies Press (US), 2009.

Jakobsen 2014

Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C.Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Medical Research Methodology* 2014;**14**(1):120.

Jensen 2017

Jensen KB, Kirsch I, Pontén M, Rosén A, Yang K, Gollub RL, et al.Certainty of genuine treatment increases drug responses among intellectually disabled patients. *Neurology* 2017;**88**(20):1912-8.

Juul 2020

Juul S, Gluud C, Simonsen S, Frandsen FW, Irving K, Jakobsen JC.Blinding in randomised clinical trials of psychological interventions: a retrospective study of published trial reports. BMJ Evidence-Based Medicine Online First: 30 September 2020. [DOI: doi: 10.1136/bmjebm-2020-111407]

Karst 2007 (pers comm)

Karst M.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: EG Faltinsen 5th of September 2019.

Karterud 2020

Karterud S, Kongerslev MT.Psychotherapy of personality disorders needs an integrative theory of personality. *Journal of Psychotherapy Integration* 2020;**31**:34-53.

Kazdin 2015

Kazdin AE.Treatment as usual and routine care in research and clinical practice. *Clinical Psychology Review* 2015;**42**:168.

Kazdin 2016

Kazdin AE.Methodological Issues & Strategies in Clinical Research. 4th edition. Washington DC: American Psychological Association, 2016.

Kelley 2012 (pers comm)

Kelley J.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 12th of August 2019.

Kienle 1997

Kienle GS, Kiene H.The powerful placebo effect - fact or fiction? *Journal of Clinical Epidemiology* 1997;**50**(12):1311-8.

Kilmann 1985

Kilmann PR, Milan RJ, Boland JP, Nankin HR, Sablis RF, West MO, et al.The sexual interactions of men with secondary erectile dysfunction. Paper presented at the annual meeting of the American Psychological Association in Los Angeles 1985.

Kilmann 1988

Kilmann PR, Milan RJ, Boland JP, Davidson E, West MO, Sablis RF et al.The sexual interactions of men with secondary erectile dysfunction. *Journal of Sex Education and Therapy* 1988;**14**(1):46-7.

Kjaergard 2001

Kjaergard LL, Villumsen J, Gluud C.Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Annals of Internal Medicine* 2001;**135**(11):982.

Kongerslev 2015

Kongerslev MT, Chanen AM, Simonsen E.Personality disorder in childhood and adolescence comes of age: a review of the current evidence and prospects for future research. *Scandinavian Journal of Child and Adolescent Psychiatry and Psychology* 2015;**3**(1):31-48.

Kube 2017

Kube T, Rief W.Are placebo and drug-specific effects additive? Questioning basic assumptions of double-blinded randomized clinical trials and presenting novel study designs. *Drug Discovery Today* 2017;**22**:729-35.

Kwan 2014

Kwan RY, Leung MC, Lai CK.Acupressure for agitation in nursing home residents with dementia:study protocol for a randomized controlled trial. *Trials* 2014;**15**(410):1-7.

Kwan 2017 (pers comm)

Kwan R.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 26th of August 2019.

Lai 2004 (pers comm)

Lai K.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 6th of September 2019.

Legrand 2016 (pers comm)

Legrand F.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 14th of August 2019.

Leichsenring 2019

Leichsenring F, Steinert C, Ioannidis JPA.Toward a paradigm shift in treatment and research of mental disorders. *Psychological Medicine* 2019;**49**(13):2111-7.

Lilienfeld 2007

Lilienfeld SO.Psychological treatments that cause harm. *Perspectives on Psychological Science* 2007;**2**(1):53-70.

Linden 2014

Linden M, Schermuly-Haupt M.Definition, assessment and rate of psychotherapy side effects. *World Psychiatry* 2014;**13**(3):306-9.

Locher 2018

Locher C, Gaab J, Blease C.When a placebo is not a placebo: problems and solutions to the gold standard in psychotherapy research. *Frontiers in Psychology* 2018;**9**:2317.

Löfholm 2013

Löfholm AC, Brännström L, Olsson M, Hansson K.Treatment-as-usual in effectiveness studies: What is it and does it matter? *International Journal of Social Welfare* 2013;**22**(1):25-34.

Lund 2014

Lund K, Vase L, Petersen GL, Jensen TS, Finnerup NB.Randomised controlled trials may underestimate drug effects: balanced placebo trial design. *PLOS One* 2014;**9**(1):1-7.

Lundh 2017

Lundh A, Lexchin J, Mintzes B, Schroll JB, Bero L.Industry sponsorship and research outcome. *Cochrane Database of Systematic Reviews* 2017, Issue 2. Art. No: MR000033. [DOI: 10.1002/14651858.MR000033.pub3]

Meissner 2011

Meissner K, Bingel U, Colloca L, Wager TD, Watson A, Flaten MA.The placebo effect: advances from different methodological approaches. *Journal of Neuroscience* 2011;**31**(45):16117-24.

Michopoulos 2021

Michopoulos I, Furukawa TA, Noma H, Kishimoto S, Onishi A, Ostinelli EG, et al.Different control conditions can produce different effect estimates in psychotherapy trials for depression. *Journal of Clinical Epidemiology* 2021;**132**:59–70.

Milich 1991

Milich R, Carlson CL, Pelham WE Jr, Licht BG.Effects of methylphenidate on the persistence of ADHD boys following

failure experiences. *Journal of Abnormal Child Psychology* 1991;**19**(5):519-36.

Miranda 1997 (pers comm)

Miranda J.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: EG Faltinsen 9th of September 2019.

Miranda 2003

Miranda J, Chung JY, Green BL, Krupnick J, Siddique J, Revicki DA, et al.Treating depression in predominantly low-income young minority women: a randomized controlled trial. *JAMA* 2003;**290**(1):57-65.

Moher 1998

Moher M, Moher D, Pham B, Jones Al, Cook DJ, Jadad AR, et al.Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998;**352**(9128):609-13.

Moher 2009

Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group.Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ (Clinical Research Ed.)* 2009;**339**:b2535. [DOI: 10.1136/bmj.b2535]

Mohr 2009

Mohr DC, Spring B, Freedland KE, Beckner V, Arean P, Hollon SD, et al.The selection and design of control conditions for randomized controlled trials of psychological interventions. *Psychotherapy and Psychosomatics* 2009;**78**(5):275-84.

Mohr 2014

Mohr DC, Ho J, Hart TL, Baron KG, Berendsen M, Beckner V, et al.Control condition design and implementation features in controlled trials: a meta-analysis of trials evaluating psychotherapy for depression. *Translational Behavioral Medicine* 2014;**4**(4):407-23.

Munder 2013

Munder T, Brütsch O, Leonhart R, Gerger H, Barth J.Researcher allegiance in psychotherapy outcome research: an overview of reviews. *Clinical Psychology Review* 2013;**33**(4):501-11.

Munder 2018

Munder T, Barth J.Cochrane's risk of bias tool in the context of psychotherapy outcome research. *Psychotherapy Research* 2018;**28**(3):347-55.

Mustafa 2013

Mustafa RA, Santesso N, Brozek J, Akl EA, Walter SD, Norman G, et al.The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses. *Journal of Clinical Epidemiology* 2013;**66**(7):736-42.e5.

Norman 2003

Norman GR, Sloan JA, Wyrwich KW.Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Medical Care* 2003;**41**(5):582-92.

O'Leary 1978

O'Leary KD, Borkovec TD.Conceptual, methodological, and ethical problems of placebo groups in psychotherapy research. *American Psychologist* 1978;**33**(9):821-30.

Pagsberg 2017

Pagsberg AK, Jeppesen P, Klauber DG, Jensen KG, Rudå D, Stentebjerg-Oleson M, et al.Quetiapine extended release versus aripiprazole in children and adolescents with first-episode psychosis: the multicentre, double-blind, randomised tolerability and efficacy of antipsychotics (TEA) trial. *Lancet Psychiatry* 2017;**4**(8):605-18.

Peterson 2002

Peterson MA, Patterson B, Pillman BM, Battista MA.Cognitive recovery following alcohol detoxification: A computerised remediation study. *Neuropsychological Rehabilitation* 2002;**12**(1):63-74.

Proudfoot 2013 (pers comm)

Proudfoot J.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 26th of August 2019.

Rapee 2006 (pers comm)

Rapee R.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 29th of August 2019.

Rapee 2007 (pers comm)

Rapee R.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 29th of August 2019.

Review Manager 5 [Computer program]

Nordic Cochrane Centre, The Cochrane Collaboration Review Manager (RevMan 5).Version 5.3. Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Rosenberg 2014

Rosenberg L.Assisted outpatient treatment: we can do better. *Journal of Behavioral Health Services and Research* 2014;**41**(3):251-3.

Rosenzweig 1936

Rosenzweig S.Some implicit common factors in diverse methods of psychotherapy. *American Journal of Orthopsychiatry* 1936;**6**(3):412-5.

Rutherford 2014

Rutherford BR, Pott E, Tandler JM, Wall MM, Roose SP, Lieberman JA.Placebo response in antipsychotic clinical trials: A meta-analysis. *JAMA Psychiatry* 2014;**71**(12):1409-21.

Savović 2012

Savović J, Jones H, Altman D, Harris R, Jüni P, Pildal J, et al.Influence of reported study design characteristics on intervention effect estimates from randomised controlled trials: combined analysis of meta-epidemiological studies. *Health Technology Assessment* 2012;**16**(35):1-82.

Savovic 2018

Savovic J, Turner RM, Mawdsley D, Jones HE, Beynon R, Higgins JP, et al.Association between Risk-of-Bias assessments and results of randomized trials in Cochrane reviews: the ROBES meta-epidemiologic study. *American Journal of Epidemiology* 2018;**187**(5):1113-22.

Scherer 2018

Scherer RW, Meerpohl JJ, Pfeifer N, Schmucker C, Schwarzer G, von Elm E.Full publication of results initially presented in abstracts. *Cochrane Database of Systematic Reviews* 2018, Issue 11. Art. No: MR000005. [DOI: 10.1002/14651858.MR000005.pub4]

Schulz 1995

Schulz KF, Chalmers I, Hayes RJ, Altman DG.Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**:408-12.

Sibbald 1998

Sibbald B, Roland M.Understanding controlled trials: why are randomised controlled trials important? *BMJ* 1998;**316**(7126):201.

Sommer 2006

Sommer J, Jansen A, Dräger B, Steinsträter O, Breitenstein C, Deppe M, et al.Transcranial magnetic stimulation - a sandwich coil design for a better sham. *Clinical Neurophysiology* 2006;**117**(2):440-6.

Spitzer 1989

Spitzer RL, William JBW, Gibbon M, First MB.Structured Clinical Interview for DSM-III-R Personality Disorders (SCID-II). 1989: Biometrics Research Department, New York State Psychiatric Institute, 1989.

Storebø 2018

Storebø OJ, Pedersen N, Ramstad E, Kielsholm ML, Nielsen SS, Krogh HB, et al.Methylphenidate for attention deficit hyperactivity disorder (ADHD) in children and adolescents – assessment of adverse events in non-randomised studies. *Cochrane Database of Systematic Reviews* 2018, Issue 5. Art. No: CD012069. [DOI: 10.1002/14651858.CD012069.pub2]

Swanson 2014

Swanson JW, Swartz MS.Why the evidence for outpatient commitment is good enough. *Psychiatric Services* 2014;**65**(6):808-11.

Szymanski 1995 (pers comm)

Szymanski J.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 20th of August 2019.

Tajika 2015

Tajika A, Ogawa Y, Takeshima N, Hayasaka Y, Furukawa TA.Replication and contradiction of highly cited research papers in psychiatry: 10-year follow-up. *British Journal of Psychiatry* 2015;**207**(4):357-62.

Teri 1997 (pers comm)

Teri L.Inclusion of study in Cochrane Systematic Review [personal communication]. Email to: A Todorovac 26th of August 2019.

Thomas 2010 [Computer program]

EPPI-Centre Software. Social Science Research Unit, UCL Institute of Education EPPI-Reviewer 4: software for research synthesis..Thomas J, Brunton J, Graziosi S, Version 4.7.1.0. London, UK: EPPI-Centre Software. Social Science Research Unit, UCL Institute of Education, 2010. Available at www.eppi.ioe.ac.uk.

Thorlund 2009

Thorlund K, Devereaux PJ, Wetterslev J, Guyatt G, Ioannidis JP, Thabane L, et al.Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses? *International Journal of Epidemiology* 2009;**38**(1):276-86.

Thorlund 2011

Thorlund K, Imberger G, Walsh M, Chu R, Gluud C, Wetterslev J, et al.The number of patients and events required to limit the risk of overestimation of intervention effects in meta-analysis - a simulation study. *PLOS One* 2011;**6**(10):e25491.

Tough 2009

Tough EA, White AR, Richards SH, Lord B, Campbell JL.Developing and validating a sham acupuncture needle. *Acupuncture in Medicine* 2009;**27**(3):118-22.

Vase 2019

Vase L.Can insights from placebo and nocebo mechanism studies help improve randomized controlled trials? *Clinical Pharmacology and Therapeutics* 2019;**106**(6):1169-71.

Wager 2015

Wager TD, Atlas LY.The neuroscience of placebo effects: connecting context, learning and health. *Nature Reviews Neuroscience* 2015;**16**(7):403-18.

Walach 2011

Walach H.Placebo controls: historical, methodological and general aspects. *Philosophical Transactions: Biological Sciences* 2011;**366**(1572):1870-8.

Wampold 2010

Wampold BE, Imiel ZE.Great Psychotherapy Debate (Counseling and Psychotherapy) - The Evidence for What Makes Psychotherapy Work. 2 edition. Taylor & Francis Inc, 2010.

Wampold 2016

Wampold BE, Frost ND, Yulish NE.Placebo effects in psychotherapy: a flawed concept and a contorted history. Psychology of Consciousness: Theory, Research, and Practice 2016;**3**(2):108-20.

Watts 2015

Watts SE, Turnell A, Kladnitski N, Newby JM, Andrews G.Treatment-as-usual (TAU) is anything but usual: a meta-analysis of CBT versus TAU for anxiety and depression. *Journal of Affective Disorders* 2015;**175**:152–67.

We 2012

We SR, Koog YH, Park MS, Min B-Il.Placebo effect was influenced by publication year in three-armed acupuncture trials. *Complementary Therapies in Medicine* 2012;**20**(1-2):83-92.

Weimer 2015

Weimer K, Colloca L, Enck P.Placebo effects in psychiatry: mediators and moderators. *Lancet Psychiatry* 2015;**2**(3):246-57.

Weissman 1974

Weissman MM, Klerman GL, Paykel ES, Prusoff B, Hanson B.Treatment effects on the social adjustment of depressed patients. *Archives of General Psychiatry* 1974;**30**(6):771-8.

Weisz 2019

Weisz JR, Kuppens S, Ng MY, Vaughn-Coaxum RA, Ugueto AM, Eckshtain D, et al.Are psychotherapies for young people growing stronger? Tracking trends over time for youth anxiety, depression, attention-deficit/hyperactivity disorder, and conduct problems. *Perspectives on Psychological Science* 2019;**14**(2):216-37.

Wetterslev 2008

Wetterslev J, Thorlund K, Brok J, Gluud C.Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. *Journal of Clinical Epidemiology* 2008;**61**(1):64-75.

Wetterslev 2009

Wetterslev J, Thorlund K, Brok J, Gluud C.Estimating required information size by quantifying diversity in random-effects model meta-analyses. *BMC Medical Research Methodology* 2009;**9**(1):86.

Wetterslev 2017

Wetterslev J, Jakobsen JC, Gluud C.Trial Sequential Analysis in systematic reviews with meta-analysis. *BMC Medical Research Methodology* 2017;**17**(1):39.

WHO 1949

Manual of the International Classification of Diseases, Sixth Revision (ICD-6). World Health Organization 1949.

WHO 1955

Manual of the International Classification of Diseases, Seventh Revision (ICD-7). World Health Organization 1955.

WHO 1967

Manual of the International Classification of Diseases, Eighth Revision (ICD-8). World Health Organization 1967.

WHO 1975

Manual of the international classification of diseases, Ninth Revision (ICD-9). World Health Organization 1975.

WHO 1993

Manual of the International Classification of Diseases, Tenth Revision (ICD-10). World Health Organization 1993.

WHO 2018

Manual of the International Classification of Diseases, Eleventh Revision (ICD-11). World Health Organization 2018. [https://www.who.int/standards/classifications/classification-of-diseases]

Wilson 1978

Wilson A, Davidson WJ, Blanchard R, WhiteJ.Disulfiram implantation. A placebo-controlled trial with two-year follow-up. *Journal of Studies on Alcohol* 1978;**39**(5):809-19.

Wood 2008

Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG, et al.Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;**336**(7644):601. [DOI: 10.1136/bmj.39465.451748.AD]

Yeung 2017

Yeung V, Sharpe L, Glozier N, Hackett ML, Colagiuri B.A systematic review and meta-analysis of placebo versus no treatment for insomnia symptoms. *Sleep Medicine Reviews* 2017;**38**:17-27.

Young 2011

Young T, Hopewell S.Methods for obtaining unpublished data. *Cochrane Database of Systematic Reviews* 2011, Issue 1110.1002/14651858.MR000027.pub2. Art. No: MR000027. [DOI: 10.1002/14651858.MR000027.pub2]

References to other published versions of this review

Faltinsen 2019

Faltinsen E, Todorovac A, Hróbjartsson A, Gluud C, Kongerslev MT, Simonsen E, Storebø OJ.Placebo, usual care and wait-list interventions for all mental health disorders.. *Cochrane Database of Systematic Reviews* 2019, Issue 1. Art. No: MR000050. [DOI: 10.1002/14651858.MR000050]

Klein 2004

Klein RG, Abikoff H, Hechtman L, Weiss G.Design and rationale of controlled study of long-term methylphenidate and multimodal psychosocial treatment in children with ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry* 2004;**43**(7):792-801.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Abikoff 2004

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Psychological placebo: methylphenidate + attention control psychosocial treatment 2. No treatment: methylphenidate 3. Active treatment: methylphenidate + multimodal psychosocial treatment (social skills training) **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 1 year **Duration of participation (trial + follow-up):** 1 year + 1 year follow-up (switched to placebo) **Setting:** outpatient **Purpose of trial:** "To test that methylphenidate combined with intensive multimodal psychosocial intervention, which includes social skills training, significantly enhances social functioning in children with attention-deficit/hyperactivity disorder (ADHD) compared with methylphenidate alone and methylphenidate plus nonspecific psychosocial treatment (attention control)" (p.820) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** 332 **Number of participants included:** 103 **Number of participants followed-up at post treatment:** 86 (after 1 year) **Number of participants randomly assigned to:** • psychological placebo: n = 34 • no-treatment: n = 34 • active treatment: n = 34 **Number of withdrawals:** n = 17 • psychological placebo: n = 5 • no-treatment: n = 6 • active treatment: n = 6 **Diagnosis:** Attention Deficit Hyperactivity Disorder (ADHD) **Diagnostic manual:** Diagnostic and Statistical Manual of Mental Disorders, 3rd edition, revised (DSM-III-R) **Means of assessment:** the Diagnostic Interview Schedule for Children (DISC-IV) **Comorbidity:** majority (55 of 103, 53.4%) met criteria for oppositional defiant disorder and 31 (30%) of 103 had one or two symptoms of conduct disorder. Relatively few (17 of 103, 16.5%) had an anxiety disorder (simple phobia, overanxious disorder, separation anxiety disorder) or major depression (four of 103, 3.9%) **Age:** 8.2 mean years (SD = 0.8) (Range 7 to 9.9) **IQ:** exclusion IQ less than 85. Mean WISC IQs were full scale, 109.5 (SD =14.5); verbal, 108.5, SD =4.0); and performance, 108.7 (SD =15.0) **Sex:** 7% female **Ethnicity:** 84% white, 13%, African American, 2% Hispanic, and 1% other.

Abikoff 2004 *(Continued)*

Country: USA

Inclusion criteria

1. DSM-III-R criteria for ADHD.
2. Grade 1 to 4.
3. Medication free 2 weeks.
4. Meaningful benefit from methylphenidate without significant side effects.

Exclusion criteria

1. Diagnosable neurological disorders.
2. Psychosis.
3. Significant medical illness.
4. Current physical or sexual abuse.
5. Chronic tic disorder or Tourette's disorder.
6. A DSM-III-R developmental reading or arithmetic disorder, defined as a standard score in reading or mathematics on the Kaufmann Test of Educational Achievement of 85 or less (i.e., at least 1 SD below the population mean) and at least 15 points (1 SD) below full-scale IQ

Comparisons

Psychological placebo

Treatment name: attention control psychosocial treatment + methylphenidate (MPH)

Description of intervention: "The attention control was designed to account for nonspecific treatment effects of the MPT intervention, including professional time and attention, extended interactions with peers, and parental attention. It contained components parallel to those of MPT but excluded specific remedial or therapeutic content." (Klein 2004 , p., 797)

Individual or group treatment: group and Individual.

Exposure/intensity to treatment: the groups' mean daily methylphenidate dose did not differ at the end of year I, or year II The percentage of positive ritalinic acid assays was 87%, without differences between groups. (Klein 2004 , p. 799)

Duration of treatment: 1 year

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: all received MPH. Otherwise not stated

No-treatment

Comparison name: methylphenidate (MPH) only (no-treatment)

Description of intervention: not stated

Exposure/intensity to treatment: "The groups' mean daily methylphenidate dose did not differ at the end of year I or year II. The percentage of positive ritalinic acid assays was 87%, without differences between groups." (Klein 2004 , p. 799)

Duration treatment: 1 year

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: all got MPH. Otherwise not stated

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported, clinical relevance
- **Outcome chosen:** social skills rating scale - subscale parents rated

Adverse events

Abikoff 2004 *(Continued)*

- Reports adverse events, but does not differentiate between groups.

Notes

Key conclusion from study authors

1. In young children with ADHD, there is no support for clinic-based social skills training as part of a long-term psychosocial intervention to improve social behaviour.
2. Significant benefits from methylphenidate were stable over 2 years.

Key limitations from study authors

1. Measures of social behaviour consisted of parent, child, and teacher ratings and school observations.
2. Other measures, such as sociometric ratings and friendship indices, might have yielded treatment effects.
3. Similarly, it is conceivable that social skills interventions for children with ADHD are not targeting appropriate social skills. It is also possible that social skills are learned during training but do not generalise to real-world settings, possibly due to a lack of reinforcement in natural settings or an underlying disturbance in the ability of children with ADHD to generalise learned social behaviours

Other notes from review authors

1. Data on adverse events not usable

Conflicts of interest

Potential industry bias: Disclosure: Dr. Abikoff is a member of the ADHD Advisory Board and a principal investigator in clinical trials, Shire Pharmaceutical Co., and a member of the Metadate CD Advisory Board of Celltech Pharmaceuticals. He is a recipient of an investigator-initiated grant from McNeil Consumer and Specialty Pharmaceuticals. Dr. Hechtman received research funding from Eli Lilly, Janssen Ortho, Purdue,Shire Pharmaceutical Co., and GlaxoSmithKline Beecham and is on the speakers roster of Shire Pharmaceutical Co., Janssen Ortho, and Eli Lilly. Dr Klein is a member of the ADHD Advisory Board of Shire Pharmaceutical Co

Judgment: no

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: “Children were randomly assigned to one of three treatments for 2 years: (1) methylphenidate alone (M), (2) methylphenidate and MPT (M + MPT), or (3) methylphenidate and attention control psychosocial treatment (M + ACT). Groups were balanced for ethnicity, sex, IQ, and oppositional defiant disorder. Assignment was done in blocks of four to enable group treatment components.” (p. 821)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Parents not blinded
Blinding of participants and personnel	No	Not possible to placebo and "no-treatment"
Incomplete outcome data	No	Attrition >15% (21.4%). No ITT
Selective outcome reported	No	Differences in outcomes (adverse events) from the published description (Klein 2004) of the methods and full report
Other sources of bias	Yes	None found

Allen 1998

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Physical placebo: nonspecific acupuncture treatment group 2. Wait-list 3. Active treatment: specific acupuncture treatment **Sample calculation:** yes **Cluster randomised:** no **Duration of trial (baseline to post):** 8 weeks. **Duration of participation (trial + follow-up):** no follow-up **Setting:** outpatient **Purpose of trial:** "This design allowed us to test whether acupuncture designed to specifically treat symptoms of depression would demonstrate efficacy compared with a wait-list control and nonspecific acupuncture treatments." (p. 397) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** not stated **Number of participants included:** 38 **Number of participants followed-up at post treatment:** 34 **Number of participants randomly assigned to:** 38 • Physical placebo: n = 12 • Wait-list: n = 12 • Active treatment: n = 14 Number of withdrawals: n = 4 (5 dropped out of the study but one of them had completed non-specific treatment and is therefore included in analyses, p. 398). • Physical placebo: n = 1 • Wait list: n= 1 • Active treatment: n= 2 **Diagnosis:** major depressive disorder **Diagnostic manual:** Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) **Means of assessment:** Structured Clinical Interview for DSM-III-R. (SCID-R) **Comorbidity:** not stated **Age:** not stated (age between 18 and 45 years) **IQ:** not stated **Sex:** 100% female **Ethnicity:** not stated **Country:** USA **Inclusion criteria** 1. Meet diagnostic criteria for current major depression as outlined in DSM-IV.

Allen 1998 *(Continued)*

Exclusion criteria

1. Dysthymia or chronic (duration greater than 2 years) major depression.
2. Any current Axis I diagnosis besides major depressive disorder.
3. History of psychosis or mania
4. Substance abuse or dependence within the past 4 months.
5. Any current treatment.
6. Endocrine abnormalities.
7. History of central nervous system lesions or any medical disorder or treatment that could cause depression.
8. Active suicidal potential necessitating immediate treatment.
9. Pregnancy.

Comparisons	**Physical placebo** **Treatment name:** nonspecific treatment group **Description of intervention:** acupuncture - "(…) a placebo-like treatment designed to treat a pattern of disharmony that was not related to the individual's depression, but was characteristic of the individual." (p. 398) "Patients in the nonspecific-treatment group received 8 weeks of nonspecific treatment first, and then 8 weeks of specific treatment." (p. 398) **Individual or group treatment:** individual **Exposure/intensity to treatment:** "Each 8-week treatment regimen (both specific and nonspecific) comprised 12 treatment sessions: 2 sessions a week for the first 4 weeks followed by 1 per week thereafter." (p. 398) **Duration of treatment:** 8 weeks **Concomitant psychotherapy:** no – excluded if any current treatment **Concomitant pharmacotherapy:** no – excluded if any current treatment **Wait-list** **Comparison name:** wait-list **Description of intervention:** "Patients in the wait-list group waited 8 weeks before receiving 8 weeks of specific treatment." (p. 398) **Exposure/intensity to treatment:** no treatment during the 8 weeks **Duration treatment:** 8 weeks. **Concomitant psychotherapy** : no – excluded if any current treatment **Concomitant pharmacotherapy:** no – excluded if any current treatment
Outcomes	**Beneficial effect** • **Hierarchy:** usable data, patient-reported • **Outcome chosen:** The Inventory of Depressive Symptomatology (IDS; a self-report version of the Hamilton Rating Scale for Depression (HRSD) **Adverse events** • No usable data reported
Notes	**Key conclusion** 1. Quote: "A comparison of the acute effect of the three 8-week treatment conditions (n=34) showed that patients receiving specific acupuncture treatments improved significantly more than those receiving the placebo-like nonspecific acupuncture treatments, and marginally more than those in the wait-list condition." (p. 397) 2. "These finding from a small sample of women with major depression suggest that acupuncture may hold sufficient efficacy to warrant a lager clinical trial." (p. 400) **Key limitations**

Allen 1998 *(Continued)*

1. Quote: “Because specific treatment did not produce significantly greater improvement than the wait list, it remains possible that the improvement during specific treatment was due to spontaneous remission (…)With greater statistical power, specific treatment would likely prove significantly more effective than a wait-list control, as the power to detect a significant difference between these two groups with the present sample size is only.31.” (p. 400)
2. “These finding are, of course, preliminary. Larger scale studies are required to provide corroboration.” (p. 400)

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: “Patients were randomly assigned to one of three conditions (specific treatment: n=12, nonspecific treatment: n=11, or wait list: n=11.” (p. 398) Quote: “In 1996, unstratified randomization was implemented by creating a master randomized order in advance” (author correspondence)
Allocation concealment	Yes	Quote: "this was concealed until each new subject needed to be randomized.” (Allen 1998 (pers comm))
Blinding of outcome assessors	Yes	”Quote: All patients were interviewed by trained raters blind to treatment condition using the previously described 31-item version of the HRSD.” (p. 399)
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	Yes	Quote: “To determine whether the results would be the same for the original intent-to-treat sample, we conducted two analyses. The simple intent-to-treat analysis took the last available DepHRSD score of each participant who dropped out and carried that score forward to count as the subsequent observation (…) The second strategy used a random regression model (Gibbons et al., 1993, implemented with BMDP 5V), which imputes missing values based on maximum-likelihood estimates of missing parameters, thereby allowing for the analysis of all intent-to-treat subjects.” (p. 399) Attrition rate <15% (Specific treatment: 14%, Nonspecific treatment: 8%, Wait-list: 8%). ITT used, but seems that they excluded non-completers from reported data.
Selective outcome reported	Unclear	No trial registry was made (Allen 1998 (pers comm))
Other sources of bias	Yes	No other sources found

Allen 2006

Study characteristics

Methods	**Parallel randomised trial with three arms**

Allen 2006 *(Continued)*

1. Physical placebo: nonspecific acupuncture treatment
2. Wait-list
3. Active treatment: Traditional Chinese Medicine (TCM)- style acupuncture with manual stimulation for depression

Sample calculation: not stated
Cluster randomised : no
Duration of trial (baseline to post): 8 weeks.
Duration of participation (trial + follow-up): "No follow-up data. Patients in WL and non-specific received specific treatment after 8 weeks"
Setting: outpatient
Purpose of trial: "The current study sought to test the efficacy of acupuncture as a monotherapy for MDD in a large randomized controlled trial of both men and women with a range in the severity of MDD by comparing the efficacy of acupuncture intervention specifically designed to target each individual's depressive symptoms with an active valid acupuncture control that was not tailored to address an individual's symptoms of depression and with a wait-list control." (p. 1666).
Open/closed placebo: closed placebo

Data

Number of participants screened : 2965

Number of participants included: 157

Number of participants followed-up at post treatment : 131

Number of participants randomly assigned to:

- Physical placebo: n = 52
- Wait-list: n = 52
- Active treatment: n = 53

Number of withdrawals: n = 26

- Physical placebo: n = 7
- Wait-list: = 8
- Active treatment: n = 11

Diagnosis: major depressive disorder (MDD)

Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV)

Means of assessment: Structured Clinical Interview for DSM-IV (SCID-R)

Comorbidity: not stated

Age: 42.1 mean years (SD = 11.0). (Range = 18 to 65)

IQ: not stated

Sex : 66.2% female

Ethnicity: physical placebo: 88% white, wait-list: 81% white, active treatment: 79% white.

Country: USA

Inclusion criteria

1. Age 18 to 65 years
2. Meet DSM-IV diagnostic criteria for current MDD
3. Score at 14 or greater on the 17-item Hamilton Rating scale for Depression

Exclusion criteria

1. Dysthymia or chronic (duration greater than 2 years) MDD

Allen 2006 *(Continued)*

2. Seasonal pattern
3. Any Current Axis I diagnosis besides MDD or any Axis II Cluster B disorder
4. History of psychosis or mania
5. Substance abuse or dependence within the past 4 months
6. Any current relevant treatment
7. Endocrine abnormalities (e.g., hypothyroidism, unstable diabetes)
8. History of central nervous system involvement (e.g., seizures, brain injury, neurological illness
9. Any medical disorder or treatment believed by the investigators to cause depression
10. Active suicidal risk necessitating immediate intervention or suicide attempt within the past year
11. Pregnancy

Comparisons	**Physical placebo** **Treatment name (type):** nonspecific acupuncture **Description of intervention:** "A placebo-like control intervention utilized a comparable number of valid acupuncture points that were not designed to treat the individual's depression." (p. 1667) "Following the initial 8 weeks in 1 of these 3 intervention groups, all patients received SPEC intervention for the next 8 weeks." (p. 1668) **Individual or group treatment:** individual **Exposure/intensity to treatment:** "Twice per week for 4 weeks, then once per week for 4 weeks for a given 8-week regimen." (p. 1667) **Duration of treatment:** 8 weeks **Concomitant psychotherapy:** no - excluded if receiving any current relevant treatment **Concomitant pharmacotherapy:** no - excluded if receiving any current relevant treatment **Wait-list** **Comparison name:** wait-list. **Description of intervention:** "Following the initial 8 weeks in 1 of these 3 intervention groups, all patients received SPEC intervention for the next 8 weeks." (p. 1668) **Exposure/intensity to treatment** : no treatment during the 8 weeks **Duration treatment:** 8 weeks **Concomitant psychotherapy:** no - excluded if receiving any current relevant treatment **Concomitant pharmacotherapy** : no - excluded if receiving any current relevant treatment
Outcomes	**Beneficial outcomes for effect** • **Hierarchy:** usable data, observer-reported • **Outcome chosen** : Hamilton Rating Scale for Depression (17-item) **Relevant outcomes for adverse events** • No adverse events measured before after the wait-list received treatment.
Notes	**Key conclusion from study authors** 1. Quote: "The results of this randomized controlled trial of acupuncture as an intervention for depression indicate that although patients receiving acupuncture demonstrated significantly greater improvement than patients assigned to waitlist, there was no evidence to support differential efficacy of the 2 types of acupuncture intervention." (p. 1672) 2. "Interventions designed to specifically target depression resulted in no better outcome than those designed to serve as a control intervention." (p. 1672) 3. "The overall low response rate achieved with acupuncture suggest that TCM-style acupuncture with manual stimulation is not likely to be an adequate monotherapy for many with depression." (p. 1672) **Key limitations from study authors** 1. Quote: "(...) differences in provider expectations between SPEC and NONSPEC interventions, although small in magnitude, were statistically significant, suggesting that the blinding strategy was not entirely successful." (p. 1672)

Allen 2006 *(Continued)*

2. "Interventions designed to specifically target depression resulted in no better outcome than those designed to serve as a control intervention. Such results could reflect that the SPEC acupuncture intervention was not particularly effective, or that the intended control of NONSPEC acupuncture was somewhat more effective than predicted (…)" (p. 1672)

Other notes from review authors

1. None

Conflicts of interest: none found

J**udgement:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: "Patients were randomly assigned to 1 of 3 intervention groups following a stratified randomization schedule based on sex and severity of depression (…)" (p. 1668). "No significant differences emerged between intervention groups with respect to age, male/female ratio, ethnicity, age at onset, number of previous episodes, or symptom severity as assessed by the HAM-D or the BDI." (p. 1668) Quote: "In 2006, the randomization was stratified by severity, with two master records created in advance." (Allen 2006 (pers comm))
Allocation concealment	Yes	Quote: "Randomization schedules were devised by the first author at study onset, with each client's assignment becoming known to the assessing acupuncturist and the study coordinator only after the completion of the intake assessment" (p. 1668) Quote: "also concealed until each new subject needed to be randomized." (Allen 2006 (pers comm))
Blinding of outcome assessors	Yes	Quote: "The primary outcome measure was the HAM-D, administered at intake and at 4-week intervals thereafter by trained raters blind to intervention condition." (p. 1668)
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	Yes	Quote: "Change in depression severity over time was examined using random regression analyses using a mixed effects linear regression model with MIXREG software (…) The random regression approach utilizes all available data, estimating rate of change for each subject based on extant observations." (p. 1668) Attrition >15% (Specific: 20.8%, Non-specific: 8.2%, Wait-list: 15.4%. Excluded 6 after randomisation due to that the did not fulfil the inclusion criteria. Uses intent-to-treat on the remaining.
Selective outcome reported	Yes	NCT00010517 Protocol found but most information was not provided. Authors provided a descriptive protocol through author correspondence (Allen 2006 (pers comm))
Other sources of bias	Yes	No other found

Alvarez 1997

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Psychological placebo: Attention placebo 2. No-treatment 3. Active treatment: Anger management **Sample calculation:** yes **Cluster randomised:** no **Duration of trial (baseline to post):** 3 weeks **Duration of participation** (trial + follow-up): 3 weeks (endpoint data only) **Setting:** drug-free, residential therapeutic community located in a large northeastern metropolitan area (outpatient) **Purpose of trial:** "This study tested the following hypotheses: (1) Anger management treatment reduces both experienced and expressed anger in a sample of drug addicts in a therapeutic community; (2) the acquisition of effective anger manager skills by these subjects increases self-esteem; (3) the acquisition of positive anger management skills by these subjects decreases depression; (4) the acquisition of positive anger management skills by these subjects decreases addiction severity." (p. 6) **Open/closed placebo**: closed
Data	**Number of participants screened:** "350 to 400 clients." (p. 49) **Number of participants included:** 119 **Number of participants followed-up at post treatment**: 76 **Number of participants randomly assigned to**: • Psychological placebo: n = 39 • No-treatment: n = 40 • Active treatment n = 40 **Completer data at post-treatment** • Psychological placebo: n =23 • No-treatment: n = 25 • Active treatment: n = 28 **Number of withdrawals:** n = • Psychological placebo: n = 16 • No-treatment: n =15 • Active treatment: n = 12 **Diagnosis:** Substance-use disorder **Diagnostic manual:** Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) **Means of assessment:** not stated **Comorbidity:** substance-use of different drugs (51.6% cocaine, 26.4% polydrug abusers, 13.2% heroin). **Age** : 33.91 mean years, (range 19 to 54) **IQ:** moderate to severe intellectual disability excluded from trial **Sex** : 47.1% female **Ethnicity:** 90% African-Americans and Latinos and 10% Asian, European, or "other."

Alvarez 1997 *(Continued)*

Country : USA

Inclusion criteria

1. Diagnosis of substance-use disorder

Exclusion criteria

1. Serious homicidal or suicidal ideation or acting out
2. Free of active psychotic processes
3. Moderate to severe mental retardation
4. Moderate to severe organic brain syndrome

Comparisons

Psychological placebo

Treatment name : attention placebo

Description of intervention: "Subjects in the attention placebo group were presented an educational lecture from another standardized protocol, a copy of which appears in Appendix H. It consisted of a presentation of the clinical pharmacology of addictive drugs based on a training course for psychologists." (p.60)

Individual or group treatment: group

Exposure/intensity to treatment: 6 hours

Duration of treatment : 3 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: "Prescription medications for medical conditions are allowed if they do not interfere with the individual's ability to fully participate in the TC program. Medications such as fluoxetine for depression are frowned upon, though sometimes tolerated." (p.65)

No-treatment

Comparison name: no-treatment

Description of intervention: "The control group (N = 40) received no treatment from the facilitators, except for the pre- and posttest batteries." (p.58)

Exposure/intensity to treatment: none

Duration treatment : 3 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: "Prescription medications for medical conditions are allowed if they do not interfere with the individual's ability to fully participate in the TC program. Medications such as fluoxetine for depression are frowned upon, though sometimes tolerated." (p.65)

Outcomes

Beneficial effect

- **Hierarchy:** patient-reported, clinical relevance, coin-toss (random.org)
- **Outcome chosen:** The State-Trait Anger Expression Inventory, Research Edition - Subscale of trait

Adverse events

- No data on adverse event reported

Notes

Key conclusion from study authors

1. The experimental group (n = 34) attended the workshop. The control group (n = 29) received an attention placebo, and the no treatment group (n = 30) was not treated

Alvarez 1997 *(Continued)*

2. No significant difference between group treatment effects pre to post-treatment were found; the hypotheses were not supported.

Key limitations from study authors

1. Workshop delivery ineffective,
2. Not intensive enough
3. No monitoring
4. High attrition

Other notes from review authors

1. None

Conflicts of interest: none found.

Judgment: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "The list generated by the registration procedure was then randomly divided into three approximately even groups and each randomly assigned to an experimental condition." (p.50) "Careful examination of procedures revealed no systematic bias that would interfere with the randomness of the groups." (p. 58)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Quote:"The staff involved in the study were not blind to the paradigm" (p. 28) Patient-reported outcomes only
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	No	Attrition >15% (36.2%). No ITT. Only reports data on completers
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Ascher 1979

Study characteristics

Methods

Parallell randomised trial with 3 interventions

1. Psychological placebo: placebo treatment
2. Wait-list
3. Active treatment: paradoxical intention

Sample calculation: not stated

Cluster randomised: no

Ascher 1979 *(Continued)*

Duration of trial (baseline to post): 4 weeks

Duration of participation (trial + follow-up): "10 days pre trial + 4 weeks of experimental phase

Setting: outpatient

Purpose of trial: "The present study focused on a comparison of paradoxical intention with appropriate control procedures in reducing sleep onset insomnia." (p. 408)

Open/closed placebo: closed placebo

Data

Number of participants screened : not stated

Number of participants included: 25

Number of participants followed-up at post treatment: not stated

Number of participants randomly assigned to:

- Psychological placebo: n = 8
- Wait-list: n= 9
- Active treatment: n = 8

Number of withdrawals: not stated

Diagnosis: sleep-wake disorder (insomnia)

Diagnostic manual: not stated

Means of assessment: interview, not otherwise specified. "Following their initial phone call, clients who described themselves as experiencing a clinically significant level of sleep disturbance were invited for a pre-treatment interview." (p. 409)

Comorbidity: not stated

Age: 39 mean years, (range 24 to 67)

IQ: not stated

Sex : 60% female

Ethnicity : not stated

Country: USA

Inclusion criteria

1. "(...) criteria of sleep disturbance necessary for selection. These included: a sleep onset latency of one hour or more at least three times per week; awakening after falling asleep at night, with or without difficulty returning to sleep, three or more evenings each week; arising uncomfortably early in the morning on three or more occasions each week." (p. 408)

Exclusion criteria

1. Participants who exhibited a sub-clinical level of sleep difficulty.
2. Secondary insomnia

Comparisons

Psychological placebo:

Treatment name: placebo

Description of intervention: " Steinmark 1974 employed a placebo condition which served as a model for the present study. During the first session each subject was required to compose a hierarchy consisting of eighteen chronologically ordered bedtime activities. Then six neutral scenes were constructed. Finally, the client imagined each of the bedtime activities in the appropriate order, each being

Ascher 1979 *(Continued)*

paired with one of the six neutral scenes. Homework involved practicing the procedure twice a day, but not within two hours of bedtime." (p. 409)

Individual or group treatment: individual

Exposure/intensity to treatment: four weekly sessions of 30-45 minutes

Duration of treatment : 4 weeks

Concomitant psychotherapy not stated

Concomitant pharmacotherapy: not stated

Wait-list

Comparison name : no treatment (in reality wait-list)

Description of intervention: "Clients in this condition were provided with no treatment throughout the four weeks of the experimental phase. Contact was maintained through brief telephone conversations once every 1 ½ weeks." (p. 409)

"Subjects in either of the control conditions who elected to continue with treatment were provided with a combination of behavioral techniques." (p. 409.)

Exposure/intensity to treatment: no treatment during the experimental phase

Duration treatment: 4 weeks.

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect

- **Hierachy:** patient-reported, clinical relevance
- **Outcome chosen:** Daily Sleep Questionnaire (DSQ) Subscale difficulty experienced in falling asleep (0-7 - 7 being no difficulties falling asleep)

Adverse events

- No data on adverse events reported

Notes

Key conclusion from study authors

1. "Specifically, in the present study, subjects in the paradoxical intention group reported a significant reduction in sleep onset latency, fewer awakenings at night with difficulty returning to sleep, a significant increase in the experience of restedness obtained from sleep, in comparison to the reports of subjects in either the placebo or no-treatment control groups." (p. 410).

Key limitations from study authors

1. Not stated

Other notes from review authors

1. None

Conflicts of interest: none found

Judgment: yes

Risk of bias

Item	Authors' judgement	Support for judgement

Ascher 1979 *(Continued)*

Random sequence generation	Unclear	Quote: "The experiment consisted of three groups of randomly assigned subjects. The paradoxical intention treatment was contrasted with no-treatment and placebo treatment control conditions." (p. 408)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	Unclear	No information Attrition unclear
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Ayen 2004

Study characteristics	
Methods	**Parallel randomised trial with four arms** 1. Psychological placebo: Unterstützende Gruppe (UGT) 2. Wait-list B 3. Active treatment: Kognitive Verhaltenstherapie (CBT) 4. Wait-list A **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial** (baseline to post): 3 months **Duration of participation (trial + follow-up):** 3 months + 12 months **Setting:** outpatient **Purpose of trial:** assess short- and long-term efficacy of CBT compared to supportive group program (UGT) and waiting list control for depressive problems among menopausal women. **Open/closed placebo: c**losed placebo
Data	**Number of participants screened** : 85 **Number of participants included** : 51 **Number of participants followed-up at post treatment** : 50 **Number of participants randomly assigned to:** • Psychological placebo: n = 20 • Wait-list B: n = 10 • Active treatment: n =11

Ayen 2004 *(Continued)*

- Wait-list A: n =10

Number of withdrawals : n = 1

- Psychological placebo: not stated
- Wait-list B: not stated
- Active treatment: not stated
- Wait-list A: not stated

Diagnosis:depression

Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV). 59 % fulfilled DSM diagnosis, 41 % had an unspecified depression

Means of assessment: The Structured Clinical Interview for DSM-IV Disorders (SCID) (German version)

Comorbidity: 6 participants (12 %) had substance dependence. 27 (53%) had anxiety disorder

Age: 51.3 means years (range = 46 – 56)

IQ: not stated

Sex: 100% females

Ethnicity: not stated

Country: Germany

Inclusion criteria

1. Irregular bleedings last 12 months
2. Menopausal difficulties
3. Between 40 and 60 years
4. Depression diagnosis

Exclusion criteria

1. Not stated

Comparisons

Psychological placebo

Treatment name : Unterstützende Gruppe (UGT)

Description of intervention: The control group incorporated none of the specific techniques from the active intervention.

Individual or group treatment: group exposure/intensity to treatment: 2 hours each session.

Duration of treatment: 3 months

Concomitant psychotherapy: not stated - but many patients had already been in psychic or psy-cho-therapeutic treatment

Concomitant pharmacotherapy: 29 did not take drugs, 6 took antidepressants and 1 took anxiety medication.

Wait-list

Comparison name: waiting list B

Description of intervention: unclear

Exposure/intensity to treatment: not stated

Duration treatment: 3 months

Ayen 2004 *(Continued)*

	Concomitant psychotherapy: not stated - but many patients had already been in psychic or psycho-therapeutic treatment **Concomitant pharmacotherapy:** 29 did not take drugs, 6 took antidepressants and 1 took anxiety medication
Outcomes	**Beneficial effect** • **Hierarchy:** Patient-reported, clinical relevance, psychometric properties • **Outcome chosen:** Beck Depression Inventory (BDI) **Adverse events** • Beschwerdenliste (BL) - complaint list
Notes	**Key conclusion from study authors** 1. Both active groups beneficial outcomes at both assessment periods. Superior to the psychological placebo. **Key limitations from study authors** 1. Initial and final examination as well as group interventions were conducted in one hand 2. Therapist diagnostic decisions at baseline 3. Many participants had already prior been involved in a psychiatric or psycho-therapeutic treatment 4. Participants reported themselves notices in medical practices or offices of health insurance companies **Other notes from review authors** 1. None **Conflicts of interest:** none found. J**udgment:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	quote: “However, initially (21 participants inside) between the KVT (N = 11) and the KG (N - 10) was randomized, later (30 participants) then between the UGT (N = 20) and the KG (N 10)” (translated from, p. 293)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and "no-treatment"
Incomplete outcome data	Yes	Attrition <15% (1.96%)
Selective outcome reported	Unclear	No information
Other sources of bias	Yes	None other sources of bias found

Berg 1983

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Pharmacological placebo: Placebo pill 2. Waitlist 3. Active treatment: Senokot **Sample calculation:** not stated **Cluster randomised** : no **Duration of trial (baseline to post):** 3 months **Duration of participation (trial + follow-up):** 3 months + 18 months **Setting:** outpatient **Purpose of trial** : "The study described in this paper was designed to see whether behaviour therapy would suffice on its own in the treatment of severe and persistent faecal soiling or would be improved by employing a laxative as well." (p.544) **Open/closed placebo:** closed
Data	**Number of participants screened**: not stated **Number of participants included**: 44 **Number of participants followed-up at post treatment** 40 **Number of participants randomly assigned to:** • Pharmacological placebo: n = 11 • Wait-list: n = 15 • Active treatment: n = 14 **Number of withdrawals:** 4 (only completed one or two visits) **Diagnosis:** faecal soiling (encopresis) **Diagnostic manual:** not stated **Means of assessment:** clinical interview (not otherwise specified) **Comorbidity:** not stated **Age** : 7.9 mean years (SD = 2.3) **IQ:** not stated **Sex** : not stated **Ethnicity:** not stated **Country:** UK **Inclusion criteria** 1. Children with soiling as main complaint 2. Initial assessment and physical examination indicating that uncomplicated functional faecal incontinence was the problem **Exclusion criteria** 1. None mentioned
Comparisons	**Pharmacological placebo** **Treatment name (type):** Group B **Description of intervention:** "The basic method of treatment given to all was behavioural, focusing on use of the toilet and freedom from soiling." (p. 544) "Group B was given placebo tablets in similar dosage…" (as group A, not specified). The children started at one tablet at night. On the next visit to the clinic, if there was no improvement in 'use of the toilet' and 'being clean' on the charts the dosage was increased to two tablets. The number of tablets was increased to three on the following visit if improvement had still not occurred by that time. When the soil-

Berg 1983 *(Continued)*

ing was getting better and the child was using the toilet the dosage was kept the same. Once the child was going regularly to the toilet and not soiling the tablets were stopped altogether." (p. 544)
Individual or group treatment: individual.
Exposure/intensity to treatment: 1-3 pills at night. "Mother and child came to see the psychologist dealing with them every fortnight for three months," (p. 544)
Duration of treatment: 3 months
Concomitant psychotherapy: not stated
Concomitant pharmacotherapy: not stated

Wait-list
Comparison name: Group C (wait-list – any child still soiling after the 3 months assessment was offered a further 3 months treatment by the same psychologist using a behavioural approach as before and explicitly labelled Senokot tablets, 1-3 at night, in addition, (p. 546).
Description of intervention: "The basic method of treatment given to all was behavioural, focusing on use of the toilet and freedom from soiling." (p. 544)
Exposure/intensity to treatment : No medical treatment. "Mother and child came to see the psychologist dealing with them every fortnight for three months," (p. 544)
Duration treatment: 3 months
Concomitant psychotherapy : not stated
Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect

- **Hierarchy:** usable data, clinical relevancy
- **Outccome chosen:** number of children soiling more than once weekly: (self-reported) - binary outcome

Adverse events

- None found

Notes

Key conclusion from study authors

1. Quote: "Significant improvement occurred following three months of outpatient treatment using behavioural approach and either Senokot, placebo or no medication. However, there was no evidence either during the trial or subsequently when Senokot was employed to supplement behavioural treatment in every child who continued with therapy that this laxative contributed in any way to relieving the problem in this group of cases." (p. 549)

Key limitations from study authors

1. Quote: "Senokot was given in doses which were not excessive, and at night. This may have helped to conceal its true nature from the psychologists". (p. 547)

Other notes from review authors

1. None

Conflicts of interest: none found

J judgment: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Cases were then randomly allocated to one of three treatment groups" (p. 544) "The process of random allocation was successful since the treatment groups did not differ on a whole variety of features" (p. 547)

Berg 1983 *(Continued)*

Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Psychiatrist were outcome assessors Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	Yes	Quote: ”Forty-four children were included in the investigation but 4 dropped out after only one or two visits.” (p. 544)” Attrition > 15%. (9.1% dropped out after one or two visits)
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Borden 1986

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Pharmacological placebo: placebo pill 2. No-treatment: No pill 3. Active treatment: methylphenidate **Sample calculation:** not stated **Cluster randomised** : no **Duration of trial (baseline to post):** 3 months **Duration of participation (trial + follow-up):** “Treatment occurred over a three-month period, with children beginning the program at various times throughout a one-year period. All testing was completed during the three weeks preceding treatment, and then during the three weeks following treatment.” (p. 42) **Setting:** outpatient **Purpose of trial:** “The present study examined the attributional effects of combining medication with cognitive behavior therapy in the treatment of children diagnosed as having Attention Deficit Disorder (ADD).” (p. iii) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** not stated **Number of participants included** : 37 (only 30 is stated in the text, but in notes at page 72, it is stated that seven additional children began the study). **Number of participants followed-up at post treatment:** 30 **Number of participants randomly assigned to**: not stated **Number of randomised completers** • Pharmacological placebo: n =10 • No-treatment: n =10 • Active treatment: n =10 **Number of withdrawals:** "7 additional children began the study" (p. 72)

Borden 1986 *(Continued)*

- Pharmacological placebo: n = 2
- No-treatment: n = 3
- Active treatment: n= 2

Diagnosis: Attention Deficit Hyperactivity Disorder (ADHD) (n = 25) or Attention Deficit Disorder without Hyperactivity (ADDNH) (n = 5).
Diagnostic manual: not stated
Means of assessment: clinical interview, laboratory tests, parent, teacher, and child questionnaires, achievement tests, intelligence tests and paediatric examination.
Comorbidity: not stated
Age: 107.1 months (SD = 20.6). (Range = 68 - 143 months)
IQ : Above 80
Sex: 16.7% female
Ethnicity: not stated
Country: USA

Inclusion criteria

1. Meets diagnostic criteria for ADD or ADDH
2. At least one Conners Rating Scale score (parent or teacher) had to be 15 or greater at the onset of the study
3. Wechsler Full Scale IQ at least 80
4. Duration of ADD or ADDH symptoms at least 6 months

Exclusion criteria

1. IQ below 80
2. An acute problematic situation in the home that might have caused ADD symptoms
3. Onset after age 6
4. Duration of symptoms for less than six months
5. Known physical, neurological, or uncorrected sensory impairment
6. Psychosis

Comparisons

Pharmacological placebo
Treatment name : Placebo (pharmacological placebo) + cognitive behavior modification
Description of intervention: “The placebos were administered to placebo children in the same way and at the same time as was the active medication. Methylphenidate and placebo doses were both packaged by a University of Illinois Hospital pharmacist in identical opaque capsules to conceal their contents.” (p. 43) “Each child was randomly assigned to a therapist who was responsible for cognitive training.” (p. 43-45)
Individual or group treatment: individual.
Exposure/intensity to treatment: two doses were administered per day: one at breakfast and one at lunch” (p. 43). Cognitive training sessions two times per week.
Duration of treatment: 3 months
Concomitant psychotherapy: cognitive behavior modification
Concomitant pharmacotherapy: not stated

No-treatment

Comparison name : No pill (no treatment) + cognitive behaviour modification
Description of intervention :

No pharmacological treatment. “Cognitive training occurred over three months, with sessions held two times per week… Each child was randomly assigned to a therapist who was responsible for cognitive training.” (p. 43-45)
Exposure/intensity to treatment: no medication. Cognitive training sessions two times per week
Duration treatment : 3 months
Concomitant psychotherapy: cognitive behavior modification
Concomitant pharmacotherapy: not stated

Borden 1986 *(Continued)*

Outcomes	**Beneficial effects** • **Hierarchy**: blinded, observer-reported, clinical relevancy, random assignment • **Outcome chosen:** The Children's Checking Task (CCT), omissions – observer-reported **Adverse events** • None reported
Notes	**Key conclusions from study authors** 1. ”Medication group was found to influence parent attributions for the causes and solutions to their children’s presenting problems.” (p. iii) 2. “While child measures did not reveal significant effects, group means were directionally similar to those of the parents.” (p. iii) 3. “No group differences were found at posttest on achievement attributions measures or on measures of behavioral or cognitive improvement.” (p. iii) **Key limitations from study authors** 1. Not stated **Other notes from review authors** 1. None **Conflicts of interest:** none found **Judgment:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: “Assignment to groups was random with the exception of stratification based upon sex and age.” (p. 33)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote:”Examiners were blind to the treatment conditions of the children they tested.” (p. 42) Tests were administered by trained masters and doctoral level students who were blind to the subjects' medication conditions.” (p. 35) “Disagreements between raters were settled by a third graduate student in psychology who was blind to the responses of the first two.” (p. 39)
Blinding of participants and personnel	No	Not possible to blind placebo and "no-treatment"
Incomplete outcome data	Unclear	Attrition > 15% (19%). No mention of ITT. Only completers included
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Borkovec 1975

Study characteristics	
Methods	**Parallel randomised trial with four arms** 1. Psychological placebo: quasi-desensitisation placebo 2. Wait-list: waiting list no treatment 3. Active treatment 1: progressive relaxation 4. Active treatment 2: relaxation without tension-release **Sample calculation** : not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 4 weeks **Duration of participation (trial + follow-up):** 4 weeks of treatment + 5 months follow-up **Setting:** outpatient **Purpose of trial** : "At the end of a counter demand period, progressive relaxation was compared to the same quasi-desensitization placebo and no treatment conditions employed in the earlier investigation. As suggested elsewhere (Borkovec,1973), a treatment procedure demonstrated to reliably produce improvement greater than placebo under such (neutral) expectancy conditions is indeed a powerful modification technique and includes active ingredients independent of demand effects." (p. 302) **Open/closed placebo** : closed placebo
Data	**Number of participants screened:** not stated **Number of participants included** : 56 **Number of participants followed-up at post treatment:** 41 N **umber of participants randomly assigned to:** 56 were randomised, but reports only data on completers • Psychological placebo: n= 11 • Wait-list: n = 10 • Active treatment 1: n = 11 • Active treatment 2: n = 9 **Number of withdrawals:** n = 15 • Psychological placebo: not stated • Wait-list: not stated • Active treatment 1: not stated • Active treatment 2: not stated **Diagnosis:** sleep-wake disorder (Insomnia) **Diagnostic manual**: not stated **Means of assessment** : clinical interview, not otherwise specified. "(…) were interviewed and screened, following the criteria of Steinmark 1974 ." (p. 302) **Comorbidity** : not stated **Age:** not stated **IQ** : not stated, but all were college students.

Borkovec 1975 *(Continued)*

Sex : not stated

Ethnicity: not stated

Country: USA

Inclusion criteria

1. Average latency to sleep onset was 31 minutes or greater

Exclusion criteria

1. Not stated

Comparisons	**Psychological placebo** **Treatment name** : quasi-desensitisation placebo **Description of intervention**: "During session 1, subjects in the P condition constructed an 18-item hierarchy of chronological bedtime activities and chose six neutral images to be paired with the hierarchy items as a substitute for the relaxation ordinarily employed in conventional desensitization. Each hierarchy item was presented three times with intervening presentations of neutral images during session 2, 3 and 4. The P subjects were instructed to practice hierarchy and neutral image visualizations twice a day." (p. 303) **Individual or group treatment** : group sessions + individual practice **Exposure/intensity to treatment** : 4 sessions of group treatment + individual practice twice a day **Duration of treatment:** 4 weeks **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:**: not stated **Wait-list** **Comparison name**: waiting list no treatment (in reality wait-list) **Description of intervention**: "Subjects in the NT condition were told by phone that current treatment groups were filled, new groups would begin in 4 wk, and they would receive priority if they filled out the daily questionnaires during the next 4 wk." (p. 303) **Exposure/intensity to treatment**: no treatment during the 4 weeks experimental period **Duration treatment**: 4 weeks **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy**: not stated
Outcomes	**Beneficial effects** • **Hierachy:** patient-reported, clinical relevance • **Outcome chosen:** Daily Sleep Questionnaire (DSQ) Subscale difficulty experienced in falling asleep (0-5 - 5 being higher being more difficult to fall asleep) **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors**

Borkovec 1975 *(Continued)*

1. "In replication of an earlier study, progressive relaxation was found to produce greater reduction in reported latency to sleep onset than both placebo and no treatment conditions during a counterdemand period." (p. 307)
2. "Reports of additional latency improvement by progressive relaxation subjects 5 mo after the conclusion of the study replicate the follow-up results of the earlier study and indicate the long-term effectiveness of that procedure." (p. 307)
3. "Second, the superiority of progressive relaxation over a control condition similar in all respects except the presence of tension-release of muscle groups suggests that attention focusing alone is not sufficient to promote sleep." (p. 307)
4. "It is noteworthy that interactions of therapist and treatment factors were isolated in the positive demand data and only among the control condition. Such results suggest that therapist characteristics (such as sex of therapist), frequently found to be unrelated to outcome in behavior therapy studies, may be potent factors in the generation of demand and placebo effects." (p. 308)

Key limitations from study authors

1. Not stated

Other notes

1. SD's were imported from Steinmark 1974 due to that the same population, outcome and intervention was used in both

Conflicts of interest: none found

Judgment: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "The subjects were ranked on latency to sleep onset obtained in the pretherapy interview and randomly assigned within levels of severity to one of the four conditions: (a) progressive relaxation (PR), (b) relaxation without tension-release (NTR), (c) quasi-desensitization placebo (P), and (d) waiting list no treatment (NT)." (p. 303)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and "no-treatment"
Incomplete outcome data	No	Attrition >15% (27%). No ITT. Only reports data on completers
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Borkovec 1976

Study characteristics

Methods	**Parallel randomised trial with three arms**

Borkovec 1976 *(Continued)*

1. Psychological placebo
2. Wait-list
3. Active treatment: progressive relaxation training

Sample calculation: not stated

Cluster randomised no

Duration of trial (baseline to post): 4 weeks (endpoint data)

Duration of participation (trial + follow-up): pre-therapy week + 4 weeks of therapy + 1 year (follow-up)

Setting: outpatient

Purpose of trial : "The purpose of the present study was to replicate the basic aspects of the earlier investigations using objective sleep measures." (p. 174)

Open/closed placebo: closed placebo

Data

Number of participants screened: not stated

Number of participants included: 36

Number of participants followed-up at post treatment: 33

Number of participants randomly assigned to:

- Psychological placebo: n = 12
- Wait-list: n = 12
- Active treatment: n = 12

Number of withdrawals: n = 3

- Psychological placebo: n = 1
- Wait-list: n = 1
- Active treatment: n = 1

Diagnosis: sleep-wake disorder (sleep disturbance)

Diagnostic manual: not stated

Means of assessment: Brief questionnaire - "A brief questionnaire on sleep behavior was given (…)" (p. 174) + pretherapy interview

Comorbidity: not stated

Age: not stated

IQ : not stated, but psychology students at University of Iowa

Sex : not stated

Ethnicity: not stated

Country: USA

Inclusion criteria

1. "Subjects indicating 31 min or greater in average latency to sleep onset and that they considered this duration to represent a problem (…)" (p. 174)

Exclusion criteria

1. Participants reporting 30 minutes or less average sleep-onset latency

Borkovec 1976 *(Continued)*

2. Current use of drugs
3. Current contact with other professional services during the interview

Comparisons	**Psychological placebo:** **Treatment name**: placebo **Description of intervention**: "The placebo condition involved a quasi-desensitization procedure. During Session 1, each subject constructed an 18-item hierarchy of chronological bedtime activities and chose six neutral images to be paired with the hierarchy items and to be used as substitutes for relaxation" (p. 175) **Individual or group treatment:** group treatment + individual practice **Exposure/intensity to treatment:** 4 weekly group sessions + individual practice twice a day **Duration of treatment:** 4 weeks **Concomitant psychotherapy**: no – excluded if current contact with other professional services during the interview **Concomitant pharmacotherapy:** not stated **Wait-list** **Comparison name**: no-treatment (in reality a wait-list) **Description of intervention**:"No-treatment subjects were told that current treatment groups were filled but that new groups would be formed in 6 weeks and that they would receive priority if they continued to fill out the daily sleep questionnaires and attended the sleep evaluation nights." (p. 174) **Exposure/intensity to treatment**: no treatment **Duration treatment**: 4 weeks **Concomitant psychotherapy**: no – excluded if current contact with other professional services during the interview **Concomitant pharmacotherapy**: not stated
Outcomes	**Beneficial effect** • **Hierachy:** observer-reported, clinical relevance • **Outcome chosen:** latency in minutes - First occurrence of Stage I EEG. **Adverse events** • No adverse events reported
Notes	**Key conclusion from study authors** 1. "In general, the results of the present study provide modest support for the efficacy of progressive relaxation in the treatment of sleep-onset disturbance. Relaxation was the only condition to produce significant improvement in Stage I onset and in reports of sleep onset during lab evaluation nights." (p. 178) 2. "Between-condition differences, however, were limited. Relaxation was significantly superior only to no-treatment on Stage I onset improvement at the positive demand period, to placebo only on self-report post questionnaire items, and to placebo and no-treatment on Stage I variance reduction at the positive demand period." (p. 178) **Key limitations from study authors** 1. "Lab setting and measurement procedures themselves might have been sufficient placebos to induce reported daily improvements" (p. 179)

Borkovec 1976 *(Continued)*

Other notes from review authors

1. None

Conflicts of interest: none found

Judgment: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "The subjects were ranked on latency to sleep onset obtained in the pretherapy interview and randomly assigned within levels of severity to one of the four conditions: (a) progressive relaxation (PR), (b) relaxation without tension-release (NTR), (c) quasi-desensitization placebo (P), and (d) waiting list no treatment (NT)." (p. 303)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote:"Three research assistants, independent of each other and "blind" to the experimental condition and evaluation night of the subjects, scores the EEG records of the three evaluation nights (pretherapy, counterdemand, and positive demand). (p. 175-6)
Blinding of participants and personnel	No	Not possible to blind placebo and "no-treatment"
Incomplete outcome data	Yes	Attrition <15% (8.5%). Equal amount from each group. No ITT
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Bornovalova 2008

Study characteristics

Methods

Parallel randomised trial with three arms

1. Psychological placebo: Supportive care
2. No-treatment
3. Active treatment: Skills for Improving Distress Intolerance (SIDI)

Sample calculation: yes

Cluster randomised: no

Duration of trial (baseline to post): 3 weeks

Duration of participation (trial + follow-up): 3 weeks

Setting: inpatient

Purpose of trial : "To develop a treatment for prevention of treatment drop-out in a residential treatment setting." (p. 1)

Bornovalova 2008 *(Continued)*

Open/closed placebo: closed placebo

Data	**Number of participants screened**: 68 **Number of participants include** : 66 **Number of participants followed-up at post treatment**: 65 **Number of participants randomly assigned to:** • Psychological placebo: n = 19 • No-treatment: n = 25 • Active treatment: n = 22 **Number of withdrawals:** n = 1 • Psychological placebo: n = 0 • No-treatment: n = 0 • Active treatment: n = 1 **Diagnosis:** substance use disorder **Diagnostic manual:** Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) **Means of assessment:** The Structured Clinical Interview for DSM-IV Disorders (SCID) **Comorbidity:** • Bipolar I or II 13.4%, Major depressive disorder (MDD) 26.9%, Past MDD 35.8%, Social Phobia 10.4%, Generalized Anxiety Disorder 13.4%, Post-Traumatic Stress Disorder 14.9%, Borderline personality disorder (BPD) 26.9%, Antisocial personality disorder (APD) 37.3%, Substance Dependence: Alcohol 32.4%, Cannabis 10.4%, Heroin 29.9%, Cocaine 58.2%, Phencyclidin (PCP) 7.5%, Dependent > one drug class 41.8%. **Age:** 43.5 mean years (SD = 9.8) **IQ** : not stated - but approximately 35.8% of the participants had an education level of “less then high school”, 34.3 % had a “high school or equivalent” level, and 30% had “some college and above” level **Sex**: 20.6% female **Ethnicity:** 90.1 % African American **Country:** USA **Inclusion criteria** 1. Only participants who were low in distress tolerance (defined as the non-completion of at least one of the two behavioural tasks), 2. Were not evidencing acute psychosis 3. Were somewhat literate were eligible for participation in the treatment protocol **Exclusion criteria** 1. Complete abstinence from drugs and alcohol is required upon entry into the centre and through the duration of the program, with the Exception of caffeine and nicotine
Comparisons	**Psychological placebo** **Treatment name:** Supportive counselling **Description of intervention**: “To control for the non-specific elements of therapist contact, approximately one-third of the patients received SC, which also consisted of six individual sessions over 3 weeks. This treatment did not follow a clearly defined theoretical model, and was best described as un-

Bornovalova 2008 *(Continued)*

conditional support, combined with information and advice on managing current problems that a given patient may be experiencing. Although the format was rather open, therapists were provided with a manual providing a script for the initial session as well as potential topics for discussion and corresponding prompts. These included (but were not limited to) day to day annoyances/issues in the treatment center likes and dislikes about the centre, discussions of drug court status and concerns related to this, discussions of families and relationships, concerns about leaving the center, spirituality, relaxation and leisure time, and employment and finances. SC specifically avoided acceptance or mood induction techniques." (p. 28)

Individual or group treatment: individual

Exposure/intensity to treatment: 6 sessions over 3 weeks

Duration of treatment: 3 weeks

Concomitant psychotherapy: "TAU is basically no treatment (from us). They are still receiving addiction TX from the Residential Facility (as are all patients regardless of condition); But they don't get additional treatment from us. We just give them a pre and post" (Bornovalova 2008 (pers comm))

Concomitant pharmacotherapy: receiving Psychotropic Medication: 13.2%. Alcohol 32.4%, Cannabis 10.4%, Heroin 29.9%, Cocaine 58.2%, Phencyclidin (PCP) 7.5%

Comparator intervention

Comparison name: no-treatment

Description of intervention: "Of note, procedure for NTC was similar for baseline and post-test. However, no therapy was given." (p. 17). Sometimes labelled TAU.

Exposure/intensity to treatment: Duration treatment: 3 weeks

Concomitant psychotherapy: TAU is basically no treatment (from us). They are still receiving addiction TX from the Residential Facility (as are all patients regardless of condition); But they do not get additional treatment from us. We just give them a pre and post" (author correspondence)

Concomitant pharmacotherapy: receiving Psychotropic Medication 13.2%. Alcohol 32.4%, Cannabis 10.4%, Heroin 29.9%, Cocaine 58.2%, Phencyclidin (PCP) 7.5%

Outcomes

Beneficial effect

- **Hierarchy**: usable data, primary outcome, patient-reported
- **Outcome chosen**: Acceptance and Action Questionnaire 16 items (AAQ-16)

Adverse events

- No data on adverse events reported

Notes

Key conclusion from study authors

1. The current results suggest that SIDI is effective in increasing distress tolerance in inner-city drug users. Additionally, the variable rates of dropout that were, nevertheless, non-siginficant suggest a need for larger-scale studies to test the effect of SIDI on dropout.

Key limitations from study authors

1. Modest sample size
2. Sample population reduced generalisability
3. Placebo group got no homework, active intervention got homework
4. Self-report measures
5. Some individuals dropped out of post-therapy assessment

Other notes from review authors

1. None found

Bornovalova 2008 *(Continued)*

Conflicts of interest: potential n on-financial conflict of interest. The principal investigator (MAB) conducted a large majority of the SIDI group .

Judgment: no

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: “I used one of the randomisation websites to pre-create a randomisation list of subject numbers and conditions to which said numbers were assigned. (Bornovalova 2008 (pers comm))
Allocation concealment	Yes	Concealed (Bornovalova 2008 (pers comm))
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and "no-treatment"
Incomplete outcome data	No	Quote: “there were no dropouts in the SIDI or SC conditions (0%), and 4 individuals (16%) dropped out of the TAU group, resulting in an overall sample dropout rate of 6.1%. In contrast, when dropout at any point in treatment (thus including any time in their contract after SIDI was completed) was used as a dependent variable, the rates were somewhat different." (p. 74) ”Thus, the positive findings were likely inflated, as treatment completer rather then intent-to- treat analyses were utilized in this study.” (p. 74) Attrition rate >15% (16% in no-treatment group, 0% in other groups). No ITT. Reports data on completers only
Selective outcome reported	No	NCT01741415. No congruent between trial registry outcomes and full report. Some outcomes added – some deleted. Control group added
Other sources of bias	Yes	No other sources found

Bramston 1985

Study characteristics

Methods

Parallel randomised trial with four arms

1. Psychological placebo: Attention-Placebo Control Group
2. Wait-list: no-treatment
3. Active treatment 1: Social-problem-solving training
4. Active treatment 2: Behavioural Social-skills training

Sample calculation: not stated

Cluster randomised: no

Duration of trial (baseline to post : 4 weeks

Bramston 1985 *(Continued)*

Duration of participation (trial + follow-up): 4 weeks + 3-month follow-up

Setting: inpatient

Purpose of trial: "This study was therefore designed to investigate the effectiveness of a traditional SST approach as compared to a cognitively based social-problem-solving (SPS) programme in enhancing the social competence of intellectually-handicapped adults." (p. 240)

Open/closed placebo: closed placebo

Data

Number of participants screened: not stated

Number of participants included: 48

Number of participants followed-up at post treatment: 48

Number of participants randomised :

- Psychological placebo: n = 12
- Wait-list: n = 12
- Active treatment 1: n = 12
- Active treatment 2: n = 12

Number of withdrawals : n = 0

Diagnosis: intellectual disability

Diagnostic manual: not stated

Means of assessment: The Wechsler Adult Intelligence Scale (WAIS)

Comorbidity: not stated

Age: 28.06 mean years (SD = 7.08), (range = 18.03 to 46.5)

IQ: a mean IQ of 40.06 and 55.20.

Sex: not stated

Ethnicity: not stated

Country: Australia

Inclusion criteria

1. Intellectual disability

Exclusion criteria

1. Not stated

Comparisons

Psychological placebo

Treatment name: Attention-Placebo Control Group

Description of intervention: “In order to control for non-specific treatment effects the 12 Ss in this condition received an equivalent degree of therapist input and small group attendance as for the BSST and SPS training groups. ” (p.242)

Individual or group treatment: group

Exposure/intensity to treatment: sessions occurred four times a week over 4 weeks, each lasting for 30 minutes

Duration of treatment: 4 weeks

Bramston 1985 *(Continued)*

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy:: not stated

Wait-list

Comparison name:: no-treatment control group (in reality a wait-list)

Description of intervention: "Twelve Ss received no direct intervention other than that ordinarily offered by the centre during the daily routine. Assessments were made before and after the training phase on all measures applied to the experimental groups. Training was then made available following the post-training assessment." (p. 242)

Exposure/intensity to treatment: no treatment provided

Duration treatment: not stated

Concomitant psychotherapy: ordinary treatment

Concomitant pharmacotherapy:: ordinary treatment

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported, psychometric properties
- **Outcome chosen:** Social Skills Assessment Chart - Behaviour Ratings

Adverse events

- No data on adverse events reported

Notes

Key conclusion from study authors

1. Significant improvement in basic social-skill performance was found for the BSST group but not for the SPS, APC or NTC groups,
2. whereas significant increases in the generation of alternative solutions were found for the cognitive SPS group but not the BSST, APC or NTC groups.

Key limitations from study authors

1. the staff rating scale used may have been insufficiently sensitive to detect any changes in social competence which occurred

Other notes from review authors

1. None

Conflicts of interest: none found

Judgment: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Subjects were randomly allocated to one of four groups: a behavioural social-skills (BSST) programme, a social-problem-solving (SPS) programme, an attention-placebo control (APC) group and a no-treatment control (NTC) group." (p. 240)
Allocation concealment	Unclear	No information

Bramston 1985 *(Continued)*

Blinding of outcome assessors	Yes	Quote:"In the present study one rater completed the scale for all Ss at each assessment session. Thisrater was highly experienced in use of the scale but was blind as to the treatment condition of Ss." (p. 241)
Blinding of participants and personnel	No	Not possible to blind placebo and "no-treatment"
Incomplete outcome data	Yes	Quote:"There was no other attrition throughout the study." (p.240) Attrition <15% (0%)
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Brill 1964a

Study characteristics

Methods

Parallel randomised trial with six arms

1. Pharmacological placebo
2. Usual care: psychotherapy
3. Wait-list
4. Active treatment 1: meprobamate
5. Active treatment 2: phenobarbital
6. Active treatment 3: prochlorperazine

Sample calculation: not stated

Cluster randomised: no

Duration of trial (baseline to post): average 5 months (up to 12 months.)

Duration of participation (trial + follow-up): 5 + 10 - 18 months follow-up

Setting: outpatient

Purpose of trial: a controlled double-blind study of 299 non-psychotic female psychiatric clinic patients divided into six groups, with members of each group dealt with in a different manner from those in other groups.

Open/closed placebo: closed placebo

Data

Number of participants screened: not stated

Number of participants included: 299

Number of participants followed-up at post treatment: 169

Number of participants randomly assigned to:

- Pharmacological placebo: n = 55
- Usual care: n = 50
- Waiting list: n = 34
- Active treatment: n = 53
- Active treatment 2: n = 53

Brill 1964a *(Continued)*

- Active treatment 3: n = 54

Number of withdrawals: n = 130

- Pharmacological placebo: n = 25
- Usual care: n = 20
- Waiting list: n = 14
- Active treatment: n = 19
- Active treatment 2: n = 25
- Active treatment 3: n = 27

Diagnosis: psychiatric patients. (The sample included patients with personality disorders, psychoneuroses, psychosomatic disturbances, and borderline schizophrenic state)

Diagnostic manual: not stated

Means of assessment:: not stated

Comorbidity: different psychiatric diagnoses

Age: range = 20-40 years

IQ: average intelligence or better

Sex: 100% female

Ethnicit::100% Caucasian

Country: USA

Inclusion criteria

1. The selection of patients was limited to Caucasian females between the ages of 20 and 40 years
2. Who were of average intelligence or better
3. Who were nonpregnant
4. And who were not psychotic
5. Drug-sensitive
6. Severely depressed
7. Or suffering from a disabling physical disease
8. The sample included patients with personality disorders, psychoneuroses, psychosomatic disturbances, and borderline schizophrenic states (of the kind normally accepted by the Clinic). Patients with severe sociopathic disorders were not included

Exclusion criteria

1. Not stated

Comparisons	**Pharmacologoical placebo**

Treatment name: placebo pill

Description of intervention: “All capsules were identical in color and size.” (p. 583)

Individual or group treatment: individual

Exposure/intensity to treatment: not stated

Duration of treatment: average 5 months

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

Brill 1964a *(Continued)*

	Wait-list **Comparison name:** no-treatment (in reality wait-list) **Description of intervention:** "The no-treatment group was assigned at random from the pool of patients who had been examined. These patients were told that they could not be accepted for treatment immediately but that treatment might be available after about four months to a year. They were told that they would be contacted in approximately that time. Four months after the date of initial evaluation, the same explanation was repeated. Some patients were re-evaluated at that time ; others were again placed on a waiting list and were recalled during the year for re-evaluation." (p. 584) **Exposure/intensity to treatment:** no treatment during waiting **Duration treatment:** average 5 months **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy** : Observer-reported • **Outcome chosen:** Symptomatic adjustment **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. The findings suggested that the widespread preference for the traditional outpatient psychotherapy is based as much on the physician's bias as on proven greater effectiveness over briefer treatment methods **Key limitations from study authors** 1. Not stated **Other notes from review authors** 1. No usable data available **Conflicts of interest:** none found **Judgment:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	No information
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Unclear	Quote: "the therapists had no knowledge of the names, types, or number of drugs involved, nor even if there was a placebo being used." (p. 585). "it was evaluated blindly and because of the prejudice in favor of psychotherapy among patients and therapists." (p. 590)
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list

Brill 1964a *(Continued)*

Incomplete outcome data	No	Attrition > 15% (43.5 %)
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Brill 1964b

Study characteristics	
Methods	See Brill 1964a
Data	See Brill 1964a
Comparisons	**Usual care** **Treatment name:** psychotherapy **Description of intervention:** "The patients treated with psychotherapy were assigned to psychiatric residents who were closely supervised by the clinical staff of The Neuropsychiatrie Institute of the UCLA Center for Health Sciences. Each resident treated one or two psychotherapy patients. These were seen for 50-minute sessions at least once a week and could be seen more often at the discretion of the resident and his supervisor. Psychotherapy, while psychoanalytically oriented and generally nondirective in keeping with the attitude of the supervisory staff, was in varying degrees supportive. The average length of treatment was five months." (p. 583) **Individual or group treatment:** individual **Exposure/intensity to treatment:** 50 minutes at least once a week **Duration of treatment:** average 5 months **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated **Wait-list** **Comparison name:** no-treatment (in reality wait-list) **Description of intervention:** "The no-treatment group was assigned at random from the pool of patients who had been examined. These patients were told that they could not be accepted for treatment immediately but that treatment might be available after about four months to a year. They were told that they would be contacted in approximately that time. Four months after the date of initial evaluation, the same explanation was repeated. Some patients were re-evaluated at that time ; others were again placed on a waiting list and were recalled during the year for re-evaluation." (p. 584) **Exposure/intensity to treatment:** no treatment during waiting **Duration treatment:** average 5 months **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated
Outcomes	See Brill 1964a
Notes	See Brill 1964a

Brill 1964b *(Continued)*

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	No information
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Unclear	Quote: ”the therapists had no knowledge of the names, types, or number of drugs involved, nor even if there was a placebo being used.” (p. 585). "it was evaluated blindly and because of the prejudice in favor of psychotherapy among patients and therapists." (p. 590)
Blinding of participants and personnel	No	Not possible to blind usual care and wait-list
Incomplete outcome data	No	Attrition > 15% (43.5 %)
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Carlson 1993

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Pharmacological placebo: placebo pill 2. No-treatment: no pill 3. Active treatment: methylphenidate (MPH) **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 6 weeks. **Duration of participation (trial + follow-up):** 6 weeks + 2 weeks **Setting:** outpatient (an 8-week day-treatment program) **Purpose of trial:** investigating MPH’s effect on the performance and perceptions of Attention Deficit Hyperactivity Disorder (ADHD) boys following solvable and insolvable puzzles. **Closed/open placebo:** closed placebo
Data	**Number of participants screened:** not stated **Number of participants included:** 28 **Number of participants followed-up at post treatment:** 28 **Number of participants randomly assigned to** : • Pharmacological placebo: n = 9 • No-treatment: n = 9 • Active treatment: n = 10 **Number of withdrawals:** n = 2

Carlson 1993 *(Continued)*

- Pharmacological placebo: not stated
- No-treatment: not stated
- Active treatment: not stated

Diagnosis: Attention Deficit Hyperactivity Disorder (ADHD)
Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition, revised (DSM-III-R)
Means of assessment: clinical interviews and standardised rating scales
Comorbidity: "Based on the interview, 17 of the boys also obtained a DSM-III-R diagnosis of oppositional/defiant disorder, and 8 boys met criteria for a DSM-III-R
diagnosis of conduct disorder." (p. 273)
Ag:: 9.35 mean years (SD = 1.33)
IQ: "On the Wechsler Intelligence Scale for Children-Revised, subjects had a mean score of 108 (SD=10), and on the Woodcock-Johnson Test of Achievement Revised, they obtained a mean reading standard score of 97 (SD =12).
Sex: 100% male
Ethnicity: not stated
Country: USA

Inclusion criteria

1. ADHD
2. Higher scores in CASQ (more adaptive attributional style)

Exclusion criteria

1. Not stated

Comparisons

Pharmacological placebo
Treatment name: placebo
Description of intervention: received placebo pill each day. "The MPH and placebo were packed in identical opaque gelatin capsules." "Medication was given in the morning and at midday." (p. 274)
Individual or group treatment: individual
Exposure/intensity to treatment: two times a day
Duration of treatment: 6 weeks.
Concomitant psychotherapy: not stated - but was a part of a Summer Treatment Program
Concomitant pharmacotherapy: not stated

No-treatment
Comparison name: no pill (no treatment)
Description of intervention: no treatment
Exposure/intensity to treatment: no treatment
Duration treatment: 6 weeks
Concomitant psychotherapy: not stated - but was a part of a Summer Treatment Program
Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effects

- **Hierarchy:** observer-rated, clinical relevancy
- **Outcome chosen:**: numbers of stopped early - subscale unsolvable – observer-rated

Adverse events

- None mentioned

Notes

Key conclusions from study authors

1. "Subjects exposed to insolvable puzzles showed greater persistence on a subsequent generalization task when receiving MPH as compared to placebo. (p. 270)
2. "They failed to find any differences between the no-pill and placebo conditions for any of the measures, whereas the placebo and MPH conditions differed consistently." (p. 282)

Carlson 1993 *(Continued)*

3. "On medication, compared to placebo, the boys solved more of the puzzles, and tended to stop early less often and to find more of the word on the final puzzle." (p. 282)

Key limitations from study authors

1. "The present study examined performance and attributions on a single- task in the context of a short-term medication assessment." (p. 185)

Other notes from review authors

1. No report of how many participants was randomised to each group. Since 28 patients was included in total, we assumed due to randomisation and ethical principles that the active arm (in this case methylphenidate) included an additional patient compared with the placebo and no-treatment group (see Table 2)

Conflicts of interest: none found

Judgment: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Order of the conditions was randomized, with the constraint that for all boys the first two sessions were either both solvable or both insolvable conditions." (Milich 1991, p. 524) No information in Carlson 1993 if the study was randomised, but replicates (Milich 1991), which was randomised.
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	No information but outcomes was objective measures (cognitive test)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Unclear	Quote: "Two were dropped from analyses because on at least one solvable day they failed to find any solvable puzzles" (p. 272) Attrition <15% (7,1%). However, only reports data on completers
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Carter 2003

Study characteristics

Methods

Parallel randomised trial with three arms

1. Psychological placebo: nonspecific self-help
2. No-treatment: wait-list
3. Active treatment: cognitive behavioural self-help

Sample calculation: yes

Carter 2003 *(Continued)*

	Cluster randomised: no **Duration of trial (baseline to post):** 8 weeks **Duration of participation (trial + follow-up):** 8 weeks (endpoint data) **Setting:** inpatient **Purpose of trial:** "The aim of the present study was to evaluate the effectiveness of a cognitive behavior self-help manual for patients with bulimia nervosa who were on a waiting list for treatment at a hospital-based specialist clinic." (p. 973) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** 245 **Number of participants included:** 85 **Number of participants followed-up at post treatment:** 65 **Number of participants randomly assigned to:** • Psychological placebo: n = 28 • No-treatment: n = 29 • Active treatment: n =28 **Number of withdrawals:** n = 20 • Psychological placebo: n = 7 • No-treatment: n = 8 • Active treatment: n = 5 **Diagnosis:** Bulimia nervosa **Diagnostic manual:** Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) **Means of assessment:** Eating Disorder Examination (EDE) **Comorbidity:** not stated **Age:** 27 mean years (SD = 8), (range=17 to 53) **IQ:** not stated **Sex:** 100% female **Ethnicity:** 83% Caucasian, 2% African Caribbean, 7% Asian, and 8% other. **Country:** Canada **Inclusion criteria** 1. Bulimia nervosa **Exclusion criteria** 1. Younger than age 17 2. Pregnant 3. Medical illness or treatment known to influence eating or weight (e.g., diabetes mellitus), 4. Current or previous specialist treatment for an eating disorder 5. Body mass index (kg/m2) under 18.
Comparisons	**Psychological placebo**

Carter 2003 *(Continued)*

Treatment name : nonspecific self-help

Description of intervention: "The second self-help condition (nonspecific self-help) was designed to control for nonspecific factors, such as receiving a self-help book, hearing a plausible rationale, and expecting to improve. It involved following the self-help manual Self-Assertion for Women. This self-help manual focuses on developing assertiveness skills and does not in any way address the specific symptoms of bulimia nervosa. This control intervention was selected because it might be regarded by patients as a credible alternative treatment, since many women with eating disorders report experiencing significant interpersonal difficulties, including inhibited self-assertion. Like the cognitive behavior self-help manual, the nonspecific self-help book contains both psychoeducational information and practical advice designed to foster behavioral change. Both books were similar in length and level of difficulty." (p. 974)

Individual or group treatment: individual (self-help)

Exposure/intensity to treatment: not stated

Duration of treatment: 8 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: "Patients who were taking an established dose of antidepressant medication were eligible to take part" (p. 974)

No-treatment

Comparison name: waiting list control group (in reality a no treatment)

Description of intervention: "Individuals assigned to the waiting list condition received no intervention." (p. 974)

Exposure/intensity to treatment: no exposure

Duration treatment: 8 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: patients who were taking an established dose of antidepressant medication were eligible to take part" (p. 974)

Outcomes

Beneficial effect

- **Hierarchy:** usable data, patient-reported, clinical relevance
- **Outcome chosen:** Eating Disorder Examination, subscale Eating Concern

Adverse events

- No data on adverse events reported

Notes

Key conclusion from study authors

1. Although the group-by-time interaction for binge eating and purging was not statistically significant, simple effects showed that there was a significant reduction in symptom frequency in both self-help conditions at posttreatment but not in the waiting list condition

Key limitations from study authors

1. A limitation of the study is that only 69.1% of those who appeared eligible to take part, according to the telephone screening interview, agreed to participate. This may limit the generalisability of the findings

Other notes from review authors

1. None

Conflicts of interest: none found

Carter 2003 *(Continued)*

Judgment: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote ”A restricted randomization procedure employing random permuted blocks of three people was used to ensure approximately equal numbers of participants in the three conditions.” (p. 974)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Quote: ”An assessor who was blind to the patients’ treatment assignment performed the posttreatment assessments 8 weeks later.” (p. 974) Patient-reported outcomes used
Blinding of participants and personnel	No	Not possible to blind placebo and "no-treatment"
Incomplete outcome data	Unclear	Quote: "Twenty participants (23.5%) dropped out of the study and did not attend the posttreatment assessment: five (17.9%) of these were from the cognitive behavior self help group (N=28), seven (25.0%) were from the nonspecific self-help group (N=28), and eight (27.6%) were from the waiting list control group (N=29). There was no statistically significant difference between the three conditions in terms of the rate of attrition. This dropout rate is similar to those reported in previous treatment studies (...). There were no significant differences between the dropouts and completers in terms of baseline characteristics.” (p. 975) Attrition >15% (23.5%). ITT analyses made
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Crisp 1991

Study characteristics

Methods

Parallel randomised trial with four arms

1. Usual care: In-patient treatment
2. No-treatment: No further treatment
3. Active treatment 1: outpatient individual and family psychotherapy plus separate dietary counselling
4. Active treatment 2: outpatient group psychotherapy (patients and parents) plus separate dietary counselling

Sample calculation: not stated

Cluster randomised: no

Duration of trial (baseline to post): mean 20 weeks

Duration of participation (trial + follow-up): 20 weeks treatment, post treatment assessment at 1 year

Crisp 1991 *(Continued)*

Setting: inpatient for usual care (outpatient for other group)

Purpose of trial: "The present study involves an extension of this established treatment approach, allowing a controlled investigation of the effects of psychotherapy directed at the development and family psychopathology, together with dietary management, provided within three different treatment settings; and comparison with a no-treatment group." (p. 327)

Data

Number of participants screened: not stated

Number of participants included: 90

Number of participants followed-up at post treatment: 73

Number of participants randomly assigned to:

- Usual care: n = 30
- No-treatment: n = 20
- Active treatment 1: n= 20
- Active treatment 2: n = 20

Number of withdrawals: n = 17

- Usual care: n = 12
- No-treatment: n = 2
- Active treatment 1: n = 3
- Active treatment 2: n = 0

Diagnosis: Anorexia nervosa

Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition, revised (DSM-III-R)

Means of assessment: diagnostic assessment interview

Comorbidity: not stated

Ag:: usual care 23.2 mean years (SD =4.9), no-treatment; 21.9 mean years (SD = 4.5)

IQ: not stated

Sex: 100% female

Ethnicity: not stated

Country: UK

Inclusion criteria

1. Patients received diagnoses of undoubted anorexia nervosa and all fulfilled DSM-III-R criteria
2. Being female
3. Having a duration of illness of less than 10 years
4. Living within out-patient reach of the service.

Exclusion criteria

1. Not stated

Comparisons

Usual care

Treatment name: inpatient treatment

Description of intervention: "It was taken to be the established treatment (Crisp, 1980) and believed to be effective (...) In-patient treatment was intensive and involved much greater patient contact than

Crisp 1991 *(Continued)*

did the other two treatment options (…) Treatment involved weight restoration to the mean matched-population weight (MMPW) at the age of onset of anorexia, supported by weekly individual therapy, family therapy, group therapy, dietary counselling and occupational therapy, including psychodrama and projective art techniques. Inpatient treatment was followed by 12 sessions of out-patient psychotherapy involving the patient and family." (p. 328)

Individual or group treatment: both individual and group treatment

Exposure/intensity to treatment : not stated

Duration of treatment: mean stay was 20 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: "Psychotropic drugs were not prescribed for or consumed by any of the patients in the three ongoing treatment groups during the period of study." (p. 328)

No-treatment

Comparison name: no further treatment

Description of intervention: "Patients allocated to option 4, no further treatment ('one-off), were referred back to their family doctor or local consultant, who received a detailed report of the assessment together with advice on further management." (p. 328)

Exposure/intensity to treatment: no treatment

Duration treatment: 12 months

Concomitant psychotherapy: not stated, but probably, as most of the participants were treated elsewhere while enrolled in the study

Concomitant pharmacotherapy: not stated, but probably, as most of the participants were treated elsewhere while enrolled in the study

Outcomes

Beneficial effect

- **Hierarchy:** Global score
- **Outcome chosen:** Morgan and Russel Mean Scores, Global Score, One-year follow-up

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. "We are left with the conclusion that all three interventions are powerful in their effect at one-year follow-up. The out-patient interventions are clearly less intensive and less expensive than the in-patient package

Key limitations from study authors

1. "This study has been fraught with difficulties that relate to the well known problem of engaging and maintaining patients with anorexia nervosa in treatment and follow-up. We believe our assessment and treatment procedures normally contain and minimize these, but the imposition of the study exacerbated them." (p. 331)

Other notes from review authors

1. None

Conflicts of interest: none found

Judgment: yes

Crisp 1991 *(Continued)*

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Following an assessment of the kind described above, 90 patients were randomly allocated to one or other of the four options (…)" (p. 327)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind usual care and no-treatment
Incomplete outcome data	No	Attrition rate: 17/90 (18.9%). In-patient: 12/30 (40%) Psychotherapy: 2/20 (10%). Group psychotherapy: 3/20 (15%). One-off: 0. Quote:"Drop-out occurred in all three treatment groups, especially the out-patient group psychotherapy. One of the patients allocated to this treatment died as a consequence of her anorexia nervosa before the treatment could begin. This was the only death" (p. 331) "We were distressed to find patients refusing treatment or dropping out because of forced allocation when they would have preferred - or we would have preferred to have offered them another therapy" (p. 331) Quote:"the sample will not include patients who refused in-patient admission in the first instance or who dropped out of in-patient treatment, refusing further intervention." (p. 326) Attrition > 15% (18.9%). Drop-out on usual care group was 40%. No ITT. Only reports data on completers
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Crouch 1988

Study characteristics

Methods	**Parallel arm with three arms** 1. Pharmacological placebo: placebo pill 2. No-treatment: no pill 3. Active treatment: propranolol **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 10 weeks **Duration of participation (trial + follow-up):** 10 weeks + 3 month follow-up **Setting:** outpatient (withdrawal clinic) **Purpose of trial:** "We therefore present some characteristics of patients referred for treatment of tranquillizer dependence to a clinic specializing in its treatment" (p. 503-4) **Open/closed placebo:** closed placebo

Crouch 1988 *(Continued)*

Data	**Number of participants screened:** 91 **Number of participants included:** 44 **Number of participants followed-up at post treatment:** 23 **Number of participants randomly assigned to** : • Pharmacological placebo: n = 10 • No-treatment: n = 8 • Active treatment: n = 5 **Number of withdrawals** : n = 21 • Pharmacological placebo: not stated • No-treatment: not stated • Active treatment: not stated **Diagnosis:** substance dependence (benzodiazepine) **Diagnostic manual:** not stated **Means of assessment:** assessment interview. All patients assessed with The State-Trait Anxiety Inventory (STAI). Otherwise, referred to by local general practitioners **Comorbidity:** Generalized anxiety disorder (GAD), n = 13; Panic disorder, n = 5; Agoraphobia; n = 1; somatic symptoms, n = 8; Insomnia, n =4; Drug abuse, n = 2; Alcohol abuse, n =1; Depression, n = 2. **Age:** 41.4 mean years (SD =10.8) **IQ:** not stated **Sex:** 14 male, 30 women **Ethnicity:** not stated **Countr:**: UK **Inclusion criteria** 1. Taking benzodiazepines regularly for at least four months 2. Wished to stop **Exclusion criteria** 1. Concurrent severe affective disorder 2. Abusing alcohol or other drugs 3. Physical illness which could be compromised by taking propranolol
Comparisons	**Pharmacological placebo** **Treatment name:** placebo pill **Description of intervention:** no information about placebo pill. “Group support was structured with active interventions, anxiety management training with a cognitive-based component” (p. 504). **Individual or group treatment:** Individual + group **Exposure/intensity to treatment:** 160 mg per day + “group meetings were held weekly for 5 weeks, then after 2 and 4 weeks respectively and lasted approximately one hour“ (p. 505) **Duration of treatment:** 10 weeks **Concomitant psychotherapy:** all patients received group support **Concomitant pharmacotherapy:** “No other medication was taken, apart from steady reduction of diazepam/lorazepam over a four-week period” (p. 504) **No-treatment** **Comparison name:** no pill

Crouch 1988 *(Continued)*

	Description of intervention: no information about no pill group. "Group support was structured with active interventions, anxiety management training with a cognitive-based component" (p. 504). **Exposure/intensity to treatment:** no pill + "group meetings were held weekly for 5 weeks, then after 2 and 4 weeks respectively and lasted approximately one hour" (p. 505) **Duration treatment:** 10 weeks **Concomitant psychotherapy:**: all patients received group support **Concomitant pharmacotherapy:** "No other medication was taken, apart from steady reduction of diazepam/lorazepam over a four week period" (p. 504)
Outcomes	**Beneficial effect** • **Hierarchy:** observer-reported • **Outcome chosen:** Hamilton Anxiety Rating Scale (no usable data) **Adverse events** • No data on adverse events reported
Notes	**Key conclusion from study authors** 1. Anxiety management training in group resulted in a considerable reduction in tranquillizer intake and half of the subjects managed to stop tranquillizers altogether despite previous failures. **Key limitations from study authors** 1. A high proportion of patients reported previous contact with psychiatric services. 2. Small sample size 3. Only 5 patients completed in the treatment in propranolol 4. 2 patients in the propranolol group revealed that they had not taken the medication. **Other notes from review authors** 1. No usable data (see Table 2) 2. 31 started treatment, only reports data on 23 completers **Conflicts of interest:** none found **Judgement:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Patients were allocated randomly to either highly structured cognitive based group therapy or offered group therapy of a more supportive non-interventionist nature. The patients were also randomly allocated to one of 3 treatment groups. Either propranolol ("Inderal LA") 16mg per day, matching placebo or no medication ("no pills")." (Hallström 1988, p. 41)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	No	Attrition >15% (52,3%). Only reports data on completers

Crouch 1988 *(Continued)*

Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Doty 1975

Study characteristics	
Methods	**Parallel randomised trial with five interventions** 1. Psychological placebo: Non-specific control 2. No treatment 3. Active treatment 1: Social skills training 4. Active treatment 2: Incentive condition 5. Active treatment 3: Combination condition **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 2 weeks **Duration of participation (trial + follow-up):** 2 weeks treatment + 2 weeks follow-up **Setting:** inpatient **Purpose of trial:** "Do training in social skills or incentives, or a combination of both contribute to either the daily social interaction rates or the social responsiveness of psychiatric patients?" (p. 677) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** 96 **Number of participants included** : 56 **Number of participants followed-up at post treatment:** 39 **Number of participants randomly assigned to:** • Psychological placebo: n = 12 • No treatment: n = 8 • Active treatment 1: n = 12 • Active treatment 2: n = 12 • Active treatment 3: n = 12 **Number of withdrawals:** n = 17 • Psychological placebo: not stated • No treatment: not stated • Active treatment 1: not stated • Active treatment 2: not stated • Active treatment 3: not stated **Diagnosis: o**pen-ward psychiatric patients

Doty 1975 *(Continued)*

- "56 male open-ward psychiatric patients at the Veterans Administration Hospital, Danville, Illinois, who were nominated for the study by nursing personnel on their wards as being noninteractive, relatively cooperative, and not engaging in active delusional or hallucinatory behavior." (p.677)

Diagnostic manual: not stated

Means of assessment: not stated

Comorbidity: not stated

Age: 47.98 mean years

IQ:: not stated

Sex:: 100% male

Ethnicity : not stated

Country: USA

Inclusion criteria

1. Psychiatric patients from hospital
2. were nominated for the study by nursing personnel on their wards as being noninteractive, relatively cooperative

Exclusion criteria

1. not engaging in active delusional or hallucinatory behavior

Comparisons

Psychological placebo

Treatment name: nonspecific control condition

Description of intervention: "This treatment was modeled after the attention-control group used by Wollersheim (1968) in that it was designed to control for the nonspecific therapy elements such as attention and positive regard from the therapist, knowledge of the target behaviors, and the expectancy that the four treatment sessions would lead to positive behavior change by the subjects. The vehicle for this control was lectures by the therapists following a transactional games analysis orientation and attempting to examine the supposed intrapsychic reasons why the subjects did not engage in more social interaction. Opportunities for the subjects to role play sample interactions or mention of concrete incentives for behavior change were specifically and intentionally avoided." (p. 678)

Individual or group treatment: group

Exposure/intensity to treatment: 4 sessions

Duration of treatment: 2 weeks

Concomitant psychotherapy: "None of the subjects were participating in other forms of active psychological treatment except chemotherapy or assignment to recreational activities at the time of the study." (p. 677)

Concomitant pharmacotherapy: "52 of the 56 subjects were receiving psychotropic medication at the time of the study." (p. 677)

No-treatment

Comparison name: no-treatment control

Description of intervention: "The only contact that the subjects in this condition had with the therapists was in the pretreatment and posttreatment social responsiveness assessments. In fact, these subjects were never told that they would receive treatment and, therefore, received none of the nonspecific treatment elements such as knowledge of the target behaviors, encouragement to change, etc." (p. 679)

Doty 1975 *(Continued)*

Exposure/intensity to treatment: not stated

Duration treatment: 4 weeks

Concomitant psychotherapy: "None of the subjects were participating in other forms of active psychological treatment except chemotherapy or assignment to recreational activities at the time of the study." (p. 677)

Concomitant pharmacotherapy "52 of the 56 subjects were receiving psychotropic medication at the time of the study." (p. 677)

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported
- **Outcome chosen:** Ward behavior observations

Adverse events

- No data on adverse event reported

Notes

Key conclusion from study authors

1. Trend analyses of ward data and post hoc
2. t-tests with the discussion data consistently indicated significant positive
3. changes at posttreatment for only those groups receiving monetary incentives.

Key limitations from study authors

1. The restriction of treatment to four sessions may have produced results with limited generality, in that very different results might have been obtained in an examination of more extended treatment
2. First, the failure of the role-playing condition subjects to demonstrate significant changes either on the ward behavior or the in session assessment data suggests that short-term treatments that fail to provide concrete incentives for behavior change outside the treatment sessions may prove fruitless
3. Second, but equally important, the behavior changes evidenced by the treatment that focused solely on incentives serves as an indictment of the incentives and encouragements typically supplied by the traditional hospital milieu. As argued earlier, continued social responsiveness and interaction is important for the consensual validation of significant events it offers the patient and because it would seem to be a prerequisite for positive response to most traditional treatment programs and for the individual's eventual release from the hospital.

Other notes from review authors

1. Usable data not available

Conflicts of interest: none found

J**udgement:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: " Subjects were randomly assigned from stratified blocks formed on the basis of pretreatment levels of daily social interaction either to one of three active treatment conditions" (p. 677) "The success of the random stratified-block subject-assignment procedure in establishing the pre-experimental group equation was checked using Treatment X Therapist analyses of variance for the group means and Bartlett's test of homogeneity of variance on the following variables both before treatment and rechecked after the data rejection mentioned above:" (p. 679)

Doty 1975 *(Continued)*

Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote: ""The observers were blind to both the nature of the dependent variable to be extracted from their recordings and the group assignments of the subjects" (p. 678) "Trained observers, unaware of the experimental group identity of the subjects,(...)" (p. 678)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	No	Quote: ""In order to make the statistical analyses as meaningful as possible it was necessary in some cases to drop some subjects' data from consideration. For instance, eight subjects were dropped for failure to adequately expose themselves to the treatments (only subjects attending three or more sessions were included), and nine subjects were dropped from the analysis of the ward data because of incomplete data. Thus, criteria for data selection were as objective as possible, and attrition rates were approximately the same across the various subject groups" (p. 680) Attrition >15% (33.9%). No ITT. Excludes non-completers
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Double 1993

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Pharmacological placebo: Placebo pill 2. No-treatment 3. Active treatment: anticholinergic medication **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** phase one = 4 weeks (total 12 weeks) **Duration of participation (trial + follow-up)** : not stated **Setting:** inpatient **Purpose of trial:**: "This study was rigorously designed to evaluate the efficacy of the long-term use of anticholinergic agents in patients maintained on neuroleptics." (p. 381) **Open/closed placebo:**: closed placebo
Data	**Number of participants screened:** 96 **Number of participants included:** 27 **Number of participants followed-up at post treatment:** 23 **Number of participants randomly assigned to** : • Pharmacological placebo: n = 9 • No-treatment: n = 9 • Active treatment: n = 9

Double 1993 *(Continued)*

Number of withdrawals : n = 4

- Pharmacological placebo: not stated
- No-treatment: not stated
- Active treatment: not stated

Diagnosis: psychiatric in-patients
Diagnostic manual: not stated
Means of assessment: not stated
Comorbidity: not stated
Age: 54 mean years (range = 22-76)
IQ: not stated
Sex: 29.6% female
Ethnicity: not stated
Country: UK

Inclusion criteria

1. Not stated

Exclusion criteria

1. No age restrictions were used in this study
2. Patients were screened for a history of prefrontal leucotomy or organic brain disease or clinical signs of dementia
3. None of the eligible patients met these exclusion criteria

Comparisons

Pharmacological placebo

Treatment name: placebo
Description of intervention: "Placebo and active medication for the trial were produced by the Pharmacy Manufacturing Department at the Royal Hallamshire
Hospital, Sheffield, from raw materials obtained from manufacturers. The capsules looked identical and were tested for quality control. They were made in the strength that they are produced commercially, so that patients in the trial, instead of receiving tablets, were given capsules in the same number as usually prescribed." (p. 382)
Individual or group treatment: individual
Exposure/intensity to treatment: up to 10 mg per day
Duration of treatment: 4 weeks
Concomitant psychotherapy: not stated
Concomitant pharmacotherapy: on concomitant antiparkinsonian and neuroleptic medication for over one year

No-treatment
Comparison name: no drug (no-treatment)
Description of intervention: not stated
Exposure/intensity to treatment: not stated
Duration treatment: 4 weeks
Concomitant psychotherapy: not stated
Concomitant pharmacotherapy: on concomitant antiparkinsonian and neuroleptic medication for over one year

Outcomes

Beneficial effects

- Hierarchy: available data
- Outcome: number of patients with relapse of parkinsonian symptoms (= need for escape medication) Extrapyramidal Symptom Rating Scale (ESRS) – Binary – observer-reported

Adverse events

- None reported

Double 1993 *(Continued)*

Notes

Key conclusion from study authors

1. The relapse rate on no medication was 14%, and if patients relapsed on no medication they also relapsed on placebo.
2. The relapse rate was not significantly different on active medication. Nor were there significant differences in ratings of Parkinsonism or dyskinesia.
3. The lack of difference between double-blind and overt withdrawal does not mean that studies that find a much higher relapse rate are necessarily unaffected by nonspecific factors, as significant unblinding may occur in clinical trials.

Key limitations from study authors

1. Unblinding occurs far more commonly in clinical trials than is generally appreciated.
2. The studies that find that anticholinergic medication seems to be needed for clinical stability may be affected by this bias.
3. The role of nonspecific factors in the studies that find a high relapse rate seems to require further investigation

Other notes from review authors

1. The outcome data was not available from the first period only, and was calculated as deriving from a parallel group trial.

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	No information
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote: ”Assessments were allocated equally between 3 assessors so that each patient was rated by a different assessor under each condition. There were, therefore, no carry-over effects from one assessment to the next. Assessors remained blind to the previous ratings by other assessors.” (p. 382) Assessors remained blind to the previous ratings by other assessors. (p. 382)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	No	Quote ”Complete assessments were unavailable for 4 patients (15%) because of drop-out from the trial at different stages” (p. 382) Attrition = 15%. Only reports data on completers
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Ehlers 2014

Study characteristics	
Methods	**Parallel randomised trial with four arms** 1. Psychological placebo: emotion-focused supportive therapy: 2. Wait-list 3. Active treatment 1: weekly cognitive therapy 4. Active treatment 2: intensive cognitive training **Sample calculation:** yes **Cluster randomised:** no **Duration of trial (baseline to post):** 14 weeks **Duration of participation (trial + follow-up):** 14 weeks + 27 weeks (follow-up 1) and 40 weeks (follow-up 2) **Setting:** outpatient **Purpose of trial:** "This clinical trial had two goals, (1) to investigate the acceptability and efficacy of a 7-day intensive version of cognitive therapy for PTSD, and (2) to investigate whether cognitive therapy has specific treatment effects by comparing intensive and standard weekly cognitive therapy with an equally credible alternative treatment." (p.1) **Open/closed placebo** : closed placebo
Data	**Number of participants screened:** 253 **Number of participants included:** 125 (only reports data on 121 completers) **Number of participants followed-up at post treatment** :112 **Number of participants randomly assigned to:** • Psychological placebo: n = 30 • Wait-list: n= 30 • Active treatment 1: n = 31 • Active treatment 2: n = 30 **Number of withdrawals:** n =9 • Psychological placebo: n = 6 • Wait-list: n = 0 • Active treatment 1: n = 0 • Active treatment 2: n = 3 **Diagnosis:** Chronic Post-traumatic Stress Disorder (PTSD) **Diagnostic manual:** Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DMSM-IV) **Means of assessment:** the Structured Clinical Interview for DSM-IV Disorders (SCID) **Comorbidity:** 63.6% had comorbid other Axis I disorders (mainly mood and anxiety disorders, substance abuse), and 19.8% had Axis II disorders (mainly obsessive-compulsive, depressive, paranoid, avoidant). **Age:** placebo: 37.8 mean years (SD = 9.9), wait-list; 36.8 mean years (SD =10.5) **IQ:** not stated **Sex:** 60% female

Ehlers 2014 *(Continued)*

Ethnicity: placebo 73.3% Caucasian, wait-list 70 % Caucasion

Country: UK

Inclusion criteria

1. Between 18-65 years old
2. Met diagnostic criteria for chronic PTSD as determined by the Structured Clinical Interview for DSM-IV
3. Their current intrusive memories were linked to one or two discrete traumatic events in adulthood
4. PTSD was the main problem

Exclusion criteria

1. History of psychosis
2. Current substance dependence
3. Borderline personality disorder
4. Acute serious suicide risk
5. Treatment could not be conducted without the aid of an interpreter

Comparisons

Pscyhological placebo

Treatment name: Emotion-focused Supportive Therapy

Description of intervention: "This non-directive treatment focused on patients' emotional reactions rather than their cognitions. It was designed to provide a credible therapeutic alternative to control for nonspecific therapeutic factors so that observed effects of cognitive therapy could be attributed to its specific effects beyond the benefits of good therapy." (p.4)

Individual or group treatment: individual

Exposure/intensity to treatment: "it comprised up to 12 weekly individual sessions (up to 20 hours in total) over three months and optional three monthly booster sessions." (p.4)

Duration of treatment: 12 weeks

Concomitant psychotherapy: 3,3% started another psychological treatment during the study.

Concomitant pharmacotherapy: "Patients taking psychotropic medication (29.8%) were required to be on a stable dose for two months before random allocation." (p.3). No one started a new medication during the trial.

Wait-list

Comparison name: wait-list

Description of intervention: patients allocated to wait-list waited for 14 weeks before receiving treatment

Exposure/intensity to treatment: no treatment during waiting

Duration treatment: 14 weeks

Concomitant psychotherapy: none reported

Concomitant pharmacotherapy: "Patients taking psychotropic medication (29.8%) were required to be on a stable dose for two months before random allocation." (p.3). No one started a new medication during the trial.

Outcomes

Beneficial effect

- **Hierarchy:** Primary, observer-reported
- **Outcome chosen:** Clinician-rated PTSD symptoms

Ehlers 2014 *(Continued)*

Adverse events

- Serious "No adverse effects (i.e., negative reactions to treatment procedures such as significant increases in dissociation, suicidal intent or hyperarousal) were reported in any of the groups." (p. 7)
- Non-serious: CAPS (deterioration)

Notes

Key conclusion from study authors

1. Cognitive therapy for PTSD delivered intensively over little more than a week is as effective as cognitive therapy delivered over 3 months.
2. Both had specific effects and were superior to supportive therapy.
3. Intensive cognitive therapy for PTSD is a feasible and promising alternative to traditional weekly treatment

Key limitations from study authors

1. Small sample size
2. The study focused on traumatic events in adulthood, and it will need to be investigated whether the results generalize the treatment of childhood trauma

Other notes from review authors

1. Only reports data on 121 completers

Conflicts of interest: potential n on-financial conflict of interest: "The treatment follows Ehlers and Clark's model of PTSD (...) ..." (p.3). First and last author developed the active treatment intervention

Judgement: no

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: "They were then randomly allocated (...) using the minimization procedure(...) to stratify for sex and severity of PTSD symptoms" (p.3)
Allocation concealment	Yes	Quote: "They were then randomly allocated to one of the four trial conditions by an independent researcher who was not involved in assessing patients (...). Assessors determining the suitability of a patient for inclusion were not informed about the stratification variables and algorithm." (p.3)
Blinding of outcome assessors	Yes	Quote: "Assessments of treatment outcome were conducted by independent evaluators without knowledge of the patient's treatment condition. Patients were asked not to reveal their group assignment to the evaluators. "(p. 3)
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	Yes	Quote: "Dropouts were defined as attending fewer than 8 sessions (...), unless the earlier completion was agreed with the therapist. Dropout rates were low and did not differ between conditions (Table 2). Only one patient in the supportive therapy group reported symptom deterioration on the Posttraumatic Diagnostic Scale (Table 2). On the CAPS, fewer patients treated with intensive and cognitive therapy were rated as having deteriorated than those in the wait-list condition. The supportive therapy group did not statistically differ from the other groups." (p.7) Attrition <15% (4.1%).
Selective outcome reported	Yes	Trial registry: ISRCTN 48524925.

Ehlers 2014 *(Continued)*

		No apparent differences in reporting between trial registry and full report.
Other sources of bias	Yes	No other sources found

Espie 1989a

Study characteristics

Methods	**Parallel randomised trial with six arms** 1. Psychological placebo: imagery relief placebo 2. Wait-list: no treatment 3. Active treatment 1: relaxation 4. Active treatment 2: stimulus control 5. Active treatment 3: paradoxical intention 6. Active treatment 4: tailored treatment **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 8 weeks **Duration of participation (trial + follow-up):** 8 weeks treatment + follow-up at 6 weeks, 3 months, 6 months and 17 months **Setting:** outpatient **Purpose of trial:** "Treatment process and outcome were investigated in terms of mean and standard deviation (night to night variability) measures of sleep pattern and sleep quality. (p. 80) **Open/closed placebo:** closed
Data	**Number of participants screened:** 141 **Number of participants included:** 101 **Number of participants followed-up at post treatment:** 84 (completers) **Number of participants randomly assigned to** : • Psychological placebo: n = 14 • Wait-list: n= 13 • Active treatment 1: n = 14 • Active treatment 2: n = 14 • Active treatment 3: n = 15 • Active treatment 4: n=14 **Number of withdrawals:**: n = 17 • Psychological placebo: not stated • Wait-list: not stated • Active treatment 1: not stated • Active treatment 2: not stated • Active treatment 3: not stated • Active treatment 4: not stated **Diagnosis:**: chronic insomniacs

Espie 1989a *(Continued)*

Diagnostic manual:: not stated

Means of assessment: not stated

Comorbidity: not stated

Age: 45.5 mean years (SD = 15.9), (range 17 to 82)

IQ: not stated

Sex: 67.9% female

Ethnicity: not stated

Country: UK

Inclusion criteria

1. Sleep-onset latency greater than 30 minutes on average per night.That is a total latency of at least 3.5 hours over the week
2. Chronic initial insomnia present for minimum of one year
3. Previous advice-seeking as evidence of clinical relevance of insomnia. In practice this constituted written physician referral
4. Legitimate to treat insomnia in isolation as the main presenting or primary problem
5. Able to ensure the withdrawal of all drugs which might interfere with the experimental design, or to maintain the patient on the same dosage of the same drugs throughout the period of study
6. No other ongoing therapy for insomnia, anxiety or depression

Exclusion criteria

1. Exclusion of patients presenting as clinically depressed at initial interview or with scores of 60 or higher on the Zung Depression Scale
2. Exclusion also if anti-depressant medication had been prescribed at any time during the 6 months preceding referral
3. Exclusion of patients considered to have drink problems
4. Exclusion of insomnia problems possibly related to medical conditions, and of sleep disorders not conforming to categories 1,2 or 9 of the Diagnostic Classification of Sleep and Arousal Disorders
5. Exclusion of patients non-compliant with either treatment instruction or adequate record-keeping

Comparisons

Psychological placebo

Treatment name: Imagery relief placebo

Description of intervention: “Patients in this group were treated in accordance with the quasi—desensitisation placebo instructions commonly used in past research (Steinmark 1974). The term "imagery relief" was, however, coined by the author. (...)The programme was, therefore, analogous to relaxation therapy/desensitisation but with the important omission of any known active ingredient, either theoretically or practically. Patientsreceived no instruction in dealing with sleeplessness per se.” (Espie 1989b p. 105-6,)

Individual or group treatment: individual

Exposure/intensity to treatment: not stated, but at least once weekly

Duration of treatment: 8 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: “Of the 84 patients, 58 (69%) were stated, by their GPs, to be "drugfree". Only 14 of these patients, however, had never been on hypnotic medication, and it transpired upon further assessment, that approximately one third of the remainder had not entirely discontinued medication. In most cases drug use was occasional and low dose. (…) Twenty—six patients were

Espie 1989a *(Continued)*

	referred with persisting sleep difficulties who were also on nightly sleep medication. ” (Espie 1989b , p. 101 - 2) **Wait-list** **Comparison name:** : no treatment (in reality a wait-list) **Description of intervention: "**This group functioned as a waiting list control and had minimal therapist contact. Patients were seen after referral for the purposes of training in the use of the DSQ. This typically involved two appointments. Occasional contact by telephone was also made to ensure that sleep diaries were being completed as required, but at no time was advice or treatment offered. Subjects were seen again at the end of the ten week data collection period having thus provided data for the entire duration of the experimental period (...). Waiting—list subjects were, therefore, treated on an ad hoc individualised basis.” (Espie 1989b , p. 107). **Exposure/intensity to treatment:** no treatment during waiting **Duration treatment:**: 10 weeks **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** "Of the 84 patients, 58 (69%) were stated, by their GPs, to be "drug free". Only 14 of these patients, however, had never been on hypnotic medication, and it transpired upon further assessment, that approximately one third of the remainder had not entirely discontinued medication. In most cases drug use was occasional and low dose. (…)Twenty—six patients were referred with persisting sleep difficulties who were also on nightly sleep medication. ” (Espie 1989b , p. 101-2)
Outcomes	**Beneficial effect** • **Hierachy:** patient-reported, clinical relevance, coin toss (random.org) • **Outcome chosen:** Self-report Daily Sleep Questionnaire (DSQ), subscale SOL. **Adverse events** • No data on adverse events reported.
Notes	**Key conclusion from study authors** 1. Only active treatments were associated with significant improvement, but the nature of treatment gains varied. 2. In particular, stimulus control improved sleep pattern, whereas relaxation affected perception of sleep quality. 3. All improvements were maintained at 17 month follow-up. **Key limitations from study authors** 1. Not stated **Other notes from review authors** 1. None **Conflicts of interest:** none found **Judgement:** no

Risk of bias

Item	Authors' judgement	Support for judgement

Espie 1989a *(Continued)*

Random sequence generation	Unclear	Quote: ”Ss were allocated according to a predetermined list of random numbers to either progressive relaxation (PR), stimulus control (SC), paradoxical intention (PI), imagery relief placebo (IR) or no treatment (NT).” (p. 81).
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported measures
Blinding of participants and personnel	No	Not possible to blind placebo and "no-treatment"
Incomplete outcome data	No	Quote: ”The 84 included subjects represented 60% of the 141 referrals received. (...) In summary, half of the subject loss was due to the operation of strict selection criteria, and half due to subjects dropping out. Of the "drop-outs", however, considerably more than one third failed to attend even the first appointment. The true "drop-out" rate amongst assessed and suitable subjects was only 17 out of 101 patients. This attrition rate is at least comparable to clinical research studies in any field of application.” (Espie 1989b , p. 180) Attrition >15% (15.5%). No ITT. Only reports data on completers
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Etringer 1982

Study characteristics	
Methods	**Parallel randomised trial with three interventions** 1. Psychological placebo: graduated subliminal modelling 2. No-treatment 3. Active treatment: participant modelling **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 1 session (1 day – a maximum of 90 minutes) **Duration of participation (trial + follow-up):** 1 day + 4 week follow-up (but problems with assessment) **Setting:** outpatient (local community & college) **Purpose of trial:** “The present study compared PM to an attention-placebo/treatment element control group that had been rated as initially equally credible for the treatment of severe snake-avoidant behavior.” (p. 477) **Open/Closed placebo:** closed placebo
Data	**Number of participants screened:** 41 **Number of participants included:** 38 **Number of participants followed-up at post treatment:** not stated

Etringer 1982 *(Continued)*

Number of participants randomly assigned to:

- Psychological placebo: n = 13
- No-treatment: n = 12
- Active treatment: n = 13

Number of withdrawals:: not stated

Diagnosis: Specific anxiety (chronic fear of snakes)

Diagnostic manual:: not stated

Means of assessment: not stated

Comorbidity: not stated

Age: 23.7 mean years

IQ: not stated- but 18/38 were college students

Sex: : 81.6% female

Ethnicity : not stated

Country: USA

Inclusion criteria

1. Chronic fear of snakes

Exclusion criteria

1. Not stated

Comparisons

Psychological placebo

Treatment name: Graduated subliminal modelling

Description of intervention:

"The subjects were given a theoretically neutral but highly credible placebo treatment couched in terms of modeling. The subjects were told that GSM was an effective method for getting rid of common fears. The stated rationale was that the method works by subjects' viewing tachistoscopically exposed slides of a model interacting with a snake. The model was ostensibly performing a graduated series of increasingly more threatening interaction behaviors with a fox snake. The subjects were told that by viewing these slides they would gradually come to learn that harmful consequences do not follow from interaction with the snake and that they would gradually overcome their fear. Subjects were also told that in order to optimize their progression through the interaction hierarchy, their heart rate and muscle tension would be monitored, thereby enabling the therapist to present the slides at such a pace that the subject would never become overly anxious." (p. 478)

Individual or group treatment: not stated

Exposure/intensity to treatment : max 90 minutes

Duration of treatment : 1 session (1 day)

Concomitant psychotherapy : not stated

Concomitant pharmacotherapy : not stated

Wait-list

Comparison name: wait-list

Description of intervention:

Etringer 1982 *(Continued)*

	Subjects in the NTC condition participated in all assessment procedures without receiving any intervening treatment. Following pretreatment assessment, the subjects in this condition sat in the experimental room for the appropriate time period and were urged to read popular magazines that were made available.These subjects were given PM in the supplementary treatment phase of the experiment if they so desired." (p. 479) **Exposure/intensity to treatment** : No treatment during waiting period **Duration treatment:** 1 session (1 day) **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy:** observer-reported • **Outcome chosen:**: Behavioral avoidance test (BAT) **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. Although initially equivalent across treatments, credibility increased significantly for the participant modelling group and stayed virtually the same for the placebo group. **Key limitations from study authors** 1. Not stated **Other notes from review authors** 1. None **Conflicts of interest:** none found J**udgement:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Random assignment of subjects to conditions produced the following groups: PM contained 2 males and I I females (4 community and 9 college subjects) (p. 478)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Therapists were kept blind to the results of all assessment procedures. (p. 479)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Unclear	Attrition unclear
Selective outcome reported	Unclear	Protocol not found

Etringer 1982 *(Continued)*

Other sources of bias	Yes	No other sources found

Foa 1991

Study characteristics	
Methods	**Parallel randomised trial with four arms** 1. Psychological placebo: supportive counselling 2. Wait-list 3. Active treatment 1: Stress inoculation (SIT) 4. Active treatment 2: Prolonged exposure (PE) **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 4.5 weeks **Duration of participation (trial + follow-up:** 4.5 week. No follow-up **Setting:** outpatient **Purpose of trial:** "We predicted that both PE and SIT would significantly reduce PTSD symptoms, more than would SC and WL." (p. 716) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** not stated **Number of participants included:** 55 **Number of participants followed-up at post treatment:** 45 **Number of participants randomly assigned to:** • Psychological placebo: n = 14 • Wait-list: n = 10 • Active treatment 1: n = 17 • Active treatment 2: n= 14 **Number of patients reported in full report:** • Psychological placebo: n = 11 • Wait-list: n = 10 • Active treatment 1: n = 14 • Active treatment 2: n = 10 **Number of withdrawals:** n = 10 • Psychological placebo: n = 3 • Wait-list: n = 0 • Active treatment 1: n = 3 • Active treatment 2: n = 4 **Diagnosis:** Post-traumatic stress-disorder (PTSD) **Diagnostic manual:** Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition, Revised (DSM-III-R)

Foa 1991 *(Continued)*

Means of assessment: clinical interview

Comorbidity: not stated, but several mental health diagnoses were excluded

Age: 31.8 mean years (SD = 8.2)

IQ:: not stated

Sex: 100% female

Ethnicity: Black 25%, White 72.7%, Hispanic 2.3%.

Country: USA

Inclusion criteria

1. PTSD
2. Raped at least 3 months before participation

Exclusion criteria

1. Current or previous diagnosis of organic mental disorder
2. Schizophrenia, or paranoid disorders as denned in the DSM-III-R
3. Depression severe enough to require immediate psychiatric treatment, bipolar depression, or depression accompanied by delusions, hallucinations, or bizarre behavior
4. Current alcohol or drug abuse
5. Assault by spouse or other family member
6. Lliteracy in English

Comparisons

Psychological placebo

Treatment name:: Supportive counselling

Description of intervention: "Supportive counseling followed the nine-session format, gathering information nine-session format, gathering information through the initial interview in the first session and presenting the rationale for treatment in the second session. During the remaining sessions, patients were taught a general problem-solving technique. Therapists played an indirect and unconditionally supportive role. Homework consisted of the patient's keeping a diary of daily problems and her attempts at problem solving. Patients were immediately redirected to focus on current daily problems if discussions of the assault occurred. No instructions for exposure or anxiety management were included." (p. 718)

Individual or group treatment: Individual

Exposure/intensity to treatment: nine biweekly 90-minute sessions

Duration of treatment: 4.5 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

Wait-list

Comparison name: wait-list (WL)

Description of intervention: "WL subjects were informed that they would receive treatment in 5 weeks. During this period, they were contacted by a therapist between assessments to determine whether emergency services were required. Following an assessment at the end of the waitlist period, patients were randomly assigned to either PE or SIT." (p. 718)

Exposure/intensity to treatment: no treatment during waiting

Duration treatment: 5 weeks

Foa 1991 *(Continued)*

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

Outcomes	**Beneficial effect** • **Hierarchy:** observer-reported, clinical relevance, global score • **Outcome chosen:** PTSD severity **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. All conditions produced improvement on all measures immediately post-treatment and at follow-up 2. However, SIT produced significantly more improvement on PTSD symptoms than did SC and WL immediately following treatment 3. At follow-up, PE produced superior outcome on PTSD symptoms 4. The implications of these findings and direction for treatment and future research are discussed. **Key limitations from study authors** 1. First, the use of only female therapists in the study limits its generalisability. However, this issue may not pose a serious limitation because most rape victims' treatment centres employ primarily women as therapists. 2. More important, the fact that the principal authors provided training and supervision in all of the treatments may have introduced experimental bias effects. Also, it is difficult to assess the impact of the fact that therapists conducted therapies that may have been contrary to their preferences. **Other notes from review authors** 1. Only reports data on 45 patients **Conflicts of interest:** non-financial conflict of interest: The first author (Foa EB) is the developer of the experimental intervention (prolonged exposure) **Judgement:** no

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: " After 10 patients were entered into the wait-list condition, subsequent admissions were randomly assigned to one of the three treatment groups" (p. 716)
Allocation concealment	Unclear	No information.
Blinding of outcome assessors	Yes	Assessments at pretreatment, posttreatment, and follow-up consisted of clinical interviews conducted by an independent assessor, who was blinded to treatment conditions, and self-report questionnaires. (p. 717)
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	No	Quote: "Dropout rates were not significantly different across the treatment groups, x^2 (3, N= 55) = 3.34, p > .30, and were as follows: PE 28.6%, SIT 17.6%, SC 21.4%, and WL 0%. (...) Subsequent analyses were conducted on data from the 45 completers." (p. 718)

Foa 1991 *(Continued)*

		Attrition >15% (PE 28.6%, SIT 17.6%, SC 21.4%, and WL 0%). No ITT. Only reports data on completers
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Foa 2018

Study characteristics	
Methods	**Parallel randomised trial with four arms** 1. Psychological placebo: person-centered therapy 2. Waitlist: minimal contact control 3. Active treatment 1: massed prolonged exposure therapy 4. Active treatment 2: paced prolonged exposure therapy **Sample calculation:** yes **Cluster randomised:** no **Duration of trial (baseline to post):** 2 weeks (available post-treatment data) **Duration of participation (trial + follow-up**): 8 weeks + 2 weeks follow-up. There are also 12 weeks and 6 months follow-up but no data on the minimal contact control group. **Setting:** outpatient **Purpose of trial:** to examine the effects of massed prolonged exposure therapy (massed therapy), spaced prolonged exposure therapy (spaced therapy), present-centered therapy (PCT), and a minimal-contact control (MCC) on PTSD severity **Open/closed placebo** : Closed placebo
Data	**Number of participants screened:** 526 **Number of participants included** : 370 **Number of participants followed-up at post treatment:** 245 **Number of participants randomly assigned to:** • Psychological placebo: n =110 • Waitlist: n = 40 • Active treatment 1: n =110 • Active treatment 2: n =110 **Number of withdrawals before treatment:** n = 55 • Psychological placebo: n =13 • Waitlist: n = 0 • Active treatment 1: n =15 • Active treatment 2: n = 27 **Number of withdrawals post treatment:** n = 70 • Psychological placebo: n = 22

Foa 2018 *(Continued)*

- Waitlist: n = 0
- Active treatment 1: n =17
- Active treatment 2: n =31

Diagnosis: Post-traumatic stress-disorder (PTSD)

Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision (DSM-IV-TR)

Means of assessment: clinical interview (not otherwise stated)

Comorbidity: depressive symptoms

Age:: mean psychological placebo: 32.54 years (SD = 7.45), wait-list: 32.70 years (SD = 7.68)

IQ:: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition

Sex: 11.6% female

Ethnicity: 32 Hispanic, 115 non-Hispanic. 0 Asians, 29 Blacks, 95 whites, 23 other

Country: USA

Inclusion criteria

1. active duty military, activated Reservist, activated National Guard, or veterans who had deployed to Operation Enduring Freedom/Operation Iraqi Freedom/ Operation New Dawnages
2. 18 to 65 years
3. PTSD diagnosis according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR).

Exclusion criteria

1. Current bipolar or psychotic disorders
2. Alcohol dependence
3. Moderate to severe traumatic brain injury
4. Suicidal ideation
5. Other disorders warranting immediate attention

Comparisons

Psychological placebo

Treatment name: Present-centered therapy

Description of intervention: "Present-centered therapy is a non–trauma-focused, manualised treatment that controls for nonspecific therapeutic factors Ten 90-minute sessions were scheduled similarly to spaced therapy and focused on current life problems that may or may not be PTSD-related. Therapists helped participants identify stressors and discussed them in a supportive, nondirective manner." (p. 356)

Individual or group treatment: not stated

Exposure/intensity to treatment: 10 sessions over 8 weeks for full treatment

Duration of treatment: 2 weeks (data used). Full treatment was 8 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: 47.3% received 1 psychotropic medication or more

Wait-list

Comparison name: Minimal contact control (wait-list)

Foa 2018 *(Continued)*

	Description of intervention: "The MCC condition consisted of 10- to 15-minute therapist telephone calls once weekly for 4 weeks. Participants were asked about their well-being, offered support as needed, and received contact information in case symptoms worsened." (p. 356). "After the 2-week follow-up, participants in the MCC group were offered their choice of the other treatments" (p. 355) **Exposure/intensity to treatment:** minimal contact for 2 weeks. (post-treatment) **Duration treatment:** 2 weeks + 2 weeks follow-up **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** 30% received 1 psychotropic medication or more
Outcomes	**Beneficial effect** • **Hierarchy:** primary outcome, observer-reported • **Outcome chosen** : PTSD Symptom Scale–Interview (PSS-I) **Adverse events** • Count data/spontaneous reporting of serious and non-serious.
Notes	**Key conclusion from study authors** 1. Among active duty military personnel with PTSD, massed therapy (10 sessions over 2 weeks) reduced PTSD symptom severity more than MCC at 2-week follow-up and was non inferior to spaced therapy (10 sessions over 8weeks), 2. There was no significant difference between spaced therapy and PCT. 3. The reductions in PTSD symptom severity with all treatments were relatively modest, suggesting that further research is needed to determine the clinical importance of these findings **Key limitations from study authors** 1. First, the design did not include an active 2-week comparison treatment for massed therapy. However, because this was the first study to evaluate intensive prolonged exposure therapy in military personnel, the use of an MCC condition was required by the Department of Defense external advisory board and supported by the respective institutional review boards. 2. Second, participants in the massed therapy group may have lacked time to sufficiently practice homework assignments. 3. Third, because participants were treatment seeking, the results are limited to military personnel seeking treatment for PTSD. 4. Fourth, the dropout rate during treatment rangedfrom12.1% (PCT) to 24.8% (spaced therapy), and only 59% of randomised participants completed the full study. Treatment effects likely would have been larger if a greater proportion of participants had completed the treatment portion of the study. **Other notes from review authors** 1. None **Conflicts of interest:** non-financial conflict of interest: the first author (Foa EB) is the developer of the experimental intervention (prolonged exposure) J**udgment:** no

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: "Randomization was originally planned as 3:11:11:11 for MCC: massed therapy: spaced therapy: PCT. On January 5, 2012, enrolment in MCC was accelerated by changing the ratio to 1:1:1:1 to allow for preliminary massed therapy vs MCC comparison per Department of Defense request. After 40 partici-

Foa 2018 *(Continued)*

		pants were randomized to receive MCC, randomization to MCC was discontinued on March 19, 2014, and subsequent participants were assigned. To receive massed therapy, spaced therapy, or PCT(1:1:1).” (p. 355) Randomization pattern was dummy coded and then added as a moderator to the analyses. (p. 355)
Allocation concealment	Unclear	Quote: “The randomization sequence was entered by a study statistician into a secure, web-based application using SAS version 9.4 (SAS Institute Inc), which was accessed by the project coordinator on enrollment of each participant.” (p. 355). Unclear whether project coordinator could have influenced allocation concealment.
Blinding of outcome assessors	Yes	Quote: “PTSD symptom severity was assessed by independent evaluators blinded to treatment condition, before and after treatment, and at 2-week, 12-week, and 6-month follow-up.” (p. 355)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Unclear	Linear mixed models and generalised linear mixed models were used to analyse the data, using SPSS version 23 (IBM SPSS). These models are intent-to-treat and calculate results based on available data without imputation of missing data (p. 356) Attrition >15% (Active treatment 41% - but only 12.1% in placebo and 0% in wait-list). No ITT
Selective outcome reported	No	Trial registry: NCT01049516. Veterans RAND 12-items HealthSurvey and Adverse events not mentioned in Clinical Trial Registry.
Other sources of bias	Yes	No other sources found

Freire 2007

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Physical placebo: sham acupuncture 2. Wait-list 3. Active treatment: acupuncture **Sample calculation:** yes **Cluster randomised:** no **Duration of trial (baseline to post):** 10 weeks **Duration of participation (trial + follow-up):** 10 + 2 weeks = 12 weeks post-treatment **Setting:** outpatient **Purpose of trial:** “To investigate the efficacy of acupuncture in the treatment of moderate obstructive sleep apnea syndrome (OSAS), assessed by polysomnography (PSG) and questionnaires of functional quality of life (SF-36) and excessive daytime sleepiness (Epworth)” (p. 43) **Open or closed placebo:** closed placebo
Data	**Number of participants screened:** 38

Freire 2007 *(Continued)*

Number of participants included: 36
Number of participants followed-up at post treatment: 26

Number of participants randomly assigned to:

- Physical placebo: n = 12
- Wait-list: n = 12
- Active treatment: n = 12

Number of withdrawals: n = 10

- Physical placebo: n = 5
- Wait-list: n = 3
- Active treatment: n = 2

Diagnosis: obstructive sleep apnoea (sleep-wake disorder)
Diagnostic manual:: not stated
Means of assessment: diagnosis "confirmed by a full polysomnographic (PSG) study with an apnea/hypopnea index (AHI) >15/hour and <30/hour (moderate OSAS" (p. 44) + clinical interview
Comorbidity: not stated
Age: 54.67 mean years, (Range = 49 to 54)
IQ :not stated – but participants with intellectual deficits were not eligible
Sex: 55.6% female
Ethnicity: not stated
Country: Brazil

Inclusion criteria

1. A diagnosis of moderate OSAS was eligible.

Exclusion criteria

1. Patients with a high alcohol intake (> 80 g/day)
2. Morbid obesity
3. Significant lung disease
4. Neurological disease
5. Intellectual deficits
6. Skeletal facial framework problems
7. Central apnoea
8. Patients who were taking any hypnotic drugs
9. Patients who had undergone oropharyngeal surgery
10. Patients who had been treated with CPAP
11. Patients with oral devices were excluded.

Comparisons

Physical placebo
Treatment name: sham acupuncture
Description of intervention: "The sham acupuncture group was stimulated with the same number of needles as the acupuncture group, and the points were localized 1 cun from the real point, in a region not related to any acupoints or meridians and was done following the standards of minimal acupuncture. For the sham acupuncture group, the needles were inserted and no manipulation was done. (…) All acupuncture procedures, as well as sham acupuncture, were performed by an experienced physician, who was a specialist in acupuncture, according to traditional Chinese acupuncture methods. All procedures were performed in the afternoon between 3 p.m. and 6 p.m. Body needles were left in situ for 30 min." (p. 45) "Finally, patients were informed that at the end of the study all patients allocated to the sham acupuncture group would receive 10 sessions of acupuncture treatment if they so wanted." (p. 44)
Individual or group treatment: individual
Exposure/intensity to treatment: once a week for 10 weeks
Duration of treatment: 10 weeks
Concomitant psychotherapy:: not stated

Freire 2007 *(Continued)*

Concomitant pharmacotherapy: not stated – but excluded if taking hypnotic drugs

Wait-list
Comparison name: control group (wait-list)
Description of intervention: "Patients assigned to the control group were offered weight reduction advice if overweight and sleep hygiene counseling. Given the usual waiting list for nCPAP at our service, which is about 6 months, the waiting time was not a matter of ethical concern." (p. 44)
Exposure/intensity to treatment: nothing during waiting
Duration treatment: 3 months (12 weeks)
Concomitant psychotherapy: not stated
Concomitant pharmacotherapy: not stated – but excluded if taking hypnotic drugs

Outcomes

Beneficial outcomes :

- **Hierarchy:** observer-reported, clinical relevance, global score
- **Outcome chosen:** apnoea-hypopnoea index

Adverse events

- "No adverse events occurred during the trial." (p. 45)

Notes

Key conclusion

1. Twenty-six patients completed the study.
2. The AHI (P = 0.005), the apnoea index (AI) (P = 0.008) and the number of respiratory events (P = 0.005) decreased significantly in the acupuncture group but not in the sham group.
3. On the other hand, the control group displayed significant deterioration in some of the polysomnographic parameters, with a significant increase in the number of respiratory events (P = 0.025).
4. Acupuncture treatment significantly improved (before vs. after treatment) several dimensions of the SF-36 and Epworth questionnaires. There was no significant association between changes in the body mass index (BMI) and AHI.
5. Conclusions: Acupuncture is more effective than sham acupuncture in ameliorating the respiratory events of patients presenting with moderate OSAS

Key limitations

1. Although this protocol did not include oesophageal balloon, the gold standard method to detect the respiratory effort, we measured the airflow through the nasal cannula. This measurement provides information about airflow limitation [23], which has been correlated with elevated upper airway resistance and increased oesophageal pressure. Since, the PSG scorers did not find any flattened inspiratory airflow associated with micro arousals or paradoxal breathing, we could rule out upper airway resistance syndrome (UARS).

Other notes from review authors

1. Follow-up data is post-treatment data

Conflicts of interest: none

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: "Randomisation was done by a blinded independent researcher and was conducted by selecting a closed piece of paper out of a box, with a treatment order written on it. (p. 44)

Freire 2007 *(Continued)*

Allocation concealment	Yes	Only the physician applying the treatments (acupuncture/ sham acupuncture) was aware of which group each patient had been assigned to and did not participate in any phase of the subsequent evaluation." (p. 44)
Blinding of outcome assessors	Yes	Quote:"All of the PSG recordings were assessed by two experienced sleep physicians (S.M. Togeiro and F.S.Chrispin), who were blind to the groups to which the patients had been assigned." (p. 46)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	No	Attrition >15% (27.8%). No ITT. Only reports data on completers
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Fuchs 1977

Study characteristics

Methods

Parallel randomised trial with three arms

1. Psychological placebo: non-specific
2. Wait-list
3. Active treatment: self-control

Sample calculation: not stated

Cluster randomised: no

Duration of trial (baseline to post): 6 weeks

Duration of participation (trial + follow-up): 6 weeks (Post-treatment assessment at 7 weeks) + 6 weeks follow-up (WL not assessed at follow-up)

Setting: outpatient

Purpose of trial: "A behavior therapy program based on a self-control model of depression was evaluated against a nonspecific group therapy condition and a waiting list control group." (p. 206)

Open/closed placebo : Closed placebo

Data

Number of participants screened: not stated

Number of participants included: 36

Number of participants followed-up at post treatment: 28

N **umber of participants randomly assigned to:**

- Psychological placebo: n= 10
- Wait-list: n = 10
- Active treatment: n = 8

Number of withdrawals : n = 8

- Psychological placebo: not stated

Fuchs 1977 *(Continued)*

- Wait-list: not stated
- Active treatment: not stated

Diagnosis: depression

Diagnostic manual: not stated

Means of assessment: Minnesota Multiphasic Personality Inventory (MMPI)

Comorbidity: not psychotic, suicidal, no history of psychiatric hospitalisation

Age: 28.8 years (range = 18 to 48)

IQ: not stated

Sex: 100% female

Ethnicity: not stated

Country: USA

Inclusion criteria

1. On the Minnesota Multiphasic Personality Inventory (MMPI), scores met these criteria: F <, 80, L <. 60, D ^ 70, D > Hy, and D > Pt), and D was among the highest two elevations on the profile,
2. Screening questionnaire and interview responses revealed no history of psychiatric hospitalisation, serious suicidal ideation or attempts, and no involvement in any other therapy for problems related to psychological functioning within the past month,
3. Clinical judgment, based on MMPI profile and interview data, was that the clients were not psychotic or suicidal.

Exclusion criteria

1. Not stated

Comparisons	**Psychological placebo** **Treatment name (type):** nonspecific therapy **Description of intervention:** "Session 1 began in the same way as the self-control procedure with introductions, collection of deposits, a review of confidentiality issues, and a 10-minute group interaction assessment procedure. As in the other groups, subjects were given an information sheet and a general introduction to group therapy concepts, generally from a nondirective framework. From that point on and throughout the ensuing sessions, therapists in this condition attempted to elicit discussion of past and current problems, to encourage group interaction, and to reflect and clarify feelings in an empathic manner. Although therapists at times suggested simple exercises within the group to facilitate open discussion, they were specifically instructed neither to recommend out-of-therapy activity nor explicitly to teach behavioral principles. These sessions lasted approximately 2 hours weekly, as did self-control therapy sessions." (p. 209) **Individual or group treatment:** group **Exposure/intensity to treatment:** 2 hours weekly **Duration of treatment:**:6 weeks **Concomitant psychotherapy:** no involvement in any other therapy for problems related to psychological functioning within the past month **Concomitant pharmacotherapy:** not stated **Wait-list** **Comparison name:** wait-list control

Fuchs 1977 *(Continued)*

	Description of intervention : "Subjects in this condition were informed by phone that they had been accepted into the research program but that our present groups were filled, so that they would have to wait about 8 weeks before their groups would start. They were also told that they would be required to retake some of the screening tests just prior to beginning therapy; however, they were assured of being seen." (p. 209) **Exposure/intensity to treatment:** no treatment during waiting **Duration treatment:** 6 weeks **Concomitant psychotherapy:** no involvement in any other therapy for problems related to psychological functioning within the past month **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy:** patient-reported, clinical relevance • **Outcome chosen** : the Beck Depression Inventory **Adverse events** • No data on adverse events reported
Notes	**Key conclusion from study authors** 1. Self-control therapy patients showed significantly greater reduction in depression on self-report and behavioural measures 2. Self-control patients also showed greater improvement in overall pathology on the Minnesota Multiphasic Personality Inventory. **Key limitations from study authors** 1. Population limited to females 2. Not able to isolate specific effects **Other notes from review authors** 1. None **Conflicts of interest: n**on-financial conflict of interest: Authors developed the experimental intervention **Judgement:** no

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Except where necessary to balance experimental conditions for mean age and severity of depression, subjects were randomly assigned to one of two therapists and one of three treatment conditions—self-control therapy, nonspecific therapy, or waiting list control." (p. 209)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list

Fuchs 1977 *(Continued)*

Incomplete outcome data	No	Quote: "Eight of the original 36 subjects dropped out of the study, all within the first 2 weeks. (...) Drop-out rate did not differ significantly between conditions, x2 (2) = .29, p < .80. Dropouts did not differ from remainders on age, Depression Inventory, MMPl D, or MMPI total elevation scores." (p. 210) Attrition >15% (22.2%). No ITT. Only reports data on completers
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Glogowska 2000

Study characteristics	
Methods	**Parallel randomised trial with two arms** 1. Usual care: therapy 2. Wait-list: watchful waiting **Sample calculation:** yes **Cluster randomised:** no **Duration of trial (baseline to post):** mean treatment 8,4 months (0,9 to 12 months). post treatment data at 12 months **Duration of participation (trial + follow-up):**: 12 months - no follow-up **Setting:** outpatient **Purpose of trial:** "To compare routine speech and language therapy in preschool children with delayed speech and language against 12 months of "watchful waiting". (p. 1)
Data	**Number of participants screened:** 507 **Number of participants included:** 159 **Number of participants followed-up at post treatment:** 155 **Number of participants randomly assigned to:** • Usual care: n = 71 • Wait-list: n = 88 **Number of withdrawals:** n= 4 • Usual care: n = 4 • Wait-list: n = 0 **Diagnosis:** delayed speech and language **Diagnostic manual:** not stated **Means of assessment:** not stated **Comorbidity:** 13 diagnosed with hearing loss **Age:** age in months, usual care 34.2 months (range = 18 to 42), wait-list 34.2 months (range = 24 to 42)

Glogowska 2000 *(Continued)*

IQ: not stated

Sex: usual care: 23% female. wait-list: 26% female

Ethnicity: not stated

Country: UK

Inclusion criteria

1. General selection criteria: Newly referred singleton children acquiring English, in a monolingual home. Aged under 3 ½ years at initial attendance for speech and language therapy assessment. No diagnosis of severe learning difficulties or autism. No oromotor deficits. No primary diagnosis of dysfluency (stammering) or dysphonia (voice disorders). No siblings currently receiving speech and language therapy. Children had to satisfy on of the clinical criteria. Be considered to have significant clinical difficulties by the speech and language therapist. A "carer" had to attend sessions. Parents had to give consent.
2. Clinical criteria: general language group: a standardised score < 1.2 SD (standard deviation) below the mean on the auditory comprehension part of the preschool language scale. Expressive language group: a standardised score >1.2 SD below the mean on auditory comprehension but <1.2 SD below the mean on the expressive language part of the preschool language scale. Phonology group: auditory comprehension and expressive language scores >1.2 SD below the mean but with an error rate of at least 40% in the production of fricative consonant (for example, f and s) and/or velar consonants (for example, "hard" c, "hard" g, and ng) and/or sounds occurring after a vowel among the 22 words included in the phonological analysis.

Exclusion criteria

1. Not stated

Comparisons

Usual care

Treatment name: therapy

Description of intervention: "Therapy provided in the study tended to focus on several areas of language simultaneously. Therapy techniques included Derbyshire language scheme tasks, as well as everyday play and games used as contexts for modelling language for the child. Goals covered a wide range of language stages – for example, understanding and building single words, using narratives, and identifying consonants in words." (p. 4)

Individual or group treatment: individual. "Children randomised to the therapy group received the one-to-one speech and language therapy (…)" (p. 2)

Exposure/intensity to treatment : 6.2 (0-15) hours of therapy. Frequency of therapy were once a month (range once a week to once every two and a half months)

Duration of treatment: 8.4 months (range = 0.9-12) - number of months over which the therapy took place

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

Wait-list

Comparison name: watchful waiting

Description of intervention: "Parents of children in the watchful waiting group could request therapy at any time if they were concerned about their child's progress. All children in the study were reassessed by the research therapists after 12 months; if a child in the watchful waiting group were still experiencing difficulties, two research therapists (SR and MG) provided up to 12 therapy sessions." (p. 2)

Exposure/intensity to treatment: no treatment

Glogowska 2000 *(Continued)*

Duration treatment: 12 months (after all received treatment and assessments)

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect

- **Outcome hierarchy: p**rimary outcome, observer-reported, continuous outcome, clinical relevance
- **Outcome chosen:** Bristol language development scales

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. "Improvement in the therapy group was significant (compared with the watchful waiting group) for only one of the five primary outcomes – auditory comprehension." (p. 4)
2. "Most children in this study still had important clinical difficulties at 12 months, regardless of trial allocation; indeed, many remained eligible for the trial, with little evidence of "spontaneous resolution." This study provides little evidence for the effectiveness of speech and language therapy when compared with "watchful waiting" over 12 months." (p. 5)

Key limitations from study authors

1. "Overall, the impacts of therapy in this trail was small, perhaps because of the relatively low level of therapy provided – considerably lower than levels reported in previous studies." (p. 5)
2. Although the children were stratified according to their broad entry criteria, which ensures similar groups in this respect, the sample size of the clinical groupings was too small to detect significant differential effects." (p. 5)
3. Blinding was maintained for all baseline assessments and for the language sample at follow-up. Although every effort was made to retain blinding at the follow up assessments, in the presence of parents strict blinding was inevitably not always feasible for the other outcomes." (p. 5)

Other notes from review authors

1. 12 months follow-up is the end of the intervention (post-treatment data)

Conflicts of interest: none found

Judgement: no

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Randomisation was stratified by the 16 clinics and by the three clinical criteria (general language, expressive language, and phonology) (…) The sequence of random numbers was generated before the trial independently of the therapists." (p. 2)
Allocation concealment	Yes	Quote: "The allocation was implemented by the therapists opening sealed opaque envelopes (coloured according to the three clinical criteria) in the presence of the parents." (p. 2)
Blinding of outcome assessors	Yes	Quote: "Assessors were blind to previous results, and every attempt was made to maintain blindness in terms of allocation. The presence of the parent meant that this was often inevitably compromised, but each child was seen by a different therapist for the two follow ups, and the language sample for the Bristol language development scales was analysed in a fully blinded manner." (p. 2)

Glogowska 2000 *(Continued)*

Blinding of participants and personnel	No	Not possible to blind usual care and wait-list
Incomplete outcome data	Yes	Quote: "The trial arms were compared on an "intention to treat" basis." (p. 2) "*Data were missing for all measures in both groups: analyses were based on 64 (therapy group) and 80 children (watchful waiting group) for auditory comprehension; 63 and 77 for expressive language; 57 and 62 for the phonology error rate; and 71 and 84 for improvement by 12 months." (p. 4) Attrition <15% (2.5%). ITT used
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Goldstein 2000

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Psychological placebo: attention placebo 2. Wait-list 3. Active treatment: EMDR **Sample calculation:** yes **Cluster randomised:** no **Duration of trial (baseline to post):** 4 weeks **Duration of participation (trial + follow-up):** 4 weeks + 1 month follow-up **Setting** : Outpatient **Purpose of trial:** "Accordingly, the purposes of the present study were twofold: (a) to conduct a replication of Feske and Goldstein's comparison of EMDR to a waiting list control group for PDA and (b) to contrast EMDR with a credible attention-placebo." (p. 948) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** not stated **Number of participants included:** 46 **Number of participants followed-up at post treatment:** 45 **Number of participants randomly assigned to:** • Psychological placebo: n = 13 • Wait-list: n = 14 • Active treatment: n = 18 **Number of withdrawals:** n = 1 • Psychological placebo: n = 0 • Wait-list: n = 1

Goldstein 2000 *(Continued)*

- Active treatment: n = 0

Diagnosis: agoraphobia

Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV)

Means of assessment: the Structured Clinical Interview for DSM-IV Disorders (SCID)

Comorbidity: 20 participants had at least one comorbid Axis I diagnosis: specific phobia (7), generalised anxiety disorder (6), social phobia (5), or obsessive—compulsive disorder (2). Of these, 5 had more than one Axis I comorbid condition. Three participants met criteria for obsessive— compulsive personality disorder, and 4 for avoidant personality disorder

Age: 38.16 mean years, (range =22 to 63)

IQ:: not stated, but 38 had attended at least some college

Sex: 80.4% female

Ethnicity: two were African American, and one was Asian American; the remainder were European American

Country: USA

Inclusion criteria

1. Agoraphobia according to Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV)

Exclusion criteria

1. age less than 18 or greater than 65 and being in therapy elsewhere if not willing to suspend that treatment until the end of the study.
2. Potential participants on dosages of alprazolam in excess of 1.5 mg daily (or similar dosages for other benzodiazepines) were excluded, as were those who had been taking antidepressant or antianxiety medication for less than 6 months or who had changed their medication within the last 12 weeks
3. Potential participants were also excluded if they had comorbid diagnoses of thought disorder, major depression (n = 5), bipolar disorder, or substance dependence (n = 1); if another anxiety disorder was more severe than the PDA (n = 3); or if they met full criteria for any of the following Axis II disorders: paranoid (n = 1), schizoid, schizotypal, antisocial, or borderline (n = 3).

Comparisons

Psychological placebo

Treatment name: attention placebo

Description of intervention: “ART, the attention-placebo treatment, included a combination of two relatively inert treatment procedures: 30—45 min of progressive muscle relaxation training and 45 60 min of association therapy” (p. 952)

Individual or group treatment: Individual

Exposure/intensity to treatment: six 90-minute sessions held over an average of 4 weeks

Duration of treatment: 4 weeks

Concomitant psychotherapy: no other treatment. Excluded if they had

Concomitant pharmacotherapy: excluded if taking medication. “Participants excluded on the basis of recent medication changes were eligible for reconsideration once medications were stabilized in appropriate limits.” (p. 950)

Wait-list

Comparison name: wait-list

Goldstein 2000 *(Continued)*

Description of intervention: "For 2 weeks prior to and after treatment or waiting list, as well as throughout the course of treatment or waiting period, participants completed anxiety forms every morning and evening and at the close of each week." (p. 950)

Exposure/intensity to treatment: waiting for treatment

Duration treatment: 4 weeks. "Once the waiting list period ended, all those assigned to waiting list were randomized to EMDR (n = 6) or attention-placebo (n = 7)." (p. 949)

Concomitant psychotherapy: no other treatment. Excluded if they had

Concomitant pharmacotherapy: excluded if taking medication. "Participants excluded on the basis of recent medication changes were eligible for reconsideration once medications were stabilized in appropriate limits." (p. 949)

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported
- **Outcome chosen:** Panic Disorder Severity Scale

Adverse events

- Count data/spontaneous reporting of serious and non-serious.

Notes

Key conclusion from study authors

1. EMDR was significantly better than waiting list for some outcome measures (questionnaire, diary, and interview measures of severity of anxiety, panic disorder, and agoraphobia) but not for others (panic attack frequency and anxious cognitions.
2. Differences between EMDR and the attention-placebo control condition were not statistically significant on any measure, and, in this case, the effect sizes were generally small (η = 2 = .00 to.06), suggesting the poor results for EMDR were not due to lack of power.

Key limitations from study authors

1. However, low power and, for panic frequency, floor effects may account for these negative results
2. The effect sizes were generally small (η = 2 = .00 to.06), suggesting the poor results for EMDR were not due to lack of power.

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: no

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Participants were initially randomly assigned to one of three groups: waiting list (n = 14), EMDR (n = 18), or an attention-placebo condition (n = 13) involving the same amount of therapist contact as EMDR (p. 3)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Quote: "Raters were not blind to group assignment." (950)

Goldstein 2000 *(Continued)*

Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	Yes	Quote: "Dropouts were replaced with the next participant to enter the study" (p. 949) "Intention-to-treat analyses were conducted at each assessment period by repeating ANOVAs and ANCOVAs with pretest scores carried forward to serve as posttest or follow-up scores for those who failed to provide posttest data or who dropped out before the conclusion of treatment or before the follow-up assessment. The findings of the EMDR versus waiting list and EMDR versus ART comparisons were unchanged." (p. 955) Quote: "Fisher;s exact tests indicated that attrition was not significantly different across groups (EMDR vs. attention-placebo p= .242; EMDR vs. waiting list p= 1.00). Of the 42 participants who completed treatment, 37 provided follow-up data. Of those who dropped from follow-up after EMDR, one required medical attention for an unrelated condition, one terminated because of increased distress during treatment, and a third refused assessment without explanation. Of those who dropped from attention- placebo, one dropped because of his disappointment with treatment and the other without explanation." (950) Attrition > 15% (19.6%). ITT used
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Goldwasser 1987

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Psychological placebo: support 2. No-treatment 3. Active treatment: reminiscence **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 5 weeks **Duration of participation (trial + follow-up):** 5 weeks + 6 weeks follow-up **Setting:** inpatient (resident population at Beth Sholom Home in Richmond, Virginia) **Purpose of trial:** "This article presents a controlled study designed to determine the degree to which reminiscence group therapy influences affective, cognitive, and behavioral functioning in demented elderly." (p. 210) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** 30 **Number of participants included:** 27

Goldwasser 1987 *(Continued)*

Number of participants followed-up at post treatment: 24

Number of participants randomly assigned to:

- Psychological placebo: n = 9
- No-treatment: n = 9
- Active treatment: n = 9

Number of withdrawals : n = 3

- Psychological placebo: n = 1
- No-treatment: n = 1
- Active treatment: n = 1

Diagnosis: dementia

Diagnostic manual: not stated

Means of assessment: not stated

Comorbidity: Alzheimers's multi-infarct, dementia secondary to medical disorder

Age: 82.3 mean years (range = 70 to 97)

IQ: not stated

Sex: 74.1% female

Ethnicity: not stated

Country: USA

Inclusion criteria

1. Clinical diagnosis of dementia,
2. Presence of symptoms associated with dementia (i.e. confusion, disorientation, cognitive dysfunction, etc.)
3. The ability to communicate verbally,
4. Tthe ability to function within a group without causing excessive disruption.

Exclusion criteria

1. Not stated

Comparisons

Psychological placebo

Treatment name: attention-placebo support group

Description of intervention: "A second group consisted of a support group that focused on present or future events and problems. This group also met for a half hour twice weekly for a period of five weeks." (p. 212); "In order to ensure that the reminiscence component of the intervention accounted for any observed changes, an attention-placebo "support" group and a "no-treatment" control group were also used." (p. 210)

Individual or group treatment:: group

Exposure/intensity to treatment: 30 minutes twice weekly

Duration of treatment: 5 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

No-treatment

Goldwasser 1987 *(Continued)*

Comparison name: no-treatment

Description of intervention: The third group served as a "no-treatment" control group, and consequently did not participate in any group activity during the same period of time." (p. 212)

Exposure/intensity to treatment: no treatment

Duration treatment: 5 weeks

Concomitant psychotherapy:: not stated

Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported, clinical relevance
- **Outcome chosen:** Mini-Mental State (MMS)

Adverse events

- No data reported on adverse events. However, it is mentioned that one patient died during the trial.

Notes

Key conclusion from study authors

1. The self-reported level of depression in participants given reminiscence therapy was positively affected compared to participants in the supportive therapy and control groups, but no significant effects were found for cognitive or behavioral functioning

Key limitations from study authors

1. The question may be raised as to whether the less impaired individuals can truly be considered to be demented. Although they were clearly confused and their MMS scores generally fell at or below the criterion level for dementia, their confusion may not have been primarily due to organicity, but rather to factors such as medications, environmental factors, or depression.

Other notes from review authors

1. No usable data

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "(...) randomly assigned to three groups of ten people each. (p. 210)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Unclear	No information
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Unclear	Quote: "Since one participant in the reminiscence group died during the course of the study, one participant from each of the other treatment groups was randomly dropped from data analyses" (p. 210)

Goldwasser 1987 *(Continued)*

		Attrition <15% (11.1%).
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Hekmat 1984

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Psychological placebo: attention placebo 2. No-treatment: wait-list 3. Active treatment: semantic desensitisation **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 3 sessions in total **Duration of participation (trial + follow-up):** follow up occurred within 4 weeks. (p. 465) **Setting**: outpatient (college) **Purpose of trial:** "This study explored the clinical effectiveness of semantic desensitization in the treatment of public speaking anxiety. (p. 463) **Closed/open placebo** : Closed placebo
Data	**Number of participants screened:** 239 **Number of participants included:** 30 **Number of participants followed-up at post treatment:** not stated **Number of participants randomly assigned:** not stated **Number of withdrawals:** not stated **Diagnosis:** social anxiety (speech-anxious students) **Diagnostic manual:** not stated **Means of assessment:** speech anxiety was measured by the following scales: Personal report of confidence as a speaker (PRCS), Affect Adjective Checklist (ACL), S-R Inventory of Anxiousness, and Timed Behavior Checklist (BCL) **Comorbidity:** not stated **Age:** not stated **IQ:** not stated - but university students **Sex:**: 60% female **Ethnicity:**: not stated **Country:** USA

Hekmat 1984 *(Continued)*

Inclusion criteria

1. Public speaking anxiety

Exclusion criteria

1. Not stated

Comparisons	**Psychological placebo** **Treatment name:** attention placebo **Description of intervention:** "Ss in this group were informed that they would receive a novel therapy called "systematic ventilization." Ss were instructed that awareness of anxiety and the ways in which it would manifest itself in behavior is essential for cure." (p. 464) **Individual or group treatment:** individual **Exposure/intensity to treatment:** 3 sessions **Duration of treatment:** not stated **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated **No-treatment** **Comparison name:** waiting list control. (in reality no-treatment) **Description of intervention:** "Ss in the no treatment waiting list control were instructed that periodic measurement of their anxiety reaction was essential to procure a reliable assessment of their problem. The no treatment waiting list control Ss also were given the pretreatment, posttreatment, and follow-up anxiety measures." (p. 464) **Exposure/intensity to treatment:** not stated **Duration treatment:** not stated **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy:**: observer-reported • **Outcome chosen:** Timed Behavior Checklist (BCL) **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. Semantic desensitisation therapy resulted in significant reductions of both the affective and behavioral components of anxiety as compared to the two controls. 2. The placebo control also showed Improvement in several indices of subjective anxiety as compared to the no-treatment waiting-list control. 3. The beneficial effects of semantic desensitisation therapy were maintained on follow-up. **Key limitations from study authors** 1. Not stated **Other notes from review authors**

Hekmat 1984 *(Continued)*

1. Usable data not available

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Ss were volunteers who were matched on the basis of their pretreatment anxiety scores and randomly assigned to one of the following treatments: Group I, semantic desensitization therapy; Group 11, placebo control; and Group 111, waiting list control." (p. 463)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote: ”The two behavioral assessors were blind to treatment assignments.” (p. 464)
Blinding of participants and personnel	No	Not possible to blind placebo and "wait-list"
Incomplete outcome data	Unclear	Attrition unclear
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Hippman 2016

Study characteristics

Methods

Parallel randomised trial with three arms

1. Educational booklet: psychological placebo
2. No-treatment
3. Active treatment: genetic counselling

Sample calculation: yes

Cluster randomised: no

Duration of trial (baseline to post): not stated

Duration of participation (trial + follow-up:: not stated, but was done during September 2008–November 2011

Setting: inpatient and outpatient

Purpose of trial: “We hypothesized that 1) mean scores for knowledge, risk perception accuracy, and perceived control over illness would be higher, and scores for internalized stigma would be lower for the GC group compared to an intervention group provided with an educational booklet (EB), and 2) mean differences in scale scores between outcome (T3) and baseline (T1) for the two intervention groups (GC, EB) would be significantly different than waitlist, with GC/EB mean scores being higher for knowledge, risk perception accuracy, and perceived control over illness, and lower for internalized stigma.” (p. 3)

Hippman 2016 *(Continued)*

	Open/closed placebo: closed placebo
Data	**Number of participants screened:** not stated **Number of participants included:** 120 **Number of participants followed-up at post treatment:** 112 **Number of participants randomly assigned to:** • Psychological placebo: n = 40 • No-treatment: n = 40 • Active treatment: n = 40 **Number of withdrawals (post-treatment):** n = 8 • Psychological placebo: n = 4 • No-treatment: n = 0 • Active treatment: n = 4 **Diagnosis:** serious mental illness: Bipolar disorder (69.2%), Schizophrenia (16.7%), Schizoaffective Disorder (10.8%), Other (Major depression and Major depression with psychosis) (3.3) **Diagnostic manual:** Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) **Means of assessment:** the Structured Clinical Interview for DSM-IV Disorders (SCID) **Comorbidity:** a variety of different mental health disorders **Age:** 41.6 mean years (range = 17 to 73) **IQ:** >70, 76.5% attended college or university **Sex:**: 60.7% female **Ethnicity:** not stated **Country:** Canada **Inclusion criteria** 1. Individuals were enrolled if they reported a diagnosis of schizophrenia, bipolar disorder, or schizoaffective disorder 2. Were fluent in English 3. Had the capacity to provide informed and autonomous consent (e.g. ≥19 years of age). **Exclusion criteria** 1. Individuals were ineligible if their SMI diagnosis was substance-induced, 2. or their ability to provide autonomous informed consent was compromised (e.g. intellectual disability (IQ < 70), 3. or currently floridly psychotic and/or intoxicated).
Comparisons	**Psychological placebo** **Treatment name:** educational booklet **Description of intervention:** "The EB intervention was designed as a rigorous control intervention; it was face-to-face and provided the same general information as GC, but without the 'active ingredient' of personalization of information/counseling by a BC/EGC." (p. 4) "EB sessions (~30 minutes) were provided by the research coordinator (AR), who answered questions regarding literal interpretations of text, but responded to participants' queries that aimed to make personal meaning of the material with responses such as: "I'm sorry, but I'm afraid I'm unable to answer

Hippman 2016 *(Continued)*

	that. If you'd like to meet with someone who can help you with questions like that, we can set up a GC appointment after you finish the study". Thus, EB sessions did not evolve into GC, yet were a stringent control intervention. Through observation, the research coordinator confirmed participant adherence to the intervention. The booklet (16 color pages, reading grade level 8) was designed in collaboration with individuals with SMI and included: a graphical depiction of the concepts of vulnerability (genetic and environmental) and resilience (the "mental illness jar"), with specific examples and a table of general RRs for relatives of people with SMI." (p. 4-5) **Individual or group treatment:** individual **Exposure/intensity to treatment:** 30 minutes **Duration of treatment:** not stated **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated **No-treatment** **Comparison name:** wait-list (in reality no-treatment) **Description of intervention:** "For the waitlist group, baseline and T1 occurred on the same day. Participants had the option of bringing a support person with them to appointments if they wished. In-person visits were arranged for some participants to complete the outcome measures at one month follow-up at their request. One of the participants in the waitlist group had received GC for SMI prior to the study. The trial was stopped once the pre-determined number of participants had been recruited and those who were not lost to follow up had completed the study. The full protocol can be obtained from the corresponding author." (p. 14) **Exposure/intensity to treatment:** none **Duration treatment:** not stated **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy:** usable data, Self-reported, clinical relevance, global score • **Outcome chosen:** the Internalized Stigma of Mental Illness scale (ISMI) **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. Genetic counselling and the educational booklet improved knowledge; and genetic counselling, but not the educational booklet, improved risk perception accuracy for this population. 2. The impact of genetic counselling on internalised stigma and perceived control is worth further investigation. 3. Genetic counselling should be considered for patients with serious mental illnesses. **Key limitations from study authors** 1. However, importantly, our sample size was underpowered to detect the observed effect sizes for internalised stigma and perceived control. 2. Additionally, blinding was not possible; due to the nature of the study, participants were aware of the group to which they had been randomised.

Hippman 2016 *(Continued)*

3. Furthermore, the risk range used in the educational booklet was narrower than that typically provided on the basis of a family history evaluation, thus biasing towards less accurate results for the EB group. However, the ranges for the GC and WL groups were comparable.
4. We excluded individuals not fluent in English; our findings, therefore, may not be generalisable to other cultural contexts.

Other notes from review authors

1. Due to no data provided for the wait-list condition for post-treatment, and no response from authors, we used the 1 month follow-up data (T3)

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: “For the randomization procedure, equally-sized laminated cards were sorted into two opaque envelopes (one for males, containing 18 GC, and 17 of each EB and WL, and one for females, containing 22 GC, and 23 of each EB and WL). .” (p. 4)
Allocation concealment	Unclear	Quote: "Participants were asked to choose a card from the appropriate (male/female) envelope without looking (under the supervision of AR or AI).” (p. 4)
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and "wait-list"
Incomplete outcome data	Yes	Quote: “While the nature of the study and interventions precluded blinding for participants or providers, an independent party blind to group status conducted data analyses.” (p. 4) Attrition <15% (6.7%). Analysed data on everyone that received the treatment. Linear mixed effects models used. Blinded data analyst
Selective outcome reported	Yes	NCT00713804. No differences in trial registry and full report
Other sources of bias	Yes	No other sources found

Howlin 2007

Study characteristics

Methods

Cluster-randomised trial with three arms

1. Wait-list: Delayed Treatment Group
2. No treatment
3. Active treatment: Immediate Treatment Group

Sample calculation: yes

Cluster randomised: yes

Howlin 2007 *(Continued)*

Duration of trial (baseline to post): 5 months

Duration of participation (trial + follow-up): 1 week + 5 months

Setting: outpatient (school classroom)

Purpose of trial: "To assess the effectiveness of expert training and consultancy for teachers of children with autism spectrum disorder in the use of the Picture Exchange Communication System (PECS)." (p. 473)

Data

Number of participants screened: 38 classes

Number of participants included: 18 classes (88 participants)

Number of participants followed-up at post treatment: not stated (84 participants)

Number of participants randomly assigned to :

- Wait-list: n = 29
- No treatment: n = 29
- Active treatment: n = 30

Number of withdrawals : n = 4

- Wait-list: n = 4
- No treatment: n= 0
- Active treatment: n = 0

Diagnosis: autism

Diagnostic manual: not stated, but "All children had received a clinical diagnosis of autism prior to enrolment in the study" (p. 476)

Means of assessment: the Autism Diagnostic Observation Schedule-Generic (ADOS-G)

Comorbidity: not stated

Age: 6.8 mean years

IQ::not stated

Sex: 17% female

Ethnicity: not stated

Country: UK

Inclusion criteria

1. Have a formal clinical diagnosis of autism and to meet criteria for autism or autism spectrum disorder on the Autism Diagnosis Observation Schedule – Generic Module 1 (ADOS-G)
2. Have little or no functional language (i.e., not exceeding single words/word approximations)
3. Have no evidence of sensory impairment
4. Be aged between 4 and 11 years; not be using PECS beyond Phase 1 (i.e., able to exchange symbols only if prompted
5. Each class was required to have a minimum of 3 children meeting the above criteria

Exclusion criteria

1. Not stated

Comparisons

Wait-list

Treatment name: delayed treatment group

Howlin 2007 *(Continued)*

Description of intervention: DTG: Receiving PECS training 2 terms after initial baseline assessment.

Exposure/intensity to treatment: no treatment during waiting

Duration of treatment: 5 months

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

No-treatment

Comparison name :no-treatment group

Description of intervention: receiving no PECS training

Exposure/intensity to treatment: no treatment

Duration treatment: 5 months

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported, clinical relevance
- **Outcome chosen: t**he Autism Diagnostic Observation Schedule-Generic (ADOS-G) – subscale Reciprocal Social Interaction

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. The results indicate modest effectiveness of PECS teacher training/consultancy. Rates of pupils' initiations and use of symbols in the classroom increased, although there was no evidence of improvement in other areas of communication.
2. Treatment effects were not maintained once active intervention ceased.

Key limitations from study authors

1. Firstly, there were significant restrictions on financial resources and personnel (both in terms of researchers and consultants) as well as time (most children were to move classrooms at the end of the school year in which training took place).
2. Secondly, we relied on only one measurement point at each assessment period for each child.
3. Furthermore, although the classroom observation assessments had high ecological validity, in order to ensure a degree of comparability across schools the primary measures were restricted to snack times.
4. Thirdly, it was not possible to collect ongoing measures of treatment fidelity – either with regard to the PECS consultants or with regard to the practice of class teachers.
5. Fourthly, the assessors were not blinded to group allocation or treatment phase, as financial limitations precluded the use of additional blinded raters to code all the video recordings.
6. Finally, while our use of ordinal data was driven by the highly skewed distribution of our primary outcome variables, this might reduce sensitivity to detect change compared to continuous quantitative data.

Other notes from review authors

1. Usable data not available

Conflicts of interest: none found

Howlin 2007 *(Continued)*

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: "In each stratum, classes were randomly allocated to one of the three treatment conditions using an online randomisation programme (http://www.random.org)" (p. 475)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Quote:"Fourthly, the assessors were not blinded to group allocation or treatment phase, as financial limitations precluded the use of additional blinded raters to code all the video recordings." (p. 479)
Blinding of participants and personnel	No	Not possible to blind wait-list and no-treatment
Incomplete outcome data	Yes	Quote:“Following random assignment, one class (ITG) subsequently withdrew from the study. One girl entered a DTG class one year into the study; thus her data were available from Time 2–Time 3 only. At baseline, one other girl (NTG) failed to meet criteria for ASD." (p. 477) Excluded from further analysis. Seven children moved out of the DTG during the watching waiting period and did not receive treatment but they were assessed at Times 2 and 3 and their data included in the analyses on an intention-to-treat basis. The final groups were: ITG (5 classes, 26 children, 21 boys, 5 girls); DTG (6 classes, 30 children, 27 boys, 3 girls); NTG (6 classes, 28 children, 25 boys, 3 girls).” (p. 478) Attrition <15% (5.6%).
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources of bias found

Karst 2007

Study characteristics

Methods

Parallel randomised trial with four arms

1. Physical placebo: placebo auricular acupuncture group
2. No treatment
3. Active treatment 1: midazolam group
4. Active treatment 2: auricular acupuncture group

Sample calculation:: yes
Cluster randomised:: no
Duration of trial (baseline to post): 1 treatment (1 day)
Duration of participation (trial + follow-up): 1 treatment (1 day)
Setting: outpatient
Purpose of trial : “Therefore, we designed a study to determine whether auricular acupuncture can decrease acute dental anxiety and compared it with the standard pharmacological sedative medication midazolam, noninvasive placebo auricular acupuncture, and no treatment.” (p. 295)

Karst 2007 *(Continued)*

	Open or closed placebo: closed placebo
Data	**Number of participants screened:** 81 **Number of participants included:** 67 **Number of participants followed-up at post treatment:** 67 **Number of participants randomly assigned to:** • Physical placebo: n = 19 • No treatment: n = 10 • Active treatment 1: n = 19 • Active treatment 2: n = 19 **Number of withdrawals:** 0 **Diagnosis:** specific anxiety (dental anxiety) **Diagnostic manual:** not stated **Means of assessment:** not stated **Comorbidity:** not stated **Age:** 38 to 49 mean years (SD = 13.09) **IQ:** not stated **Sex:** 44.8% female **Ethnicity:** not stated **Country:** Germany **Inclusion criteria** 1. Inclusion criteria were dental extraction 2. Age of 1 to –65 years 3. German speaking 4. Informed consent **Exclusion criteria** 1. Exclusion criteria were allergy to benzodiazepines, 2. Addiction to any drugs or alcohol or the use of such substances preoperatively 3. Any major psychiatric, neurologic, or cardiopulmonary disorder 4. Previous acupuncture treatment 5. Anticoagulation 6. Pregnant or lactating
Comparisons	**Physical placebo** **Treatment name:** Placebo auricular acupuncture group **Description of intervention:** "In addition, patients in the placebo auricular acupuncture group were told that the needles would only be inserted gently and superficially and that an elastic cube would, therefore, be necessary to support the needle" (p. 296) "This group received placebo ear acupuncture by using the finger and liver points, which do not have any documented effects on anxiety reduction. A placebo needle system was used, in which the tip of the needle is blunt so as to cause a pricking sensation mimicking real acupuncture without actually puncturing the skin. To support the needle, an elastic foam was used which was fixed upon the area of the acupoint. In contrast to superficial sham acupuncture, this form of control may be associated with less unspecific physiological effects." (p. 296) **Individual or group treatment:** individual **Exposure/intensity to treatment:** 1 treatment **Duration of treatment:** 1 day **Concomitant psychotherapy:**: not stated **Concomitant pharmacotherapy:** not stated

Karst 2007 *(Continued)*

No-treatment
Comparison name: no treatment
Description of intervention: not stated
Exposure/intensity to treatment: no treatment
Duration treatment: not stated
Concomitant psychotherapy: not stated
Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported
- **Outcome chosen:** sedation score - follow-up 2 (after dental treatment)

Adverse events

- "Some patients (n=7, 36.8%) complained of nasal burning for a few minutes after intranasal midazolam administration. No adverse effects were reported in the other groups"

Notes

Key conclusion from study authors

1. With the no treatment group as control, the auricular acupuncture group, and the midazolam group were significantly less anxious at 30 minutes compared with patients in the placebo acupuncture group (Spielberger Stait-Trait Anxiety Inventory X1, P = 0.012 and <;0.001, respectively)
2. In addition, patient compliance assessed by the dentist was significantly improved if auricular acupuncture or application of intranasal midazolam had been performed (P = 0.032 and 0.049, respectively)
3. In conclusion, both, auricular acupuncture and intranasal midazolam were similarly effective for the treatment of dental anxiety

Key limitations from study authors

1. However, placebo auricular acupuncture also decreased anxiety somewhat; these effects may have been caused by a placebo system that was not totally inert or by psychological effects, such as patients' expectations and beliefs which, in acupuncture trials especially, can not only modulate treatment effects and neuronal substrates, but also baseline values. (…) However, we tried to reduce such effects to a minimum by having a dental student (B.F.) do the interventions. He was carefully trained for each procedure, but was not instructed on the theoretical background of acupuncture or the pharmacologic therapy of anxiety
2. Furthermore, communication between investigator and patients was restricted to a minimum. In addition, baseline assessment and all follow- up assessments were done by an independent investigator (A.H.) who was unaware of the treatment
3. Although patients were blinded regarding both acupuncture procedures, blinding was not achieved from the patients' perspective whether intranasal midazolam, auricular acupuncture, or no treatment were given. This may have been a source of significant bias. On the other hand, placebo or sham acupuncture may exert potential physiologic effects, which make it difficult to use such procedures for a double-dummy technique
4. Furthermore, it is not uncommon for acupuncture trials to compare against standard care, that is, no specific treatment for dental anxiety
5. Another potential limitation is that patients were included consecutively, regardless of heterogenous groups regarding general dental anxiety. However, the STAI baseline scores indicate that tooth extraction creates a specific anxiety, the awareness of which may be used to explore the consequences of dental anxiety in general. Additionally, the STAI baseline scores are about the same as those in the Hollenhorst et al's study which investigated intranasal midazolam to prevent claustrophobia induced by magnetic resonance imaging
6. Further potential limitations of our study are the relatively small population size, the small no treatment control group, and the lack of assessing pain that might be a potential source of bias in the settings of this study, although dental extraction was performed under local anaesthesia

Other notes from review authors

1. None

Karst 2007 *(Continued)*

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: "The names of the recruited patients were transmitted to the Department of Biometrics, Hannover Medical School. A list with random numbers was prepared by one of its members (L.H.)." (p. 296)
Allocation concealment	Yes	Author L.H. made the list and randomised. "statistical procedures (L.H.) (…) were blind to treatment condition" (p. 297)
Blinding of outcome assessors	Yes	Quote: "Both the investigators performing follow-up examinations (A.H.) and statistical procedures (L.H.), and the dentist were blind to treatment condition" (p. 297)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Attrition <15% (0%)
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Kelley 2012

Study characteristics

Methods

Parallel randomised trial with two arms

1. Pharmacological placebo: open-label placebo
2. Wait-list

Sample calculation: not stated
Cluster randomised: no
Duration of trial (baseline to post): 2 weeks
Duration of participation (trial + follow-up): 2 weeks (post treatment data) + 2/4 weeks of placebo treatment (no follow-up).
Setting: outpatient
Purpose of trial: investigating if open-label placebo can be used as a first-line treatment for depression
Open or closed placebo: open placebo

Data

Number of participants screened: not stated

Number of participants included: 20

Number of participants followed-up at post treatment: 15

Number of participants randomly assigned to:

- Pharmacological placebo: n = 11

Kelley 2012 *(Continued)*

- Wait-list: n = 9

Number of withdrawals: n = 5

- Pharmacological placebo: not stated
- Wait-list: not stated

Diagnosis: non-psychotic Major Depressive Disorder
Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV)
Means of assessment: the Structured Clinical Interview for DSM-IV Disorders (SCID)
Comorbidity: "While some comorbid conditions resulted in patients being excluded (e.g., schizophrenia), many other comorbid conditions were allowed (e.g., Generalised anxiety disorder (GAD), so long as the GAD was not primary over major depressive disorder (MDD))." (Kelley 2012 (pers comm))
Age: 38.8 mean years (SD = 12.6)
IQ: not stated
Sex: 70% female
Ethnicity:: not stated
Country: USA

Inclusion criteria

1. Men or women aged 18-60 years old
2. Current Major Depressive Disorder (MDD)
3. Written informed consent
4. A score of 11 or greater on the Quick Inventory of Depressive
5. Symptomatology – Self-Rated (QIDS-SR)
6. For wait-list/no treatment group: Patient must continue to meet criteria for
7. current MDD at baseline. Patient must have Clinical Global Impression
8. Improvement (CGI) scores; 2 (i.e. less than much or very much
9. improved) from the screen to the baseline visit

Exclusion criteria

1. A score of greater than 25 on the HAMD-17 and/or a score of 6 or greater on the CGI-Severity scale
2. Pregnant women or women of child bearing potential not using a medically accepted means of contraception
3. Patients who are a serious suicide or homicide risk
4. Unstable medical illness including cardiovascular, hepatic, renal, respiratory, endocrine, neurological, or hematological disease
5. The following Diagnostic and Statistical Manual of Mental Disorders-IV diagnoses: a) organic mental disorders; b) substance use disorders, including alcohol, active within the last year; c) schizophrenia; d) delusional disorder; e) psychotic disorders not elsewhere classified; f) bipolar disorder; g) acute bereavement; h) severe borderline or antisocial personality disorder; i) current primary diagnoses of panic disorder, social phobia, generalised anxiety disorder (GAD), or obsessive compulsive disorder (OCD) (disorders that present as chief complaint and/or have their onset preceding the onset of major depressive disorder)
6. Uncontrolled seizure disorder
7. Patients with mood congruent or mood incongruent psychotic features
8. Current use of other psychotropic drugs. Exception: Patients who have been on a stable dose for 30 days of classes of medications such as non-benzodiazepine sedatives, anxiolytic benzodiazepines, non-narcotic analgesics may be included. Flexibility will be allowed based on physician discretion
9. Clinical or laboratory evidence of hypothyroidism
10. Patients who have taken an investigational psychotropic drug within the last year.
11. Patients who have not responded to two or more antidepressant trials of adequate doses (e.g., fluoxetine 40 mg/day or higher) and duration (e.g.,for six weeks or more) over the past five years
12. Any concomitant form of psychotherapy (depression focused)

Comparisons	**Pharmacological placebo**

Kelley 2012 *(Continued)*

Treatment name: Open-label placebo
Description of intervention: Patients were instructed to take two placebo pills, twice daily. The placebos were blue capsules containing microcrystalline cellulose" (p. 1).
Individual or group treatment individual
Exposure/intensity to treatment: 2 placebo pills twice daily
Duration of treatment: 2 weeks (post treatment data) + 2 weeks
Concomitant psychotherapy: not allowed (see exclusion criteria)
Concomitant pharmacotherapy: not allowed (see exclusion criteria)

Wait-list
Comparison name: wait-list control
Description of intervention: waiting for treatment/placebo
Exposure/intensity to treatment: waiting for treatment
Duration treatment :2 weeks
Concomitant psychotherapy: not allowed (see exclusion criteria)
Concomitant pharmacotherapy: not allowed (see exclusion criteria)

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported, clinician-rated
- **Outcome chosen:** 17-item Hamilton Scale for Depression (HAM-D-17)

Adverse events

- No adverse events mentioned

Notes

Key conclusion from study authors

1. The results do not support the hypothesis that open-label placebo is an effective treatment for depression, however small statistically significant improvements were found.

Key limitations from study authors

1. Small sample size, larger trials for open-label placebo for MDD are warranted
2. Low statistical power
3. Short duration

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: "The randomization itself was created by the biostatistician using a computer to generate the sequence of assignments." (Kelley 2012 (pers comm))
Allocation concealment	Yes	Quote:" prior to enrollment and the revelation of treatment assignment, the randomization was concealed from both the clinician and patient." (Kelley 2012 (pers comm))
Blinding of outcome assessors	Yes	Quote:"Blinded clinicians assessed patients at baseline and every two weeks thereafter. The primary outcome was the clinician-rated 17-item Hamilton Scale for Depression." (p. 1).

Kelley 2012 *(Continued)*

		"In addition, all assessments were conducted by assessors who were blinded to treatment allocation." (Kelley 2012 (pers comm))
Blinding of participants and personnel	No	Open-labelled placebo. Quote:"(1) Since this was a trial of open-label placebo vs. no treatment control, patients and clinicians were not blinded during treatment." (Kelley 2012 (pers comm))
Incomplete outcome data	Unclear	Attrition >15% (25%). No mention of ITT
Selective outcome reported	No	NCT01103271 Different primary outcome measure – feasibility (timeframe; one year) in protocol, while HAMD-17 in the full report
Other sources of bias	Yes	No other sources found

Kennedy 1974

Study characteristics	
Methods	**Parallel randomised trial with six arms** 1. Psychological placebo: pseudo-desensitisation 2. No-treatment: untreated 3. Active treatment 1: desensitization group 1 4. Active treatment 2: desensitization group 2 5. Active treatment 3: desensitization group 3 6. Active treatment 4: desensitization group 4 **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 1 year (max. of 6 sessions) **Duration of participation (trial + follow-up):** 1 year **Setting:** outpatient (college) **Purpose of trial:** "Thus, another purpose of the present study was to obtain a more accurate picture of the relationship between anxiety decrements and approach behavior at various stages of performance on the behavioral avoidance test." (p. 722-3) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** 81 **Number of participants included:** 74 **Number of participants followed-up at post treatment:** 60 **Number of participants randomly assigned to:** • Psychological placebo: n = 10 • No-treatment: n = 10 • Active treatment 1: n = 10 • Active treatment 2: n = 10 • Active treatment 3: n = 10

Kennedy 1974 *(Continued)*

- Active treatment 4: n = 10

Number of withdrawals : n = 14

- Psychological placebo: not stated
- No-treatment: not stated
- Active treatment 1: not stated
- Active treatment 2: not stated
- Active treatment 3: not stated
- Active treatment 4: not stated

Diagnosis: specific anxiety (snake)

Diagnostic manual: not stated

Means of assessment:: behavioural avoidance test

Comorbidity: not stated

Age: not stated

IQ: not stated – but college students

Sex: 100% female

Ethnicity: not stated

Country: USA

Inclusion criteria

1. Seventy-four students satisfied the pretreatment Behavior Avoidance Test (BAT) criterion of not being able to reach Step 10

Exclusion criteria

1. Not stated

Comparisons

Psychological placebo

Treatment name: Pseudo-desensitisation

Description of intervention: "Subjects in this control condition received the same type and amount of relaxation training as subjects in the desensitization groups. In contrast to the latter groups, however, relaxation was paired with snake-irrelevant stimuli during the subsequent therapy sessions. That is, pseudo-desensitization subjects were instructed to relax and imagine neutral, pleasant scenes such as walking in the mountains, sailing, frolicking at the beach, etc. Pseudo-desensitization subjects were matched with subjects in the 100% desensitization group in terms of the number of treatment sessions and the time of the posttreatment assessment." (p. 722)

Individual or group treatment: individual

Exposure/intensity to treatment: 40-minute sessions for 6 sessions

Duration of treatment: 1 year

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

No-treatment

Comparison name: untreated (no treatment)

Kennedy 1974 *(Continued)*

	Description of intervention: "Untreated subjects participated only in the pretreatment and posttreatment assessment procedures. Posttreatment evaluation was conducted at approximately the same interval as for subjects in the 100% desensitization group." (p. 722) **Exposure/intensity to treatment:** no treatment **Duration treatment:** 1 year **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy:** only one outcome • **Outcome chosen:** Behavior Avoidance Test (BAT) **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. Degree of transfer and fear change associated with four levels of desensitization, pseudodesensitisation, and no treatment were assessed in snake-phobic students. 2. Only participants desensitised to 75% or more of the hierarchy demonstrated reliably greater reductions in avoidance behaviour than controls. 3. However,participants completing 50% or less of the hierarchy showed smaller transfer decrements than those who finished the hierarchy. 4. Evidence also suggested that repeated exposure tends to improve transfer efficiency. On the posttest, desensitisation participants reported significantly less anxiety than no-treatment controls when repeating their highest pretreatment responses, but were no different from either control group when performing new approach responses, suggesting that behavioural improvement is not dependent upon the elimination or inhibition of conditioned emotional arousal. **Key limitations from study authors** 1. Not stated **Other notes from review authors** 1. Terminal = post-treatment 2. SD was generated from Etringer 1982 and Rosen 1976 (same outcome, same population, same scale) 3. Only reports data on completers **Conflicts of interest:** none found J**udgement:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Assignment to groups was random, with the constraint that extreme pretreatment BAT scores be distributed in order to keep pretreatment means reasonably equal." (p. 722)
Allocation concealment	Unclear	No information

Kennedy 1974 *(Continued)*

Blinding of outcome assessors	Yes	Quote:"The experimental assistant who administered the BAT participated only in the assessment procedures and had no knowledge of which group the subjects represented." (p. 722)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	No	Attrition >15% (18.92%). No ITT. Only reports data on completers
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Kilmann 1987

Study characteristics	
Methods	**Parallel randomised trial with five arms** 1. Psychological placebo: attention-placebo 2. Wait-list: no-treatment 3. Active treatment 1: communication Technique Training 4. Active treatment 2: Sexual Technique Training 5. Active treatment 3: Combination Treatment **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial:** 4 weeks **Duration of participation (trial + follow-up):**: 4 weeks + 6 months follow-up. However, the same participants were used in Kilmann 1985 and Kilmann 1988 . In these they were compared to a healthy sample. **Setting:**: outpatient (WJB Dorn Veterans Hospital in Columbia) **Purpose of trial:** Testing the effect of Group treatment (Communication Technique Training, Sexual Technique Training, Combination Treatment) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** not stated **Number of participants included:** 21 **Number of participants followed-up at post treatment:** 20 **Number of participants randomly assigned to:** • Psychological placebo: n = 4 • Wait-list: n = 4 • Active treatment 1: n= 4 • Active treatment 2: n = 4 • Active treatment 3: n = 4

Kilmann 1987 *(Continued)*

Number of withdrawals: n = 1, one man dropped out after the first week of treatment due to an unexpected illness which required hospitalisation.

Diagnosis: erectile dysfunction

Diagnostic manual: not stated

Mean of assessment: Clinical Interview (Sexual Interaction Inventory)

Comorbidity: not stated

Age:: 51 mean years, (range = 31 to 67)

IQ: not stated – their education average was 15.3 years, with a range of 12 to 20.

Sex: 100% male (and with their respective partners)

Ethnicity: not stated

Country: USA

Inclusion criteria

1. the man must have reported an inability to experience successful penetration of the vagina and subsequent ejaculation in 20% or more of his attempts during the past 5 months;
2. the man must not have been more than 70 years old;
3. the man must have been in a committed relationship with his partner for at least the past six months;
4. the man's partner was willing to participate in treatment;
5. neither the man nor his partner had debilitating levels of anxiety, depression, or hostility as determined from extreme scores on the Multiple Affect Adjective Checklist;
6. neither the man nor his partner had any disturbances in reality testing as assessed from a clinical interview and a score of 10 or more on the Whitaker Index of Schizophrenic Thinking";
7. both partners agreed to participate in 20 hours of group treatment and to respond to a battery of measures before, during, and after treatment, and at a 6-month follow-up;
8. the man agreed to undergo extensive medical/physiological screening and testing at a cost of up to $100;
9. and the man must have been judged from the medical/physiological screening to be able to experience improvement in his ability to gain and maintain erections in his sexual interactions.

Exclusion criteria

1. Exclude focal neurologic disease

Comparisons

Pscyhological placebo

Treatment name: Attention placebo control

Description of intervention: “This format was conducted for eight 2-hour sessions for a total of 16 hours. The format was designed to control for the therapist and treatment variables thought to be inherent in the Communication Technique Training, Sexual Technique Training, and the Combination Treatment formats. (p. 172)

Individual or group treatment: group-format.

Exposure/intensity to treatment: 8x 2 hours sessions (total 16 sessions)

Duration of treatment: 4 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: “Nine of the 20 men were taking some type of medication on a regular basis; the medication was not presumed to interfere with the ability to gain and maintain an erection in sexual situations.” (p. 175)

Kilmann 1987 *(Continued)*

Wait-list

Comparison name: no treatment (in reality wait-list)

Description of intervention: "The couples in the No-Treatment Control group did not receive any treatment for a 5-week waiting list period; they responded to the outcome measures in the same 5-week pre- to posttesting time interval as the couples in the other experimental groups. After posttesting, these couples received two 2-hour sessions of sex education followed by the Combination Treatment format described above." (p. 172)

Individual or group treatment: group-format

Exposure/intensity to treatment: none (wait-listed)

Duration treatment: 5 weeks of waiting

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: "Nine of the 20 men were taking some type of medication on a regular basis; the medication was not presumed to interfere with the ability to gain and maintain an erection in sexual situations." (p. 175)

Outcomes

Beneficial effect

- **Hierarchy:** patient-reported, clinical relevance
- **Outcome chosen:** Sexual Interaction Inventory (SII)

Adverse events

- No data on adverse events reported

Notes

Key conclusion from study authors

1. All three treatment groups fostered substantial gains so that between format differences were not statistically significant
2. Subject variables which predicted success! experience ratio gains included age of the male partner, perceived level of relationship adjustment, and the male partner's success experience ratio prior to treatment
3. Eighty-one percent of the treated men reached the criterion of 80% or greater success! experience ratio (successful penetration and subsequent ejaculation) at the 6-month follow-up
4. Good nocturnal tumescence prior to treatment was correlated with a better treatment outcome than poor tumescence

Key limitations from study authors

1. Across all treatments, the man's age, income, success of prior sexual functioning, perceived level of relationship adjustment, and ability to rupture two or more bands on the Snap Gauge (DACOMED) test predicted outcome, regardless of the treatment format
2. The men of the couples who reflected more pretreatment relationship discord reported greater gains in sexual harmony and in their success/experience ratios after treatment.
3. The intensive (i.e. 4 hours each week for 4 weeks) treatment formats used in this study may be more beneficial for men with these pretreatment characteristics
4. A larger sample of relatively homogeneous men should be recruited and screened
5. The findings of the present study suggest that these men should be matched on demographic, relationship, and physiological variables prior to assignment to experimental and control conditions

Other notes from review authors

1. None

Conflicts of interest: none found

Kilmann 1987 *(Continued)*

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote:"Their assignment to a particular group would be done on a random, statistical basis and would have nothing to do with them personally" (p. 175) "A table of random numbers determined that the therapist would conduct the treatment formats in the following sequence: Attention-Placebo control, Combination Treatment, Communication Technique Training, No-Treatment control, Sexual Technique Training." (p. 176)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and "no-treatment"
Incomplete outcome data	Yes	Attrition <15% (5%). Only 1 person dropped out after 1 week of trial. Not stated which group he was a part of. No ITT
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Klein 1977

Study characteristics

Methods

Parallel randomised trial with three arms

1. Psychological placebo: attention-placebo control
2. No-treatment control
3. Active treatment 1: exercise training
4. Active treatment 2: relaxation

Sample calculation: not stated
Cluster randomised: no
Duration of trial (baseline to post): 3 weeks
Duration of participation (trial + follow-up:: 3 weeks + 1 week follow-up
Setting: outpatient
Purpose of trial: "This study explored the effectiveness of progressive relaxation and large muscle exercise in improving the cognitive performance of hyperactive, impulsive males.." (p. 1159)
Open/closed placebo: closed placebo

Data

Number of participants screened: not stated

Number of participants included: 24

Number of participants followed-up at post treatment: not stated

Number of participants randomly assigned to:

Klein 1977 *(Continued)*

1. Psychological placebo: n = 6
2. No-treatment: n = 6
3. Active treatment 1: n = 6
4. Active treatment 2: n = 6

Number of withdrawals: not stated

Diagnosis: hyperactivity/impulsivity

Diagnostic manual: not stated

Means of assessment: “The rating scale was modeled after Conners' (1969) scale and consisted of short behavioral definitions of restlessness, impulsivity, distractibility, and short attention span. Employing a 1 to 4 rating scale (not at all, small degree, generally, very much) teachers rated the degree to which each behavioral description characterized the child's typical classroom behavior. Scores were summed across the four behaviors with greater scores representing greater hyperactivity.” (p. 1160)

Comorbidity: not stated

Age: not stated – but third grade

IQ: not stated

Sex:: 100% male

Ethnicity: not stated

Country: USA

Inclusion criteria

1. The 24 most "hyperactive, impulsive" males (3 from each class)

Exclusion criteria

1. Not stated

Comparisons

Psychological placebo

Treatment name: Attention-placebo

Description of intervention: “The task of making various objects from "Play Doh" was assigned to the attentional-training group. This task was chosen because it appeared neither especially rewarding nor boring. Each child in this control group worked by himself at a table inside the van.” (p. 1160)

Individual or group treatment: individual

Exposure/intensity to treatment: five sessions, 20 minutes each

Duration of treatment: 3 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

No-treatment

Comparison name: No-treatment

Description of intervention : “Subjects in the no-treatment and nonhyperactive controls received no intervention throughout the 3-wk. period.” (p. 1160)

Exposure/intensity to treatment: no treatment

Duration treatment:: 3 weeks

Klein 1977 *(Continued)*

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported, clinical relevance
- **Outcome chosen:** Matching Familiar Figures Test

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. Progressive relaxation and large muscle exercise were compared to an attentional-training placebo, a no-treatment control, and a non-hyperactive control.
2. While no differences were found on the continuous Performance Task, relaxation, exercise, and non-hyperactive control groups performed significantly better on the Matching Familiar Figures test than the no-treatment control.
3. Results were suggestive of the effectiveness of both progressive relaxation and large muscle exercise in treating hyperactive, impulsive youngsters

Key limitations from study authors

1. These conclusions must be interpreted in light of the experimental design. The conclusions apply only to training manipulations which not only train subjects in relaxation or exercise skills but also have participants apply them just before testing.
2. If some of these children differ greatly in the level of body awareness and ability to differentiate internal feelings of relaxation and tension, they might conceivably profit from longer and more individualised relaxation training or biofeedback assisted relaxation which provides continuous visual or auditory feedback

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: The 24 most "hyperactive, impulsive" males (3 from each class) were randomly assigned to one of four treatments: (a) muscle relaxation, (b) large muscle exercise, (c) attention-placebo control, (d) no- treatment control. (p. 1160)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote:“The testing took place at the subject's school and was administered by an examiner who was not familiar with the subjects or the hypotheses of the study.” (p. 1160)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Unclear	Attrition unclear

Klein 1977 *(Continued)*

Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Klerman 1974a

Study characteristics	
Methods	**Parallel randomised trial with three arms** High interpersonal contact: 1. Pharmacological placebo:placebo 2. No-treatment: no pill 3. Active treatment: active drug group **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 8 months. **Duration of participation (trial + follow-up:**: 8 months (no follow-up) **Setting:** outpatients **Purpose of trial:** evaluate maintenance treatment of depression including feasibility, efficacy and safety. **Closed or open placebo:** closed placebo
Data	**Number of participants screened:** 278 **Number of participants included:** 150 **Number of participants followed-up at post treatment:** 139 (completed or relapsed). **Number of participants randomly assigned to:** 1. Pharmacological placebo: n =2 5 2. No-treatment: n = 25 3. Active treatment: n= 25 **Number of withdrawals** : n = 6 1. Pharmacological placebo: n = 0 No-treatment: n = 3 1. Active treatment: n= 3 **Diagnosis:** depression **Diagnostic manual:** Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition (DSM-III) **Means of assessment:** Brief Psychiatric Rating Scale (BPRS) **Comorbidity:** ”85 percent of our patients were diagnosed as having neurotic depressions, according to criteria from the American Psychiatric Association’s Diagnostic and Statistical Manual of Mental Disorders” p. 187. **Ag:**: 39 mean years **IQ:** not stated **Sex:** 100% female **Ethnicity:** 83.3% white **Country:** USA **Inclusion criteria**

Klerman 1974a *(Continued)*

1. For entrance into the study (preliminary phase) it requires that the patients score at least 7 on the Raskin Depression Scale.
2. For completion of the preliminary phase and entrance into the maintenance phase it requires a 50-percent decrease in the patient initial score on the Raskin Depression scale (symptomatically responsive to amitriptyline).
3. Female.

Exclusion criteria

1. Patients were excluded if the depression appeared secondary to another predominant syndrome, such as schizophrenia.
2. The following patient populations were also excluded: alcoholics, drug addicts, patients with subnormal intelligence or serious physical illnesses, patients receiving ongoing psychotherapy, or patients who had failed to respond to an adequate course of tricyclic antidepressants in the last six months.

Comparisons

Pharmacological placebo
Treatment name: Placebo group + high interpersonal contact
Description of intervention: no description of pharmacological placebo "The high contact group met with a social worker for a minimum of one hour a week" (p. 187).
Individual or group treatment: individual
Exposure/intensity to treatment: exposure to pharmacological placebo is not stated. High contact group: met with a social work for a minimum of one hour a week.
Duration of treatment: 8 months.
Concomitant psychotherapy: not allowed (exclusion criteria)
Concomitant pharmacotherapy: not stated

No-treatment
Comparison name: No pill group (no treatment) + high interpersonal contact
Description of intervention: "(…) a no pill group (a control for placebo effect), which received regular monthly visits from a psychiatrist and the same battery of rating scales as the other groups, but which did not receive any medication or pills." (p. 187)
Exposure/intensity to treatment: no pharmacological treatment. High contact group: met with a social work for a minimum of one hour a week
Duration treatment: 8 months
Concomitant psychotherapy: not allowed (exclusion criteria)
Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect

- **Hierarchy:** usable data
- **Outcome chosen: c**linical relapses

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. "On the basis of our findings and those of other recent studies, we conclude that maintenance therapy is effective, feasibly, and relatively safe, is reasonably effective in reducing relapse, and that the efficacy depends upon the diagnostic and historical backgrounds of the patients." (p. 190)
2. ".. the most conclusive current findings support the efficacy of maintenance drug therapy for selected depressions." (p. 190)
3. "Although the value of the drug effect is significantly confirmed in our study there remain many unanswered questions about the value of psychotherapy." (p. 190)

Key limitations from study authors

1. "Because there were only 25 patients in each cell, the interpretation of the relapse rates may be difficult." (p. 188)

Klerman 1974a *(Continued)*

2. Since the preliminary phase did not have a control group, it is not possible to conclude that the improvements was due specifically to the drug.

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "In the maintenance phase of our project, patients were assigned to treatment in a six-cell design in a double-blind controlled manner (…)" (p. 186)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Unclear	Quote: "Assessment ratings were carried out during a 1- to 1,5-hour semi-structured interview with the patient, by two Bachelor Degree-level research assistants not involved in the treatment." (Weissman 1974 , p. 773)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Quote: "Patients terminated prior to month 9 for reasons other than relapse were counted for each completed month but for only one-half of the month of termination." Attrition <15% (7.3%)
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Klerman 1974b

Study characteristics

Methods

Parallel randomised trial with three arms

Low interpersonal contact:

1. Pharmacological placebo: placebo
2. No-treatment: no pill
3. Active treatment: active drug group

Sample calculation: not stated
Cluster randomised: no
Duration of trial (baseline to post): 8 months.
Duration of participation (trial + follow-up): 8 months (no follow-up)
Setting : outpatients
Purpose of trial: to evaluate maintenance treatment of depression including feasibility, efficacy and safety.
Closed or open placebo: closed placebo

Klerman 1974b *(Continued)*

Data	**Number of participants screened:** 278 **Number of participants included:** 150 **Number of participants followed-up at post treatment:** 139 (completed or relapsed) **Number of participants randomly assigned to:** 1. Pharmacological placebo: n = 25 2. No-treatment: n = 25 3. Active treatment: n = 25 **Number of withdrawals** : n = 5 1. Pharmacological placebo: n = 3 2. No-treatment: n = 0 3. Active treatment: n= 2 **Diagnosis:** depression **Diagnostic manual** : Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition (DSM-III) **Means of assessment** : Brief Psychiatric Rating Scale (BPRS) **Comorbidity:** ”85 percent of our patients were diagnosed as having neurotic depressions, according to criteria from the American Psychiatric Association’s Diagnostic and Statistical Manual of Mental Disorders” p. 187. **Age:** 39 mean years **IQ:** not stated **Sex:** 100% female **Ethnicity:**: 83.3% white **Country:** USA **Inclusion criteria** 1. For entrance into the study (preliminary phase) it requires that the patients score at least 7 on the Raskin Depression Scale. 2. For completion of the preliminary phase and entrance into the maintenance phase it requires a 50-percent decrease in the patient initial score on the Raskin Depression scale (symptomatically responsive to amitriptyline). 3. Female. **Exclusion criteria** 1. Patients were excluded if the depression appeared secondary to another predominant syndrome, such as schizophrenia. 2. The following patient populations were also excluded: alcoholics, drug addicts, patients with subnormal intelligence or serious physical illnesses, patients receiving ongoing psychotherapy, or patients who had failed to respond to an adequate course of tricyclic antidepressants in the last six months.
Comparisons	**Pharmacological placebo** **Treatment name:** Placebo group + low interpersonal contact **Description of intervention:** no description of pharmacological placebo The low contact group "saw the project psychiatrist for 15 minutes once a month for the completion of the rating scale, management of drug doses, and questions about possible side-effects.” (p. 187) **Individual or group treatment:** individual **Exposure/intensity to treatment:** exposure to pharmacological placebo is not stated. low contact group: met with a project psychiatrist for 15 minutes once a month **Duration of treatment:** 8 months. **Concomitant psychotherapy:** not allowed (exclusion criteria) **Concomitant pharmacotherapy:** not stated **No-treatment** **Comparison name:**No pill group (no treatment) + low interpersonal contact

Klerman 1974b *(Continued)*

	Description of intervention: "(…) a no pill group (a control for placebo effect), which received regular monthly visits from a psychiatrist and the same battery of rating scales as the other groups, but which did not receive any medication or pills." (p. 187) **Exposure/intensity to treatment:** no pharmacological treatment. low contact group: met with a project psychiatrist for 15 minutes once a month **Duration treatment:** 8 months **Concomitant psychotherapy:** not allowed (exclusion criteria) **Concomitant pharmacotherapy:** not stated
Outcomes	See Klerman 1974a
Notes	See Klerman 1974a

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "In the maintenance phase of our project, patients were assigned to treatment in a six-cell design in a double-blind controlled manner (…)" (p. 186)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Unclear	Quote: "Assessment ratings were carried out during a 1- to 1,5-hour semi-structured interview with the patient, by two Bachelor Degree-level research assistants not involved in the treatment." (Weissman 1974 , p. 773)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Quote: "Patients terminated prior to month 9 for reasons other than relapse were counted for each completed month but for only one-half of the month of termination." Attrition <15% (7,3%)
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Klosko 1990

Study characteristics

Methods	**Parallel randomised trial with four arms** 1. Pharmacological placebo 2. Wait-list 3. Active treatment 1: alprazolam 4. Active treatment 2: procalcitonin **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 15 weeks. **Duration of participation (trial + follow-up):** 15 weeks (no follow-up) **Setting:** outpatient

Klosko 1990 *(Continued)*

Purpose of trial: "The purpose of this study is to evaluate the relative effectiveness of each treatment in one setting using a group of clients diagnosed in an identical manner with outcome measured in precisely the same way (…) This study was seen as a precursor to studying possible integration or coordination of these treatments in panic disorder patients." (p. 78)
Closed or open placebo: closed placebo

Data

Number of participants screened: not stated

Number of participants included:: 69

Number of participants followed-up at post treatment: 57

Number of participants randomly assigned to:

- Pharmacological placebo: n = 18
- Wait-list: n = 16
- Active treatment 1: n = 17
- Active treatment 2: n = 18

Number of withdrawals: n = 12

- Pharmacological placebo: n = 7
- Wait-list: n = 1
- Active treatment 1: n = 1
- Active treatment 2: n = 3

Diagnosis: Panic Disorder (with clinician's severity rating of at least 4 on a 0-8 scale).
Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition (DSM-III)
Means of assessment: Anxiety Disorders Interview Schedule-Revised (ADIS-R).
Comorbidity: not stated
Age: 37 mean years (SD = 11.04) (Range: 18 to 65)
IQ: not stated
Sex: 74% female
Ethnicity: "One dropout from the placebo group was Black; all other subjects were White." (p. 80)
Country: USA

Inclusion criteria

1. Panic disorder.
2. Age between 18 and 65 years.
3. At least moderate severity.
4. "Only subjects who were panicking actively were included; that is, subjects who reported at least one panic attack in the week before starting treatment on the weekly record self-monitoring form." (p. 78)

Exclusion criteria

1. Patients who had begun pharmacotherapy or psychotherapy in the past 6 months.
2. Patients who had been either in drug or psychotherapeutic treatment more than 6 months, unless they agreed to stop such treatment for the duration of the study.
3. Patient who had been on 4 mg or more of alprazolam for any 3-week period and were non-responders, who displayed evidence of benzodiazepine hypersensitivity, or who had undergone cognitive-behavioral therapy for anxiety at any time.
4. Females who were pregnant or lactating or at risk to become pregnant.
5. Patients with significant medical problems, as determined by history, medical report, and laboratory values.
6. Patients with a history of psychotic disorder or dementia.
7. Patients with a history of alcohol or other substance abuse within the last 6 months.
8. Patients with current or past bipolar disorder.
9. Patients with depression only if depression predominated over panic disorder at the time of presentation and if depression preceded panic disorder chronologically.

Klosko 1990 *(Continued)*

	10.Patients with acute suicidal ideation
Comparisons	**Pharmacological placebo** **Treatment name:** Placebo **Description of intervention:** "Subjects received 15 individual treatment sessions in weekly meeting with a study psychiatrist experienced in alprazolam treatment of panic disorder. Medication was supplied by the Upjohn Company in matching 1-mg tablets, packaged in matching bottles containing sufficient medication for 1 week (…) Medication was gradually increased following a standardized but flexible schedule until maximum benefit was achieved or dose-limiting side effects occurred (…) The psychiatrist was instructed to limit interactions with subjects to discussion of clinical history, explanation of panic disorder, discussion of medication effects and side effects, and general support." (p. 78-9) **Individual or group treatment** : individual **Exposure/intensity to treatment:** not stated, but gradually increased dose to a maximum of 10 mg per day if required and at least three attempts made to titrate the medication to at least 6 mg per day. **Duration of treatment:** 15 weeks. **Concomitant psychotherapy:** "subjects who had been either in drug or psychotherapeutic treatment more than 6 months were excluded unless they agreed to stop such treatment for the duration of the study." (p. 78) **Concomitant pharmacotherapy:** no – "Subjects in the three treatment groups withdrew from prestudy medications under the supervision of the study psychiatrist. Adherence to drug withdrawal was determined by analyses of plasma benzodiazepine screens." (p. 78) **Wait-list** **Comparison name** : Waiting list **Description of intervention:** "Subjects were placed on a 15-week waiting list for treatment. They were told that they might contact the clinic by telephone during this time if they felt the need and that we would contact them approximately weekly by telephone." (p. 79) **Exposure/intensity to treatment:** no treatment **Duration treatment:** 15 weeks **Concomitant psychotherapy:** "subjects who had been either in drug or psychotherapeutic treatment more than 6 months were excluded unless they agreed to stop such treatment for the duration of the study." (p. 78) **Concomitant pharmacotherapy:** yes – "Although inclusion criteria required all subjects to have been stabilized on medication, waiting-list subjects were not required to withdraw from medication". Therefore, this waiting-list group might also be considered a minimal treatment condition, thereby providing a more conservative comparison with other treatment conditions." (p. 79)
Outcomes	**Beneficial effect** • **Hierarchy:** observer-rated, clinical relevancy • **Outcome chosen:** Anxiety Disorder Interview Schedule-Revised (ADIS-R) **Adverse events** • Spontaneous reporting
Notes	**Key conclusion from study authors** 1. "Patterns of results on measures of panic attacks, generalized anxiety, and global clinical ratings reveal that PCT was significantly more effective that placebo and waiting-list conditions on most measures." (p. 77) 2. "The percentage of clients completing the study who were free of panic attacks following PCT was 87%, compared with 50% for alprazolam, 36% for placebo, and 33% for the waiting-list group." (p. 77) 3. "The alprazolam group differed significantly from neither PCT nor placebo." (p. 77) **Key limitations from study authors** 1. "In additional caution we have alluded to concerns the fact that posttreatment measures in the alprazolam group were taken after attempts to withdraw these clients from alprazolam. When this did not prove feasible, subjects were quickly restablized on study dosage, and data would suggest that with-

Klosko 1990 *(Continued)*

drawal symptoms did not adversely influence posttreatment measures. Nevertheless, the possibility still exist that this adverse influence occurred in some patients." (p. 84)

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Subjects remained off medication for at least 7 days before administration of psychophysiological, self-report, and self-monitoring measures and random assignment to one of the three treatment groups." (p. 78)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote: Patients"Posttreatment clinical assessment measures were gathered through administration of a short form of the ADIS-R. The ADIS-R administrators were blind to group assignment." (p. 81)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	No	Quote: "A higher rate of dropout was observed in the placebo group compared with the other three groups." "In any case, if one considers only completers, then 45% of placebo completers achieved high end state functioning posttreatment; but if one includes dropouts, only 28% of placebo subjects achieved high end state functioning." (p. 84) Attrition >15% (17,4%). Only reports data on completers
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Krapfl 1970

Study characteristics

Methods

Parallel randomised trial with five arms

1. Psychological placebo: pseudodesensitisation
2. No-treatment
3. Active treatment 1: Traditional desensitisation
4. Active treatment 2: Desensitisation reversed
5. Active treatment 3: Random order desensitisation

Sample calculation: not stated

Cluster randomised: no

Duration of trial (baseline to post): 5 weeks (5 treatment sessions about a week apart)

Krapfl 1970 *(Continued)*

Duration of participation (trial + follow-up): 5 weeks + 6 weeks follow-up

Setting: outpatient (University setting)

Purpose of trial: "The present experiment attempted to examine whether or not SD would be successful if the aversive imagery were presented in random or decreasingly aversive hierarchical order." (p. 333)

Open/Closed: closed placebo

Data

Number of participants screened: 1200

Number of participants included : 50

Number of participants followed-up at post treatment: not stated

Number of participants randomly assigned to:

- Psychological placebo: n = 10
- No-treatment: n = 10
- Active treatment 1: n = 10
- Active treatment 2: n = 10
- Active treatment 3: n = 10

Number of withdrawals: not stated

Diagnosis::specific anxiety (Snake phobia)

Diagnostic manual: not stated

Means of assessment:: exposure of snake

Comorbidity: not stated

Age:not stated

IQ: not stated. College undergraduates

Sex: 100% female

Ethnicity: not stated

Country: USA

Inclusion criteria

1. Inability to touch the snake (in test) with the gloved hand
2. Affirmative answers to the three questions relative to the imagery test

Exclusion criteria

1. Not undergoing any psychiatric treatment.

Comparisons

Psychological placebo

Treatment name: Pseudodesensitisation

Description of intervention: "The imaginal stimuli used in treating this group were descriptions of 20 different snake-irrelevant and pleasant landscape scenes. This group was included to control nonspecific factors, such as expectancy and commitment on the part of 5, relationship, attention, and suggestion." (p. 334)

Individual or group treatment: not stated

Exposure/intensity to treatment: 30 minutes relaxation training and five sessions/week

Krapfl 1970 *(Continued)*

	Duration of treatment: 5 weeks **Concomitant psychotherapy:** exclusion criteria not undergoing psychiatric treatment. Otherwise not stated **Concomitant pharmacotherapy:** not stated **No-treatment** **Comparison name:** No-treatment **Description of intervention:** "These 5s received only the pretreatment, posttreatment, and follow-up test batteries which were administered to the other four groups" (p. 334) **Exposure/intensity to treatment:** none **Duration treatment** : 5 weeks **Concomitant psychotherapy:** exclusion criteria not undergoing psychiatric treatment. Otherwise not stated **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy:** observer-reported • **Outcome chosen:** the Behavioral Avoidance Test (BAT) **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. While no differences were found between 5s exposed to an increasingly aversive hierarchy and 5s who received a decreasing order, the random order tended to be less effective than the other two. An ascending aversive order of stimulus presentations is not an essential and integral part of successful desensitization. **Key limitations from study authors** 1. Not stated **Other notes from review authors** 1. Means generated from Figure 1 2. SD's generated fromRosen 1976orEtringer 1982 **Conflicts of interest:** none found **Judgement**: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Assignment of a matched 5 to one of the five groups was then done on a random basis. (p. 334)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Unclear	No information

Krapfl 1970 *(Continued)*

Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Unclear	Attrition unclear
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Kwan 2017

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Physical placebo: sham group 2. Wait-list 3. Active treatment: acupressure **Sample calculation:** yes **Cluster randomised:** no **Duration of trial (baseline to post):** 3 weeks **Duration of participation (trial + follow-up):** 3 weeks + 6 weeks follow-up **Setting:** outpatient. Nursing Home **Purpose of trial:** "(…) the purpose of the present study is to examine the effect of acupressure on agitation and salivary cortisol by testing the following null hypotheses: (1) there is no difference in the level of agitation between the acupressure group and the control groups among nursing home residents with dementia over time; and (2) there is no significant difference in the salivary cortisol level between the acupressure group and the control groups among agitated nursing home residents with dementia over time." (p. 93) **Closed or open placebo:** closed placebo
Data	**Number of participants screened:** 2014 **Number of participants included:** 121 **Number of participants followed-up at post treatment:** 118 **Number of participants randomly assigned to:** • Physical placebo: n = 41 • Wait-list: n = 40 • Active treatment: n = 40 **Number of withdrawals** : n = 3 • Physical placebo: n = 0 • Wait-list: n = 1 • Active treatment: n = 2 **Diagnosis:** dementia **Diagnostic manual:** Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision (DSM IV-4-TR) (Kwan 2017 (pers comm)) **Means of assessment:** not stated **Comorbidity:** number of chronic illness 4.1 (SD = 1.9) in the total population. Otherwise not stated **Age:** 86.5 mean years (SD = 6.3) **IQ:** not stated

Kwan 2017 *(Continued)*

Sex:: 78% female
Ethnicity: not stated
Country: China

Inclusion criteria

1. Aged over 65 years
2. Had been documented in their medical records as having dementia
3. Had been displaying agitated behaviours for at least 1 month before recruitment according to the criteria for agitation stated in the CMAI

Exclusion criteria

1. Those with skin/musculoskeletal problems on their acupoints
2. Those who had received acupuncture/acupressure within 8 weeks prior to the day of recruitment

Comparisons	**Physical placebo** **Treatment name** : Sham group + usual care Description of intervention: “The sham protocol was identical to the acupressure protocol, except for the points on which pressure was applied. (…) They applied pressure on nonacupoints that were in an adequate distance from the acupoints. (…) These were located on (1) the nasal bone, (2) the olecranon, (3) the styloid process of the ulna, (4) the medial malleolus over the ankle, and (5) the head of the fibula.” (p. 95) **Individual or group treatment:** individual **Exposure/intensity to treatment** two 10-minute acupressure sessions a day (morning session at 08.00-12.00 and an afternoon session at 14.00-18.00), 5 days per week **Duration of treatment:** 2 weeks **Concomitant psychotherapy:** received usual care provided by the RCH to manage agitated residents every day (not specified) **Concomitant pharmacotherapy:** received usual care provided by the RCH to manage agitated residents every day (not specified). Number of psychotropic drugs used in the whole sample: 1.5 (SD =1.1) **Wait-list** **Comparison name:** usual-care group (in reality wait-list) **Description of intervention:** “The participants in this group received only the usual care provided by the RCH to manage agitated residents every day, such as activity programs and the use of restraints if needed as judged by the nursing home staff. Such care was also provided to the participants in the acupressure and sham group.” (p 95) "In this group, participants receive no acupressure- related intervention. They receive a free course of acupressure sessions, identical to the one stipulated in the AG, after completing the study." (Kwan 2014 , p. 5) **Exposure/intensity to treatment:** no treatment but the usual care **Duration treatment:** 2 weeks **Concomitant psychotherapy:** received usual care provided by the RCH to manage agitated residents every day (not specified) **Concomitant pharmacotherapy:** received usual care provided by the RCH to manage agitated residents every day (not specified). Number of psychotropic drugs used in the whole sample: 1.5 (SD =1.1)
Outcomes	**Beneficial effects** • **Hierarchy:** usable data,observer-reported, clinical relevance, global score • **Outcome chosen:** locally-validated version of the CMAI **Adverse events** • Spontaneous reporting
Notes	**Key conclusion from study authors**

Kwan 2017 *(Continued)*

1. 1”(...) when the group that received acupressure was compared to the control groups, no significant difference between the groups could be observed. (…) acupoint activation is not a significantly effective component for reducing agitation.” (p. 101)
2. 2“A significant reduction in cortisol was seen in the acupressure group compared to the control groups.” (p. 101)

Key limitations from study authors

1. 1. ”The major limitation of this study was the relatively low collection rate of valid saliva samples.” (p. 102)
2. 2. “Another limitation is that only 18 RCHs (6.3%) agreed to participate out of the 284 RCHs that were invited.” (p. 102)
3. 3. “The post hoc power analysis showed that 190 participants would be needed to demonstrate the size of the effect on agitation calculated from this study compared to the 119 participant (i.e., post hoc power 0.56) that were actually recruited.” (p. 101)

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: ”The participants were allocated by permuted block randomization to 3 parallel groups in a 1:1:1 ratio (...) The permuted block randomization list was generated by the web-based generator at Randomization.com.” (p. 93)
Allocation concealment	Yes	Quote:”An independent research assistant, who did not participate in any other parts of the research and was blinded to participants’ demographics and clinical characteristics, allocated participants to groups according to the randomization list and subject codes provided by the data collectors.” (p. 93)
Blinding of outcome assessors	Yes	Quote:“Participants, RCH staff members, and data collectors who were not involved in providing any care services in the participating RCHs were blind to the group labels.”) (p. 93)
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	Yes	Quote:“Data analysis was performed using SPSS version 21.0 based on a modified intention-to-treat (mITT) principle, which means that all subjects were included after randomization except for those who withdrew from the study before undergoing the first session of the intervention. A generalized estimating equation (GEE) was used to answer the 2 hypotheses.(…) Participants with missing data points were included in the mITT as estimated by the GEE by following the missing-at-random assumption.” (p. 96) Attrition <15% (2.5%)
Selective outcome reported	Yes	CUHK_CCT003347 No apparent differences in reporting between trail registry and full report
Other sources of bias	Yes	No other sources found

Lacy 1990

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Physical placebo: self-relaxation placebo 2. Wait-list 3. Active treatment: Autogenic Training Group **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial:** 5 months in total **Duration of participation (trial + follow-up):** 5 months. Post-treatment data **Setting:** inpatient (hospital) **Purpose of trial:** "The purpose of the current investigation was to determine the effectiveness of short-term autogenic training as a stress management intervention with hospitalized emotionally disturbed adolescents" (p. 96) **Open/Closed placebo:** closed placebo
Data	**Number of participants screened:** 58 **Number of participants included:** 45 **Number of participants followed-up at post treatment:** 45 **Number of participants randomly assigned to:** • Physical placebo: n = 15 • Wait-list: n = 15 • Active treatment: n = 15 **Number of withdrawals:** n = 0 **Diagnosis:** several psychiatric disorders. **Diagnostic manual:** not stated **Means of assessment:** Clinical interview (Global Assessment of Functioning; GAF): ranged from 28 to 55 years, average 37 **Comorbidity:** " Adolescents admitted to the psychiatric hospital during the investigation exhibited various forms of dysfunctional behavior and were diagnosed with diverse psychiatric classifications. Many were determined to be a danger to self or others" (p. 59) "All subjects in the investigation were diagnosed with moderate to severe emotional problems; however, a number of subjects rated their anxiety levels quite low, below expected levels generally associated with normal adolescence." (p. 102) **Age:**: 15.22 mean years (SD = 1.38). (range= 13 to 17) **IQ:** "The mean composite intelligence score of subjects was 95, falling within the average range 63-125" (p. 61). **Sex**: 46.7% female **Ethnicity:** "Thirty-eight subjects participating in the investigation were white. Three Native Americans, three Blacks (African Americans) and one Asian American also served as subjects" (p. 61) **Country:** USA **Inclusion criteria** 1. In order to participate as a participant in the stress management program inpatient adolescents were required to have a written permission form signed by a parent or legal representative 2. Also, to be eligible to participate adolescents had to score at the fourth grade level or higher on a reading achievement test to demonstrate the ability to read items on a self-report anxiety scale used to assess stress **Exclusion criteria**

Lacy 1990 *(Continued)*

1. Actively psychotic adolescents were excluded from participating as subjects
2. Also, adolescents with reading levels determined as too limited to adequately comprehend written items on the anxiety scale did not participate as subjects
3. Only four adolescents with signed parental permission forms were not used as subjects
4. One adolescent appeared to be experiencing a drug-induced psychotic episode and three others lacked adequate reading skills

Comparisons	**Physical placebo** **Treatment name:**: a Self-relaxation placebo control group **Description of intervention:** "Subjects practicing self-relaxation assumed a similar horizontal posture under the same experimental conditions as the autogenic trainees." (p. 73) **Individual/group** : Individual treatment **Exposure/intensity to treatment:** six sessions were conducted over a two-day period following the same time frame as the autogenic training sessions. Subjects practiced self-relaxation for no longer than 10 minutes in one session **Duration of treatment:** 5 months **Concomitant psychotherapy:** "Treatment received prior to admission to the hospital varied for each subject. Most subjects had been involved in some level of counseling. Some subjects had been previously hospitalized for treatment." (p. 59) **Concomitant pharmacotherapy:** "Several subjects had histories of medication intervention. None of the subjects in the investigation were prescribed psychoactive medication during the evaluation period. Also, there were no controls for residual effects of prior medication." (p. 59) **Wait-list** **Comparison name:** No-treatment (in reality wait-list) **Description of intervention:** "Subjects in this control group participated in only pre- and posttesting measurement activity. No relaxation training was provided in the interim. Subjects were informed relaxation training was to begin with the completion of two measurement sessions. Control group subjects received brief instruction in autogenic training with the completion of posttesting. (p. 73) **Exposure/intensity to treatment:** no treatment during waiting **Duration treatment:** 5 months **Concomitant psychotherapy**: "Treatment received prior to admission to the hospital varied for each subject. Most subjects had been involved in some level of counseling. Some subjects had been previously hospitalized for treatment." (p. 59) **Concomitant pharmacotherapy:** "Several subjects had histories of medication intervention. None of the subjects in the investigation were prescribed psychoactive medication during the evaluation period. Also, there were no controls for residual effects of prior medication." (p. 59)
Outcomes	**Beneficial effects** • **Hierarchy:** observer-reported, clinical relevance, psychometric properties • **Outcome chosen:** Thought Technology Temp/SC 201 T Biofeedback System - subscale skin conductance levels) **Adverse events** • Spontaneous reporting
Notes	**Key conclusion from study authors** 1. In this investigation short term autogenic training proved an effective method of stress reduction for hospitalised emotionally-disturbed adolescents 2. Significant post-test differences were evidenced between the autogenic training group and both the self-relaxation (placebo control) and no treatment control groups 3. Autogenic training was more effective than either control group in increasing peripheral skin temperature indicative of stress reduction **Key limitations from study authors** 1. In the present investigation the residual effects of prior medication or illicit substances are unknown

Lacy 1990 *(Continued)*

2. Subjective feedback of participants concerning autogenic training is not available in the current investigation

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: A total of 45 adolescents were randomly assigned to participate as subjects in one of the three experimental conditions." (p. 62)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Unclear	Quote: "No other assistance in completing the scale was provided. The scale was not scored by the investigator until the completion of posttesting." (p. 65)
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	Yes	Quote: "Attrition was not a problem as all 45 subjects complied with all treatment and measurement demands." (p. 62) Attrition < 15% (0%)
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Lai 2004

Study characteristics

Methods

Parallel randomised trial with three arms

1. Psychological placebo: comparison group
2. No-treatment: control group
3. Active treatment: intervention group

Sample calculation: yes

Cluster randomised: no

Duration of trial (baseline to post): 6 weeks

Duration of participation (trial + follow-up): 6 weeks. No follow-up

Setting: inpatient (nursing homes)

Purpose of trial: "This study is a randomized controlled trial aimed at finding out whether a specific reminiscence program would lead to any changes in social well-being for nursing home residents with dementia." (p. 34)

Lai 2004 *(Continued)*

	Open/closed placebo:: closed placebo
Data	**Number of participants screened:** 127 **Number of participants included:** 101 **Number of participants followed-up at post treatment:** 89 **Number of participants randomly assigned to** • Psychological placebo: n = 35 • No-treatment: n = 30 • Active treatment: n = 36 **Number of withdrawals** : n =12 • Psychological placebo: not stated • No-treatment: not stated • Active treatment: not stated **Diagnosis:** dementia **Diagnostic manual:** Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) **Means of assessment:** "The physicians in charge of the homes confirmed that those residents diagnosed as having dementia met the criteria as specified by the DSM-IV." (p. 38) **Comorbidity:** number of other diagnoses: psychological placebo: 3.7 (SD=1.7). No-treatment 4.5 (SD = 1.9). Not specified which **Age:** 85.45 mean years (SD = 7.35) **IQ:**: not stated, but years of education included **Sex:**: 77.5% female **Ethnicity:** not stated **Country**: China/USA **Inclusion criteria** 1. Participants included were residents diagnosed as suffering from dementia (DSM-IV) 2. Able to communicate most of the time (according to the Resident Assessment Instrument [RAI] communication scale) 3. Able to understand and speak Cantonese. **Exclusion criteria** 1. Excluded were residents with any active major psychiatric disorders (schizophrenia, major affective disorders) 2. Any acute or unstable chronic medical conditions including cardiac or lung diseases. 3. Blindness (RAI – vision scale) 4. Inability to hear even with hearing aids (deafness) (RAI – hearing scale)
Comparisons	**Psychological placebo** **Treatment name :** Comparison group **Description of intervention:** "To control for the possibility that a resident's improvement might have been the result of the attention and social contacts resulting from the intervention itself, the comparison program was designed to provide social contacts." (p. 35).

Lai 2004 *(Continued)*

Individual or group treatment:: individual

Exposure/intensity to treatment: a 30-minute session weekly

Duration of treatment: 6 weeks

Concomitant psychotherapy: both groups receive their regular group sessions (Nursing home resident). It consisted of exercise, card games etc.

Concomitant pharmacotherapy: not stated but regular medicine

No-treatment

Comparison name: Control group (No treatment)

Description of intervention: "Subjects assigned to the control group received no intervention." (p. 36)

Exposure/intensity to treatment: no intervention

Duration treatment:: 6 weeks

Concomitant psychotherapy: both groups received their regular group sessions (Nursing home resident). It consisted of exercise, card games etc.

Concomitant pharmacotherapy: not stated, but regular medicine

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported, clinical relevance
- **Outcome chosen:** WIB-Well-being/Ill-being Scale

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. Although the intervention did not lead to significant differences between the three groups over time, there was a significant improvement in psychosocial well-being for the intervention group

Key limitations from study authors

1. First, the sample size was too small for a repeated-measures multivariate analysis. Only two nursing homes were used as study sites. Indeed the use of a few study settings facilitated standardised sampling, data collection, protocol adherence and control for a number of confounding variables; however, it also posed restrictions on the adequate recruitment of subjects
2. Second, the "dosage" of the intervention might have been weak. The intervention program consisted of only six 30-minute weekly sessions
3. Third, regardless of the number of precautionary steps that had been exercised, it would be impossible to prevent people from having preconceived notions about the intervention and comparison programs
4. Last, it was likely that the measures were not sensitive enough

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Lai 2004 *(Continued)*

Item	Authors' judgement	Support for judgement
Random sequence generation	No	Quote: "It was also unclear whether a random assignment was used in group allocation" (p. 46)
Allocation concealment	Unclear	Concealed from authors Chi & Kayser-Jones, but not Lai
Blinding of outcome assessors	Yes	Quote: "The group of RAs who collected data on the participants included both raters (who rated only the WIB of the DCM) and assessors (who performed the rest of the assessments), and both groups were blinded to subject assignment." (p. 37)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Quote: " In the ITT sample, (Table 1) the percentage of data missing for the outcome variables of T0, T1, and T2 was 0.5%, whereas the missing data for two controlling variables, the MMSE and the MDS-ADL, was 0.9%. Missing data for all other variables constituted 1.1%. In total, 2.5% of the data was missing in this dataset. For the per protocol sample, the percentages of missing data for the outcome variables was 0.2%, the MMSE and the MDS-ADL–0.5%, and all other variables –0.8%. The total percentage of missing data for the per protocol sample was 1.5%. The mean value of the outcome variables for each respective group was used as a replacement for the missing data." (p. 39) Attrition <15% /(2.5%). ITT used
Selective outcome reported	Unclear	Not registered beforehand (Lai 2004 (pers comm))
Other sources of bias	Unclear	Quote: "Concerning the per protocol population, significant differences were found between the control and comparison group in the number of medical diagnoses other than dementia and in whether they had any regular programs. It was difficult to explain the meaning of the differences, as these two groups were not significantly different in terms of their baseline (T0) and T1, C-MMSE and MDS-ADL scores." (p. 45).

Lang 1965

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Psychological placebo: pseudotherapy 2. No-treatment: untreated controls 3. Active treatment: desensitisation **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 8-16 weeks **Duration of participation (trial + follow-up):** 8-16 weeks **Settin:** : outpatient (University setting)

Lang 1965 *(Continued)*

Purpose of trial: "44 snake phobic Ss participated in laboratory experiments assessing the degree of fear change associated with systematic desensitization, no treatment, placebo treatment, and the trait of suggestibility" (p. 395)

Open/closed placebo: closed placebo

Data

Number of participants screened: not stated

Number of participants included: 44

Number of participants followed-up at post treatment: 44

Number of participants randomly assigned to:

- Psychological placebo: n = 10
- No-treatment: n = 11
- Active treatment: n = 23

Number of withdrawals: n = 0

Diagnosis: specific anxiety (snake phobia)

Diagnostic manual: not stated

Means of assessment: "They rated their fear of nonpoisonous snakes as "intense," on a fear questionnaire, and were included in this research only if a psychological interview corroborated this statement." (p. 397)

Comorbidity: not stated

Age:: not stated – Introductory psychology students

IQ:: not stated, Introductory psychology students

Sex:: not stated

Ethnicity:: not stated

Country: USA

Inclusion criteria

1. Specific anxiety towards snakes

Exclusion criteria

1. Participants who appeared to have impairing physical disabilities
2. Latent psychosis (based on the psychotic scales of the MMPI or the clinical judgment of the interviewer)

Comparisons

Psychological placebo

Treatment name: Pseudotherapy

Description of intervention: "An effort was made to involve the subject in a treatment procedure, which was therapeutically neutral except for the therapist-client relationship. Because desensitization was to be evaluated, all procedures employed in that method were included in pseudotherapy." (p. 396)

Individual or group treatment: Individual

Exposure/intensity to treatment: the participants first experienced the same 5 training sessions as in desensitisation, followed by 11 pseudotherapy sessions.

Lang 1965 *(Continued)*

Duration of treatment : 16 sessions (8-16 weeks) “Sessions lasted approximately 45 minutes at the rate of 1 or 2 per week” (p. 396)

Concomitant psychotherapy : None of the participants in this study were being seen elsewhere because of psychological problems.

Concomitant pharmacotherapy: not stated

No-treatment

Comparison name:: No-treatment

Description of intervention: “Untreated subjects were not seen, except for evaluation sessions.” (p. 397)

Exposure/intensity to treatment: no treatment

Duration treatment:: not stated

Concomitant psychotherapy: none of the participants in this study were being seen elsewhere because of psychological problems

Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect

- **Hierarchy:** Observer-reported
- **Outcome chosen:** Avoidance test (observer-reported)

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. Desensitisation Ss showed significantly greater fear reduction than controls, while placebo Ss changed no more than did untreated Ss.
2. Successful desensitisation was relatively independent of suggestibility.
3. Desensitisation of specific fears generalised positively to other fears, and among desensitization Ss, degree of fear change could be predicted from measurable aspects of therapy process.

Key limitations from study authors

1. The current experiment was limited to a brief 11 desensitisation sessions, and it is, of course, possible that all subjects would have improved with a sufficient exposure.
2. generalisation to the clinic must be cautiously undertaken. Many issues are raised that need more intensive investigation

Other notes from review authors

1. No usable data
2. Author pools placebo and no-treatment

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: ”Assignment to groups was essentially random, although some pretreatment effort to balance control variables was made. A more elaborate de-

Lang 1965 *(Continued)*

		scription of the selection battery has already been reported (Lang 1963).” (p. 397)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Unclear	No information
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Attrition <15% (0%)
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Legrand 2016

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Physical placebo: stretching 2. No-treatment: no intervention 3. Active treatment: aerobic exercise **Sample calculation:** yes **Cluster randomised:** no. **Duration of trial (baseline to post):** 10 days. **Duration of participation (trial + follow-up):** 10 days (no follow-up) **Setting:** inpatient. **Purpose of trial:** “(…) in the present study we examined the efficacy of a 10-days long aerobic exercise program as an add-on treatment in severely depressed patients being treated with antidepressant medication (and no other form of therapy) for less than two weeks.” (p. 140) **Closed/open placebo:**: closed placebo
Data	**Number of participants screened:** 124 **Number of participants included:** 35 **Number of participants followed-up at post treatment:** 31 **Number of participants randomly assigned to:** 1. Physical placebo: n = 11 2. No-treatment: n = 10 3. Active treatment: n = 14 **Number of withdrawals** : n = 4 1. Physical placebo: n = 2 2. No-treatment: n = 1 3. Active treatment: n = 1 **Diagnosis:** Major depressive disorder (MDD) **Diagnostic Manual:** Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision (DSM-IV-TR)

Legrand 2016 *(Continued)*

Means of assessment: "Patients were assessed using the Beck Depression Inventory (BDI-II)" (Legrand 2016 (pers comm))
Comorbidity : not stated
Age:: 45.3 mean years (SD = 13.2)
IQ: not stated
Sex: 71.4% female
Ethnicity: not stated
Country:: France

Inclusion criteria

1. Diagnosis of MDD according to the DSM-IV-TR
2. Antidepressant drug therapy initiated for < two weeks
3. Score of 29 or more on the Beck Depression Inventory
4. Ability to run or walk briskly and to understand written French

Exclusion criteria

1. Had a medical contraindication for exercise practice
2. Had MDD with psychotic features
3. Receiving beta-blocking drugs or another form of therapy (e.g. sleep deprivation, electroconvulsive therapy)

Comparisons

Physical placebo
Treatment name: Placebo Stretching exercise
Description of intervention: "Patients in the stretching (ST) group also performed a daily 30 min exercise program for 10 consecutive days, but this consisted of stretching exercises instead of endurance training. Several muscle groups (thighs, calves, gluteal, shoulders, back) were stretched for 60 s. with equivalent resting intervals between stretching series. Training sessions were carried out in a room of the hospital restricted to these activities and were also supervised by the first author." (p. 140-2)

Individual or group treatment: "(…) the format of delivery was mostly individual (only 4 of the 85 stretching sessions included 2 patients)." (p. 141)
Exposure/intensity to treatment: 30 minutes daily
Duration of treatment: 10 days
Concomitant psychotherapy: no
Concomitant pharmacotherapy: antidepressant medication

No-treatment
Comparison name: No-intervention (no treatment).
Description of intervention: "(…) participants in the control (NI) group received no intervention other than the prescribed medication." (p. 141)
Exposure/intensity to treatment : no treatment other than the prescribed medication
Duration treatment: 10 days
Concomitant psychotherapy: no
Concomitant pharmacotherapy: antidepressant medication

Outcomes

Beneficial effects

- Hierarchy: usable outcome, patient-reported
- Outcome chosen: Beck Depression Inventory (BDI)

Adverse events

- Spontaneous reporting

Notes

Key conclusion from study authors

1. "Our results indicate that a short endurance training intervention can lead to a substantial reduction of depressive symptoms in hospitalized patients with severe depression, with more than 50% of exercisers achieving a more than 50% reduction in their BDI-II score." (p. 142)

Legrand 2016 *(Continued)*

2. "We found an average depression score decrease of 47.6% in the AE group (vs 24.8% in the ST group and 18.0% in the NI group)." (p. 142)

Key limitations from study authors

1. "A first important limitation that should be considered when interpreting the results is the lack of blinding of study participants to interventions." (p. 143)
2. "A second limitation is that no individual information on dosage of antidepressant drugs was available." (p. 143)
3. "Third, this study was a short-term (10 days) RCT with no follow-up. Therefore, no conclusion can be drawn about the duration of the antidepressant effect of exercise reported here (…) Moreover, it may be that the sudden interruption of the exercise program resulted in a rebound effect with an increase in depressive symptoms." (p. 143)
4. "A fourth limitation of our study is that participants in the AE and ST groups exercised under close supervision (in order to standardize and control exercise intensity or movement execution), which resulted in a lengthy interaction time between patients from these groups and the first author." (p. 143)
5. "Finally, because patients in the AE group exercised outdoors and those in the ST group exercised indoors, daylight exposure could explain or partly explain our finding." (p. 143)

Other notes from review authors

1. "Adverse events in the AE group included transient muscular/joint soreness (n = 3),headache (n = 1), and fatigue (n = 2)." (p. 141) No adverse events mentioned for stretching and no treatment."Indeed adverse events were scrutinized for participants in both groups (Aerobic Exercise, Stretching, No Intervention). Some minor problems were detected (headaches, fatigue)... but this was seen only among patients in the Aerobic Exercise group." (author correspondence)

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: "One of the three study arms (aerobic exercise, stretching, no intervention) was randomly chosen for each participant at the end of an initial individual visit (…) This was done by running the rand between function of Microsoft Excel on our laptop, which generated a random number between 1 and 3: 1=aerobic exercise (AE), 2=stretching (ST), 3=no intervention." (p. 140)
Allocation concealment	Yes	Quote: "Yes..... A can be understood from our text (p. 140, right-hand column, 3rd paragraph) we did not know whether the next study patient would receive Aerobic Exercise, Stretching (sham), or No Intervention;...Consequently, it is possible to state that our random allocation procedure was concealed." (Legrand 2016 (pers comm))
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Quote: "These analyses were performed using an intent-to-treat (ITT) approach in which all patients with baseline measures were included in analyses, even if they missed more than two training sessions. Missing data were imputed using the average change in depression score from baseline to post-intervention in the control group. When drop-out rates are less than 20% (which

Legrand 2016 *(Continued)*

		is the case in our study), this method keeps statistical power at higher levels compared to the last-observation-carried-forward method." (p. 141) Attrition <15% (11.5%)
Selective outcome reported	Yes	NCT02612142 No apparent differences in reporting between trial registry and full report
Other sources of bias	Yes	No other sources found

Lick 1975

Study characteristics	
Methods	**Parallel randomised trial with four arms** 1. Psychological placebo 2. Wait-list 3. Active treatment 1: Sytematic desensitisation 4. Active treatment 2: Placebo with feedback **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial:** 6 weeks **Duration of participation (trial + follow-up):** 6 weeks. 4-week follow-up, and 4-month follow-up **Setting:** outpatient **Purpose of trial:** "The investigation reported here compares two versions of (...) placebo manipulation with Wolpian systematic desensitization in the modification of snake and spider fear in volunteer adult subjects who manifest real-life inhibitions as a result of their fear." (p. 558) **Open/Closed placebo** : closed placebo
Data	**Number of participants screened:** 48 **Number of participants included:** 36 **Number of participants followed-up at post treatment:** post-treatment not stated. (31 after 4 months) **Number of participants randomly assigned to:** • Psychological placebo: n = 9 • Wait-list: n = 9 • Active treatment 1: n =9 • Active treatment 2: n = 9 **Number of withdrawals:** post-treatment: not stated, (5 in follow-up) **Diagnosis:** specific anxiety (snakes & spiders) **Diagnostic manual:** not stated **Means of assessment:**: clinical interview **Comorbidity:** not stated

Lick 1975 *(Continued)*

Age: 29.77 mean years (range =18 to 59)

IQ: not stated

Sex: 100% females

Ethnicity: not stated

Country: USA

Inclusion criteria

1. Evidence during the initial interview that the subject manifested fairly intense behavioural inhibitions regarding snakes or spiders, i.e. the person could not go camping, clean closets, or kill spiders
2. Failure to reach into an aquarium containing a live snake or spider with a gloved hand during a behavioral pretest
3. Agreeing to receive, after hearing a brief description, any of the three treatments being evaluated or no treatment at all if "therapist time proved to be inadequate"
4. Submitting a $20 deposit to be returned after treatment and assessments were completed

Exclusion criteria

- Not stated

Comparisons	**Psychological placebo** **Treatment name:** Placebo **Description of intervention: "**This treatment was exactly the same as placebo feedback described above except that (a) the subject was not shown her GSR printouts, nor did the therapist comment about them at any point in treatment. If the subject asked about her GSR printouts, the therapist responded that they were not interpretable until analyzed by the computer, (b) The subject was also told that the equipment reduced the number of shocks it delivered on each session automatically and independently of the subject's responses" (p. 559-560) **Individual or group treatment:** individual **Exposure/intensity to treatment:** 20 minutes per session (except first one) **Duration of treatment:** 8 biweekly sessions for 6 weeks **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated **Wait-list** **Comparison name:** No-treatment (in reality wait-list) **Description of intervention:** "The subjects in this group were sent a letter stating that because of staff shortages they could not be treated at this time but that they would receive treatment in several months when therapist time became available." (p. 560) **Exposure/intensity to treatment:** no treatment during the period **Duration treatment:** 6 weeks **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effects** • Hierarchy: observer-reported • Outcome chosen: behavioral approach **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. Overall, the results of this experiment were interpreted as contradicting a traditional conditioning explanation of systematic desensitisation

Lick 1975 *(Continued)*

2. An alternate explanation for the operation of systematic desensitisation emphasising the motivational as opposed to conditioning aspects of the procedure is discussed

Key limitations from study authors

1. Validity (two therapists were students and not clinical experienced). The third one was the author
2. A more serious problem with the systematic desensitisation procedure was treatment duration

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "(…) the subjects were stratified into three blocks of four snake phobics and six blocks of four spider phobics. One subject from each block was then randomly assigned to each of the four conditions, and the three therapists were randomly assigned one snake phobic and two spider phobics from each condition." (p. 559)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote:"The pretreatment Behavioral Approach Test was administered by the author; the posttreatment Behavioral Approach Test was administered by another male experimenter who was blind about the conditions to which subjects had been assigned". (p. 560)
Blinding of participants and personnel	No	Not possible to blind placebo and "no-treatment"
Incomplete outcome data	Unclear	Attrition unclear. 5 were lost to follow-up. Unclear how many that were lost to post-treatment.
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	No	Author (J.L.) was a therapist himself in the study

Lick 1977

Study characteristics

Methods	**Parallel randomised trial with four arms** 1. Physical placebo 2. No-treatment 3. Active treatment 1: relaxation 4. Active treatment 2: relaxation plus tape **Sample calculation:** not stated **Cluster randomised:** no

Lick 1977 *(Continued)*

Duration of trial (baseline to post): 6 weeks
Duration of participation (trial + follow-up): 6 weeks of treatment + 1 month follow-up
Setting : Outpatient. University setting
Purpose of trial: "The present study is designed to further evaluate the relative efficacy of relaxation training and attention placebo procedures in the treatment of severe insomnia." (p. 154)
"Finally, this study assesses whether giving people a tape with relaxation instructions to play in bed before retiring adds to the efficacy of relaxation training conducted in the clinic." (p. 154)

Open/closed placebo: closed placebo

Data

Number of participants screened: 65
Number of participants included: 40

Number of participants followed-up at post treatment: 35
Number of participants randomly assigned to :

- Physical placebo: n = 10
- No-treatment: n = 10
- Active treatment 1: n = 10
- Active treatment 2: n = 10

Number of withdrawals: 5 – "Subjects approximately equivalent in age, sex, and time to fall asleep were substituted for these latter subjects." (p. 155)

- Physical placebo: not stated
- No-treatment: not stated
- Active treatment 1: not stated
- Active treatment 2: not stated

Diagnosis: Insomnia
Diagnostic manual : none
Means of assessment: averaging 50 minutes or longer to fall asleep during a 20-day pre-tratment baseline period
Comorbidity: not stated
Age: 47.48 mean years (SD = 10.88), (range = 29 to 72 years)
IQ: not stated
Sex: 65% female
Ethnicity: not stated
Country: USA

Inclusion criteria

1. Averaging 50 minutes or longer to fall asleep during a 20-day pretreatment baseline period
2. Agreeing to receive any of several treatments or no treatment at all "if therapist time proved to be inadequate"
3. Submitting a $20 deposit to be returned after treatment and assessments were completed

Exclusion criteria

1. Severe physical difficulties or psychotic behaviour

Comparisons

Physical placebo
Treatment name: Placebo
Description of intervention: "The placebo procedures are a variant of those found credible and effective in the treatment of phobic behavior and were expected to work only because of placebo and other nonspecific factors." (p. 154). "This procedure was a modified form of "T-scope therapy". Subjects were told that this procedure was designed to reduce autonomic arousal that was incompatible with sleep. Subjects were told that their autonomic arousal would be monitored with a polygraph and that whenever the polygraph detected a "high-intensity autonomic response," it would trigger a shock generator that would deliver a mildly unpleasant shock to the subject's left index finger (…) Subjects received

Lick 1977 *(Continued)*

	a decreasing number of shocks over the six sessions (…) Over the six sessions these printouts showed less and less autonomic activity (…)" (p. 155) **Individual or group treatment:** individual. **Exposure/intensity to treatment:** six sessions **Duration of treatment:** 6 weeks. **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** participants were allowed to take sleep-inducing drugs **No-treatment** **Comparison name:** No-treatment **Description of intervention:** no treatment – no other description of intervention. **Exposure/intensity to treatment:** no treatment. **Duration treatment:** 6 weeks **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** participants were allowed to take sleep-inducing drug
Outcomes	**Beneficial effects** • Hierarchy: primary outcome, self-reported, clinical relevance • Outcome chosen: time to fall asleep or observer-rated sleep postcards. **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. "The major finding from this investigation was the superiority of the two relaxation training procedures over placebo and no-treatment control groups (…) Indeed, the present study found that relaxation training influenced several sleep parameters in addition to the latency-of-sleep-onset measure" (p. 159) 2. "There is also evidence from this study that relaxation training can influence self-report measures of trait anxiety in chronic insomniacs and reduce their consumption of sleep-inducing drugs, a finding of particular significance given reports of massive barbiturate abuse in our society." (p. 159) 3. "Somewhat surprisingly, there was no difference in the effectiveness of the relaxation and relaxation-plus-tape treatments. Posttreatment interviews suggested general satisfaction with and use of the relaxation tape, but this adjunct did not influence any of the dependent measures used." (p. 159) 4. "The ineffectiveness of the placebo manipulation relative to relaxation training and no-treatment control (…) suggest that improvement produced by relaxation training procedures in the treatment of severe insomnia is not mediated primarily by expectation of therapeutic gain or other nonspecific influences." (p. 160) **Key limitations from study authors** 1. "Demonstrating that a treatment procedure produces statistically significant improvement on outcome measures relative to control procedures does not imply that the treatment has clinical utility. To make a determination about the clinical value of treatment effects, the researcher should optimally have information about (a) the relationship between the dependent measures used to assess treatment outcome and the distressing target behaviors that subjects want to modify, (b) how "normal" subjects perform on these dependent measures, and (c) subjects' degree of satisfaction with the changes experienced in treatment." (p. 161) **Other notes from review authors** 1. None **Conflicts of interest:** none found **Judgement:** yes

Risk of bias

Lick 1977 *(Continued)*

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: ”Within each wave, subjects were blocked according to sex, time to fall asleep, and whether they took sleeping medication. They were then randomly assigned to the following conditions: (a) progressive relaxation; (b) progressive relaxation plus taped relaxation; (c) placebo control; or (d) no-treatment control. (p. 155)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Quote: "(…) 5 subjects dropped out of treatment; 2 dropped out because they repeatedly missed appointments, and 3 dropped out because of severe emotional or physical difficulties unrelated to treatment. Subjects approximately equivalent in age, sex, and time to fall asleep were substituted for these latter subjects.” (p. 155) Attrition <15% (12.5%)
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Liddle 1990

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Psychological placebo: attention placebo 2. No treatment 3. Active treatment: social competence training **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial:** 8 weeks **Duration of participation (trial + follow-up):** 8 weeks + 5 week follow-up **Setting:** Outpatient **Purpose of trial:** “The present study differs from other published studies cited in combining both a follow-up assessment and a non-specific treatment control group which will make it possible to determine if any beneficial effects are specific to the treatment.” (p. 88) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** 380 **Number of participants included:**: 33

Liddle 1990 *(Continued)*

Number of participants followed-up at post treatment: 31

Number of participants randomly assigned to: 31

- Psychological placebo: n = 10
- No treatment: n = 10
- Active treatment: n = 11

Number of withdrawals: n = 2

- Psychological placebo: not stated
- No treatment: not stated
- Active treatment: not stated

Diagnosis: depression

Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition (DSM-III)

Means of assessment: structured clinical interview

Comorbidity: not stated

Age:: 9.2 mean years (SD =1.15)

IQ: not stated, but intellectually handicapped were excluded

Sex:: 36.4% female

Ethnicity: not stated

Country:: Australia

Inclusion criteria

1. Children were aged from 7 to 12 years, were enrolled in mainstream classes
2. and were fluent in the English language

Exclusion criteria

1. Intellectually handicapped children were excluded from the study

Comparisons

Psychological placebo

Treatment name: attention placebo

Description of intervention: "The APC group consisted of a drama programme which was adapted by the experimenter from one devised by Milneand Spence (1987) from educational literature on the teaching of drama to primary school children (...). Drama was selected for the APC because the following elements were shared with the social competence condition: withdrawal from class during school time; small group interaction;an equal amount of time and attention and homework assignments. Nospecific skill-based instruction regarding the perception of interpersonal cues or methods of dealing with interpersonal situations was given. There is no evidence that drama sessions per se should improve social competence" (p. 93)

Individual or group treatment: group

Exposure/intensity to treatment:: eight, weekly, one-hour group sessions

Duration of treatment: 8 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

No-treatment

Liddle 1990 *(Continued)*

Comparison name : No-treatment

Description of intervention: “Children in this condition did not participate in any "special activities". They attended regular school classes and were withdrawn for the pre-, post- and follow-up assessments only, on the pretext that they were helping the University with research into "how children feel about things".” (p. 94)

Exposure/intensity to treatment: no treatment

Duration treatment: 8 weeks of no treatment

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy:: not stated

Outcomes

Beneficial effect

- **Hierarchy:**: usable data, post-treatment, patient-reported, clinical relevance
- **Outcome chosen:** Children's Depression Inventory

Adverse events

- No data on adverse events reported

Notes

Key conclusion from study authors

1. The results indicated a decline in depression scores during the treatment period for subjects in all conditions, and this continued during the two month follow-up period to within the normal range
2. The SCT programme did not produce significantly greater reductions in depression than either the APC or NTC conditions and was not effective in producing improvements on measures of social competence

Key limitations from study authors

1. In terms of age, the children in the (...) another study were several years older than those involved in the present programme, making comparability of data rather dubious.
2. Furthermore, the programme of eight, once-weekly sessions was not as long or as intense as those used in previous studies
3. The issue of group versus individual therapy for depression also requires consideration

Other notes from review authors

1. Only report data on completers

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "The children were then assigned (it is unclear if this was random) to one of four conditions, namely role-play to train interpersonal problem solving skills, cognitive restructuring, attention—placebo and classroom control. " (p. 86)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes

Liddle 1990 *(Continued)*

Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Attrition <15% (0% during treatment)
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Matson 1980

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Usual care: standard treatment 2. No-treatment 3. Active treatment: independence training **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 6 weeks **Duration of participation (trial + follow-up):** 6 weeks + 8 weeks (14 weeks) follow-up **Setting** : outpatient **Purpose of trial:** "The present study was designed to compare variants of self-control, self-evaluation, and self-reinforcement (independence training) to both a conventional treatment and a no-treatment control group. (p. 488)
Data	**Number of participants screened:** 164 **Number of participants included:** 75 **Number of participants followed-up at post treatment:** 75 **Number of participants randomly assigned to:** • Usual care: n = 25 • No-treatment: n =25 • Active treatment: n = 25 **Number of withdrawals** : n = 0 **Diagnosis:** mental retardation **Diagnostic manual:** not stated **Means of assessment:** not stated, but range of retardation was based on administration of Stanford-Binet Intelligence Scale, Form L-M, and the American Association on Mental Deficiency Adaptive Behavior Scale **Comorbidity:** not stated **Age:** 33.9 mean years (range= 22 to 57)

Matson 1980 *(Continued)*

IQ: not stated, but participants were mentally retarded. Moderate to profound range

Sex: not stated

Ethnicity: not stated

Country: USA

Inclusion criteria

1. Mental retardation

Exclusion criteria

1. Not stated

Comparisons	**Usual care** **Treatment name:** standard-treatment group **Description of intervention:** "The standard treatment consisted of procedures of choice for training of self-help skills to retarded persons, determined by the frequency of cited research articles in the training of self-help behavior. Training methods included verbal prompts, modeling, manual guidance, social reinforcement, shaping, fading, and chaining." (p. 491) **Individual or group treatment:** individual **Exposure/intensity to treatment:** five training session weekly. "Each subject received five training sessions over a one-week period." (p. 491) **Duration of treatment:** 6 weeks **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated **No-treatment** **Comparison name:** No-treatment control group **Description of intervention:** "Target behaviors for these subjects was measures but not trained. The target behaviors were recorded at time periods in correspondence with the treatment groups without delivery of instructions, feedback, reinforcement, etc., specific to the behavior of being measured." (p. 491) **Exposure/intensity to treatment:** no treatment **Duration treatment:** 6 weeks **Concomitant psychotherapy:** not stated, but possibly part of ambulatory treatment **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy:** observer-reported • **Outcome chosen:** steps on target behaviours passed **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors**

Matson 1980 *(Continued)*

- "Independence training was the most effective procedure. No improvement was observed for the no-treatment control subjects, and persons in the standard-training group improved only gradually. The relative lack of effectiveness with the standard treatment was somewhat puzzling. Given data on past performance of residents at the facility where the study was carried out and from previous studies (...), we hypothesized that the standard procedure was an effective method but that treatment may not have been sustained for a long enough time period to produce the desired results." (p. 493)

Key limitations from study authors

- The relative lack of effectiveness with the standard treatment was somewhat puzzling. Given data on past performance of residents at the facility where the study was carried out and from previous studies (...), we hypothesized that the standard procedure was an effective method but that treatment may not have been sustained for a long enough time period to produce the desired results." (p. 493)
- In addition, given the long-standing nature of the problem areas that were trained, it was apparent that desired changes would be slow and that some persons would not meet optimum training levels. (p. 493)

Other notes from review authors

- Usable data not available

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Twenty-five subjects were randomly assigned to each of three conditions: no-treatment control, standard treatment, and independence training." (p. 491)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote:"(…) and waiting until data collection had been completed before debriefing raters on the purpose of the study." (p. 490)
Blinding of participants and personnel	No	Not possible to blind usual care and no-treatment
Incomplete outcome data	Yes	Quote:"Group mean pre-, post-, and follow-up scores for the 75 subjects who completed the experiment (25 in each group) are shown in Figure 1." (p. 492) Attrition <15% (0%)
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

McLachlan 1991

Study characteristics

Methods	**Parallel randomised trial with three arms** 1. Pharmacological placebo

McLachlan 1991 *(Continued)*

2. No-treatment
3. Active treatment: desferrioxamine

Sample calculation: not stated
Cluster randomised: no
Duration of trial (baseline to post): 24 months
Duration of participation (trial + follow-up): 24 months (no follow-up)
Setting: inpatient (nursing home)
Purpose of trial: "We have completed a two year, singleblind study to investigate whether the progression of dementia could be slowed by the trivalent ion chelator, desferrioxamine." (p. 1304)
Closed/open placebo: closed placebo

Data

Number of participants screened: 1510
Number of participants included:: 63
Number of participants followed-up at post treatment: 41
Number of participants randomly assigned to:

- Pharmacological placebo: n = 9
- No-treatment: n = 14
- Active treatment: n = 25

Number of withdrawals: n = 22

- Pharmacological placebo: not stated
- No-treatment: not stated
- Active treatment: not stated

Diagnosis: Alzheimer's disease
Diagnostic manual: not stated
Means of assessment: brain tissue + cognitive tests
Comorbidity: not stated
Age: 63.1 mean years (SD = 6.3)
IQ: 79.2 to -81.4
Sex: 39.7% female
Ethnicity: not stated
Country: USA

Inclusion criteria

1. Inclusion criteria for probable AD were those of a work group established by the National Institute of Neurological and Communicative Disorders and Stroke, and the Alzheimer's Disease and Related Disorders Association ll and included computed tomographic scans that showed atrophy and no infarcts, together with Hachinski ischaemic scores 12 of four or less.
2. Identification of the principal caregiver was essential.

Exclusion criteria

1. History of myocardial infarction, angina pectoris, tachycardia, or a sustained systolic pressure of greater than 170 mm Hg, diastolic pressure greater than 100 mm Hg, or a history of congestive heart failure
2. History of hyperthyroidism, hypothyroidism, or thyrotoxicosis
3. History of schizophrenia, manic-depressive psychosis, involutional depression, electroconvulsive therapy, a current diagnosis of a psychosis, endogenous depression, paranoid features, post-traumatic brain syndrome, post-infection brain disease, cerebral neoplasm, mental retardation, alcoholic brain syndrome, tardive dyskinesia, or dementia accompanied by cerebral infarcts
4. History of psychoactive chemotherapy in doses comparable with or more than 200 mg/day of chlorpromazine
5. History of epilepsy or intake of anticonvulsant medication
6. Clinically important liver or renal disease

McLachlan 1991 *(Continued)*

7. Glaucoma, blindness, deafness, and language difficulties other than those associated with Alzheimer's disease, or any other disability sufficiently severe to prevent the subject from participating in all evaluation measurements
8. Clinically significant chronic respiratory insufficiency
9. Diabetes requiring insulin, significant anaemia, positive serological tests for syphilis or AIDS, and pathological conditions such as malignant tumours or metastases
10. Malnutrition, emaciation, malabsorption syndrome, or a history of peptic ulcer within the last three years
11. Participants from whom informed consent could not be obtained, who were participating in another study, or who do not have a caregiver living in the same residence with the participant

Comparisons	**Pharmacological Placebo** **Treatment name:** placebo **Description of intervention:** oral lecithin **Individual or group treatment:** Individual **Exposure/intensity to treatment:** 500 mg lecithin twice daily, **Duration of treatment:** 24 months **Concomitant psychotherapy:** not stated – but lived in nursing home **Concomitant pharmacotherapy:** not stated – but lived in nursing home **No-treatment** **Comparison name:** No-treatment **Description of intervention:** not stated **Exposure/intensity to treatment:** not stated **Duration treatment:** 24 months **Concomitant psychotherapy:** not stated – but lived in nursing home **Concomitant pharmacotherapy:** not stated – but lived in nursing home
Outcomes	**Beneficial effect** • **Hierarchy:** observer-reported • **Outcome chosen:** activities of daily living (ADL, rate of decline of video recorder home-behavioural assessment) **Adverse events** • Spontaneous reporting
Notes	**Key conclusion from study authors** 1. No significant differences in baseline measures of intelligence, memory, or speech ability existed between groups. Activities of daily living were assessed and video recorded at 6, 12, 18, and 24 month intervals 2. There were no differences in the rate of deterioration of patients receiving either placebo or no treatment. Desferrioxamine treatment led to significant reduction in the rate of decline of daily living skills as assessed by both group means (P = 003) and variances 3. The mean rate of decline was twice as rapid for the no-treatment group. Appetite (n = 4) and weight (n = 1) loss were the only reported side-effects 4. We conclude that sustained administration of desferrioxamine may slow the clinical progression of the dementia associated with AD. **Key limitations from study authors** 1. "Some participants may not have had AD since the diagnostic criteria did have limitations. The accuracy of our clinical diagnosis of AD was 87%, based on the necropsy results of 70 brains from a separate unpublished study. The diagnostic accuracy for participants in the present study was likely to be similar and hence up to 6 of 48 patients may not have had AD. This uncertainty would increase the likelihood of a type II error, and would reduce the apparent efficacy of the drug" (p. 1307) **Other notes from review authors**

McLachlan 1991 *(Continued)*

1. ”Some participants may not have had AD since the diagnostic criteria did have limitations. The accuracy of our clinical diagnosis of AD was 87%, based on the necropsy results of 70 brains from a separate unpublished study. The diagnostic accuracy for participants in the present study was likely to be similar and hence up to 6 of 48 patients may not have had AD. This uncertainty would increase the likelihood of a type II error, and would reduce the apparent efficacy of the drug” (p. 1307)

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: ”Those eligible to enter the study were assigned to one of three groups from a table of random numbers:” (p. 1305)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote: ”Test results collected by the behaviour evaluation team were not revealed to the medical management team; the behaviour evaluation team remained blind to the group to which participants were assigned until statistical analysis began.” (p. 1305) Quote: “Subsequent analysis was by trained raters who were not told about the nature of the study” (p. 1305)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	No	Attrition >15% (23,8%). Only reports data on completers
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Mealiea 1971

Study characteristics

Methods

Parallel randomised trial with five arms

1. Psychological placebo: pseudotherapy
2. No-treatment
3. Active treatment 1: systematic desensitisation
4. Active treatment 2: implosive therapy
5. Active treatment 3: implosive-desensitisation (ID)

Sample calculation: not stated

Cluster randomised: no

Duration of trial (baseline to post): 5 weeks

Duration of participation (trial + follow-up): 5 weeks + 4 weeks follow-up

Setting: outpatient (University setting)

Mealiea 1971 *(Continued)*

Purpose of trial: "Systematic desensitization (SD) and implosive therapy (IT) were compared for their effectiveness in modifying snake phobic behavior." (p. 85)

Open/closed placebo: closed placebo

Data

Number of participants screened: 1200

Number of participants included: 50

Number of participants followed-up at post treatment: 48

Number of participants randomly assigned to:

- Psychological placebo: n = 10
- No-treatment: n = 10
- Active treatment 1: n = 10
- Active treatment 2: n = 10
- Active treatment 3: n = 10

Number of withdrawals: n =1

- Psychological placebo: n = 0
- No-treatment: n = 0
- Active treatment 1: n = 0
- Active treatment 2: n = 1
- Active treatment 3: n = 0

Diagnosis: specific anxiety (snake phobia)

Diagnostic manual: not stated

Means of assessment: not stated

Comorbidity: not stated

Ag:: not stated

IQ: not stated - but university students

Sex: 100% females

Ethnicity: not stated

Country: USA

Inclusion criteria

1. That the individual not be undergoing or have undergone any form of psychiatric treatment
2. That she be able to experience vivid imagery
3. That she have an intense fear of snakes as measured by the Behavioral Avoidance Test.

Exclusion criteria

1. Not stated

Comparisons

Psychological placebo

Treatment name: Pseudotherapy

Description of intervention: "The PT group received treatment similar to that of the SD group except that relaxation was paired with 20 snake irrelevant scenes; i.e. relaxation was paired with descriptions of landscapes, beaches, clouds, etc. The PT group was included to control for possible placebo or expectancy effects." (p. 88)

Mealiea 1971 *(Continued)*

Individual or group treatment: group

Exposure/intensity to treatment: 5 sessions, 30 minutes each

Duration of treatment:: 5 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

No-treatment

Comparison name: No-treatment

Description of intervention: “The Ss in the NT group participated only in the pre-, post-, and follow-up treatment assessment sessions. The same programmed instructions regarding assessment procedure given the other groups were given the NT group." (p. 88)

Exposure/intensity to treatment: no treatment

Duration treatment: 5 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect

- **Hierarchy:** usable data, observer-reported
- **Outcome chosen: n**umber of people who picked up snake

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. The results support a counterconditioning view of SD and indicate that further research must be conducted before any conclusions can be made concerning it’s efficacy

Key limitations from study authors

1. The programmed IT treatment could be criticized on the basis that without a live therapist interacting with the Ss, it would be easy for the Ss to engage in avoidance behavior, thereby mitigating the implosive effect

Other notes

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: ”The Ss were matched on the basis of both their BAT and FT scores and randomly assigned to one of five equal groups.” (p. 87)
Allocation concealment	Unclear	No information

Mealiea 1971 *(Continued)*

Blinding of outcome assessors	Yes	Outcomes are objective
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Attrition < 15% (4%)
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources of bias found

Milby 1980

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Usual care: Group II – therapy 2. Wait-list: Group III - no treatment 3. Active treatment: Group I – therapy + urine surveillance **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 3 months of treatment **Duration of participation (trial + follow-up):** no follow-up **Setting:** outpatient **Purpose of trial:** "This study assessed the efficacy of urine surveillance as an adjunct to outpatient psychotherapy." (p. 994)
Data	**Number of participants screened:** not stated **Number of participants included:** 29 **Number of participants followed-up at post treatment:** 29 **Number of participants randomly assigned to** : • Usual care: n = 13 • Wait-list: n = 13. • Active treatment: n = 13 **Number of withdrawals** : n = 0 **Diagnosis:s**ubstance use disorder **Diagnostic manual:** not stated **Means of assessment:** Menlo Park VA Drug Abuse Program Questionnaire, MMPI. **Comorbidit**: not stated - previously addicted to narcotics or barbiturate-like drugs **Age:** 25.8 mean years (range =16 to 54)

Milby 1980 *(Continued)*

IQ: not stated

Sex: 34.5% females

Ethnicity: 22 Whites, 7 Blacks

Country: USA

Inclusion criteria

1. Did not meet FDA criteria for methadone maintenance
2. Was addicted to narcotics or barbiturate-like drugs
3. Was able and willing to complete the test battery
4. Was a first admission
5. Agreed to participate in the study and signed appropriate consent forms

Exclusion criteria

1. eliminated during study if Attended less than 50 % of treatment
2. Got more than three hours of psychotherapy elsewhere
3. Failed to return on follow-up testing

Comparisons	**Usual care** **Treatment name:** group II - therapy **Description of intervention:** "Group I received 3 months of outpatient individual or group therapy (…) Group II received the same treatment (…)" (p. 995) **Individual or group treatment:** individual and group therapy **Exposure/intensity to treatment:** not stated **Duration of treatment:** 3 months **Concomitant psychotherapy:** not stated. See exclusion criteria **Concomitant pharmacotherapy:** not stated **Wait-list** **Comparison name:** No treatment (in reality a wait-list) **Description of intervention:** "Group III was not seen for outpatient therapy or urine surveillance until after 3 months, thus serving as a control group." (p. 996) **Exposure/intensity to treatment** : no treatment during waiting **Duration treatment:** 3 months **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy: u**sable data, clinical relevance • **Outcome chosen: n**umber of participants who spent more, less or zero hours per day in illegal activities before versus after treatment - subscale spent zero **Adverse event** • No data reported on adverse events

Milby 1980 *(Continued)*

Notes	**Key conclusion from study authors** 1. All results considered, one must conclude that urine surveillance was somewhat helpful. Main results supporting this were significant differences between the surveillance group and the other groups in decreased drug-related friendships and barbiturate use. However, these results were somewhat effaced by the rather broad changes that occurred in all groups from pre- to posttreatment. These broad changes can at least partially be accounted for by the intense intervention that all subjects received during the inpatient detoxification and evaluation phase **Key limitations from study authors** 1. Not stated **Other notes from review authors** 1. None **Conflicts of interest:** none found **Judgement:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: ”After the battery, patients were randomly assigned to one of three treatment groups.” (p. 995)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Unclear	No information
Blinding of participants and personnel	No	Not possible to blind usual care and wait-list
Incomplete outcome data	Unclear	Quote: “Subjects were eliminated from the study if they: (1) attended less than 50% of the outpatient therapy sessions scheduled, (2) got more than 3 hours psychotherapy or other treatment elsewhere when they had been assigned to Group 3, and (3) failed to return for follow-up testing (…) No subjects needed to be eliminated from the study.” (p. 996) Attrition unclear
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Miranda 1997

Study characteristics

Methods	**Parallel randomised trial with four arms** 1. Psychological placebo: placebo control 2. No-treatment 3. Active treatment 1: attributional training 1

Miranda 1997 *(Continued)*

4. Active treatment 2: attributional training 2

Sample calculation: not stated

Cluster randomised: no

Duration of trial (baseline to post): 2.5 months

Duration of participation (trial + follow-up): 2.5 months - unclear if followed up

Setting: community outpatient

Purpose of trial: This study investigates the efficiency of attributional retraining associated to a programme aimed at teaching problem-solving strategies to students with learning disabilities (LD)

Open/closed placebo: closed placebo

Data

Number of participants screened: 173

Number of participants included: 41

Number of participants followed-up at post treatment: 41

Number of participants randomly assigned to

- Psychological placebo: n = 11
- No-treatment: n = 10
- Active treatment 1: n = 10
- Active treatment 2: n = 10

Number of withdrawals: 0

Diagnosis: learning disabilities

Diagnostic manual: not stated

Means of assessment: arithmetic verbal task

Comorbidity: not stated

Age: mean range 10.1 to 11.9

IQ: within normality

Sex: 47.6% female

Ethnicity: not stated

Country: Spain

Inclusion criteria

1. Normal intelligence
2. Absence of socio-cultural deprivation
3. Regular school attendance
4. Poor performance on arithmetic verbal problems compared to norm group

Exclusion criteria

1. Not stated

Comparisons

Psychological placebo

Treatment name: Placebo

Miranda 1997 *(Continued)*

Description of intervention: "Carried out problem-solving activities but without training in strategies." Used the same materials as work in experimental group, but the instruction procedure consistent of independent practice of problems.

Individual or group treatment: group

Exposure/intensity to treatment: 3 weekly sessions lasting 50 minutes

Duration of treatment: 2.5 months

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

No-treatment

Comparison name: No-treatment

Description of intervention: received no special training

Exposure/intensity to treatment: no treatment

Duration treatment: 2.5 months

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported, clinical relevance
- **Outcomes chosen:** teacher-rated problems list by Achenbach

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. Results indicated that children in the two experimental groups following the strategy instruction programme obtained higher test scores on all measures, but improvements, specially in the follow-up phase, were greater for the group that also received attributional retraining

Key limitations from study authors

1. Not stated

Other notes from review authors

1. Study is in Spanish. Translation was required

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "The selected subjects were randomly assigned to each of the four groups: (p. 37)
Allocation concealment	Unclear	No information

Miranda 1997 *(Continued)*

Blinding of outcome assessors	Unclear	No information
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Attrition <15% (0%)
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	None other sources found

Mitchell 2008

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Physical placebo 2. Wait-list: control 3. Active treatment: experimental group **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 1 month **Duration of participation (trial + follow-up):** 1 month (no follow-up) **Setting:** outpatient **Purpose of trial:** "The current study investigated the effects of lens tinting by using a design which provided controls for various effects (e.g., Hawthorne, maturational, practice, and experimenter effects." (p. 517) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** 66 **Number of participants included:** 49 **Number of participants followed-up at post treatment:** 49 N **umber of participants randomly assigned to:** 1. Physical placebo: n = 15 2. Wait-list: n = 17 3. Active treatment: n = 17 **Number of withdrawals** : n = 0 **Diagnosis:** visuoperceptual reading disabilities **Diagnostic manual:** not stated **Means of assessment:** Neale Analysis of Reading Ability **Comorbidity:** not stated **Age:** range 7 to 11 years **IQ:** not stated **Sex:** 28.6% female **Ethnicity:** not stated **Country:** South Africa **Inclusion criteria** 1. Visuoperceptual reading disabilities

Mitchell 2008 *(Continued)*

2. No ophthalmic problems which could lead to the reading disability
3. Informed consent received from the parent of the child and the child

Exclusion criteria

1. Undiagnosed ophthalmological problems

Comparisons	**Physical placebo** **Treatment name:** Placebo **Description of intervention:** "(…) the placebo group underwent the colour testing and received coloured filters in a colour complementary to that which was specified as the optimal colour for that child." (p. 523) "All children who received no or placebo filters were supplied with correctly matched coloured filters." (p. 524) **Individual or group treatment:** individual **Exposure/intensity to treatment:** "During this time the placebo and experimental groups wore their filters for all reading activities including homework (…)" (p. 523) **Duration of treatment:** one month **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated **Wait-list** **Comparison name:** Control group (wait-list). **Description of intervention:** "(…) the control group received no filters (…)." (p. 523) "(…) the control group received no special treatment." (p. 523) "All children who received no or placebo filters were supplied with correctly matched coloured filters." (p. 523) **Exposure/intensity to treatment:** no treatment **Duration treatment:** month **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy:** observer-reported, clinical relevance, psychometric properties • **Outcomes chosen:** Neale Analysis of Reading Ability: age **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. "The results on the cognitive efficiency measure (Symbol Digit Modalities Test) are of interest. The coloured lenses had no significant effect on scores, indicating that the lenses did not necessarily assist the participants to decode the visual stimuli more efficiently or that the lenses were implicated in the visuoperceptual task of the written response." (p. 529) 2. "The current study yielded no significant improvements in reading accuracy, comprehension, or rate". (p. 529) **Key limitations from study authors** 1. "That group differences were seen on three of the six dependent measures may suggest problems with randomization. No specific problems with the randomization protocol could be identified; however, this possibility cannot be dismissed." (p. 529) 2. "A lack of a clear definition of visuoperceptual reading disabilities makes operationalization of variables and selection of subjects difficult." (p. 529) **Other notes from review authors**

Mitchell 2008 *(Continued)*

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "The researcher randomly placed the participants into three groups, as matching can create uncontrollable discrepancies." (p. 524)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Unclear	No information
Blinding of participants and personnel	No	Not possible to placebo and no-treatment
Incomplete outcome data	Yes	No mention of attrition rate or imputation methods, but it seems from table 2 that all participants is included in the analyses
Selective outcome reported	Unclear	No protocol found
Other sources of bias	No	Quote: "That group differences were seen on three of the six dependent measures may suggest problems with randomization." (p. 529)

Nandi 1976

Study characteristics

Methods

Parallel randomised trial with three arms

1. Pharmacological placebo
2. No-treatment: natural process
3. Active treatment: antidepressive drug treatment

Sample calculation: not stated
Cluster randomised: no
Duration of trial (baseline to post): 28 days.
Duration of participation (trial + follow-up): not stated
Setting: outpatient
Purpose of trial: "To find an answer to the question, whether all of those who were labelled as depressives in this door-to-door study were treatable depressives as compared to those who sought treatment in clinics, a clinical trial with the following design was undertaken." (p. 524)
Open/closed placebo: closed placebo

Data

Number of participants screened: 1078
Number of participants included: 41
Number of participants followed-up at post treatment: 35
Number of participants randomly assigned to:

- Pharmacological placebo: n = 10
- No-treatment: n = 11

Nandi 1976 *(Continued)*

- Active treatment: n = 20

Number of withdrawals : n = 6

- Pharmacological placebo: n = 0
- No-treatment: n = 3
- Active treatment: n = 3

Diagnosis: depression

Diagnostic manual: not stated (see other notes)

Means of assessment: clinical interview. “The labelling of a case with an appropriate diagnosis was done on the basis of diagnostic criterion accepted for each diagnostic entity.” (p. 524)

Comorbidity: “The patients were thoroughly examined and all were found to be free from any physical illness.” (p. 524)

Age: not stated

IQ: not stated

Sex: not stated

Ethnicity: not stated

Country: India

Inclusion criteria

1. Depression

Exclusion criteria

1. None mentioned

Comparisons	**Pharmacological placebo** **Treatment name:** Placebo **Description of intervention:** “The second group, consisting of 10 patients, was given placebo and will be called the Placebo group. The placebo used was lactose in the form of tablets, which were administered in exactly the same manner as the drugs prescribed in the Medicine group” (p. 524) “One 25 mg tablet twice daily for 2 days, then two tablets twice daily to continue for the remaining days; time of administration being morning and afternoon after some meal.” (p. 524) **Individual or group treatment:** individual **Exposure/intensity to treatment:** one 25 mg tablet twice daily for 2 days, then two tablets daily the remaining 26 days **Duration of treatment:** 28 days **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated **No-treatment** **Comparison name** : Natural Process (no treatment) **Description of intervention:** “The third group consisting of 11 patients was given no treatment (and was left to nature for any change that might be observed in them after a given interval of time.” (p. 524) **Exposure/intensity to treatment:** no treatment **Duration treatment:** 28 days. **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy:** primary outcome, observer-reported • **Outcomes chosen:** Hamilton’s Depressive Rating Scale

Nandi 1976 *(Continued)*

Adverse events

- Spontaneous reporting

Notes

Key conclusion from study authors

1. "The comparison between Placebo and Natural Process groups did not yield a significant result on the 28th day."(p. 526)
2. "Thus it may be stated that the change seen in the Medicine group was definitely different from those observed in the Placebo and Natural Process groups". (p. 526)
3. "Thus we may say that the rural population who were labelled as depressives were not different from those who attended clinics and hospitals and sought treatment, so far as their response to treatment is concerned." (p. 526)

Key limitations from study authors

1. None mentioned

Other notes from review authors

1. "An affective disorder characterized essentially by morbid changes of mood in the form of depression which is unprovoked by any physical or environmental cause, and expressed the feeling of misery, gloom and wretchedness often tinged with anxiety. Self-reproach, moral worthlessness (guilt feeling), and suicidal tendency, are quite common. When occurring for the first time in late forties, strong paranoid component may be present. Hypochondriacal ideas which in extreme cases may be nihilistic and bizarre, are frequent.The mood tends to be worse in the morning. Biological symptoms like disturbances of sleep pattern,early morning waking being the rule, loss of weight, appetite and libido are almost invariably present. There is retardation of thinking and Endogenous Depression, depressed type of MDP, action which may proceed to the level of stupor. This Reactive Depression and Involutional Depression have all psychotic state has a tendency to recur and is often been included. But Neurotic Depression must be classified self-limiting, separately." (p. 526-7)

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "They were randomly divided into four groups. Three of these groups had 10 patients each and the fourth group had 11 patients. Two of them were randomly chosen and merged together which consisted of 20 patients" (p. 524)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote:"The placebo tablets and the drugs used in the Medicine group did not look alike, but this dissimilarity was not a source of bias, as neither the patients nor the raters knew who were given what." (p. 524) Quote:"All the assessments were made by two raters and inter-rater agreement was found to be high. Both the raters worked independently, and the mean of their ratings was taken in each case." (p. 524)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	No	Attrition rate >15% (15% in medicine. 0% in Placebo, 27% in Natural process. Only reports data on completers.

Nandi 1976 *(Continued)*

		No mention of ITT
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Nicassio 1974

Study characteristics	
Methods	**Parallel randomised trial with four arms** 1. Physical placebo: self-relaxation 2. No-treatment 3. Active treatment 1: progressive relaxation 4. Active treatment 2: autogenic training **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 4 weeks **Duration of participation (trial + follow-up):** 4 + 6 months follow-up **Setting** : Outpatient **Purpose of trial** : "the objective of the present study is to compare and evaluate two direct, short-term treatments for insomnia." (p. 253) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** 60 **Number of participants included:** 32 **Number of participants followed-up at post treatment:** 30 **Number of participants randomly assigned to** : • Physical placebo: n = 8 • No-treatment: n = 8 • Active treatment 1: n = 8 • Active treatment 2: n = 8 **Number of withdrawals** : n = 2 • Physical placebo: not stated • No-treatment: not stated • Active treatment 1: not stated • Active treatment 2: not stated **Diagnosis:** sleep-wake disorder (insomnia) **Diagnostic manual:** not stated **Means of assessment:** daily sleep records (not otherwise specified) **Comorbidity:** not stated **Age:** 45.1 mean years (SD = 14.57 (range: 22 to 71 years) **IQ:** not stated **Sex:** 70% female **Ethnicity:** not stated **Country:** USA **Inclusion criteria**

Nicassio 1974 *(Continued)*

1. Average daily time to fall asleep exceeded 30 minutes.

Exclusion criteria

1. None mentioned

Comparisons	**Physical placebo** **Treatment name** : Self-relaxation control **Description of intervention:** "(…) subjects received four one-hour individual treatment sessions and a posttreatment session at which final assessments were made." (p. 255) "Although subjects were seen for the same number of treatment sessions as those receiving bona fide relaxation instruction, no technique of relaxation was taught. Subjects were told that everyone knows how to relax; it is just a matter of scheduling time to do so." (p. 255) **Individual or group treatment:** individual treatment. **Exposure/intensity to treatment:** "As in the other treatment groups, subjects were instructed to relax for one 20- to 30-minute session at home during the day and again in bed while preparing for sleep." (p. 255) **Duration of treatment:** 4 weeks. **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** "An additional requirement was that all subjects stop taking sleeping pills or other forms of medication with soporific effects after obtaining the permission of their physicians." (p. 254) **Wait-list** **Comparison name:** No treatment control (wait-list) **Description of intervention:** "Subjects in this group were seen at the beginning of the treatment period and were told that an extended baseline measurement period would be needed to accurately assess the nature and stability of their sleeping difficulties. For the next four weeks, subjects mailed in data and were not seen again until the posttreatment session at which several assessments were mad. Subjects then received either progressive relaxation or autogenic training for their insomnia." (p. 255) **Exposure/intensity to treatment:** no treatment **Duration treatment:** 4 weeks **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** "An additional requirement was that all subjects stop taking sleeping pills or other forms of medication with soporific effects after obtaining the permission of their physicians." (p. 254)
Outcomes	**Beneficial effect** • **Hierarchy:** usable data • **Outcomes chosen:** time to fall asleep **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. "At posttest, treated subjects reported less time to fall asleep than subjects in the two control groups and they reported more global improvement." (p. 258) 2. "The two treatments were equally effective with the one exception that subjects in progressive relaxation showed significant improvement by the third week, whereas subjects in autogenic training did not show significant improvement until the fourth week." (p. 25) 3. "As measured by daily reports from the subjects, a reduction in time to fall asleep was not accompanied by improvement on other sleep variables such as hours slept, feeling on awakening, quality of sleep, and number of awakenings per night." (p. 259) 4. "As measured by daily reports from the subjects, a reduction in time to fall asleep was not accompanied by improvement on other sleep variables such as hours slept, feeling on awakening, quality of sleep, and number of awakenings per night." (p. 259)

Nicassio 1974 *(Continued)*

Key limitations from study authors

1. "Interestingly, at follow-up, the daily log of time to fall asleep indicated that subjects had maintained their improvement, but the global reports were no longer as enthusiastic. This may be due to either the possibility that posttest global reports were spuriously high (due, for example, to a "hello goodbye" effect) or to the possibility that subjects no longer remembered how severe their problem was before treatment began." (p. 259.)

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote:"Within each wave, subjects were blocked according to time to fall asleep and then randomly assigned to the following conditions: (a) autogenic training, (b) progressive relaxation, (c) self-relaxation control, and (d) no-treatment control." (p. 254)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Quote:"2 participants dropped out of the study for reasons unrelated to treatment." (p. 254) Attrition <15% (6.25%)
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Pearl 1956

Study characteristics

Methods

Parallel randomised trial with three arms

1. Pharmacological placebo: placebo pill
2. No-treatment
3. Active treatment: reserpine

Sample calculation: not stated
Cluster randomised: no
Duration of trial (baseline to post): 2 months
Duration of participation (trial + follow-up): 2 months. No follow-up
Setting: Inpatient (closed hospital ward)

Pearl 1956 *(Continued)*

Purpose of trial: "The present investigation was undertaken to study the effectiveness of reserpine in the treatment of schizophrenic male patients within a controlled setting" (p. 198)

Open/closed placebo: closed placebo

Data

Number of participants screened: not stated
Number of participants included: 133
Number of participants followed-up at post treatment: not stated

Number of participants randomly assigned to:

- Pharmacological placebo: n = 52
- No-treatment: n = 22
- Active treatment: n = 59

Number of withdrawals: not stated

Diagnosis: schizophrenia
Diagnostic manual: not stated
Means of assessment: Diagnostic interview: Lorr Multidimensional Scale for Rating Psychiatric Patients
Comorbidity: different symptomatology of schizophrenia

Age: not stated
IQ: not stated
Sex: 100% male
Ethnicity: not stated
Country: USA

Inclusion criteria

1. Only patients were selected whose diagnoses were uncomplicated by findings of organici nvolvement
2. who within the previous year had not received insulin coma therapy, electroconvulsive treatment, or individual or group psychotherapy
3. and who were not consistently receiving adjunctive medication, such as barbiturates

Exclusion criteria

1. None mentioned

Comparisons

Pharmacological placebo
Treatment name: Placebo tablets
Description of intervention: "As a control factor, placebo tablets were obtained which were identical in physical characteristics with the drug tablets, e.g., shape, size, color, and taste. All personnel dispensing the placebo were told that it was a variant of reserpine named Plasepine. An equivalent number of placebo tablets was given to control patients, as were reserpine tablets to experimental subjects." (p. 199)
Individual or group treatment: individual treatment.
Exposure/intensity to treatment: "The actual reserpine dose varied from 2 to 10 mg. daily, the majority of patients receiving 4-5 mg. a day." (p. 199)
Duration of treatment: 2 months
Concomitant psychotherapy: not stated, but inpatients
Concomitant pharmacotherapy: no adjunctive medication

No-treatment
Comparison name: No-treatment
Description of intervention: not stated
Exposure/intensity to treatment: no treatment
Duration treatment: 2 months
Concomitant psychotherapy: not stated, but inpatients
Concomitant pharmacotherapy: no adjunctive medication

Pearl 1956 *(Continued)*

Outcomes	**Beneficial effect** • **Hierarchy:** primary outcome • **Outcomes chosen:** Multidimensional Scale for Rating Psychiatric Patient **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. "Comparisons between reserpine and control patients disclosed that the most chronically regressed patients treated by reserpine showed no significant improvement, whereas less chronic and more disturbed patients improved in various areas of psychopathology." (p. 204) 2. "The extent of improvement, however, was not sufficiently great, except in individual patients, that the bulk of reserpine-treated schizophrenic patients could be eligible for early hospital discharge." (p. 204) **Key limitations from study authors** 1. Not stated **Other notes from review authors** 1. No usable data **Conflicts of interest:** none found **Judgement:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote:"Within each ward, patients were randomly divided into three treatment groups." (p. 199)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Unclear	Quote:"In order to obtain as reliable judgments of condition as possible, in this study, the diagnostic interviews were jointly conducted by a team composed of psychiatrist, clinical psychologist, and psychological trainee. Each then rendered independent ratings based on his interview observations and inferences. Independent ratings of ward behavior were also secured from the nurse and charge aide on each of the experimental wards" (p. 199) No mention of blinding
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Unclear	Attrition unclear
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Peck 1976

Study characteristics

Methods	**Parallel randomised trial with five arms** 1. Psychological placebo + pharmacological placebo: Attention-placebo + pill 2. No-treatment 3. Active treatment 1: contact desensitisation 4. Active treatment 2: vicarious symbolic desensitisation 5. Active treatment 3: systematic desensitisation **Sample calculation:** yes **Cluster randomised:** no **Duration of trial (baseline to post):** 5 weeks **Duration of participation (trial + follow-up):** not stated **Setting:** inpatient **Purpose of trial:** "This study is an attempt to apply three types of desensitization procedures (Systematic Densitization, Vicarious Symbolic Desensitization, and Contact Densitization) to the mentally retarded." (p. 138) **Open/Closed placebo:** closed placebo
Data	**Number of participants screened:** 67 **Number of participants included:** 20 **Number of participants followed-up at post treatment:** not stated **Number of participants randomly assigned to:** • Psychological placebo + pharmacological placebo: n = 4 • No treatment: n = 4 • Active treatment 1: n = 4 • Active treatment 2: n = 4 • Active treatment 3: n = 4 **Number of withdrawals:** not stated **Diagnosis:** mild retardation + specific anxiety (heights/rats) **Diagnostic manual:** American Association of Mental Deficiency 1973 **Means of assessment:** Modified Fear Survey Schedule, Behavior Avoidance Test (BAT) **Comorbidity:** not stated **Age:** range = 19 to 61 years **IQ:**: 52 to 74 **Sex:** not stated **Ethnicity:** not stated **Country:** USA I**nclusion criteria** 1. Mild retardation

Peck 1976 *(Continued)*

2. Would not touch the rat
3. or would not climb higher than three flights on a fire escape

Exclusion criteria

1. Not stated

Comparisons	**Psychological and pharmacological placebo** **Treatment name:** Attention placebo + pill **Description of intervention:** "modified version of the "stress training" placebo rationale (...) was used. Ss were told that by learning to handle stress in one situation, they would learn not to be fearful in the presence of the phobic stimulus, Ss were instructed to take a "'pill" (a passive placebo was used) which would relax them and left alone in the room for 5 minutes. The therapist subsequently reentered the room, checked the S's pulse and eyelids and announced that he/she was relaxed. S was instructed to lie back, relax, and watch a cartoon on the videotape. The cartoons shown were interrupted visually and auditorily at the most exciting parts and a number was inserted as the presumed "stress training". Ss were asked to read off these numbers as they appeared, and the therapist appeared to collect data on the number and latency of response. At the end of each session, S was told that he/she was progressing well and would soon be unafraid of the phobic stimulus (height or rat)." (p.139-140) **Individual or group treatment:** individual **Exposure/intensity to treatment:** 15 sessions of 1,5 hours, three times a week **Duration of treatment:** 5 weeks **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** "Fifteen of the Ss received tranquilizing medication of various types before and during the experiment." (p. 139) **No-treatment** **Comparison name:** No treatment **Description of intervention:** "No-Treatment Control Ss were not contacted during the treatment phase." (p. 140) **Exposure/intensity to treatment:** no contact **Duration treatment:** 5 weeks **Concomitant psychotherapy:** not stated Concomitant pharmacotherapy: "Fifteen of the Ss received tranquilizing medication of various types before and during the experiment." (p. 139)
Outcomes	**Beneficial effect** • **Hierarchy:** usable data, observer-reported • **Outcome chosen:** Behavior Avoidance Test (BAT) **Adverse events** • No data on adverse events reported
Notes	**Key conclusion from study authors** 1. Post-treatment data on the Behavior Avoidance Test, Fear Thermometer and Behavior Checklist showed that generally contact desensitisation was most effective and most efficient with this population. 2. Results show that the mildly retarded are able to follow slightly simplified desensitisation procedures.

Peck 1976 *(Continued)*

Key limitations from study authors

1. Small sample

Other notes from review authors

1. SDs were generated from Etringer 1982 and Rosen 1976 due to that the same population and outcome was used in all trials

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote:"Ss were then randomly assigned within blocking variables (high and low anxiety) and height or rat stimuli to one of five conditions. … " (p. 139)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Unclear	No information
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Unclear	Attrition unclear
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Pelham 1992

Study characteristics

Methods

Parallel randomised trial with three arms

1. Pharmacological placebo
2. No-treatment: no pill
3. Active treatment: methylphenidate

Sample calculation: not stated
Cluster randomised: no
Duration of trial (baseline to post): 6 weeks
Duration of participation (trial + follow-up): not stated
Setting: outpatient
Purpose of trial: "The present studies were designed to examine ADHD boys' causal attributions in a double-blind, within-subject, placebo-controlled medication trial." (p. 283)
Open/closed placebo: closed placebo

Data

Number of participants screened: not stated
Number of participants included: 38
Number of participants followed-up at post treatment: 38

Pelham 1992 *(Continued)*

Number of participants randomly assigned to:

- Pharmacological placebo: n = 13
- No-treatment: n = 12
- Active treatment: n = 13

Number of withdrawals : n= 0
Diagnosis: Attention Deficit/hyperactivity Disorder (ADHD)
Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition, Revised (DSM-III-R)
Means of assessment: structured parent interview, standardised parent and teacher rating scales
Comorbidity: "Eighteen also met criteria for conduct oppositional/defiant disorder,and another 8 met criteria for conduct disorder (…)" (p. 286)

Age: 9 mean years and 11 months, (range =7 years and 3 months to 13 years and 9 months)
IQ: Wechsler Intelligence Scale for children – mean: 106.2 (SD =13.1)
Sex: 100% male
Ethnicity: not stated
Country: USA

Inclusion criteria

1. Meet the criteria for a DSM-III-R diagnosis of ADHD

Exclusion criteria

1. None mentioned

Comparisons

Pharmacological placebo
Treatment name : Placebo
Description of intervention:
"Each subject received an identical capsule of either a low dose, a high dose, or placebo before 8.00 a.m. and at midday, with the medication condition randomized daily." (p. 283)
"A day in the STP lasted from 8.00 a.m. until 5.00 p.m. on weekdays and was divided into the following activities: two academic classroom periods, each staffed by a special education teacher and an aide; an art class; swimming; three supervised, group, outdoor recreational activities (e.g., soccer); and lunch." (p. 283)
Individual or group treatment: individual
Exposure/intensity to treatment: before 8.00 a.m. and at midday (dose not mentioned)
Duration of treatment: 6 weeks
Concomitant psychotherapy: not stated - but was a part of a Summer Treatment Program
Concomitant pharmacotherapy: 22 of the 38 participants received pemoline, dexedrine spansule, slow-release ritalin on days not included in the experiment. No carry-over effect was present as concluded by the authors.

No-treatment
Comparison name: No pill
Description of intervention: "No treatment but attendance in the Summer Treatment Program.
"A day in the STP lasted from 8.00 a.m. until 5.00 p.m. on weekdays and was divided into the following activities: two academic classroom periods, each staffed by a special education teacher and an aide; an art class; swimming; three supervised, group, outdoor recreational activities (e.g., soccer); and lunch." (p. 283)
Exposure/intensity to treatment: no pill
Duration treatment: 6 weeks
Concomitant psychotherapy: not stated - but was a part of a Summer Treatment Program
Concomitant pharmacotherapy: 22 of the 38 participants received pemoline, dexedrine spansule, slow-release ritalin on days not included in the experiment. No carry-over effect was present as concluded by the authors.

Outcomes

Beneficial effect

Pelham 1992 *(Continued)*

- **Hierarchy:** patient-reported, clinically relevant
- **Outcome chosen:** rating Scale Items (10-point scale): subscale behaviour (question: Did you have a good or bad day?)

Adverse events

- No data on adverse events reported

Notes

Key conclusion from study authors

1. "Simply taking a pill (no-pill vs. placebo comparison) did not show significant effects (…)." (p. 282)
2. 2. "Across drug conditions a self-enhancing attributional pattern was obtained; the majority of attributions for success were to ability or effort, whereas attributions for failure were to the pill or to counselors." (p. 282)

Key limitations from study authors

1. (…) results were obtained in a day-treatment setting with a highly structured behavioral point system, and the effects of medication on attributions and perceptions must be interpreted in that light." (p. 291)
2. "A second limitation concerns the fact that we assessed only children of elementary-school age. Because response to MPH is inversely correlated with age, alder ADHD children may have fewer opportunities to make effort attributions for medication-induced success."(p. 291)
3. "(…) although larger and more comprehensive than previous investigations, out studies assessed attributions in an acute medication trial. Because ADHD children typically receive stimulant for years rather than days, it would be critical to study attributions after extended pharmacotherapy. Perhaps consistent and long-term association between medication and success/failure outcomes may induce or magnify the dysfunctional style shown by some of out HPA/LPA subgroups" (p. 291)

Other notes from review authors

1. Since 38 patients was included in total, we assumed due to randomisation and ethical principles that the active arm (methylphenidate) and placebo included an additional patient each

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	No information
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Outcome was patient-reported
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Attrition < 15% (0%)
Selective outcome reported	Unclear	No protocol found
Other sources of bias	No	Quote: "Each subject received an identical capsule of either a low dose, a high dose, or placebo before 8.00 a.m. and at midday, with the medication condition randomized daily." (p. 283)

Pelham 1992 *(Continued)*

	Potential carry-over effect

Pendleton 1983

Study characteristics	
Methods	**Parallel randomised trial with five arms** 1. Psychological placebo: relaxation training expectancy control 2. Waiting list 3. Active treatment 1: negative practice 4. Active treatment 2: desensitisation **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 6 weeks **Duration of participation (trial + follow-up):** 6 weeks + 2 weeks post-assessment **Setting:**outpatient (University setting) **Purpose of trial** : "The purpose of this study was to evaluate the effectiveness of a symptom scheduling technique in reducing excessive fear." (p. 317) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** not stated **Number of participants included:** 62 **Number of participants followed-up at post treatment:** 58 **Number of participants randomly assigned to** : • Psychological placebo: n =15 • Waiting list: n =15 • Active treatment 1: n =16 • Active treatment 2: n =16 **Number of withdrawals**: n= 4 • Psychological placebo: n =2 • Waiting list: n =0 • Active treatment 1: n =2 • Active treatment 2: n = 0 **Diagnosis** : Specific anxiety (Acrophobia) **Diagnostic manual:** not stated **Means of assessment:** Acrophobia Behavioral Test **Comorbidity:** not stated **Age:** 25.1 mean years (SD = 6.75) (R =17 to 54) **IQ:** not stated **Sex:** 74.2% females

Pendleton 1983 *(Continued)*

Ethnicity: not stated

Country : USA

Inclusion criteria

1. Those who reported both strong fear of heights and strong interest in overcoming their fear were scheduled for pre-testing
2. Only participants who additionally failed to complete the Acrophobia Behavioral Test (described below) at pre-testing were selected for the study

Exclusion criteria

1. Not stated

Comparisons	**Psychological placebo** **Treatment name:** Relaxation training expectancy control **Description of intervention:** "Subjects in the relaxation only treatment condition received a treatment rationale in which relaxation training was described as a passive, automatic process that acts in a cumulative manner." (p. 319) "Although the relaxation only condition was included as an expectancy control, it appears in retrospect to have inadvertently contained specific treatment components." (p. 321) **Individual or group treatment:** group **Exposure/intensity to treatment:** 45 minutes x 6 sessions **Duration of treatment:** 6 weeks **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated **Wait-list** **Comparison name:** Wait-list **Description of intervention:** "Subjects in the waiting list control condition were contacted during the week following the pre-testing session and informed that their treatment would be delayed. They were then retested during the regular posttesting period. After posttesting, these subjects were debriefed and then treated with the desensitization procedure." (p. 319) **Exposure/intensity to treatment:** no treatment during waiting **Duration treatment:** 6 weeks **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy:** observer-reported • **Outcome chosen:** Acrophobia behavioral test **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors**

Pendleton 1983 *(Continued)*

1. It was concluded that symptom scheduling can be an effective treatment for acrophobia, and that these results are likely to generalise to clinical populations

Key limitations from study authors

1. Not stated

Other notes from review authors

1. Information on how many patients was randomised to each group was lacking. Since 62 patients was included in total, we assumed (due to randomisation and ethical principles) that the two active arms included an additional patient each

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Acrophobic subjects were randomly assigned to one of three treatment conditions." (p. 318)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote:" Posttesting session. Posttesting appointments were scheduled one to two weeks after completion of all treatment sessions. All pretreatment measures were re- administered in the same manner as in pre-testing by an experimenter who was blind to the subjects' pretreatment performances and condition assignments." (p. 319)
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	Yes	Quote:"Two community and two student subjects (one each from both the negative practice and relaxation only conditions) dropped out of the study due to scheduling conflicts." (p. 318) Attrition <15% (6.5%). No ITT. Reports data on completers only
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	None other sources found

Pillman 2001

Study characteristics

Methods

Parallel randomised trial with three arms

1. Physical placebo
2. Wait-list: no-treatment
3. Active treatment: experimental group

Sample calculation : not stated
Cluster randomised: no

Pillman 2001 *(Continued)*

Duration of trial (baseline to post): participants averaged 22 days to complete the study (range between 18 to 35 days from pre- to post-test)
Duration of participation (trial + follow-up): "All tests given on the 1 st day of testing (…) The final day of testing was administered 28 days after the initial testing." (p. 56)
Setting: inpatient substance abuse rehabilitation
Purpose of trial: "Twenty-one subjects with a primary diagnosis of alcohol abuse in a 30-day inpatient treatment program were placed into three groups to determine the effectiveness of the computerized cognitive remediation treatment, NeurXercise." (p. v)
Closed/open placebo: closed placebo

Data

Number of participants screened: 45
Number of participants included: 38
Number of participants followed-up at post treatment: 22
Number of participants randomly assigned to:

- Physical placebo: n = 13
- Wait-list: n = 12
- Active treatment: n = 13

Number of withdrawals: n = 20

- Physical placebo: n = 7
- Wait-list: n = 7
- Active treatment: n = 6

Diagnosis: alcohol dependency
Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV)
Means of assessment: not stated
Comorbidity: not stated (but excluded if they were polysubstance abusers)
Age: Pharmacological placebo: 48.43 mean years (SD = 7.43), no-treatment: 47.25 mean years (SD = 7.34)
IQ: Wechsler Memory Scale, 117.8
Sex:: 100% male
Ethnicity: only median given, 1 = Caucasian. It is not stated whether the other possibilities (2 = African American, 3 = Asia, 4 = Hispanic, and 5 = other)
Country**:** USA
Country of treatment (where did the treatment take place): USA

Inclusion criteria

1. "Although many alcoholics are polysubstance abusers, only those who did not habitually abuse other substances were recruited." (p. 42)
2. All subjects had their last drink no longer than 7 to 14 days prior to admissionin to treatment program.

Exclusion criteria

1. "Any subjects with a history of psychosis, head trauma, major medical problems, or CNS abnormalities were excluded." (p. 42-3)
2. "Those with a history of seizures, penetrating head injuries, neurosurgery, and those who have had blunt head trauma causing a loss of consciousness for over 24 hours were excluded from participating in this study." (p. 53)
3. "Those individuals with a long-standing psychiatric history of inpatient treatment were excluded." (p 53)

Comparisons

Physical placebo

Treatment name: Placebo
Description of intervention:"Those subjects in the placebo treatment group listened to several novels on audiotape. During the first two 1-hour training sessions, the examiner supervised the patients to make certain that they were aware of the procedures and could operate the tape machine independently (p. 57)

Pillman 2001 *(Continued)*

Individual or group treatment: individual
Exposure/intensity to treatment: a total of 15 hours. "The examiners saw the patients in the placebo treatment group every 2 to 3 treatment hours to briefly discuss their progress. After 15 hours of placebo treatment, the patients were given the final battery of neuropsychological tests (…)" (p. 58)
Duration of treatment: 22 days
Concomitant psychotherapy: not stated
Concomitant pharmacotherapy: not stated – but the demographic questionnaire obtained information concerning current medications

Wait-list

Comparison name: no treatment (in reality wait-list)

Description of intervention: "All subjects were informed of the 66% possibility that they would not be placed in the treatment group and that if the treatment was found to be successful, they had the right to return to undergo cognitive remediation treatment." (p. 55)

"Subjects in the no-treatment group were given the initial battery of neuropsychological tests upon their agreement to participate. After 14 days, they were administered ANAM v3.11, and after 28 days they were administered the full battery of neuropsychological measures (…) They were allowed to ask questions and were informed as to their right to undergo cognitive remediation if the cognitive remediation was found to be effective." (p. 58)
Exposure/intensity to treatment: no treatment
Duration treatment: 22 days
Concomitant psychotherapy: not stated
Concomitant pharmacotherapy: not stated – but the demographic questionnaire obtained information concerning current medications

Outcomes

Beneficial effect

- **Hierarchy:** usable data, observer-reported, clinical relevance, coin-toss(random.org) between ANAM Efficiency Scores, ANAM Accuracy Scores
- **Outcome chosen:** ANAM continuous performance test

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. "The results did not reveal support for the between-groups effects speculated in the hypotheses. Additionally, the accuracy and efficiency scores showed no statistically significant differences between the three groups over the three trials of cognitive testing." (p. 71)
2. "Previous research has revealed evidence that alcoholics' performance on neurocognitive tests improves significantly over a 30-day period following detoxification. The efficiency data support this conclusion; all three groups improved significantly over the three test trial (…) It is likely that the absence of alcohol in the brain aids in a return of functioning." (p. 71)

Key limitations from study authors

1. "The reasons why the hypotheses were not significant may have come from several sources, including sampling, duration of the study, return of functioning, and use of the cognitive remediation program. The small sample size utilized may not have provided adequate power for the analysis performed. Also, although there was no statistical significance between subjects who completed the study and those who terminated early, clinical observation revealed that several lower functioning individuals terminated early and other lower functioning individuals refused participation in comparison to higher functioning individuals. Thus, the sample used for this study may be slightly higher functioning than the average patient on the inpatient unit." (p. 76-7)
2. "As a result, all testing occurred after 7 days from patients' last period of alcohol use, and recovery of functioning was likely to have begun prior to testing." (p. 77)

Pillman 2001 *(Continued)*

3. "Although the remediation was designed to target deficits specific to alcoholics, the cognitive remediation is typically used over a period of several months. It is possible that the treatment was not utilized long enough to be effective." (p. 77)

Other notes from review authors

1. Usable data is reported in Peterson 2002

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	No	Quote: "Subjects who met the initial criteria for participation in this study were systematically assigned to one of the three groups on an ongoing basis determined by when they entered the treatment program (i.e., Subject 1 was placed in Group 1, Subject 2 was placed in Group 2, Subject 3 was placed in group 3, and Subject 4 was placed in Group 1, etc.)." (p. 55)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	The outcome assessment is blinded, as the ANAM is computer-based
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	No	Quote: "(…) the chi-square analysis of the demographic variables age, ethnicity, years of schooling, and handedness showed that those subjects completing the study did not differ significantly from those who terminated early." (p 60) Attrition >15% (52.6%). No ITT used.
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Poland 2013

Study characteristics

Methods

Parallel randomised trial with three arms

1. Physical placebo: touch
2. No-treatment
3. Active treatment: Swedish massage

Sample calculation: yes

Cluster randomised: no

Duration of trial (baseline to post): 8 weeks

Duration of participation (trial + follow-up): 8 weeks of participation – no follow-up

Poland 2013 *(Continued)*

Setting: outpatient

Purpose of trial: "The study objectives were to determine whether massage therapy reduces symptoms of depression in subjects with human immunodeficiency virus (HIV) disease." (p. 334)

Data

Number of participants screened: 81

Number of participants included: 54

Number of participants followed-up at post treatment: 37

Number of participants randomly assigned to:

- Physical placebo: n = 19
- No-treatment: n = 14
- Active treatment: n = 21

Number of withdrawals: n =17

- Physical placebo: n = 8
- No-treatment: n = 2
- Active treatment: n = 7

Diagnosis: major depressive disorder (in HIV infected subjects)

Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV)

Means of assessment: The Structured Clinical Interview for DSM-IV Disorders (SCID)

Comorbidity: HIV. Otherwise not stated

Age: Pharmacological placebo: 42.6 mean years (SD = 5.1), and No-treatment 42.6 mean years (SD = 4.9)

IQ: not stated

Sex: 14.8% female

Ethnicity: from ITT: African American: n = 6, White: n = 26, Hispanic: n = 8, other: n = 10

Country: USA

Inclusion criteria

1. At least 16 years of age
2. HIV-seropositive
3. Diagnosed with major depressive disorder
4. A score ≥ 15 on the HAM-D at screening
5. Participants had to be on a stable neuropsychiatric, analgesic, and antiretroviral regimen for at least 4 weeks and planning to remain on the same regimen for the 8-week duration of the study
6. Participants had been on a stable antidepressant regimen for > 30 days and the regimen remained fixed for the duration of the study
7. All participants were medically stable as determined by physical examination, full chemistry panels, thyroid function test, electrocardiograms, and urine drug tests

Exclusion criteria

1. Been unable to provided informed consent
2. An unstable medical condition (new opportunistic infection, malignancies, or acute hospitalizations during the past 30 days)
3. Active suicidal ideation or a recent suicide attempt

Poland 2013 *(Continued)*

4. Current or previous diagnosis of anorexia/bulimia nervosa, primary anxiety disorder, bipolar disorder or psychotic disorder
5. Taking any growth hormone or adrenocorticoid preparations
6. Massage therapy or new alternative medicine use in the preceding 30 days
7. History of intolerance to or contraindication to massage.

Comparisons

Physical placebo

Treatment name: Light touch

Description of intervention: "Using a novel dual-control group design,16 which included a light ''touch'' group to control for some of the nonspecific effects of massage and therapist–subject interaction,'' (p. 335)

"The touch group had a massage therapist place both hands on the subject with slight pressure, but no massage, in a uniform distribution for 1 hour twice per week in the same pattern used for the massage subjects. Subjects were told that for study purposes, verbal communication between them and the therapist should be kept to a minimum (...)." (p. 335)

Individual or group treatment: individual

Exposure/intensity to treatment: twice per week for one hour (Monday/Thursday or Tuesday/Friday)

Duration of treatment: 8 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: yes. "Subjects had to be on a stable neuropsychiatric, analgesic, and antiretroviral regimen for at least 4 weeks and planning to remain on the same regimen for the 8-week duration of the study." (p. 335) "Approximately 40% of the subjects also were currently taking antidepressants. As with the antiretroviral regimens, subjects had been on a stable antidepressant regimen for > 30 days and the regimen remained fixed for the duration of the study." (p. 335)

No-treatment

Comparison name: No intervention (no treatment)

Description of intervention: no description of intervention

Exposure/intensity to treatment: no treatment

Duration treatment: 8 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: Yes - "Subjects had to be on a stable neuropsychiatric, analgesic, and antiretroviral regimen for at least 4 weeks and planning to remain on the same regimen for the 8-week duration of the study." (p. 335) "Approximately 40% of the subjects also were currently taking antidepressants. As with the antiretroviral regimens, subjects had been on a stable antidepressant regimen for > 30 days and the regimen remained fixed for the duration of the study." (p. 335)

Outcomes

Beneficial effect

- **Hierarchy:** primary, observer-reported
- **Outcome chosen:** Hamilton Depression Rating Scale (HAM-D)

Adverse events

- No data on adverse events reported

Notes

Key conclusion from study authors

Poland 2013 *(Continued)*

1. ”The trial showed highly significant improvements in relief of depression in those receiving massage versus touch or NI. In general, there was little difference between the ITT and completer analyses for the NI and touch group” (p. 337)
2. “Severity of depression was also assessed using two rating scales: the HAM-D (clinician administered) and the BDI (self-report), and general concordance was found between the results from the two instruments, which further bolstered these findings.” (p. 337)
3. “(…) touch had very little effect over time, particularly in comparison to massage.” (p. 337)
4. “In the present study, about half of the subjects showed a response to massage, defined as a reduction in HAM-D ≥ 50%.” (p. 338)

Key limitations from study authors

1. ”rigorous trials of many CAM interventions including massage are somewhat limited and difficult to interpret because participants in the study are usually not “blind” to the interventions.” (p. 337)
2. “Although subjects received massage for 1 hour twice weekly, the optimal frequency has not yet been defined. (…) Based on the available literature, an intermediate frequency was chosen.” (p. 337)
3. “Most of the subjects in this study were males. There is no a priori reason to believe that there would be sex differences in response, as none of the previous studies of massage described females as being less responsive.” (p. 338)
4. “Some other limitations are the relatively modest sample size and the differences, albeit small, in baseline HAM-D between the NI and touch groups. (…) Finally, since the study required subjects to be able to come to the facility twice a week, it is not clear how selection bias might have impacted the results.” (p. 338)

Other notes from review authors

1. None

Conflicts of interest: one found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: ”Subjects were enrolled by the study coordinators and then randomized using a random numbers table by a nurse practitioner whose sole role in the study was to randomize subjects.” (p. 335)
Allocation concealment	Yes	Quote:(…) randomized using a random numbers table by a nurse practitioner whose sole role in the study was to randomize subjects. Subjects were randomized 1:1:1 into one of three parallel groups” (p. 335)
Blinding of outcome assessors	Unclear	No information
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	No	Quote:“50 subjects completed at least 1 week of the protocol (intent-to-treat; ITT) (…) The last information carried forward (LOCF) method was used for analyses of the ITT group.” (p. 335). Attrition >15% (31.5%)
Selective outcome reported	Unclear	NCT00033852. No apparent differences in reporting between trail registry and full report, but information about outcome measures and eligibility criteria was not available.

Poland 2013 *(Continued)*

Other sources of bias	Yes	No other sources found

Powers 2004

Study characteristics

Methods

Parallel randomised trial with five arms

1. Psychological placebo: credible psychological
2. Wait-list
3. Active treatment 1: exposure only
4. Active treatment 2: exposure with safety-behaviour utilisation
5. Active treatment 3: exposure with safety-behaviour availability

Sample calculation: not stated

Cluster randomised: no

Duration of trial (baseline to post): 1 day: ("They were also told that one effective strategy for reducing their fear is to be exposed to the feared situation repeatedly until the anxiety decreases. Participants in the three exposure conditions received a total of 30 min of self-guided in vivo exposure to the claustrophobic chamber used for BAT 1." (p. 450))

Duration of participation (trial + follow-up): 1 day + 2 weeks follow-up

Setting: outpatient

Purpose of trial: "The primary aim of the current study was to further investigate the deleterious effects of safety-seeking behaviors on fear reduction by disentangling the effects of perceived availability of threat-relevant safety behaviors during treatment versus their actual use." (p. 449)

Open/closed placebo: closed placebo

Data

Number of participants screened: pool of approximately 5000

Number of participants included: 72

Number of participants followed-up at post treatment: 71

Number of participants randomly assigned to: 71

- Psychological placebo: n = 12
- Wait-list: n= 15
- Active treatment 1: n = 17
- Active treatment 2: n = 16
- Active treatment 3: n = 11

Number of withdrawals: n = 1

- Psychological placebo: not stated
- Wait-list: not stated
- Active treatment 1: not stated
- Active treatment 2: not stated
- Active treatment 3: not stated

Diagnosis: claustrophobia

Diagnostic manual : 75% met full Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV), 25% met all DSM-IV criteria with the exception of Criterion E (i.e.the person must experience

Powers 2004 *(Continued)*

	significant interference in social, academic, or work functioning or experience marked distress about having the phobia) **Means of assessment:** first potential participants rated their overall fear on a 5-point Likert scale, afterwards participants were administered the CIDI-Auto **Comorbidity:** not stated **Age:** 21.06 mean years (SD = 5.08), (range 18 to 49 years) **IQ:** not stated **Sex:** 86% female **Ethnicity:** 74% Caucasian, 13% Mexican American, 7% African American, 5% Asian American, and 1% Indian American **Country:** USA **Inclusion criteria** 1. Claustrophobia **Exclusion criteria** 1. "Individuals who refused to attempt either BAT or reported only mild fear during either BAT 1 or BAT 2 (i.e., less than 50 on a 100-point Likert Scale) were deemed insufficiently phobic and excluded from the study." (p. 450)
Comparisons	**Psychological placebo** **Treatment name:** Placebo **Description of intervention:** "Participants in the PL group returned 2 weeks after completing screening and received a similar rationale (…)". (p. 450). "The DAVID developed by Comptronic Devices is used by health care professionals as a relaxation device. I is a small soundboard about the size of a stereo receiver, which includes a headset and plastic mask. The headset emits controllable ticking sounds, similar to those made by a metronome. The plastic mask resembles ski goggles and delivers pulsed orange lights at controllable rates. In this study, the audio and video stimulus frequency was set at 12 Hz (cycles per second), which is the rate at which the device is suggested to maximally produce relaxation and meditative states." (p. 450) **Individual or group treatment:** individual **Exposure/intensity to treatment:** one treatment **Duration of treatment:** one treatment **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated **Wait-list** **Comparison name:** Wait-list **Description of intervention:** "This group was informed that they had been placed on a WL. They returned for assessment 2 weeks later and completed the post-assessment. Following assessment, they received exposure treatment." (p. 450) **Exposure/intensity to treatment:** no treatment **Duration treatment:** no treatment **Concomitant psychotherapy:** not stated

Powers 2004 *(Continued)*

	Concomitant pharmacotherapy: not stated
Outcomes	**Beneficial effect** • **Hierarchy:** patient-reported, clinical relevance, global score, psychometric properties • **Outcome chosen:** CLQ: total **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. ”Consistent with previous research making safety behaviors available to claustrophobic individuals during in vivo exposure had a marked disruptive effect on fear reduction. The magnitude of this effect at posttreatment was considerable, as evidenced by the 94% versus 45% treatment response rate for those in the EO condition versus the two exposure-plus-safety-behavior conditions.” (p. 453) **Key limitations from study authors** 1. “Several limitations deserve comment. First, although we used a stringent two-stage screening procedure to ensure that study participants display marked phobicity (…), 25% of the participants did not meet DSM-IV criteria for specific phobia.” (p. 453) 2. “Finally, the follow-up period of 2 weeks was too brief to make inferences about the stability of treatment effects over the long term.” (p. 453) **Other notes from review authors** 1. None **Conflicts of interest: n**one found **Judgement:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote:”Eligible participants were randomly assigned to one of five conditions: (a) EO, (b) exposure with SBU, (c) exposure with SBA, (d) credible psychological PL, or (e) WL control.” (p. 449)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	Yes	Attrition <15% (1,41%)
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Powers 2008a

Study characteristics

Methods	**Parallel randomised trial with four arms** 1. Psychological placebo 2. Wait-list 3. Pharmacological placebo: exposure + inactive pill 4. No-treatment: exposure only **Sample calculation:** yes **Cluster randomised:** no **Duration of trial (baseline to post):** 1 session (1 day) **Duration of participation (trial + follow-up):** 1 day + 1-week follow-up **Setting:** outpatient **Purpose of trial:** "On the basis of the available evidence, we hypothesized the following: (a) Participants led to believe that they ingested a sedating herb with anxiety dampening effects would show significantly greater return of fear compared with those led to believe that they ingested a placebo; (b) participants led to believe that they ingested a stimulating herb with anxiety enhancing effects would show significantly enhanced maintenance of treatment gains at follow-up compared with those led to believe that they ingested a placebo; and (c) the effects of the pill expectancy manipulation on changes in fear during the follow-up period would be mediated by changes in coping self-efficacy." (p. 479) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** 5326 **Number of participants included:** 95 **Number of participants followed-up at post treatment:** 95 **Number of participants randomly assigned to:** • Psychological placebo: n = 15 • Wait-list: n = 15 • Pharmacological placebo: n = 55 • No-treatment: n = 15 **Number of withdrawals:** n = 0 **Diagnosis** : Claustrophobia **Diagnostic manual:** Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) **Means of assessment:** Composite International Diagnostic Interview (CIDI-Auto). Most participants (74%) met full Diagnostic and Statistical Manual of Mental Disorders (4th ed.; DSM–IV; APA 1994) criteria for claustrophobia, whereas 26% met all DSM–IV criteria with the exception of Criterion E, which requires that the person experience significant interference in social, academic, or work functioning or marked distress about having the phobia **Comorbidity** : not stated **Age:** 20.11 mean years (SD = 6.23), (range = 18 to 60) **IQ:** not stated **Sex:** 71 % female

Powers 2008a *(Continued)*

Ethnicity: the ethnic breakdown of the sample was 73% Caucasian, 12% Hispanic, 9% African American, 4% Asian, and 2% Native American.

Country: the Netherlands

Inclusion criteria

1. Reporting moderate or greater fear of enclosed spaces as defined by a rating of 2 or higher

Exclusion criteria

1. Individuals who refused to attempt either BAT (n 4) or who reported a fear level less than 50 during either BAT-1 or BAT-2 (n 57) were excluded from the study

Comparisons	**Psychological placebo** **Treatment name:** Credible Psychological Placebo Treatment **Description of intervention:** "Participants assigned to the psychological placebo condition returned 1 week following pretreatment assessment to receive 30 min of pulsed audio-photic stimulation with a device called the Digital Audio Integration Device (DAVID) Paradise XL (Mind Alive Inc., Edmonton, Alberta, Canada). It consists of a headset, which emits controllable pulsing sounds, and plastic goggles, which produce pulsing lights at controllable rates. The number of treatment trials (N 6), the size and layout of the treatment room, the position of the participant (supine), and the duration of each trial (5 min) were equivalent to those receiving the exposure treatment. However, they received no exposure treatment." (p. 482) **Individual or group treatment:** individual **Exposure/intensity to treatment:** 1 session - 30 minutes **Duration of treatment:** 1 day (1 session) **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated **Wait-list** **Comparison name** : Wait-list **Description of intervention:** "Participants in the waitlist condition completed assessments at each of the three time points and were offered exposure treatment following study completion." (p. 482) **Exposure/intensity to treatment:** 1 session **Duration treatment:** 1 day **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy** : patient-reported, clinical relevance, global score, psychometric properties • **Outcome chosen:** CLQ: total **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. Return of fear rates for the 3 conditions were 39%, 0%, and 0%, respectively 2. Moreover, the deleterious effects of the sedation instructions were mediated by reduced self-efficacy

Powers 2008a *(Continued)*

3. These findings highlight the importance of assessing patient attributions regarding the improvements achieved with combined exposure-based and pharmacological treatments for anxiety disorders

Key limitations from study authors

1. First and foremost, findings from this analogue investigation need to be replicated within the context of a randomized controlled trial employing more severe clinical samples, higher treatment doses, and an expanded range of outcomes
2. The follow-up period of 1 week was too brief to make inferences about the stability of the effects over time
3. Although a generalization probe was included (BAT-2), it is unclear to what extent these findings would generalize to other claustrophobic situations, such as riding elevators and subways
4. The primary manipulation took place after the posttreatment assessment. Although the posttreatment manipulation has the advantage of disentangling expectancy and attribution effects, it deviates markedly from common clinical practice in which patients are provided expectations about the effects of medications at the commencement of treatment

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote:"Research assistants enrolled and randomized participants by cycling through a list consisting of a computer generated random sequence of the four treatment conditions (...). Three times as many participants were randomized to the exposure treatment plus inactive pill condition in anticipation of later randomization to the three perceived pill effect conditions." (p. 482) " Analyses showed no significant differences between groups at baseline on any of these measures (all ps = .20), suggesting that randomization was successful." (p. 483)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Attrition <15% (0%)
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Powers 2008b

Study characteristics

Methods	See Powers 2008a

Powers 2008b *(Continued)*

Data	See Powers 2008a
Comparisons	**Pharmacological placebo** **Treatment name:** Exposure + inactive pill **Description of intervention:** "Prior to the start of exposure treatment, they were administered an inactive pill of 250 mg of Vitamin C and told that the experiment would be investigating an anxiety treatment while simultaneously examining the effects of an herbal supplement—"Adomoxin" (a fictitious name)— on memory." (p. 482) **Individual or group treatment:** individual **Exposure/intensity to treatment** : Six 5-minute trials **Duration of treatment:** 1 day (1 session) **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated **No-treatment** **Comparison name:** Exposure + no pill **Description of intervention:** "In brief, this treatment consisted of several elements, including (a) brief education about the nature of claustrophobia, (b) rationale for exposure treatment, (c) six 5-min trials of in vivo exposure to a claustrophobic chamber identical to that used in the BAT-1 assessment, and (d) completion of treatment process ratings before and after each exposure trial." (p. 482) **Exposure/intensity to treatment:** Six 5 min trials **Duration treatment:** 1 day **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated
Outcomes	See Powers 2008a
Notes	See Powers 2008a

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote:"Research assistants enrolled and randomized participants by cycling through a list consisting of a computer generated random sequence of the four treatment conditions (...). Three times as many participants were randomized to the exposure treatment plus inactive pill condition in anticipation of later randomization to the three perceived pill effect conditions." (p. 482) " Analyses showed no significant differences between groups at baseline on any of these measures (all ps = .20), suggesting that randomization was successful." (p. 483)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment

Powers 2008b *(Continued)*

Incomplete outcome data	Yes	Attrition <15% (0%)
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Proudfoot 2013

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Physical placebo: attention control 2. Wait-list 3. Active treatment: myCompass intervention **Sample calculation:** yes **Cluster randomised:** no **Duration of trial (baseline to post):** 7 weeks **Duration of participation (trial + follow-up):** "Eligible participants completed a baseline questionnaire prior to randomization, a post-intervention questionnaire administered at eight weeks, and a follow-up questionnaire administered 12 weeks later for participants in the myCompass and AC groups, and 19 weeks later for the WL group." (p. 4) **Setting:** outpatient **Purpose of trial:** "The aim of this paper is to report the outcomes of a CONSORT-compliant randomised controlled trial (RCT) to evaluate the efficacy of the myCompass program in a large community sample of people experiencing mild-to-moderate depression, anxiety and/or stress. We predicted that symptoms of depression, anxiety and stress would reduce in participants randomly allocated to receive myCompass, relative to both attention control (AC) and waitlist (WL) conditions. We also predicted that use of myCompass would increase work and social functioning relative to the AC and WL conditions." (p. 2) **Open/closed placebo:** Closed placebo
Data	**Number of participants screened:** 2955 **Number of participants included:** 720 **Number of participants followed-up at post treatment:** 515 **Number of participants randomly assigned to:** • Psychological placebo: n = 248 • Waiting list: n = 230 • Active treatment: n = 242 **Number of withdrawals:** n = 205 • Psychological placebo: n = 55 • Waiting list: n = 34 • Active treatment: n = 116 **Diagnosis:** depression, anxiety or/and stress

Proudfoot 2013 *(Continued)*

Diagnostic manual: not stated

Means of assessment: Depression, Anxiety and Stress Scales (DASS)

Comorbidity: 65% had comorbid symptoms. Not specified which

Age: psychological placebo: 40 (SD = 11.42), wait-list: 38 (SD = 10.26)

IQ: not stated

Sex: 68.2% female

Ethnicity: not stated

Country: Australia

Inclusion criteria

1. Australian resident aged 18 to 75
2. Own an internet-enabled mobile phone
3. Have access to a desk-top computer with internet capability
4. Have a valid email address
5. Report symptoms of mild-to-moderate depression, anxiety and/or stress, defined as a total score of 27-63 inclusive on the Depression Anxiety and Stress Scales (DASS)

Exclusion criteria

1. Score at 64 or more on the DASS (severe symptomatology)
2. Answered positively to questions asking about suicidal thoughts, intent and/or previous suicide attempts
3. Met criteria for psychotic symptoms, as measured by the Psychosis Screening Questionnaire

Comparisons

Physical placebo

Treatment name : Attention control

Description of intervention: "myCompass is a fully-automated, self-help, public health intervention, that is tailored to the user and has no therapist input." (p. 3)

“Attention control participants received a control mental health program matched to the active intervention on duration and mode of delivery. Each week for seven weeks, they received a fact sheet containing information about depression, anxiety or stress sent to their email address. The information was designed to be read on computer in approximately 10 minutes, and to be credible but void of management advice or treatment strategies. They also received on their mobile phones weekly SMS messages containing brief factual statements about depression, anxiety and stress. The mobile phone statements were also therapeutically inactive, but chosen to ensure that the control program had face validity.” (p. 3-4)

Individual or group treatment: individual

Exposure/intensity to treatment: each week they received a fact sheet and weekly SMS messages containing brief factual statements

Duration of treatment: 7 weeks

Concomitant psychotherapy: not excluded if they did, but no records (author correspondence)

Concomitant pharmacotherapy: probably but not stated. “(…) and use of antidepressant and anxiolytic medication were also assessed.” (p. 4)

Wait-list

Comparison name: Wait-list

Proudfoot 2013 *(Continued)*

Description of intervention: "Waitlist participants did not receive emails or SMSs during the intervention phases, but received full access to the myCompass program at the end of the seven weeks." (p. 3)

Exposure/intensity to treatment: no treatment

Duration treatment:: 7 weeks

Concomitant psychotherapy: not excluded if they did, but no records (author correspondence)

Concomitant pharmacotherapy: probably but not stated. "(…) and use of antidepressant and anxiolytic medication were also assessed." (p. 4)

Outcomes

Beneficial effect

- **Hierachy:** primary, patient-reported, total score
- **Outcome chosen:** the Depression, Anxiety and Stress Scales (DASS)

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. "At post-intervention, the myCompass group showed significantly reduced symptoms of depression, anxiety and stress, and significantly improved levels of work and social functioning (…) Scores reduced to the normal range by post-intervention and treatment gains were maintained at 3-month follow-up. Participants in the AC condition showed gradual improvement over the post-intervention period and no differences were observed between myCompass and AC participants at 3 month follow-up" (p. 9)
2. "The pattern of symptom improvement observed for the AC group is similar to the natural course of symptom remission over several months observed in untreated depression and anxiety. In contrast, the myCompass intervention accelerated symptom remission to within two months, producing rapid benefit for those-in-need, and with effect sizes predominantly in the moderate range." (p. 10)

Key limitations from study authors

1. (…) dropout attrition was high, especially for the myCompass group, and rates of engagement for myCompass participants with the program content were highly variable (and in some instances minimal). Inspection of possible biases due to attrition showed that dropouts were more likely to be male and employed, thus reducing our confidence in generalising to these groups." (p. 10)
2. "Future research is needed to isolate the relative contributions of the mobile phone (e.g., self-monitoring, SMS messages) and computer-based (e.g., psychoeducational modules) elements of the intervention." (p. 11)

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote:"A research assistant not involved in the RCT randomised participants after baseline using computerised random numbers. Allocation was either to the myCompass, AC or WL condition." (p. 3)
Allocation concealment	Yes	Quote:"A research assistant not involved in the RCT randomised participants after baseline using computerised random numbers." (p. 3)

Proudfoot 2013 *(Continued)*

Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	Unclear	Quote:"Effects of the myCompass intervention on study outcomes were evaluated using intention-to-treat (ITT) analyses that included data from all participants who completed the baseline assessment and any follow-up assessment. Strategies for dealing with missing data in longitudinal studies vary, so we adopted two recommended techniques for analysing incomplete datasets, namely, mixed models repeated measures (MMRM) and multiple imputation (…)" (p. 4) Attrition >15% (MyCompass: 47.9%, Attention control: 22.2%, Wait-list: 14.8%)
Selective outcome reported	No	ACTRN 12610000625077 Differences between protocol and report are found 1) 6 weeks of treatment in the protocol, 7 weeks in the report. 2) According to the protocol the assessments are made at baseline, and at 3, 6, 9, 12, 15 and 18 weeks after intervention commencement – in the report assessments are reported for baseline, 8 weeks (post treatment) and a 12 weeks follow-up 3) It is stated in the protocol that masking is used – but this is not stated in the report
Other sources of bias	Yes	No other sources found

Quayhagen 1995

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Psychological Placebo: passive cognitive stimulation 2. Wait-list: no stimulation, but post study introduction to treatment 3. Active treatment: active cognitive stimulation **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 12 weeks **Duration of participation (trial + follow-up):** 12 weeks + 6 months **Setting:** outpatient **Purpose of trial:** "The objective of this study was to determine the impact over time on the cognitive and behavioral functioning of the care recipient from a home-based intervention program of active cognitive stimulation implemented by the family caregiver." (p. 154) **Open/closed:** closed placebo
Data	**Number of participants screened:** 132 family units **Number of participants included:** 95 family units **Number of participants followed-up at post treatment:** 79 (only 78 retained for analysis – one family unit was eliminated due to data inconsistency)

Quayhagen 1995 *(Continued)*

Number of participants randomly assigned to:

- Psychological Placebo: n = 28
- Wait-list: n = 25
- Active treatment: n = 25

Number of withdrawals: n = 16

- Psychological Placebo: not stated
- Wait-list: not stated
- Active treatment: not stated

Diagnosis : Alzheimer's disease

Diagnostic manual: not stated

Means of assessment : Mattis Dementia Rating Scale

Comorbidity: not stated

Age: 73.6 mean years (SD = 8.0)

IQ: not stated

Sex: 28.4% female

Ethnicity: white (85%), African American (3%), Hispanic (11%)

Country: USA

Inclusion criteria

1. Having a care recipient with a confirmed diagnosis of possible or probable Alzheimer's disease, with mild to moderate decline, in accord with the clinical criteria for these stages, and a score of 90 or better on the Mattis Dementia Rating Scale

Exclusion criteria

1. Not stated

Comparisons

Psychological placebo

Treatment name: placebo

Description of intervention:"With neuropsychological consultation, the decision was made to select activities similar to those in the experimental condition, but using a passive approach. Caregiver implementation relied heavily of the modeling work of Bandura (1977), where the impaired member was exposed to passive observation of the activity without enforced participation." (p. 156)

Individual or group treatment: individual

Exposure/intensity to treatment: "All families in the treatment groups attended 12 consecutive weekly in-home sessions with members of the intervention team. The caregiver and the care recipient were trained together in program implementation techniques." (p. 155). "Following each weekly instruction session, the intervention was executed in the home by the family caregiver." (p. 153)

Duration of treatment: 12 weeks.

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: "None were participating in pharmacological clinical trials." (p. 154)

Wait-list

Comparison name: Wait-list control

Quayhagen 1995 *(Continued)*

Description of intervention: "The families assigned to the control group were placed on a waiting list for complementary sessions on the cognitive stimulation program once the wave of the study in which they were participating was completed." (p. 156)

Exposure/intensity to treatment: no treatment

Duration treatment: 12 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: "None were participating in pharmacological clinical trials." (p. 154)

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported, clinical relevance, global score
- **Outcome chosen:** The Mattis Dementia Rating Scale (general cognitive functioning)

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. "The hypothesis was supported in that the experimental group was at or near baseline by the 9th month on the cognitive and behavioral outcomes. Also as predicted, the control group consistently declined, while, contrary to expectation, the placebo group only noted decline on selected outcomes and maintained at baseline on others." (p. 156)
2. What was not anticipated was the finding of cognitive improvement in the experimental care recipients post treatment; only maintenance was hypothesized because of the trajectory of decline in dementia." (p. 156)

Key limitations from study authors

1. "Through log recordings of the caregivers and the observation of the investigative team, it became apparent early in the study that a group of the care recipients in the placebo condition had exceeded the passivity parameters and were working toward self-initiated improvement. A plausible explanation for the self-initiation is that the active and passive intervention protocols may have been too similar despite all attempts to keep them distinctly active versus distinctly passive in mode and execution." (p. 156)

Other notes from review authors

1. Only completers retained for analysis not all included in the study

Conflicts of interest: one found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "(...) the care recipients were stratified by degree of cognitive impairment to maintain initial comparability of functioning across groups. They were then randomly assigned to one of three conditions: (a) intervention (active cognitive stimulation), (b) placebo (passive cognitive stimulation), and (c) wait-list control (no stimulation, but poststudy introduction to treatment)." (p. 155)
Allocation concealment	Unclear	No information

Quayhagen 1995 *(Continued)*

Blinding of outcome assessors	Unclear	Quote: "To control for experimenter bias, assessments were conducted by research assistant who, with rare exception, were blinded to the condition to which the family had been assigned." (p. 155)
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	No	Attrition >15% (17%). No mention of ITT
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	No	Quote: "Return demonstrations by caregivers were required to validate training. At each session, the caregiver was reminded to give positive feedback and to complete a weekly log that included success or problems in implementation and the amount of time spent each day with the intervention." (p. 155). Quote: "Through log recordings of the caregivers and the observation of the investigative team, it became apparent early in the study that a group of the care recipients in the placebo condition had exceeded the passivity parameters and were working toward self-initiated improvement." (p. 156)

Rabkin 1990

Study characteristics	
Methods	**Parallel randomised trial with two arms** 1. Pharmacological placebo 2. No-treatment: placebo stopped **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 6 weeks **Duration of participation (trial + follow-up):** 12 week follow-up **Setting:** outpatient **Purpose of trial:** to evaluate pill-taking as determinant of placebo response **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** not stated **Number of participants included:** 50 **Number of participants followed-up at post treatment:** 50 **Number of participants randomly assigned to:** • Pharmacological placebo: n = 27 • No-treatment: n = 23 **Number of withdrawals:** n = 0 **Diagnosis:** major depression **Diagnostic manual:** Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition (DSM III) **Means of assessment:** Hamilton rating scale for depression. **Comorbidity:** dysthymia = 18, major depressive disorder (MDD) =17, dysthymia + MDD = 15 **Age:** 37 mean years (SD = 10) (Range = 15 to 55) **IQ:** not stated **Sex:** 54% female **Ethnicity:** not stated **Country:** USA

Rabkin 1990 *(Continued)*

	Inclusion criteria 1. DSM III for major depression or dysthymia 2. Reactive mood 3. Hamilton Depression scores at least 12 **Exclusion criteria** 1. Not stated
Comparisons	**Pharmacological placebo** **Treatment name:**: Placebo **Description of intervention:** single-blind pill placebo responders **Individual or group treatment:** Individual **Exposure/intensity to treatment:** not stated **Duration of treatment:** 6 weeks. **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated **No-treatment** **Comparison name:** No-treatment **Description of intervention:** received no intervention **Exposure/intensity to treatment:** no treatment **Duration treatment:** 6 weeks **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy:** usable data, clinical relevance • **Outcome chosen:** relapse **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. Half of the patients in each condition relapsed within 6 weeks, indicating that pill-taking itself does not influence maintenance of placebo response. Placebo response was more likely to be maintained in patients who were currently married. At the end of 3 months, the overall relapse rate was 58% **Key limitations from study authors** 1. Not include severely depressed, melancholic, acute depression **Other notes from review authors** 1. None **Conflicts of interest:** none found **Judgement:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: "The randomization process followed the order provided by our statistician in blocks of 4. I'm really not sure, but likely used a predefined list. (author correspondence)

Rabkin 1990 *(Continued)*

Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote: "All assessors were blinded" (author correspondence)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Quote: "Over the 5-year period of the study, 58 patients were rated as 10-day placebo responders; these patients constituted 10% of all patients who received single- blind medication in clinical trials during this period. Four of the 58 patients refused to participate in the study, one moved out of the area during the trial, one refused to stop taking the single-blind placebo pills, one dropped out, and one denied 2 weeks later that he had been a placebo responder in the first place. The eight patients who did not complete the study did not differ from the 50 who did complete the study with respect to demographic characteristics, illness history, baseline illness characteristics, baseline symptoms, or severity of depression." (p. 1623) Attrition <15% (0% in the second trial)
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Rapee 2006

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Usual care: standard group treatment 2. Wait-list 3. Active treatment: bibliotherapy **Sample calculation:** yes **Cluster randomised:** no **Duration of trial (baseline to post):** 12 weeks of treatment **Duration of participation (trial + follow-up):** 12 weeks of treatment + 24 weeks follow-up **Setting:** outpatient **Purpose of trial:** "The aims of the current study was to examine the impact of using parents as therapists for their own child in a trial of bibliotherapy materials for parents of children with anxiety disorders." (p. 437)
Data	**Number of participants screened:** not stated **Number of participants included:** 267 **Number of participants followed-up at post treatment:** 212 **Number of participants randomly assigned to** • Usual care: n = 90

Rapee 2006 *(Continued)*

- Waitlist: n = 87
- Active treatment: n = 90

Number of withdrawals : n = 55

- Usual care: n = 14
- Waitlist: n = 12
- Active treatment: n = 29

Diagnosis: anxiety disorder

Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV)

Means of assessment: Anxiety Disorders Interview Schedule, Parent and Child Versions (ADIS-CP)

Comorbidity: main comorbid diagnoses = anxiety disorder (N = 219, 82%), externalising disorder (N = 72, 27%), mood disorder (N = 23,8.6%). Generalised anxiety disorder (N = 103). Social phobia (N = 64). Separation anxiety disorder (N = 51). Specific phobia (N = 33). Obsessive-compulsive disorder (N = 13) Panic disorder (N = 3)

Age: age in months: usual care: 113.7 (SD = 20.4) Waitlist: 114.1 (SD = 19.1)

IQ: not stated

Sex: usual care: 53.3% female. Wait-list: 29.9% female

Ethnicity: not stated

Country: Australia

Inclusion criteria

1. Years 1 through 6 at school (ages 6 to 12 years)
2. Met criteria for an anxiety disorder as their principal (most interfering) disorder
3. Their parent or parents were able to read a standard, English-language newspaper

Exclusion criteria

1. Children with non-anxiety disorder if these demanded immediate attention

Comparisons

Usual care

Treatment name: Standard group treatment

Description of intervention: “Group treatment was based on the Cool Kids Program, a nine-session cognitive-behavioral program for the management of broad-based childhood anxiety disorders.” (p. 437)

“Parents and children attend all nine sessions of the program on a weekly basis over 12 weeks (the final few sessions are biweekly) and cover recognition of emotion and anxiety, realistic thinking, child management strategies, exposure to feared cues, and additional skills such as assertiveness and dealing with teasing. (p. 438)

Individual or group treatment: group treatment

Exposure/intensity to treatment: nine sessions, approximately 2 hours per session

Duration of treatment: 12 weeks

Concomitant psychotherapy: “Generally we allow concomitant medication after stabilization; and we allow psychotherapy for unrelated problems also after stabilization” (author correspondence)

Concomitant pharmacotherapy: yes. “Children on medication were included if the medication had been stable for the previous month.” (p. 437) “Generally we allow concomitant medication after sta-

Rapee 2006 *(Continued)*

bilization; and we allow psychotherapy for unrelated problems also after stabilization" (Rapee 2006 (pers comm))

Wait-list

Comparison name: Wait-list

Description of intervention: "Participants in the waitlist were simply told that they had been randomly assigned to wait for treatment and that they would be recontacted for additional assessments in 3 months' time, after which they would be offered the next available treatment group." (p. 438)

Exposure/intensity to treatment: no treatment during waiting

Duration treatment: 12 weeks

Concomitant psychotherapy: "Generally we allow concomitant medication after stabilization; and we allow psychotherapy for unrelated problems also after stabilization" (author correspondence)

Concomitant pharmacotherapy: yes. "Children on medication were included if the medication had been stable for the previous month." (p. 437). "Generally we allow concomitant medication after stabilization; and we allow psychotherapy for unrelated problems also after stabilization" (Rapee 2006 (pers comm))

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported, clinical relevance
- **Outcome chosen:** ADIS-CP – diagnostic severity

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. Overall, the results of the current trial demonstrate that children whose parents received bibliotherapy with no therapist contact improved somewhat more than children on wait-list after 12 weeks and that these results maintained up to 3 months (…) On the basis on structured clinical interviews, bibliotherapy was significantly better than no treatment according to both completer and intention-to-treat analyses
2. The results also show that standard cognitive-behavioral group treatment with a therapist resulted in greater change than bibliotherapy according to both clinician and parent reports. Therefore, these results do not suggest a replacement of traditional models of therapy but do suggest a potential alternate model of treatment delivery under appropriate circumstances.

Key limitations from study authors

1. First, the structured interviews were not technically administered
2. More important to note is that the lack of separate diagnoses meant that there was no independent validation for the self-reports from parents and children
3. An additional limitation is the fact that parents did not complete data on compliance or preference for bibliotherapy. Such data would be important to more fully understanding the implications and benefits of self-help and should be included in any future studies
4. Perhaps the main limitation of the study is the fact that the sample for the study came from a traditional, specialist anxiety clinic. This was necessary to allow properly controlled scientific design and group treatment comparison. However, this recruitment means that we cannot be certain whether those families who do not seek traditional forms of therapy would benefit from bibliotherapy

Other notes from review authors

1. None

Conflicts of interest: none found

Rapee 2006 *(Continued)*

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: "Randomization occurred in block of eight to allow allocation to group treatment based on a predetermined random number schedule known only to the study coordinator." (p. 438)
Allocation concealment	Yes	Quote: "(...) known only to the study coordinator." (p. 438)
Blinding of outcome assessors	Yes	Quote: "Repeated interviews were conducted by clinicians who were masked to the child's allocated treatment (…)" (p. 437)
Blinding of participants and personnel	No	Not possible to blind usual care and wait-list
Incomplete outcome data	No	Quote: "Analyses were made on completers and on all randomized (intent-to-treat). "Intention-to-treat analyses included all participants who were allocated to a condition (aside from 7 participants who did not return any data at pretreatment) and used the last-point-carried-forward method to deal with missing data." (p. 439) Attrition >15% (20.6%). LOCF used. Only reports data on completers
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Rapee 2007

Study characteristics

Methods

Parallel randomised trial with four arms

1. Usual care: standard group treatment
2. Wait-list: waiting list
3. Active treatment 1: 'pure' self-help
4. Active treatment 2: self-help augmented by therapist assistance

Sample calculation: not stated

Cluster randomised: no

Duration of trial (baseline to post): 12 weeks

Duration of participation (trial + follow-up): 12 weeks of treatment + 24 weeks follow-up

Setting : outpatient

Purpose of trial: "Our study was designed to determine the value of two forms of self-help through the use of bibliotherapeutic materials in the reduction of social phobia: pure bibliotherapy that involved almost no contact with the researchers, and therapist-augmented bibliotherapy in which printed material was supplemented with five group sessions conducted by a therapist. Benchmarks for these condi-

Rapee 2007 *(Continued)*

	tions were provided by comparisons with a no-treatment waiting list and standard ten-session group therapy conducted by therapist." (p. 246)
Data	**Number of participants screened:** not stated **Number of participants included:** 224 **Number of participants followed-up at post treatment:** 177 **Number of participants randomly assigned to:** • Usual care: n = 59 • Wait-list: n = 52 • Active treatment 1: n = 56 • Active treatment 2: n = 57 **Number of withdrawals** : n = 47 • Usual care: n = 14 • Wait-list: n = 10 • Active treatment 1: n = 8 • Active treatment 2: n = 15 **Diagnosis:** 95.7% met criteria for generalized subtype of social phobia and 55.8% met criteria for a diagnosis of avoidant personality disorder **Diagnostic manual:** Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) **Means of assessment:** Anxiety Disorders Interview Schedule + International Statistical Classification of Diseases and Related Health Problems, 10th edition (ICD-10) International Personality Disorder Examination **Comorbidity:** 42.9% met criteria for an additional anxiety disorder, 33.9% met criteria for an additional mood disorder and 4.0% met criteria for an additional substance use or alcohol disorder **Age:** 35.5 mean years (SD = 11.0) **IQ:** not stated **Se:**: 50.4% female **Ethnicity:** not stated **Country:** Australia **Inclusion criteria** 1. Aged 20 to 65 years 2. Met criteria for social phobia as their main (or most interfering) disorder 3. Had sufficient English and education to read a tabloid newspaper in English **Exclusion criteria** 1. Problems requiring immediate attention such as clear suicidal intent, severe substance misuse or dependence, or florid psychosis
Comparisons	**Usual care** **Treatment name:**Standard group treatment **Description of intervention:** "Treatment was conducted in groups of approximately six participants (..) Therapy extended for ten 2 h sessions across 12 weeks. (...) Components included those typically found in empirically validated treatments for social phobia including cognitive restructuring of nega-

Rapee 2007 *(Continued)*

tive evaluation beliefs, exposure to feared social situations, realistic feedback of social performance, and attention training. Participants engaged in home exercise and received various handouts as relevant." (p. 247)

Individual or group treatment: group

Exposure/intensity to treatment: 10 two-hour sessions

Duration of treatment: 12 weeks

Concomitant psychotherapy: no participant was in concurrent psychotherapy but they were allowed to. "Concurrent pharmacotherapy or psychotherapy was allowed as long as dosage had been consistent for 3 months and there was no plan to change." (p. 246)

Concomitant pharmacotherapy: "Concurrent pharmacotherapy or psychotherapy was allowed as long as dosage had been consistent for 3 months and there was no plan to change." (p. 246)."(…) 6.8% were taking benzodiazepines or other anxiolytics, 21.2% were taking selective serotonin reuptake inhibitors or other antidepressants and 9.9% were taking other prescription medications." (p. 246)

Wait-list

Comparison name: Waiting list

Description of intervention: "Participants on the waiting list were told that they had been randomly allocation to receive no treatment for 12 weeks. At the end of the 12-week period they were offered out best available treatment." (p. 247)

Exposure/intensity to treatment: no treatment during waiting

Duration treatment: 12 weeks

Concomitant psychotherapy: no participant was in concurrent psychotherapy but they were allowed to. "Concurrent pharmacotherapy or psychotherapy was allowed as long as dosage had been consistent for 3 months and there was no plan to change." (p. 246)

Concomitant pharmacotherapy: "Concurrent pharmacotherapy or psychotherapy was allowed as long as dosage had been consistent for 3 months and there was no plan to change." (p. 246) "(…) 6.8% were taking benzodiazepines or other anxiolytics, 21.2% were taking selective serotonin reuptake inhibitors or other antidepressants and 9.9% were taking other prescription medications." (p. 246)

Outcomes

Beneficial effect

- **Hierarchy:** clinical relevance, psychometric properties, random (between SPS and SIAS)
- **Outcome chosen:** Social Phobia Scale (SPS)

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. The results provided mixed support for the value of bibliotherapy in reducing both social fears and the degree of life interference caused by social anxiety. Specifically, the extent of the reductions was markedly influenced by the method of delivering bibliotherapy (…) Hence as a clinical intervention, pure bibliotherapy appears to show limited value for social phobia
2. In contrast to pure self-help, augmentation of self-help with five therapist-led group sessions resulted in marked improvements in symptoms of social phobia and life interference that were as great as those produced by standard group treatment

Key limitations from study authors

1. Not stated

Other notes from review authors

Rapee 2007 *(Continued)*

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: "Randomisation was done using a pre-assigned random number generator in blocks of eight to allow for group delivery." (p. 247)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote: "As much as possible. That is – we do not tell assessor about the condition a patient is in and we instruct patients not to describe their treatments. But of course in some cases they do talk about it" (Rapee 2007 (pers comm))
Blinding of participants and personnel	No	Not possible to blind usual care and wait-list
Incomplete outcome data	Unclear	Quote: "Interpolation was used if post-treatment data only were not available. As a precaution against biasing effects of these methods of handling missing data, analyses were conducted with and without missing data substituted. Analyses with missing data substituted are equivalent to intent-to-treat analyses." (p. 248) "Means are calculated with missing data substituted by the last observed value or the interpolation of adjacent values (described in more detail in the method section)". (p. 250) Attrition >15% (21%)
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Robin 1976

Study characteristics

Methods

Parallel randomised trial with two arms

1. Usual care: immediate
2. Wait-list: delayed

Sample calculation: not stated

Cluster randomised: no

Duration of trial (baseline to post): 1 week in immediate groups, but participants in the wait-list group were on the list for 12 weeks (range 6 to 18 weeks)

Duration of participation (trial + follow-up): no follow-up

Setting: outpatient

Robin 1976 *(Continued)*

	Purpose of trial: "This study examines the effect of the waiting list in a prospective controlled trial." (p. 138)
Data	**Number of participants screened:** not stated **Number of participants included:** 234 **Number of participants followed-up at post treatment:** 234 **Number of participants randomly assigned to:** • Usual care: n = 116 • Wait-list: n = 118 **Number of withdrawals:** n = 0 **Diagnosis:** psychiatric patients with different diagnoses: affective disorder, neurosis, personality problem, alcoholism/drug dependence, schizophrenia, organic cerebral disease, mental handicap **Diagnostic manual:** not stated **Means of assessment:** not stated **Comorbidity:** different diagnoses **Ag:**: not stated **IQ:** 3% were mentally handicapped **Sex:** 50.4% female **Ethnicity:** not stated **Country:** UK **Inclusion criteria** 1. Referred to a clinic by a general practitioner for the first time or re-referred at least six months after discharge. **Exclusion criteria** 1. Not stated
Comparisons	**Usual care** **Treatment name:** immediate appointments **Description of intervention:** "(…) so that patient referred by general practitioners in the week before such a clinic could be offered appointments within seven days ('immediate appointments')" (p. 138) **Individual or group treatment:** not stated– various treatments **Exposure/intensity to treatment:** not stated – various treatments **Duration of treatment:** 1 week **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** not stated **Wait-list** **Comparison name:** Delayed appointment

Robin 1976 *(Continued)*

Description of intervention: "Those referred in the following weeks ('delayed appointments') would join a waiting list of, on average, 12 weeks' duration (range 6-18 weeks)." (p. 138)

Exposure/intensity to treatment: no treatment during the wait-list period

Duration treatment: 12 weeks – range 6 to 18 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect

- **Hierarchy:** usable data
- **Outcome chosen:** number attended out-patients clinics

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. Seventy-six per cent of those offered immediate appointments attended, as against 53 per cent of those offered delayed appointment – almost half as many again
2. To determine possible reasons for non-attendance a survey of services supplied three months on either side of the date of the missed appointment was undertaken (…) Of 55 (47 per cent) delayed non-attenders, five (4 per cent) were admitted before their appointment (one after a domiciliary visit). A sixth patient was seen at home, and 12 others (11 per cent in all) were given earlier outpatient appointments because of urgency reported by the referring doctor. Eleven (9 per cent) of the delayed group were seen as outpatients later than their original appointment, and in all 26 (22 per cent) delayed referrals received no local service
3. The main finding of the present study is that a delayed appointment results in significantly fewer patients attending, and argues that a reduction in attendances is a rational objective in present circumstances if, as seems probable, no harmful effects ensue.

Key limitations from study authors

1. Not stated

Other notes from review authors

1. Differences between in treatment duration between usual care and wait-list

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	No	Quote: "Without specific announcement, alternate out-patient clinics were kept vacant so that patients referred by general practitioners in the week before such a clinic could be offered appointment within seven days ('immediate appointments'). Those referred in the following weeks ('delayed appointments') would join a waiting list of, on average, 12 weeks' duration (range 6-18 weeks). Referrals were thus randomized between 'immediate' and 'delayed' appointments, but where urgency was specifically stressed by the family doctor after an appointment date had been offered, and effort was made to provide an earlier appointment for delayed patients." (p. 138)
Allocation concealment	Unclear	No information

Robin 1976 *(Continued)*

Blinding of outcome assessors	Yes	Outcome is 'attended/did not attend'
Blinding of participants and personnel	No	Not possible to blind usual care and wait-list
Incomplete outcome data	Yes	Not relevant
Selective outcome reported	Unclear	No protocol found
Other sources of bias	No	Time bias in outcome measure. Patients in the immediate group was "assessed" for attenending immediately, but patients in the delayed was assessed after 12 weeks of wait-list treatment. Attention bias: patients in TAU received treatment for 1 week, while WL received treatment for appr. 12 weeks (Range = 6-18 weeks)

Roehrich 1993

Study characteristics	
Methods	**Parallel randomised trial with four arms** 1. Psychoogical placebo: alcohol opinions and attitudes (PBO-REM) 2. No treatment: no remediation 3. Active treatment 1: standard neuropsychological remediation (NEURO-REM) 4. Active treatment 2: ecologically relevant remediation (ECO-REM) **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 2 weeks **Duration of participation (trial + follow-up):** 2 weeks (no follow-up) **Setting:** inpatient alcohol treatment program **Purpose of trial:** "The current investigation focused on treatment-relevant remediation (acquisition of the content of a relapse-prevention [RP] program) using task administered by self-guided workbooks." (p. 812) **Closed/open placebo:** closed placebo
Data	**Number of participants screened:** not stated **Number of participants included:** 80 **Number of participants followed-up at post treatment** : 61 **Number of participants randomly assigned to** : • Psychoogical placebo: n = 16 • No treatment: n = 15 • Active treatment 1: n= 15 • Active treatment 2: n = 15 **Number of withdrawals:** n = 19

Roehrich 1993 *(Continued)*

Diagnosis: substance use disorder (alcohol dependence)

Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition (DSM III)

Means of assessment: not stated

Comorbidity: "Forty percent of the subjects admitted to past drug abuse or dependence, and 21% of the sample had received prior treatment for depression, drug dependency, or posttraumatic stress disorder." (p. 814)

Age: psychological placebo: 42.38 mean years (SD = 10.65), no-treatment: 43.13 mean years (SD =11.54)

IQ: Wechsler Adult Intelligence Scale –, psychological placebo: 101.31 (SD = 12.85), no-treatment: 96.07 (SD = 8.61)

Sex: 100% male

Ethnicity: 82% white, 16% black, 2% listed themselves as "other"

Country: USA

Inclusion criteria

1. Alcohol dependence

Exclusion criteria

1. Patients with drug-positive urine samples at intake were excluded from the overall treatment program
2. Patients with visual impairments requiring use of large-print books or cassette tapes were also excluded
3. Admitted patients were further screened for: epilepsy or other seizure history. History of thought disorder. History of bipolar disorder. Severe head injury. History of brain illness, brain surgery, or organic brain syndrome. Impairments of the dominant hand such as broken fingers or severe arthritis

Comparisons

Psychological placebo

Treatment name: Alcohol opinions and attitudes (PBO-REM)

Description of intervention: "Subjects received workbooks according to their treatment group status 10 to 12 days after admission. Instructions for the workbooks appeared inside each booklet and were not presented by the experimenter." (p. 814)

"The PBO-REM workbook also contained elements of repetition and feedback. However, in this case, subjects were asked to respond to statements about alcoholism and then to provide a written rationale for their choices. This workbook relied heavily on automatic verbal skills that have been noted to be relatively unimpaired in alcoholic patients." (p. 815).

States on p. 814 that it is a placebo intervention.

Individual or group treatment: individual – "All subjects were instructed to work independently and not share answers with one another." (p. 814)

Exposure/intensity to treatment: 4 sessions. "Four 1-hr workbook sessions were spread out over the final 2 weeks of inpatient treatment." (p. 814)

Duration of treatment: 2 weeks. "A four-group, pretest-posttest design was overlaid on an operating 28-day, inpatient alcohol treatment program." (p. 813). "Four 1-hr workbook sessions were spread out over the final 2 weeks of inpatient treatment." (p. 814)

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: "All patients had been abstinent for at least 7 days and were medication-free except for vitamins, common analgesics (e.g., acetaminophen), and some use of antihypertensives and anti-inflammatories (for arthritis)." (p. 813)

Roehrich 1993 *(Continued)*

	No-treatment **Comparison name:** No remediation **Description of intervention:** no description of the control group – only pre- and post-testing **Exposure/intensity to treatment:** no treatment **Duration treatment:** 2 weeks **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** "All patients had been abstinent for at least 7 days and were medication-free except for vitamins, common analgesics (e.g., acetaminophen), and some use of antihypertensives and anti-inflammatories (for arthritis)." (p. 813)
Outcomes	**Beneficial effect** • **Hierarchy:** observer-reported, global score • **Outcome chosen:** digit symbol **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. "Results showed that exposure to both types of remediation produced significant cognitive recovery, with skills transferring to posttest neuropsychological measures and RP acquisition. Hence, cognitive remediation may facilitate alcoholism treatment." (p. 812) 2. "Comparison of pretest scores for Digit Symbol, Trails A, and Trails B at pretest with the corresponding posttest scores for each group shows that the NEURO-REM and ECO-REM groups evidenced considerable more improvement for each measure than PBO-REM or control groups (which showed little or no change) (…) There were, however, no differences between the standard neuropsychological and the ecologically valid strategies, suggesting that either procedure could be used with equal effectiveness." (p. 817) **Key limitations from study authors** 1. Not stated **Other notes from review authors** 1. Only numbers of completers stated, not the total number of participants randomly assigned **Conflicts of interest:** none found **Judgement:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Each participant was then randomly assigned to one of four treatment groups; subjects younger than 40 years and those 40 years of age and older were assigned separately to each group to balance for age both within and across groups." (p. 814)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Unclear	No information

Roehrich 1993 *(Continued)*

Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	No	Attrition >15% (24%). No ITT used
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Rosa-Alcatraz 2009

Study characteristics	
Methods	**Parallel randomised trial with four arms** 1. Psychological Placebo 2. Wait-list 3. Active treatment 1: educational group 4. Active treatment 2: IAFS multi-component package **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 3 months **Duration of participation (trial + follow-up):** 3 months + 12 months follow-up **Setting:** outpatient **Purpose of trial:** "The aim of this study is to analyze the specific effects of the Intervention in Adolescents with Social Phobia (IAFS) program together with the nonspecific factors of the interventions used in the treatment of adolescents with social phobia" (p. 44) **Closed/open placebo:** closed placebo
Data	**Number of participants screened:** 2650 **Number of participants included:** 77 **Number of participants followed-up at post-treatment:** 77 **Number of participants randomly assigned to:** • Psychological Placebo: n = 18 • Wait-list: n = 20 • Active treatment 1: n = 19 • Active treatment 2: n = 20 **Number of withdrawals:** n = 0 **Diagnosis:** Social Anxiety disorder (SAD) + Generalized social phobia (GAD) **Diagnostic manual:** Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision (DSM-IV-TR) **Means of assessment:** Anxiety Disorders Interview Schedule for Children, 4th edition (ADIS-IV-C)

Rosa-Alcatraz 2009 *(Continued)*

Comorbidity: Panic Disorder (4/77), Agoraphobia (7/77), Selective mutism (2/77), Generalized Anxiety Disorder (12/77), obsessive compulsive disorder (OCD) (2/77), Obsessive–compulsive personality disorder (10/77), Specific phobia (37/77), Post-traumatic Stress disorder (PTSD) (2/77), Dysthymia (11/77), Substance use disorder: alcohol (11/77), Substance use disorder: other substances (4/77)

Age: 14.87 mean years (SD = 0.80). (range = 14 to 17)

IQ: not stated

Sex: 71.4% female

Ethnicity: not stated

Country: Spain

Inclusion criteria

1. SAD diagnosis
2. Requirement of written parental consent for their children to participate in the research and authorization to make audiovisual recordings for strictly clinical purposes

Exclusion criteria

1. Not stated

Comparisons

Psychological placebo

Treatment name: Placebo group

Description of intervention: “Control Group Placebo (…) in order to provide empirical evidence both regarding the role of information transmission and in relation to the extent to which the effects generated by the treatment are due to the so-called "spontaneous remission", that is, to specific elements of the treatment or to non-specific factors such as, for example, of the patient or therapist, expectations towards treatment or care and support provided by the therapist.” (translated, p. 45)

“The group Placebo received information on adequate nutrition (2 sessions), consumption of psychoactive substances (4 sessions), hygiene (1 session), sports (2 sessions), AIDS prevention (2 sessions), prevention of unwanted pregnancies (1 session). It was controlled that in no case were they taught or indicated how they should or could act in the face of the problems from which they received information. After the presentation of the contents, the adolescents commented as a group (of about 6 subjects) and discussed the problems that were presented in real life on the different topics, reaching agreements that should be presented to the large group (the 18 subjects together to the therapist) before the end of the session. The treatments were applied by therapists with experience in the application of the IAFS and in the transmission of specific and qualified information.” (translated, p. 50)

Individual or group treatment: group

Exposure/intensity to treatment: 12 sessions, 90 minutes weekly

Duration of treatment: 12 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

Wait-list

Comparison name: Wait-list

Description of intervention: “Control Group (…) Waiting List, in order to provide empirical evidence both regarding the role of information transmission and in relation to the extent to which the effects generated by the treatment are due to the so-called "spontaneous remission", that is, to specific elements of the treatment or to non-specific factors such as, for example, of the patient or therapist, expectations towards treatment or care and support provided by the therapist.” (translated, p. 50)

Rosa-Alcatraz 2009 *(Continued)*

Exposure/intensity to treatment: no treatment during waiting

Duration treatment: 3 months

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported, (clinical relevance, psychometric properties)
- **Outcome chosen:** Anxiety Disorders Interview Schedule for Children (ADIS-IV-C) – (Nº situaciones sociales fóbicas)

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. The results show the short- and medium-term effectiveness (12 months) of the IAFS according to specific measures assessing social anxiety and avoidance as well as other related constructs (assertiveness, social skills and adjustment).
2. The placebo group achieved important improvements in some of the mentioned conditions, whereas that of information transmission did not obtained significant changes, with the exception of the self-esteem variable.

Key limitations from study authors

1. One variable that would explain the improvement of the Placebo Group is that the participants of this condition had to interact in groups and expose themselves to present the results of their reflections in both small and large groups, so that this forced the children to interact with each other orally. This variable, group exposure, is the one that best explains the improvements achieved, from our point of view, so that, although we originally labelled this group of placebo, it really incorporates an active element (translated, p. 57)
2. Regarding future research that took into account both the limitations of this work and the questions that can be derived from the results thereof, we believe that it would be important to use different types of placebo groups, eliminating any active ingredient of treatment for phobia. (translated) in order to really analyse the nonspecific effects of interventions in generalized social phobia. On the other hand, it would be relevant to obtain more measures from independent observers that would help to evaluate the effects of the treatment in a more objective way, as well as include the partners (sociometric test) in the group of social agents (together with parents and teachers)
3. In order to continue assessing the social validity of the changes generated. Another interesting aspect to have in (translated)

Other notes from review authors

1. Quotes were translated from Spanish

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "77 adolescents were selected at random and randomly distributed between four experimental conditions: psychological treatment group (IAFS multi-component package), transmission of information or educational group, placebo and waiting list control group" (p. 44)

Rosa-Alcatraz 2009 *(Continued)*

Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote:"The evaluation was carried out by three independent groups of collaborators coordinated by the third author, previously trained for this purpose. The first group performed the preliminary evaluation and follow-up at twelve months. The second, the posttest and the third, the follow- up at six months. The team was only aware of all the data referring to the different measures when the collection of information related to the second follow-up ended." (translated from p. 47)
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	Yes	Attrition <15% (0%, All patients completed posttreatment)
Selective outcome reported	Unclear	SEJ2004-01471/PSIC Protocol: Olivares 2005: Programa IAFS. Protocolo para el tratamiento de la fobia social en adolescentes. Not able to locate trial registry or published protocol
Other sources of bias	Yes	No other sources found

Rosen 1976

Study characteristics	
Methods	**Parallel randomised trial with five arms** 1. Psychological placebo 2. Wait-list 3. Active treatment 1: desensitisation therapist: 4. Active treatment 2: desensitisation calls 5. Active treatment 3: desensitisation manual **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 8 weeks **Duration of participation (trial + follow-up):** 8 weeks + 2 months follow-up **Setting:** outpatient **Purpose of trial:** "The present study evaluates the clinical efficacy of self-administered desensitization in the context of a controlled outcome study" (p. 209) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** not stated **Number of participants included:** 55 **Number of participants followed-up at post-treatment:** 43 **Number of participants randomly assigned to** • Psychological placebo: n = 11 • Wait-list n= 11

Rosen 1976 *(Continued)*

- Active treatment 1: n = 11
- Active treatment 2: n = 11
- Active treatment 3: n = 11

Number of withdrawals: n = 12

- Psychological placebo: n = 1
- Wait-list n = 4
- Active treatment 1: n = 2
- Active treatment 2: n= 2
- Active treatment 3: n= 3

Diagnosis: specific anxiety (snakes)

Diagnostic manual: not stated

Means of assessment: self-referred, scored high on Snake Attitude Questionnaire (SNAQ) and Fear Survey Schedule (FSS)

Comorbidity: not stated

Age: 33.5 mean years

IQ: not stated

Sex: 92.7% female

Ethnicity: not stated

Country: USA

Inclusion criteria

Subjects were self-referred snake phobics who:

1. responded to a local newspaper announcement that offered treatment for individuals "truly terrified of snakes";
2. refused to touch a snake during a behavior approach test;
3. scored 19 or above on the Snake Attitude Questionnaire (SNAQ);
4. rated "very much fear" or "terror" on the snake item of Fear Survey Schedule (FSS) ;
5. specified a target situation such as gardening or camping that was significantly affected by fear of snakes;
6. and were not currently receiving treatment for their phobia.

Exclusion criteria

1. Not stated

Comparisons

Psychological placebo

Treatment name: Placebo control

Description of intervention: “To control for initial therapeutic expectancies and other nonspecific treatment factors, a totally self-administered bibliotherapy placebo called systematic relearning was included in the present study. The basic rationale of the treatment program was that people could substantially reduce their fears by replacing inaccurate perceptions with more accurate information about the feared object. To accomplish this goal, each individual studies a manual that organizes factual information about snakes into 10 chapters. Each chapter contains questions at the end to help subjects assess their mastery of the materials. As subjects work on the program they construct an "information hierarchy" from the relevant information in each chapter. Because of recent findings that question the adequacy of many placebo procedures (...), a number of steps were taken to assure adequate experimental control over subjects' expectancies. In addition, possible therapist expectancy effects were

Rosen 1976 *(Continued)*

avoided by sending all self-instructional materials through the mail. In effect, systematic relearning and self-administered desensitization as previously described were administered under double-blind conditions. " (p. 210-11)

Individual or group treatment: individual

Exposure/intensity to treatment: generally twice weekly

Duration of treatment: up to 8 weeks

Concomitant psychotherapy: were not currently receiving treatment for their phobia

Concomitant pharmacotherapy: were not currently receiving treatment for their phobia

Wait-list

Comparison name: No-treatment (in reality wait-list)

Description of intervention: "No-treatment control. Subjects in this group were informed that the large number of clients in the project necessitated a delay in treatment for some individuals. After posttesting, untreated controls were offered treatment." (p. 211)

Exposure/intensity to treatment: no treatment

Duration treatment: up to 8 weeks

Concomitant psychotherapy: were not currently receiving treatment for their phobia

Concomitant pharmacotherapy: were not currently receiving treatment for their phobia

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported, clinical relevance, global score
- **Outcome chosen:** Behavior Avoidance Test (BAT)

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. It was concluded that within the context of moderate treatment effects the present study provides support for the clinical efficacy of totally self-administered desensitization
2. Implications of these findings for the clinical management of specific fears arc discussed

Key limitations from study authors

1. High attrition

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote:"After pretreatment assessments had been completed, the subjects were matched on behavior approach scores and assigned by block randomization to one of four treatment groups or a no- treatment control." (p. 210)

Rosen 1976 *(Continued)*

Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote:"Pretreatment to posttreatment and follow-up assessments of subjects' attitudes and reactions toward snakes were conducted by assistants blind to subjects' group assignment" (p. 209)
Blinding of participants and personnel	No	Not possible to blind placebo and "no-treatment"
Incomplete outcome data	No	Quote:"Two subjects in each of the therapist-aided treatment groups dropped out of their programs during the first week of therapy. Three self-administered desensitization subjects, 1 placebo control, and 4 untreated controls could not be reached at time of posttesting primarily because of address changes. Accordingly, a final N of 43 was achieved, and group sizes were not equal (…)" (p. 211) Attrition >15% (38%). No ITT. Only reports data on completers. Very low sample size
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Roth 1964

Study characteristics	
Methods	**Parallel randomised trial with six arms** 1. Pharmacological placebo 2. Wait-list 3. Active treatment 1: librium 4. Pharmacological placebo + psychotherapy 5. Wait-list+ psychotherapy 6. Active treatment 2: librium + psychotherapy **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 4 weeks **Duration of participation (trial + follow-up):** 4 weeks + 6 months follow-up **Setting:** outpatient **Purpose of trial:** "The broad aim of the study was to determine some of the early effects of an ataractic agent, chlordiazepoxide, on anxiety and tension.." (p. 257) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** 311 **Number of participants included:** 181 **Number of participants followed-up at post-treatment:** 150 **Number of participants randomly assigned to** • Pharmacological placebo: n = 25 • Wait-list: n = 25 • Active treatment 1: n = 25 • Pharmacological placebo + psychotherapy: n = 25 • Wait-list+ psychotherapy: n = 25

Roth 1964 *(Continued)*

- Active treatment 2 + psychotherapy: n = 25

Number of withdrawals: n = 31

- Pharmacological placebo: n = 3
- Wait-list: n =13
- Active treatment 1: n = 5
- Pharmacological placebo + psychotherapy: n = 5
- Wait-list+ psychotherapy: n = 2
- Active treatment 2 + psychotherapy: n = 3

Diagnosis: Psychiatric outpatient, different diagnoses. “The therapists’ initial diagnosis classified 42 per cent of the total sample as Neurotics, 26 per cent as Personality Disorders, 15 per cent as Psychophysiologic Disorders, and 18 per cent as Psychotics. While there were some differences between treatment groups, these proved to be non-significant. The mean global severity of illness rating for the sample was “Moderately Ill”.” p. 262.
Diagnostic manual: not stated
Means of assessment: intake interviews (clinical structured interview)
Comorbidity: different disorders
Age: 37.8 mean years
IQ: not stated
Sex: 100% male
Ethnicity: not stated
Country: USA

Inclusion criteria

1. Veterans newly accepted for individual psychotherapy in six Veterans Administration Mental Hygiene Clinics

Exclusion criteria

1. Those who had been hospitalised for a psychiatric illness during the previous three months
2. Those who had been in group or individual psychotherapy anywhere during the previous three months
3. Those with a history of central nervous system disorders or seizures; 4. those with symptomatic addiction to alcohol
4. Those who could not discontinue medication while in the study
5. 5Those 55 years of age or over

Comparisons

Pharmacological placebo
Treatment name: Placebo
Description of intervention: ”The study medication which consisted of chlordiazepoxide (Librium) in 10 mg capsules or placebo were prescribed in the following manner. During the first week all patients scheduled for medication received a total daily dosage of four capsules. In the second week the physician prescribed four capsules if adjustment was not required. When required the dosage could be lowered to three or increased to five or six capsules per day. At the beginning of the third week, dosage could be lowered to two capsules or increased to eight capsules per day.
This final dosage remained fixed for a given patient for the last two weeks.” (p. 260)
Individual or group treatment: individual.
Exposure/intensity to treatment: between 2 and 8 capsules per day
Duration of treatment: 4 weeks
Concomitant psychotherapy: half of the patients started psychotherapy concomitantly with the medication treatment
Concomitant pharmacotherapy: not allowed

Wait-list
Comparison name: Wait group
Description of intervention: “All patients completed an inventory on the same day or within a week of initiation of treatment, but always before the first treatment. They were administered a IO-mm adjective rating scale and completed a global improvement rating just before the second, third and fourth

Roth 1964 *(Continued)*

treatment visits to the therapist or prescribing physician. Just before the fifth treatment, each patient was re-examined on a modified inventory. Since the Wait Group received no treatment for four weeks, initial testing was completed early. The adjective rating scale was given only once, a week after the initial inventory was completed. Wait Group cases then waited three weeks for a final re-testing, given prior to the initial psychotherapeutic
interview. The treatment groups did not differ on the number of days between their initial and final tests" (p. 260)
Exposure/intensity to treatment: no pharmacological treatment. Some patients received psychotherapy.
Duration treatment: 4 weeks
Concomitant psychotherapy: half of the patients started psychotherapy.
Concomitant pharmacotherapy: not allowed

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported, clinical relevance
- **Outcome chosen:** global estimate of the severity of the patient's illness and a global measure of over-all improvement

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

A four-week double-blind study was made of the effects of chlordiazepoxide on a group of 150 male outpatients newly accepted for individual psychotherapy. Three major hypotheses were tested. The findings with respect to patient criteria were:

1. patients receiving the drug reported themselves no better than placebo patients with but one exception;
2. patients receiving either capsule (chlordiazepoxide or placebo) reported a greater reduction in tension, anxiety and depression, greater over-all improvement, and more social changes than patients not receiving capsules;
3. treatment groups receiving psychotherapy combined with drug or placebo reported themselves no differently from patients not receiving psychotherapy (except for the Wait group)
4. The findings for patients receiving psychotherapy as reported by their therapists were: patients receiving the drug were significantly less severely ill, in better rapport with others, and better able to express affection;
5. groups receiving either the drug or a placebo as compared to the group receiving only psychotherapy showed significantly greater reduction in anxiety, self blame, physical complaints, and greater overall improvement
6. From the standpoint of patient self reports, only after the first week did patients receiving the drug show greater changes than those receiving placebos. They reported less tension anxiety, greater vigor and more over-all improvement. These effects wash out by the end of four weeks. After four weeks placebo patients report as much change as chlordiazepoxide patients

Key limitations from study authors

1. First, it should be noted that the experiment was not intended to evaluate psychotherapy per se, only its short-term influence on anxiety and its interactive effect with the drug.
2. But it is important to note that the Wait Group did not improve at all, while the improvement reported by Psychotherapy Only patients was significant and in a pattern very similar to the group also receiving the drug.
3. The Psychotherapy Only patients did not exhibit this pattern. In what way if any, can the known pharmacological properties of chlordiazepoxide account for the present findings (...) Unlike chlorpromazine it does not produce autonomic blocking. It also lacks hypnotic effects at high doses. The drug has appetite-stimulating effects on rats and dogs.

Roth 1964 *(Continued)*

4. The reportedEffects of muscle-relaxation, taming and appetite stimulation are consistent with the findings reported here of early reduction in tension-anxiety, the heightened sense of well-being (Vigour score), and improved rapport with others
5. The results reported here are also consistent, for the most part, with clinical and controlled reports.@-10) The major novel finding is the extent to which anxiety can be reduced simply by giving a patient a capsule

Other notes from the review authors

1. Data not usable. The standard deviation (SD) was not reported on either global estimate of the severity of the patient's illness and a global measure of over-all improvement

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Patients were randomly assigned to six treatment groups receiving the combinations of psychotherapy and medication treatment (p. 283)
Allocation concealment	Yes	Quote:"Each clinic pharmacist dispensed the medication, which was delivered in individual bottles containing 250 capsules identical in size, color, appearance, and taste. Each bottle was labeled with the unique code number assigned to the patient. The identification label of the study medication sealed in small envelopes was placed in the hands of the pharmacist. Codes were not broken until after completion of the study." (p. 260)
Blinding of outcome assessors	Yes	The therapist were outcome assessors- and were blinded. Quote:"Although therapists were eventually permitted to ascertain whether their patient had initially been on active drug or placebo, this was not allowed until many months after the medication was administered." (p. 284)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	No	Attrition >15% (17.1%) Only reports data on completers. Eliminates patients to ensure equal groups
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Rupert 1978

Study characteristics

Methods

Parallel randomised trial with seven arms

1. Physical placebo 1: increase heart rate: placebo feedback
2. Physical placebo 2: decrease heart rate: placebo feedback
3. No-treatment
4. Active treatment 1: Increase heart rate: true biofeedback

Rupert 1978 *(Continued)*

5. Active treatment 2: Decrease heart rate: true biofeedback
6. Other control 1: Increase heart rate: no biofeedback
7. Other control 1: Decrease heart rate: no biofeedback

Sample calculation: not stated
Cluster randomised: no
Duration of trial (baseline to post): 4 to 7 days
Duration of participation (trial + follow-up): 4 to 7 days. No follow-up
Setting: inpatient
Purpose of trial: "The present experiment examined the effects of multiple sessions of heart rate biofeedback training on the heart rate control and anxiety levels of anxious psychiatric patients." (p. 583)
Open/Closed placebo: closed

Data

Number of participants screened: not stated
Number of participants included: 56
Number of participants followed-up at post-treatment: not stated
Number of participants randomly assigned to:

- Physical placebo 1: n = 8
- Physical placebo 2: n = 8
- No-treatment: n = 8
- Active treatment 1: n = 8
- Active treatment 2: n = 8
- Other control 1: n = 8
- Other control 1: n = 8

Number of withdrawals: not stated
Diagnosis: anxious psychiatric patients
Diagnostic manual: not stated
Means of assessment: not stated
Comorbidity: not stated
Age: (range 19 to 55)
IQ: not stated
Sex: 100% male
Ethnicity: not stated
Country: USA

Inclusion criteria

1. Judged by their physicians to be suffering from a high degree of anxiety
2. Either on no medication or on a relatively low and stable dosage of medication

Exclusion criteria

1. Free from serious mental ilness
2. Free from coronary disease or other health problems that might make the procedures dangerous

Comparisons

Physical placebo 1 and Physical placebo 2
Treatment name: Placebo biofeedback
Description of intervention: "The other meter was constructed to give placebo (false positive) feedback and was a voltmeter wired in a circuit that generated a positive or negative signal. This signal was used to give the needle on the meter on the appropriate direction to indicate a steadily increasing or decreasing heart rate. The meters were either covered or uncovered depending on whether the subjects were in a biofeedback (true or placebo) or a no-biofeedback condition." (p. 584)
"Subjects in the true biofeedback and placebo biofeedback conditions were also told that the meter in front of them would give them information about changes in their heart rates and that they were to use this information to help them control their heart rates. It was explained that movements of the needle on the meter to the right and left reflected increases and decreases in their heart rates, respectively." (p. 585)

Rupert 1978 *(Continued)*

Individual or group treatment: individual
Exposure/intensity to treatment: "All subjects participated in 4 expediential sessions which were conducted within a period of 4 to 7 days. Each session included a total of 25 min of heart rate training (or recording time) plus time required for directions, acquisition of initial heart rate levels, and rest periods." (p. 584)
Duration of treatment: 4 to 7 days
Concomitant psychotherapy: not stated
Concomitant pharmacotherapy: no to little medication

No-treatment
Comparison name: No-treatment
Description of intervention: assessed before and after treatment
Exposure/intensity to treatment: no treatment
Duration treatment: 4 to 7 days
Concomitant psychotherapy: not stated
Concomitant pharmacotherapy: no to little medication

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported, clinical relevance
- **Outcome chosen:** heart rate

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. Results indicated that: a) neither instructions alone nor the combinations of instructions and true or placebo biofeedback were more effective than simply sitting quietly (adaptation) for decreasing heart rate,
2. instructions plus true biofeedback was more effective than instructions alone or instructions plus placebo biofeedback for increasing heart rate,
3. multiple sessions of training did not enhance the level of control achieved early in the first session,
4. the control achieved with biofeedback did not transfer to a subsequent no-biofeedback situation,
5. and biofeedback training did not influence subjects'; subjective anxiety levels

Key limitations from study authors

1. Not stated

Other notes from review authors

1. No usable data - nor possible to generate data

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Subjects were randomly assigned in equal numbers to the six experimental conditions and the no-treatment control condition." (p. 584)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Unclear	No information

Rupert 1978 *(Continued)*

Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Unclear	No information
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Shalev 2012

Study characteristics	
Methods	**Parallel randomised trial with five arms** 1. Pharmacological placebo 2. Wait-list 3. Active treatment 1: PE (prolonged exposure) 4. Active treatment 2: CT (cognitive therapy) 5. Active treatment 3: SSRI (escitalopram) **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 12 weeks of treatment and post treatment data at 5 months. **Duration of participation (trial + follow-up):** 12 weeks of treatment + post treatment data at 5 month + follow-up at 9 months. **Setting:** outpatient **Purpose of trial:** "To compare early and delayed exposure based, cognitive, and pharmacological interventions for preventing PTSD." (p. 166) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** 5286 **Number of participants included:** 242 **Number of participants followed-up at post-treatment:** 207 **Number of participants randomly assigned to:** • Pharmacological placebo: n= 23 • Wait-list: n=93 • Active treatment 1: : n = 63 • Active treatment 2: n = 40 • Active treatment 3: n = 23 **Number of withdrawals:** n = 35 • Pharmacological placebo: n = 5 • Wait-list: n =14 • Active treatment 1: : n =7 • Active treatment 2: n =7 • Active treatment 3: n = 2 **Diagnosis:** post-traumatic stress-disorder (PTSD) or acute stress disorder (ASD) **Diagnostic manual:** Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV). **Means of assessment:** The PTSD Symptom Scale-Interviewer Version (PSS-I) and the ASD Scale (ASDS)

Shalev 2012 *(Continued)*

Comorbidity: not stated
Age: Pharmacological placebo: 36.26 mean years (SD = 12.39), wait-list: 37.28 mean years (SD = 11.91)
IQ: not stated
Sex: not stated
Ethnicity: not stated
Country: Israel

Inclusion criteria

1. If participant resides within a 1-hour drive from Jerusalem
2. Meets diagnostic criteria

Exclusion criteria

1. Sustained an injury that required more than 7 days of hospital stay
2. Were unconscious on admission to emergency services
3. Had medical or surgical conditions that interfered with their ability to participate or provide informed consent
4. Not fluent enough in Hebrew, Arabic, or English to answer questions and/or interact during clinical assessments
5. Current or past psychosis or bipolar disorder
6. Current substance abuse problem
7. Conditions requiring urgent attention (e.g, suicidal ideations or acute grief)
8. Chronic PTSD
9. Started treatment elsewhere

Comparisons

Pharmacological placebo

Treatment name: Placebo
Description of intervention: "Concealed tablets of either 10 mg escitalopram or placebo were prepared and coded by Lundbeck Pharmaceuticals (Copenhagen, Denmark) and were supplied to clinicians by a research associate. An initial dose of 1 tablet daily was increased to 2 tablets after 2 weeks of treatment. Trained psychiatrists provided 4 weekly sessions (weeks 1-4) followed by 4 biweekly sessions (weeks 6-12)." (p. 168)
Individual or group treatment: individual
Exposure/intensity to treatment: 1 tablet daily, increased to 2 tablets daily after 2 weeks of treatment
Duration of treatment: 12 weeks
Concomitant psychotherapy: not stated
Concomitant pharmacotherapy: not stated

Wait-list
Comparison name: Waiting list
Description of intervention: "The WL participants who met PTSD diagnostic criteria at 5 months received PE at that time (hereafter referred to as delayed PE." (p. 167). "A telephone interview briefly contacted participant on the WL every 2 weeks to inquire about emerging needs or possible emergencies. These calls did not contain elements of PE or CT." (p. 169)
Exposure/intensity to treatment: no treatment
Duration treatment: 12 weeks
Concomitant psychotherapy: not stated
Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported, clinical relevance, continuous
- **Outcome chosen:** CAPS Total Score T2,

Adverse events

- No data reported on adverse events

Shalev 2012 *(Continued)*

Notes

Key conclusion from study authors

1. "The results of our study show that there are significant and similar preventive effects of PE and CT." (p. 174)

2. "The escitalopram subgroup did not differ from the placebo subgroup or the WL group at 5 months; however, the escitalopram subgroup fared worse than all the other groups at 9 months." (p. 174)

3. "Delaying PE did not affect the 9-month outcome (…) Our finding suggests that delaying the intervention does not increase the risk of chronic PTSD. Delaying treatment somewhat reduced the number of treatment candidates: about a third of those with initial PTSD recovered by 5 months (…) Thus, a delayed intervention is an acceptable option when early clinical interventions cannot be provided (e.g. during wars, disasters, or continuous hostilities" (p. 174)

Key limitations from study authors

1. Sample of civilian, survivors of single, short traumatic events
2. Sample includes referrals from emergency services and thus a number of participants who had a physical injury
3. Our study group sizes did not allow us to further explore the factors underlying the heterogeneity of treatment responses or the effect of treatment completion

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	No	Quote:"The equipoise-stratified randomization is a method for randomly allocating participant to interventions in treatment studies that include more than two arms (…) In our study, participants who agreed to start treatment (n=269) were informed about the 4 treatment options (PE, CT, treatment with SSRI vs placebo, and WL and subsequent delayed PE) could decline up to 2 treatment options (including the WL for delayed PE), and were randomly assigned to the remaining treatment option." (P. 168)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote:"The clinical assessments were made by clinical psychology interns (…) They remained blind to treatment attendance and adherence." (p. 167)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Quote:"To account for missing observations and the groups' heterogeneities, we used a linear mixed model with covariance for significant initial group differences and for the time lag between the traumatic event of each participant and subsequent assessment." (p. 171) Attrition <15% (PE: 11%, CT:17,5%, SSRI: 9%, Placebo: 22%. Waitlist, 15%
Selective outcome reported	Yes	NCT00146900 No apparent differences in reporting between trial registry and full report

Shalev 2012 *(Continued)*

Other sources of bias	Yes	No other sources found

Shealy 1979

Study characteristics

Methods

Parallel randomised trial with five arms

1. Psychological placebo
2. Wait-list
3. Active treatment 1: relaxation without muscle-tension
4. Active treatment 2: stimulus control plus relaxation without muscle tension
5. Active treatment 3: self-monitoring

Sample calculation: NS

Cluster randomised: (yes/no): no

Duration of trial (baseline to post): 6 weeks of treatment

Duration of participation (trial + follow-up): 6 weeks treatment + 6 months follow-up

Setting: outpatient (college undergraduates)

Purpose of trial: "There has been a relatively recent effort to develop behavioral treatments for sleep-onset insomnia. The present research was designed as an extension of this previous work in order to examine four current methodological and theoretical issues" (p. 541)

Closed/open: closed

Data

Number of participants screened: not stated

Number of participants included: 70

Number of participants followed-up at post-treatment: not stated

Number of participants randomly assigned to:

- Psychological placebo: n = 14
- Waiting list: n = 14
- Active treatment 1: n = 14
- Active treatment 2: n = 14
- Active treatment 3: n = 14

Number of withdrawals: not stated

Diagnosis: sleep-wake disorder (insomnia)

Diagnostic manual: not stated, but probably due to information provided: "Insomnia was considered mild if there was a sleep onset latency greater than 30 min present at least three nights per week and moderate if there was a sleep onset latency greater than 45 min present at least four nights per week. In both instances. a subjective indication of difficulty in falling asleep was also required, Three levels of insomnia duration were also examined: 3-11 months, 11 years, and greater than 4 years. However, only mild insomniacs were eventually classified on the duration factor." (p. 542)

Means of assessment: not stated

Comorbidity: not stated

Age: 19.4 mean years (range =17 to 30)

Shealy 1979 *(Continued)*

IQ: not stated, but all college students

Sex: 100% female

Ethnicity: not stated

Countr:: USA

Inclusion criteria

1. Insomnia

Exclusion criteria

1. Subjects were excluded from this study if they were currently seeking other sources of psychological help
2. currently using hypnotic drugs to control insomnia
3. or experienced in relaxation training

Comparisons

Psychological placebo

Treatment name: Placebo

Description of intervention: "Insomniacs in this group also lay on their backs with heads on pillows in a dimly lighted room. Topics related to sleep such as sleep disturbances and dreams were discussed. There was no discussion on how to resolve sleep disturbances. It was assumed that this treatment would have little therapeutic value. Insomniacs in this group completed the DSQ every morning throughout the study." (p. 543)

Individual or group treatment: not stated, but seems like a group

Exposure/intensity to treatment: self-monitoring for 2 weeks assessing own sleeping patterns + two one-half hours conducted for 3 weeks + 1 week of positive demand

Duration of treatment: 6 weeks

Concomitant psychotherapy: not allowed

Concomitant pharmacotherapy: no hypnotic drugs

Wait-list

Comparison name: Waiting list

Description of intervention: "Subjects in these groups were told that all treatment groups were filled and that they would receive treatment as soon as possible. (…) The insomniacs in the waiting list group were required to fill out only two questionnaires- one at pretherapy and one at the end of the positive demand period." (p. 543)

Exposure/intensity to treatment: waiting for treatment

Duration treatment: 6 weeks

Concomitant psychotherapy: not allowed

Concomitant pharmacotherapy: no hypnotic drugs

Outcomes

Beneficial effect

- **Hierarchy:** usable data, patient-reported
- **Outcome chosen:** a Daily Sleep and Relaxation Practice Questionnaire (DSRQ) – positive demand

Adverse events

Shealy 1979 *(Continued)*

- No data reported on adverse events

Notes

Key conclusion from study authors

1. During this counterdemand period. the two relaxation groups showed significantly greater decreases in sleep onset latency than the control conditions
2. It appeared that duration affected treated outcome
3. The effectiveness of treatment packages and self-monitoring in alleviating insomnia is briefly discussed

Key limitations from study authors

1. Finally. approximately 84%; of the insomniacs in the present study reported obsessive thoughts while lying in bed at night before falling asleep
2. Finally, the duration of insomnia may have an effect on treatment outcome. The duration of 3-l1 months may really be a situational, insomnia rather than some 'trait' characteristic and thus more amenable to intervention

Other notes from review authors

1. Positive demand are post-treatment
2. We pooled the means from mild and moderate due to no information of patients in each group. A mean score was generated from the two groups.
3. SD's was generated in from Ascher 1979 and Steinmark 1974 (studies with same outcomes, population, scale)

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote:"First, insomniacs agreeing to participate were randomly assigned to the five groups on the basis of their answers to the pretherapy questionnaire. Group assignment occurred before baseline. because self-monitoring itself was an independent variable. Each group consisted of nine mild (three of each duration) and five moderate insomniacs." (p. 542)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	Unclear	Attrition unclear
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Sibilio 1957

Study characteristics

Methods	**Parallel randomised cross-over trial with three arms** 1. Pharmacological placebo 2. No-treatment 3. Active treatment: drug **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 28 days **Duration of participation (trial + follow-up):** 28 days + 3 months of other cross-over phases **Setting:** inpatient **Purpose of trial** : "This paper gives the experimental design for the study and includes a report of the effects of promazine on the behavioral adjustment of patients. Future papers will report the effects of promazine on attention, as measured by the reaction-time test, and on daily ward behavior, blood pressure, blood counts, and bone marrow." (p. 419) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** not stated **Number of participants included:** 93 **Number of participants followed-up at post-treatment:** not stated **Number of participants randomly assigned to:** • Pharmacological placebo: n = 31 • No-treatment: n = 31 • Active treatment: Drug: n = 31 **Number of withdrawals:** not stated **Diagnosis:** chronic schizophrenia **Diagnostic manual:** not stated **Means of assessment:** not stated **Comorbidity:** not stated **Age:** (range = 32 to 61 years) **IQ:** all were considered to be of at least average intelligence (clinically determined) **Sex:** 100% female **Ethnicity:** not stated **Country:** USA **Inclusion criteria** 1. The criterion for chronicity was a continuous period of hospitalisation of at least five years **Exclusion criteria** 1. No patient in whom an organic pathology was primary was included in the study 2. Finally, no patient received any additional or adjunctive therapy during the course of the study
Comparisons	**Pharmacological placebo** **Treatment name:** Placebo **Description of intervention:** "The pharmacologically inactive placebo was in every respect similar to the drug in appearance." (p. 421) **Individual or group treatment:** individual **Exposure/intensity to treatment:** 4 times daily **Duration of treatment:** 28 days **Concomitant psychotherapy:** not allowed **Concomitant pharmacotherapy:** not stated **No-treatment** **Comparison name:** no treatment **Description of intervention:** not stated

Sibilio 1957 *(Continued)*

	Exposure/intensity to treatment: no pill **Duration of treatment:** 28 days **Concomitant psychotherapy:** not allowed **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy:** observer-reported • **Outcome chosen:** Gardner Behavior Chart **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. Changes in behavioural adjustment of female chronic schizophrenic patients produced by administration of promazine failed to attain statistical significance 2. In general, it can be concluded that there were no behavioural changes occurring in any group of patients which could be associated with administration of promazine or a placebo or with no treatment 3. No difference between regular and irregular patterns of administering medication with regard to amount and frequency of dosage was obtained 4. It is felt that, while no changes in behavior adjustment were obtained from using promazine in what is considered an adequate test of its effectiveness, further research should be directed toward the development of procedures and techniques that might identify specific clinical characteristics which respond positively or are refractory to promazine therapy **Key limitations from study authors** 1. Not stated **Other notes from review authors** 1. No usable data **Conflicts of interest:** none found **Judgement:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote:"To arrive at the research population for the study, each patient was randomly assigned to one of three groups in a manner which assured equivalence of the groups with respect to scores obtained on a behavioral adjustment rating scale." (p. 419)
Allocation concealment	Yes	Quote:"The double-blind technique was employed throughout the study. Patients received their medication (drug or placebo) in individual, sealed envelopes. Neither the patient nor the attendant dispensing the envelopes was cognizant of the kind and amount of medication contained in each envelope." (p. 420-1)
Blinding of outcome assessors	Yes	Quote:"Those attendants who rated the patients' behavioral adjustment did not dispense medication and were unaware of the experimental group to which a patient was assigned." (p. 421)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment

Sibilio 1957 *(Continued)*

Incomplete outcome data	Unclear	Attrition unclear
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Sommerness 1955

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Pharmacological placebo 2. No-treatment 3. Active treatment: reserpine **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 12 weeks **Duration of participation (trial + follow-up):** 12 weeks treatment + 4 weeks follow-up **Setting:** inpatient (hospital) **Purpose of trial:** "The effect of suggestion must be measured as well as the effect of an intercession of a break in ward routine when one is attempting to evaluate a drug. For these reasons, it was felt that experiments with the drug, placebo, and control groups must be carried out simultaneously in order to obtain reliable scientific results." (p. 316) **Closed/open placebo:** closed placebo
Data	**Number of participants screened:** 2000-bed hospital **Number of participants included:** 90 **Number of participants followed-up at post-treatment:** not stated **Number of participants randomly assigned to:** • Pharmacological placebo: n = 30 • No-treatment: n = 30 • Active treatment: n = 30 **Number of withdrawals:** not stated **Diagnosis:** chronic mental illness **Diagnostic manual:** not stated **Means of assessment:** "Patients were rated once every two weeks during the 20-week period by two psychiatric aides independently, using the L-M Fergus Falls Behavior Rating Scale" (p. 317) **Comorbidity:** not stated **Age:** not stated **IQ:** not stated **Sex:** 100% male **Ethnicity:** not stated **Country:** USA **Inclusion criteria** 1. The 90 most chronically disturbed male patients in the hospital were placed on one ward **Exclusion criteria**

Sommerness 1955 *(Continued)*

1. Not stated

Comparisons	**Pharmacological placebo** **Treatment name:** Placebo **Description of intervention:** "The second group was given an identical-appearing placebo under identical conditions." (p. 316) **Individual or group treatment:** individual **Exposure/intensity to treatment:** not stated **Duration of treatment:** 12 weeks **Concomitant psychotherapy:** "The usual hospital routine was continued on the ward. Other therapies were neither increased nor decreased during the period of the experiment" (p. 317) **Concomitant pharmacotherapy:** "The usual hospital routine was continued on the ward. Other therapies were neither increased nor decreased during the period of the experiment" (p. 317) **No-treatment** **Comparison name:** No-treated control group **Description of intervention:** "The third group received neither reserpine nor placebo but was otherwise under identical conditions." (p. 316) **Exposure/intensity to treatment:** no treatment **Duration of treatment:** 12 weeks **Concomitant psychotherapy:** "The usual hospital routine was continued on the ward. Other therapies were neither increased nor decreased during the period of the experiment" (p. 317) **Concomitant pharmacotherapy:** "The usual hospital routine was continued on the ward. Other therapies were neither increased nor decreased during the period of the experiment" (p. 317)
Outcomes	**Beneficial effect** • **Hierarchy:** observer-reported • **Outcome chosen:** Fergus Falls Behavior Rating Scale **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. Reserpine in oral dose of 1 mg. twice daily did not effect a behavioral improvement (as measured by the L-M Fergus Falls Behavior Rating Scale) in chronic disturbed male patients 2. Reserpine effected a lowering of blood pressure 3. Reserpine effected a slight weight gain. The greater attention to patients inherent in taking blood pressures and weights and increasing the interest of ward personnel resulted in behavioural improvement in all three groups under study **Key limitations from study authors** 1. The fact that the drug group did not show any more improvement behaviorally than the other two groups may be due to one of two factors: reserpine may not have a positive effect on the behaviour of long-term disturbed patients ; 2 mg. orally per day may not be an effective dose for chronically disturbed patients 2. We have reached the conclusion that having a doctor take blood pressure on patients every two weeks, having the patients weighed regularly, and arousing the interests of the ward personnel seem to account for the favorable behavioral results uniformly obtained in all three groups **Other notes from review authors** 1. No usable data. Not possible to generate data due to an unspecified population **Conflicts of interest:** none found **Judgement:** yes

Sommerness 1955 *(Continued)*

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote:"They were divided into three groups by a random numbers table. This randomization resulted in essential equality for diagnosis, behavior, weight, and blood pressure." (p. 316)
Allocation concealment	Yes	Quote:"The hospital pharmacist alone knew which group received reserpine or placebo. This information was not available to the other experimenters until a complete analysis of the results of the experiment had been made." (p. 316)
Blinding of outcome assessors	Yes	See above
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Unclear	Attrition unclear
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Steinmark 1974

Study characteristics	
Methods	**Parallel randomised trial with four arms** 1. Psychological placebo: quasi-desensitisation placebo 2. Wait-list: Waiting-list no treatment 3. Active treatment 1: progressive relaxation 4. Active treatment 2: single-item desensitisation **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 4 weeks **Duration of participation (trial + follow-up):** 4 weeks + 5 months follow-up **Setting:** outpatient (psychology students) **Purpose of trial:** "The present study was designed to critically test the demand and placebo interpretations of outcome improvement among subjects trained in relaxation and to evaluate whether any additional benefit would be achieved by the use of single-item desensitization." (p. 157-8) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** 519 **Number of participants included:** 52 **Number of participants followed-up at post-treatment:** 48

Steinmark 1974 *(Continued)*

Number of participants randomly assigned to:

- Psychological placebo: n = 13
- Wait-list: n = 13
- Active treatment 1: n = 13
- Active treatment 2: n = 13

Number of withdrawals : n = 4

- Psychological placebo: n = 1
- Wait-list: n = 1
- Active treatment 1: n = 1
- Active treatment 2: n = 1

Diagnosis: sleep-wake disorder

Diagnostic manual: not stated

Means of assessment: brief questionnaire on sleep behaviour

Comorbidity: not stated

Age: not stated, but psychology students

IQ: not stated, but psychology students

Sex: not stated

Ethnicity: not stated

Country: USA

Inclusion criteria

1. Participants indicating 31 minutes or greater in latency of steep onset
2. and willingness to participate in the study were contacted by phone by a female graduate assistant and scheduled for a pre-therapy interview
3. Interviews were conducted by two male research assistants unassociated with the remainder of the study

Exclusion criteria

1. Any participant reporting 30 minutes or less average sleep onset latency,
2. current use of drugs,
3. current contact with other professional services,
4. or whose sleep disturbance was shorter than six months in duration was excluded from the study

Comparisons

Psychological placebo

Treatment name : Quasi-desensitization placebo

Description of intervention: "The placebo condition involved a quasi-desensitization procedure. During Session 1 each subject constructed an IS-item hierarchy of chronological bedtime activities and chose six neutral images to be paired with the hierarchy items and to be used as substitutes for relaxation. Viewing sleep disturbance as a problem in which bedtime stimuli elicit responses (physiological and/or cognitive) incompatible with sheep, then the imaginal pairing of such stimuli with varied, neutral images should not theoretically change that functional relationship. In Sessions 2, 3,and 4, each item was presented six times with intervening presentations of neutral images. The subjects in this condition were told to practice hierarchy and neutral image visualizations twice a day, the Iasi practice being at least two hours prior to retiring. The hatter instruction was included to insure that practice would not increase sleep disturbance." (p. 159)

Individual or group treatment: group

Steinmark 1974 *(Continued)*

Exposure/intensity to treatment: 4 sessions – 1 weekly

Duration of treatment: 4 weeks

Concomitant psychotherapy: not allowed

Concomitant pharmacotherapy: not allowed

Wait-list

Comparison name: Waiting list no-treatment (wait-list)

Description of intervention: not stated

Exposure/intensity to treatment: no treatment during waiting

Duration of treatment: 4 weeks

Concomitant psychotherapy: not allowed

Concomitant pharmacotherapy: not allowed

Outcomes

Beneficial effect

- **Hierarchy:** patient-reported, clinical relevance
- **Outcome chosen:** the daily sleep questionnaires – Subscale difficulty experienced in falling asleep (0-5 - 5 being higher being more difficult to fall asleep)

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. Relaxation and desensitisation procedures produced significantly greater reports of improvement in latency of sleep onset than placebo and no treatment during the counterdemand period, while all three treated groups reported significantly greater improvement than no treatment after the fourth (positive demand) session
2. The results supported the effectiveness of relaxation therapy in the treatment of moderate insomnia
3. Demand characteristics may contribute to subject reports, but the use of counterdemand instructions allows for valid comparisons among therapy conditions.

Key limitations from study authors

1. It should be noted that while the counterdemand procedure allows valid comparisons among conditions, the issue of the validity of self- report data remains
2. Identification of the specific active ingredient in the relaxation procedure is left unanswered

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote:"The subjects were ranked on latency of sleep onset obtained in the pretreatment interview and were randomly assigned within severity blocks to one of four treatment conditions" (p. 159)

Steinmark 1974 *(Continued)*

Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	Yes	Attrition <15% (7.7%). One from each condition. Only reports data on completers
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	No	Quote:"The first author served as a therapist." (p.158)

Szymanski 1995

Study characteristics	
Methods	**Parallel randomised trial with six arms** 1. Psychological placebo: spider facts placebo 2. No-treatment 3. Active treatment 1: cognitive restructuring 4. Active treatment 2:cognitive restructuring in the presence of a spider 5. Active treatment 3:cognitive restructuring in the presence of a snake 6. Active treatment 4: an in vivo exposure condition with the spider **Sample calculation:** yes **Cluster randomised:** no **Duration of trial (baseline to post):** 2 weeks **Duration of participation (trial + follow-up):** 2 weeks. No follow-up **Setting:** outpatient (college) **Purpose of trial:** "The first hypothesis was concerned with whether cognitive restructuring is an effective intervention with spider phobics. The second hypothesis was whether the presence of anxiety in subjects while learning the cognitive restructuring techniques would be more effective than teaching the techniques while the subjects were in a neutral emotional state." "Finally, it was hypothesized that the addition of a cognitive intervention to a behavioral intervention would be more effective than the behavioral intervention alone" (p.134-35) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** 203 **Number of participants included:** 32 **Number of participants follow-up at post-treatment:** 32 **Number of participants randomly assigned to:** • Psychological placebo: n = 5 • No-treatment: n = 7

Szymanski 1995 *(Continued)*

- Active treatment 1: n = 7
- Active treatment 2: n = 5
- Active treatment 3: n = 3
- Active treatment 4: n = 5

Number of withdrawals: n = 0

Diagnosis: specific phobia (snake)

Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 3rd, Revised (DSM-III-R)

Means of assessment: Clinical interview

Comorbidity: not stated

Age: 18.4 mean years

IQ: not stated, but college students

Sex: 71.9% female

Ethnicity : not stated

Country: USA

Inclusion criteria

1. Snake phobics

Exclusion criteria

1. First, to protect subjects, volunteers were not allowed to participate if they reported being allergic to bee or wasp stings (n = 15).
2. Second, volunteers who did not meet the criterion of "phobic avoidance" on a behavioral avoidance test were eliminated (n = 46). That is, if a potential subject could touch the spider with his or her bare hand during the initial assessment session, they were not considered "phobic."
3. Third, volunteers who did not meet the criteria for snake phobia were excluded (n= 25).
4. Fourth,volunteers who indicated that they either did not want to continue in the experiment or could not meet during the group times were excluded (n = 83).
5. Finally, 2 subjects were dropped from the experiment (There were two reasons for dropping these subjects: 1) both subjects were the only ones in their treatment "group," and 2) these subjects' group leaders became ill and were unable to finish conducting the group meetings

Comparisons

Psychological placebo

Treatment name: Spider facts placebo

Description of intervention: “This condition consisted of a spider facts lecture. It was included to test whether the content of the cognitive restructuring conditions was responsible for the treatment effects. In a related vein, the credibility of the placebo groups (as well as the cognitive restructuring groups)was monitored.” (p. 139)

Individual or group treatment: group

Exposure/intensity to treatment: “subjects were asked to participate in three, approximately one-hour group sessions over a two-week period.” (p. 139)

Duration of treatment: 2 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy:: not stated

No-treatment

Szymanski 1995 *(Continued)*

Comparison name: No treatment

Description of intervention: "Subjects in the no-treatment control condition, the sixth condition, only participated in the two assessment sessions (i.e., pre- and post-tests) at a similar interval to the experimental conditions." (p. 139)

Exposure/intensity to treatment: no treatment

Duration of treatment: 2 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect outcome:

- **Hierarchy:** observer-reported
- **Outcome chosen:** Behavior Avoidance Test (BAT)

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. Results indicated that when working with spider phobics, three sessions of cognitive restructuring, in vivo exposure, or a facts lecture resulted in equal effectiveness immediately following treatment, but are more effective than no-treatment at all

Key limitations from study authors

1. Finally, it is important to point out the limitations of this study and the generalizability of these results
2. First, only three treatment sessions were used. Consequently, there were low to moderate effect sizes for the various dependent measures
3. Second, due to the low number of subjects completing this experiment, statistical power was poor for the SPQ (.44) and the BAT (.28), but high for the ES (.92)
4. Third, as previously mentioned, the death of the spider and its more active replacement may not have accurately reflected treatment gains for some of the subjects
5. Fourth, since a follow-up was not conducted it is unclear whether or not the treatment gains made by the different groups would have maintained, improved, or deteriorated over time
6. Fifth, due to the nature of the population used, i.e., college students not seeking treatment for their spider phobia, it is unclear how these interventions would affect a clinical population. In addition, there may have been a floor effect in this study. That is, since these subjects were not necessarily clinically phobic, they may not have started out high enough on the dependent measures (i.e., phobic enough) to show enough differential effect at the post-tests. This is supported by the fact that subjects were indistinguishable from the non-phobic subjects on some of the measures after only three sessions

Other notes from review authors

1. Only data on the completers

Conflicts of interest: one found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: "Subjects who were identified as snake phobics were then randomly assigned to one of the six conditions." (p. 139)

Szymanski 1995 *(Continued)*

		The groups were each given a number 1-5. After each subject completed the screening they were then added to the next group. So the first subject went into group 1, the second subject went into group 2, etc. We did not try to match or even review information from their screen. (Szymanski 1995 (pers comm))
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote:"As they signed up they were randomly assigned to conditions by a research assistant who was blind to the conditions." (Szymanski 1995 (pers comm))
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	No	Quote:"Due to the high levels of attrition it was important to determine whether the subjects who dropped out of the experiment were significantly different from those who remained in the study. Using two tailed t-tests we found that subjects who had initially agreed to continue in the study but subsequently dropped out did not have significantly different pre-test scores from subjects who completed the study." (p.144) Attrition unclear. Unclear number of randomised participants. The authors report themselves high levels of attrition
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Tan 1986

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Psychological placebo: supportive counselling 2. Wait-list 3. Active treatment: cognitive behavioural therapy (CBT) **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 8 weeks **Duration of participation (trial + follow-up):** 8 weeks treatment + 4 month follow-up **Setting:** outpatient **Purpose of trial:** "The present study evaluated the efficacy of group cognitive-behavior therapy for the alleviation of psychosocial problems and reduction of seizures with adult epileptic patients". (p.225) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** not stated **Number of participants included:** 30 **Number of participants followed-up at post-treatment:** 27

Tan 1986 *(Continued)*

Number of participants randomly assigned to:

- Psychological placebo: n= 10
- Wait-list: n= 10
- Active treatment: n = 10

Number of withdrawals: n = 3

- Psychological placebo: n = 0
- Wait-list: n= 1
- Active treatment: n = 2

Diagnosis: depression/anxiety

Diagnostic manual: not stated, but were referred by a neurologic clinician

Means of assessment: clinical interview - The Minnesota Multiphasic Personality Inventory (MMPI), Beck Depression Inventory (BDI)

Comorbidity: epilepsy

Age: 33.4 mean years (SD = 11.1)

IQ: not stated, but mentally retarded was excluded

Sex: 63% female

Ethnicity: not stated

Country: Canada

Inclusion criteria

1. The inclusion criteria for participation in the present study (i.e., adult epileptic patients with significant psychosocial problems and inadequate seizure control) were defined by the referring neurologist according to his clinical judgment

Exclusion criteria

1. Mentally retarded patients
2. Psychotic patients were excluded from the study

Comparisons

Psychological placebo

Treatment name: Supportive care

Description of intervention: "The SC group as an attention-placebo control group also received a total of eight 2-h sessions of group counselling or discussion that was mainly supportive in nature. Techniques such as reflection and clarification of feelings were used, but no specific cognitive-behavioral strategies were taught. This SC group intervention was meant to provide the "nonspecific" factors of any group psychological intervention or therapy, such as therapist attention, suggestion, group support," (p.227)

Individual or group treatment: group

Exposure/intensity to treatment: 8 x 2 hours sessions

Duration of treatment: 8 weeks

Concomitant psychotherapy: 60% received concomitant therapy. "Six patients in each of the three groups did not receive any concomitant or other professional counselling or psychiatric treatment during the present study. The remaining patients received treatment from other therapists while participating in the present study, but such ongoing treatment could not be terminated for obvious ethical reasons" (p. 227)

Tan 1986 *(Continued)*

Concomitant pharmacotherapy: all patients received anticonvulsant medication for their epilepsy

Wait-list

Comparison name: Wait-list

Description of intervention: "The WL group did not receive any group therapy until after the present study was completed. Patients assigned to this group were seen for the three assessments at about the time the CBT and SC groups had them (i.e., before therapy or pre, after therapy or post, and at a 4-month follow-up)" (p. 227)

Exposure/intensity to treatment: waiting for therapy

Duration of treatment: 8 weeks

Concomitant psychotherapy: 60% received concomitant therapy. "Six patients in each of the three groups did not receive any concomitant or other professional counselling or psychiatric treatment during the present study. The remaining patients received treatment from other therapists while participating in the present study, but such ongoing treatment could not be terminated for obvious ethical reasons" (p. 227)

Concomitant pharmacotherapy: all patients received anticonvulsant medication for their epilepsy

Outcomes

Beneficial effect

- **Hierarchy:** continuous, patient-reported, clinical relevance
- **Outcome chosen:** Beck depression inventory (BDI)

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. Overall, little support was found for the efficacy of group cognitive behavior therapy (eight 2-h weekly sessions) for the reduction of psychosocial difficulties or seizures.

Key limitations from study authors

1. First, eight 2-hour weekly sessions of cognitive-behavior therapy may not have been sufficiently long or comprehensive to produce significant therapeutic change on more of the outcome measures used
2. Individual cognitive-behaviour therapy may be more effective than group cognitive-behavior therapy with epileptic patients
3. An active, coping-skills-oriented group cognitive-behavior therapy like the one used in the present study may not be equally effective or helpful to all adult epileptic patients, especially those who may not be particularly oriented to self-control skills
4. The inclusion criteria for participation in the present study (i.e. adult epileptic patients with significant psychosocial problems and inadequate seizure control) were defined by the referring neurologist according to his clinical judgment
5. It should be pointed out that the present study yielded one positive finding on therapist's global ratings of patients' psychological adjustment in the expected direction

Other notes from review authors

1. Only data on the completers
2. SD was imported from Fuchs 1977 (same scale and population)

Conflicts of interest: one found

Judgement: yes

Risk of bias

Tan 1986 *(Continued)*

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	"Quote - Twenty-seven outpatients were randomly assigned to one of three groups: Cognitive-Behavior Therapy, Supportive Counseling (attention-placebo control), and Waiting list (no treatment control)". (p. 225)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Attrition <15% (10%). No ITT. On reports data on completers
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Teri 1997

Study characteristics	
Methods	**Parallel randomised trial with four arms** 1. Usual care: typical care control (TCC) 2. No-treatment: wait-list 3. Active treatment 1: Behavior Therapy-Pleasant Events (BT-PE) 4. Active treatment 2: Behavior Therapy-Problem-solving (BT-PS) **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 9 weeks of treatment **Duration of participation (trial + follow-up):** 9 weeks of treatment, follow-up at 6 months. **Setting:** outpatient **Purpose of trial:** “The current study is a controlled clinical investigation of two non-pharmacological treatments of depression in patients with Alzheimer’s disease. Two active behavioral treatments, one emphasizing patient pleasant events and one emphasizing caregiver problem solving, were compared to an equal-duration typical care condition and a wait list control.” (p. 159)
Data	**Number of participants screened:** not stated **Number of participants included:** 88 **Number of participants followed-up at post-treatment:** 72 **Number of participants randomly assigned to:** • Usual care: n = 10 • No-treatment: n= 20 • Active treatment 1: n= 23

Teri 1997 *(Continued)*

- Active treatment 2: n= 19

Number of withdrawals : n = 16

- Usual care: not stated
- No-treatment: not stated
- Active treatment 1: not stated
- Active treatment 2: not stated

Diagnosis: depression (in dementia patients) - 75% were diagnosed with major depressive disorder, 25% were diagnosed with minor depressive disorder

Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition, Revised (DSM-III-R)

Means of assessment: clinical interview – Schedule for Affective Disorders and Schizophrenia

Comorbidity: dementia

Age: 76.4 mean years (SD = 8.2)

IQ: not stated

Sex: 47% female

Ethnicity: not stated

Country: USA

Inclusion criteria

1. Meet National Institute of Neurological and Communicative Disorders and Stroke (NINCDS-ADRDA) criteria for probable Alzheimer's Disease
2. Have at least a six-month history of cognitive problems
3. Live with their caregivers in the community
4. Meet Research Diagnostic Criteria (RCD) and Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R) criteria for major or minor depressive disorder (not including the exclusionary criteria for dementia)
5. Have a Hamilton Depression Rating Scale score of at least 10

Exclusion criteria

1. Not stated

Comparisons

Usual care

Treatment name: typical care control (TCC)

Description of intervention: “Subjects in this condition were given information, advice, and support with their efforts to manage patient problems. No specified homework assignments or record keeping were provided. Specific problem solving or behavioral strategies were not implemented. Therapists have suggestions and advice of an unstructured nature.” (p. 161)

Individual or group treatment: individual

Exposure/intensity to treatment: one 60-minute session per week. Duration of treatment: 9 weeks

Concomitant psychotherapy: allowed to receive other treatment

Concomitant pharmacotherapy: allowed to receive other treatment

No-treatment

Comparison name: Wait list control (in reality no-treatment)

Teri 1997 *(Continued)*

	Description of intervention: "Subjects in this condition received no contact with therapists. Following assignment, they were informed that they would receive no active intervention during the 9-week period. Immediately following the 9 weeks, they were post-tested." (p. 161) **Exposure/intensity to treatment:** no treatment **Duration treatment:** 9 weeks **Concomitant psychotherapy:** aAllowed to receive other treatment **Concomitant pharmacotherapy:** allowed to receive other treatment
Outcomes	**Beneficial effect** • **Hierarchy:** observer-rated, clinical relevance • **Outcome chosen:** Cornell Scale for Depression in Dementia – caregiver **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. "The results of this study support the effectiveness of behavioral treatment of depression in patients with AD. Patients and caregivers receiving behavioral treatment with either a pleasant event or problem-solving focus demonstrated significant reductions in their level of depression following treatment, and those reductions were maintained at follow-up. Patients and caregivers receiving behavioral treatment improved significantly more than those receiving an equal duration typical care or wait list control." (p. 165) 2. "Patients with major depressive disorder, who were experiencing symptoms of dysphoria, loss of interest, suicidal ideation, appetite change, and sleep disturbance, were most likely to benefit from treatment. Those experiencing predominantly dementia-related symptoms, such as difficulty concentrating, were less likely to show improvement (…) Depression symptoms were hypothesized to improve; cognitive symptoms were not." (p. 165) **Key limitations from study authors** 1. First, the assessment of depression in patients with dementia is not yet perfected 2. Second, the factors influencing successful treatment outcome are likely to be multifactorial **Other notes from review authors** 1. Only reports data on completers **Conflicts of interest:** none found **Judgement:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Yes	Quote: "Subjects were randomly assigned to one of four treatment conditions: behavior therapy-pleasant events (BT-PE); behavior therapy-problem solving (BT-PS); typical care control (TCC); and wait list control (WLC)." (p. 160) "Conducted independently by statistician using computer program" (Teri 1997 (pers comm))
Allocation concealment	Yes	Quote: "Randomisation concealed from the researchers" (Teri 1997 (pers comm))

Teri 1997 *(Continued)*

Blinding of outcome assessors	Yes	Quote: "(…) assessed at pre-, post-, and 6- month follow-up intervals by interviewers blind to treatment assignment." (p. 160) Quote: "Interviews were conducted by experienced master's- and PhD-level clinical geriatric interviewers, blind to treatment condition." (p. 160)
Blinding of participants and personnel	No	Not possible to blind usual care and wait-list
Incomplete outcome data	No	Quote: "Eighty-eight patient- caregiver pairs began the study; 72 (82%) completed the pretest, 9-week intervention, and posttest. Subjects who discontinued treatment did so for the following reasons: serious medical illness (n = 4), change in living situation (n = 4), exclusionary medication prescribed during the intervention stage (n = 2), and caregiver stopped participating (n = 6). No significant differences were obtained on baseline measures between subjects who did and did not discontinue treatment." (p. 160) Attrition >15% (18.2%). No ITT used. Reports data on completers only
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	Yes	No other sources found

Tori 1973

Study characteristics	
Methods	**Parallel randomised trial with five arms** 1. Physical placebo: high expectancy placebo: 9. 2. No-treatment control: 10. 3. Active treatment 1: specific cognitive: 10. 4. Active treatment 2: general cognitive: 9. 5. Active treatment 3: counterconditioning: 9. **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 2 weeks **Duration of participation (trial + follow-up):** 2 weeks treatment + follow-up at 30 days. **Setting:** outpatient **Purpose of trial:** "The present study was designed in order that comparisons could be made between counterconditioning, expectancy, and cognitive-coping variables in the reduction of the fear and anxiety of snake-phobic college students." (p. 270) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** initial pool of 700 undergraduate students **Number of participants included:** 47 **Number of participants followed-up at post-treatment:** not stated **Number of participants randomly assigned to:** • Physical placebo: n = 9 • No-treatment control: n= 10 • Active treatment 1: n = 10 • Active treatment 2: n = 9 • Active treatment 3: n = 9

Tori 1973 *(Continued)*

Number of withdrawals: not stated

Diagnosis: specific anxiety (snakes)

Diagnostic manual: not stated

Means of assessment: self-rating on a 5-point scale + approach test (excluded if they went beyond Point 8)

Comorbidity: not stated

Age: not stated

IQ: not stated, college students

Sex: 70.2% female

Ethnicity: not stated

Country: USA

Inclusion criteria

1. Rated themselves on a 5-point scale as being “much” or “very much” afraid of harmless snakes

Exclusion criteria

1. ”Anyparticipant who went beyond Point 8, touching the snake, was disqualified from the study.” (p. 271)

Comparisons

Physical placebo

Treatment name: High expectancy placebo
Description of intervention: “Electrodes were then attached to the subject’s wrists, left ankle, and left index finger. A polygraph and several other impressive but nonfunctional machines (all with flashing lights or moving dial indicators) were activated. Deep muscle relaxation was then induced, following the principles of progressive relaxation (...). The subjects then spent about 10 minutes imagining psychologically pleasant scenes, with the entire session lasting approximately 40 minutes” (p. 272)
Individual or group treatment : individual
Exposure/intensity to treatment: “Days spent in treatment were held constant by having the subjects meet either on Monday and Wednesday or Tuesday and Thursday for a two-week period. (...) Each subject served individually in four treatment sessions lasting from 30 to 50 minutes each.” (p. 271)
Duration of treatment: 2 weeks
Concomitant psychotherapy: not stated
Concomitant pharmacotherapy: not stated

No-treatment
Comparison name: No-treatment control
Description of intervention:
“The pretest, posttest, and follow-up batteries were administered to these subjects (n=10, 7 females and 3 males) in order to determine the amount of improvement due to repeated exposure to the snakes and to control for any nonspecific changes that might occur over the time of the experiment.” (p. 271)
Exposure/intensity to treatment : no treatment
Duration of treatment: 2 weeks
Concomitant psychotherapy: not stated
Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect

- **Hierarchy:** observer-rated
- **Outcome chosen:** approach test

Adverse events

Tori 1973 *(Continued)*

- No data reported on adverse events

Notes

Key conclusion from study authors

1. "First, the nonsignificant differences between the counterconditioning and the no-treatment control groups suggest that there are no automatic reductions in human fear and avoidant behavior when graded aversive stimuli are contiguously paired with anxiety-competing deep muscle relaxation." (p. 276)
2. "Second, cognitive-expectancy variables appear to be the critical factors in reducing human avoidant behavior." (p. 276)
3. "(…) cognitive-expectancy model were confirmed in that the combined posttreatment and follow-up outcomes measures of the high-expectancy placebo and cognitive-coping groups were not significantly different from each other and were superior to those of the counterconditioning and no-treatment control groups. This result is taken as strong support for the supposition that the efficacy of systematic desensitization therapy is due to cognitive-expectancy variables rather than to the conditioning of antagonistic responses." (p. 276)

Key limitations from study authors

1. 1. "We can begin by stating that all of the counterconditioning procedures (...) were carefully followed. Still, one possible criticism might be that there were "gaps" in the stimulus hierarchy that we employed." (p. 276)
2. 2. "It should be noted, however, that in our high-expectancy placebo treatment, both demand and expectancy variables were operative. This means that in addition to convincing the subjects that they would be able to approach the phobic object with greater ease, they were also placed under a great deal of pressure to be cooperative and to show the desired therapeutic improvements (…) Since the separation of demand and expectancy variables was not attempted in our experiment, it was not possible to separately ascertain the amount of behavioral and self-report changes that could be attributed to altered expectancy concerning the feared stimulus and to the demands of the experimental situation." (p. 277)

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "The first 40 subjects were randomly assigned to the five groups employed in the experiment. In order to more closely equate groups for initial snake approach behavior, the last 7 subjects were assigned on the basis of their pretest approach scores." (p. 271)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote: "The female experimenter who conducted post-testing and follow-up testing was not aware of the subjects' treatment conditions." (p. 273)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Unclear	No information
Selective outcome reported	Unclear	No protocol found

Tori 1973 *(Continued)*

Other sources of bias	Yes	No other sources found

Trexler 1972

Study characteristics

Methods

Parallel randomised trial with three arms

1. Psychological placebo: attention control
2. Wait-list: no-treatment
3. Active treatment: rational-emotive therapy

Sample calculation: not stated

Cluster randomised: no

Duration of trial (baseline to post): 1 week (4 sessions spaced several days apart)

Duration of participation (trial + follow-up): 1 week. No follow-up

Setting: outpatient (University)

Purpose of trial: "In a partial replication and refinement of an earlier study by the authors, 33 college student volunteers reporting high levels of public-speaking anxiety received rational-emotive therapy (RET), attention placebo (AP), or no treatment (NT)." (p. 60)

Open/closed placebo: closed placebo

Data

Number of participants screened: not stated

Number of participants included: 33

Number of participants followed-up: not stated

Number of participants randomly assigned to:

- Psychological placebo: n = 10
- Wait-list: n = 12
- Active treatment: n =11

Number of withdrawals: not stated

Diagnosis: specific anxiety (public-speaking anxiety)

Diagnostic manual: not stated

Means of assessment: level 4 on a 10-point behaviorally anchored rating scale of anxiety

Comorbidity: not stated

Age: not stated, undergraduates

IQ: not stated - but college students

Sex: 51.5% female

Ethnicity: not stated

Country: USA

Inclusion criteria

Trexler 1972 *(Continued)*

1. Level 4 on a 10-point behaviorally anchored rating scale of anxiety. The anchor at Level 4 read, "My anxiety is a little more than most people feel, and I have been reluctant to give speeches and somewhat hesitant to comment or ask questions in class as well." No 5s were phobic, if this is defined as a complete avoidance of public-speaking situations (Level 10 on the preliminary scale)

Exclusion criteria

1. Not stated

Comparisons	**Psychological placebo** **Treatment name:** Attention placebo **Description of intervention:** ”AP treatment consisted mainly of the typical training in relaxation employed in systematic desensitization studies, without presentation of stimulus hierarchies. The expectancy was communicated continually that this procedure was a well-regarded treatment for general anxiety and as such would be helpful in overcoming public-speaking anxiety. Training was effected both with a tape recording and viva voce. Homework assignments for all sessions consisted of practicing relaxation skills and reading specially prepared materials (...), emphasizing both the purpose and techniques of relaxation.” (p. 61) **Individual or group treatment:** group **Exposure/intensity to treatment:** 4 sessions spaced several days apart **Duration of treatment:** 1 week **Concomitant psychotherapy:**: not stated **Concomitant pharmacotherapy:** not stated **Wait-list** **Comparison name:** No-treatment (in reality wait-list) **Description of intervention:** “Those in the NT group were informed that there would be a short delay before they would begin therapy” (p. 61) **Exposure/intensity to treatment:** no treatment during waiting **Duration of treatment:** 1 week **Concomitant psychotherapy** : not stated **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy:** observer-reported, clinical relevance, psychometric properties • **Outcome chosen:**a Finger Sweat Print (FSP) **Adverse events** • No data on adverse events reported
Notes	**Key conclusion from study authors** 1. Primary analyses of pre-therapy to post-therapy changes as assessed with a variety of self-report and observational measures tended to support the conclusion that RET is more effective than either NT or the AP treatment used (relaxation training) **Key limitations from study authors** 1. The use of a single therapist and target symptom approach limits the generalizations to be made from this study.

Trexler 1972 *(Continued)*

Other notes from review authors

1. Usable data not available

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Treatment groups were matched approximately on the basis of the preliminary self- rating and randomly assigned to treatment conditions at the initial speech session" (p. 61)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote: "Pretherapy and posttherapy evaluations were made by a graduate student and an advanced undergraduate psychology honors student trained in observation prior to the study and kept blind as to 5s' treatment." (p. 60)
Blinding of participants and personnel	No	Not possible to blind placebo and "no treatment"
Incomplete outcome data	Unclear	Attrition unclear
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	No	The first author served as therapist

Turner 1979

Study characteristics

Methods

Parallel randomised trial with five arms

1. Psychological placebo
2. Wait-list
3. Active treatment 1: paradoxical intention
4. Active treatment 2: stimulus control
5. Active treatment 3: progressive relaxation

Sample calculation: yes

Cluster randomised: no

Duration of trial (baseline to post): 4 weeks

Duration of participation (trial + follow-up): 4 weeks. No follow-up

Setting: outpatient

Purpose of trial: "In light of these criticisms, the present experiment incorporates appropriate modifications in contrasting the effectiveness of stimulus control with progressive relaxation in ameliorating the complaints of severe insomniacs. A second purpose of the present study is to provide an initial experimental investigation of the efficacy of paradoxical intention for the treatment of insomnia and

Turner 1979 *(Continued)*

to compare this technique with the two more conventional procedures for reducing disorders of sleep (i.e., stimulus control and progressive relaxation)" (p. 501)

Open/closed placebo: closed placebo

Data	**Number of participants screened:** 115 **Number of participants included:** 50 **Number of participants followed-up:** 50 **Number of participants randomly assigned to:** • Psychological placebo: n = 10 • Wait-list: n = 10 • Active treatment 1: n = 10 • Active treatment 2: n = 10 • Active treatment 3: n = 10 **Number of withdrawals** : n = 0 **Diagnosis:** sleep-wake disorder **Diagnostic manual:** not stated **Duration of treatment:** 1 week **Means of assessment:** Clinical interview (Monroe's Daily Sleep Questionnaire) **Comorbidity:** no physiological or psychological complaints **Age:** 39 mean years (range = 24 to 79) **IQ:** not stated **Sex:** 50% female **Ethnicity:** not stated **Country:** USA **Inclusion criteria** 1. Sleep disturbance **Exclusion criteria** 1. Of 74 persons who were interviewed, 9 were rejected due to subclinical levels of sleep disturbance 2. Eight were judged unsuitable because their sleep difficulties were secondary to other physiological (e.g. arthritis) or psychological (e.g. depression) complaints 3. Any volunteer demonstrating secondary insomnia (i.e., a complication of other physical or psychological problems) was excluded from the study
Comparisons	**Psychological placebo** Treatment name: Placebo control Description of intervention: "The placebo condition used was that described by Steinmark 1974 . Their "quasi-desensitization" condition required clients to construct an individualized 18-item hierarchy of chronological bedtime activities to be paired with six neutral images. During the first session clients were taught how to construct the hierarchy and develop six neutral images. Sessions 2, 3, and 4 consisted of imagining the hierarchies and alternating neutral images between scenes. Clients were instructed to practice the procedure twice a day, the last practice not being within 2 hours of bedtime." (p. 504)

Turner 1979 *(Continued)*

Individual or group treatment: individual

Exposure/intensity to treatment: 30 to 45 minutes once per week

Duration of treatment: 4 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

Wait-list

Comparison name: Wait-list

Description of intervention:"The no-treatment control clients were asked to forego treatment for 5 weeks and to serve as control subjects. These subjects were assured of receiving treatment in 4 weeks" (p. 504)

Exposure/intensity to treatment: Waiting for treatment

Duration treatment: 4 to 5 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: not stated

Outcomes

Beneficial effect

- **Hierarchy:** patient-reported, clinical relevance
- **Outcome chosen:** daily Sleep Questionnaire (DSQ) Subscale difficulty experienced in falling asleep (0-7 - 7 being no difficulties falling asleep)

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. The results indicated that each of the therapeutic procedures significantly reduced sleep complaints in contrast to placebo and waiting list control groups
2. No differences were observed among the three active techniques

Key limitations from study authors

1. Two shortcomings with the credibility assessment procedure should be noted. One deficiency centers on the method of assessment. The self-report questionnaire used to assess credibility has not been subjected to psychometric validity and reliability examination and is therefore subject to the same demand characteristic problems from which all paper-and pencil measures suffer Second, the practice of assessing credibility of treatment rationales following therapy vitiates an unambiguous interpretation of the results. Since clients had 5 therapeutic hours invested in their respective programs, they may have been unwilling to provide negative credibility ratings, or the ratings may reflect the degree of success of treatment and not the credibility and expectancy for improvement generated by the therapeutic rationales
2. An additional methodological flaw of the present study was the use of a single therapist to conduct all of the treatment sessions
3. No significant effect for the therapist factor has been detected either for treatment type, symptom type, or sex of the client. Thus, although the use of only one therapist clearly limits the external validity of the present study, previous relevant experiments and additional data gathered by us suggest these procedures can be successfully applied by any therapist
4. Thus, the random assignment of clients to the treatment conditions employed in the present experiment might have obscured the differences among the therapeutic procedures by failing to match specific sleep difficulty with appropriate treatment

Turner 1979 *(Continued)*

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Following the 10-day baseline period, 10 clients were randomly assigned to each of the five treatments." (p. 502)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and "no-treatment"
Incomplete outcome data	Yes	Attrition <15% (0%).
Selective outcome reported	Unclear	Protocol not found
Other sources of bias	No	Quote: “the first author served as the therapist in all conditions.” (p. 502-3)

Vanderplate 1983

Study characteristics

Methods

Parallel randomised trial with four arms

1. Physical placebo: EMG Pseudo-feedback Training Group
2. Wait-list: No-treatment Waiting-list Control group
3. Active treatment 1: EMG Biofeedback Training Group
4. Active treatment 2: Self-monitoring Group

Sample calculation: yes
Cluster randomised: not stated
Duration of trial (baseline to post): for EMG-groups: 3 weeks
Duration of participation (trial + follow-up): 3 weeks of treatment +2 months follow-up
Setting: outpatient
Purpose of trial: “The purpose of the present study was to assess the efficacy of frontalis EMG biofeedback as a treatment for insomnia within a more adequately designed and rigorously controlled investigation. No previous study in this area has controlled for the reactive effects of sleep behavior self-monitoring, or employed a pseudo-treatment control group sufficiently equivalent on all non-specific treatment variables. Therefore, success of EMG biofeedback in alleviating insomnia to date may attribute to self-monitoring effects or demand and expectancy effects.” (p. 32)
Open/closed placebo: closed placebo

Data

Number of participants screened: 625
Number of participants included: 36
Number of participants followed-up at post-treatment: 24

Vanderplate 1983 *(Continued)*

Number of participants randomly assigned to:

- Physical placebo: n = 9
- Wait-list: n = 9
- Active treatment 1: n = 9
- Active treatment 2: n = 9

Number of withdrawals: n = 12

- Physical placebo: n = 3
- Wait-list: n = 3
- Active treatment 1: n = 3
- Active treatment 2: n = 3

Diagnosis: insomnia (sleep-wake disorder)
Diagnostic manual: not stated
Means of assessment: self-report – Sleep Behavior Questionnaire
Comorbidity: "Thirteen coeds reported currently experiencing severe headaches, and all but one coed reported that these headaches were stress related." (p. 40)
Age: 20.38 mean years, (rRange = 18 to 25 years of age)
IQ: not stated
Sex: 100% female
Ethnicity: all were Caucasian
Country: USA

Inclusion criteria

1. "This study employed females who suffered from insomnia, defined as a sleep onset latency greater than 30 minutes, on at least three nights per week, present for at least three consecutive months prior to the study." (p. 38)

Exclusion criteria

1. If participants reported having taken any prescription medication for sleep induction within five weeks prior to the beginning of the study
2. Any coed having previous experience in biofeedback training
3. Currently under other treatment for insomnia
4. Having medical of psychological problem capable of affecting sleep

Comparisons

Physical placebo
Treatment name (type): EMG Pseudo-feedback Training Group
Description of intervention: "All coeds of this group received treatment identical to that of the EMG biofeedback training group with the exception that feedback provided was noncontingent upon their EMG levels (...) During the training phase, each daily training session conformed to the following procedure: the coed spent 2 ½ minutes resting quietly in the reclining chair in the training room, followed immediately by 2 ½ minutes with the electrodes attached to the forehead. Then, the tape recorder was turned on and each coed heard feedback recorded from the
training session of a randomly chosen biofeedback group coed." (p. 47)
Individual or group treatment: individual
Exposure/intensity to treatment: "During the training phase, each daily training session conformed to the following procedure: the coed spent 2 ½ minutes resting quietly in the reclining chair in the training room, followed immediately by 2 ½ minutes with the electrodes attached to the forehead." (p. 47)
Duration of treatment: 3 weeks
Concomitant psychotherapy: no - exclusion if subject was under other treatment for insomnia
Concomitant pharmacotherapy: no – "Subjects were excluded from this study if they reported having taking any prescription medication for sleep induction within five weeks prior to the beginning of the study." (p. 40)

Wait-list
Comparison name: No-treatment Waiting-list Control group

Vanderplate 1983 *(Continued)*

Description of intervention: "Coeds of this group were informed at the initial pre-treatment meeting that all treatment groups were filled, and that they would receive treatment at the earliest possible time. They were requested to complete only the Sleep Behavior Questionnaire and the STAI following completion of this study
Exposure/intensity to treatment: no treatment
Duration treatment: not stated
Concomitant psychotherapy: no – exclusion if participant was under other treatment for insomnia
Concomitant pharmacotherapy: no – "Subjects were excluded from this study if they reported having taking any prescription medication for sleep induction within five weeks prior to the beginning of the study." (p. 40)

Outcomes

Beneficial effect

- **Hierarchy** : patient-reported, clinical relevance
- **Outcome chosen:** the Sleep Behavior Questionnaire, difficulty falling asleep

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. "The conclusion seems warranted, therefore, that anxiety or muscle tension played some role in the maintenance of sleep disturbance at least in this female student population. These finding suggest that treatment strategies which reduce anxiety may be efficacious in alleviating sleep disturbance, regardless of the anxiety reduction mechanisms involved. Such treatment strategies thus provide a viable alternative to pharmacological treatments and their dire consequences." (p. 115)
2. "Perhaps the most interesting and unexpected finding of this investigation was the relative lack of difference between the biofeedback group and pseudo-feedback group on sleep improvement across treatment. Both groups reported significant improvement over treatment and remained improved at follow-up. This finding raises the serious question of whether sophisticated equipment and the delivery of physiological information to an individual is actually necessary to alleviate sleep onset insomnia." (p. 116)

Key limitations from study authors

1. Not stated

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: (…) 36 were assigned to one of four groups, with nine coeds per group. An attempt was made to make each group as equivalent as possible with regard to sleep behavior indices." (p. 44)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list

Vanderplate 1983 *(Continued)*

Incomplete outcome data	No	Quote: "Data from one coed of the biofeedback group were randomly excluded from the results to allow an equal N per group." (p. 38) Attrition >15% (33,3%). No ITT. Only reports data on completers
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Watzl 1988

Study characteristics	
Methods	**Parallel randomised trial with four arms** 1. Pharmacological placebo: Placebo injections 2. No-treatment 3. Active treatment 1: surveillance 4. Active treatment 2: surveillance and placebo injections **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 3 months **Duration of participation (trial + follow-up):** 3 months + 1 year follow-up **Setting:** inpatient **Purpose of trial:** "This study investigates the impact of both the strict surveillance of patient's intake and of nonspecific medication on relapse rates." (p. 197) **Open or closed placebo:** closed placebo
Data	**Number of participants screened:** not stated **Number of participants included:** 80 **Number of participants followed-up at post-treatment:** 70 **Number of participants randomly assigned to** : • Pharmacological placebo: n = 18 • No-treatment: n =16 • Active treatment 1: n = 20 • Active treatment 2: n = 16 **Number of withdrawals** : n = 0 **Diagnosis** : alcohol substance use **Diagnostic manual:** International Statistical Classification of Diseases and Related Health Problems, 9th edition (ICD-9) **Means of assessment:** not stated. The duration of dependence was 7.2 years on the average. Ninety per cent of the women reported daily or nearly daily consumption **Comorbidity:** not stated **Age:** 36.7,mean years (range = 21 to 59) I **Q:** not stated

Watzl 1988 *(Continued)*

Sex: 100% female

Ethnicity: not stated

Country: Germany

Inclusion criteria

1. Inpatients at public psychiatric hospital

Exclusion criteria

1. Not stated

Comparisons	**Pharmacological placebo** **Treatment name:** Placebo Injection **Description of intervention** : "A placebo condition is employed to assess the psychological sequelae of nonspecific medication." (p. 197) **Individual or group treatment** : Individual **Exposure/intensity to treatment:** 2 injections every week. **Duration of treatment** : 3 months **Concomitant psychotherapy** : Behavioral treatment inpatient **Concomitant pharmacotherapy** : not stated **No-treatment** **Comparison name:** No-treatment **Description of intervention:** only assessment before and after. **Exposure/intensity to treatment:** no treatment **Duration of treatment:** 3 months **Concomitant psychotherapy:** behavioral treatment inpatient **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy** : global score, psychometric properties • **Outcome chosen:** evident relapse, combined **Adverse events** • No data reported on adverse events
Notes	**Key conclusion from study authors** 1. The amount of surveillance did not influence relapse rates 2. However, during inpatient treatment more relapses occurred among patients who received the placebo injections than among those who did not **Key limitations from study authors** 1. Not stated **Other notes from review authors** 1. None **Conflicts of interest:** none found **Judgement:** yes
Risk of bias	

Watzl 1988 *(Continued)*

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "To assign patients to one of the four conditions, quadruples matched for age and number of previous detoxifications were formed. First, two of these four patients were randomly assigned to the "placebo"- condition, the remaining two patients to the "no placebo"-condition. Patients were informed that only those were to receive the medication (called EWOCA), for whom a certain blood factor indicates that they would profit from such a treatment." (p. 191)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Unclear	No information
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Yes	Quote: "Of these, 10 were not considered in the final evaluation. Four of them were husbands of female patients. Six patients had dropped out of treatment due to reasons not related to the investigation." (p. 198) Attrition <15% (12.5). Outcome were relapses
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Weingaertner 1971

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Physical placebo 2. No-treatment 3. Active treatment: self-shock **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 2 weeks **Duration of participation (trial + follow-up):** 2 weeks. No follow up **Setting:** inpatient **Purpose of trial** : "The present study employed a three-group design in which hospitalized schizophrenics were given a shock box and told to shock themselves each time they experienced hallucinating voices." (p. 422) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** 95 **Number of participants included:** 45 **Number of participants followed-up at post-treatment:** not stated **Number of participants randomly assigned to:** • Physical placebo: n = 15 • No-treatment: n = 15 • Active treatment: n = 15

Weingaertner 1971 *(Continued)*

Number of withdrawals: not stated
Diagnosis: schizophrenia
Diagnostic manual: not stated
Means of assessment: not stated
Comorbidity: none were physically handicapped. Two patients with brain damage.
Age: 37.2 mean years
IQ: not stated
Sex: 100% male
Ethnicity: not stated
Country: USA

Inclusion criteria

1. The patient's admission to experiencing auditory hallucinations
2. his expressed willingness to carry out the self-imposed shock procedure for a two-week period

Exclusion criteria

1. Not stated

Comparisons

Physical placebo
Treatment name: Shock placebo
Description of intervention: "Patients in the placebo group carried a box which gave no shock." (p. 422)
Individual or group treatment: individual
Exposure/intensity to treatment: all patients in the self-shock and placebo groups were told to retain the box for two weeks and to get in touch with E if they had any trouble. Any adjustments necessary during the two weeks were made promptly.
Duration of treatment: 2 weeks
Concomitant psychotherapy: not stated
Concomitant pharmacotherapy: no hallucinogenic drugs

No-treatment
Comparison name: No-treatment
Description of intervention: "The no-treatment group received only the pre and post evaluations which were given all Ss Exposure/intensity to treatment: "(p. 422)
Duration of treatment: 2 weeks
Concomitant psychotherapy: not stated
Concomitant pharmacotherapy: no hallucinogenic drugs

Outcomes

Beneficial effect

- **Hierarchy:** observer-reported, clinical relevance
- **Outcome chosen:** Brief Psychiatric Rating Scale: hallucination scale

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. No significant differences between groups were found
2. It was concluded that placebo was the primary agent of change
3. Conscious cognitive factors seemed central to the improvement

Key limitations from study authors

1. Several implications arise from this study. First, in the absence of evidence of symptom substitution or other ill effects (no patient showed any overall deterioration), many assumptions about the putatively inhuman nature of aversive methods can again be questioned. Indeed, shock-group patients pressed significantly more times than placebo-group patients

Weingaertner 1971 *(Continued)*

2. Second, the effects attributable to shock per se, if any, were elusive, so large was the change due to other factors. It is evident that use of single-S designs for investigation of similar phenomena may show change which is incorrectly interpreted
3. Third, verbal pre-post evaluation interviewing may sensitize patients in such a way as to constitute or mimic a therapeutic effect

Other notes from review authors

1. Distribution of patients randomised to the groups were unclear. We assumed that it was equally distributed
2. Transformative data taken from Hróbjartsson 2010

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "The 5s were randomly assigned to three groups: self-shock, placebo, and no treatment" (p. 423)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote: ”The Brief Psychiatric Rating Scale (...) was filled out by two raters independently on the basis of an 18-minute interview. (...) The raters were not aware of the project for which the patient was being evaluated and in general did not know whether a rating was the first, last, or one of a series of routine ratings as required in some of the studies conducted simultaneously with this one.” (p. 424)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment
Incomplete outcome data	Unclear	No information
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Whittaker 1963

Study characteristics

Methods

Parallel randomised trial with three arms

1. Pharmacological placebo: placebo elixir
2. No-treatment: no medication
3. Active treatment: active elixir

Sample calculation: not stated
Cluster randomised: no
Duration of trial (baseline to post): 10 weeks
Duration of participation (trial + follow-up): not stated
Setting: inpatient

Whittaker 1963 *(Continued)*

Purpose of trial: "Whilst a number of these studies was controlled, in none of the investigations was a separate control group receiving no medication employed. This is a technique which would help to gauge the relative importance of placebo and pharmacological effects." (p. 422)
Open/closed placebo: closed placebo

Data

Number of participants screened: not stated

Number of participants included: 39

Number of participants followed-up at post-treatment: 39

Number of participants randomly assigned to:

- Pharmacological placebo: n = 13
- No-treatment: n = 13
- Active treatment: n = 13

Number of withdrawals : n= 0
Diagnosis: chronic schizophrenia
Diagnostic manual: not stated, but diagnosed by at least two psychiatrists
Means of assessment: clinical interview
Comorbidity: all had a paranoid condition, but two had also catatonic tendencies, and another had hebephrenic features. Six had been leucotomised prior to 1952
Age: 50 mean years, (range = 27 to 66)
IQ: not stated
Sex: 100% male
Ethnicity: not stated
Country: Scotland

Inclusion criteria

1. They had all been diagnosed by at least two psychiatrists as having chronic schizophrenia they had all a paranoid condition, but two had also catatonic tendencies, and another had hebephrenic features
2. Six had been leucotomised prior to 1952. Their behaviour and symptomatology over a long period were well known to the doctor and to the four charge nurses who were making the assessments

Exclusion criteria

1. Not stated

Comparisons

Pharmacological placebo
Treatment name: Placebo elixir
Description of intervention:: not stated
Individual or group treatment: individual
Exposure/intensity to treatment: 8 mg (3 patients received 20/12/12mg)
Duration of treatment: 10 weeks
Concomitant psychotherapy: not stated – but inpatient
Concomitant pharmacotherapy: "They had all been receiving maintenance doses of perphenazine (Fentazin) for periods ranging from 8 to 30 months and the average duration was 16 months. Their condition had shown no significant recent changes: no E.C.T. had been used - at least since perphenazine had been introduced, and none had an early prospect of discharge." (p. 422)
"If a patient had required an anti-Parkinsonism agent, orphenadrine (Disipal), this was continued in the same dosage irrespective of the trial group to which the patient was allocated." (p. 422)
"No other psychotropic drugs were used during the trial which lasted for ten weeks from 4 October, 1961 until 12 December, 1961 (inclusive)." (p. 422)

No-treatment
Comparison name: No-medication
Description of intervention: not stated
Exposure/intensity to treatment: no medication
Duration of treatment: 10 weeks
Concomitant psychotherapy: not stated – but inpatient

Whittaker 1963 *(Continued)*

Concomitant pharmacotherapy: "They had all been receiving maintenance doses of perphenazine (Fentazin) for periods ranging from 8 to 30 months and the average duration was 16 months. Their condition had shown no significant recent changes: no E.C.T. had been used - at least since perphenazine had been introduced, and none had an early prospect of discharge." (p. 422)
"If a patient had required an anti-Parkinsonism agent, orphenadrine (Disipal), this was continued in the same dosage irrespective of the trial group to which the patient was allocated." (p. 422)
"No other psychotropic drugs were used during the trial which lasted for ten weeks from 4 October, 1961 until 12 December, 1961 (inclusive)." (p. 422)

Outcomes

Beneficial effect

- **Hierarchy** : usable data, global scale
- **Outcome chosen:** combined of major and minor relapses

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. In the present trial 6 per cent. of those taken off perphenazine had not restarted the drug at ten weeks. If the minor relapses or an equivalent number of patients had been put on to the active drug at the end of the trial, or if one assumes arbitrarily that this would have taken place by the end of the twelfth week, then only 36 per cent of those off perphenazine would have remained so at the end of the three month period

Key limitations from study authors

1. Only one patient commented upon a slight difference in taste of his medicine. Considering that 13 changed from Fentazin elixir to Placebo elixir, this may be taken as a tribute to the supplier's ability to make a satisfactory simulation of the active preparation
2. Only 6 of the major relapses were apparent at interview at their onset. This emphasizes the need for close co-operation with nursing and artisan staff in such studies

Other notes from review authors

1. None

Conflicts of interest: none found

J**udgement:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "Each patient was arbitrarily allocated to one of three groups of I 3 by the hospital pharmacist." (p. 423)
Allocation concealment	Yes	Quote: "Individual bottles were used and, as far as the latter two groups were concerned, the trial was blind in that only the pharmacist knew which bottles were active." (p. 423)
Blinding of outcome assessors	No	Quote: "Within the ten weeks of the trial, eleven patients were considered to have relapsed into their previous pattern (…) This was done by the doctor (C.B.W) acting upon his own observations, or upon those of the charge nurses." (p. 424)
Blinding of participants and personnel	No	Not possible to blind placebo and no-treatment

Whittaker 1963 *(Continued)*

Incomplete outcome data	Unclear	Attrition not clear. Outcome were relapses
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Wilson 1980

Study characteristics	
Methods	**Parallel randomised trial with three arms** 1. Pharmacological placebo: placebo implant 2. No-treatment: no-operation control 3. Active treatment: disulfiram implant: **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 1 day (one operation – 10 to 15 minutes). **Duration of participation (trial + follow-up):** one operation + 48 weeks of follow-up **Setting:** outpatient **Purpose of trial:** "In a recent study, we compared the effect of disulfiram and placebo implants in alcoholic patients. The present study is a partial replication using larger samples and including no-operation controls and pseudocontrols (…) We designed the present study to evaluate the effectiveness of disulfiram implantation by comparing the duration of abstinence in patients receiving disulfiram or placebo implants and patient knowingly receiving neither type of implant" (p. 429) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** 100 **Number of participants included:** 90 **Number of participants followed-up at post-treatment:** 90 **Number of participants randomly assigned to:** • Pharmacological placebo n = 40 • No-treatment: n = 10 • Active treatment: n = 40 **Number of withdrawals:** not stated **Diagnosis:** substance use disorder (alcoholic) **Diagnostic manual:** not stated **Means of assessment:** not stated **Comorbidity:** not stated **Age:** 36.1 mean years (SD = 4.9) **IQ:** not stated **Sex:** 11% female **Ethnicity:** not stated

Wilson 1980 *(Continued)*

Country: Canada

Inclusion criteria

1. Alchoholics

Exclusion criteria

1. Not stated

Comparisons	**Pharmacological placebo** **Treatment name** : Placebo implant **Description of intervention:** "The disulfiram and placebo implant operations were carried out as in our previous study." (p. 430) From Wilson 1978 : "The operations were carried out in an outpatient operating room using standard aseptic surgical technique (…) For patients implanted with placebo, 0.5 cc of physiological saline was deposited in the tract using a 10-cc syringe and 19-gauge needle. The procedure was repeated to a total of 8 insertions in 8 separate tracts radiating from the central incision." (p. 812) **Individual or group treatment:** individual **Exposure/intensity to treatment:** one day one operation **Duration of treatment:** one operation – 10 to 15 minutes **Concomitant psychotherapy** : not stated **Concomitant pharmacotherapy:** not stated **No-treatment** **Comparison name** : No-operation control **Description of intervention** : "No-operation control patients were told that an implant did not appear to be appropriate for them at this time and were asked to participate as though they had received an implant (…)" (p. 4309 **Exposure/intensity to treatment:** no treatment **Duration treatment:** one day (no treatment) **Concomitant psychotherapy** : not stated **Concomitant pharmacotherapy:** not stated
Outcomes	**Beneficial effect** • **Hierarchy** : usable data, observer-reported • **Outcome chosen:** absence from alcohol **Adverse events** • Incidence of disulfiram-ethanol reaction (DER) in disulfiram and a few placebo participant, when patients resumed drinking (dyspnoea, hypertension, nausea, patchy erythema, pyrexia, tachycardia and vomiting)
Notes	**Key conclusion from study authors** 1. "The finding that the disulfiram and placebo patients remained abstinent significantly longer than the no-operation control and pseudocontrol patients confirms the efficacy of the procedure reported previously." (p. 433) 2. "The suggestion that the effectiveness of disulfiram implantation lies partly in its psychological component is supported by the present results (…) Patients remained abstinent because they feared the consequences of drinking."(p. 433) 3. "In our previous studies there were no significant differences between disulfiram and placebo patients on the time between intervention and relapse (…) However, this explanation is not supported by the present data, and since the samples are larger (40 vs 10), the results demand more credence" (p. 434) 4. "The inability to detect circulating disulfiram in the blood of disulfiram implant patients presents a problem for the interpretation of the relative superiority of disulfiram over a placebo implant in terms of patients' postintervention abstinence." (p. 434) **Key limitations from study authors**

Wilson 1980 *(Continued)*

1. None mentioned

Other notes from review authors

1. None

Conflicts of interest none found

J**udgement:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "The 90 patients who met the selection criteria were block randomly assigned to a disulfiram implant (N=40), placebo implant (N=40), or no-operation control (N=10) condition." (p. 430)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	No	Patient-reported outcomes
Blinding of participants and personnel	No	Not possible to blind placebo and no-operation
Incomplete outcome data	Unclear	Attrition unclear
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Wolitzky 2009

Study characteristics

Methods	**Parallel randomised trial with four arms** 1. Physical placebo: A credible placebo control consisting of pulsed audio-photic stimulation (APS) 2. Wait-list 3. Active treatment 1: exposure augmented with oppositional actions (E+OA) 4. Active treatment 2: in vivo exposure only (EO) **Sample calculation:** yes **Cluster randomised:** no **Duration of trial (baseline to post):** 1 treatment: **Duration of participation (trial + follow-up):** 1 treatment + 1 month follow-up **Setting:** outpatient **Purpose of trial:** "The principal aim of the current study was to investigate the hypothesized facilitative effects of having acrophobic individuals engage in actions (i.e., running toward the rail of a balcony, spinning in place in the phobic situation to induce dizziness, holding one's hands behind one's back while looking over the edge of a railing) that are in direct opposition to their threat-relevant fear action tendencies." (p. 59) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:**: not stated

Wolitzky 2009 *(Continued)*

Number of participants included: 89

Number of participants followed-up at post-treatment: 88

Number of participants randomly assigned to:

- Physical placebo: n = 25
- Wait-list: n = 11
- Active treatment 1: n = 28
- Active treatment 2: n = 24

Number of withdrawals : n=1

- Physical placebo: not stated
- Wait-list: not stated
- Active treatment 1: not stated
- Active treatment 2: not stated

Diagnosis: specific phobia (acrophobia)
Diagnostic manual:: Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV)
Means of assessment: Composite International Diagnostic Interview (CIDI-Auto).
Comorbidity: not stated
Age: 20.08 mean years (range =18 to 64)
IQ: not stated, but 82% were university students
Sex: 69% female
Ethnicity: 49% Caucasians, 15% Hispanic/Latino, 12.5% African-American, 19% Asian-American, 3% Native American, and 2% multi-racial or other race.
Country: USA

Inclusion criteria

1. Meet DSM-IV criteria for specific phobia (acrophobia) based on the Composite International Diagnostic Interview
2. Display at least moderate fear (50 or higher on a 0 = no fear to 100 = extreme fear rating scale) on two consecutive behavioral approach tests.
3. Report moderate fear or avoidance on a modified version of the Acrophobia Questionnaire

Exclusion criteria

1. Presented with a history of seizures due to the slight increased risk of seizure for those participants randomized to the pulsed audio-photic stimulation placebo treatment
2. Presented with a medical condition that precluded them from safely climbing stairs

Comparisons

Physical placebo

Treatment name: Credible placebo, pulsed audio-photic stimulation
Description of intervention:"APS (Seiver, Mind Alive, Inc.) is typically used by health professionals to induce relaxation. The APS device resembles a small soundboard and is about the size of a MP3 player. The device consist of a headset, which emits controllable pulsing sounds like a metronome, and a plastic mask, which produces orange lights at controllable rates (…) Participants were told that introducing these lights and sounds would relax them by inducing alpha waves in the brain, which are typically associated with relaxation and meditation" (p. 62)
Individual or group treatment: individual
Exposure/intensity to treatment :"Treatment was delivered in six, 6-min trials, for a total of one 36-min session." (p. 61-2)
Duration of treatment: one treatment (one day)
Concomitant psychotherapy: not stated – but probably not as they were not in or seeking treatment
Concomitant pharmacotherapy\; not stated – but probably not as they were not in or seeking treatment

Wait-list

Wolitzky 2009 *(Continued)*

Comparison name: Wait list control
Description of intervention: "Participants received no treatment until after the posttreatment assessment was completed. At that time, they received on 36-min session of in vivo exposure treatment. In order to provide treatment as soon as possible, participants in this condition did not complete a follow-up assessment." (p. 63)
Exposure/intensity to treatment: no treatment.
Duration treatment: one day waiting
Concomitant psychotherapy: not stated – but probably not as they were not in or seeking treatment
Concomitant pharmacotherapy: not stated – but probably not as they were not in or seeking treatment

Outcomes

Beneficial effect

- **Hierarchy** : usable data, observer-reported
- **Outcome chosen:** BAT – subscale BAT2 – Raw Heart rate

Adverse events

- No data reported on adverse events

Notes

Key conclusion from study authors

1. "A relatively consistent pattern of findings emerged with respect to treatment outcome. Participants receiving exposure while enacting threat- relevant oppositional action tendencies showed significantly greater improvement at the end of treatment and at a 1-month follow-up assessment relative to participants receiving exposure only." (p. 66)
2. Whereas those receiving exposure only did not show significantly greater improvement than placebo on out two primary outcome measures assessed in the untrained context (i.e., generalization probe), treatment gains in the untrained context for participants in the E + OA group were marked and significantly greater than those observed in the other three treatment groups." (p. 66)

Key limitations from study authors

1. "First, our two behavioral approach tests used in the behavioral assessment of acrophobia were only modestly challenging." (p. 68)
2. "Second, our experimental design does not disentangle the effects of enacting oppositional actions from the effects of safety behavior fading." (p. 68)
3. "Caution is warranted when attempting to generalize the findings to clinical practice. The exposure treatment used in this study is not representative of exposure therapy as it is typically conducted in the real world." (p. 69)
4. "Fourth, although our study participants did reveal significant severity on several independent measures of acrophobic fear, most (82%) were not seeking treatment and thus our findings cannot be generalized to a treatment-seeking clinical sample." (p. 69)

Other notes from review authors

1. None

Conflicts of interest: none found

Judgement: yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: "89 participants were randomized to one of the four treatment conditions (14 participants declined to participate in the treatment phase of the study)." (p. 60)

Wolitzky 2009 *(Continued)*

		"One-way ANOVAs were used to assess whether randomization was successful in achieving equivalent groups at baseline. As seen in Table 1, the four experimental groups did not differ on the measures at baseline with the exception of BAT-2 peak fear ratings (…)" (p. 63)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Yes	Quote: "During each treatment trial, a trained undergraduate research assistant blind to the study hypotheses coded the presence or absence of nine possible oppositional actions." (p. 61)
Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	Yes	Quote: One participant dropped out during treatment. Thus, 88 participants completed treatment, which was conducted one week after the pretreatment assessment." (p. 60) Attrition < 15% (1.1%).
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

Wollersheim 1991

Study characteristics	
Methods	**Parallel randomised trial with four arms** 1. Psychological placebo: supporting therapy 2. Wait-list: delayed treatment 3. Active treatment 1: bibliotherapy 4. Active treatment 2: coping **Sample calculation:** not stated **Cluster randomised:** no **Duration of trial (baseline to post):** 11 weeks **Duration of participation (trial + follow-up):** 11 weeks + 6 months follow-up **Setting:** outpatient **Purpose of trial:** "To evaluate differentially the effectiveness of cognitive-behavioral group treatment the present investigation included a supportive group approach, a bibliotherapeutic approach, and a delayed treatment condition." (p. 496-7) **Open/closed placebo:** closed placebo
Data	**Number of participants screened:** 99 **Number of participants included:** 32 **Number of participants followed-up at post-treatment:** 25 **Number of participants randomly assigned to:**

Wollersheim 1991 *(Continued)*

- Psychological placebo: n = 8
- Wait-list: n = 8
- Active treatment 1: n = 8
- Active treatment 2: n = 8

Number of withdrawals : n =7

- Psychological placebo: n = 2
- Wait-list: n = 0
- Active treatment 1: n = 3
- Active treatment 2: n = 2

Diagnosis: depression

Diagnostic manual: Diagnostic and Statistical Manual of Mental Disorders, 3rd Edition (DSM-III)

Means of assessment: Clinical interview

Comorbidity: not stated

Age: 39.4 mean years (range= 22 to 68 years)

IQ: not stated

Sex: 71.9% female

Ethnicity: not stated

Country: USA

Inclusion criteria

1. Depression diagnosis

Exclusion criteria

1. Current medication or psychological treatment
2. Evidence of a thought disorder, bipolar depression, organicity, or melancholia based on history, clinical interview, and assessment data
3. Suicide risk in the moderate to high range
4. recent history of alcoholism or drug abuse

Comparisons

Psychological placebo

Treatment name: Supportive therapy

Description of intervention: "The supportive treatment was designed to parallel the group format used in coping therapy, but supportive therapy techniques were much less directive and did not emphasize therapist-generated problem-solving tactics (p. 497)

Individual or group treatment: Group

Exposure/intensity to treatment: 10 weekly treatment sessions of 2 hours

Duration of treatment: 11 weeks

Concomitant psychotherapy: not stated

Concomitant pharmacotherapy: medication not allowed

Wait-list

Comparison name: Delayed treatment group

Wollersheim 1991 *(Continued)*

	Description of intervention: "Patients assigned to this condition were informed that their treatment would begin 11 weeks after their initial assessment session. Members of the delayed treatment group were asked to call if they needed more immediate treatment and told that in such cases treatment would be arranged but they would then not be involved in the treatment research program. In addition to receiving individual assessment sessions with a clinician, they were also given reassurance regarding prognosis and hope for improvement and were informed that they would receive the best approach available at the start of their treatment. After post-treatment assessment, these patients were given group coping treatment." (p. 498) **Exposure/intensity to treatment:** waiting for treatment **Duration of treatment:** 11 weeks **Concomitant psychotherapy:** not stated **Concomitant pharmacotherapy:** medication not allowed
Outcomes	**Beneficial effect** • **Hierarchy:** observer-reported • **Outcome chosen:** Clinician Rating of Depression (CR) **Adverse events** • Spontaneous reporting
Notes	**Key conclusion from study authors** 1. Statistical analyses and quantitative analyses of clinical significance demonstrated improvement for all conditions and maintenance of therapeutic gains at follow-up **Key limitations from study authors** 1. Small sample size is a distinct problem that limits discrimination across groups 2. Another limitation is that group treatment consisted of only 10 2-hour sessions 3. The 20 hours of treatment received by patients in this study is unlikely to be exceeded by most outpatients 4. Thus, further research investigating the interaction of length of treatment, therapy format, and outcome success is warranted **Other notes from review authors** 1. None **Conflicts of interest:** none found **Judgement:** yes

Risk of bias

Item	Authors' judgement	Support for judgement
Random sequence generation	Unclear	Quote: ”All 32 patients were blocked for age and sex and then randomly assigned to one of four treatment groups: coping therapy, supportive therapy, bibliotherapy, or a delayed treatment group.” (p. 498)
Allocation concealment	Unclear	No information
Blinding of outcome assessors	Unclear	No information

Wollersheim 1991 *(Continued)*

Blinding of participants and personnel	No	Not possible to blind placebo and wait-list
Incomplete outcome data	No	Quote: "A total of seven patients (two each from coping and supportive group treatments, three from bibliotherapy) could not be located at follow up" (p. 499) Attrition >15% (21.9%). No ITT.
Selective outcome reported	Unclear	No protocol found
Other sources of bias	Yes	No other sources found

CBT: Cognitive behavioral therapy;**CPAP:** continuous positive airway pressure; **GAF:** (Global Assessment of Functioning; **IQ:** intelligence quotient; **ITT:** intention-to-treat;**MDD:** Major depressive disorder; **NCPAP:** (nasal continuous positive airway pressure; **SD:** standard deviation;**TAU:** treatmnt as usual; **WISC:** Wechsler Intelligence Scale for children

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Adelman 1990	Not a randomised clinical trial
Altenhofer 2007	Not a randomised clinical trial
Andreasson 1990	Not a randomised clinical trial
Banasiak 2005	Did not fulfil inclusion criteria for a usual care
Bass 2016	Did not fulfil inclusion criteria for a usual care
de Jongh 1995	Did not fulfil inclusion criteria for a usual care
Dolhanty 2005	Did not fulfil inclusion criteria for a physical placebo
Drozdek 2010	Not a randomised clinical trial
Endicott 1964	Not a randomised clinical trial
Ertl 2011	Did not fulfil inclusion criteria for a usual care
Gitlin 2013	Did not fulfil inclusion criteria for a usual care
Gournay 1994	Wrong patient population (family distress)
Haertzen 1964	Not a randomised clinical trial, and wrong population (post-addicts)
Hartelius 1968	Not a randomised clinical trial
Hinsberger 2017	Some participants were not randomised but placed on a second wait list called "no camp".
Keltner 1976	Not a randomised clinical trial
Laessle 1988	Not a randomised clinical trial

Study	Reason for exclusion
Litrownik 1978	Did not fulfil inclusion criteria for a physical placebo
Mak 2008	Did not fulfil inclusion criteria for a physical placebo
McCall 2005	No wait-list or no-treatment group included
McNeil 1991	No placebo group included
Otteson 1979	Did not fulfil inclusion criteria for a usual care
Propst 1992	Did not fulfil inclusion criteria for a usual care
Sackley 2018	Did not fulfil inclusion criteria for a usual care
Salt 2002	Not a randomised clinical trial
Schultz 2002	Wrong patient population quote: ("to be interested in stress and depression"
Sheikh 1986	Did not fulfil inclusion criteria for a physical placebo
Steiner 2011	Did not fulfil inclusion criteria for a usual care
Timko 1995	Not a randomised clinical trial
Tomporowski 1984	Pools no-treatment, usual care and placebo in one group.
Torkan 2014	Did not fulfil inclusion criteria for a psychological placebo
Weber 1975	Not a randomised clinical trial
Wilson 2002	Did not fulfil inclusion criteria for a usual care

Characteristics of studies awaiting classification *[ordered by study ID]*

Bommert 1978

Methods	Randomised controlled trial
Data	Not available **Condition: Learning Disabled students**
Comparisons	**Four-armed randomised trial** 1. Verbal group Therapy 2. Client-centred behavior modification 3. Combination of the two 4. No-treatment
Outcomes	Not stated
Notes	Not able to locate the full report

Brandes 2010

Methods	Not stated
Data	Not stated
Comparisons	Not stated
Outcomes	Not stated
Notes	Only the title located. Not possible to locate the full report. Authors did not respond to correspondence

McLachlan 1993

Methods	Randomsied single-blind, placebo-controlled clinical trial
Data	Not available **Condition** : Alzheimer's disease
Comparisons	**Three-armed randomised trial** 1. Desferrioxamine 2. Placebo (oral lecithin) 3. No-treatment
Outcomes	**Outcome** • Videotaped home behavior (VHB/) assessment instruments
Notes	Not able to locate the full report

Newton-Cross 2017

Methods	Not stated
Data	Not stated
Comparisons	Not stated
Outcomes	Not stated
Notes	Not able to locate the full report. Not able to find contact information of the authors

Schwarzler 1999

Methods	Not stated
Data	Not stated

Schwarzler 1999 *(Continued)*

Comparisons	Not stated
Outcomes	Not stated
Notes	Not able to locate the full report

Trianes Torres 1991

Methods	Not stated
Data	Not available **Condition** : Inhibited and impulsive preschool children
Comparisons	**Three-armed randomised trial** 1. Maximum Intervention 2. Minimum Intervention 3. Control
Outcomes	**Outcome** • Interpersonal cognitive problem-solving skills measure
Notes	Not able to locate the full report. Not able to find contact information of the authors

Characteristics of ongoing studies *[ordered by study ID]*

Heitman 2017

Study name	Internet-based attentional bias modification training as add-on to regular treatment in alcohol and cannabis dependent outpatients: a study protocol of a randomised control trial
Methods	Randomised controlled trial
Data	Not available **Condition:** substance dependent (alcohol/cannabis)
Comparisons	**Two-armed randomised trial** 1. internet-based ABM (iABM) intervention + treatment as usual (TAU) 2. treatment as usual (TAU)
Outcomes	**Primary outcomes** 1. substance use, craving, and rates of relapse 2. changes in attentional bias will be measured to investigate whether changes in primary outcome measures can be attributed to the modification of attentional bias' **Secondary outcomes** 1. indices of cost-effectiveness 2. secondary physical and psychological complaints (depression, anxiety, and stress)

Heitman 2017 *(Continued)*

Starting date	18/05/2015
Contact information	Janika Heitmann email: J.Heitmann@vnn.nl
Notes	Quote: "We are currently processing the data, but unfortunately I do not yet have data to share (in case this is still of interest in a couple of moths it might be possible to share some results with you). Further, after reading your protocol I doubt whether our study is eligible for your review. That is, all our participants received psychological treatment (cognitive behavioural therapy), and only for the add-on treatment we had three groups from which one is a group that received a placebo treatment and one was a treatment-as-usual group (only received cognitive behavioural therapy and no add-on)." Heitmann 2017 (pers comm)

ISRCTN21392756

Study name	Standardised stress management mental health training: does it have a beneficial effect?
Methods	Single-centre open randomised controlled trial
Data	Not available **Condition:** adjustment disorder or mild depression
Comparisons	**Two-armed randomised trial** 1. The standardised Stress Management Mental Health Training 2. Wait-list
Outcomes	**Primary outcome** 1. General Health Questionnaire-28 score recorded as a continuous variable. All outcomes will be assessed at the start of the trial (visit 2), 6 weeks (visit 3) and 12 weeks (visit 4) **Secondary outcomes** 1. Beck Depression Inventory (BDI-II) 2. Beck Anxiety Inventory (BAI) 3. Clinical Global Impression (CGI) All outcomes will be assessed at the start of the trial (visit 2), 6 weeks (visit 3) and 12 weeks (visit 4)
Starting date	06/09/2010
Contact information	Dr Trevor Hicks email: trevor@hicks65.freeserve.co.uk
Notes	No response to author correspondence

ISRCTN35717198

Study name	Cognitive training with non-invasive brain stimulation to treat binge eating disorder
Methods	Randomised controlled trial

ISRCTN35717198 *(Continued)*

Data	Not available **Condition** : Binge Eating Disorder
Comparisons	**Three-armed randomised trial** 1. Approach bias modification training + transcranial direct current stimulation (tDCS) (ABM + real tDCS) 2. Approach bias modification training + sham transcranial direct current stimulation (tDCS)(ABM + sham tDCS) 3. inactive wait-list control group
Outcomes	**Primary outcomes** The primary outcome for this proof-of-concept feasibility study is to establish the feasibility of adding concurrent [ABM + tDCS] to treatment as usual in a binge eating disorder patient group and acquire key information to inform the development of a large-scale randomised sham-controlled trial (RCT). In line with recommendations of Eldridge et al. (2016). The primary outcomes of the proposed feasibility study are to: 1. establish the feasibility of conducting a large-scale RCT of [ABM and real/sham tDCS] with a wait-list control in patients with BED by assessing recruitment, attendance, and retention rates 2. determine the feasibility of administering both ABM and tDCS simultaneously; 3. determine the best instruments for measuring outcomes in a full trial by examining the quality, completeness, and variability in the data; 4. estimate the treatment effect sizes and standard deviations for outcome measures to inform the sample size calculation for a large-scale RCT; 5. evaluate whether the treatment is operating as it is designed by analysing process measures, such as within session visual analogue scales; 6. determine whether patients with BED view [ABM and real/sham tDCS] as acceptable and credible 7. Obtain information about patients' willingness to undergo random allocation to ABM paired with real or sham tDCS, and to the wait-list control. **Secondary outcomes** 1. Scores in the EDE-Q, specifically indicators of change in the frequency of objective binge eating episodes from baseline, post-assessment and the 4-week follow-up time points. 2. Differences between pre [ABM and real/sham tDCS] VAS scores and post [ABM and real/sham tDCS] VAS scores (per each of the 6 treatment sessions and cumulatively over sessions). 3. Changes in scores on the questionnaires and performance in tasks regarding eating disorder and general psychopathology, food, hunger and craving, impulsivity, delayed gratification and inhibitory control, measured at baseline, post-assessment and the 4-week follow-up.
Starting date	03/05/2018
Contact information	Gemma Gordon Email: gemma.gordon@kcl.ac.uk
Notes	No response to author correspondence

NCT00044629

Study name	Combined behavioral/pharmacologicaltTherapy for insomnia

NCT00044629 *(Continued)*

Methods	Randomised clinical trial
Data	Not available **Condition:** insomnia
Comparisons	**Three-armed randomised trial** 1. Behavioral: Cognitive-Behavioral Therapy for Insomnia + placebo 2. Drug: zolpidem tartrate (Ambien) 3. Drug: Placebo
Outcomes	Not stated
Starting date	05/09/2002
Contact information	Duke University
Notes	Not able to find any contact information

DATA AND ANALYSES

Comparison 1. Wait-list versus no-treatment for people with mental health disorder

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 Wait-list versus no-treatment for people with mental health disorder	1		Other data	No numeric data

Analysis 1.1. Comparison 1: Wait-list versus no-treatment for people with mental health disorder, Outcome 1: Wait-list versus no-treatment for people with mental health disorder

Wait-list versus no-treatment for people with mental health disorder

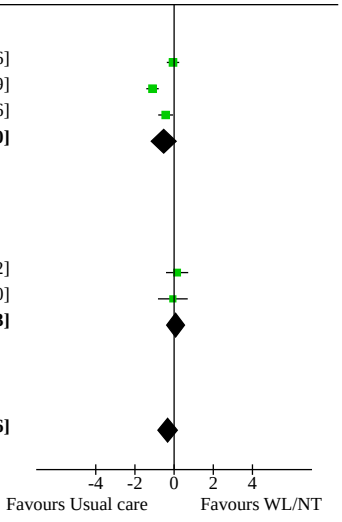
Study	Heading 1
Howlin 2007	Only one cluster-randomised trial compared a wait-list intervention to a no-treatment intervention was included (Howlin 2007). However, no usable data was provided in the full report, and the authors did engage in correspondence. 84 elementary school children with a autism spectrum disorder was randomised to either, i) immediate treatment, ii) delayed treatment (wait-list), and iii) no-treatment. Conclusion were that Picture Exchange Communication System (PECS) training indicated modest effectiveness for children with autism spectrum disorder. In general there were no differences on across outcome measures between the wait-list and no-treatment intervention groups.

Comparison 2. Usual care versus wait-list or no-treatment for all mental health disorders

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.1 Usual care compared with wait-list/no-treatment for continuous data	5	523	Std. Mean Difference (IV, Random, 95% CI)	-0.33 [-0.83, 0.16]
2.1.1 Usual care compared with wait-list	3	443	Std. Mean Difference (IV, Random, 95% CI)	-0.53 [-1.17, 0.10]
2.1.2 Usual care compared with no-treatment	2	80	Std. Mean Difference (IV, Random, 95% CI)	0.08 [-0.38, 0.53]
2.2 Usual care compared with wait-list/no-treatment for dichotomous data	2	260	Risk Ratio (M-H, Random, 95% CI)	0.93 [0.34, 2.57]
2.2.1 Usual care compared with wait-list	2	260	Risk Ratio (M-H, Random, 95% CI)	0.93 [0.34, 2.57]
2.2.2 Usual care compared with no-treatment	0	0	Risk Ratio (M-H, Random, 95% CI)	Not estimable

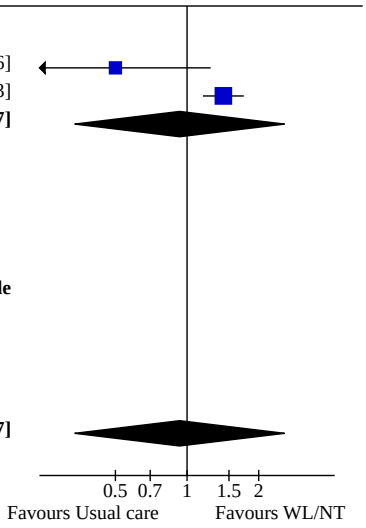
Analysis 2.1. Comparison 2: Usual care versus wait-list or no-treatment for all mental health disorders, Outcome 1: Usual care compared with wait-list/no-treatment for continuous data

	Usual care			Wait-list/no-treatment				Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
2.1.1 Usual care compared with wait-list									
Glogowska 2000	-4.9	1.7	71	-4.8	1.8	84	22.3%	-0.06 [-0.37 , 0.26]	
Rapee 2006	3.4	2.6	90	5.8	1.6	87	22.3%	-1.10 [-1.42 , -0.79]	
Rapee 2007	28.63	15.01	59	35.337	15.83	52	21.4%	-0.43 [-0.81 , -0.06]	
Subtotal (95% CI)			**220**			**223**	**65.9%**	**-0.53 [-1.17 , 0.10]**	
Heterogeneity: Tau² = 0.28; Chi² = 21.42, df = 2 (P < 0.0001); I² = 91%									
Test for overall effect: Z = 1.64 (P = 0.10)									
2.1.2 Usual care compared with no-treatment									
Crisp 1991	-5.5	0.6	30	-5.6	0.7	20	18.5%	0.15 [-0.41 , 0.72]	
Teri 1997	13.9	4.14	10	14.1	2.32	20	15.5%	-0.06 [-0.82 , 0.70]	
Subtotal (95% CI)			**40**			**40**	**34.1%**	**0.08 [-0.38 , 0.53]**	
Heterogeneity: Tau² = 0.00; Chi² = 0.20, df = 1 (P = 0.65); I² = 0%									
Test for overall effect: Z = 0.33 (P = 0.74)									
Total (95% CI)			**260**			**263**	**100.0%**	**-0.33 [-0.83 , 0.16]**	
Heterogeneity: Tau² = 0.26; Chi² = 27.62, df = 4 (P < 0.0001); I² = 86%									
Test for overall effect: Z = 1.32 (P = 0.19)									
Test for subgroup differences: Chi² = 2.33, df = 1 (P = 0.13), I² = 57.1%									



Analysis 2.2. Comparison 2: Usual care versus wait-list or no-treatment for all mental health disorders, Outcome 2: Usual care compared with wait-list/no-treatment for dichotomous data

Study or Subgroup	Usual care Events	Usual care Total	Wait-list/no-treatment Events	Wait-list/no-treatment Total	Weight	Risk Ratio M-H, Random, 95% CI
2.2.1 Usual care compared with wait-list						
Milby 1980	4	13	8	13	40.5%	0.50 [0.20 , 1.26]
Robin 1976	88	116	63	118	59.5%	1.42 [1.17 , 1.73]
Subtotal (95% CI)		**129**		**131**	**100.0%**	**0.93 [0.34 , 2.57]**
Total events:	92		71			
Heterogeneity: Tau² = 0.44; Chi² = 4.81, df = 1 (P = 0.03); I² = 79%						
Test for overall effect: Z = 0.14 (P = 0.89)						
2.2.2 Usual care compared with no-treatment						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
Total (95% CI)		**129**		**131**	**100.0%**	**0.93 [0.34 , 2.57]**
Total events:	92		71			
Heterogeneity: Tau² = 0.44; Chi² = 4.81, df = 1 (P = 0.03); I² = 79%						
Test for overall effect: Z = 0.14 (P = 0.89)						
Test for subgroup differences: Not applicable						



Comparison 6. All placebos versus waitl-list or no-treatment for people with mental health disorders

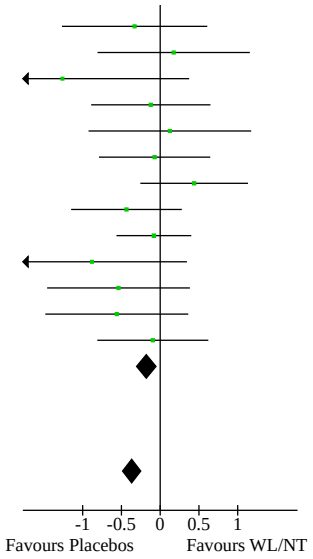
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
6.1 All placebos compared with wait-list/no-treatment for continuous data	65	2446	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.49, -0.25]
6.1.1 All placebos compared with wait-list	31	1410	Std. Mean Difference (IV, Random, 95% CI)	-0.55 [-0.76, -0.35]
6.1.2 All placebos compared with no-treatment	34	1036	Std. Mean Difference (IV, Random, 95% CI)	-0.18 [-0.30, -0.05]
6.2 All placebos compared with wait-list/no-treatment for dichotomous data	9	385	Risk Ratio (M-H, Random, 95% CI)	1.05 [0.74, 1.48]
6.2.1 All placebos compared with wait-list	0	0	Risk Ratio (M-H, Random, 95% CI)	Not estimable
6.2.2 All placebos compared with no-treatment	9	385	Risk Ratio (M-H, Random, 95% CI)	1.05 [0.74, 1.48]

Analysis 6.1. Comparison 6: All placebos versus waitl-list or no-treatment for people with mental health disorders, Outcome 1: All placebos compared with wait-list/no-treatment for continuous data

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
6.1.1 All placebos compared with wait-list								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	1.3%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	2.8%	-0.72 [-1.15 , -0.29]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	1.1%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	1.2%	-1.71 [-2.60 , -0.82]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	1.2%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	1.3%	-0.25 [-1.09 , 0.59]
Bramston 1985	-78	14.21	12	-72	13.82	12	1.4%	-0.41 [-1.22 , 0.40]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	2.3%	-0.64 [-1.16 , -0.12]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	1.5%	-0.65 [-1.43 , 0.13]
Foa 1991	18.09	7.13	11	19.5	7.18	10	1.3%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	3.0%	-1.10 [-1.48 , -0.71]
Fuchs 1977	14.3	7	10	21.4	7	10	1.1%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	1.5%	-0.80 [-1.59 , -0.01]
Kelley 2012	-1.64	4.52	11	0.67	4	9	1.2%	-0.51 [-1.41 , 0.38]
Kilmann 1987	69	27	4	90	14	4	0.5%	-0.85 [-2.35 , 0.65]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	2.7%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	1.6%	-0.48 [-1.21 , 0.25]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	1.1%	-0.71 [-1.67 , 0.25]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	1.4%	-1.33 [-2.16 , -0.49]
Powers 2004	46	20.09	12	41.6	13.15	15	1.5%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	1.4%	-0.10 [-0.90 , 0.70]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	3.9%	-0.11 [-0.31 , 0.09]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	2.3%	-0.24 [-0.78 , 0.30]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	1.6%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	1.3%	-1.10 [-1.97 , -0.23]
Turner 1979	-4.1	2.644	10	-4	2.449	10	1.3%	-0.04 [-0.91 , 0.84]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	1.1%	-0.52 [-1.46 , 0.43]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	1.7%	-0.11 [-0.82 , 0.60]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	1.1%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**755**			**655**	**49.1%**	**-0.55 [-0.76 , -0.35]**
Heterogeneity: Tau² = 0.17; Chi² = 79.17, df = 30 (P < 0.00001); I² = 62%								
Test for overall effect: Z = 5.40 (P < 0.00001)								
6.1.2 All placebos compared with no-treatment								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	2.5%	-0.45 [-0.93 , 0.03]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	2.2%	-0.09 [-0.66 , 0.48]
Borden 1986	29.7	17.5	10	32	26	10	1.3%	-0.10 [-0.98 , 0.78]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	2.0%	-0.11 [-0.72 , 0.49]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	1.2%	-0.09 [-1.01 , 0.84]
Carter 2003	3.8	1.2	28	3.8	1.3	29	2.3%	0.00 [-0.52 , 0.52]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	1.5%	-0.33 [-1.12 , 0.46]
Freire 2007	24.6	11	7	28.2	18	9	1.1%	-0.22 [-1.21 , 0.77]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	2.7%	-0.08 [-0.51 , 0.36]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	1.5%	-0.53 [-1.30 , 0.25]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	1.2%	-0.59 [-1.49 , 0.31]
Klein 1977	10.17	5.49	6	15	7.32	6	0.8%	-0.69 [-1.87 , 0.49]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	1.4%	-0.81 [-1.63 , 0.00]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	1.3%	-0.35 [-1.24 , 0.53]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	2.5%	0.03 [-0.46 , 0.52]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	1.3%	-0.14 [-1.00 , 0.72]
Lick 1977	65.86	25.45	10	62.84	31.78	10	1.3%	0.10 [-0.78 , 0.98]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	1.3%	0.34 [-0.55 , 1.22]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	1.2%	-0.04 [-0.93 , 0.86]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	1.3%	-0.45 [-1.32 , 0.42]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	1.7%	-0.17 [-0.87 , 0.53]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	1.1%	-0.33 [-1.27 , 0.61]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.1%	0.17 [-0.81 , 1.16]

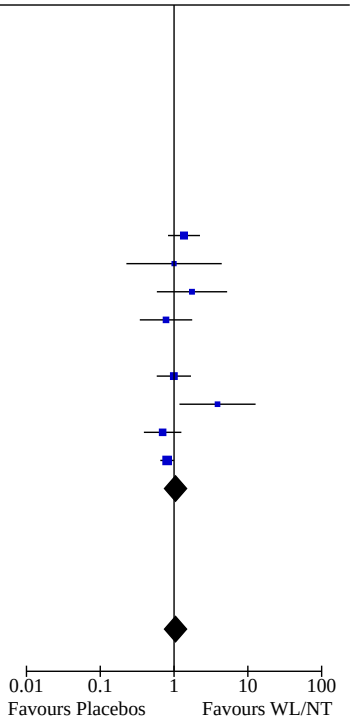
Analysis 6.1. (Continued)

Nandi 1976	53.2	11.2	10	57.5	13.8	8	1.1%	-0.33 [-1.27 , 0.61]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.1%	0.17 [-0.81 , 1.16]
Peck 1976	-12.5	4.25	4	-7	3.27	4	0.5%	-1.26 [-2.90 , 0.37]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	1.5%	-0.12 [-0.89 , 0.65]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.0%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	1.7%	-0.07 [-0.79 , 0.65]
Powers 2008b	64	20	18	51	37	15	1.7%	0.44 [-0.26 , 1.13]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	1.7%	-0.43 [-1.15 , 0.28]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	2.5%	-0.08 [-0.56 , 0.40]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	0.8%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	1.2%	-0.54 [-1.46 , 0.38]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	1.2%	-0.56 [-1.48 , 0.36]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	1.7%	-0.10 [-0.81 , 0.62]
Subtotal (95% CI)			**491**			**545**	**50.9%**	**-0.18 [-0.30 , -0.05]**
Heterogeneity: Tau² = 0.00; Chi² = 18.83, df = 33 (P = 0.98); I² = 0%								
Test for overall effect: Z = 2.81 (P = 0.005)								
Total (95% CI)			**1246**			**1200**	**100.0%**	**-0.37 [-0.49 , -0.25]**
Heterogeneity: Tau² = 0.08; Chi² = 108.28, df = 64 (P = 0.0005); I² = 41%								
Test for overall effect: Z = 6.16 (P < 0.00001)								
Test for subgroup differences: Chi² = 9.63, df = 1 (P = 0.002), I² = 89.6%								



Analysis 6.2. Comparison 6: All placebos versus waitl-list or no-treatment for people with mental health disorders, Outcome 2: All placebos compared with wait-list/no-treatment for dichotomous data

Study or Subgroup	All placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
6.2.1 All placebos compared with wait-list						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
6.2.2 All placebos compared with no-treatment						
Berg 1983	9	11	9	15	16.9%	1.36 [0.83 , 2.24]
Double 1993	3	22	3	22	4.4%	1.00 [0.23 , 4.42]
Klerman 1974a	7	25	4	25	7.1%	1.75 [0.58 , 5.24]
Klerman 1974b	7	25	9	25	10.5%	0.78 [0.34 , 1.76]
Mealiea 1971	0	9	0	10		Not estimable
Rabkin 1990	14	27	12	23	16.0%	0.99 [0.58 , 1.70]
Watzl 1988	11	34	3	36	6.3%	3.88 [1.18 , 12.73]
Whittaker 1963	7	13	10	13	14.9%	0.70 [0.39 , 1.26]
Wilson 1980	31	40	10	10	23.8%	0.80 [0.65 , 0.99]
Subtotal (95% CI)		**206**		**179**	**100.0%**	**1.05 [0.74 , 1.48]**
Total events:	89		60			
Heterogeneity: Tau² = 0.12; Chi² = 16.68, df = 7 (P = 0.02); I² = 58%						
Test for overall effect: Z = 0.26 (P = 0.79)						
Total (95% CI)		**206**		**179**	**100.0%**	**1.05 [0.74 , 1.48]**
Total events:	89		60			
Heterogeneity: Tau² = 0.12; Chi² = 16.68, df = 7 (P = 0.02); I² = 58%						
Test for overall effect: Z = 0.26 (P = 0.79)						
Test for subgroup differences: Not applicable						



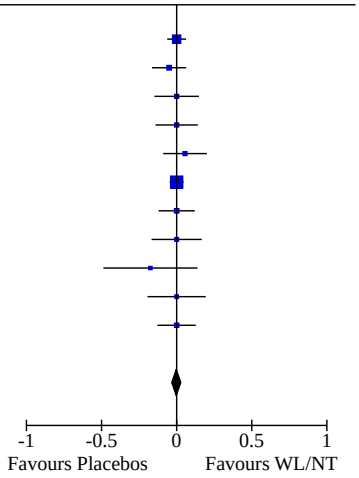
Comparison 7. Serious adverse events of placebos versus wait-list or no-treatment for people with mental health disorders

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
7.1 All placebos compared with wait-list/no-treatment for dichotomous data	11	517	Risk Difference (IV, Random, 95% CI)	-0.00 [-0.03, 0.03]
7.2 Psychological placebos compared with wait-list/no-treatment for dichotomous data	2	207	Risk Difference (IV, Random, 95% CI)	-0.01 [-0.07, 0.04]
7.3 Pharmacological placebos compared with wait-list/no-treatment for dichotomous data	4	125	Risk Difference (M-H, Random, 95% CI)	0.01 [-0.08, 0.09]
7.4 Physical placebos compared with wait-list/no-treatment for dichotomous data	5	185	Risk Difference (M-H, Random, 95% CI)	0.00 [-0.04, 0.04]

Analysis 7.1. Comparison 7: Serious adverse events of placebos versus wait-list or no-treatment for people with mental health disorders, Outcome 1: All placebos compared with wait-list/no-treatment for dichotomous data

Study or Subgroup	Placebos Events	Placebos Total	Wait-list/no-treatment Events	Wait-list/no-treatment Total	Weight	Risk Difference IV, Random, 95% CI
Ehlers 2014	0	30	0	30	22.5%	0.00 [-0.06 , 0.06]
Foa 2018	8	107	5	40	6.8%	-0.05 [-0.16 , 0.06]
Freire 2007	0	12	0	12	4.0%	0.00 [-0.15 , 0.15]
Karst 2007	0	19	0	10	4.5%	0.00 [-0.14 , 0.14]
Klosko 1990	1	18	0	16	4.2%	0.06 [-0.09 , 0.20]
Kwan 2017	0	41	0	40	40.1%	0.00 [-0.05 , 0.05]
Lacy 1990	0	15	0	15	6.1%	0.00 [-0.12 , 0.12]
Legrand 2016	0	11	0	10	3.2%	0.00 [-0.17 , 0.17]
McLachlan 1991	1	9	4	14	0.9%	-0.17 [-0.49 , 0.14]
Nandi 1976	0	10	0	8	2.4%	0.00 [-0.19 , 0.19]
Wilson 1980	0	40	0	10	5.4%	0.00 [-0.13 , 0.13]
Total (95% CI)		**312**		**205**	**100.0%**	**-0.00 [-0.03 , 0.03]**
Total events:	10		9			

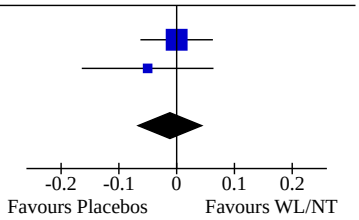
Heterogeneity: Tau² = 0.00; Chi² = 2.47, df = 10 (P = 0.99); I² = 0%
Test for overall effect: Z = 0.18 (P = 0.86)
Test for subgroup differences: Not applicable



Analysis 7.2. Comparison 7: Serious adverse events of placebos versus wait-list or no-treatment for people with mental health disorders, Outcome 2: Psychological placebos compared with wait-list/no-treatment for dichotomous data

Study or Subgroup	Placebos Events	Placebos Total	Wait-list/no-treatment Events	Wait-list/no-treatment Total	Weight	Risk Difference IV, Random, 95% CI
Ehlers 2014	0	30	0	30	76.8%	0.00 [-0.06 , 0.06]
Foa 2018	8	107	5	40	23.2%	-0.05 [-0.16 , 0.06]
Total (95% CI)		**137**		**70**	**100.0%**	**-0.01 [-0.07 , 0.04]**
Total events:	8		5			

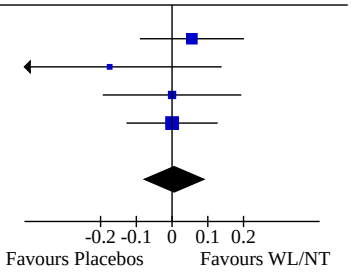
Heterogeneity: Tau² = 0.00; Chi² = 0.57, df = 1 (P = 0.45); I² = 0%
Test for overall effect: Z = 0.42 (P = 0.68)
Test for subgroup differences: Not applicable



Analysis 7.3. Comparison 7: Serious adverse events of placebos versus wait-list or no-treatment for people with mental health disorders, Outcome 3: Pharmacological placebos compared with wait-list/no-treatment for dichotomous data

Study or Subgroup	Placebos Events	Placebos Total	Wait-list/no-treatment Events	Wait-list/no-treatment Total	Weight	Risk Difference M-H, Random, 95% CI
Klosko 1990	1	18	0	16	32.5%	0.06 [-0.09 , 0.20]
McLachlan 1991	1	9	4	14	7.0%	-0.17 [-0.49 , 0.14]
Nandi 1976	0	10	0	8	18.3%	0.00 [-0.19 , 0.19]
Wilson 1980	0	40	0	10	42.2%	0.00 [-0.13 , 0.13]
Total (95% CI)		**77**		**48**	**100.0%**	**0.01 [-0.08 , 0.09]**
Total events:	2		4			

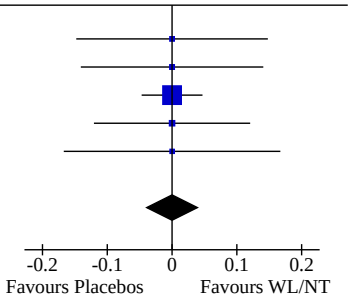
Heterogeneity: Tau² = 0.00; Chi² = 2.06, df = 3 (P = 0.56); I² = 0%
Test for overall effect: Z = 0.14 (P = 0.89)
Test for subgroup differences: Not applicable



Analysis 7.4. Comparison 7: Serious adverse events of placebos versus wait-list or no-treatment for people with mental health disorders, Outcome 4: Physical placebos compared with wait-list/no-treatment for dichotomous data

Study or Subgroup	Placebos Events	Placebos Total	Wait-list/no-treatment Events	Wait-list/no-treatment Total	Weight	Risk Difference M-H, Random, 95% CI
Freire 2007	0	12	0	12	7.0%	0.00 [-0.15 , 0.15]
Karst 2007	0	19	0	10	7.7%	0.00 [-0.14 , 0.14]
Kwan 2017	0	41	0	40	69.3%	0.00 [-0.05 , 0.05]
Lacy 1990	0	15	0	15	10.5%	0.00 [-0.12 , 0.12]
Legrand 2016	0	11	0	10	5.5%	0.00 [-0.17 , 0.17]
Total (95% CI)		**98**		**87**	**100.0%**	**0.00 [-0.04 , 0.04]**
Total events:	0		0			

Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 4 (P = 1.00); I² = 0%
Test for overall effect: Z = 0.00 (P = 1.00)
Test for subgroup differences: Not applicable

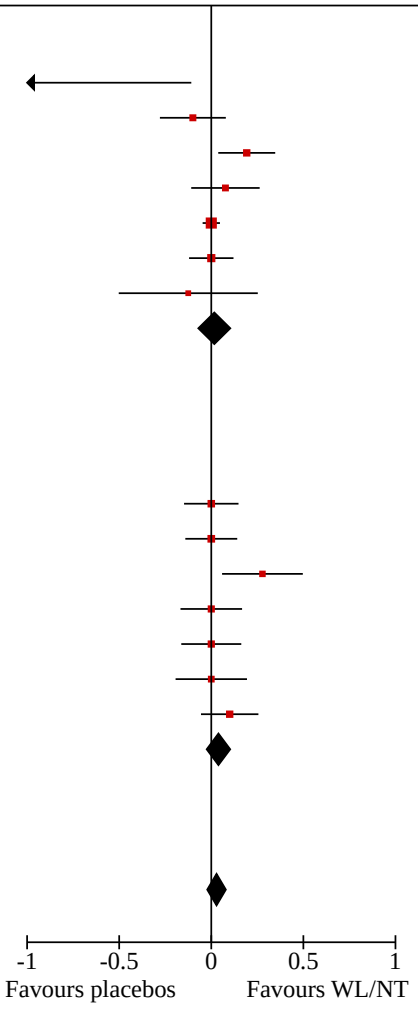


Comparison 8. Non-serious adverse events of all placebos versus wait-list or no- treatment for people with mental health disorders

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
8.1 All placebos compared with wait-list/no-treatment for dichotomous data	14		Risk Difference (IV, Random, 95% CI)	0.03 [-0.02, 0.08]
8.1.1 All placebos compared with wait-list	7		Risk Difference (IV, Random, 95% CI)	0.02 [-0.07, 0.10]
8.1.2 All placebos compared with no-treatment	7		Risk Difference (IV, Random, 95% CI)	0.04 [-0.03, 0.10]
8.2 Psychological placebos compared with wait-list/no-treatment for dichotomous data	5		Risk Difference (IV, Random, 95% CI)	0.01 [-0.18, 0.19]
8.2.1 Psychological placebos compared with wait-list	5		Risk Difference (IV, Random, 95% CI)	0.01 [-0.18, 0.19]
8.2.2 Psychological placebos compared with no-treatment	0		Risk Difference (IV, Random, 95% CI)	Not estimable
8.3 Pharmacological placebos compared with wait-list/no-treatment for dichotomous data	4	125	Risk Difference (M-H, Random, 95% CI)	0.08 [-0.04, 0.21]
8.3.1 Pharmacological placebos compared with wait-list	0	0	Risk Difference (M-H, Random, 95% CI)	Not estimable
8.3.2 Pharmacological placebos compared with no-treatment	4	125	Risk Difference (M-H, Random, 95% CI)	0.08 [-0.04, 0.21]
8.4 Physical placebos compared with wait-list/no-treatment for dichotomous data	5	185	Risk Difference (M-H, Random, 95% CI)	0.00 [-0.04, 0.04]
8.4.1 Physical placebos compared with wait-list	2	111	Risk Difference (M-H, Random, 95% CI)	0.00 [-0.04, 0.04]
8.4.2 Physical placebos compared with no-treatment	3	74	Risk Difference (M-H, Random, 95% CI)	0.00 [-0.09, 0.09]

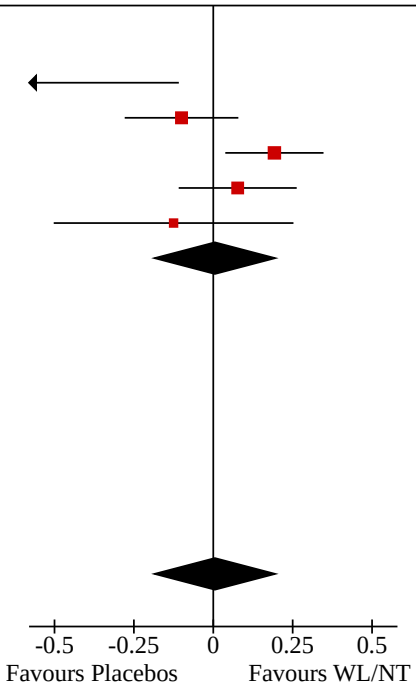
Analysis 8.1. Comparison 8: Non-serious adverse events of all placebos versus wait-list or no- treatment for people with mental health disorders, Outcome 1: All placebos compared with wait-list/no-treatment for dichotomous data

Study or Subgroup	RD	SE	Weight	Risk Difference IV, Random, 95% CI
8.1.1 All placebos compared with wait-list				
Ayen 2004	-1.5385	0.7296	0.1%	-1.54 [-2.97 , -0.11]
Ehlers 2014	-0.1	0.0913	6.1%	-0.10 [-0.28 , 0.08]
Foa 2018	0.1925	0.0789	7.5%	0.19 [0.04 , 0.35]
Goldstein 2000	0.0769	0.0948	5.7%	0.08 [-0.11 , 0.26]
Kwan 2017	0	0.024	20.9%	0.00 [-0.05 , 0.05]
Lacy 1990	0	0.0615	10.4%	0.00 [-0.12 , 0.12]
Wollersheim 1991	-0.125	0.1926	1.7%	-0.13 [-0.50 , 0.25]
Subtotal (95% CI)			**52.5%**	**0.02 [-0.07 , 0.10]**
Heterogeneity: Tau² = 0.01; Chi² = 12.49, df = 6 (P = 0.05); I² = 52%				
Test for overall effect: Z = 0.37 (P = 0.71)				
8.1.2 All placebos compared with no-treatment				
Freire 2007	0	0.0754	8.0%	0.00 [-0.15 , 0.15]
Karst 2007	0	0.0719	8.6%	0.00 [-0.14 , 0.14]
Klosko 1990	0.2778	0.1118	4.4%	0.28 [0.06 , 0.50]
Legrand 2016	0	0.0853	6.7%	0.00 [-0.17 , 0.17]
McLachlan 1991	0	0.0831	7.0%	0.00 [-0.16 , 0.16]
Nandi 1976	0	0.0989	5.4%	0.00 [-0.19 , 0.19]
Wilson 1980	0.1	0.0795	7.5%	0.10 [-0.06 , 0.26]
Subtotal (95% CI)			**47.5%**	**0.04 [-0.03 , 0.10]**
Heterogeneity: Tau² = 0.00; Chi² = 6.30, df = 6 (P = 0.39); I² = 5%				
Test for overall effect: Z = 1.19 (P = 0.23)				
Total (95% CI)			**100.0%**	**0.03 [-0.02 , 0.08]**
Heterogeneity: Tau² = 0.00; Chi² = 19.41, df = 13 (P = 0.11); I² = 33%				
Test for overall effect: Z = 1.11 (P = 0.27)				
Test for subgroup differences: Chi² = 0.17, df = 1 (P = 0.68), I² = 0%				



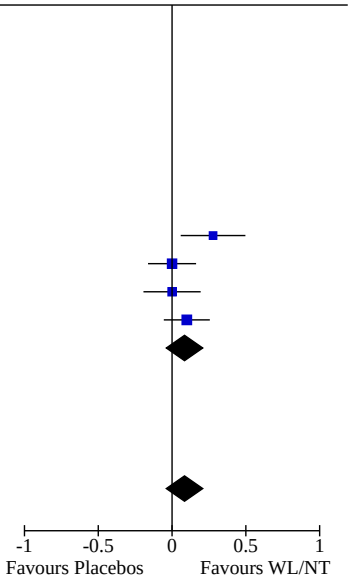
Analysis 8.2. Comparison 8: Non-serious adverse events of all placebos versus wait-list or no- treatment for people with mental health disorders, Outcome 2: Psychological placebos compared with wait-list/no-treatment for dichotomous data

Study or Subgroup	RD	SE	Weight	Risk Difference IV, Random, 95% CI
8.2.1 Psychological placebos compared with wait-list				
Ayen 2004	-1.5385	0.7296	1.7%	-1.54 [-2.97 , -0.11]
Ehlers 2014	-0.1	0.0913	27.4%	-0.10 [-0.28 , 0.08]
Foa 2018	0.1925	0.0789	29.2%	0.19 [0.04 , 0.35]
Goldstein 2000	0.0769	0.0948	26.9%	0.08 [-0.11 , 0.26]
Wollersheim 1991	-0.125	0.1926	14.9%	-0.13 [-0.50 , 0.25]
Subtotal (95% CI)			**100.0%**	**0.01 [-0.18 , 0.19]**
Heterogeneity: Tau² = 0.03; Chi² = 11.61, df = 4 (P = 0.02); I² = 66%				
Test for overall effect: Z = 0.05 (P = 0.96)				
8.2.2 Psychological placebos compared with no-treatment				
Subtotal (95% CI)				**Not estimable**
Heterogeneity: Not applicable				
Test for overall effect: Not applicable				
Total (95% CI)			**100.0%**	**0.01 [-0.18 , 0.19]**
Heterogeneity: Tau² = 0.03; Chi² = 11.61, df = 4 (P = 0.02); I² = 66%				
Test for overall effect: Z = 0.05 (P = 0.96)				
Test for subgroup differences: Not applicable				



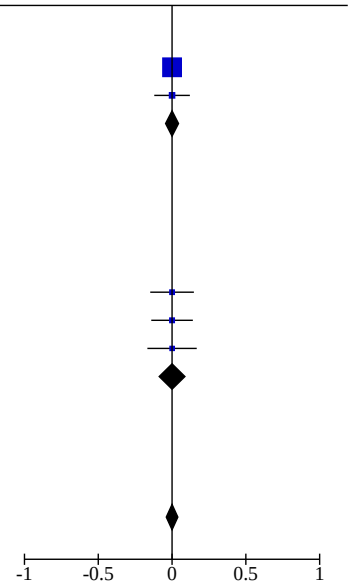
Analysis 8.3. Comparison 8: Non-serious adverse events of all placebos versus wait-list or no- treatment for people with mental health disorders, Outcome 3: Pharmacological placebos compared with wait-list/no-treatment for dichotomous data

Study or Subgroup	All Placebos Events	All Placebos Total	Wait-list/no-treatment Events	Wait-list/no-treatment Total	Weight	Risk Difference M-H, Random, 95% CI
8.3.1 Pharmacological placebos compared with wait-list						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
8.3.2 Pharmacological placebos compared with no-treatment						
Klosko 1990	5	18	0	16	20.0%	0.28 [0.06 , 0.50]
McLachlan 1991	0	9	0	14	27.9%	0.00 [-0.16 , 0.16]
Nandi 1976	0	10	0	8	23.2%	0.00 [-0.19 , 0.19]
Wilson 1980	4	40	0	10	29.0%	0.10 [-0.06 , 0.26]
Subtotal (95% CI)		**77**		**48**	**100.0%**	**0.08 [-0.04 , 0.21]**
Total events:	9		0			
Heterogeneity: Tau² = 0.01; Chi² = 5.59, df = 3 (P = 0.13); I² = 46%						
Test for overall effect: Z = 1.34 (P = 0.18)						
Total (95% CI)		**77**		**48**	**100.0%**	**0.08 [-0.04 , 0.21]**
Total events:	9		0			
Heterogeneity: Tau² = 0.01; Chi² = 5.59, df = 3 (P = 0.13); I² = 46%						
Test for overall effect: Z = 1.34 (P = 0.18)						
Test for subgroup differences: Not applicable						



Analysis 8.4. Comparison 8: Non-serious adverse events of all placebos versus wait-list or no- treatment for people with mental health disorders, Outcome 4: Physical placebos compared with wait-list/no-treatment for dichotomous data

Study or Subgroup	All Placebos Events	All Placebos Total	Wait-list/no-treatment Events	Wait-list/no-treatment Total	Weight	Risk Difference M-H, Random, 95% CI
8.4.1 Physical placebos compared with wait-list						
Kwan 2017	0	41	0	40	69.3%	0.00 [-0.05 , 0.05]
Lacy 1990	0	15	0	15	10.5%	0.00 [-0.12 , 0.12]
Subtotal (95% CI)		**56**		**55**	**79.8%**	**0.00 [-0.04 , 0.04]**
Total events:	0		0			
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 1.00); I² = 0%						
Test for overall effect: Z = 0.00 (P = 1.00)						
8.4.2 Physical placebos compared with no-treatment						
Freire 2007	0	12	0	12	7.0%	0.00 [-0.15 , 0.15]
Karst 2007	0	19	0	10	7.7%	0.00 [-0.14 , 0.14]
Legrand 2016	0	11	0	10	5.5%	0.00 [-0.17 , 0.17]
Subtotal (95% CI)		**42**		**32**	**20.2%**	**0.00 [-0.09 , 0.09]**
Total events:	0		0			
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 2 (P = 1.00); I² = 0%						
Test for overall effect: Z = 0.00 (P = 1.00)						
Total (95% CI)		**98**		**87**	**100.0%**	**0.00 [-0.04 , 0.04]**
Total events:	0		0			
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 4 (P = 1.00); I² = 0%						
Test for overall effect: Z = 0.00 (P = 1.00)						
Test for subgroup differences: Chi² = 0.00, df = 1 (P = 1.00), I² = 0%						

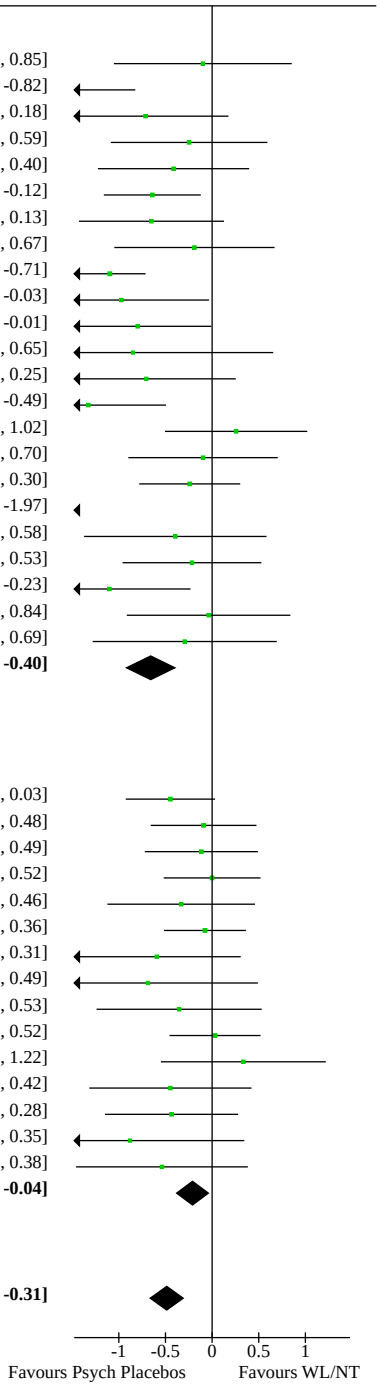


Comparison 9. Psychological placebos versus wait-list or no-treatment for people with mental health disorders

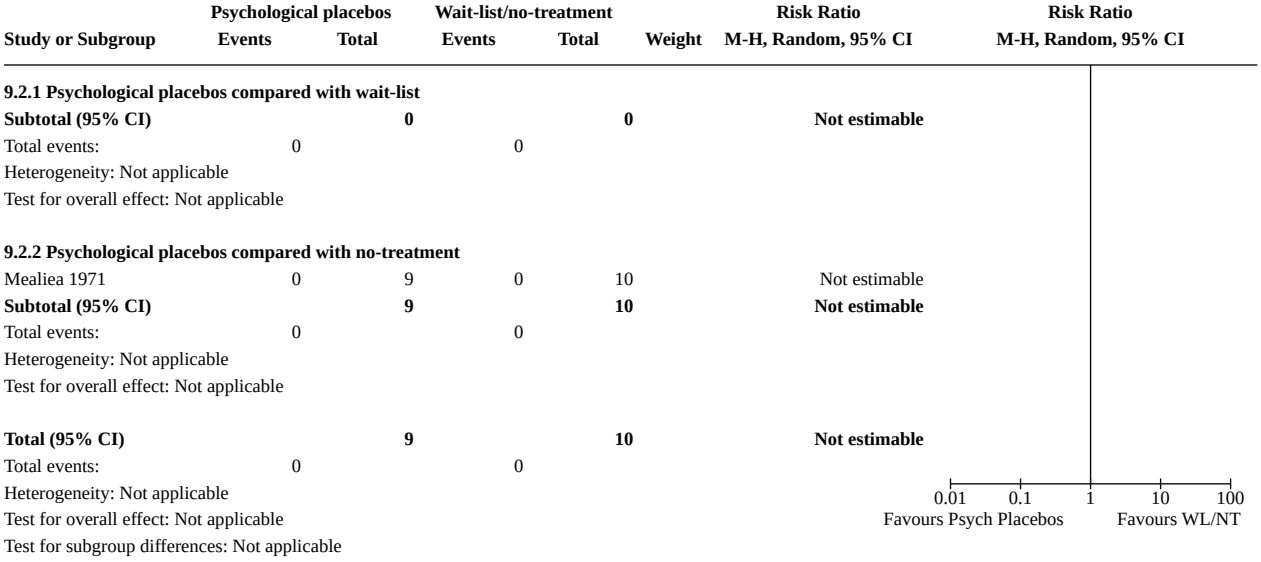
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
9.1 Psychological placebos compared with wait-list/no-treatment for continuous data	38	1263	Std. Mean Difference (IV, Random, 95% CI)	-0.49 [-0.66, -0.31]
9.1.1 Psychological placebos compared with wait-list	23	721	Std. Mean Difference (IV, Random, 95% CI)	-0.66 [-0.92, -0.40]
9.1.2 Psychological placebos compared with no-treatment	15	542	Std. Mean Difference (IV, Random, 95% CI)	-0.21 [-0.38, -0.04]
9.2 Psychological placebos compared with wait-list/no-treatment for dichotomous data	1	19	Risk Ratio (M-H, Random, 95% CI)	Not estimable
9.2.1 Psychological placebos compared with wait-list	0	0	Risk Ratio (M-H, Random, 95% CI)	Not estimable
9.2.2 Psychological placebos compared with no-treatment	1	19	Risk Ratio (M-H, Random, 95% CI)	Not estimable

Analysis 9.1. Comparison 9: Psychological placebos versus wait-list or no-treatment for people with mental health disorders, Outcome 1: Psychological placebos compared with wait-list/no-treatment for continuous data

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
9.1.1 Psychological placebos compared with wait-list								
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	2.1%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	2.3%	-1.71 [-2.60 , -0.82]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	2.3%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	2.4%	-0.25 [-1.09 , 0.59]
Bramston 1985	-78	14.21	12	-72	13.82	12	2.5%	-0.41 [-1.22 , 0.40]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	3.7%	-0.64 [-1.16 , -0.12]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	2.6%	-0.65 [-1.43 , 0.13]
Foa 1991	18.09	7.13	11	19.5	7.18	10	2.4%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	4.3%	-1.10 [-1.48 , -0.71]
Fuchs 1977	14.3	7	10	21.4	7	10	2.1%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	2.6%	-0.80 [-1.59 , -0.01]
Kilmann 1987	69	27	4	90	14	4	1.1%	-0.85 [-2.35 , 0.65]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	2.1%	-0.71 [-1.67 , 0.25]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	2.5%	-1.33 [-2.16 , -0.49]
Powers 2004	46	20.09	12	41.6	13.15	15	2.7%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	2.6%	-0.10 [-0.90 , 0.70]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	3.6%	-0.24 [-0.78 , 0.30]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	2.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	2.0%	-0.39 [-1.37 , 0.58]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	2.8%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	2.3%	-1.10 [-1.97 , -0.23]
Turner 1979	-4.1	2.644	10	-4	2.449	10	2.3%	-0.04 [-0.91 , 0.84]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	2.0%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**403**			**318**	**57.7%**	**-0.66 [-0.92 , -0.40]**
Heterogeneity: Tau² = 0.22; Chi² = 54.09, df = 22 (P = 0.0002); I² = 59%								
Test for overall effect: Z = 4.98 (P < 0.00001)								
9.1.2 Psychological placebos compared with no-treatment								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	3.9%	-0.45 [-0.93 , 0.03]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	3.5%	-0.09 [-0.66 , 0.48]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	3.3%	-0.11 [-0.72 , 0.49]
Carter 2003	3.8	1.2	28	3.8	1.3	29	3.7%	0.00 [-0.52 , 0.52]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	2.6%	-0.33 [-1.12 , 0.46]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	4.1%	-0.08 [-0.51 , 0.36]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	2.3%	-0.59 [-1.49 , 0.31]
Klein 1977	10.17	5.49	6	15	7.32	6	1.6%	-0.69 [-1.87 , 0.49]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	2.3%	-0.35 [-1.24 , 0.53]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	3.8%	0.03 [-0.46 , 0.52]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	2.3%	0.34 [-0.55 , 1.22]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	2.3%	-0.45 [-1.32 , 0.42]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	2.9%	-0.43 [-1.15 , 0.28]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	1.5%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	2.2%	-0.54 [-1.46 , 0.38]
Subtotal (95% CI)			**270**			**272**	**42.3%**	**-0.21 [-0.38 , -0.04]**
Heterogeneity: Tau² = 0.00; Chi² = 8.42, df = 14 (P = 0.87); I² = 0%								
Test for overall effect: Z = 2.42 (P = 0.02)								
Total (95% CI)			**673**			**590**	**100.0%**	**-0.49 [-0.66 , -0.31]**
Heterogeneity: Tau² = 0.15; Chi² = 78.11, df = 37 (P < 0.0001); I² = 53%								
Test for overall effect: Z = 5.40 (P < 0.00001)								
Test for subgroup differences: Chi² = 8.03, df = 1 (P = 0.005), I² = 87.5%								



Analysis 9.2. Comparison 9: Psychological placebos versus wait-list or no-treatment for people with mental health disorders, Outcome 2: Psychological placebos compared with wait-list/no-treatment for dichotomous data

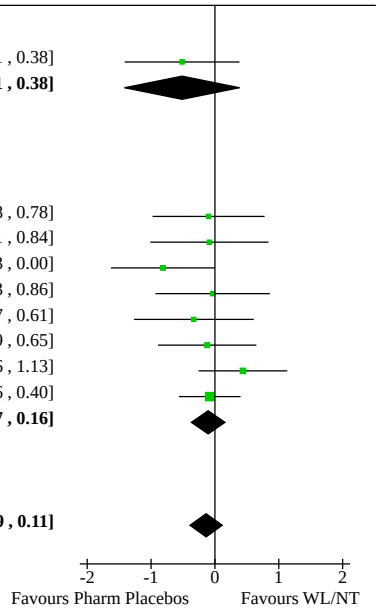


Comparison 10. Pharmacological placebos versus wait-list or no-treatment for people with mental health disorders

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
10.1 Pharmacological placebos compared with wait-list/no-treatment for continuous data	9	279	Std. Mean Difference (IV, Fixed, 95% CI)	-0.14 [-0.39, 0.11]
10.1.1 Pharmacological placebos compared with wait-list	1	20	Std. Mean Difference (IV, Fixed, 95% CI)	-0.51 [-1.41, 0.38]
10.1.2 Pharmacological placebos compared with no-treatment	8	259	Std. Mean Difference (IV, Fixed, 95% CI)	-0.11 [-0.37, 0.16]
10.2 Pharmacological placebos compared with no-treatment for continuous data	8	366	Risk Ratio (M-H, Random, 95% CI)	1.05 [0.74, 1.48]
10.2.1 Pharmacological placebos compared with wait-list	0	0	Risk Ratio (M-H, Random, 95% CI)	Not estimable
10.2.2 Pharmacological placebos compared with no-treatment	8	366	Risk Ratio (M-H, Random, 95% CI)	1.05 [0.74, 1.48]

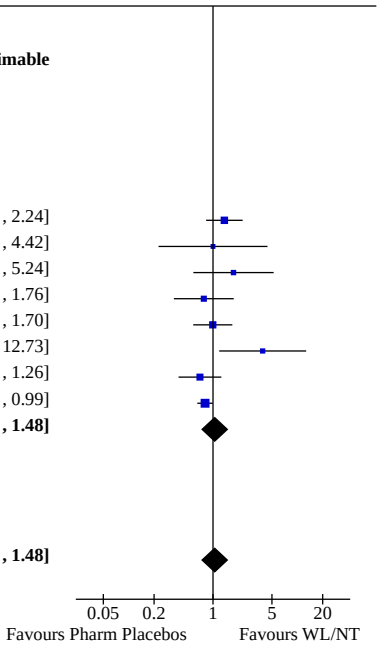
Analysis 10.1. Comparison 10: Pharmacological placebos versus wait-list or no-treatment for people with mental health disorders, Outcome 1: Pharmacological placebos compared with wait-list/no-treatment for continuous data

Study or Subgroup	Pharmacological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
10.1.1 Pharmacological placebos compared with wait-list								
Kelley 2012	-1.64	4.52	11	0.67	4	9	7.9%	-0.51 [-1.41 , 0.38]
Subtotal (95% CI)			**11**			**9**	**7.9%**	**-0.51 [-1.41 , 0.38]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.12 (P = 0.26)								
10.1.2 Pharmacological placebos compared with no-treatment								
Borden 1986	29.7	17.5	10	32	26	10	8.3%	-0.10 [-0.98 , 0.78]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	7.5%	-0.09 [-1.01 , 0.84]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	9.6%	-0.81 [-1.63 , 0.00]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	8.0%	-0.04 [-0.93 , 0.86]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	7.3%	-0.33 [-1.27 , 0.61]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	10.8%	-0.12 [-0.89 , 0.65]
Powers 2008b	64	20	18	51	37	15	13.2%	0.44 [-0.26 , 1.13]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	27.4%	-0.08 [-0.56 , 0.40]
Subtotal (95% CI)			**100**			**159**	**92.1%**	**-0.11 [-0.37 , 0.16]**
Heterogeneity: Chi² = 5.52, df = 7 (P = 0.60); I² = 0%								
Test for overall effect: Z = 0.79 (P = 0.43)								
Total (95% CI)			**111**			**168**	**100.0%**	**-0.14 [-0.39 , 0.11]**
Heterogeneity: Chi² = 6.26, df = 8 (P = 0.62); I² = 0%								
Test for overall effect: Z = 1.07 (P = 0.28)								
Test for subgroup differences: Chi² = 0.73, df = 1 (P = 0.39), I² = 0%								



Analysis 10.2. Comparison 10: Pharmacological placebos versus wait-list or no-treatment for people with mental health disorders, Outcome 2: Pharmacological placebos compared with no-treatment for continuous data

Study or Subgroup	Pharmacological placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
10.2.1 Pharmacological placebos compared with wait-list						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
10.2.2 Pharmacological placebos compared with no-treatment						
Berg 1983	9	11	9	15	16.9%	1.36 [0.83 , 2.24]
Double 1993	3	22	3	22	4.4%	1.00 [0.23 , 4.42]
Klerman 1974a	7	25	4	25	7.1%	1.75 [0.58 , 5.24]
Klerman 1974b	7	25	9	25	10.5%	0.78 [0.34 , 1.76]
Rabkin 1990	14	27	12	23	16.0%	0.99 [0.58 , 1.70]
Watzl 1988	11	34	3	36	6.3%	3.88 [1.18 , 12.73]
Whittaker 1963	7	13	10	13	14.9%	0.70 [0.39 , 1.26]
Wilson 1980	31	40	10	10	23.8%	0.80 [0.65 , 0.99]
Subtotal (95% CI)		**197**		**169**	**100.0%**	**1.05 [0.74 , 1.48]**
Total events:	89		60			
Heterogeneity: Tau² = 0.12; Chi² = 16.68, df = 7 (P = 0.02); I² = 58%						
Test for overall effect: Z = 0.26 (P = 0.79)						
Total (95% CI)		**197**		**169**	**100.0%**	**1.05 [0.74 , 1.48]**
Total events:	89		60			
Heterogeneity: Tau² = 0.12; Chi² = 16.68, df = 7 (P = 0.02); I² = 58%						
Test for overall effect: Z = 0.26 (P = 0.79)						
Test for subgroup differences: Not applicable						

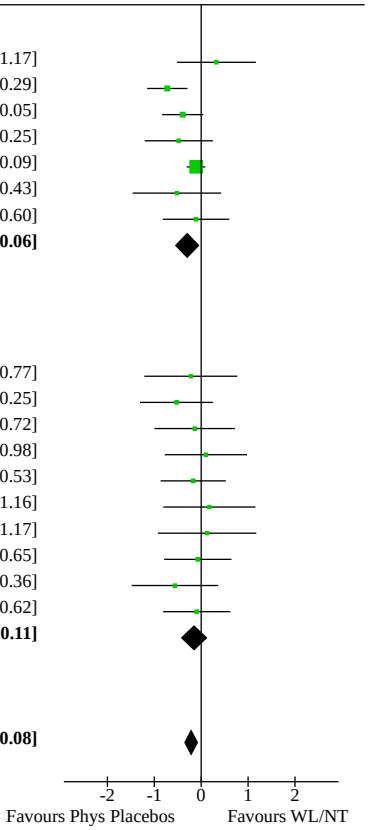


Comparison 11. Physical placebo versus wait-list or no-treatment for people with mental health disorders

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
11.1 Physical placebos compared with wait-list/no-treatment for continuous data	17	896	Std. Mean Difference (IV, Random, 95% CI)	-0.21 [-0.35, -0.08]
11.1.1 Physical placebos compared with wait-list	7	669	Std. Mean Difference (IV, Random, 95% CI)	-0.30 [-0.54, -0.06]
11.1.2 Physical placebos compared with no-treatment	10	227	Std. Mean Difference (IV, Random, 95% CI)	-0.15 [-0.42, 0.11]

Analysis 11.1. Comparison 11: Physical placebo versus wait-list or no-treatment for people with mental health disorders, Outcome 1: Physical placebos compared with wait-list/no-treatment for continuous data

Study or Subgroup	Physical placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
11.1.1 Physical placebos compared with wait-list								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	2.5%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	9.5%	-0.72 [-1.15 , -0.29]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	9.1%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	3.3%	-0.48 [-1.21 , 0.25]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	44.9%	-0.11 [-0.31 , 0.09]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	2.0%	-0.52 [-1.46 , 0.43]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	3.5%	-0.11 [-0.82 , 0.60]
Subtotal (95% CI)			**341**			**328**	**74.7%**	**-0.30 [-0.54 , -0.06]**
Heterogeneity: Tau² = 0.04; Chi² = 9.56, df = 6 (P = 0.14); I² = 37%								
Test for overall effect: Z = 2.41 (P = 0.02)								
11.1.2 Physical placebos compared with no-treatment								
Freire 2007	24.6	11	7	28.2	18	9	1.8%	-0.22 [-1.21 , 0.77]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	2.9%	-0.53 [-1.30 , 0.25]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	2.4%	-0.14 [-1.00 , 0.72]
Lick 1977	65.86	25.45	10	62.84	31.78	10	2.3%	0.10 [-0.78 , 0.98]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	3.6%	-0.17 [-0.87 , 0.53]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.8%	0.17 [-0.81 , 1.16]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.6%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	3.4%	-0.07 [-0.79 , 0.65]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	2.1%	-0.56 [-1.48 , 0.36]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	3.4%	-0.10 [-0.81 , 0.62]
Subtotal (95% CI)			**117**			**110**	**25.3%**	**-0.15 [-0.42 , 0.11]**
Heterogeneity: Tau² = 0.00; Chi² = 2.74, df = 9 (P = 0.97); I² = 0%								
Test for overall effect: Z = 1.13 (P = 0.26)								
Total (95% CI)			**458**			**438**	**100.0%**	**-0.21 [-0.35 , -0.08]**
Heterogeneity: Tau² = 0.00; Chi² = 12.59, df = 16 (P = 0.70); I² = 0%								
Test for overall effect: Z = 3.17 (P = 0.002)								
Test for subgroup differences: Chi² = 0.64, df = 1 (P = 0.42), I² = 0%								



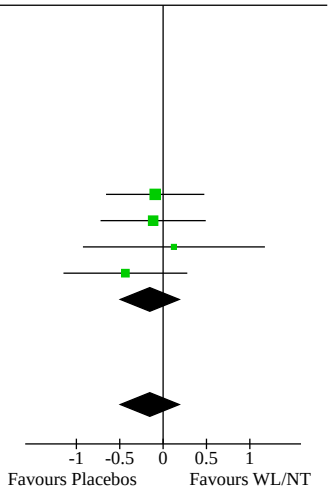
Comparison 12. All placebos versus wait-list or no-treatment for people with specific mental health disorders for continuous data

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
12.1 Efficacy of all placebos for people with substance use disorders	4	136	Std. Mean Difference (IV, Random, 95% CI)	-0.15 [-0.49, 0.19]
12.1.1 All placebos compared with wait-list	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
12.1.2 All placebos compared with no-treatment	4	136	Std. Mean Difference (IV, Random, 95% CI)	-0.15 [-0.49, 0.19]
12.2 Efficacy of all placebos for people with sleep-wake disorders	11	229	Std. Mean Difference (IV, Random, 95% CI)	-0.34 [-0.60, -0.07]
12.2.1 All placebos compared with wait-list	3	70	Std. Mean Difference (IV, Random, 95% CI)	-0.58 [-1.10, -0.05]
12.2.2 All placebos compared with no-treatment	8	159	Std. Mean Difference (IV, Random, 95% CI)	-0.24 [-0.55, 0.08]
12.3 Efficacy of all placebos for people with depression	10	286	Std. Mean Difference (IV, Random, 95% CI)	-0.42 [-0.78, -0.05]
12.3.1 All placebos compared with wait-list	6	197	Std. Mean Difference (IV, Random, 95% CI)	-0.65 [-1.15, -0.15]
12.3.2 All placebos compared with no-treatment	4	89	Std. Mean Difference (IV, Random, 95% CI)	-0.05 [-0.47, 0.37]
12.4 Efficacy of all placebos for people with post-traumatic stress disorder (PTSD)	4	329	Std. Mean Difference (IV, Random, 95% CI)	-0.54 [-1.06, -0.02]
12.4.1 All placebos compared with wait-list	3	231	Std. Mean Difference (IV, Random, 95% CI)	-0.75 [-1.23, -0.27]
12.4.2 All placebos compared with no-treatment	1	98	Std. Mean Difference (IV, Random, 95% CI)	-0.08 [-0.56, 0.40]
12.5 Efficacy of all placebos for people with anxiety disorders	16	401	Std. Mean Difference (IV, Random, 95% CI)	-0.57 [-0.93, -0.21]
12.5.1 All placebos compared with wait-list	6	181	Std. Mean Difference (IV, Random, 95% CI)	-0.81 [-1.66, 0.05]
12.5.2 All placebos compared with no-treatment	10	220	Std. Mean Difference (IV, Random, 95% CI)	-0.39 [-0.66, -0.11]
12.6 Efficacy of all placebos for people with attention deficit hyperactivity disorder (ADHD) or attention deficit disorder (ADD)	5	145	Std. Mean Difference (IV, Random, 95% CI)	-0.31 [-0.64, 0.02]
12.6.1 All placebos compared with wait-list	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
12.6.2 All placebos compared with no-treatment	5	145	Std. Mean Difference (IV, Random, 95% CI)	-0.31 [-0.64, 0.02]
12.7 Efficacy of all placebos for people with neurodegenerative disorders	4	231	Std. Mean Difference (IV, Random, 95% CI)	-0.21 [-0.47, 0.05]
12.7.1 All placebos compared with wait-list	2	134	Std. Mean Difference (IV, Random, 95% CI)	-0.33 [-0.67, 0.01]
12.7.2 All placebos compared with no-treatment	2	97	Std. Mean Difference (IV, Random, 95% CI)	-0.03 [-0.43, 0.36]

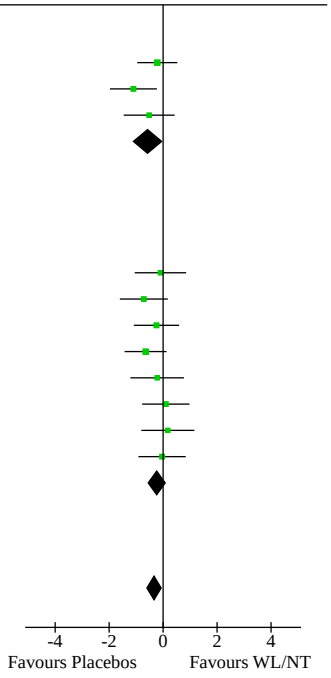
Analysis 12.1. Comparison 12: All placebos versus wait-list or no-treatment for people with specific mental health disorders for continuous data, Outcome 1: Efficacy of all placebos for people with substance use disorders

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
12.1.1 All placebos compared with wait-list								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
12.1.2 All placebos compared with no-treatment								
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	35.8%	-0.09 [-0.66 , 0.48]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	31.2%	-0.11 [-0.72 , 0.49]
Pillman 2001	-85.3	14	7	-87.6	20	7	10.4%	0.12 [-0.92 , 1.17]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	22.5%	-0.43 [-1.15 , 0.28]
Subtotal (95% CI)			**64**			**72**	**100.0%**	**-0.15 [-0.49 , 0.19]**
Heterogeneity: Tau² = 0.00; Chi² = 0.93, df = 3 (P = 0.82); I² = 0%								
Test for overall effect: Z = 0.89 (P = 0.37)								
Total (95% CI)			**64**			**72**	**100.0%**	**-0.15 [-0.49 , 0.19]**
Heterogeneity: Tau² = 0.00; Chi² = 0.93, df = 3 (P = 0.82); I² = 0%								
Test for overall effect: Z = 0.89 (P = 0.37)								
Test for subgroup differences: Not applicable								



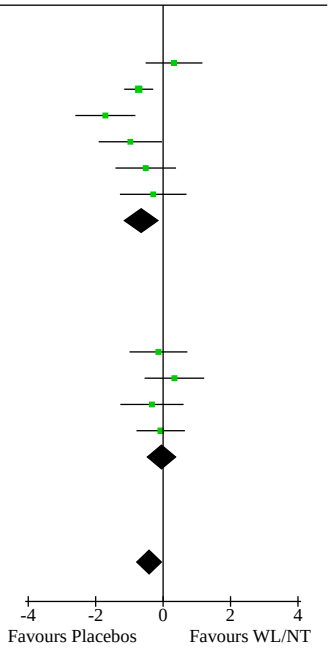
Analysis 12.2. Comparison 12: All placebos versus wait-list or no-treatment for people with specific mental health disorders for continuous data, Outcome 2: Efficacy of all placebos for people with sleep-wake disorders

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
12.2.1 All placebos compared with wait-list								
Shealy 1979	29.6	30.82	14	35.78	24.4	14	12.6%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	9.2%	-1.10 [-1.97 , -0.23]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	7.8%	-0.52 [-1.46 , 0.43]
Subtotal (95% CI)			**35**			**35**	**29.7%**	**-0.58 [-1.10 , -0.05]**
Heterogeneity: Tau² = 0.03; Chi² = 2.31, df = 2 (P = 0.31); I² = 14%								
Test for overall effect: Z = 2.16 (P = 0.03)								
12.2.2 All placebos compared with no-treatment								
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	7.7%	-0.10 [-1.05 , 0.85]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	8.8%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	9.9%	-0.25 [-1.09 , 0.59]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	11.5%	-0.65 [-1.43 , 0.13]
Freire 2007	24.6	11	7	28.2	18	9	7.1%	-0.22 [-1.21 , 0.77]
Lick 1977	65.86	25.45	10	62.84	31.78	10	9.1%	0.10 [-0.78 , 0.98]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	7.2%	0.17 [-0.81 , 1.16]
Turner 1979	-4.1	2.644	10	-4	2.449	10	9.1%	-0.04 [-0.91 , 0.84]
Subtotal (95% CI)			**79**			**80**	**70.3%**	**-0.24 [-0.55 , 0.08]**
Heterogeneity: Tau² = 0.00; Chi² = 3.71, df = 7 (P = 0.81); I² = 0%								
Test for overall effect: Z = 1.48 (P = 0.14)								
Total (95% CI)			**114**			**115**	**100.0%**	**-0.34 [-0.60 , -0.07]**
Heterogeneity: Tau² = 0.00; Chi² = 7.30, df = 10 (P = 0.70); I² = 0%								
Test for overall effect: Z = 2.50 (P = 0.01)								
Test for subgroup differences: Chi² = 1.19, df = 1 (P = 0.28), I² = 15.7%								



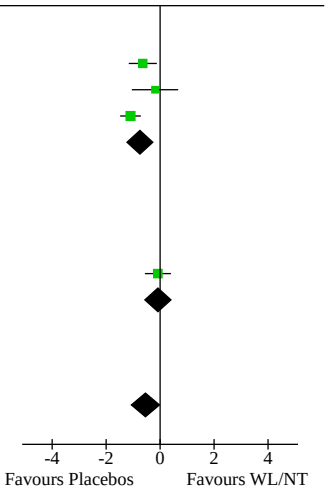
Analysis 12.3. Comparison 12: All placebos versus wait-list or no-treatment for people with specific mental health disorders for continuous data, Outcome 3: Efficacy of all placebos for people with depression

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
12.3.1 All placebos compared with wait-list								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	9.8%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	15.9%	-0.72 [-1.15 , -0.29]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	9.2%	-1.71 [-2.60 , -0.82]
Fuchs 1977	14.3	7	10	21.4	7	10	8.7%	-0.97 [-1.91 , -0.03]
Kelley 2012	-1.64	4.52	11	0.67	4	9	9.1%	-0.51 [-1.41 , 0.38]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	8.2%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**105**			**92**	**61.0%**	**-0.65 [-1.15 , -0.15]**
Heterogeneity: Tau² = 0.21; Chi² = 11.74, df = 5 (P = 0.04); I² = 57%								
Test for overall effect: Z = 2.57 (P = 0.01)								
12.3.2 All placebos compared with no-treatment								
Legrand 2016	28.43	7.46	11	29.9	12.57	10	9.6%	-0.14 [-1.00 , 0.72]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	9.3%	0.34 [-0.55 , 1.22]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	8.7%	-0.33 [-1.27 , 0.61]
Poland 2013	22	8.2	16	22.6	8	14	11.4%	-0.07 [-0.79 , 0.65]
Subtotal (95% CI)			**47**			**42**	**39.0%**	**-0.05 [-0.47 , 0.37]**
Heterogeneity: Tau² = 0.00; Chi² = 1.12, df = 3 (P = 0.77); I² = 0%								
Test for overall effect: Z = 0.23 (P = 0.82)								
Total (95% CI)			**152**			**134**	**100.0%**	**-0.42 [-0.78 , -0.05]**
Heterogeneity: Tau² = 0.17; Chi² = 18.48, df = 9 (P = 0.03); I² = 51%								
Test for overall effect: Z = 2.24 (P = 0.03)								
Test for subgroup differences: Chi² = 3.31, df = 1 (P = 0.07), I² = 69.8%								



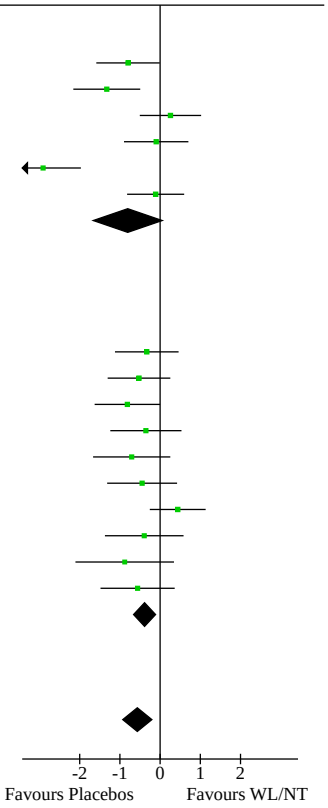
Analysis 12.4. Comparison 12: All placebos versus wait-list or no-treatment for people with specific mental health disorders for continuous data, Outcome 4: Efficacy of all placebos for people with post-traumatic stress disorder (PTSD)

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
12.4.1 All placebos compared with wait-list								
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	25.9%	-0.64 [-1.16 , -0.12]
Foa 1991	18.09	7.13	11	19.5	7.18	10	17.9%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	29.4%	-1.10 [-1.48 , -0.71]
Subtotal (95% CI)			**151**			**80**	**73.1%**	**-0.75 [-1.23 , -0.27]**
Heterogeneity: Tau² = 0.10; Chi² = 4.49, df = 2 (P = 0.11); I² = 55%								
Test for overall effect: Z = 3.07 (P = 0.002)								
12.4.2 All placebos compared with no-treatment								
Shalev 2012	47.11	20.13	21	49.18	26.4	77	26.9%	-0.08 [-0.56 , 0.40]
Subtotal (95% CI)			**21**			**77**	**26.9%**	**-0.08 [-0.56 , 0.40]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.33 (P = 0.74)								
Total (95% CI)			**172**			**157**	**100.0%**	**-0.54 [-1.06 , -0.02]**
Heterogeneity: Tau² = 0.20; Chi² = 11.66, df = 3 (P = 0.009); I² = 74%								
Test for overall effect: Z = 2.05 (P = 0.04)								
Test for subgroup differences: Chi² = 3.70, df = 1 (P = 0.05), I² = 73.0%								



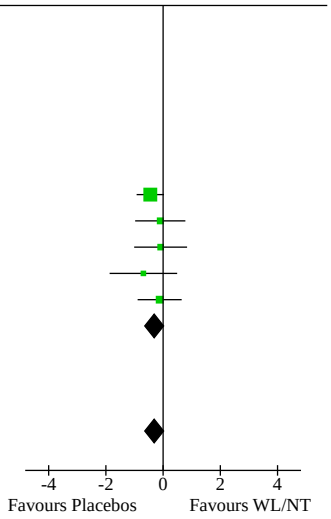
Analysis 12.5. Comparison 12: All placebos versus wait-list or no-treatment for people with specific mental health disorders for continuous data, Outcome 5: Efficacy of all placebos for people with anxiety disorders

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
12.5.1 All placebos compared with wait-list								
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	6.6%	-0.80 [-1.59 , -0.01]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	6.4%	-1.33 [-2.16 , -0.49]
Powers 2004	46	20.09	12	41.6	13.15	15	6.7%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	6.5%	-0.10 [-0.90 , 0.70]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	5.8%	-2.91 [-3.85 , -1.97]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	7.0%	-0.11 [-0.82 , 0.60]
Subtotal (95% CI)			**96**			**85**	**39.1%**	**-0.81 [-1.66 , 0.05]**
Heterogeneity: Tau² = 0.98; Chi² = 34.35, df = 5 (P < 0.00001); I² = 85%								
Test for overall effect: Z = 1.84 (P = 0.07)								
12.5.2 All placebos compared with no-treatment								
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	6.6%	-0.33 [-1.12 , 0.46]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	6.6%	-0.53 [-1.30 , 0.25]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	6.5%	-0.81 [-1.63 , 0.00]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	6.1%	-0.35 [-1.24 , 0.53]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	5.7%	-0.71 [-1.67 , 0.25]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	6.2%	-0.45 [-1.32 , 0.42]
Powers 2008b	64	20	18	51	37	15	7.1%	0.44 [-0.26 , 1.13]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	5.6%	-0.39 [-1.37 , 0.58]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	4.6%	-0.88 [-2.11 , 0.35]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	5.9%	-0.56 [-1.48 , 0.36]
Subtotal (95% CI)			**115**			**105**	**60.9%**	**-0.39 [-0.66 , -0.11]**
Heterogeneity: Tau² = 0.00; Chi² = 7.83, df = 9 (P = 0.55); I² = 0%								
Test for overall effect: Z = 2.77 (P = 0.006)								
Total (95% CI)			**211**			**190**	**100.0%**	**-0.57 [-0.93 , -0.21]**
Heterogeneity: Tau² = 0.35; Chi² = 44.01, df = 15 (P = 0.0001); I² = 66%								
Test for overall effect: Z = 3.07 (P = 0.002)								
Test for subgroup differences: Chi² = 0.83, df = 1 (P = 0.36), I² = 0%								



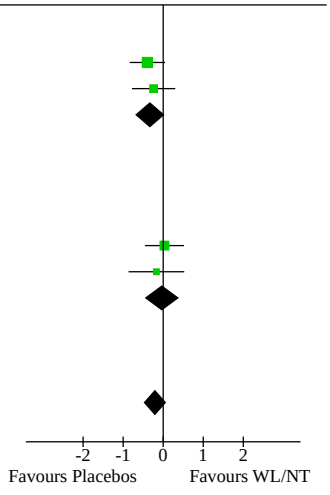
Analysis 12.6. Comparison 12: All placebos versus wait-list or no-treatment for people with specific mental health disorders for continuous data, Outcome 6: Efficacy of all placebos for people with attention deficit hyperactivity disorder (ADHD) or attention deficit disorder (ADD)

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
12.6.1 All placebos compared with wait-list								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
12.6.2 All placebos compared with no-treatment								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	47.3%	-0.45 [-0.93 , 0.03]
Borden 1986	29.7	17.5	10	32	26	10	14.1%	-0.10 [-0.98 , 0.78]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	12.7%	-0.09 [-1.01 , 0.84]
Klein 1977	10.17	5.49	6	15	7.32	6	7.8%	-0.69 [-1.87 , 0.49]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	18.3%	-0.12 [-0.89 , 0.65]
Subtotal (95% CI)			**73**			**72**	**100.0%**	**-0.31 [-0.64 , 0.02]**
Heterogeneity: Tau² = 0.00; Chi² = 1.39, df = 4 (P = 0.85); I² = 0%								
Test for overall effect: Z = 1.86 (P = 0.06)								
Total (95% CI)			**73**			**72**	**100.0%**	**-0.31 [-0.64 , 0.02]**
Heterogeneity: Tau² = 0.00; Chi² = 1.39, df = 4 (P = 0.85); I² = 0%								
Test for overall effect: Z = 1.86 (P = 0.06)								
Test for subgroup differences: Not applicable								



Analysis 12.7. Comparison 12: All placebos versus wait-list or no-treatment for people with specific mental health disorders for continuous data, Outcome 7: Efficacy of all placebos for people with neurodegenerative disorders

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
12.7.1 All placebos compared with wait-list								
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	34.8%	-0.39 [-0.83 , 0.05]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	23.0%	-0.24 [-0.78 , 0.30]
Subtotal (95% CI)			**69**			**65**	**57.8%**	**-0.33 [-0.67 , 0.01]**
Heterogeneity: Tau² = 0.00; Chi² = 0.19, df = 1 (P = 0.66); I² = 0%								
Test for overall effect: Z = 1.91 (P = 0.06)								
12.7.2 All placebos compared with no-treatment								
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	28.3%	0.03 [-0.46 , 0.52]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	13.9%	-0.17 [-0.87 , 0.53]
Subtotal (95% CI)			**50**			**47**	**42.2%**	**-0.03 [-0.43 , 0.36]**
Heterogeneity: Tau² = 0.00; Chi² = 0.22, df = 1 (P = 0.64); I² = 0%								
Test for overall effect: Z = 0.17 (P = 0.86)								
Total (95% CI)			**119**			**112**	**100.0%**	**-0.21 [-0.47 , 0.05]**
Heterogeneity: Tau² = 0.00; Chi² = 1.64, df = 3 (P = 0.65); I² = 0%								
Test for overall effect: Z = 1.56 (P = 0.12)								
Test for subgroup differences: Chi² = 1.24, df = 1 (P = 0.27), I² = 19.2%								

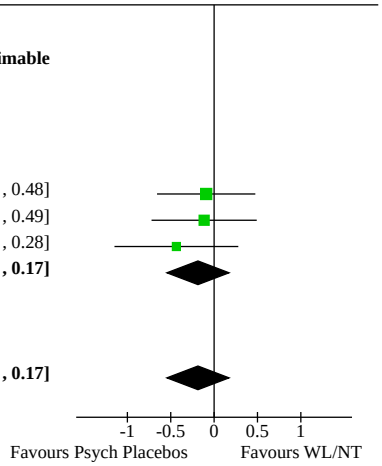


Comparison 13. Psychological placebos versus wait-list or no-treatment for people with specific mental health disorders continuous data

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
13.1 Efficacy of all placebos for people with substance use disorders	3	122	Std. Mean Difference (IV, Random, 95% CI)	-0.19 [-0.54, 0.17]
13.1.1 Psychological placebos compared with wait-list	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
13.1.2 Psychological placebos compared with no-treatment	3	122	Std. Mean Difference (IV, Random, 95% CI)	-0.19 [-0.54, 0.17]
13.2 Efficacy of all placebos for people with sleep-wake disorders	7	159	Std. Mean Difference (IV, Random, 95% CI)	-0.44 [-0.76, -0.12]
13.2.1 Psychological placebos compared with wait-list	2	52	Std. Mean Difference (IV, Random, 95% CI)	-0.63 [-1.49, 0.24]
13.2.2 Psychological placebos compared with no-treatment	5	107	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.75, 0.02]
13.3 Efficacy of all placebos for people with depression	4	86	Std. Mean Difference (IV, Random, 95% CI)	-0.66 [-1.56, 0.23]
13.3.1 Psychological placebos compared with wait-list	3	66	Std. Mean Difference (IV, Random, 95% CI)	-1.01 [-1.82, -0.21]
13.3.2 Psychological placebos compared with no-treatment	1	20	Std. Mean Difference (IV, Random, 95% CI)	0.34 [-0.55, 1.22]
13.4 Efficacy of all placebos for people with post-traumatic stress disorder (PTSD)	3	231	Std. Mean Difference (IV, Random, 95% CI)	-0.75 [-1.23, -0.27]
13.4.1 Psychological placebos compared with wait-list	3	231	Std. Mean Difference (IV, Random, 95% CI)	-0.75 [-1.23, -0.27]
13.4.2 Psychological placebos compared with no-treatment	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
13.5 Efficacy of all placebos for people with anxiety disorders	11	258	Std. Mean Difference (IV, Random, 95% CI)	-0.71 [-1.19, -0.22]
13.5.1 Psychological placebos compared with wait-list	6	157	Std. Mean Difference (IV, Random, 95% CI)	-0.94 [-1.83, -0.05]
13.5.2 Psychological placebos compared with no-treatment	5	101	Std. Mean Difference (IV, Random, 95% CI)	-0.43 [-0.83, -0.04]

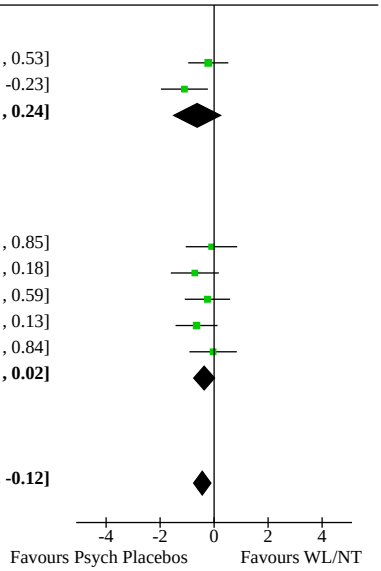
Analysis 13.1. Comparison 13: Psychological placebos versus wait-list or no-treatment for people with specific mental health disorders continuous data, Outcome 1: Efficacy of all placebos for people with substance use disorders

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
13.1.1 Psychological placebos compared with wait-list								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
13.1.2 Psychological placebos compared with no-treatment								
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	40.0%	-0.09 [-0.66 , 0.48]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	34.9%	-0.11 [-0.72 , 0.49]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	25.2%	-0.43 [-1.15 , 0.28]
Subtotal (95% CI)			**57**			**65**	**100.0%**	**-0.19 [-0.54 , 0.17]**
Heterogeneity: Tau² = 0.00; Chi² = 0.63, df = 2 (P = 0.73); I² = 0%								
Test for overall effect: Z = 1.02 (P = 0.31)								
Total (95% CI)			**57**			**65**	**100.0%**	**-0.19 [-0.54 , 0.17]**
Heterogeneity: Tau² = 0.00; Chi² = 0.63, df = 2 (P = 0.73); I² = 0%								
Test for overall effect: Z = 1.02 (P = 0.31)								
Test for subgroup differences: Not applicable								



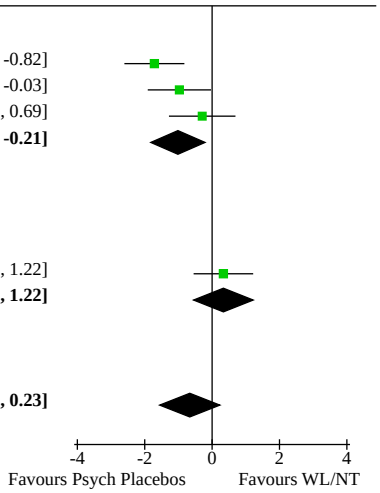
Analysis 13.2. Comparison 13: Psychological placebos versus wait-list or no-treatment for people with specific mental health disorders continuous data, Outcome 2: Efficacy of all placebos for people with sleep-wake disorders

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
13.2.1 Psychological placebos compared with wait-list								
Shealy 1979	29.6	30.82	14	35.78	24.4	14	18.3%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	13.4%	-1.10 [-1.97 , -0.23]
Subtotal (95% CI)			**26**			**26**	**31.7%**	**-0.63 [-1.49 , 0.24]**
Heterogeneity: Tau² = 0.22; Chi² = 2.30, df = 1 (P = 0.13); I² = 56%								
Test for overall effect: Z = 1.42 (P = 0.15)								
13.2.2 Psychological placebos compared with no-treatment								
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	11.2%	-0.10 [-1.05 , 0.85]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	12.8%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	14.4%	-0.25 [-1.09 , 0.59]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	16.7%	-0.65 [-1.43 , 0.13]
Turner 1979	-4.1	2.644	10	-4	2.449	10	13.2%	-0.04 [-0.91 , 0.84]
Subtotal (95% CI)			**54**			**53**	**68.3%**	**-0.37 [-0.75 , 0.02]**
Heterogeneity: Tau² = 0.00; Chi² = 2.02, df = 4 (P = 0.73); I² = 0%								
Test for overall effect: Z = 1.88 (P = 0.06)								
Total (95% CI)			**80**			**79**	**100.0%**	**-0.44 [-0.76 , -0.12]**
Heterogeneity: Tau² = 0.00; Chi² = 4.72, df = 6 (P = 0.58); I² = 0%								
Test for overall effect: Z = 2.70 (P = 0.007)								
Test for subgroup differences: Chi² = 0.29, df = 1 (P = 0.59), I² = 0%								



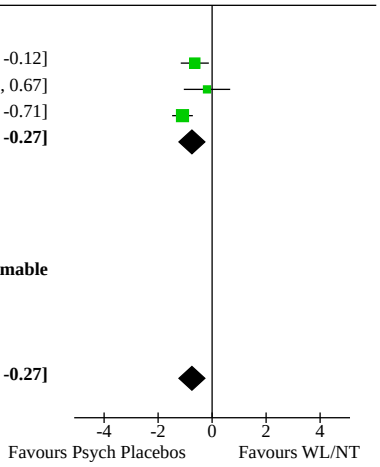
Analysis 13.3. Comparison 13: Psychological placebos versus wait-list or no-treatment for people with specific mental health disorders continuous data, Outcome 3: Efficacy of all placebos for people with depression

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
13.3.1 Psychological placebos compared with wait-list								
Ayen 2004	14.6	4.5	20	22.1	3.7	10	25.5%	-1.71 [-2.60 , -0.82]
Fuchs 1977	14.3	7	10	21.4	7	10	24.8%	-0.97 [-1.91 , -0.03]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	24.1%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**38**			**28**	**74.4%**	**-1.01 [-1.82 , -0.21]**
Heterogeneity: Tau² = 0.28; Chi² = 4.41, df = 2 (P = 0.11); I² = 55%								
Test for overall effect: Z = 2.47 (P = 0.01)								
13.3.2 Psychological placebos compared with no-treatment								
Liddle 1990	19.3	6.93	10	16.9	6.79	10	25.6%	0.34 [-0.55 , 1.22]
Subtotal (95% CI)			**10**			**10**	**25.6%**	**0.34 [-0.55 , 1.22]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.74 (P = 0.46)								
Total (95% CI)			**48**			**38**	**100.0%**	**-0.66 [-1.56 , 0.23]**
Heterogeneity: Tau² = 0.61; Chi² = 11.19, df = 3 (P = 0.01); I² = 73%								
Test for overall effect: Z = 1.46 (P = 0.15)								
Test for subgroup differences: Chi² = 4.90, df = 1 (P = 0.03), I² = 79.6%								

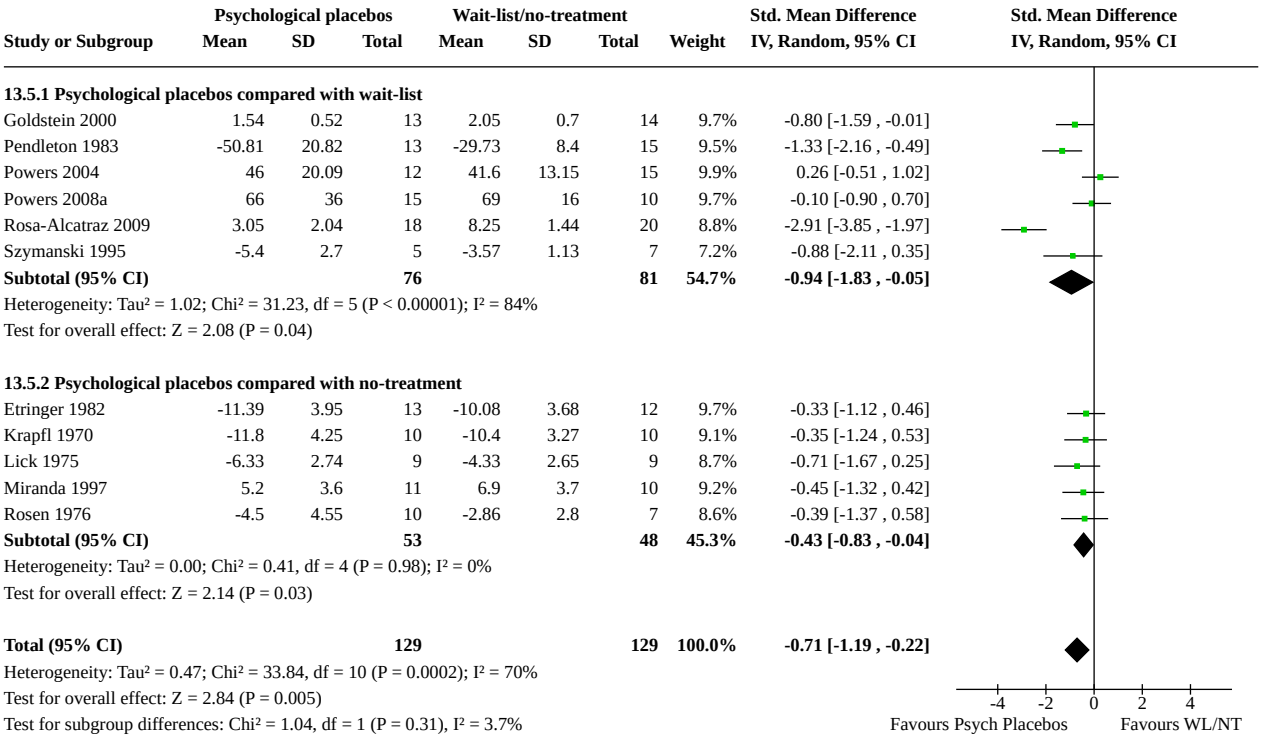


Analysis 13.4. Comparison 13: Psychological placebos versus wait-list or no-treatment for people with specific mental health disorders continuous data, Outcome 4: Efficacy of all placebos for people with post-traumatic stress disorder (PTSD)

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
13.4.1 Psychological placebos compared with wait-list								
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	35.5%	-0.64 [-1.16 , -0.12]
Foa 1991	18.09	7.13	11	19.5	7.18	10	20.6%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	43.9%	-1.10 [-1.48 , -0.71]
Subtotal (95% CI)			**151**			**80**	**100.0%**	**-0.75 [-1.23 , -0.27]**
Heterogeneity: Tau² = 0.10; Chi² = 4.49, df = 2 (P = 0.11); I² = 55%								
Test for overall effect: Z = 3.07 (P = 0.002)								
13.4.2 Psychological placebos compared with no-treatment								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
Total (95% CI)			**151**			**80**	**100.0%**	**-0.75 [-1.23 , -0.27]**
Heterogeneity: Tau² = 0.10; Chi² = 4.49, df = 2 (P = 0.11); I² = 55%								
Test for overall effect: Z = 3.07 (P = 0.002)								
Test for subgroup differences: Not applicable								



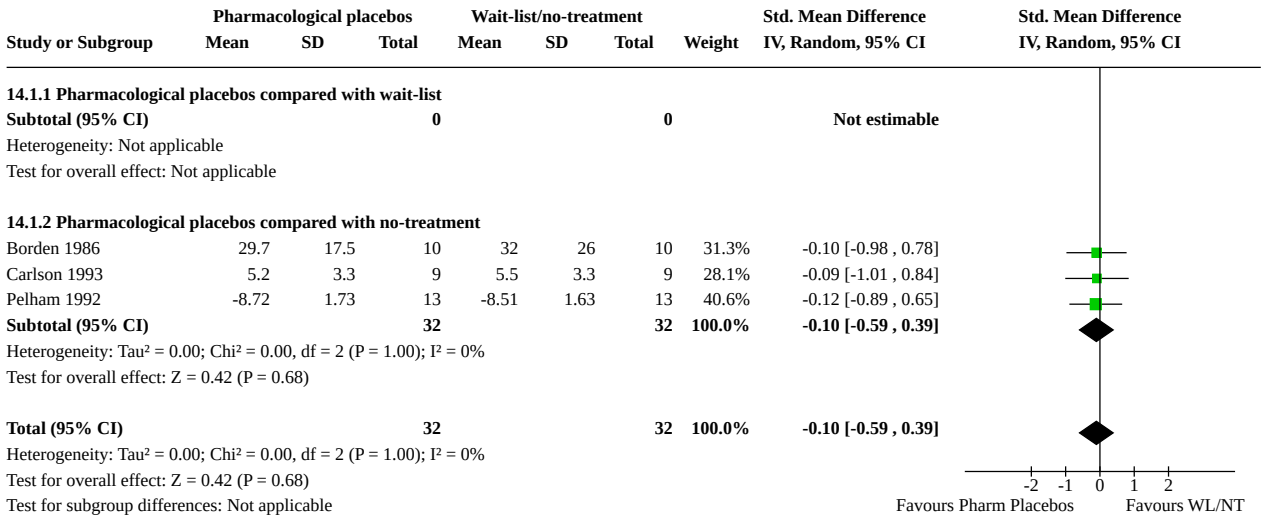
Analysis 13.5. Comparison 13: Psychological placebos versus wait-list or no-treatment for people with specific mental health disorders continuous data, Outcome 5: Efficacy of all placebos for people with anxiety disorders



Comparison 14. Pharmacological placebos versus wait-list or no-treatment for people with specific mental health disorders continuous data

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
14.1 Efficacy of all placebos for people with attention deficit hyperactivity disorder (ADHD) or attention deficit disorder (ADD)	3	64	Std. Mean Difference (IV, Random, 95% CI)	-0.10 [-0.59, 0.39]
14.1.1 Pharmacological placebos compared with wait-list	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
14.1.2 Pharmacological placebos compared with no-treatment	3	64	Std. Mean Difference (IV, Random, 95% CI)	-0.10 [-0.59, 0.39]

Analysis 14.1. Comparison 14: Pharmacological placebos versus wait-list or no-treatment for people with specific mental health disorders continuous data, Outcome 1: Efficacy of all placebos for people with attention deficit hyperactivity disorder (ADHD) or attention deficit disorder (ADD)

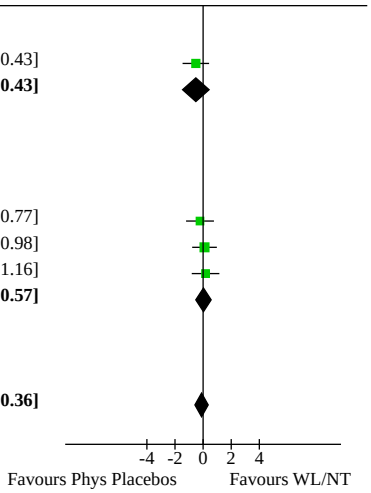


Comparison 15. Physical placebos versus wait-list or no-treatment for people with specific mental health disorders continuous data

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
15.1 Efficacy of all placebos for people with sleep-wake disorders	4	70	Std. Mean Difference (IV, Random, 95% CI)	-0.11 [-0.58, 0.36]
15.1.1 Physical placebos compared with wait-list	1	18	Std. Mean Difference (IV, Random, 95% CI)	-0.52 [-1.46, 0.43]
15.1.2 Physical placebos compared with no-treatment	3	52	Std. Mean Difference (IV, Random, 95% CI)	0.03 [-0.52, 0.57]
15.2 Efficacy of all placebos for people with depression	4	162	Std. Mean Difference (IV, Random, 95% CI)	-0.24 [-0.73, 0.25]
15.2.1 Physical placebos compared with wait-list	2	111	Std. Mean Difference (IV, Random, 95% CI)	-0.27 [-1.28, 0.75]
15.2.2 Physical placebos compared with no-treatment	2	51	Std. Mean Difference (IV, Random, 95% CI)	-0.10 [-0.65, 0.45]
15.3 Efficacy of all placebos for people with anxiety disorders	3	84	Std. Mean Difference (IV, Random, 95% CI)	-0.36 [-0.82, 0.09]
15.3.1 Physical placebos compared with wait-list	1	36	Std. Mean Difference (IV, Random, 95% CI)	-0.11 [-0.82, 0.60]
15.3.2 Physical placebos compared with no-treatment	2	48	Std. Mean Difference (IV, Random, 95% CI)	-0.54 [-1.14, 0.06]

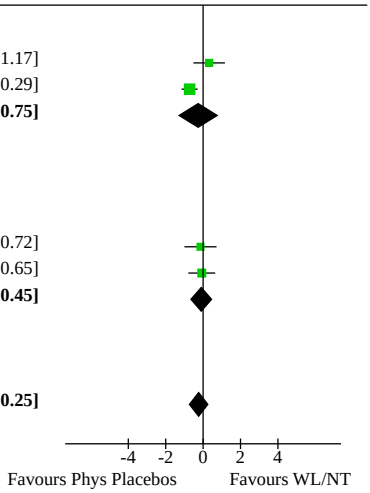
Analysis 15.1. Comparison 15: Physical placebos versus wait-list or no-treatment for people with specific mental health disorders continuous data, Outcome 1: Efficacy of all placebos for people with sleep-wake disorders

Study or Subgroup	Physical placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
15.1.1 Physical placebos compared with wait-list								
Vanderplate 1983	2	0.6	9	2.3	0.5	9	25.1%	-0.52 [-1.46 , 0.43]
Subtotal (95% CI)			**9**			**9**	**25.1%**	**-0.52 [-1.46 , 0.43]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.07 (P = 0.28)								
15.1.2 Physical placebos compared with no-treatment								
Freire 2007	24.6	11	7	28.2	18	9	22.7%	-0.22 [-1.21 , 0.77]
Lick 1977	65.86	25.45	10	62.84	31.78	10	29.0%	0.10 [-0.78 , 0.98]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	23.1%	0.17 [-0.81 , 1.16]
Subtotal (95% CI)			**25**			**27**	**74.9%**	**0.03 [-0.52 , 0.57]**
Heterogeneity: Tau² = 0.00; Chi² = 0.35, df = 2 (P = 0.84); I² = 0%								
Test for overall effect: Z = 0.09 (P = 0.93)								
Total (95% CI)			**34**			**36**	**100.0%**	**-0.11 [-0.58 , 0.36]**
Heterogeneity: Tau² = 0.00; Chi² = 1.31, df = 3 (P = 0.73); I² = 0%								
Test for overall effect: Z = 0.46 (P = 0.65)								
Test for subgroup differences: Chi² = 0.95, df = 1 (P = 0.33), I² = 0%								

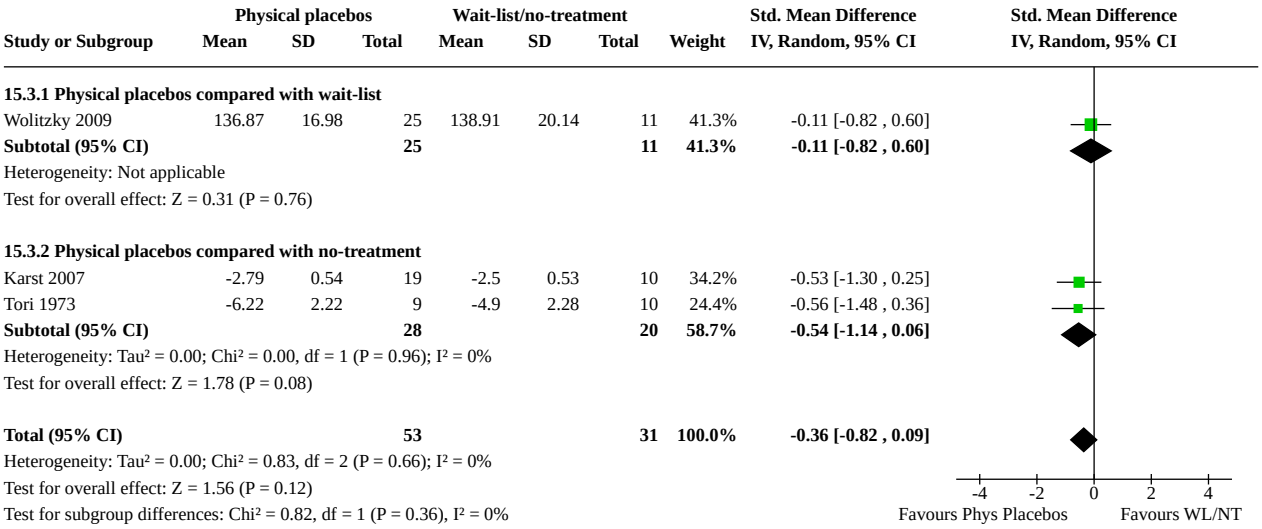


Analysis 15.2. Comparison 15: Physical placebos versus wait-list or no-treatment for people with specific mental health disorders continuous data, Outcome 2: Efficacy of all placebos for people with depression

Study or Subgroup	Physical placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
15.2.1 Physical placebos compared with wait-list								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	20.2%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	36.0%	-0.72 [-1.15 , -0.29]
Subtotal (95% CI)			**56**			**55**	**56.2%**	**-0.27 [-1.28 , 0.75]**
Heterogeneity: Tau² = 0.43; Chi² = 4.70, df = 1 (P = 0.03); I² = 79%								
Test for overall effect: Z = 0.51 (P = 0.61)								
15.2.2 Physical placebos compared with no-treatment								
Legrand 2016	28.43	7.46	11	29.9	12.57	10	19.7%	-0.14 [-1.00 , 0.72]
Poland 2013	22	8.2	16	22.6	8	14	24.1%	-0.07 [-0.79 , 0.65]
Subtotal (95% CI)			**27**			**24**	**43.8%**	**-0.10 [-0.65 , 0.45]**
Heterogeneity: Tau² = 0.00; Chi² = 0.01, df = 1 (P = 0.91); I² = 0%								
Test for overall effect: Z = 0.35 (P = 0.72)								
Total (95% CI)			**83**			**79**	**100.0%**	**-0.24 [-0.73 , 0.25]**
Heterogeneity: Tau² = 0.13; Chi² = 6.14, df = 3 (P = 0.11); I² = 51%								
Test for overall effect: Z = 0.96 (P = 0.34)								
Test for subgroup differences: Chi² = 0.08, df = 1 (P = 0.78), I² = 0%								



Analysis 15.3. Comparison 15: Physical placebos versus wait-list or no-treatment for people with specific mental health disorders continuous data, Outcome 3: Efficacy of all placebos for people with anxiety disorders



Comparison 16. Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
16.1 Type of active intervention	64	2413	Std. Mean Difference (IV, Random, 95% CI)	-0.38 [-0.50, -0.27]
16.1.1 Psychological intervention	38	1215	Std. Mean Difference (IV, Random, 95% CI)	-0.48 [-0.66, -0.30]
16.1.2 Pharmacological intervention	7	148	Std. Mean Difference (IV, Random, 95% CI)	-0.29 [-0.62, 0.04]
16.1.3 Physical intervention	13	798	Std. Mean Difference (IV, Random, 95% CI)	-0.21 [-0.35, -0.07]
16.1.4 Other or combination intervention	6	252	Std. Mean Difference (IV, Random, 95% CI)	-0.32 [-0.59, -0.04]
16.2 Type of active intervention	9	385	Risk Ratio (IV, Random, 95% CI)	1.01 [0.76, 1.35]
16.2.1 Psychological intervention	1	19	Risk Ratio (IV, Random, 95% CI)	Not estimable
16.2.2 Pharmacological intervention	7	316	Risk Ratio (IV, Random, 95% CI)	1.04 [0.74, 1.48]
16.2.3 Physical intervention	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
16.2.4 Other or combination intervention	1	50	Risk Ratio (IV, Random, 95% CI)	0.99 [0.58, 1.70]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
16.3 Risk of bias	65	2446	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.49, -0.25]
16.3.1 Low risk of bias	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
16.3.2 High risk of bias	65	2446	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.49, -0.25]
16.4 Risk of bias	9	385	Risk Ratio (IV, Random, 95% CI)	1.01 [0.76, 1.35]
16.4.1 Low risk of bias	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
16.4.2 High risk of bias	9	385	Risk Ratio (IV, Random, 95% CI)	1.01 [0.76, 1.35]
16.5 Type of outcome domain	56	2205	Std. Mean Difference (IV, Random, 95% CI)	-0.38 [-0.51, -0.24]
16.5.1 Blinded observer-reported	29	1046	Std. Mean Difference (IV, Random, 95% CI)	-0.50 [-0.70, -0.31]
16.5.2 Non-blinded observer-reported	2	96	Std. Mean Difference (IV, Random, 95% CI)	-0.54 [-0.95, -0.13]
16.5.3 Patient-reported	25	1063	Std. Mean Difference (IV, Random, 95% CI)	-0.19 [-0.35, -0.03]
16.6 Type of outcome domain	5	189	Risk Ratio (IV, Random, 95% CI)	0.94 [0.73, 1.21]
16.6.1 Blinded observer-reported	3	113	Risk Ratio (IV, Random, 95% CI)	0.99 [0.60, 1.64]
16.6.2 Non-blinded observer-reported	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
16.6.3 Patient-reported	2	76	Risk Ratio (IV, Random, 95% CI)	1.00 [0.60, 1.65]
16.7 Awareness of placebo intervention	65	2446	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.49, -0.25]
16.7.1 Open placebo	1	20	Std. Mean Difference (IV, Random, 95% CI)	-0.51 [-1.41, 0.38]
16.7.2 Closed placebo	64	2426	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.49, -0.25]
16.8 Awareness of placebo intervention	9	385	Risk Ratio (IV, Random, 95% CI)	1.01 [0.76, 1.35]
16.8.1 Open placebo	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
16.8.2 Closed placebo	9	385	Risk Ratio (IV, Random, 95% CI)	1.01 [0.76, 1.35]
16.9 The trial objective	65	2446	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.49, -0.25]

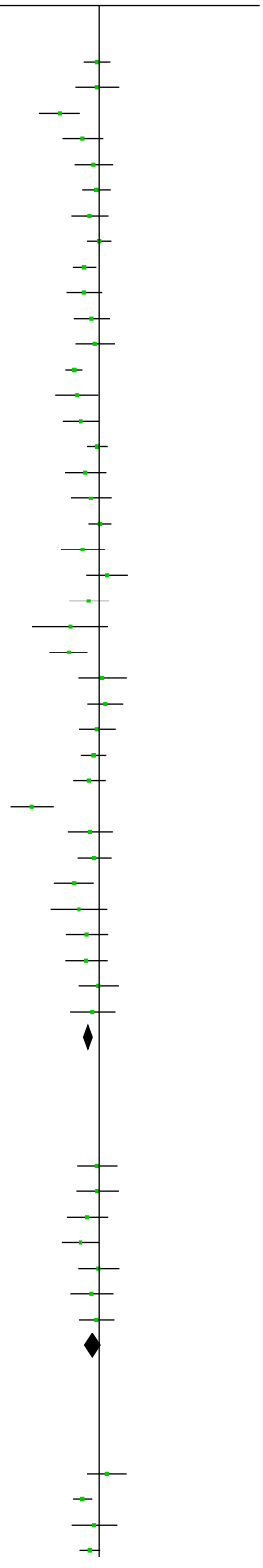
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
16.9.1 A trial's objective is clearly to assess the effects of placebo	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
16.9.2 No such objectives are stated	65	2446	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.49, -0.25]
16.10 The trial objective	9	385	Risk Ratio (IV, Random, 95% CI)	1.01 [0.76, 1.35]
16.10.1 A trial's objective is clearly to assess the effects of placebo	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
16.10.2 No such objectives are stated	9	385	Risk Ratio (IV, Random, 95% CI)	1.01 [0.76, 1.35]
16.11 Mean age of participants	57	2301	Std. Mean Difference (IV, Random, 95% CI)	-0.35 [-0.47, -0.22]
16.11.1 Below 18 years	9	274	Std. Mean Difference (IV, Random, 95% CI)	-0.47 [-0.98, 0.04]
16.11.2 18 to 50 years	40	1735	Std. Mean Difference (IV, Random, 95% CI)	-0.31 [-0.44, -0.18]
16.11.3 Above 50 years	8	292	Std. Mean Difference (IV, Random, 95% CI)	-0.41 [-0.75, -0.06]
16.12 Mean age of participants	7	340	Risk Ratio (IV, Random, 95% CI)	0.94 [0.70, 1.27]
16.12.1 18 to 50 years	5	270	Risk Ratio (IV, Random, 95% CI)	1.07 [0.70, 1.62]
16.12.2 Above 50 years	2	70	Risk Ratio (IV, Random, 95% CI)	0.73 [0.43, 1.27]
16.13 Duration of treatment	64	2366	Std. Mean Difference (IV, Random, 95% CI)	-0.38 [-0.50, -0.26]
16.13.1 Above 3 months	11	464	Std. Mean Difference (IV, Random, 95% CI)	-0.68 [-1.09, -0.27]
16.13.2 Below 3 months	53	1902	Std. Mean Difference (IV, Random, 95% CI)	-0.32 [-0.43, -0.20]
16.14 Duration of treatment	8	366	Risk Ratio (IV, Random, 95% CI)	1.01 [0.76, 1.35]
16.14.1 Above 3 months	4	196	Risk Ratio (IV, Random, 95% CI)	1.45 [0.84, 2.49]
16.14.2 Below 3 months	4	170	Risk Ratio (IV, Random, 95% CI)	0.82 [0.68, 0.98]
16.15 Mental health diagnoses	65	2446	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.49, -0.25]
16.15.1 Formal diagnosis according to DSM/ICD	29	1256	Std. Mean Difference (IV, Random, 95% CI)	-0.42 [-0.62, -0.21]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
16.15.2 Fulfil symptoms of disorder ICD/DSM while not stating systems	14	326	Std. Mean Difference (IV, Random, 95% CI)	-0.30 [-0.52, -0.08]
16.15.3 Population is classified as having a mental disorder, but full diagnostic criteria not reported	22	864	Std. Mean Difference (IV, Random, 95% CI)	-0.27 [-0.42, -0.12]
16.16 Mental health diagnoses	9	385	Risk Ratio (IV, Random, 95% CI)	1.01 [0.76, 1.35]
16.16.1 Formal diagnosis according to DSM/ICD	4	220	Risk Ratio (IV, Random, 95% CI)	1.31 [0.72, 2.37]
16.16.2 Fulfill symptoms of disorder ICD/DSM while not stating systems	3	96	Risk Ratio (IV, Random, 95% CI)	1.01 [0.62, 1.63]
16.16.3 Population is classified as having a mental disorder, but full diagnostic criteria not reported	2	69	Risk Ratio (IV, Random, 95% CI)	0.80 [0.65, 0.99]
16.17 Affiliation bias	65	2446	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.49, -0.25]
16.17.1 Risk of affiliation, industry, and allegiance bias	6	363	Std. Mean Difference (IV, Random, 95% CI)	-0.61 [-0.95, -0.26]
16.17.2 No risk found of affiliation, industry, and allegiance bias	59	2083	Std. Mean Difference (IV, Random, 95% CI)	-0.33 [-0.45, -0.21]
16.18 Risk of bias (participants and personnel excluded)	65	2446	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.49, -0.25]
16.18.1 Low risk of bias	3	230	Std. Mean Difference (IV, Random, 95% CI)	-0.58 [-0.85, -0.32]
16.18.2 Unclear risk of bias	2	51	Std. Mean Difference (IV, Random, 95% CI)	-0.12 [-0.95, 0.71]
16.18.3 High risk of bias	60	2165	Std. Mean Difference (IV, Random, 95% CI)	-0.36 [-0.48, -0.23]
16.19 Imputed data	63	2416	Std. Mean Difference (IV, Random, 95% CI)	-0.36 [-0.48, -0.24]
16.19.1 Available data	45	1236	Std. Mean Difference (IV, Random, 95% CI)	-0.39 [-0.52, -0.25]
16.19.2 Intent-to-treat	9	891	Std. Mean Difference (IV, Random, 95% CI)	-0.22 [-0.44, -0.01]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
16.19.3 No attrition	9	289	Std. Mean Difference (IV, Random, 95% CI)	-0.50 [-1.00, 0.00]

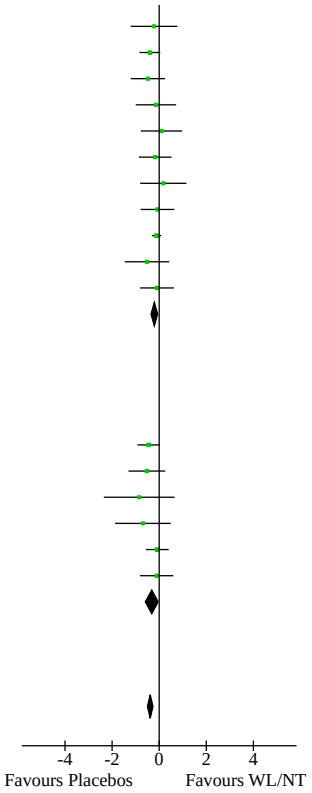
Analysis 16.1. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 1: Type of active intervention

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
16.1.1 Psychological intervention								
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	2.2%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	1.1%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	1.2%	-1.71 [-2.60 , -0.82]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	1.3%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	1.4%	-0.25 [-1.09 , 0.59]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	2.1%	-0.11 [-0.72 , 0.49]
Bramston 1985	-78	14.21	12	-72	13.82	12	1.4%	-0.41 [-1.22 , 0.40]
Carter 2003	3.8	1.2	28	3.8	1.3	29	2.4%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	2.4%	-0.64 [-1.16 , -0.12]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	1.5%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	1.5%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	1.3%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	3.1%	-1.10 [-1.48 , -0.71]
Fuchs 1977	14.3	7	10	21.4	7	10	1.2%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	1.5%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	2.8%	-0.08 [-0.51 , 0.36]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	1.2%	-0.59 [-1.49 , 0.31]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	1.3%	-0.35 [-1.24 , 0.53]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	2.6%	0.03 [-0.46 , 0.52]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	1.1%	-0.71 [-1.67 , 0.25]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	1.3%	0.34 [-0.55 , 1.22]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	1.3%	-0.45 [-1.32 , 0.42]
Peck 1976	-12.5	4.25	4	-7	3.27	4	0.5%	-1.26 [-2.90 , 0.37]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	1.4%	-1.33 [-2.16 , -0.49]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.0%	0.12 [-0.92 , 1.17]
Powers 2004	46	20.09	12	41.6	13.15	15	1.6%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	1.5%	-0.10 [-0.90 , 0.70]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	2.3%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	1.7%	-0.43 [-1.15 , 0.28]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.2%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	1.6%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	1.3%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	0.8%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	1.2%	-0.54 [-1.46 , 0.38]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	1.2%	-0.56 [-1.48 , 0.36]
Turner 1979	-4.1	2.644	10	-4	2.449	10	1.3%	-0.04 [-0.91 , 0.84]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	1.1%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**648**			**567**	**57.5%**	**-0.48 [-0.66 , -0.30]**
Heterogeneity: Tau² = 0.16; Chi² = 79.86, df = 37 (P < 0.0001); I² = 54%								
Test for overall effect: Z = 5.17 (P < 0.00001)								
16.1.2 Pharmacological intervention								
Borden 1986	29.7	17.5	10	32	26	10	1.3%	-0.10 [-0.98 , 0.78]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	1.2%	-0.09 [-1.01 , 0.84]
Kelley 2012	-1.64	4.52	11	0.67	4	9	1.2%	-0.51 [-1.41 , 0.38]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	1.4%	-0.81 [-1.63 , 0.00]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	1.2%	-0.04 [-0.93 , 0.86]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	1.2%	-0.33 [-1.27 , 0.61]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	1.5%	-0.12 [-0.89 , 0.65]
Subtotal (95% CI)			**72**			**76**	**9.0%**	**-0.29 [-0.62 , 0.04]**
Heterogeneity: Tau² = 0.00; Chi² = 2.70, df = 6 (P = 0.85); I² = 0%								
Test for overall effect: Z = 1.75 (P = 0.08)								
16.1.3 Physical intervention								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	1.4%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	2.8%	-0.72 [-1.15 , -0.29]
Freire 2007	24.6	11	7	28.2	18	9	1.1%	-0.22 [-1.21 , 0.77]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	2.8%	-0.39 [-0.83 , 0.05]



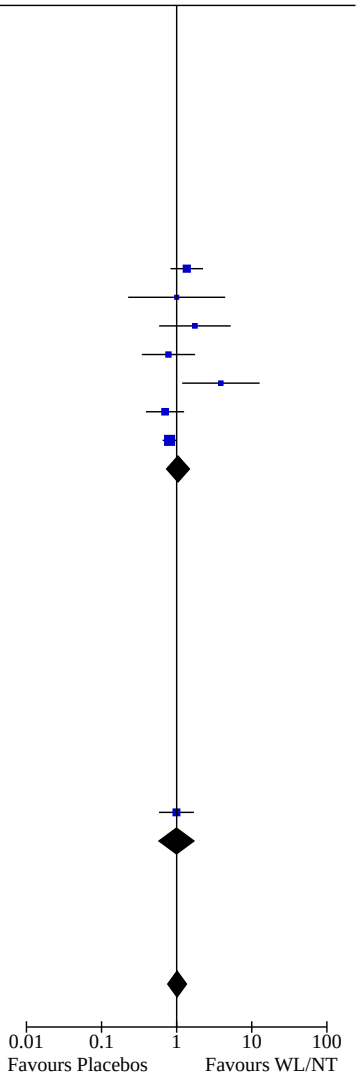
Analysis 16.1. (Continued)

Freire 2007	24.6	11	7	28.2	18	9	1.1%	-0.22 [-1.21 , 0.77]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	2.8%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	1.7%	-0.48 [-1.21 , 0.25]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	1.3%	-0.14 [-1.00 , 0.72]
Lick 1977	65.86	25.45	10	62.84	31.78	10	1.3%	0.10 [-0.78 , 0.98]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	1.7%	-0.17 [-0.87 , 0.53]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.1%	0.17 [-0.81 , 1.16]
Poland 2013	22	8.2	16	22.6	8	14	1.7%	-0.07 [-0.79 , 0.65]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	4.1%	-0.11 [-0.31 , 0.09]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	1.1%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	1.7%	-0.10 [-0.81 , 0.62]
Subtotal (95% CI)			**398**			**400**	**23.7%**	**-0.21 [-0.35 , -0.07]**
Heterogeneity: Tau² = 0.00; Chi² = 10.95, df = 12 (P = 0.53); I² = 0%								
Test for overall effect: Z = 2.90 (P = 0.004)								
16.1.4 Other or combination intervention								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	2.6%	-0.45 [-0.93 , 0.03]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	1.5%	-0.53 [-1.30 , 0.25]
Kilmann 1987	69	27	4	90	14	4	0.5%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	0.8%	-0.69 [-1.87 , 0.49]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	2.6%	-0.08 [-0.56 , 0.40]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	1.7%	-0.11 [-0.82 , 0.60]
Subtotal (95% CI)			**110**			**142**	**9.7%**	**-0.32 [-0.59 , -0.04]**
Heterogeneity: Tau² = 0.00; Chi² = 2.66, df = 5 (P = 0.75); I² = 0%								
Test for overall effect: Z = 2.28 (P = 0.02)								
Total (95% CI)			**1228**			**1185**	**100.0%**	**-0.38 [-0.50 , -0.27]**
Heterogeneity: Tau² = 0.08; Chi² = 103.46, df = 63 (P = 0.0010); I² = 39%								
Test for overall effect: Z = 6.42 (P < 0.00001)								
Test for subgroup differences: Chi² = 5.46, df = 3 (P = 0.14), I² = 45.1%								



Analysis 16.2. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 2: Type of active intervention

Study or Subgroup	All placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio IV, Random, 95% CI
16.2.1 Psychological intervention						
Mealiea 1971	0	9	0	10		Not estimable
Subtotal (95% CI)		**9**		**10**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
16.2.2 Pharmacological intervention						
Berg 1983	9	11	9	15	17.1%	1.36 [0.83 , 2.24]
Double 1993	3	22	3	22	3.4%	1.00 [0.23 , 4.42]
Klerman 1974a	7	25	4	25	5.7%	1.75 [0.58 , 5.24]
Klerman 1974b	7	25	9	25	9.1%	0.78 [0.34 , 1.76]
Watzl 1988	11	34	3	36	5.0%	3.88 [1.18 , 12.73]
Whittaker 1963	7	13	10	13	14.3%	0.70 [0.39 , 1.26]
Wilson 1980	31	40	10	10	29.6%	0.80 [0.65 , 0.99]
Subtotal (95% CI)		**170**		**146**	**84.1%**	**1.04 [0.74 , 1.48]**
Total events:	75		48			
Heterogeneity: $Tau^2 = 0.09$; $Chi^2 = 11.82$, df = 6 (P = 0.07); $I^2 = 49\%$						
Test for overall effect: Z = 0.23 (P = 0.82)						
16.2.3 Physical intervention						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
16.2.4 Other or combination intervention						
Rabkin 1990	14	27	12	23	15.9%	0.99 [0.58 , 1.70]
Subtotal (95% CI)		**27**		**23**	**15.9%**	**0.99 [0.58 , 1.70]**
Total events:	14		12			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.02 (P = 0.98)						
Total (95% CI)		**206**		**179**	**100.0%**	**1.01 [0.76 , 1.35]**
Total events:	89		60			
Heterogeneity: $Tau^2 = 0.06$; $Chi^2 = 11.95$, df = 7 (P = 0.10); $I^2 = 41\%$						
Test for overall effect: Z = 0.09 (P = 0.93)						
Test for subgroup differences: $Chi^2 = 0.02$, df = 1 (P = 0.89), $I^2 = 0\%$						

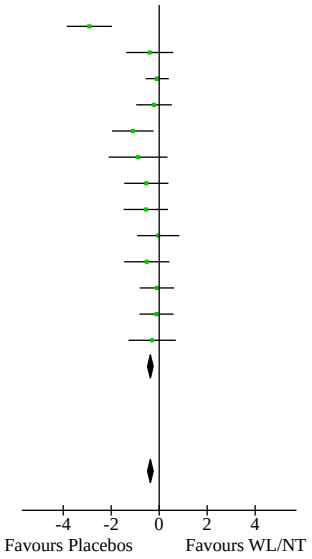


Analysis 16.3. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 3: Risk of bias

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
16.3.1 Low risk of bias								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
16.3.2 High risk of bias								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	2.5%	-0.45 [-0.93 , 0.03]
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	1.3%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	2.8%	-0.72 [-1.15 , -0.29]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	2.2%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	1.1%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	1.2%	-1.71 [-2.60 , -0.82]
Borden 1986	29.7	17.5	10	32	26	10	1.3%	-0.10 [-0.98 , 0.78]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	1.2%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	1.3%	-0.25 [-1.09 , 0.59]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	2.0%	-0.11 [-0.72 , 0.49]
Bramston 1985	-78	14.21	12	-72	13.82	12	1.4%	-0.41 [-1.22 , 0.40]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	1.2%	-0.09 [-1.01 , 0.84]
Carter 2003	3.8	1.2	28	3.8	1.3	29	2.3%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	2.3%	-0.64 [-1.16 , -0.12]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	1.5%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	1.5%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	1.3%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	3.0%	-1.10 [-1.48 , -0.71]
Freire 2007	24.6	11	7	28.2	18	9	1.1%	-0.22 [-1.21 , 0.77]
Fuchs 1977	14.3	7	10	21.4	7	10	1.1%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	1.5%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	2.7%	-0.08 [-0.51 , 0.36]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	1.5%	-0.53 [-1.30 , 0.25]
Kelley 2012	-1.64	4.52	11	0.67	4	9	1.2%	-0.51 [-1.41 , 0.38]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	1.2%	-0.59 [-1.49 , 0.31]
Kilmann 1987	69	27	4	90	14	4	0.5%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	0.8%	-0.69 [-1.87 , 0.49]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	1.4%	-0.81 [-1.63 , 0.00]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	1.3%	-0.35 [-1.24 , 0.53]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	2.7%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	1.6%	-0.48 [-1.21 , 0.25]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	2.5%	0.03 [-0.46 , 0.52]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	1.3%	-0.14 [-1.00 , 0.72]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	1.1%	-0.71 [-1.67 , 0.25]
Lick 1977	65.86	25.45	10	62.84	31.78	10	1.3%	0.10 [-0.78 , 0.98]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	1.3%	0.34 [-0.55 , 1.22]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	1.2%	-0.04 [-0.93 , 0.86]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	1.3%	-0.45 [-1.32 , 0.42]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	1.7%	-0.17 [-0.87 , 0.53]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	1.1%	-0.33 [-1.27 , 0.61]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.1%	0.17 [-0.81 , 1.16]
Peck 1976	-12.5	4.25	4	-7	3.27	4	0.5%	-1.26 [-2.90 , 0.37]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	1.5%	-0.12 [-0.89 , 0.65]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	1.4%	-1.33 [-2.16 , -0.49]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.0%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	1.7%	-0.07 [-0.79 , 0.65]
Powers 2004	46	20.09	12	41.6	13.15	15	1.5%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	1.4%	-0.10 [-0.90 , 0.70]
Powers 2008b	64	20	18	51	37	15	1.7%	0.44 [-0.26 , 1.13]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	3.9%	-0.11 [-0.31 , 0.09]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	2.3%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	1.7%	-0.43 [-1.15 , 0.28]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]

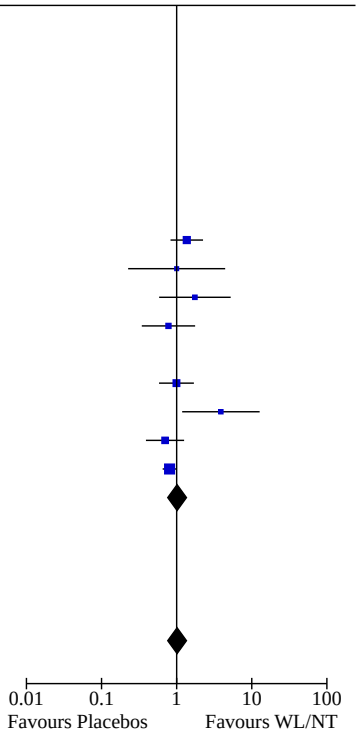
Analysis 16.3. (Continued)

Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	2.5%	-0.08 [-0.56 , 0.40]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	1.6%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	1.3%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	0.8%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	1.2%	-0.54 [-1.46 , 0.38]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	1.2%	-0.56 [-1.48 , 0.36]
Turner 1979	-4.1	2.644	10	-4	2.449	10	1.3%	-0.04 [-0.91 , 0.84]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	1.1%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	1.7%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	1.7%	-0.11 [-0.82 , 0.60]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	1.1%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**1246**			**1200**	**100.0%**	**-0.37 [-0.49 , -0.25]**
Heterogeneity: Tau² = 0.08; Chi² = 108.28, df = 64 (P = 0.0005); I² = 41%								
Test for overall effect: Z = 6.16 (P < 0.00001)								
Total (95% CI)			**1246**			**1200**	**100.0%**	**-0.37 [-0.49 , -0.25]**
Heterogeneity: Tau² = 0.08; Chi² = 108.28, df = 64 (P = 0.0005); I² = 41%								
Test for overall effect: Z = 6.16 (P < 0.00001)								
Test for subgroup differences: Not applicable								



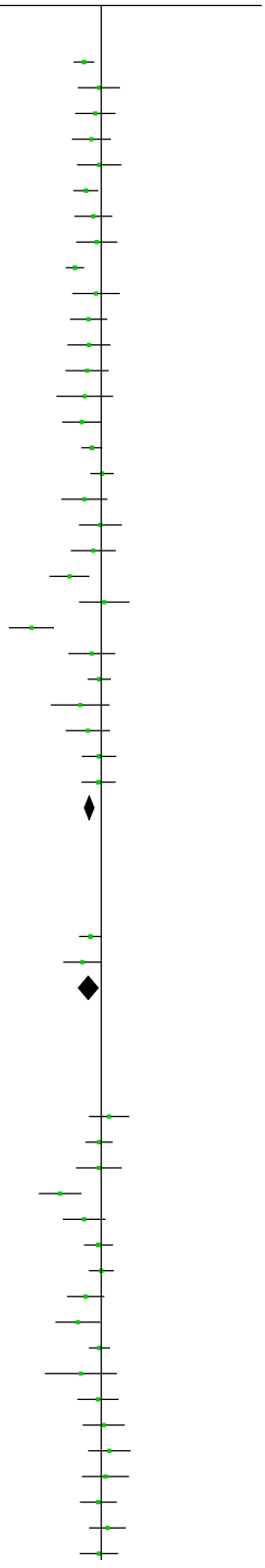
Analysis 16.4. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 4: Risk of bias

Study or Subgroup	All placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio IV, Random, 95% CI
16.4.1 Low risk of bias						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
16.4.2 High risk of bias						
Berg 1983	9	11	9	15	17.1%	1.36 [0.83 , 2.24]
Double 1993	3	22	3	22	3.4%	1.00 [0.23 , 4.42]
Klerman 1974a	7	25	4	25	5.7%	1.75 [0.58 , 5.24]
Klerman 1974b	7	25	9	25	9.1%	0.78 [0.34 , 1.76]
Mealiea 1971	0	9	0	10		Not estimable
Rabkin 1990	14	27	12	23	15.9%	0.99 [0.58 , 1.70]
Watzl 1988	11	34	3	36	5.0%	3.88 [1.18 , 12.73]
Whittaker 1963	7	13	10	13	14.3%	0.70 [0.39 , 1.26]
Wilson 1980	31	40	10	10	29.6%	0.80 [0.65 , 0.99]
Subtotal (95% CI)		**206**		**179**	**100.0%**	**1.01 [0.76 , 1.35]**
Total events:	89		60			
Heterogeneity: Tau² = 0.06; Chi² = 11.95, df = 7 (P = 0.10); I² = 41%						
Test for overall effect: Z = 0.09 (P = 0.93)						
Total (95% CI)		**206**		**179**	**100.0%**	**1.01 [0.76 , 1.35]**
Total events:	89		60			
Heterogeneity: Tau² = 0.06; Chi² = 11.95, df = 7 (P = 0.10); I² = 41%						
Test for overall effect: Z = 0.09 (P = 0.93)						
Test for subgroup differences: Not applicable						



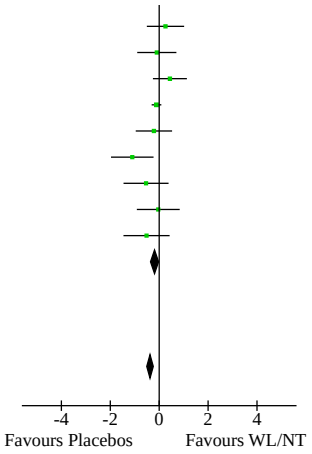
Analysis 16.5. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 5: Type of outcome domain

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
16.5.1 Blinded observer-reported								
Allen 2006	12	9.6	45	19	9.6	44	3.0%	-0.72 [-1.15 , -0.29]
Borden 1986	29.7	17.5	10	32	26	10	1.5%	-0.10 [-0.98 , 0.78]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	1.6%	-0.25 [-1.09 , 0.59]
Bramston 1985	-78	14.21	12	-72	13.82	12	1.7%	-0.41 [-1.22 , 0.40]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	1.4%	-0.09 [-1.01 , 0.84]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	2.6%	-0.64 [-1.16 , -0.12]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	1.7%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	1.5%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	3.2%	-1.10 [-1.48 , -0.71]
Freire 2007	24.6	11	7	28.2	18	9	1.3%	-0.22 [-1.21 , 0.77]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	1.7%	-0.53 [-1.30 , 0.25]
Kelley 2012	-1.64	4.52	11	0.67	4	9	1.5%	-0.51 [-1.41 , 0.38]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	1.5%	-0.59 [-1.49 , 0.31]
Klein 1977	10.17	5.49	6	15	7.32	6	1.0%	-0.69 [-1.87 , 0.49]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	1.7%	-0.81 [-1.63 , 0.00]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	2.9%	-0.39 [-0.83 , 0.05]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	2.7%	0.03 [-0.46 , 0.52]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	1.3%	-0.71 [-1.67 , 0.25]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	1.5%	-0.04 [-0.93 , 0.86]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	1.4%	-0.33 [-1.27 , 0.61]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	1.6%	-1.33 [-2.16 , -0.49]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.2%	0.12 [-0.92 , 1.17]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.4%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.3%	-0.39 [-1.37 , 0.58]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	2.8%	-0.08 [-0.56 , 0.40]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	0.9%	-0.88 [-2.11 , 0.35]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	1.4%	-0.56 [-1.48 , 0.36]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	1.9%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	1.9%	-0.11 [-0.82 , 0.60]
Subtotal (95% CI)			**541**			**505**	**51.0%**	**-0.50 [-0.70 , -0.31]**
Heterogeneity: Tau² = 0.13; Chi² = 56.48, df = 28 (P = 0.001); I² = 50%								
Test for overall effect: Z = 5.01 (P < 0.00001)								
16.5.2 Non-blinded observer-reported								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	2.8%	-0.45 [-0.93 , 0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	1.7%	-0.80 [-1.59 , -0.01]
Subtotal (95% CI)			**48**			**48**	**4.5%**	**-0.54 [-0.95 , -0.13]**
Heterogeneity: Tau² = 0.00; Chi² = 0.55, df = 1 (P = 0.46); I² = 0%								
Test for overall effect: Z = 2.60 (P = 0.009)								
16.5.3 Patient-reported								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	1.6%	0.32 [-0.52 , 1.17]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	2.4%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	1.3%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	1.5%	-1.71 [-2.60 , -0.82]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	1.5%	-0.71 [-1.60 , 0.18]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	2.3%	-0.11 [-0.72 , 0.49]
Carter 2003	3.8	1.2	28	3.8	1.3	29	2.6%	0.00 [-0.52 , 0.52]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	1.7%	-0.65 [-1.43 , 0.13]
Fuchs 1977	14.3	7	10	21.4	7	10	1.4%	-0.97 [-1.91 , -0.03]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	2.9%	-0.08 [-0.51 , 0.36]
Kilmann 1987	69	27	4	90	14	4	0.7%	-0.85 [-2.35 , 0.65]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	1.5%	-0.14 [-1.00 , 0.72]
Lick 1977	65.86	25.45	10	62.84	31.78	10	1.5%	0.10 [-0.78 , 0.98]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	1.5%	0.34 [-0.55 , 1.22]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.3%	0.17 [-0.81 , 1.16]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	1.8%	-0.12 [-0.89 , 0.65]
Powers 2004	46	20.09	12	41.6	13.15	15	1.8%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	1.7%	-0.10 [-0.90 , 0.70]



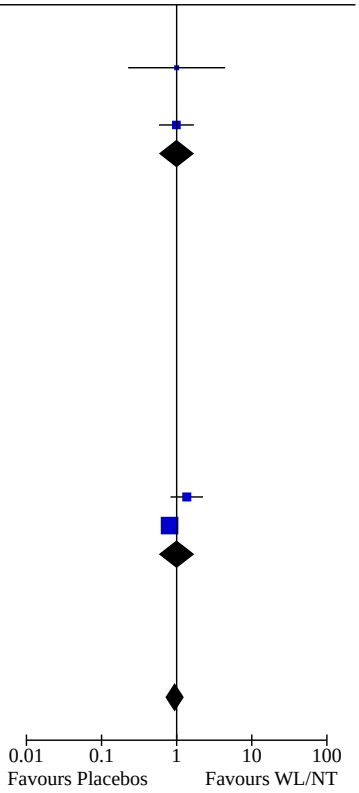
Analysis 16.5. (Continued)

Powers 2004	46	20.09	12	41.6	13.15	15	1.8%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	1.7%	-0.10 [-0.90 , 0.70]
Powers 2008b	64	20	18	51	37	15	2.0%	0.44 [-0.26 , 1.13]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	4.0%	-0.11 [-0.31 , 0.09]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	1.8%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	1.5%	-1.10 [-1.97 , -0.23]
Tan 1986	9.5	7	10	13.44	7	9	1.4%	-0.54 [-1.46 , 0.38]
Turner 1979	-4.1	2.644	10	-4	2.449	10	1.5%	-0.04 [-0.91 , 0.84]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	1.4%	-0.52 [-1.46 , 0.43]
Subtotal (95% CI)			**534**			**529**	**44.5%**	**-0.19 [-0.35 , -0.03]**
Heterogeneity: Tau² = 0.04; Chi² = 32.30, df = 24 (P = 0.12); I² = 26%								
Test for overall effect: Z = 2.27 (P = 0.02)								
Total (95% CI)			**1123**			**1082**	**100.0%**	**-0.38 [-0.51 , -0.24]**
Heterogeneity: Tau² = 0.11; Chi² = 105.90, df = 55 (P < 0.0001); I² = 48%								
Test for overall effect: Z = 5.54 (P < 0.00001)								
Test for subgroup differences: Chi² = 6.73, df = 2 (P = 0.03), I² = 70.3%								



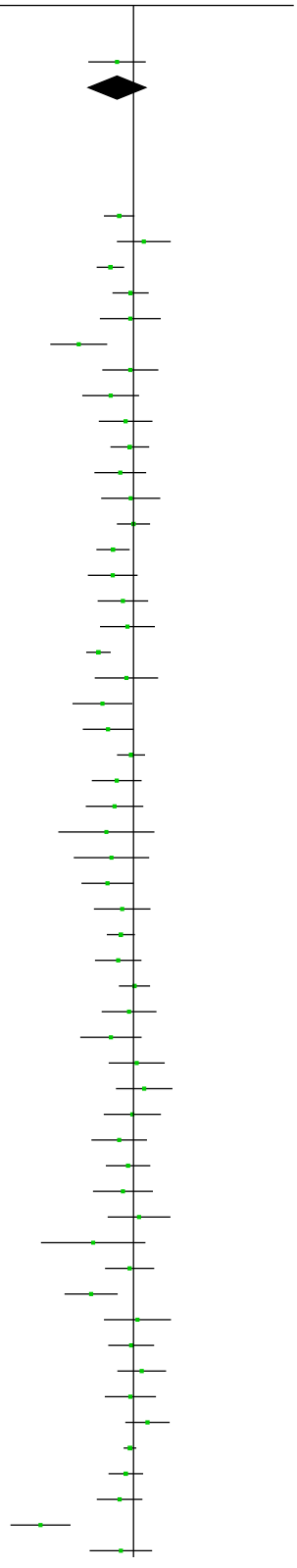
Analysis 16.6. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 6: Type of outcome domain

	All placebos		Wait-list/no-treatment			Risk Ratio	Risk Ratio
Study or Subgroup	**Events**	**Total**	**Events**	**Total**	**Weight**	**IV, Random, 95% CI**	**IV, Random, 95% CI**
16.6.1 Blinded observer-reported							
Double 1993	3	22	3	22	2.9%	1.00 [0.23 , 4.42]	
Mealiea 1971	0	9	0	10		Not estimable	
Rabkin 1990	14	27	12	23	18.5%	0.99 [0.58 , 1.70]	
Subtotal (95% CI)		**58**		**55**	**21.3%**	**0.99 [0.60 , 1.64]**	
Total events:	17		15				
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 0.99); I² = 0%							
Test for overall effect: Z = 0.02 (P = 0.98)							
16.6.2 Non-blinded observer-reported							
Subtotal (95% CI)		**0**		**0**		**Not estimable**	
Total events:	0		0				
Heterogeneity: Not applicable							
Test for overall effect: Not applicable							
16.6.3 Patient-reported							
Berg 1983	9	11	9	15	20.6%	1.36 [0.83 , 2.24]	
Wilson 1980	31	40	10	10	58.1%	0.80 [0.65 , 0.99]	
Subtotal (95% CI)		**51**		**25**	**78.7%**	**1.00 [0.60 , 1.65]**	
Total events:	40		19				
Heterogeneity: Tau² = 0.10; Chi² = 3.64, df = 1 (P = 0.06); I² = 73%							
Test for overall effect: Z = 0.01 (P = 0.99)							
Total (95% CI)		**109**		**80**	**100.0%**	**0.94 [0.73 , 1.21]**	
Total events:	57		34				
Heterogeneity: Tau² = 0.02; Chi² = 3.87, df = 3 (P = 0.28); I² = 22%							
Test for overall effect: Z = 0.49 (P = 0.63)							
Test for subgroup differences: Chi² = 0.00, df = 1 (P = 1.00), I² = 0%							



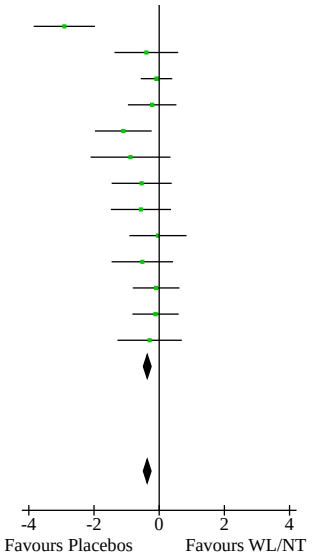
Analysis 16.7. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 7: Awareness of placebo intervention

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
16.7.1 Open placebo								
Kelley 2012	-1.64	4.52	11	0.67	4	9	1.2%	-0.51 [-1.41 , 0.38]
Subtotal (95% CI)			**11**			**9**	**1.2%**	**-0.51 [-1.41 , 0.38]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.12 (P = 0.26)								
16.7.2 Closed placebo								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	2.5%	-0.45 [-0.93 , 0.03]
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	1.3%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	2.8%	-0.72 [-1.15 , -0.29]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	2.2%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	1.1%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	1.2%	-1.71 [-2.60 , -0.82]
Borden 1986	29.7	17.5	10	32	26	10	1.3%	-0.10 [-0.98 , 0.78]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	1.2%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	1.3%	-0.25 [-1.09 , 0.59]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	2.0%	-0.11 [-0.72 , 0.49]
Bramston 1985	-78	14.21	12	-72	13.82	12	1.4%	-0.41 [-1.22 , 0.40]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	1.2%	-0.09 [-1.01 , 0.84]
Carter 2003	3.8	1.2	28	3.8	1.3	29	2.3%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	2.3%	-0.64 [-1.16 , -0.12]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	1.5%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	1.5%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	1.3%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	3.0%	-1.10 [-1.48 , -0.71]
Freire 2007	24.6	11	7	28.2	18	9	1.1%	-0.22 [-1.21 , 0.77]
Fuchs 1977	14.3	7	10	21.4	7	10	1.1%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	1.5%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	2.7%	-0.08 [-0.51 , 0.36]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	1.5%	-0.53 [-1.30 , 0.25]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	1.2%	-0.59 [-1.49 , 0.31]
Kilmann 1987	69	27	4	90	14	4	0.5%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	0.8%	-0.69 [-1.87 , 0.49]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	1.4%	-0.81 [-1.63 , 0.00]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	1.3%	-0.35 [-1.24 , 0.53]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	2.7%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	1.6%	-0.48 [-1.21 , 0.25]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	2.5%	0.03 [-0.46 , 0.52]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	1.3%	-0.14 [-1.00 , 0.72]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	1.1%	-0.71 [-1.67 , 0.25]
Lick 1977	65.86	25.45	10	62.84	31.78	10	1.3%	0.10 [-0.78 , 0.98]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	1.3%	0.34 [-0.55 , 1.22]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	1.2%	-0.04 [-0.93 , 0.86]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	1.3%	-0.45 [-1.32 , 0.42]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	1.7%	-0.17 [-0.87 , 0.53]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	1.1%	-0.33 [-1.27 , 0.61]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.1%	0.17 [-0.81 , 1.16]
Peck 1976	-12.5	4.25	4	-7	3.27	4	0.5%	-1.26 [-2.90 , 0.37]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	1.5%	-0.12 [-0.89 , 0.65]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	1.4%	-1.33 [-2.16 , -0.49]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.0%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	1.7%	-0.07 [-0.79 , 0.65]
Powers 2004	46	20.09	12	41.6	13.15	15	1.5%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	1.4%	-0.10 [-0.90 , 0.70]
Powers 2008b	64	20	18	51	37	15	1.7%	0.44 [-0.26 , 1.13]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	3.9%	-0.11 [-0.31 , 0.09]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	2.3%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	1.7%	-0.43 [-1.15 , 0.28]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]



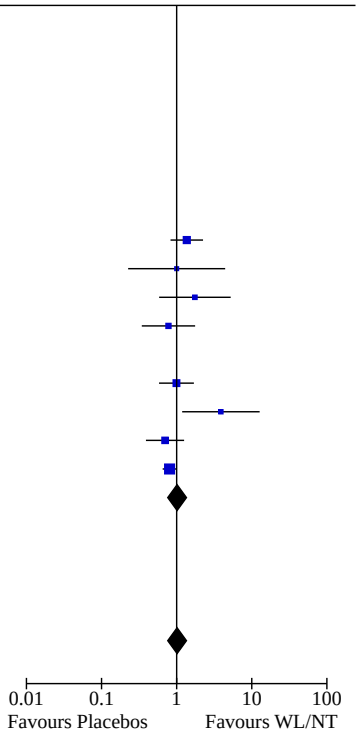
Analysis 16.7. (Continued)

Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	2.5%	-0.08 [-0.56 , 0.40]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	1.6%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	1.3%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	0.8%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	1.2%	-0.54 [-1.46 , 0.38]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	1.2%	-0.56 [-1.48 , 0.36]
Turner 1979	-4.1	2.644	10	-4	2.449	10	1.3%	-0.04 [-0.91 , 0.84]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	1.1%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	1.7%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	1.7%	-0.11 [-0.82 , 0.60]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	1.1%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**1235**			**1191**	**98.8%**	**-0.37 [-0.49 , -0.25]**
Heterogeneity: Tau² = 0.08; Chi² = 108.12, df = 63 (P = 0.0004); I² = 42%								
Test for overall effect: Z = 6.06 (P < 0.00001)								
Total (95% CI)			**1246**			**1200**	**100.0%**	**-0.37 [-0.49 , -0.25]**
Heterogeneity: Tau² = 0.08; Chi² = 108.28, df = 64 (P = 0.0005); I² = 41%								
Test for overall effect: Z = 6.16 (P < 0.00001)								
Test for subgroup differences: Chi² = 0.10, df = 1 (P = 0.75), I² = 0%								



Analysis 16.8. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 8: Awareness of placebo intervention

	All placebos		Wait-list/no-treatment			Risk Ratio	Risk Ratio
Study or Subgroup	**Events**	**Total**	**Events**	**Total**	**Weight**	**IV, Random, 95% CI**	**IV, Random, 95% CI**
16.8.1 Open placebo							
Subtotal (95% CI)		**0**		**0**		**Not estimable**	
Total events:	0		0				
Heterogeneity: Not applicable							
Test for overall effect: Not applicable							
16.8.2 Closed placebo							
Berg 1983	9	11	9	15	17.1%	1.36 [0.83 , 2.24]	
Double 1993	3	22	3	22	3.4%	1.00 [0.23 , 4.42]	
Klerman 1974a	7	25	4	25	5.7%	1.75 [0.58 , 5.24]	
Klerman 1974b	7	25	9	25	9.1%	0.78 [0.34 , 1.76]	
Mealiea 1971	0	9	0	10		Not estimable	
Rabkin 1990	14	27	12	23	15.9%	0.99 [0.58 , 1.70]	
Watzl 1988	11	34	3	36	5.0%	3.88 [1.18 , 12.73]	
Whittaker 1963	7	13	10	13	14.3%	0.70 [0.39 , 1.26]	
Wilson 1980	31	40	10	10	29.6%	0.80 [0.65 , 0.99]	
Subtotal (95% CI)		**206**		**179**	**100.0%**	**1.01 [0.76 , 1.35]**	
Total events:	89		60				
Heterogeneity: Tau² = 0.06; Chi² = 11.95, df = 7 (P = 0.10); I² = 41%							
Test for overall effect: Z = 0.09 (P = 0.93)							
Total (95% CI)		**206**		**179**	**100.0%**	**1.01 [0.76 , 1.35]**	
Total events:	89		60				
Heterogeneity: Tau² = 0.06; Chi² = 11.95, df = 7 (P = 0.10); I² = 41%							
Test for overall effect: Z = 0.09 (P = 0.93)							
Test for subgroup differences: Not applicable							

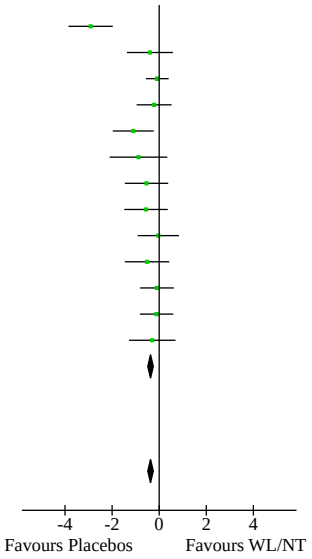


Analysis 16.9. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 9: The trial objective

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
16.9.1 A trial's objective is clearly to assess the effects of placebo								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
16.9.2 No such objectives are stated								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	2.5%	-0.45 [-0.93 , 0.03]
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	1.3%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	2.8%	-0.72 [-1.15 , -0.29]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	2.2%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	1.1%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	1.2%	-1.71 [-2.60 , -0.82]
Borden 1986	29.7	17.5	10	32	26	10	1.3%	-0.10 [-0.98 , 0.78]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	1.2%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	1.3%	-0.25 [-1.09 , 0.59]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	2.0%	-0.11 [-0.72 , 0.49]
Bramston 1985	-78	14.21	12	-72	13.82	12	1.4%	-0.41 [-1.22 , 0.40]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	1.2%	-0.09 [-1.01 , 0.84]
Carter 2003	3.8	1.2	28	3.8	1.3	29	2.3%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	2.3%	-0.64 [-1.16 , -0.12]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	1.5%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	1.5%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	1.3%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	3.0%	-1.10 [-1.48 , -0.71]
Freire 2007	24.6	11	7	28.2	18	9	1.1%	-0.22 [-1.21 , 0.77]
Fuchs 1977	14.3	7	10	21.4	7	10	1.1%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	1.5%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	2.7%	-0.08 [-0.51 , 0.36]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	1.5%	-0.53 [-1.30 , 0.25]
Kelley 2012	-1.64	4.52	11	0.67	4	9	1.2%	-0.51 [-1.41 , 0.38]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	1.2%	-0.59 [-1.49 , 0.31]
Kilmann 1987	69	27	4	90	14	4	0.5%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	0.8%	-0.69 [-1.87 , 0.49]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	1.4%	-0.81 [-1.63 , 0.00]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	1.3%	-0.35 [-1.24 , 0.53]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	2.7%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	1.6%	-0.48 [-1.21 , 0.25]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	2.5%	0.03 [-0.46 , 0.52]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	1.3%	-0.14 [-1.00 , 0.72]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	1.1%	-0.71 [-1.67 , 0.25]
Lick 1977	65.86	25.45	10	62.84	31.78	10	1.3%	0.10 [-0.78 , 0.98]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	1.3%	0.34 [-0.55 , 1.22]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	1.2%	-0.04 [-0.93 , 0.86]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	1.3%	-0.45 [-1.32 , 0.42]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	1.7%	-0.17 [-0.87 , 0.53]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	1.1%	-0.33 [-1.27 , 0.61]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.1%	0.17 [-0.81 , 1.16]
Peck 1976	-12.5	4.25	4	-7	3.27	4	0.5%	-1.26 [-2.90 , 0.37]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	1.5%	-0.12 [-0.89 , 0.65]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	1.4%	-1.33 [-2.16 , -0.49]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.0%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	1.7%	-0.07 [-0.79 , 0.65]
Powers 2004	46	20.09	12	41.6	13.15	15	1.5%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	1.4%	-0.10 [-0.90 , 0.70]
Powers 2008b	64	20	18	51	37	15	1.7%	0.44 [-0.26 , 1.13]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	3.9%	-0.11 [-0.31 , 0.09]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	2.3%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	1.7%	-0.43 [-1.15 , 0.28]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]

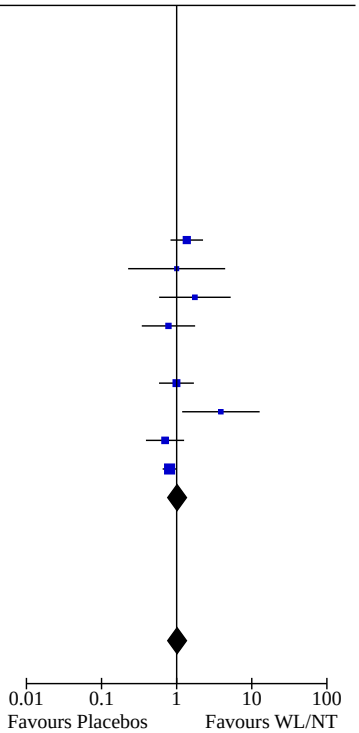
Analysis 16.9. (Continued)

Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	2.5%	-0.08 [-0.56 , 0.40]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	1.6%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	1.3%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	0.8%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	1.2%	-0.54 [-1.46 , 0.38]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	1.2%	-0.56 [-1.48 , 0.36]
Turner 1979	-4.1	2.644	10	-4	2.449	10	1.3%	-0.04 [-0.91 , 0.84]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	1.1%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	1.7%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	1.7%	-0.11 [-0.82 , 0.60]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	1.1%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**1246**			**1200**	**100.0%**	**-0.37 [-0.49 , -0.25]**
Heterogeneity: Tau² = 0.08; Chi² = 108.28, df = 64 (P = 0.0005); I² = 41%								
Test for overall effect: Z = 6.16 (P < 0.00001)								
Total (95% CI)			**1246**			**1200**	**100.0%**	**-0.37 [-0.49 , -0.25]**
Heterogeneity: Tau² = 0.08; Chi² = 108.28, df = 64 (P = 0.0005); I² = 41%								
Test for overall effect: Z = 6.16 (P < 0.00001)								
Test for subgroup differences: Not applicable								



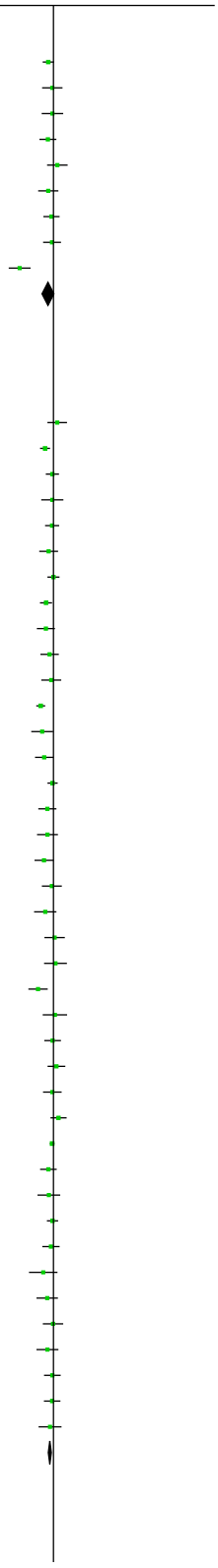
Analysis 16.10. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 10: The trial objective

Study or Subgroup	All placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio IV, Random, 95% CI
16.10.1 A trial's objective is clearly to assess the effects of placebo						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
16.10.2 No such objectives are stated						
Berg 1983	9	11	9	15	17.1%	1.36 [0.83 , 2.24]
Double 1993	3	22	3	22	3.4%	1.00 [0.23 , 4.42]
Klerman 1974a	7	25	4	25	5.7%	1.75 [0.58 , 5.24]
Klerman 1974b	7	25	9	25	9.1%	0.78 [0.34 , 1.76]
Mealiea 1971	0	9	0	10		Not estimable
Rabkin 1990	14	27	12	23	15.9%	0.99 [0.58 , 1.70]
Watzl 1988	11	34	3	36	5.0%	3.88 [1.18 , 12.73]
Whittaker 1963	7	13	10	13	14.3%	0.70 [0.39 , 1.26]
Wilson 1980	31	40	10	10	29.6%	0.80 [0.65 , 0.99]
Subtotal (95% CI)		**206**		**179**	**100.0%**	**1.01 [0.76 , 1.35]**
Total events:	89		60			
Heterogeneity: Tau² = 0.06; Chi² = 11.95, df = 7 (P = 0.10); I² = 41%						
Test for overall effect: Z = 0.09 (P = 0.93)						
Total (95% CI)		**206**		**179**	**100.0%**	**1.01 [0.76 , 1.35]**
Total events:	89		60			
Heterogeneity: Tau² = 0.06; Chi² = 11.95, df = 7 (P = 0.10); I² = 41%						
Test for overall effect: Z = 0.09 (P = 0.93)						
Test for subgroup differences: Not applicable						



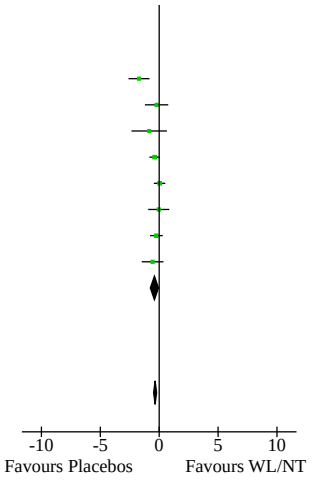
Analysis 16.11. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 11: Mean age of participants

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
16.11.1 Below 18 years								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	2.7%	-0.45 [-0.93 , 0.03]
Borden 1986	29.7	17.5	10	32	26	10	1.4%	-0.10 [-0.98 , 0.78]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	1.3%	-0.09 [-1.01 , 0.84]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	1.8%	-0.48 [-1.21 , 0.25]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	1.4%	0.34 [-0.55 , 1.22]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	1.4%	-0.45 [-1.32 , 0.42]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	1.9%	-0.17 [-0.87 , 0.53]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	1.7%	-0.12 [-0.89 , 0.65]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.3%	-2.91 [-3.85 , -1.97]
Subtotal (95% CI)			**136**			**138**	**14.9%**	**-0.47 [-0.98 , 0.04]**
Heterogeneity: Tau² = 0.45; Chi² = 31.89, df = 8 (P < 0.0001); I² = 75%								
Test for overall effect: Z = 1.82 (P = 0.07)								
16.11.2 18 to 50 years								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	1.5%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	2.9%	-0.72 [-1.15 , -0.29]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	2.3%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	1.3%	-0.10 [-1.05 , 0.85]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	2.2%	-0.11 [-0.72 , 0.49]
Bramston 1985	-78	14.21	12	-72	13.82	12	1.6%	-0.41 [-1.22 , 0.40]
Carter 2003	3.8	1.2	28	3.8	1.3	29	2.5%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	2.5%	-0.64 [-1.16 , -0.12]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	1.7%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	1.6%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	1.4%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	3.2%	-1.10 [-1.48 , -0.71]
Fuchs 1977	14.3	7	10	21.4	7	10	1.3%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	1.6%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	2.9%	-0.08 [-0.51 , 0.36]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	1.6%	-0.53 [-1.30 , 0.25]
Kelley 2012	-1.64	4.52	11	0.67	4	9	1.4%	-0.51 [-1.41 , 0.38]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	1.6%	-0.81 [-1.63 , 0.00]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	1.5%	-0.14 [-1.00 , 0.72]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	1.2%	-0.71 [-1.67 , 0.25]
Lick 1977	65.86	25.45	10	62.84	31.78	10	1.4%	0.10 [-0.78 , 0.98]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.2%	0.17 [-0.81 , 1.16]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	1.5%	-1.33 [-2.16 , -0.49]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.1%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	1.8%	-0.07 [-0.79 , 0.65]
Powers 2004	46	20.09	12	41.6	13.15	15	1.7%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	1.6%	-0.10 [-0.90 , 0.70]
Powers 2008b	64	20	18	51	37	15	1.9%	0.44 [-0.26 , 1.13]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	4.0%	-0.11 [-0.31 , 0.09]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	1.8%	-0.43 [-1.15 , 0.28]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.2%	-0.39 [-1.37 , 0.58]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	2.7%	-0.08 [-0.56 , 0.40]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	1.7%	-0.22 [-0.96 , 0.53]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	0.9%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	1.3%	-0.54 [-1.46 , 0.38]
Turner 1979	-4.1	2.644	10	-4	2.449	10	1.4%	-0.04 [-0.91 , 0.84]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	1.3%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	1.8%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	1.8%	-0.11 [-0.82 , 0.60]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	1.2%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**884**			**851**	**71.2%**	**-0.31 [-0.44 , -0.18]**
Heterogeneity: Tau² = 0.05; Chi² = 56.71, df = 39 (P = 0.03); I² = 31%								
Test for overall effect: Z = 4.69 (P < 0.00001)								
16.11.3 Above 50 years								



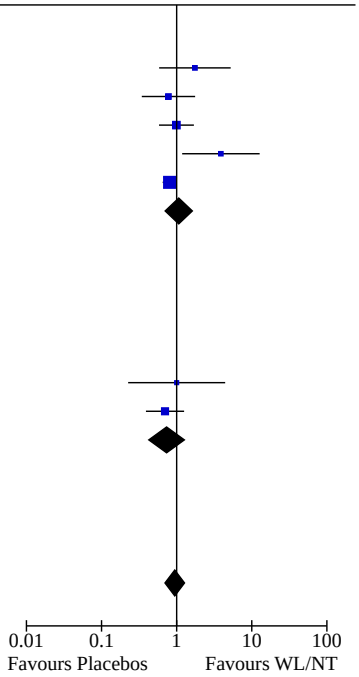
Analysis 16.11. (Continued)

16.11.3 Above 50 years								
Ayen 2004	14.6	4.5	20	22.1	3.7	10	1.4%	-1.71 [-2.60 , -0.82]
Freire 2007	24.6	11	7	28.2	18	9	1.2%	-0.22 [-1.21 , 0.77]
Kilmann 1987	69	27	4	90	14	4	0.6%	-0.85 [-2.35 , 0.65]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	2.9%	-0.39 [-0.83 , 0.05]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	2.7%	0.03 [-0.46 , 0.52]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	1.4%	-0.04 [-0.93 , 0.86]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	2.4%	-0.24 [-0.78 , 0.30]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	1.3%	-0.56 [-1.48 , 0.36]
Subtotal (95% CI)			**152**			**140**	**13.9%**	**-0.41 [-0.75 , -0.06]**
Heterogeneity: Tau² = 0.10; Chi² = 12.72, df = 7 (P = 0.08); I² = 45%								
Test for overall effect: Z = 2.29 (P = 0.02)								
Total (95% CI)			**1172**			**1129**	**100.0%**	**-0.35 [-0.47 , -0.22]**
Heterogeneity: Tau² = 0.09; Chi² = 102.50, df = 56 (P = 0.0001); I² = 45%								
Test for overall effect: Z = 5.39 (P < 0.00001)								
Test for subgroup differences: Chi² = 0.54, df = 2 (P = 0.76), I² = 0%								



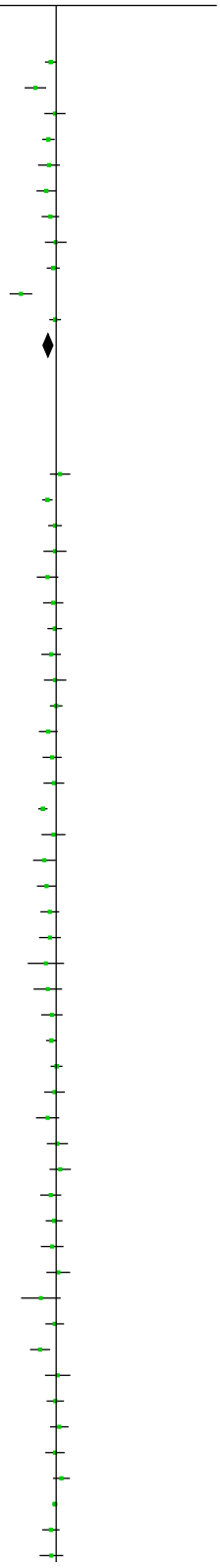
Analysis 16.12. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 12: Mean age of participants

	All placebos		Wait-list/no-treatment			Risk Ratio	Risk Ratio
Study or Subgroup	**Events**	**Total**	**Events**	**Total**	**Weight**	**IV, Random, 95% CI**	**IV, Random, 95% CI**
16.12.1 18 to 50 years							
Klerman 1974a	7	25	4	25	6.4%	1.75 [0.58 , 5.24]	
Klerman 1974b	7	25	9	25	10.4%	0.78 [0.34 , 1.76]	
Rabkin 1990	14	27	12	23	18.8%	0.99 [0.58 , 1.70]	
Watzl 1988	11	34	3	36	5.6%	3.88 [1.18 , 12.73]	
Wilson 1980	31	40	10	10	38.2%	0.80 [0.65 , 0.99]	
Subtotal (95% CI)		**151**		**119**	**79.4%**	**1.07 [0.70 , 1.62]**	
Total events:	70		38				
Heterogeneity: Tau² = 0.11; Chi² = 8.48, df = 4 (P = 0.08); I² = 53%							
Test for overall effect: Z = 0.30 (P = 0.76)							
16.12.2 Above 50 years							
Double 1993	3	22	3	22	3.7%	1.00 [0.23 , 4.42]	
Whittaker 1963	7	13	10	13	16.8%	0.70 [0.39 , 1.26]	
Subtotal (95% CI)		**35**		**35**	**20.6%**	**0.73 [0.43 , 1.27]**	
Total events:	10		13				
Heterogeneity: Tau² = 0.00; Chi² = 0.19, df = 1 (P = 0.66); I² = 0%							
Test for overall effect: Z = 1.11 (P = 0.27)							
Total (95% CI)		**186**		**154**	**100.0%**	**0.94 [0.70 , 1.27]**	
Total events:	80		51				
Heterogeneity: Tau² = 0.05; Chi² = 9.03, df = 6 (P = 0.17); I² = 34%							
Test for overall effect: Z = 0.38 (P = 0.70)							
Test for subgroup differences: Chi² = 1.14, df = 1 (P = 0.29), I² = 12.2%							



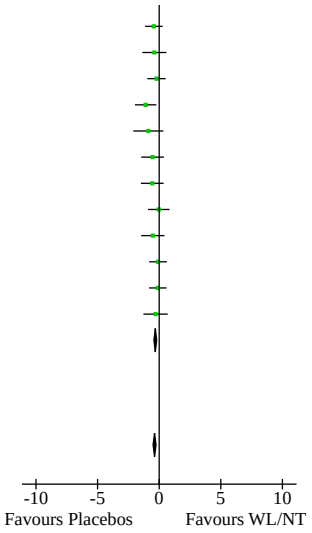
Analysis 16.13. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 13: Duration of treatment

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
16.13.1 Above 3 months								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	2.6%	-0.45 [-0.93 , 0.03]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	1.3%	-1.71 [-2.60 , -0.82]
Borden 1986	29.7	17.5	10	32	26	10	1.3%	-0.10 [-0.98 , 0.78]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	2.4%	-0.64 [-1.16 , -0.12]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	1.3%	-0.59 [-1.49 , 0.31]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	1.4%	-0.81 [-1.63 , 0.00]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	1.7%	-0.48 [-1.21 , 0.25]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	1.3%	-0.04 [-0.93 , 0.86]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	2.3%	-0.24 [-0.78 , 0.30]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.2%	-2.91 [-3.85 , -1.97]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	2.6%	-0.08 [-0.56 , 0.40]
Subtotal (95% CI)			**206**			**258**	**19.3%**	**-0.68 [-1.09 , -0.27]**
Heterogeneity: Tau² = 0.33; Chi² = 38.70, df = 10 (P < 0.0001); I² = 74%								
Test for overall effect: Z = 3.28 (P = 0.001)								
16.13.2 Below 3 months								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	1.4%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	2.8%	-0.72 [-1.15 , -0.29]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	2.2%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	1.2%	-0.10 [-1.05 , 0.85]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	1.3%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	1.4%	-0.25 [-1.09 , 0.59]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	2.1%	-0.11 [-0.72 , 0.49]
Bramston 1985	-78	14.21	12	-72	13.82	12	1.5%	-0.41 [-1.22 , 0.40]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	1.2%	-0.09 [-1.01 , 0.84]
Carter 2003	3.8	1.2	28	3.8	1.3	29	2.4%	0.00 [-0.52 , 0.52]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	1.5%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	1.5%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	1.3%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	3.0%	-1.10 [-1.48 , -0.71]
Freire 2007	24.6	11	7	28.2	18	9	1.1%	-0.22 [-1.21 , 0.77]
Fuchs 1977	14.3	7	10	21.4	7	10	1.2%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	1.5%	-0.80 [-1.59 , -0.01]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	1.5%	-0.53 [-1.30 , 0.25]
Kelley 2012	-1.64	4.52	11	0.67	4	9	1.3%	-0.51 [-1.41 , 0.38]
Kilmann 1987	69	27	4	90	14	4	0.6%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	0.8%	-0.69 [-1.87 , 0.49]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	1.3%	-0.35 [-1.24 , 0.53]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	2.8%	-0.39 [-0.83 , 0.05]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	2.5%	0.03 [-0.46 , 0.52]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	1.3%	-0.14 [-1.00 , 0.72]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	1.1%	-0.71 [-1.67 , 0.25]
Lick 1977	65.86	25.45	10	62.84	31.78	10	1.3%	0.10 [-0.78 , 0.98]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	1.3%	0.34 [-0.55 , 1.22]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	1.3%	-0.45 [-1.32 , 0.42]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	1.8%	-0.17 [-0.87 , 0.53]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	1.2%	-0.33 [-1.27 , 0.61]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.1%	0.17 [-0.81 , 1.16]
Peck 1976	-12.5	4.25	4	-7	3.27	4	0.5%	-1.26 [-2.90 , 0.37]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	1.6%	-0.12 [-0.89 , 0.65]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	1.4%	-1.33 [-2.16 , -0.49]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.0%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	1.7%	-0.07 [-0.79 , 0.65]
Powers 2004	46	20.09	12	41.6	13.15	15	1.6%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	1.5%	-0.10 [-0.90 , 0.70]
Powers 2008b	64	20	18	51	37	15	1.8%	0.44 [-0.26 , 1.13]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	3.9%	-0.11 [-0.31 , 0.09]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	1.7%	-0.43 [-1.15 , 0.28]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]



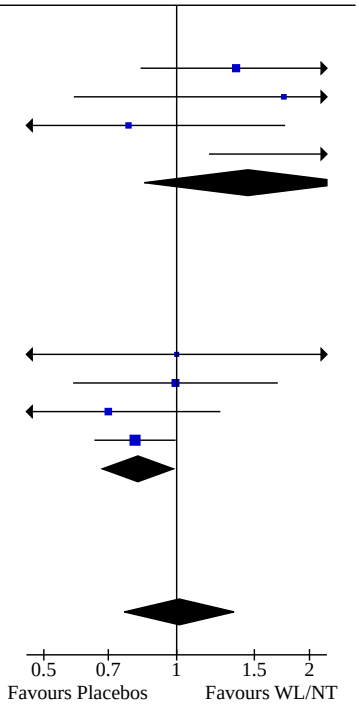
Analysis 16.13. (Continued)

Roehrich 1993	-53	10.18	16	-48.93	7.81	15	1.7%	-0.43 [-1.15 , 0.28]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	1.6%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	1.3%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	0.8%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	1.2%	-0.54 [-1.46 , 0.38]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	1.2%	-0.56 [-1.48 , 0.36]
Turner 1979	-4.1	2.644	10	-4	2.449	10	1.3%	-0.04 [-0.91 , 0.84]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	1.2%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	1.7%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	1.7%	-0.11 [-0.82 , 0.60]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	1.1%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**1000**			**902**	**80.7%**	**-0.32 [-0.43 , -0.20]**
Heterogeneity: Tau² = 0.03; Chi² = 63.05, df = 52 (P = 0.14); I² = 18%								
Test for overall effect: Z = 5.57 (P < 0.00001)								
Total (95% CI)			**1206**			**1160**	**100.0%**	**-0.38 [-0.50 , -0.26]**
Heterogeneity: Tau² = 0.08; Chi² = 106.90, df = 63 (P = 0.0005); I² = 41%								
Test for overall effect: Z = 6.18 (P < 0.00001)								
Test for subgroup differences: Chi² = 2.86, df = 1 (P = 0.09), I² = 65.1%								



Analysis 16.14. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 14: Duration of treatment

Study or Subgroup	All placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio IV, Random, 95% CI
16.14.1 Above 3 months						
Berg 1983	9	11	9	15	17.1%	1.36 [0.83 , 2.24]
Klerman 1974a	7	25	4	25	5.7%	1.75 [0.58 , 5.24]
Klerman 1974b	7	25	9	25	9.1%	0.78 [0.34 , 1.76]
Watzl 1988	11	34	3	36	5.0%	3.88 [1.18 , 12.73]
Subtotal (95% CI)		**95**		**101**	**36.9%**	**1.45 [0.84 , 2.49]**
Total events:	34		25			
Heterogeneity: Tau² = 0.12; Chi² = 4.99, df = 3 (P = 0.17); I² = 40%						
Test for overall effect: Z = 1.35 (P = 0.18)						
16.14.2 Below 3 months						
Double 1993	3	22	3	22	3.4%	1.00 [0.23 , 4.42]
Rabkin 1990	14	27	12	23	15.9%	0.99 [0.58 , 1.70]
Whittaker 1963	7	13	10	13	14.3%	0.70 [0.39 , 1.26]
Wilson 1980	31	40	10	10	29.6%	0.80 [0.65 , 0.99]
Subtotal (95% CI)		**102**		**68**	**63.1%**	**0.82 [0.68 , 0.98]**
Total events:	55		35			
Heterogeneity: Tau² = 0.00; Chi² = 0.88, df = 3 (P = 0.83); I² = 0%						
Test for overall effect: Z = 2.14 (P = 0.03)						
Total (95% CI)		**197**		**169**	**100.0%**	**1.01 [0.76 , 1.35]**
Total events:	89		60			
Heterogeneity: Tau² = 0.06; Chi² = 11.95, df = 7 (P = 0.10); I² = 41%						
Test for overall effect: Z = 0.09 (P = 0.93)						
Test for subgroup differences: Chi² = 3.88, df = 1 (P = 0.05), I² = 74.2%						

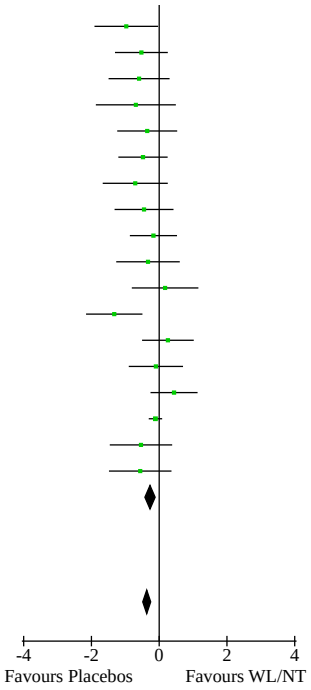


Analysis 16.15. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 15: Mental health diagnoses

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
16.15.1 Formal diagnosis according to DSM/ICD								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	2.5%	-0.45 [-0.93 , 0.03]
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	1.3%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	2.8%	-0.72 [-1.15 , -0.29]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	2.2%	-0.09 [-0.66 , 0.48]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	1.2%	-1.71 [-2.60 , -0.82]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	2.0%	-0.11 [-0.72 , 0.49]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	1.2%	-0.09 [-1.01 , 0.84]
Carter 2003	3.8	1.2	28	3.8	1.3	29	2.3%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	2.3%	-0.64 [-1.16 , -0.12]
Foa 1991	18.09	7.13	11	19.5	7.18	10	1.3%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	3.0%	-1.10 [-1.48 , -0.71]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	1.5%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	2.7%	-0.08 [-0.51 , 0.36]
Kelley 2012	-1.64	4.52	11	0.67	4	9	1.2%	-0.51 [-1.41 , 0.38]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	1.4%	-0.81 [-1.63 , 0.00]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	2.7%	-0.39 [-0.83 , 0.05]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	2.5%	0.03 [-0.46 , 0.52]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	1.3%	-0.14 [-1.00 , 0.72]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	1.3%	0.34 [-0.55 , 1.22]
Peck 1976	-12.5	4.25	4	-7	3.27	4	0.5%	-1.26 [-2.90 , 0.37]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	1.5%	-0.12 [-0.89 , 0.65]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.0%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	1.7%	-0.07 [-0.79 , 0.65]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	1.7%	-0.43 [-1.15 , 0.28]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.1%	-2.91 [-3.85 , -1.97]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	2.5%	-0.08 [-0.56 , 0.40]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	0.8%	-0.88 [-2.11 , 0.35]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	1.7%	-0.11 [-0.82 , 0.60]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	1.1%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**645**			**611**	**50.1%**	**-0.42 [-0.62 , -0.21]**
Heterogeneity: Tau² = 0.18; Chi² = 75.68, df = 28 (P < 0.00001); I² = 63%								
Test for overall effect: Z = 3.98 (P < 0.0001)								
16.15.2 Fulfil symptoms of disorder ICD/DSM while not stating systems								
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	1.1%	-0.10 [-1.05 , 0.85]
Borden 1986	29.7	17.5	10	32	26	10	1.3%	-0.10 [-0.98 , 0.78]
Bramston 1985	-78	14.21	12	-72	13.82	12	1.4%	-0.41 [-1.22 , 0.40]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	1.5%	-0.65 [-1.43 , 0.13]
Kilmann 1987	69	27	4	90	14	4	0.5%	-0.85 [-2.35 , 0.65]
Lick 1977	65.86	25.45	10	62.84	31.78	10	1.3%	0.10 [-0.78 , 0.98]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	1.2%	-0.04 [-0.93 , 0.86]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	2.3%	-0.24 [-0.78 , 0.30]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	1.6%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	1.3%	-1.10 [-1.97 , -0.23]
Turner 1979	-4.1	2.644	10	-4	2.449	10	1.3%	-0.04 [-0.91 , 0.84]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	1.1%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	1.7%	-0.10 [-0.81 , 0.62]
Subtotal (95% CI)			**164**			**162**	**18.6%**	**-0.30 [-0.52 , -0.08]**
Heterogeneity: Tau² = 0.00; Chi² = 7.12, df = 13 (P = 0.90); I² = 0%								
Test for overall effect: Z = 2.65 (P = 0.008)								
16.15.3 Population is classified as having a mental disorder, but full diagnostic criteria not reported								
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	1.2%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	1.3%	-0.25 [-1.09 , 0.59]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	1.5%	-0.33 [-1.12 , 0.46]
Freire 2007	24.6	11	7	28.2	18	9	1.1%	-0.22 [-1.21 , 0.77]
Fuchs 1977	14.3	7	10	21.4	7	10	1.1%	-0.97 [-1.91 , -0.03]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	1.5%	-0.53 [-1.30 , 0.25]

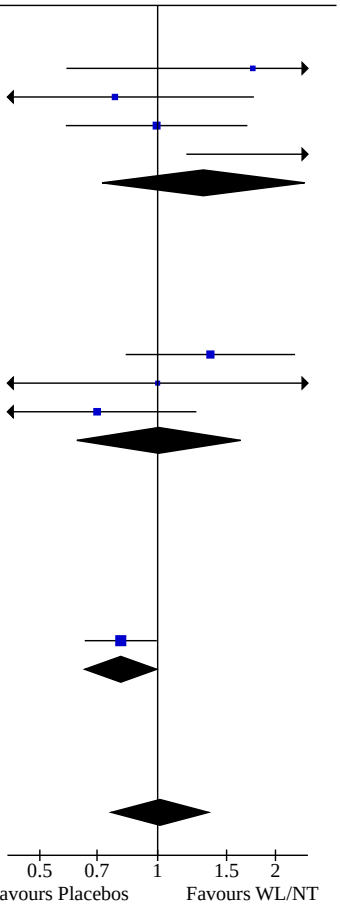
Analysis 16.15. (Continued)

Fuchs 1977	14.3	7	10	21.4	7	10	1.1%	-0.97 [-1.91 , -0.03]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	1.5%	-0.53 [-1.30 , 0.25]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	1.2%	-0.59 [-1.49 , 0.31]
Klein 1977	10.17	5.49	6	15	7.32	6	0.8%	-0.69 [-1.87 , 0.49]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	1.3%	-0.35 [-1.24 , 0.53]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	1.6%	-0.48 [-1.21 , 0.25]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	1.1%	-0.71 [-1.67 , 0.25]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	1.3%	-0.45 [-1.32 , 0.42]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	1.7%	-0.17 [-0.87 , 0.53]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	1.1%	-0.33 [-1.27 , 0.61]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.1%	0.17 [-0.81 , 1.16]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	1.4%	-1.33 [-2.16 , -0.49]
Powers 2004	46	20.09	12	41.6	13.15	15	1.5%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	1.4%	-0.10 [-0.90 , 0.70]
Powers 2008b	64	20	18	51	37	15	1.7%	0.44 [-0.26 , 1.13]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	3.9%	-0.11 [-0.31 , 0.09]
Tan 1986	9.5	7	10	13.44	7	9	1.2%	-0.54 [-1.46 , 0.38]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	1.2%	-0.56 [-1.48 , 0.36]
Subtotal (95% CI)			**437**			**427**	**31.3%**	**-0.27 [-0.42 , -0.12]**
Heterogeneity: Tau² = 0.01; Chi² = 21.98, df = 21 (P = 0.40); I² = 4%								
Test for overall effect: Z = 3.52 (P = 0.0004)								
Total (95% CI)			**1246**			**1200**	**100.0%**	**-0.37 [-0.49 , -0.25]**
Heterogeneity: Tau² = 0.08; Chi² = 108.28, df = 64 (P = 0.0005); I² = 41%								
Test for overall effect: Z = 6.16 (P < 0.00001)								
Test for subgroup differences: Chi² = 1.27, df = 2 (P = 0.53), I² = 0%								



Analysis 16.16. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 16: Mental health diagnoses

Study or Subgroup	All placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio IV, Random, 95% CI
16.16.1 Formal diagnosis according to DSM/ICD						
Klerman 1974a	7	25	4	25	5.7%	1.75 [0.58 , 5.24]
Klerman 1974b	7	25	9	25	9.1%	0.78 [0.34 , 1.76]
Rabkin 1990	14	27	12	23	15.9%	0.99 [0.58 , 1.70]
Watzl 1988	11	34	3	36	5.0%	3.88 [1.18 , 12.73]
Subtotal (95% CI)		**111**		**109**	**35.7%**	**1.31 [0.72 , 2.37]**
Total events:	39		28			
Heterogeneity: Tau² = 0.17; Chi² = 5.75, df = 3 (P = 0.12); I² = 48%						
Test for overall effect: Z = 0.89 (P = 0.38)						
16.16.2 Fulfill symptoms of disorder ICD/DSM while not stating systems						
Berg 1983	9	11	9	15	17.1%	1.36 [0.83 , 2.24]
Double 1993	3	22	3	22	3.4%	1.00 [0.23 , 4.42]
Whittaker 1963	7	13	10	13	14.3%	0.70 [0.39 , 1.26]
Subtotal (95% CI)		**46**		**50**	**34.7%**	**1.01 [0.62 , 1.63]**
Total events:	19		22			
Heterogeneity: Tau² = 0.06; Chi² = 2.90, df = 2 (P = 0.24); I² = 31%						
Test for overall effect: Z = 0.03 (P = 0.98)						
16.16.3 Population is classified as having a mental disorder, but full diagnostic criteria not reported						
Mealiea 1971	0	9	0	10		Not estimable
Wilson 1980	31	40	10	10	29.6%	0.80 [0.65 , 0.99]
Subtotal (95% CI)		**49**		**20**	**29.6%**	**0.80 [0.65 , 0.99]**
Total events:	31		10			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.01 (P = 0.04)						
Total (95% CI)		**206**		**179**	**100.0%**	**1.01 [0.76 , 1.35]**
Total events:	89		60			
Heterogeneity: Tau² = 0.06; Chi² = 11.95, df = 7 (P = 0.10); I² = 41%						
Test for overall effect: Z = 0.09 (P = 0.93)						
Test for subgroup differences: Chi² = 2.67, df = 2 (P = 0.26), I² = 25.2%						

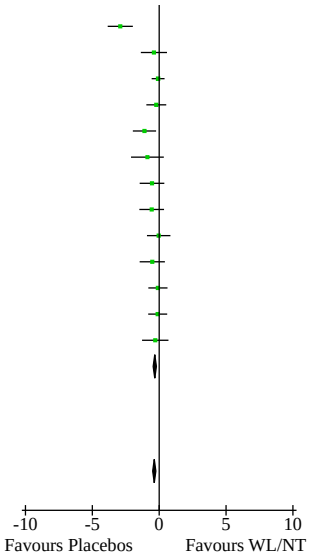


Analysis 16.17. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 17: Affiliation bias

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
16.17.1 Risk of affiliation, industry, and allegiance bias								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	2.5%	-0.45 [-0.93 , 0.03]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	2.0%	-0.11 [-0.72 , 0.49]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	2.3%	-0.64 [-1.16 , -0.12]
Foa 1991	18.09	7.13	11	19.5	7.18	10	1.3%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	3.0%	-1.10 [-1.48 , -0.71]
Fuchs 1977	14.3	7	10	21.4	7	10	1.1%	-0.97 [-1.91 , -0.03]
Subtotal (95% CI)			**214**			**149**	**12.3%**	**-0.61 [-0.95 , -0.26]**
Heterogeneity: Tau² = 0.09; Chi² = 10.43, df = 5 (P = 0.06); I² = 52%								
Test for overall effect: Z = 3.47 (P = 0.0005)								
16.17.2 No risk found of affiliation, industry, and allegiance bias								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	1.3%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	2.8%	-0.72 [-1.15 , -0.29]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	2.2%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	1.1%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	1.2%	-1.71 [-2.60 , -0.82]
Borden 1986	29.7	17.5	10	32	26	10	1.3%	-0.10 [-0.98 , 0.78]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	1.2%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	1.3%	-0.25 [-1.09 , 0.59]
Bramston 1985	-78	14.21	12	-72	13.82	12	1.4%	-0.41 [-1.22 , 0.40]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	1.2%	-0.09 [-1.01 , 0.84]
Carter 2003	3.8	1.2	28	3.8	1.3	29	2.3%	0.00 [-0.52 , 0.52]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	1.5%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	1.5%	-0.33 [-1.12 , 0.46]
Freire 2007	24.6	11	7	28.2	18	9	1.1%	-0.22 [-1.21 , 0.77]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	1.5%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	2.7%	-0.08 [-0.51 , 0.36]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	1.5%	-0.53 [-1.30 , 0.25]
Kelley 2012	-1.64	4.52	11	0.67	4	9	1.2%	-0.51 [-1.41 , 0.38]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	1.2%	-0.59 [-1.49 , 0.31]
Kilmann 1987	69	27	4	90	14	4	0.5%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	0.8%	-0.69 [-1.87 , 0.49]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	1.4%	-0.81 [-1.63 , 0.00]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	1.3%	-0.35 [-1.24 , 0.53]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	2.7%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	1.6%	-0.48 [-1.21 , 0.25]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	2.5%	0.03 [-0.46 , 0.52]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	1.3%	-0.14 [-1.00 , 0.72]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	1.1%	-0.71 [-1.67 , 0.25]
Lick 1977	65.86	25.45	10	62.84	31.78	10	1.3%	0.10 [-0.78 , 0.98]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	1.3%	0.34 [-0.55 , 1.22]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	1.2%	-0.04 [-0.93 , 0.86]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	1.3%	-0.45 [-1.32 , 0.42]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	1.7%	-0.17 [-0.87 , 0.53]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	1.1%	-0.33 [-1.27 , 0.61]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.1%	0.17 [-0.81 , 1.16]
Peck 1976	-12.5	4.25	4	-7	3.27	4	0.5%	-1.26 [-2.90 , 0.37]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	1.5%	-0.12 [-0.89 , 0.65]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	1.4%	-1.33 [-2.16 , -0.49]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.0%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	1.7%	-0.07 [-0.79 , 0.65]
Powers 2004	46	20.09	12	41.6	13.15	15	1.5%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	1.4%	-0.10 [-0.90 , 0.70]
Powers 2008b	64	20	18	51	37	15	1.7%	0.44 [-0.26 , 1.13]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	3.9%	-0.11 [-0.31 , 0.09]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	2.3%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	1.7%	-0.43 [-1.15 , 0.28]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]

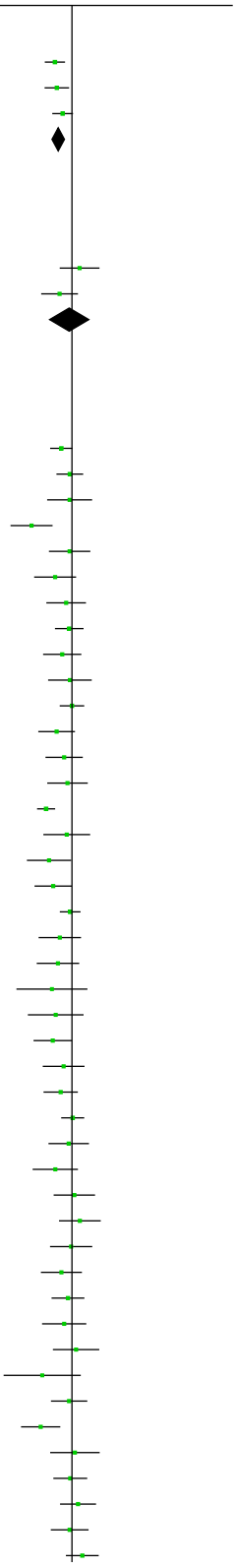
Analysis 16.17. (Continued)

Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	2.5%	-0.08 [-0.56 , 0.40]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	1.6%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	1.3%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	0.8%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	1.2%	-0.54 [-1.46 , 0.38]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	1.2%	-0.56 [-1.48 , 0.36]
Turner 1979	-4.1	2.644	10	-4	2.449	10	1.3%	-0.04 [-0.91 , 0.84]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	1.1%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	1.7%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	1.7%	-0.11 [-0.82 , 0.60]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	1.1%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**1032**			**1051**	**87.7%**	**-0.33 [-0.45 , -0.21]**
Heterogeneity: Tau² = 0.06; Chi² = 87.74, df = 58 (P = 0.007); I² = 34%								
Test for overall effect: Z = 5.41 (P < 0.00001)								
Total (95% CI)			**1246**			**1200**	**100.0%**	**-0.37 [-0.49 , -0.25]**
Heterogeneity: Tau² = 0.08; Chi² = 108.28, df = 64 (P = 0.0005); I² = 41%								
Test for overall effect: Z = 6.16 (P < 0.00001)								
Test for subgroup differences: Chi² = 2.22, df = 1 (P = 0.14), I² = 55.0%								



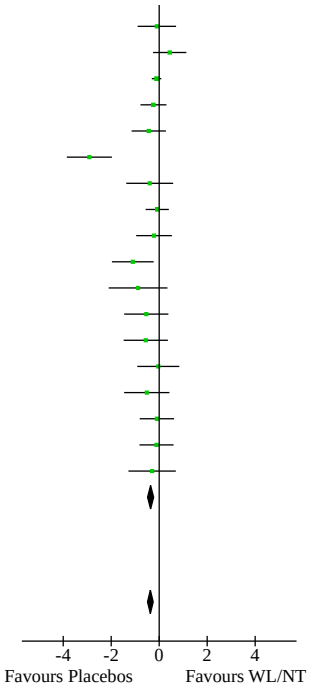
Analysis 16.18. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 18: Risk of bias (participants and personnel excluded)

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
16.18.1 Low risk of bias								
Allen 2006	12	9.6	45	19	9.6	44	2.8%	-0.72 [-1.15 , -0.29]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	2.3%	-0.64 [-1.16 , -0.12]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	2.7%	-0.39 [-0.83 , 0.05]
Subtotal (95% CI)			**116**			**114**	**7.8%**	**-0.58 [-0.85 , -0.32]**
Heterogeneity: Tau² = 0.00; Chi² = 1.16, df = 2 (P = 0.56); I² = 0%								
Test for overall effect: Z = 4.32 (P < 0.0001)								
16.18.2 Unclear risk of bias								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	1.3%	0.32 [-0.52 , 1.17]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	1.5%	-0.53 [-1.30 , 0.25]
Subtotal (95% CI)			**30**			**21**	**2.8%**	**-0.12 [-0.95 , 0.71]**
Heterogeneity: Tau² = 0.19; Chi² = 2.10, df = 1 (P = 0.15); I² = 52%								
Test for overall effect: Z = 0.27 (P = 0.78)								
16.18.3 High risk of bias								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	2.5%	-0.45 [-0.93 , 0.03]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	2.2%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	1.1%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	1.2%	-1.71 [-2.60 , -0.82]
Borden 1986	29.7	17.5	10	32	26	10	1.3%	-0.10 [-0.98 , 0.78]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	1.2%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	1.3%	-0.25 [-1.09 , 0.59]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	2.0%	-0.11 [-0.72 , 0.49]
Bramston 1985	-78	14.21	12	-72	13.82	12	1.4%	-0.41 [-1.22 , 0.40]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	1.2%	-0.09 [-1.01 , 0.84]
Carter 2003	3.8	1.2	28	3.8	1.3	29	2.3%	0.00 [-0.52 , 0.52]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	1.5%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	1.5%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	1.3%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	3.0%	-1.10 [-1.48 , -0.71]
Freire 2007	24.6	11	7	28.2	18	9	1.1%	-0.22 [-1.21 , 0.77]
Fuchs 1977	14.3	7	10	21.4	7	10	1.1%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	1.5%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	2.7%	-0.08 [-0.51 , 0.36]
Kelley 2012	-1.64	4.52	11	0.67	4	9	1.2%	-0.51 [-1.41 , 0.38]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	1.2%	-0.59 [-1.49 , 0.31]
Kilmann 1987	69	27	4	90	14	4	0.5%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	0.8%	-0.69 [-1.87 , 0.49]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	1.4%	-0.81 [-1.63 , 0.00]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	1.3%	-0.35 [-1.24 , 0.53]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	1.6%	-0.48 [-1.21 , 0.25]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	2.5%	0.03 [-0.46 , 0.52]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	1.3%	-0.14 [-1.00 , 0.72]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	1.1%	-0.71 [-1.67 , 0.25]
Lick 1977	65.86	25.45	10	62.84	31.78	10	1.3%	0.10 [-0.78 , 0.98]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	1.3%	0.34 [-0.55 , 1.22]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	1.2%	-0.04 [-0.93 , 0.86]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	1.3%	-0.45 [-1.32 , 0.42]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	1.7%	-0.17 [-0.87 , 0.53]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	1.1%	-0.33 [-1.27 , 0.61]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.1%	0.17 [-0.81 , 1.16]
Peck 1976	-12.5	4.25	4	-7	3.27	4	0.5%	-1.26 [-2.90 , 0.37]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	1.5%	-0.12 [-0.89 , 0.65]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	1.4%	-1.33 [-2.16 , -0.49]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.0%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	1.7%	-0.07 [-0.79 , 0.65]
Powers 2004	46	20.09	12	41.6	13.15	15	1.5%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	1.4%	-0.10 [-0.90 , 0.70]
Powers 2008b	64	20	18	51	37	15	1.7%	0.44 [-0.26 , 1.13]



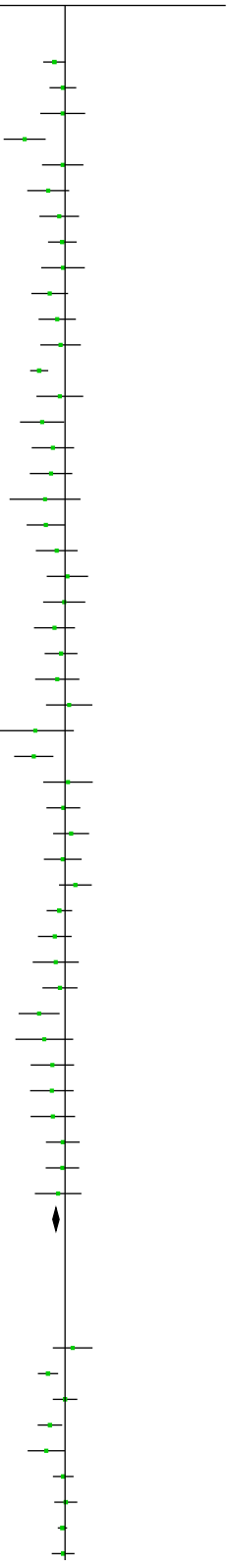
Analysis 16.18. (Continued)

Powers 2008a	66	36	15	69	16	10	1.4%	-0.10 [-0.90 , 0.70]
Powers 2008b	64	20	18	51	37	15	1.7%	0.44 [-0.26 , 1.13]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	3.9%	-0.11 [-0.31 , 0.09]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	2.3%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	1.7%	-0.43 [-1.15 , 0.28]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	2.5%	-0.08 [-0.56 , 0.40]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	1.6%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	1.3%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	0.8%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	1.2%	-0.54 [-1.46 , 0.38]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	1.2%	-0.56 [-1.48 , 0.36]
Turner 1979	-4.1	2.644	10	-4	2.449	10	1.3%	-0.04 [-0.91 , 0.84]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	1.1%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	1.7%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	1.7%	-0.11 [-0.82 , 0.60]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	1.1%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**1100**			**1065**	**89.4%**	**-0.36 [-0.48 , -0.23]**
Heterogeneity: Tau² = 0.09; Chi² = 100.89, df = 59 (P = 0.0006); I² = 42%								
Test for overall effect: Z = 5.58 (P < 0.00001)								
Total (95% CI)			**1246**			**1200**	**100.0%**	**-0.37 [-0.49 , -0.25]**
Heterogeneity: Tau² = 0.08; Chi² = 108.28, df = 64 (P = 0.0005); I² = 41%								
Test for overall effect: Z = 6.16 (P < 0.00001)								
Test for subgroup differences: Chi² = 2.70, df = 2 (P = 0.26), I² = 26.0%								



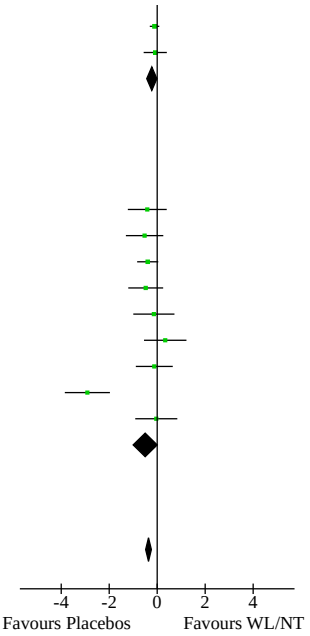
Analysis 16.19. Comparison 16: Subgroup analyses for all placebos compared with wait-list or no-treatment for continuous data, Outcome 19: Imputed data

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
16.19.1 Available data								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	2.6%	-0.45 [-0.93 , 0.03]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	2.2%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	1.2%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	1.3%	-1.71 [-2.60 , -0.82]
Borden 1986	29.7	17.5	10	32	26	10	1.3%	-0.10 [-0.98 , 0.78]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	1.3%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	1.4%	-0.25 [-1.09 , 0.59]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	2.0%	-0.11 [-0.72 , 0.49]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	1.2%	-0.09 [-1.01 , 0.84]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	1.5%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	1.5%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	1.3%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	3.0%	-1.10 [-1.48 , -0.71]
Freire 2007	24.6	11	7	28.2	18	9	1.1%	-0.22 [-1.21 , 0.77]
Fuchs 1977	14.3	7	10	21.4	7	10	1.2%	-0.97 [-1.91 , -0.03]
Kelley 2012	-1.64	4.52	11	0.67	4	9	1.3%	-0.51 [-1.41 , 0.38]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	1.2%	-0.59 [-1.49 , 0.31]
Kilmann 1987	69	27	4	90	14	4	0.5%	-0.85 [-2.35 , 0.65]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	1.4%	-0.81 [-1.63 , 0.00]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	1.3%	-0.35 [-1.24 , 0.53]
Lick 1977	65.86	25.45	10	62.84	31.78	10	1.3%	0.10 [-0.78 , 0.98]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	1.3%	-0.04 [-0.93 , 0.86]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	1.3%	-0.45 [-1.32 , 0.42]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	1.8%	-0.17 [-0.87 , 0.53]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	1.2%	-0.33 [-1.27 , 0.61]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.1%	0.17 [-0.81 , 1.16]
Peck 1976	-12.5	4.25	4	-7	3.27	4	0.5%	-1.26 [-2.90 , 0.37]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	1.4%	-1.33 [-2.16 , -0.49]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.0%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	1.7%	-0.07 [-0.79 , 0.65]
Powers 2004	46	20.09	12	41.6	13.15	15	1.6%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	1.5%	-0.10 [-0.90 , 0.70]
Powers 2008b	64	20	18	51	37	15	1.8%	0.44 [-0.26 , 1.13]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	2.3%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	1.7%	-0.43 [-1.15 , 0.28]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	1.6%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	1.3%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	0.8%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	1.2%	-0.54 [-1.46 , 0.38]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	1.2%	-0.56 [-1.48 , 0.36]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	1.2%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	1.7%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	1.7%	-0.11 [-0.82 , 0.60]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	1.1%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**664**			**572**	**63.9%**	**-0.39 [-0.52 , -0.25]**
Heterogeneity: Tau² = 0.04; Chi² = 54.31, df = 44 (P = 0.14); I² = 19%								
Test for overall effect: Z = 5.70 (P < 0.00001)								
16.19.2 Intent-to-treat								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	1.4%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	2.8%	-0.72 [-1.15 , -0.29]
Carter 2003	3.8	1.2	28	3.8	1.3	29	2.4%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	2.4%	-0.64 [-1.16 , -0.12]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	1.5%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	2.7%	-0.08 [-0.51 , 0.36]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	2.5%	0.03 [-0.46 , 0.52]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	3.9%	-0.11 [-0.31 , 0.09]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	2.5%	-0.08 [-0.56 , 0.40]



Analysis 16.19. (Continued)

Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	3.9%	-0.11 [-0.31 , 0.09]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	2.5%	-0.08 [-0.56 , 0.40]
Subtotal (95% CI)			**418**			**473**	**22.1%**	**-0.22 [-0.44 , -0.01]**
Heterogeneity: Tau² = 0.04; Chi² = 14.89, df = 8 (P = 0.06); I² = 46%								
Test for overall effect: Z = 2.08 (P = 0.04)								
16.19.3 No attrition								
Bramston 1985	-78	14.21	12	-72	13.82	12	1.4%	-0.41 [-1.22 , 0.40]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	1.5%	-0.53 [-1.30 , 0.25]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	2.7%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	1.7%	-0.48 [-1.21 , 0.25]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	1.3%	-0.14 [-1.00 , 0.72]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	1.3%	0.34 [-0.55 , 1.22]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	1.5%	-0.12 [-0.89 , 0.65]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.2%	-2.91 [-3.85 , -1.97]
Turner 1979	-4.1	2.644	10	-4	2.449	10	1.3%	-0.04 [-0.91 , 0.84]
Subtotal (95% CI)			**149**			**140**	**14.0%**	**-0.50 [-1.00 , 0.00]**
Heterogeneity: Tau² = 0.43; Chi² = 31.53, df = 8 (P = 0.0001); I² = 75%								
Test for overall effect: Z = 1.96 (P = 0.05)								
Total (95% CI)			**1231**			**1185**	**100.0%**	**-0.36 [-0.48 , -0.24]**
Heterogeneity: Tau² = 0.08; Chi² = 107.34, df = 62 (P = 0.0003); I² = 42%								
Test for overall effect: Z = 5.95 (P < 0.00001)								
Test for subgroup differences: Chi² = 1.99, df = 2 (P = 0.37), I² = 0%								



Comparison 17. Subgroup analyses for psychological placebos compared with wait-list or no-treatment for continuous data

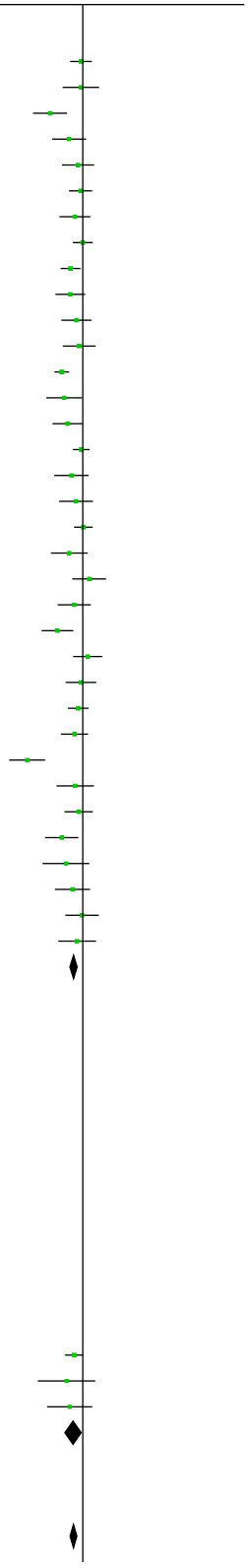
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
17.1 Type of active intervention	38	1263	Std. Mean Difference (IV, Random, 95% CI)	-0.49 [-0.66, -0.31]
17.1.1 Psychological intervention	35	1174	Std. Mean Difference (IV, Random, 95% CI)	-0.48 [-0.67, -0.29]
17.1.2 Pharmacological intervention	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
17.1.3 Physical intervention	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
17.1.4 Other or combination intervention	3	89	Std. Mean Difference (IV, Random, 95% CI)	-0.51 [-0.94, -0.09]
17.2 Risk of bias	38	1263	Std. Mean Difference (IV, Random, 95% CI)	-0.49 [-0.66, -0.31]
17.2.1 Low risk of bias	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
17.2.2 High risk of bias	38	1263	Std. Mean Difference (IV, Random, 95% CI)	-0.49 [-0.66, -0.31]
17.3 Type of outcome domain	34	1515	Std. Mean Difference (IV, Random, 95% CI)	-0.49 [-0.68, -0.30]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
17.3.1 Blinded observer-reported	14	512	Std. Mean Difference (IV, Random, 95% CI)	-0.72 [-1.08, -0.36]
17.3.2 Non-blinded observer-reported	2	96	Std. Mean Difference (IV, Random, 95% CI)	-0.54 [-0.95, -0.13]
17.3.3 Patient-reported	18	907	Std. Mean Difference (IV, Random, 95% CI)	-0.27 [-0.47, -0.08]
17.4 Awareness of placebo intervention	38	1263	Std. Mean Difference (IV, Random, 95% CI)	-0.49 [-0.66, -0.31]
17.4.1 Open placebo	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
17.4.2 Closed placebo	38	1263	Std. Mean Difference (IV, Random, 95% CI)	-0.49 [-0.66, -0.31]
17.5 The trial objective	38	1263	Std. Mean Difference (IV, Random, 95% CI)	-0.49 [-0.66, -0.31]
17.5.1 A trial's objective is clearly to assess the effects of placebo	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
17.5.2 No such objectives are stated	38	1263	Std. Mean Difference (IV, Random, 95% CI)	-0.49 [-0.66, -0.31]
17.6 Mean age of participants	33	1537	Std. Mean Difference (IV, Random, 95% CI)	-0.45 [-0.64, -0.26]
17.6.1 Below 18 years	4	148	Std. Mean Difference (IV, Random, 95% CI)	-0.85 [-2.02, 0.32]
17.6.2 18 to 50 years	25	1233	Std. Mean Difference (IV, Random, 95% CI)	-0.38 [-0.55, -0.21]
17.6.3 Above 50 years	4	156	Std. Mean Difference (IV, Random, 95% CI)	-0.59 [-1.33, 0.14]
17.7 Duration of treatment	38	1576	Std. Mean Difference (IV, Random, 95% CI)	-0.48 [-0.66, -0.31]
17.7.1 Above 3 months	6	270	Std. Mean Difference (IV, Random, 95% CI)	-1.02 [-1.68, -0.36]
17.7.2 Below 3 months	32	1306	Std. Mean Difference (IV, Random, 95% CI)	-0.38 [-0.53, -0.22]
17.8 Type of psychological placebo	38	1263	Std. Mean Difference (IV, Random, 95% CI)	-0.49 [-0.66, -0.31]
17.8.1 Interaction placebo	15	613	Std. Mean Difference (IV, Random, 95% CI)	-0.70 [-1.04, -0.36]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
17.8.2 Psychoeducational placebo	9	329	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.36, 0.07]
17.8.3 Exposure placebo	14	321	Std. Mean Difference (IV, Random, 95% CI)	-0.48 [-0.71, -0.24]
17.9 Mode of psychological placebo	35	1571	Std. Mean Difference (IV, Random, 95% CI)	-0.48 [-0.66, -0.30]
17.9.1 Individual treatment	19	1151	Std. Mean Difference (IV, Random, 95% CI)	-0.31 [-0.49, -0.12]
17.9.2 Group treatment	13	308	Std. Mean Difference (IV, Random, 95% CI)	-0.85 [-1.30, -0.39]
17.9.3 Combination of individual and group	3	112	Std. Mean Difference (IV, Random, 95% CI)	-0.45 [-0.83, -0.08]
17.10 Mental health diagnoses	39	1656	Std. Mean Difference (IV, Random, 95% CI)	-0.47 [-0.64, -0.30]
17.10.1 Formal diagnosis according to DSM/ICD	16	767	Std. Mean Difference (IV, Random, 95% CI)	-0.53 [-0.86, -0.21]
17.10.2 Fulfill symptoms of disorder ICD/DSM while not stating systems	9	218	Std. Mean Difference (IV, Random, 95% CI)	-0.39 [-0.66, -0.12]
17.10.3 Population is classified as having a mental disorder, but full diagnostic criteria not reported	14	671	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.59, -0.15]
17.11 Affiliation bias	65	2446	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.49, -0.25]
17.11.1 Risk of affiliation, industry, allegiance bias	6	363	Std. Mean Difference (IV, Random, 95% CI)	-0.61 [-0.95, -0.26]
17.11.2 No risk of affiliation, industry, allegiance bias	59	2083	Std. Mean Difference (IV, Random, 95% CI)	-0.33 [-0.45, -0.21]
17.12 Imputed data	37	1626	Std. Mean Difference (IV, Random, 95% CI)	-0.46 [-0.63, -0.29]
17.12.1 Available data	27	842	Std. Mean Difference (IV, Random, 95% CI)	-0.50 [-0.68, -0.33]
17.12.2 Intent-to-treat	6	682	Std. Mean Difference (IV, Random, 95% CI)	-0.18 [-0.39, 0.03]
17.12.3 No attrition	4	102	Std. Mean Difference (IV, Random, 95% CI)	-0.75 [-2.11, 0.62]

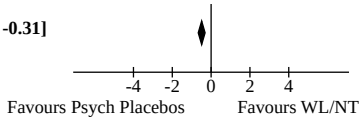
Analysis 17.1. Comparison 17: Subgroup analyses for psychological placebos compared with wait-list or no-treatment for continuous data, Outcome 1: Type of active intervention

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
17.1.1 Psychological intervention								
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	3.5%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	2.1%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	2.3%	-1.71 [-2.60 , -0.82]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	2.3%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	2.4%	-0.25 [-1.09 , 0.59]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	3.3%	-0.11 [-0.72 , 0.49]
Bramston 1985	-78	14.21	12	-72	13.82	12	2.5%	-0.41 [-1.22 , 0.40]
Carter 2003	3.8	1.2	28	3.8	1.3	29	3.7%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	3.7%	-0.64 [-1.16 , -0.12]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	2.6%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	2.6%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	2.4%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	4.3%	-1.10 [-1.48 , -0.71]
Fuchs 1977	14.3	7	10	21.4	7	10	2.1%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	2.6%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	4.1%	-0.08 [-0.51 , 0.36]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	2.3%	-0.59 [-1.49 , 0.31]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	2.3%	-0.35 [-1.24 , 0.53]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	3.8%	0.03 [-0.46 , 0.52]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	2.1%	-0.71 [-1.67 , 0.25]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	2.3%	0.34 [-0.55 , 1.22]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	2.3%	-0.45 [-1.32 , 0.42]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	2.5%	-1.33 [-2.16 , -0.49]
Powers 2004	46	20.09	12	41.6	13.15	15	2.7%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	2.6%	-0.10 [-0.90 , 0.70]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	3.6%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	2.9%	-0.43 [-1.15 , 0.28]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	2.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	2.0%	-0.39 [-1.37 , 0.58]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	2.8%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	2.3%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	1.5%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	2.2%	-0.54 [-1.46 , 0.38]
Turner 1979	-4.1	2.644	10	-4	2.449	10	2.3%	-0.04 [-0.91 , 0.84]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	2.0%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**628**			**546**	**93.4%**	**-0.48 [-0.67 , -0.29]**
Heterogeneity: Tau² = 0.17; Chi² = 77.70, df = 34 (P < 0.0001); I² = 56%								
Test for overall effect: Z = 4.99 (P < 0.00001)								
17.1.2 Pharmacological intervention								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
17.1.3 Physical intervention								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
17.1.4 Other or combination intervention								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	3.9%	-0.45 [-0.93 , 0.03]
Kilmann 1987	69	27	4	90	14	4	1.1%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	1.6%	-0.69 [-1.87 , 0.49]
Subtotal (95% CI)			**45**			**44**	**6.6%**	**-0.51 [-0.94 , -0.09]**
Heterogeneity: Tau² = 0.00; Chi² = 0.35, df = 2 (P = 0.84); I² = 0%								
Test for overall effect: Z = 2.36 (P = 0.02)								
Total (95% CI)			**673**			**590**	**100.0%**	**-0.49 [-0.66 , -0.31]**
Heterogeneity: Tau² = 0.15; Chi² = 78.11, df = 37 (P < 0.0001); I² = 53%								



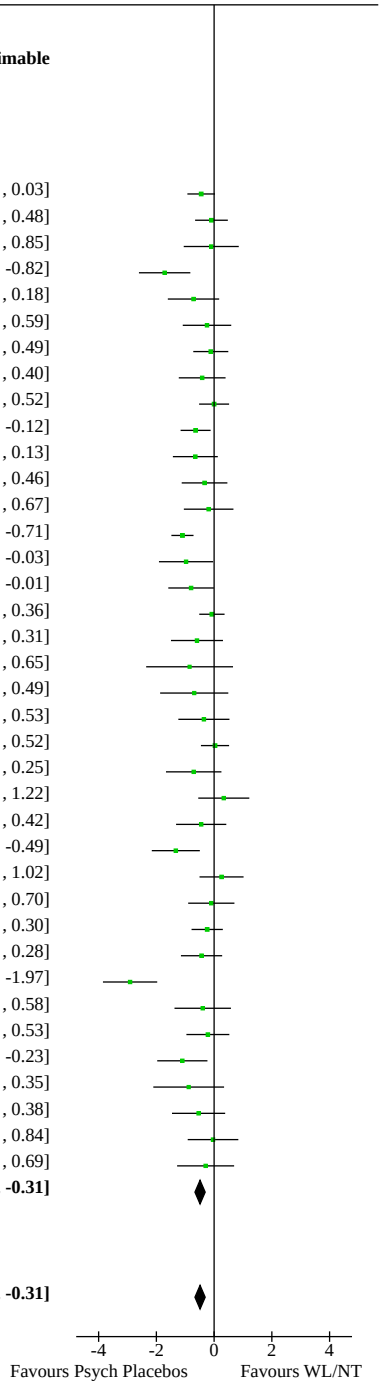
Analysis 17.1. (Continued)

Total (95% CI)	**673**	**590**	**100.0%**	**-0.49 [-0.66 , -0.31]**
Heterogeneity: Tau² = 0.15; Chi² = 78.11, df = 37 (P < 0.0001); I² = 53%				
Test for overall effect: Z = 5.40 (P < 0.00001)				
Test for subgroup differences: Chi² = 0.01, df = 1 (P = 0.91), I² = 0%				



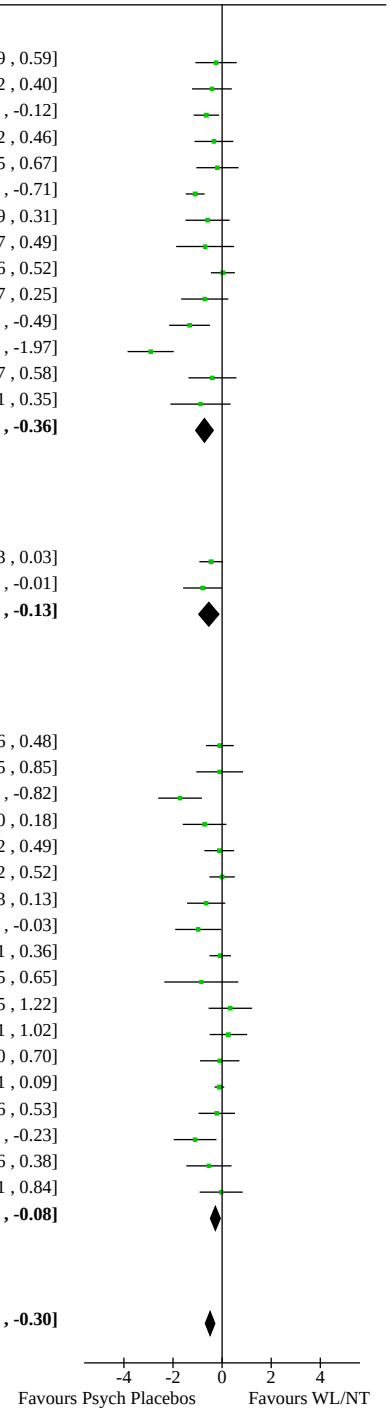
Analysis 17.2. Comparison 17: Subgroup analyses for psychological placebos compared with wait-list or no-treatment for continuous data, Outcome 2: Risk of bias

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
17.2.1 Low risk of bias								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
17.2.2 High risk of bias								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	3.9%	-0.45 [-0.93 , 0.03]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	3.5%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	2.1%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	2.3%	-1.71 [-2.60 , -0.82]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	2.3%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	2.4%	-0.25 [-1.09 , 0.59]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	3.3%	-0.11 [-0.72 , 0.49]
Bramston 1985	-78	14.21	12	-72	13.82	12	2.5%	-0.41 [-1.22 , 0.40]
Carter 2003	3.8	1.2	28	3.8	1.3	29	3.7%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	3.7%	-0.64 [-1.16 , -0.12]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	2.6%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	2.6%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	2.4%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	4.3%	-1.10 [-1.48 , -0.71]
Fuchs 1977	14.3	7	10	21.4	7	10	2.1%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	2.6%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	4.1%	-0.08 [-0.51 , 0.36]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	2.3%	-0.59 [-1.49 , 0.31]
Kilmann 1987	69	27	4	90	14	4	1.1%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	1.6%	-0.69 [-1.87 , 0.49]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	2.3%	-0.35 [-1.24 , 0.53]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	3.8%	0.03 [-0.46 , 0.52]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	2.1%	-0.71 [-1.67 , 0.25]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	2.3%	0.34 [-0.55 , 1.22]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	2.3%	-0.45 [-1.32 , 0.42]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	2.5%	-1.33 [-2.16 , -0.49]
Powers 2004	46	20.09	12	41.6	13.15	15	2.7%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	2.6%	-0.10 [-0.90 , 0.70]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	3.6%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	2.9%	-0.43 [-1.15 , 0.28]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	2.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	2.0%	-0.39 [-1.37 , 0.58]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	2.8%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	2.3%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	1.5%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	2.2%	-0.54 [-1.46 , 0.38]
Turner 1979	-4.1	2.644	10	-4	2.449	10	2.3%	-0.04 [-0.91 , 0.84]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	2.0%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**673**			**590**	**100.0%**	**-0.49 [-0.66 , -0.31]**
Heterogeneity: Tau² = 0.15; Chi² = 78.11, df = 37 (P < 0.0001); I² = 53%								
Test for overall effect: Z = 5.40 (P < 0.00001)								
Total (95% CI)			**673**			**590**	**100.0%**	**-0.49 [-0.66 , -0.31]**
Heterogeneity: Tau² = 0.15; Chi² = 78.11, df = 37 (P < 0.0001); I² = 53%								
Test for overall effect: Z = 5.40 (P < 0.00001)								
Test for subgroup differences: Not applicable								



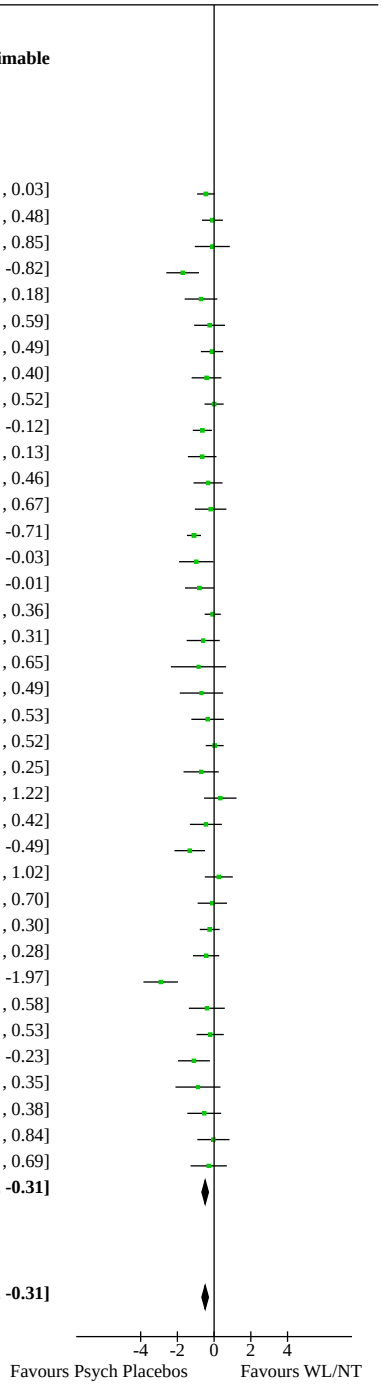
Analysis 17.3. Comparison 17: Subgroup analyses for psychological placebos compared with wait-list or no-treatment for continuous data, Outcome 3: Type of outcome domain

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
17.3.1 Blinded observer-reported								
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	2.7%	-0.25 [-1.09 , 0.59]
Bramston 1985	-78	14.21	12	-72	13.82	12	2.8%	-0.41 [-1.22 , 0.40]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	4.0%	-0.64 [-1.16 , -0.12]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	2.9%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	2.6%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	4.6%	-1.10 [-1.48 , -0.71]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	2.5%	-0.59 [-1.49 , 0.31]
Klein 1977	10.17	5.49	6	15	7.32	6	1.8%	-0.69 [-1.87 , 0.49]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	4.1%	0.03 [-0.46 , 0.52]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	2.3%	-0.71 [-1.67 , 0.25]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	2.7%	-1.33 [-2.16 , -0.49]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	2.4%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	2.3%	-0.39 [-1.37 , 0.58]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	1.7%	-0.88 [-2.11 , 0.35]
Subtotal (95% CI)			**293**			**219**	**39.2%**	**-0.72 [-1.08 , -0.36]**
Heterogeneity: Tau² = 0.29; Chi² = 40.54, df = 13 (P = 0.0001); I² = 68%								
Test for overall effect: Z = 3.93 (P < 0.0001)								
17.3.2 Non-blinded observer-reported								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	4.1%	-0.45 [-0.93 , 0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	2.9%	-0.80 [-1.59 , -0.01]
Subtotal (95% CI)			**48**			**48**	**7.0%**	**-0.54 [-0.95 , -0.13]**
Heterogeneity: Tau² = 0.00; Chi² = 0.55, df = 1 (P = 0.46); I² = 0%								
Test for overall effect: Z = 2.60 (P = 0.009)								
17.3.3 Patient-reported								
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	3.7%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	2.3%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	2.5%	-1.71 [-2.60 , -0.82]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	2.5%	-0.71 [-1.60 , 0.18]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	3.6%	-0.11 [-0.72 , 0.49]
Carter 2003	3.8	1.2	28	3.8	1.3	29	4.0%	0.00 [-0.52 , 0.52]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	2.9%	-0.65 [-1.43 , 0.13]
Fuchs 1977	14.3	7	10	21.4	7	10	2.4%	-0.97 [-1.91 , -0.03]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	4.3%	-0.08 [-0.51 , 0.36]
Kilmann 1987	69	27	4	90	14	4	1.3%	-0.85 [-2.35 , 0.65]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	2.5%	0.34 [-0.55 , 1.22]
Powers 2004	46	20.09	12	41.6	13.15	15	3.0%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	2.8%	-0.10 [-0.90 , 0.70]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	5.3%	-0.11 [-0.31 , 0.09]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	3.0%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	2.6%	-1.10 [-1.97 , -0.23]
Tan 1986	9.5	7	10	13.44	7	9	2.4%	-0.54 [-1.46 , 0.38]
Turner 1979	-4.1	2.644	10	-4	2.449	10	2.6%	-0.04 [-0.91 , 0.84]
Subtotal (95% CI)			**454**			**453**	**53.8%**	**-0.27 [-0.47 , -0.08]**
Heterogeneity: Tau² = 0.06; Chi² = 26.52, df = 17 (P = 0.07); I² = 36%								
Test for overall effect: Z = 2.71 (P = 0.007)								
Total (95% CI)			**795**			**720**	**100.0%**	**-0.49 [-0.68 , -0.30]**
Heterogeneity: Tau² = 0.17; Chi² = 86.79, df = 33 (P < 0.00001); I² = 62%								
Test for overall effect: Z = 5.01 (P < 0.00001)								
Test for subgroup differences: Chi² = 5.02, df = 2 (P = 0.08), I² = 60.2%								



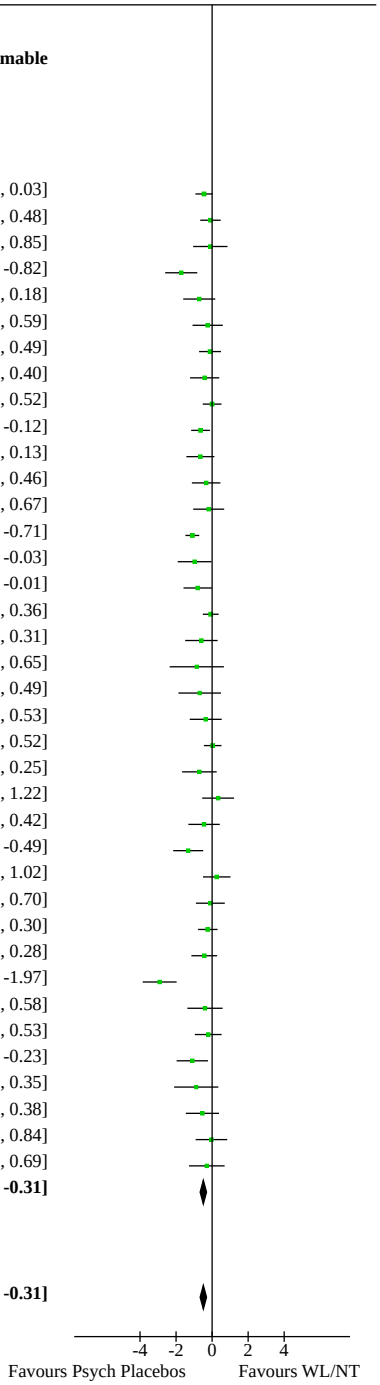
Analysis 17.4. Comparison 17: Subgroup analyses for psychological placebos compared with wait-list or no-treatment for continuous data, Outcome 4: Awareness of placebo intervention

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
17.4.1 Open placebo								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
17.4.2 Closed placebo								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	3.9%	-0.45 [-0.93 , 0.03]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	3.5%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	2.1%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	2.3%	-1.71 [-2.60 , -0.82]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	2.3%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	2.4%	-0.25 [-1.09 , 0.59]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	3.3%	-0.11 [-0.72 , 0.49]
Bramston 1985	-78	14.21	12	-72	13.82	12	2.5%	-0.41 [-1.22 , 0.40]
Carter 2003	3.8	1.2	28	3.8	1.3	29	3.7%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	3.7%	-0.64 [-1.16 , -0.12]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	2.6%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	2.6%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	2.4%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	4.3%	-1.10 [-1.48 , -0.71]
Fuchs 1977	14.3	7	10	21.4	7	10	2.1%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	2.6%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	4.1%	-0.08 [-0.51 , 0.36]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	2.3%	-0.59 [-1.49 , 0.31]
Kilmann 1987	69	27	4	90	14	4	1.1%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	1.6%	-0.69 [-1.87 , 0.49]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	2.3%	-0.35 [-1.24 , 0.53]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	3.8%	0.03 [-0.46 , 0.52]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	2.1%	-0.71 [-1.67 , 0.25]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	2.3%	0.34 [-0.55 , 1.22]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	2.3%	-0.45 [-1.32 , 0.42]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	2.5%	-1.33 [-2.16 , -0.49]
Powers 2004	46	20.09	12	41.6	13.15	15	2.7%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	2.6%	-0.10 [-0.90 , 0.70]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	3.6%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	2.9%	-0.43 [-1.15 , 0.28]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	2.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	2.0%	-0.39 [-1.37 , 0.58]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	2.8%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	2.3%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	1.5%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	2.2%	-0.54 [-1.46 , 0.38]
Turner 1979	-4.1	2.644	10	-4	2.449	10	2.3%	-0.04 [-0.91 , 0.84]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	2.0%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**673**			**590**	**100.0%**	**-0.49 [-0.66 , -0.31]**
Heterogeneity: Tau² = 0.15; Chi² = 78.11, df = 37 (P < 0.0001); I² = 53%								
Test for overall effect: Z = 5.40 (P < 0.00001)								
Total (95% CI)			**673**			**590**	**100.0%**	**-0.49 [-0.66 , -0.31]**
Heterogeneity: Tau² = 0.15; Chi² = 78.11, df = 37 (P < 0.0001); I² = 53%								
Test for overall effect: Z = 5.40 (P < 0.00001)								
Test for subgroup differences: Not applicable								



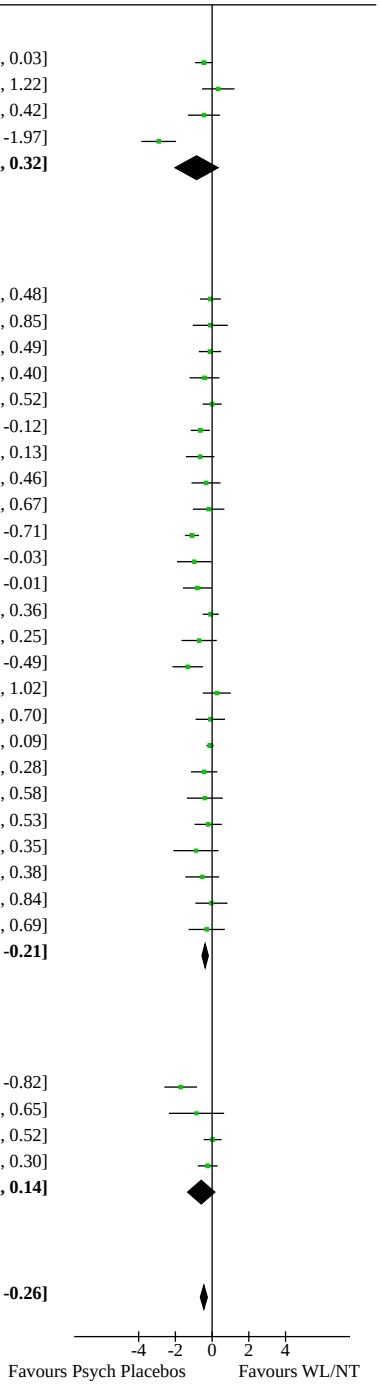
Analysis 17.5. Comparison 17: Subgroup analyses for psychological placebos compared with wait-list or no-treatment for continuous data, Outcome 5: The trial objective

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
17.5.1 A trial's objective is clearly to assess the effects of placebo								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
17.5.2 No such objectives are stated								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	3.9%	-0.45 [-0.93 , 0.03]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	3.5%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	2.1%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	2.3%	-1.71 [-2.60 , -0.82]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	2.3%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	2.4%	-0.25 [-1.09 , 0.59]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	3.3%	-0.11 [-0.72 , 0.49]
Bramston 1985	-78	14.21	12	-72	13.82	12	2.5%	-0.41 [-1.22 , 0.40]
Carter 2003	3.8	1.2	28	3.8	1.3	29	3.7%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	3.7%	-0.64 [-1.16 , -0.12]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	2.6%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	2.6%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	2.4%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	4.3%	-1.10 [-1.48 , -0.71]
Fuchs 1977	14.3	7	10	21.4	7	10	2.1%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	2.6%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	4.1%	-0.08 [-0.51 , 0.36]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	2.3%	-0.59 [-1.49 , 0.31]
Kilmann 1987	69	27	4	90	14	4	1.1%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	1.6%	-0.69 [-1.87 , 0.49]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	2.3%	-0.35 [-1.24 , 0.53]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	3.8%	0.03 [-0.46 , 0.52]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	2.1%	-0.71 [-1.67 , 0.25]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	2.3%	0.34 [-0.55 , 1.22]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	2.3%	-0.45 [-1.32 , 0.42]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	2.5%	-1.33 [-2.16 , -0.49]
Powers 2004	46	20.09	12	41.6	13.15	15	2.7%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	2.6%	-0.10 [-0.90 , 0.70]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	3.6%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	2.9%	-0.43 [-1.15 , 0.28]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	2.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	2.0%	-0.39 [-1.37 , 0.58]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	2.8%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	2.3%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	1.5%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	2.2%	-0.54 [-1.46 , 0.38]
Turner 1979	-4.1	2.644	10	-4	2.449	10	2.3%	-0.04 [-0.91 , 0.84]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	2.0%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**673**			**590**	**100.0%**	**-0.49 [-0.66 , -0.31]**
Heterogeneity: Tau² = 0.15; Chi² = 78.11, df = 37 (P < 0.0001); I² = 53%								
Test for overall effect: Z = 5.40 (P < 0.00001)								
Total (95% CI)			**673**			**590**	**100.0%**	**-0.49 [-0.66 , -0.31]**
Heterogeneity: Tau² = 0.15; Chi² = 78.11, df = 37 (P < 0.0001); I² = 53%								
Test for overall effect: Z = 5.40 (P < 0.00001)								
Test for subgroup differences: Not applicable								



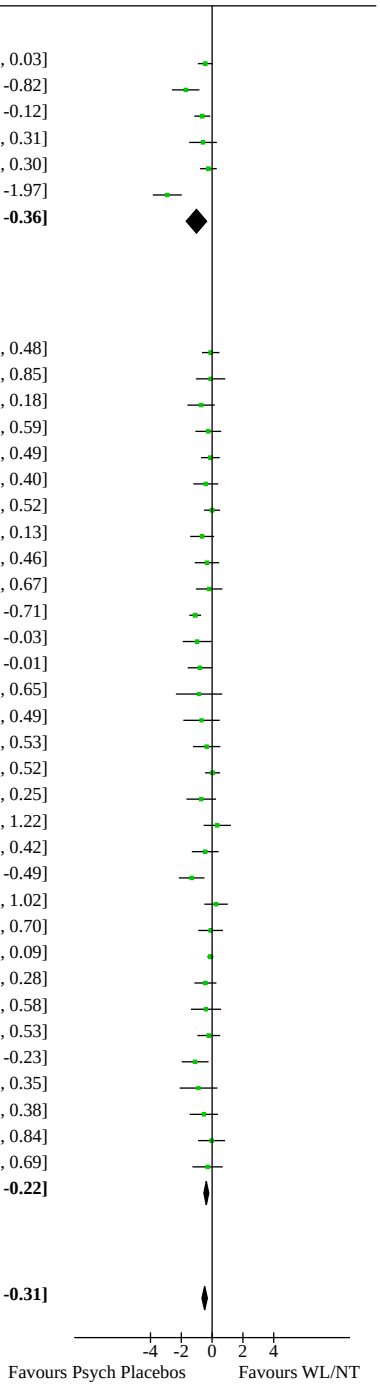
Analysis 17.6. Comparison 17: Subgroup analyses for psychological placebos compared with wait-list or no-treatment for continuous data, Outcome 6: Mean age of participants

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
17.6.1 Below 18 years								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	4.2%	-0.45 [-0.93 , 0.03]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	2.5%	0.34 [-0.55 , 1.22]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	2.6%	-0.45 [-1.32 , 0.42]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	2.4%	-2.91 [-3.85 , -1.97]
Subtotal (95% CI)			**74**			**74**	**11.7%**	**-0.85 [-2.02 , 0.32]**
Heterogeneity: Tau² = 1.25; Chi² = 27.88, df = 3 (P < 0.00001); I² = 89%								
Test for overall effect: Z = 1.42 (P = 0.16)								
17.6.2 18 to 50 years								
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	3.8%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	2.3%	-0.10 [-1.05 , 0.85]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	3.6%	-0.11 [-0.72 , 0.49]
Bramston 1985	-78	14.21	12	-72	13.82	12	2.8%	-0.41 [-1.22 , 0.40]
Carter 2003	3.8	1.2	28	3.8	1.3	29	4.0%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	4.0%	-0.64 [-1.16 , -0.12]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	2.9%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	2.8%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	2.6%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	4.6%	-1.10 [-1.48 , -0.71]
Fuchs 1977	14.3	7	10	21.4	7	10	2.4%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	2.9%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	4.4%	-0.08 [-0.51 , 0.36]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	2.3%	-0.71 [-1.67 , 0.25]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	2.7%	-1.33 [-2.16 , -0.49]
Powers 2004	46	20.09	12	41.6	13.15	15	3.0%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	2.8%	-0.10 [-0.90 , 0.70]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	5.4%	-0.11 [-0.31 , 0.09]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	3.1%	-0.43 [-1.15 , 0.28]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	2.2%	-0.39 [-1.37 , 0.58]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	3.0%	-0.22 [-0.96 , 0.53]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	1.7%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	2.4%	-0.54 [-1.46 , 0.38]
Turner 1979	-4.1	2.644	10	-4	2.449	10	2.6%	-0.04 [-0.91 , 0.84]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	2.2%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**647**			**586**	**76.6%**	**-0.38 [-0.55 , -0.21]**
Heterogeneity: Tau² = 0.06; Chi² = 39.93, df = 24 (P = 0.02); I² = 40%								
Test for overall effect: Z = 4.33 (P < 0.0001)								
17.6.3 Above 50 years								
Ayen 2004	14.6	4.5	20	22.1	3.7	10	2.5%	-1.71 [-2.60 , -0.82]
Kilmann 1987	69	27	4	90	14	4	1.2%	-0.85 [-2.35 , 0.65]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	4.2%	0.03 [-0.46 , 0.52]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	3.9%	-0.24 [-0.78 , 0.30]
Subtotal (95% CI)			**87**			**69**	**11.8%**	**-0.59 [-1.33 , 0.14]**
Heterogeneity: Tau² = 0.39; Chi² = 11.94, df = 3 (P = 0.008); I² = 75%								
Test for overall effect: Z = 1.58 (P = 0.11)								
Total (95% CI)			**808**			**729**	**100.0%**	**-0.45 [-0.64 , -0.26]**
Heterogeneity: Tau² = 0.16; Chi² = 83.13, df = 32 (P < 0.00001); I² = 62%								
Test for overall effect: Z = 4.70 (P < 0.00001)								
Test for subgroup differences: Chi² = 0.88, df = 2 (P = 0.64), I² = 0%								



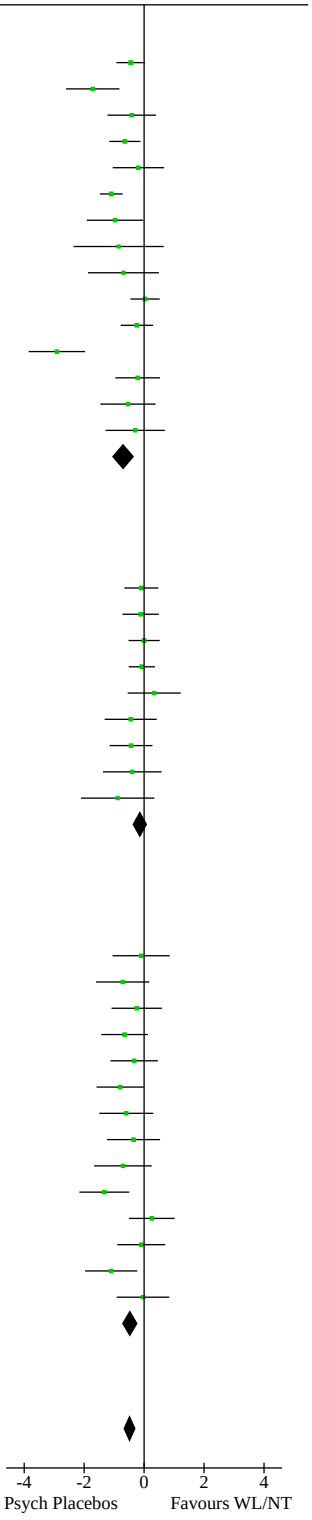
Analysis 17.7. Comparison 17: Subgroup analyses for psychological placebos compared with wait-list or no-treatment for continuous data, Outcome 7: Duration of treatment

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
17.7.1 Above 3 months								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	3.9%	-0.45 [-0.93 , 0.03]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	2.3%	-1.71 [-2.60 , -0.82]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	3.7%	-0.64 [-1.16 , -0.12]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	2.2%	-0.59 [-1.49 , 0.31]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	3.6%	-0.24 [-0.78 , 0.30]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	2.1%	-2.91 [-3.85 , -1.97]
Subtotal (95% CI)			**141**			**129**	**17.7%**	**-1.02 [-1.68 , -0.36]**
Heterogeneity: Tau² = 0.55; Chi² = 30.03, df = 5 (P < 0.0001); I² = 83%								
Test for overall effect: Z = 3.02 (P = 0.003)								
17.7.2 Below 3 months								
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	3.5%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	2.1%	-0.10 [-1.05 , 0.85]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	2.3%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	2.4%	-0.25 [-1.09 , 0.59]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	3.3%	-0.11 [-0.72 , 0.49]
Bramston 1985	-78	14.21	12	-72	13.82	12	2.5%	-0.41 [-1.22 , 0.40]
Carter 2003	3.8	1.2	28	3.8	1.3	29	3.7%	0.00 [-0.52 , 0.52]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	2.6%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	2.6%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	2.4%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	4.3%	-1.10 [-1.48 , -0.71]
Fuchs 1977	14.3	7	10	21.4	7	10	2.1%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	2.6%	-0.80 [-1.59 , -0.01]
Kilmann 1987	69	27	4	90	14	4	1.1%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	1.6%	-0.69 [-1.87 , 0.49]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	2.3%	-0.35 [-1.24 , 0.53]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	3.8%	0.03 [-0.46 , 0.52]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	2.1%	-0.71 [-1.67 , 0.25]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	2.3%	0.34 [-0.55 , 1.22]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	2.3%	-0.45 [-1.32 , 0.42]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	2.4%	-1.33 [-2.16 , -0.49]
Powers 2004	46	20.09	12	41.6	13.15	15	2.7%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	2.5%	-0.10 [-0.90 , 0.70]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	5.1%	-0.11 [-0.31 , 0.09]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	2.9%	-0.43 [-1.15 , 0.28]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	2.0%	-0.39 [-1.37 , 0.58]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	2.7%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	2.3%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	1.5%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	2.2%	-0.54 [-1.46 , 0.38]
Turner 1979	-4.1	2.644	10	-4	2.449	10	2.3%	-0.04 [-0.91 , 0.84]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	2.0%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**687**			**619**	**82.3%**	**-0.38 [-0.53 , -0.22]**
Heterogeneity: Tau² = 0.06; Chi² = 46.19, df = 31 (P = 0.04); I² = 33%								
Test for overall effect: Z = 4.70 (P < 0.00001)								
Total (95% CI)			**828**			**748**	**100.0%**	**-0.48 [-0.66 , -0.31]**
Heterogeneity: Tau² = 0.15; Chi² = 85.27, df = 37 (P < 0.0001); I² = 57%								
Test for overall effect: Z = 5.43 (P < 0.00001)								
Test for subgroup differences: Chi² = 3.42, df = 1 (P = 0.06), I² = 70.8%								



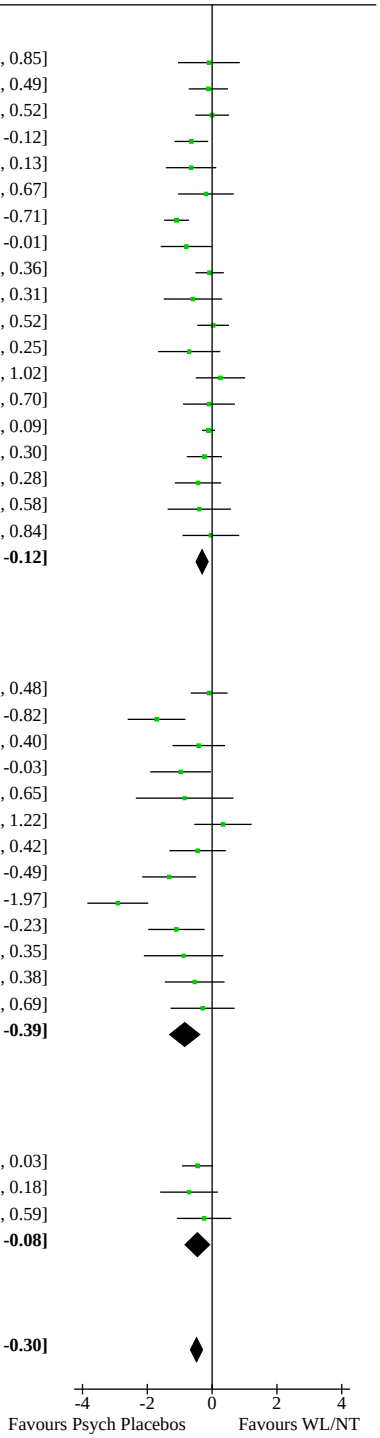
Analysis 17.8. Comparison 17: Subgroup analyses for psychological placebos compared with wait-list or no-treatment for continuous data, Outcome 8: Type of psychological placebo

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
17.8.1 Interaction placebo								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	3.9%	-0.45 [-0.93 , 0.03]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	2.3%	-1.71 [-2.60 , -0.82]
Bramston 1985	-78	14.21	12	-72	13.82	12	2.5%	-0.41 [-1.22 , 0.40]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	3.7%	-0.64 [-1.16 , -0.12]
Foa 1991	18.09	7.13	11	19.5	7.18	10	2.4%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	4.3%	-1.10 [-1.48 , -0.71]
Fuchs 1977	14.3	7	10	21.4	7	10	2.1%	-0.97 [-1.91 , -0.03]
Kilmann 1987	69	27	4	90	14	4	1.1%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	1.6%	-0.69 [-1.87 , 0.49]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	3.8%	0.03 [-0.46 , 0.52]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	3.6%	-0.24 [-0.78 , 0.30]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	2.1%	-2.91 [-3.85 , -1.97]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	2.8%	-0.22 [-0.96 , 0.53]
Tan 1986	9.5	7	10	13.44	7	9	2.2%	-0.54 [-1.46 , 0.38]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	2.0%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**351**			**262**	**40.6%**	**-0.70 [-1.04 , -0.36]**
Heterogeneity: Tau² = 0.28; Chi² = 47.14, df = 14 (P < 0.0001); I² = 70%								
Test for overall effect: Z = 4.07 (P < 0.0001)								
17.8.2 Psychoeducational placebo								
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	3.5%	-0.09 [-0.66 , 0.48]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	3.3%	-0.11 [-0.72 , 0.49]
Carter 2003	3.8	1.2	28	3.8	1.3	29	3.7%	0.00 [-0.52 , 0.52]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	4.1%	-0.08 [-0.51 , 0.36]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	2.3%	0.34 [-0.55 , 1.22]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	2.3%	-0.45 [-1.32 , 0.42]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	2.9%	-0.43 [-1.15 , 0.28]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	2.0%	-0.39 [-1.37 , 0.58]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	1.5%	-0.88 [-2.11 , 0.35]
Subtotal (95% CI)			**161**			**168**	**25.7%**	**-0.14 [-0.36 , 0.07]**
Heterogeneity: Tau² = 0.00; Chi² = 4.30, df = 8 (P = 0.83); I² = 0%								
Test for overall effect: Z = 1.29 (P = 0.20)								
17.8.3 Exposure placebo								
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	2.1%	-0.10 [-1.05 , 0.85]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	2.3%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	2.4%	-0.25 [-1.09 , 0.59]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	2.6%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	2.6%	-0.33 [-1.12 , 0.46]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	2.6%	-0.80 [-1.59 , -0.01]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	2.3%	-0.59 [-1.49 , 0.31]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	2.3%	-0.35 [-1.24 , 0.53]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	2.1%	-0.71 [-1.67 , 0.25]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	2.5%	-1.33 [-2.16 , -0.49]
Powers 2004	46	20.09	12	41.6	13.15	15	2.7%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	2.6%	-0.10 [-0.90 , 0.70]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	2.3%	-1.10 [-1.97 , -0.23]
Turner 1979	-4.1	2.644	10	-4	2.449	10	2.3%	-0.04 [-0.91 , 0.84]
Subtotal (95% CI)			**161**			**160**	**33.8%**	**-0.48 [-0.71 , -0.24]**
Heterogeneity: Tau² = 0.01; Chi² = 13.85, df = 13 (P = 0.38); I² = 6%								
Test for overall effect: Z = 4.01 (P < 0.0001)								
Total (95% CI)			**673**			**590**	**100.0%**	**-0.49 [-0.66 , -0.31]**
Heterogeneity: Tau² = 0.15; Chi² = 78.11, df = 37 (P < 0.0001); I² = 53%								
Test for overall effect: Z = 5.40 (P < 0.00001)								
Test for subgroup differences: Chi² = 8.68, df = 2 (P = 0.01), I² = 76.9%								



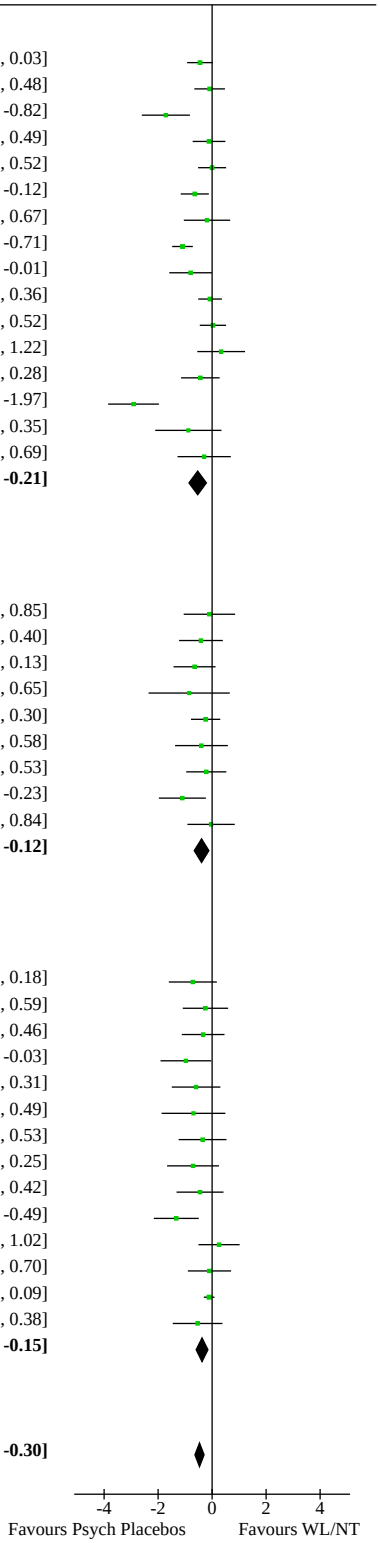
Analysis 17.9. Comparison 17: Subgroup analyses for psychological placebos compared with wait-list or no-treatment for continuous data, Outcome 9: Mode of psychological placebo

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
17.9.1 Individual treatment								
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	2.2%	-0.10 [-1.05 , 0.85]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	3.4%	-0.11 [-0.72 , 0.49]
Carter 2003	3.8	1.2	28	3.8	1.3	29	3.8%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	3.8%	-0.64 [-1.16 , -0.12]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	2.8%	-0.65 [-1.43 , 0.13]
Foa 1991	18.09	7.13	11	19.5	7.18	10	2.5%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	4.4%	-1.10 [-1.48 , -0.71]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	2.7%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	4.2%	-0.08 [-0.51 , 0.36]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	2.4%	-0.59 [-1.49 , 0.31]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	4.0%	0.03 [-0.46 , 0.52]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	2.2%	-0.71 [-1.67 , 0.25]
Powers 2004	46	20.09	12	41.6	13.15	15	2.8%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	2.7%	-0.10 [-0.90 , 0.70]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	5.2%	-0.11 [-0.31 , 0.09]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	3.7%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	3.0%	-0.43 [-1.15 , 0.28]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	2.2%	-0.39 [-1.37 , 0.58]
Turner 1979	-4.1	2.644	10	-4	2.449	10	2.4%	-0.04 [-0.91 , 0.84]
Subtotal (95% CI)			**612**			**539**	**60.6%**	**-0.31 [-0.49 , -0.12]**
Heterogeneity: Tau² = 0.06; Chi² = 32.49, df = 18 (P = 0.02); I² = 45%								
Test for overall effect: Z = 3.28 (P = 0.001)								
17.9.2 Group treatment								
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	3.6%	-0.09 [-0.66 , 0.48]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	2.4%	-1.71 [-2.60 , -0.82]
Bramston 1985	-78	14.21	12	-72	13.82	12	2.7%	-0.41 [-1.22 , 0.40]
Fuchs 1977	14.3	7	10	21.4	7	10	2.3%	-0.97 [-1.91 , -0.03]
Kilmann 1987	69	27	4	90	14	4	1.2%	-0.85 [-2.35 , 0.65]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	2.4%	0.34 [-0.55 , 1.22]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	2.5%	-0.45 [-1.32 , 0.42]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	2.6%	-1.33 [-2.16 , -0.49]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	2.3%	-2.91 [-3.85 , -1.97]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	2.5%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	1.6%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	2.3%	-0.54 [-1.46 , 0.38]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	2.1%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**156**			**152**	**30.4%**	**-0.85 [-1.30 , -0.39]**
Heterogeneity: Tau² = 0.49; Chi² = 40.66, df = 12 (P < 0.0001); I² = 70%								
Test for overall effect: Z = 3.60 (P = 0.0003)								
17.9.3 Combination of individual and group								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	4.0%	-0.45 [-0.93 , 0.03]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	2.4%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	2.6%	-0.25 [-1.09 , 0.59]
Subtotal (95% CI)			**57**			**55**	**9.0%**	**-0.45 [-0.83 , -0.08]**
Heterogeneity: Tau² = 0.00; Chi² = 0.56, df = 2 (P = 0.75); I² = 0%								
Test for overall effect: Z = 2.37 (P = 0.02)								
Total (95% CI)			**825**			**746**	**100.0%**	**-0.48 [-0.66 , -0.30]**
Heterogeneity: Tau² = 0.16; Chi² = 86.65, df = 34 (P < 0.00001); I² = 61%								
Test for overall effect: Z = 5.12 (P < 0.00001)								
Test for subgroup differences: Chi² = 4.62, df = 2 (P = 0.10), I² = 56.7%								



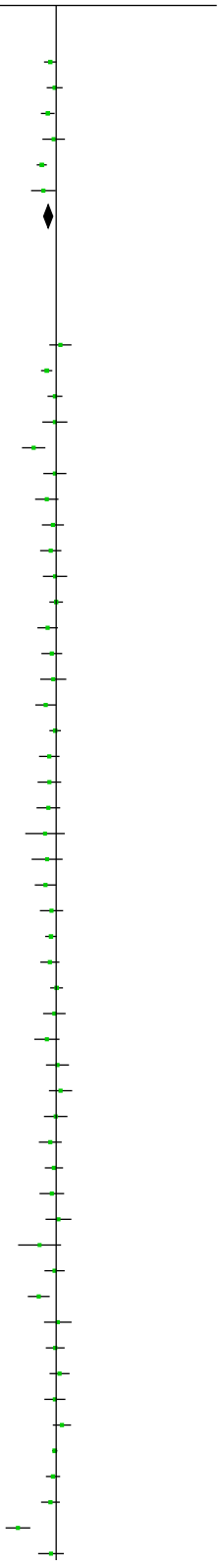
Analysis 17.10. Comparison 17: Subgroup analyses for psychological placebos compared with wait-list or no-treatment for continuous data, Outcome 10: Mental health diagnoses

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
17.10.1 Formal diagnosis according to DSM/ICD								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	3.7%	-0.45 [-0.93 , 0.03]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	3.3%	-0.09 [-0.66 , 0.48]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	2.2%	-1.71 [-2.60 , -0.82]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	3.2%	-0.11 [-0.72 , 0.49]
Carter 2003	3.8	1.2	28	3.8	1.3	29	3.6%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	3.6%	-0.64 [-1.16 , -0.12]
Foa 1991	18.09	7.13	11	19.5	7.18	10	2.3%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	4.2%	-1.10 [-1.48 , -0.71]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	2.5%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	3.9%	-0.08 [-0.51 , 0.36]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	3.7%	0.03 [-0.46 , 0.52]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	2.2%	0.34 [-0.55 , 1.22]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	2.7%	-0.43 [-1.15 , 0.28]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	2.0%	-2.91 [-3.85 , -1.97]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	1.4%	-0.88 [-2.11 , 0.35]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	1.9%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**420**			**347**	**46.3%**	**-0.53 [-0.86 , -0.21]**
Heterogeneity: Tau² = 0.31; Chi² = 62.25, df = 15 (P < 0.00001); I² = 76%								
Test for overall effect: Z = 3.25 (P = 0.001)								
17.10.2 Fulfill symptoms of disorder ICD/DSM while not stating systems								
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	2.0%	-0.10 [-1.05 , 0.85]
Bramston 1985	-78	14.21	12	-72	13.82	12	2.4%	-0.41 [-1.22 , 0.40]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	2.5%	-0.65 [-1.43 , 0.13]
Kilmann 1987	69	27	4	90	14	4	1.0%	-0.85 [-2.35 , 0.65]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	3.5%	-0.24 [-0.78 , 0.30]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.9%	-0.39 [-1.37 , 0.58]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	2.6%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	2.2%	-1.10 [-1.97 , -0.23]
Turner 1979	-4.1	2.644	10	-4	2.449	10	2.2%	-0.04 [-0.91 , 0.84]
Subtotal (95% CI)			**112**			**106**	**20.4%**	**-0.39 [-0.66 , -0.12]**
Heterogeneity: Tau² = 0.00; Chi² = 4.84, df = 8 (P = 0.77); I² = 0%								
Test for overall effect: Z = 2.82 (P = 0.005)								
17.10.3 Population is classified as having a mental disorder, but full diagnostic criteria not reported								
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	2.2%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	2.3%	-0.25 [-1.09 , 0.59]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	2.5%	-0.33 [-1.12 , 0.46]
Fuchs 1977	14.3	7	10	21.4	7	10	2.0%	-0.97 [-1.91 , -0.03]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	2.1%	-0.59 [-1.49 , 0.31]
Klein 1977	10.17	5.49	6	15	7.32	6	1.5%	-0.69 [-1.87 , 0.49]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	2.2%	-0.35 [-1.24 , 0.53]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	2.0%	-0.71 [-1.67 , 0.25]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	2.2%	-0.45 [-1.32 , 0.42]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	2.3%	-1.33 [-2.16 , -0.49]
Powers 2004	46	20.09	12	41.6	13.15	15	2.6%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	2.4%	-0.10 [-0.90 , 0.70]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	5.0%	-0.11 [-0.31 , 0.09]
Tan 1986	9.5	7	10	13.44	7	9	2.1%	-0.54 [-1.46 , 0.38]
Subtotal (95% CI)			**336**			**335**	**33.3%**	**-0.37 [-0.59 , -0.15]**
Heterogeneity: Tau² = 0.03; Chi² = 16.11, df = 13 (P = 0.24); I² = 19%								
Test for overall effect: Z = 3.35 (P = 0.0008)								
Total (95% CI)			**868**			**788**	**100.0%**	**-0.47 [-0.64 , -0.30]**
Heterogeneity: Tau² = 0.14; Chi² = 87.11, df = 38 (P < 0.0001); I² = 56%								
Test for overall effect: Z = 5.40 (P < 0.00001)								
Test for subgroup differences: Chi² = 0.70, df = 2 (P = 0.70), I² = 0%								



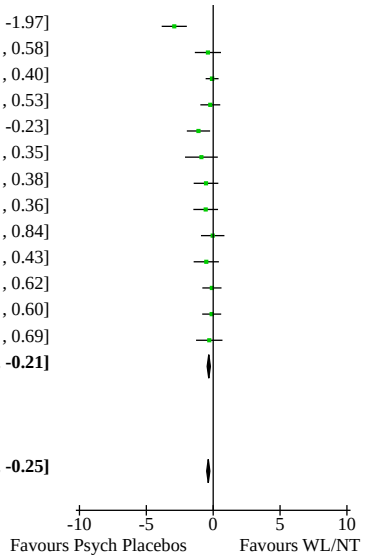
Analysis 17.11. Comparison 17: Subgroup analyses for psychological placebos compared with wait-list or no-treatment for continuous data, Outcome 11: Affiliation bias

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
17.11.1 Risk of affiliation, industry, allegiance bias								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	2.5%	-0.45 [-0.93 , 0.03]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	2.0%	-0.11 [-0.72 , 0.49]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	2.3%	-0.64 [-1.16 , -0.12]
Foa 1991	18.09	7.13	11	19.5	7.18	10	1.3%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	3.0%	-1.10 [-1.48 , -0.71]
Fuchs 1977	14.3	7	10	21.4	7	10	1.1%	-0.97 [-1.91 , -0.03]
Subtotal (95% CI)			**214**			**149**	**12.3%**	**-0.61 [-0.95 , -0.26]**
Heterogeneity: Tau² = 0.09; Chi² = 10.43, df = 5 (P = 0.06); I² = 52%								
Test for overall effect: Z = 3.47 (P = 0.0005)								
17.11.2 No risk of affiliation, industry, allegiance bias								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	1.3%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	2.8%	-0.72 [-1.15 , -0.29]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	2.2%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	1.1%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	1.2%	-1.71 [-2.60 , -0.82]
Borden 1986	29.7	17.5	10	32	26	10	1.3%	-0.10 [-0.98 , 0.78]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	1.2%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	1.3%	-0.25 [-1.09 , 0.59]
Bramston 1985	-78	14.21	12	-72	13.82	12	1.4%	-0.41 [-1.22 , 0.40]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	1.2%	-0.09 [-1.01 , 0.84]
Carter 2003	3.8	1.2	28	3.8	1.3	29	2.3%	0.00 [-0.52 , 0.52]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	1.5%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	1.5%	-0.33 [-1.12 , 0.46]
Freire 2007	24.6	11	7	28.2	18	9	1.1%	-0.22 [-1.21 , 0.77]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	1.5%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	2.7%	-0.08 [-0.51 , 0.36]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	1.5%	-0.53 [-1.30 , 0.25]
Kelley 2012	-1.64	4.52	11	0.67	4	9	1.2%	-0.51 [-1.41 , 0.38]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	1.2%	-0.59 [-1.49 , 0.31]
Kilmann 1987	69	27	4	90	14	4	0.5%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	0.8%	-0.69 [-1.87 , 0.49]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	1.4%	-0.81 [-1.63 , 0.00]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	1.3%	-0.35 [-1.24 , 0.53]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	2.7%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	1.6%	-0.48 [-1.21 , 0.25]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	2.5%	0.03 [-0.46 , 0.52]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	1.3%	-0.14 [-1.00 , 0.72]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	1.1%	-0.71 [-1.67 , 0.25]
Lick 1977	65.86	25.45	10	62.84	31.78	10	1.3%	0.10 [-0.78 , 0.98]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	1.3%	0.34 [-0.55 , 1.22]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	1.2%	-0.04 [-0.93 , 0.86]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	1.3%	-0.45 [-1.32 , 0.42]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	1.7%	-0.17 [-0.87 , 0.53]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	1.1%	-0.33 [-1.27 , 0.61]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.1%	0.17 [-0.81 , 1.16]
Peck 1976	-12.5	4.25	4	-7	3.27	4	0.5%	-1.26 [-2.90 , 0.37]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	1.5%	-0.12 [-0.89 , 0.65]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	1.4%	-1.33 [-2.16 , -0.49]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.0%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	1.7%	-0.07 [-0.79 , 0.65]
Powers 2004	46	20.09	12	41.6	13.15	15	1.5%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	1.4%	-0.10 [-0.90 , 0.70]
Powers 2008b	64	20	18	51	37	15	1.7%	0.44 [-0.26 , 1.13]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	3.9%	-0.11 [-0.31 , 0.09]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	2.3%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	1.7%	-0.43 [-1.15 , 0.28]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]



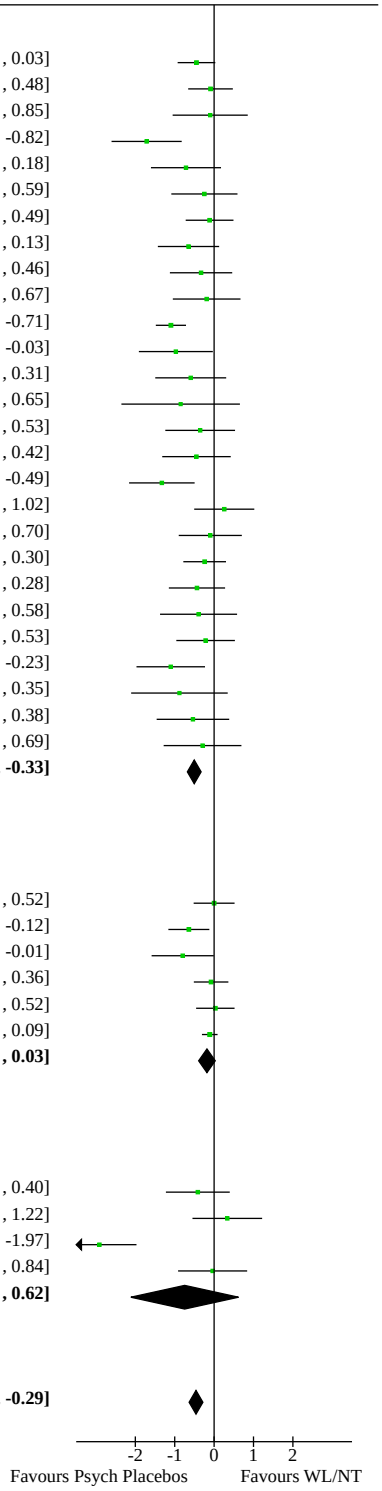
Analysis 17.11. (Continued)

Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	2.5%	-0.08 [-0.56 , 0.40]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	1.6%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	1.3%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	0.8%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	1.2%	-0.54 [-1.46 , 0.38]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	1.2%	-0.56 [-1.48 , 0.36]
Turner 1979	-4.1	2.644	10	-4	2.449	10	1.3%	-0.04 [-0.91 , 0.84]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	1.1%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	1.7%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	1.7%	-0.11 [-0.82 , 0.60]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	1.1%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**1032**			**1051**	**87.7%**	**-0.33 [-0.45 , -0.21]**
Heterogeneity: Tau² = 0.06; Chi² = 87.74, df = 58 (P = 0.007); I² = 34%								
Test for overall effect: Z = 5.41 (P < 0.00001)								
Total (95% CI)			**1246**			**1200**	**100.0%**	**-0.37 [-0.49 , -0.25]**
Heterogeneity: Tau² = 0.08; Chi² = 108.28, df = 64 (P = 0.0005); I² = 41%								
Test for overall effect: Z = 6.16 (P < 0.00001)								
Test for subgroup differences: Chi² = 2.22, df = 1 (P = 0.14), I² = 55.0%								



Analysis 17.12. Comparison 17: Subgroup analyses for psychological placebos compared with wait-list or no-treatment for continuous data, Outcome 12: Imputed data

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
17.12.1 Available data								
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	3.9%	-0.45 [-0.93 , 0.03]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	3.4%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	2.1%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	2.2%	-1.71 [-2.60 , -0.82]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	2.3%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	2.4%	-0.25 [-1.09 , 0.59]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	3.3%	-0.11 [-0.72 , 0.49]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	2.6%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	2.6%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	2.3%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	4.3%	-1.10 [-1.48 , -0.71]
Fuchs 1977	14.3	7	10	21.4	7	10	2.1%	-0.97 [-1.91 , -0.03]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	2.2%	-0.59 [-1.49 , 0.31]
Kilmann 1987	69	27	4	90	14	4	1.1%	-0.85 [-2.35 , 0.65]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	2.3%	-0.35 [-1.24 , 0.53]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	2.3%	-0.45 [-1.32 , 0.42]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	2.4%	-1.33 [-2.16 , -0.49]
Powers 2004	46	20.09	12	41.6	13.15	15	2.7%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	2.5%	-0.10 [-0.90 , 0.70]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	3.6%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	2.8%	-0.43 [-1.15 , 0.28]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	2.0%	-0.39 [-1.37 , 0.58]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	2.7%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	2.3%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	1.5%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	2.2%	-0.54 [-1.46 , 0.38]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	2.0%	-0.29 [-1.28 , 0.69]
Subtotal (95% CI)			**462**			**380**	**68.0%**	**-0.50 [-0.68 , -0.33]**
Heterogeneity: Tau² = 0.05; Chi² = 35.81, df = 26 (P = 0.10); I² = 27%								
Test for overall effect: Z = 5.67 (P < 0.00001)								
17.12.2 Intent-to-treat								
Carter 2003	3.8	1.2	28	3.8	1.3	29	3.7%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	3.7%	-0.64 [-1.16 , -0.12]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	2.6%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	4.0%	-0.08 [-0.51 , 0.36]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	3.8%	0.03 [-0.46 , 0.52]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	5.1%	-0.11 [-0.31 , 0.09]
Subtotal (95% CI)			**341**			**341**	**22.8%**	**-0.18 [-0.39 , 0.03]**
Heterogeneity: Tau² = 0.02; Chi² = 7.13, df = 5 (P = 0.21); I² = 30%								
Test for overall effect: Z = 1.69 (P = 0.09)								
17.12.3 No attrition								
Bramston 1985	-78	14.21	12	-72	13.82	12	2.5%	-0.41 [-1.22 , 0.40]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	2.3%	0.34 [-0.55 , 1.22]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	2.1%	-2.91 [-3.85 , -1.97]
Turner 1979	-4.1	2.644	10	-4	2.449	10	2.3%	-0.04 [-0.91 , 0.84]
Subtotal (95% CI)			**50**			**52**	**9.2%**	**-0.75 [-2.11 , 0.62]**
Heterogeneity: Tau² = 1.74; Chi² = 29.18, df = 3 (P < 0.00001); I² = 90%								
Test for overall effect: Z = 1.07 (P = 0.28)								
Total (95% CI)			**853**			**773**	**100.0%**	**-0.46 [-0.63 , -0.29]**
Heterogeneity: Tau² = 0.15; Chi² = 86.36, df = 36 (P < 0.00001); I² = 58%								
Test for overall effect: Z = 5.16 (P < 0.00001)								
Test for subgroup differences: Chi² = 5.65, df = 2 (P = 0.06), I² = 64.6%								



Comparison 18. Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
18.1 Type of active intervention	9	279	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.39, 0.11]
18.1.1 Psychological intervention	1	33	Std. Mean Difference (IV, Random, 95% CI)	0.44 [-0.26, 1.13]
18.1.2 Pharmacological intervention	7	148	Std. Mean Difference (IV, Random, 95% CI)	-0.29 [-0.62, 0.04]
18.1.3 Physical intervention	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
18.1.4 Other or combination intervention	1	98	Std. Mean Difference (IV, Random, 95% CI)	-0.08 [-0.56, 0.40]
18.2 Type of active intervention	8	366	Risk Ratio (M-H, Random, 95% CI)	1.05 [0.74, 1.48]
18.2.1 Psychological intervention	0	0	Risk Ratio (M-H, Random, 95% CI)	Not estimable
18.2.2 Pharmacological intervention	7	316	Risk Ratio (M-H, Random, 95% CI)	1.09 [0.71, 1.67]
18.2.3 Physical intervention	0	0	Risk Ratio (M-H, Random, 95% CI)	Not estimable
18.2.4 Other or combination intervention	1	50	Risk Ratio (M-H, Random, 95% CI)	0.99 [0.58, 1.70]
18.3 Risk of bias	9	279	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.39, 0.11]
18.3.1 Low risk of bias	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
18.3.2 High risk of bias	9	279	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.39, 0.11]
18.4 Risk of bias	8	366	Risk Ratio (IV, Random, 95% CI)	1.01 [0.76, 1.35]
18.4.1 Low risk of bias	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
18.4.2 High risk of bias	8	366	Risk Ratio (IV, Random, 95% CI)	1.01 [0.76, 1.35]
18.5 Type of outcome domain	9	279	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.39, 0.11]
18.5.1 Blinded observer-reported	7	220	Std. Mean Difference (IV, Random, 95% CI)	-0.24 [-0.53, 0.05]
18.5.2 Non-blinded observer-reported	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable

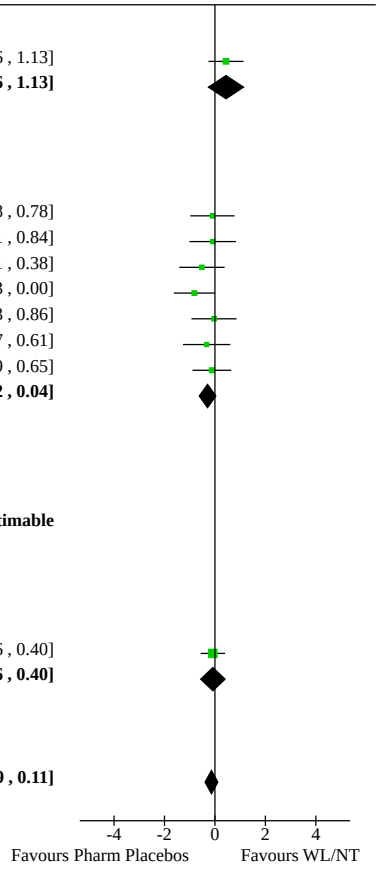
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
18.5.3 Patient-reported	2	59	Std. Mean Difference (IV, Random, 95% CI)	0.18 [-0.36, 0.73]
18.6 Type of outcome domain	4	170	Risk Ratio (IV, Random, 95% CI)	0.94 [0.73, 1.21]
18.6.1 Blinded observer-reported	2	94	Risk Ratio (IV, Random, 95% CI)	0.99 [0.60, 1.64]
18.6.2 Non-blinded observer-reported	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
18.6.3 Patient-reported	2	76	Risk Ratio (IV, Random, 95% CI)	1.00 [0.60, 1.65]
18.7 Awareness of placebo intervention	9	279	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.39, 0.11]
18.7.1 Open placebo	1	20	Std. Mean Difference (IV, Random, 95% CI)	-0.51 [-1.41, 0.38]
18.7.2 Closed placebo	8	259	Std. Mean Difference (IV, Random, 95% CI)	-0.11 [-0.37, 0.16]
18.8 Awareness of placebo intervention	9	385	Risk Ratio (IV, Random, 95% CI)	1.01 [0.76, 1.35]
18.8.1 Open placebo	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
18.8.2 Closed placebo	9	385	Risk Ratio (IV, Random, 95% CI)	1.01 [0.76, 1.35]
18.9 The trial objective	9	279	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.39, 0.11]
18.9.1 A trial's objective is clearly to assess the effects of placebo	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
18.9.2 No such objectives are stated	9	279	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.39, 0.11]
18.10 The trial objective	9	385	Risk Ratio (IV, Random, 95% CI)	1.01 [0.76, 1.35]
18.10.1 A trial's objective is clearly to assess the effects of placebo	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
18.10.2 No such objectives are stated	9	385	Risk Ratio (IV, Random, 95% CI)	1.01 [0.76, 1.35]
18.11 Mean age of participants	7	241	Std. Mean Difference (IV, Random, 95% CI)	-0.13 [-0.41, 0.14]
18.11.1 Below 18 years	3	64	Std. Mean Difference (IV, Random, 95% CI)	-0.10 [-0.59, 0.39]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
18.11.2 18 to 50 years	4	177	Std. Mean Difference (IV, Random, 95% CI)	-0.19 [-0.68, 0.31]
18.11.3 Above 50 years	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
18.12 Mean age of participants	7	340	Risk Ratio (IV, Random, 95% CI)	0.94 [0.70, 1.27]
18.12.1 Below 18 years	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
18.12.2 18 to 50 years	5	270	Risk Ratio (IV, Random, 95% CI)	1.07 [0.70, 1.62]
18.12.3 Above 50 years	2	70	Risk Ratio (IV, Random, 95% CI)	0.73 [0.43, 1.27]
18.13 Duration of treatment	9	279	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.39, 0.11]
18.13.1 Above 3 months	4	164	Std. Mean Difference (IV, Random, 95% CI)	-0.21 [-0.56, 0.14]
18.13.2 Below 3 months	5	115	Std. Mean Difference (IV, Random, 95% CI)	-0.06 [-0.43, 0.31]
18.14 Duration of treatment	8	366	Risk Ratio (IV, Random, 95% CI)	1.01 [0.76, 1.35]
18.14.1 Above 3 months	4	196	Risk Ratio (IV, Random, 95% CI)	1.45 [0.84, 2.49]
18.14.2 Below 3 months	4	170	Risk Ratio (IV, Random, 95% CI)	0.82 [0.68, 0.98]
18.15 Mental health diagnoses	9	279	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.39, 0.11]
18.15.1 Formal diagnosis according to DSM/ICD	5	188	Std. Mean Difference (IV, Random, 95% CI)	-0.25 [-0.57, 0.06]
18.15.2 Fulfill symptoms of disorder ICD/DSM while not stating systems	2	40	Std. Mean Difference (IV, Random, 95% CI)	-0.07 [-0.69, 0.56]
18.15.3 Population is classified as having a mental disorder, but full diagnostic criteria not reported	2	51	Std. Mean Difference (IV, Random, 95% CI)	0.12 [-0.62, 0.86]
18.16 Mental health diagnoses	7	316	Risk Ratio (IV, Random, 95% CI)	1.12 [0.79, 1.59]
18.16.1 Formal diagnosis according to DSM/ICD	4	220	Risk Ratio (IV, Random, 95% CI)	1.31 [0.72, 2.37]
18.16.2 Fulfill symptoms of disorder ICD/DSM while not stating systems	3	96	Risk Ratio (IV, Random, 95% CI)	1.01 [0.62, 1.63]
18.16.3 Population is classified as having a mental disorder,	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
but full diagnostic criteria not reported				
18.17 Imputed data	9	279	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.39, 0.11]
18.17.1 Available data	7	155	Std. Mean Difference (IV, Random, 95% CI)	-0.17 [-0.49, 0.16]
18.17.2 Intent-to-treat	1	98	Std. Mean Difference (IV, Random, 95% CI)	-0.08 [-0.56, 0.40]
18.17.3 No attrition	1	26	Std. Mean Difference (IV, Random, 95% CI)	-0.12 [-0.89, 0.65]

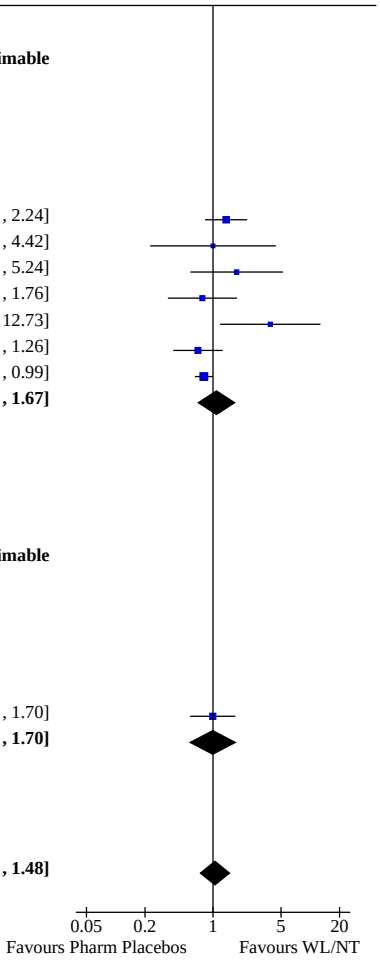
Analysis 18.1. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 1: Type of active intervention

Study or Subgroup	Pharmacological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
18.1.1 Psychological intervention								
Powers 2008b	64	20	18	51	37	15	13.2%	0.44 [-0.26 , 1.13]
Subtotal (95% CI)			**18**			**15**	**13.2%**	**0.44 [-0.26 , 1.13]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.24 (P = 0.22)								
18.1.2 Pharmacological intervention								
Borden 1986	29.7	17.5	10	32	26	10	8.3%	-0.10 [-0.98 , 0.78]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	7.5%	-0.09 [-1.01 , 0.84]
Kelley 2012	-1.64	4.52	11	0.67	4	9	7.9%	-0.51 [-1.41 , 0.38]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	9.6%	-0.81 [-1.63 , 0.00]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	8.0%	-0.04 [-0.93 , 0.86]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	7.3%	-0.33 [-1.27 , 0.61]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	10.8%	-0.12 [-0.89 , 0.65]
Subtotal (95% CI)			**72**			**76**	**59.3%**	**-0.29 [-0.62 , 0.04]**
Heterogeneity: Tau² = 0.00; Chi² = 2.70, df = 6 (P = 0.85); I² = 0%								
Test for overall effect: Z = 1.75 (P = 0.08)								
18.1.3 Physical intervention								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
18.1.4 Other or combination intervention								
Shalev 2012	47.11	20.13	21	49.18	26.4	77	27.4%	-0.08 [-0.56 , 0.40]
Subtotal (95% CI)			**21**			**77**	**27.4%**	**-0.08 [-0.56 , 0.40]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.33 (P = 0.74)								
Total (95% CI)			**111**			**168**	**100.0%**	**-0.14 [-0.39 , 0.11]**
Heterogeneity: Tau² = 0.00; Chi² = 6.26, df = 8 (P = 0.62); I² = 0%								
Test for overall effect: Z = 1.07 (P = 0.28)								
Test for subgroup differences: Chi² = 3.55, df = 2 (P = 0.17), I² = 43.7%								



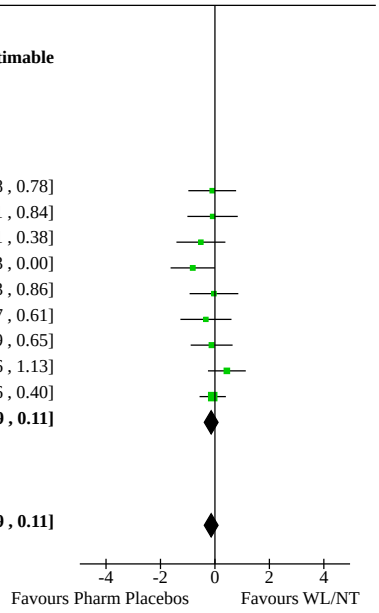
Analysis 18.2. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 2: Type of active intervention

Study or Subgroup	Pharmacological placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
18.2.1 Psychological intervention						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
18.2.2 Pharmacological intervention						
Berg 1983	9	11	9	15	16.9%	1.36 [0.83 , 2.24]
Double 1993	3	22	3	22	4.4%	1.00 [0.23 , 4.42]
Klerman 1974a	7	25	4	25	7.1%	1.75 [0.58 , 5.24]
Klerman 1974b	7	25	9	25	10.5%	0.78 [0.34 , 1.76]
Watzl 1988	11	34	3	36	6.3%	3.88 [1.18 , 12.73]
Whittaker 1963	7	13	10	13	14.9%	0.70 [0.39 , 1.26]
Wilson 1980	31	40	10	10	23.8%	0.80 [0.65 , 0.99]
Subtotal (95% CI)		**170**		**146**	**84.0%**	**1.09 [0.71 , 1.67]**
Total events:	75		48			
Heterogeneity: Tau² = 0.18; Chi² = 17.65, df = 6 (P = 0.007); I² = 66%						
Test for overall effect: Z = 0.38 (P = 0.71)						
18.2.3 Physical intervention						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
18.2.4 Other or combination intervention						
Rabkin 1990	14	27	12	23	16.0%	0.99 [0.58 , 1.70]
Subtotal (95% CI)		**27**		**23**	**16.0%**	**0.99 [0.58 , 1.70]**
Total events:	14		12			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.02 (P = 0.98)						
Total (95% CI)		**197**		**169**	**100.0%**	**1.05 [0.74 , 1.48]**
Total events:	89		60			
Heterogeneity: Tau² = 0.12; Chi² = 16.68, df = 7 (P = 0.02); I² = 58%						
Test for overall effect: Z = 0.26 (P = 0.79)						
Test for subgroup differences: Chi² = 0.06, df = 1 (P = 0.80), I² = 0%						



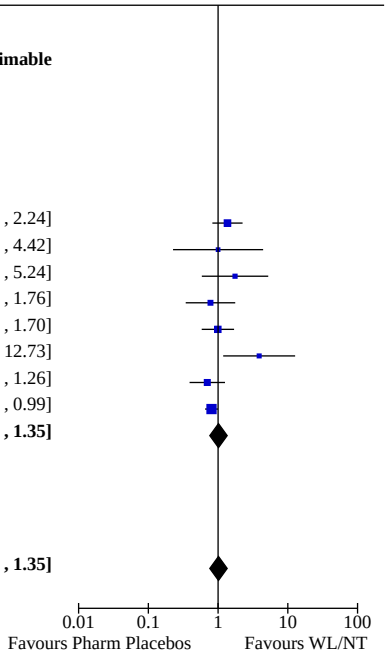
Analysis 18.3. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 3: Risk of bias

Study or Subgroup	Pharmacological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
18.3.1 Low risk of bias								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
18.3.2 High risk of bias								
Borden 1986	29.7	17.5	10	32	26	10	8.3%	-0.10 [-0.98 , 0.78]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	7.5%	-0.09 [-1.01 , 0.84]
Kelley 2012	-1.64	4.52	11	0.67	4	9	7.9%	-0.51 [-1.41 , 0.38]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	9.6%	-0.81 [-1.63 , 0.00]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	8.0%	-0.04 [-0.93 , 0.86]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	7.3%	-0.33 [-1.27 , 0.61]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	10.8%	-0.12 [-0.89 , 0.65]
Powers 2008b	64	20	18	51	37	15	13.2%	0.44 [-0.26 , 1.13]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	27.4%	-0.08 [-0.56 , 0.40]
Subtotal (95% CI)			**111**			**168**	**100.0%**	**-0.14 [-0.39 , 0.11]**
Heterogeneity: Tau² = 0.00; Chi² = 6.26, df = 8 (P = 0.62); I² = 0%								
Test for overall effect: Z = 1.07 (P = 0.28)								
Total (95% CI)			**111**			**168**	**100.0%**	**-0.14 [-0.39 , 0.11]**
Heterogeneity: Tau² = 0.00; Chi² = 6.26, df = 8 (P = 0.62); I² = 0%								
Test for overall effect: Z = 1.07 (P = 0.28)								
Test for subgroup differences: Not applicable								



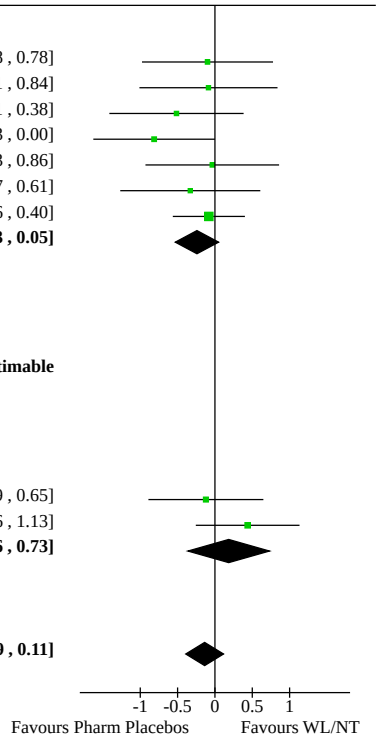
Analysis 18.4. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 4: Risk of bias

Study or Subgroup	Pharmacological placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio IV, Random, 95% CI
18.4.1 Low risk of bias						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
18.4.2 High risk of bias						
Berg 1983	9	11	9	15	17.1%	1.36 [0.83 , 2.24]
Double 1993	3	22	3	22	3.4%	1.00 [0.23 , 4.42]
Klerman 1974a	7	25	4	25	5.7%	1.75 [0.58 , 5.24]
Klerman 1974b	7	25	9	25	9.1%	0.78 [0.34 , 1.76]
Rabkin 1990	14	27	12	23	15.9%	0.99 [0.58 , 1.70]
Watzl 1988	11	34	3	36	5.0%	3.88 [1.18 , 12.73]
Whittaker 1963	7	13	10	13	14.3%	0.70 [0.39 , 1.26]
Wilson 1980	31	40	10	10	29.6%	0.80 [0.65 , 0.99]
Subtotal (95% CI)		**197**		**169**	**100.0%**	**1.01 [0.76 , 1.35]**
Total events:	89		60			
Heterogeneity: Tau² = 0.06; Chi² = 11.95, df = 7 (P = 0.10); I² = 41%						
Test for overall effect: Z = 0.09 (P = 0.93)						
Total (95% CI)		**197**		**169**	**100.0%**	**1.01 [0.76 , 1.35]**
Total events:	89		60			
Heterogeneity: Tau² = 0.06; Chi² = 11.95, df = 7 (P = 0.10); I² = 41%						
Test for overall effect: Z = 0.09 (P = 0.93)						
Test for subgroup differences: Not applicable						



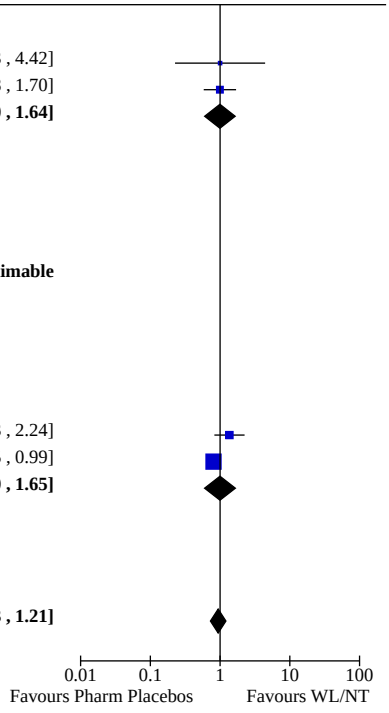
Analysis 18.5. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 5: Type of outcome domain

Study or Subgroup	Pharmacological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
18.5.1 Blinded observer-reported								
Borden 1986	29.7	17.5	10	32	26	10	8.3%	-0.10 [-0.98 , 0.78]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	7.5%	-0.09 [-1.01 , 0.84]
Kelley 2012	-1.64	4.52	11	0.67	4	9	7.9%	-0.51 [-1.41 , 0.38]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	9.6%	-0.81 [-1.63 , 0.00]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	8.0%	-0.04 [-0.93 , 0.86]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	7.3%	-0.33 [-1.27 , 0.61]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	27.4%	-0.08 [-0.56 , 0.40]
Subtotal (95% CI)			**80**			**140**	**76.0%**	**-0.24 [-0.53 , 0.05]**
Heterogeneity: Tau² = 0.00; Chi² = 3.12, df = 6 (P = 0.79); I² = 0%								
Test for overall effect: Z = 1.63 (P = 0.10)								
18.5.2 Non-blinded observer-reported								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
18.5.3 Patient-reported								
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	10.8%	-0.12 [-0.89 , 0.65]
Powers 2008b	64	20	18	51	37	15	13.2%	0.44 [-0.26 , 1.13]
Subtotal (95% CI)			**31**			**28**	**24.0%**	**0.18 [-0.36 , 0.73]**
Heterogeneity: Tau² = 0.02; Chi² = 1.12, df = 1 (P = 0.29); I² = 11%								
Test for overall effect: Z = 0.66 (P = 0.51)								
Total (95% CI)			**111**			**168**	**100.0%**	**-0.14 [-0.39 , 0.11]**
Heterogeneity: Tau² = 0.00; Chi² = 6.26, df = 8 (P = 0.62); I² = 0%								
Test for overall effect: Z = 1.07 (P = 0.28)								
Test for subgroup differences: Chi² = 1.82, df = 1 (P = 0.18), I² = 45.0%								



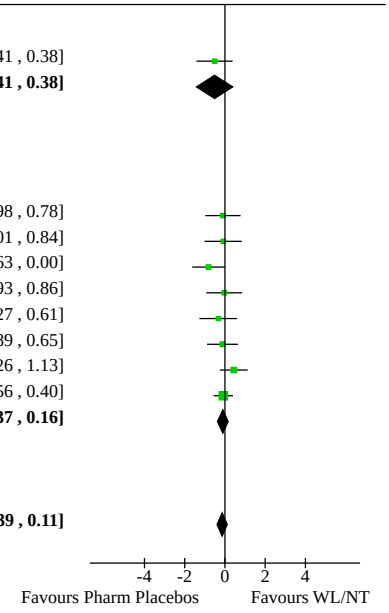
Analysis 18.6. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 6: Type of outcome domain

Study or Subgroup	Pharmacological placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio IV, Random, 95% CI
18.6.1 Blinded observer-reported						
Double 1993	3	22	3	22	2.9%	1.00 [0.23 , 4.42]
Rabkin 1990	14	27	12	23	18.5%	0.99 [0.58 , 1.70]
Subtotal (95% CI)		**49**		**45**	**21.3%**	**0.99 [0.60 , 1.64]**
Total events:	17		15			
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 0.99); I² = 0%						
Test for overall effect: Z = 0.02 (P = 0.98)						
18.6.2 Non-blinded observer-reported						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
18.6.3 Patient-reported						
Berg 1983	9	11	9	15	20.6%	1.36 [0.83 , 2.24]
Wilson 1980	31	40	10	10	58.1%	0.80 [0.65 , 0.99]
Subtotal (95% CI)		**51**		**25**	**78.7%**	**1.00 [0.60 , 1.65]**
Total events:	40		19			
Heterogeneity: Tau² = 0.10; Chi² = 3.64, df = 1 (P = 0.06); I² = 73%						
Test for overall effect: Z = 0.01 (P = 0.99)						
Total (95% CI)		**100**		**70**	**100.0%**	**0.94 [0.73 , 1.21]**
Total events:	57		34			
Heterogeneity: Tau² = 0.02; Chi² = 3.87, df = 3 (P = 0.28); I² = 22%						
Test for overall effect: Z = 0.49 (P = 0.63)						
Test for subgroup differences: Chi² = 0.00, df = 1 (P = 1.00), I² = 0%						



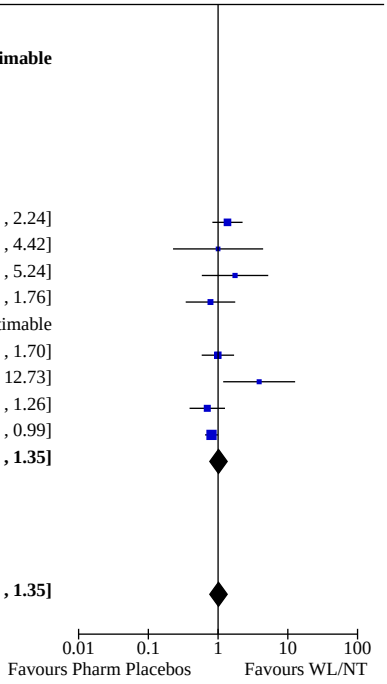
Analysis 18.7. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 7: Awareness of placebo intervention

Study or Subgroup	Pharmacological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
18.7.1 Open placebo								
Kelley 2012	-1.64	4.52	11	0.67	4	9	7.9%	-0.51 [-1.41 , 0.38]
Subtotal (95% CI)			**11**			**9**	**7.9%**	**-0.51 [-1.41 , 0.38]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.12 (P = 0.26)								
18.7.2 Closed placebo								
Borden 1986	29.7	17.5	10	32	26	10	8.3%	-0.10 [-0.98 , 0.78]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	7.5%	-0.09 [-1.01 , 0.84]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	9.6%	-0.81 [-1.63 , 0.00]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	8.0%	-0.04 [-0.93 , 0.86]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	7.3%	-0.33 [-1.27 , 0.61]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	10.8%	-0.12 [-0.89 , 0.65]
Powers 2008b	64	20	18	51	37	15	13.2%	0.44 [-0.26 , 1.13]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	27.4%	-0.08 [-0.56 , 0.40]
Subtotal (95% CI)			**100**			**159**	**92.1%**	**-0.11 [-0.37 , 0.16]**
Heterogeneity: Tau² = 0.00; Chi² = 5.52, df = 7 (P = 0.60); I² = 0%								
Test for overall effect: Z = 0.79 (P = 0.43)								
Total (95% CI)			**111**			**168**	**100.0%**	**-0.14 [-0.39 , 0.11]**
Heterogeneity: Tau² = 0.00; Chi² = 6.26, df = 8 (P = 0.62); I² = 0%								
Test for overall effect: Z = 1.07 (P = 0.28)								
Test for subgroup differences: Chi² = 0.73, df = 1 (P = 0.39), I² = 0%								



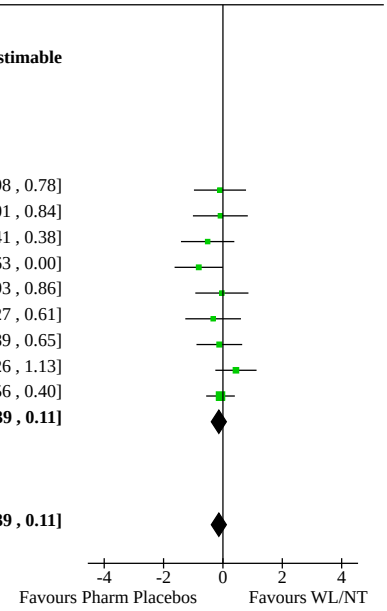
Analysis 18.8. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 8: Awareness of placebo intervention

Study or Subgroup	Pharmacological placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio IV, Random, 95% CI
18.8.1 Open placebo						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
18.8.2 Closed placebo						
Berg 1983	9	11	9	15	17.1%	1.36 [0.83 , 2.24]
Double 1993	3	22	3	22	3.4%	1.00 [0.23 , 4.42]
Klerman 1974a	7	25	4	25	5.7%	1.75 [0.58 , 5.24]
Klerman 1974b	7	25	9	25	9.1%	0.78 [0.34 , 1.76]
Mealiea 1971	0	9	0	10		Not estimable
Rabkin 1990	14	27	12	23	15.9%	0.99 [0.58 , 1.70]
Watzl 1988	11	34	3	36	5.0%	3.88 [1.18 , 12.73]
Whittaker 1963	7	13	10	13	14.3%	0.70 [0.39 , 1.26]
Wilson 1980	31	40	10	10	29.6%	0.80 [0.65 , 0.99]
Subtotal (95% CI)		**206**		**179**	**100.0%**	**1.01 [0.76 , 1.35]**
Total events:	89		60			
Heterogeneity: $Tau^2 = 0.06$; $Chi^2 = 11.95$, df = 7 (P = 0.10); $I^2 = 41\%$						
Test for overall effect: Z = 0.09 (P = 0.93)						
Total (95% CI)		**206**		**179**	**100.0%**	**1.01 [0.76 , 1.35]**
Total events:	89		60			
Heterogeneity: $Tau^2 = 0.06$; $Chi^2 = 11.95$, df = 7 (P = 0.10); $I^2 = 41\%$						
Test for overall effect: Z = 0.09 (P = 0.93)						
Test for subgroup differences: Not applicable						



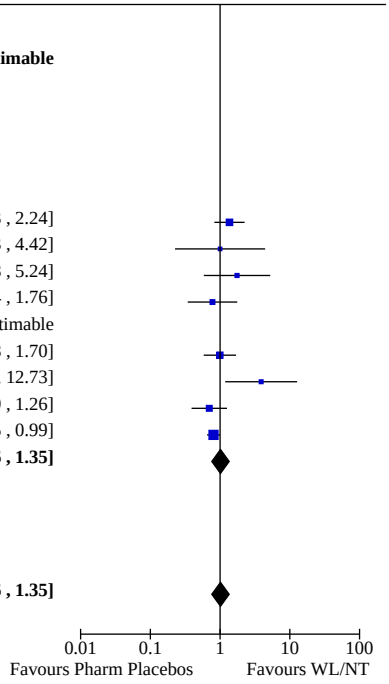
Analysis 18.9. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 9: The trial objective

Study or Subgroup	Pharmacological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
18.9.1 A trial's objective is clearly to assess the effects of placebo								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
18.9.2 No such objectives are stated								
Borden 1986	29.7	17.5	10	32	26	10	8.3%	-0.10 [-0.98 , 0.78]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	7.5%	-0.09 [-1.01 , 0.84]
Kelley 2012	-1.64	4.52	11	0.67	4	9	7.9%	-0.51 [-1.41 , 0.38]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	9.6%	-0.81 [-1.63 , 0.00]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	8.0%	-0.04 [-0.93 , 0.86]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	7.3%	-0.33 [-1.27 , 0.61]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	10.8%	-0.12 [-0.89 , 0.65]
Powers 2008b	64	20	18	51	37	15	13.2%	0.44 [-0.26 , 1.13]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	27.4%	-0.08 [-0.56 , 0.40]
Subtotal (95% CI)			**111**			**168**	**100.0%**	**-0.14 [-0.39 , 0.11]**
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 6.26$, df = 8 (P = 0.62); $I^2 = 0\%$								
Test for overall effect: Z = 1.07 (P = 0.28)								
Total (95% CI)			**111**			**168**	**100.0%**	**-0.14 [-0.39 , 0.11]**
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 6.26$, df = 8 (P = 0.62); $I^2 = 0\%$								
Test for overall effect: Z = 1.07 (P = 0.28)								
Test for subgroup differences: Not applicable								



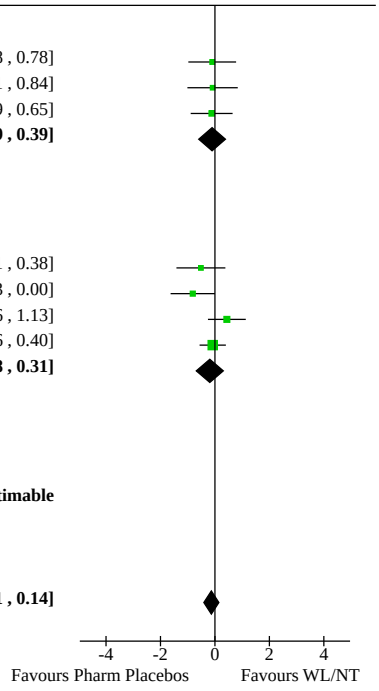
Analysis 18.10. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 10: The trial objective

Study or Subgroup	Pharmacological placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio IV, Random, 95% CI
18.10.1 A trial's objective is clearly to assess the effects of placebo						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
18.10.2 No such objectives are stated						
Berg 1983	9	11	9	15	17.1%	1.36 [0.83 , 2.24]
Double 1993	3	22	3	22	3.4%	1.00 [0.23 , 4.42]
Klerman 1974a	7	25	4	25	5.7%	1.75 [0.58 , 5.24]
Klerman 1974b	7	25	9	25	9.1%	0.78 [0.34 , 1.76]
Mealiea 1971	0	9	0	10		Not estimable
Rabkin 1990	14	27	12	23	15.9%	0.99 [0.58 , 1.70]
Watzl 1988	11	34	3	36	5.0%	3.88 [1.18 , 12.73]
Whittaker 1963	7	13	10	13	14.3%	0.70 [0.39 , 1.26]
Wilson 1980	31	40	10	10	29.6%	0.80 [0.65 , 0.99]
Subtotal (95% CI)		**206**		**179**	**100.0%**	**1.01 [0.76 , 1.35]**
Total events:	89		60			
Heterogeneity: $Tau^2 = 0.06$; $Chi^2 = 11.95$, df = 7 (P = 0.10); $I^2 = 41\%$						
Test for overall effect: Z = 0.09 (P = 0.93)						
Total (95% CI)		**206**		**179**	**100.0%**	**1.01 [0.76 , 1.35]**
Total events:	89		60			
Heterogeneity: $Tau^2 = 0.06$; $Chi^2 = 11.95$, df = 7 (P = 0.10); $I^2 = 41\%$						
Test for overall effect: Z = 0.09 (P = 0.93)						
Test for subgroup differences: Not applicable						



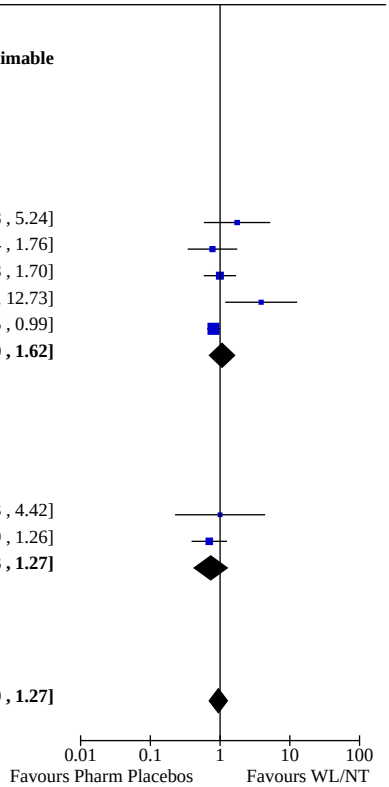
Analysis 18.11. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 11: Mean age of participants

Study or Subgroup	Pharmacological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
18.11.1 Below 18 years								
Borden 1986	29.7	17.5	10	32	26	10	9.8%	-0.10 [-0.98 , 0.78]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	8.9%	-0.09 [-1.01 , 0.84]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	12.8%	-0.12 [-0.89 , 0.65]
Subtotal (95% CI)			**32**			**32**	**31.5%**	**-0.10 [-0.59 , 0.39]**
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 2 (P = 1.00); I² = 0%								
Test for overall effect: Z = 0.42 (P = 0.68)								
18.11.2 18 to 50 years								
Kelley 2012	-1.64	4.52	11	0.67	4	9	9.4%	-0.51 [-1.41 , 0.38]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	11.4%	-0.81 [-1.63 , 0.00]
Powers 2008b	64	20	18	51	37	15	15.6%	0.44 [-0.26 , 1.13]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	32.1%	-0.08 [-0.56 , 0.40]
Subtotal (95% CI)			**61**			**116**	**68.5%**	**-0.19 [-0.68 , 0.31]**
Heterogeneity: Tau² = 0.13; Chi² = 6.02, df = 3 (P = 0.11); I² = 50%								
Test for overall effect: Z = 0.75 (P = 0.46)								
18.11.3 Above 50 years								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
Total (95% CI)			**93**			**148**	**100.0%**	**-0.13 [-0.41 , 0.14]**
Heterogeneity: Tau² = 0.00; Chi² = 6.04, df = 6 (P = 0.42); I² = 1%								
Test for overall effect: Z = 0.94 (P = 0.35)								
Test for subgroup differences: Chi² = 0.06, df = 1 (P = 0.81), I² = 0%								



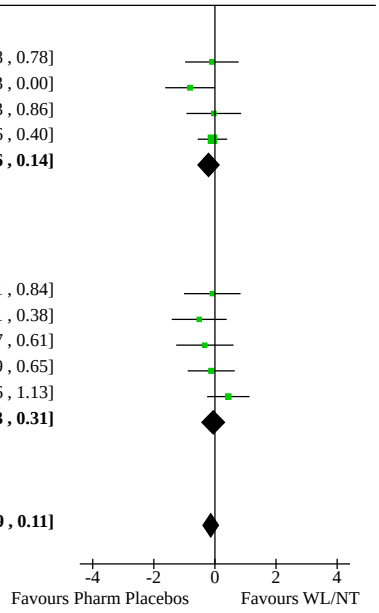
Analysis 18.12. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 12: Mean age of participants

Study or Subgroup	Pharmacological placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio IV, Random, 95% CI
18.12.1 Below 18 years						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
18.12.2 18 to 50 years						
Klerman 1974a	7	25	4	25	6.4%	1.75 [0.58 , 5.24]
Klerman 1974b	7	25	9	25	10.4%	0.78 [0.34 , 1.76]
Rabkin 1990	14	27	12	23	18.8%	0.99 [0.58 , 1.70]
Watzl 1988	11	34	3	36	5.6%	3.88 [1.18 , 12.73]
Wilson 1980	31	40	10	10	38.2%	0.80 [0.65 , 0.99]
Subtotal (95% CI)		**151**		**119**	**79.4%**	**1.07 [0.70 , 1.62]**
Total events:	70		38			
Heterogeneity: $Tau^2 = 0.11$; $Chi^2 = 8.48$, df = 4 (P = 0.08); $I^2 = 53\%$						
Test for overall effect: Z = 0.30 (P = 0.76)						
18.12.3 Above 50 years						
Double 1993	3	22	3	22	3.7%	1.00 [0.23 , 4.42]
Whittaker 1963	7	13	10	13	16.8%	0.70 [0.39 , 1.26]
Subtotal (95% CI)		**35**		**35**	**20.6%**	**0.73 [0.43 , 1.27]**
Total events:	10		13			
Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 0.19$, df = 1 (P = 0.66); $I^2 = 0\%$						
Test for overall effect: Z = 1.11 (P = 0.27)						
Total (95% CI)		**186**		**154**	**100.0%**	**0.94 [0.70 , 1.27]**
Total events:	80		51			
Heterogeneity: $Tau^2 = 0.05$; $Chi^2 = 9.03$, df = 6 (P = 0.17); $I^2 = 34\%$						
Test for overall effect: Z = 0.38 (P = 0.70)						
Test for subgroup differences: $Chi^2 = 1.14$, df = 1 (P = 0.29), $I^2 = 12.2\%$						



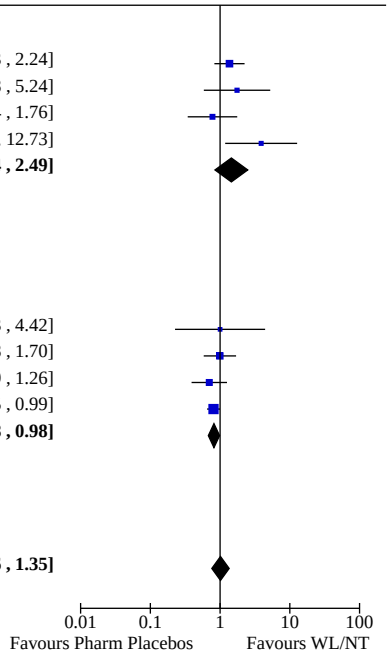
Analysis 18.13. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 13: Duration of treatment

Study or Subgroup	Pharmacological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
18.13.1 Above 3 months								
Borden 1986	29.7	17.5	10	32	26	10	8.3%	-0.10 [-0.98 , 0.78]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	9.6%	-0.81 [-1.63 , 0.00]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	8.0%	-0.04 [-0.93 , 0.86]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	27.4%	-0.08 [-0.56 , 0.40]
Subtotal (95% CI)			**50**			**114**	**53.3%**	**-0.21 [-0.56 , 0.14]**
Heterogeneity: Tau² = 0.00; Chi² = 2.59, df = 3 (P = 0.46); I² = 0%								
Test for overall effect: Z = 1.19 (P = 0.23)								
18.13.2 Below 3 months								
Carlson 1993	5.2	3.3	9	5.5	3.3	9	7.5%	-0.09 [-1.01 , 0.84]
Kelley 2012	-1.64	4.52	11	0.67	4	9	7.9%	-0.51 [-1.41 , 0.38]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	7.3%	-0.33 [-1.27 , 0.61]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	10.8%	-0.12 [-0.89 , 0.65]
Powers 2008b	64	20	18	51	37	15	13.2%	0.44 [-0.26 , 1.13]
Subtotal (95% CI)			**61**			**54**	**46.7%**	**-0.06 [-0.43 , 0.31]**
Heterogeneity: Tau² = 0.00; Chi² = 3.31, df = 4 (P = 0.51); I² = 0%								
Test for overall effect: Z = 0.30 (P = 0.77)								
Total (95% CI)			**111**			**168**	**100.0%**	**-0.14 [-0.39 , 0.11]**
Heterogeneity: Tau² = 0.00; Chi² = 6.26, df = 8 (P = 0.62); I² = 0%								
Test for overall effect: Z = 1.07 (P = 0.28)								
Test for subgroup differences: Chi² = 0.35, df = 1 (P = 0.55), I² = 0%								



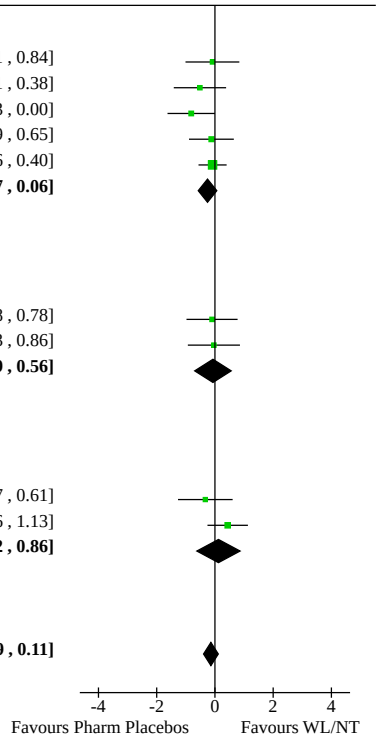
Analysis 18.14. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 14: Duration of treatment

Study or Subgroup	Pharmacological placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio IV, Random, 95% CI
18.14.1 Above 3 months						
Berg 1983	9	11	9	15	17.1%	1.36 [0.83 , 2.24]
Klerman 1974a	7	25	4	25	5.7%	1.75 [0.58 , 5.24]
Klerman 1974b	7	25	9	25	9.1%	0.78 [0.34 , 1.76]
Watzl 1988	11	34	3	36	5.0%	3.88 [1.18 , 12.73]
Subtotal (95% CI)		**95**		**101**	**36.9%**	**1.45 [0.84 , 2.49]**
Total events:	34		25			
Heterogeneity: Tau² = 0.12; Chi² = 4.99, df = 3 (P = 0.17); I² = 40%						
Test for overall effect: Z = 1.35 (P = 0.18)						
18.14.2 Below 3 months						
Double 1993	3	22	3	22	3.4%	1.00 [0.23 , 4.42]
Rabkin 1990	14	27	12	23	15.9%	0.99 [0.58 , 1.70]
Whittaker 1963	7	13	10	13	14.3%	0.70 [0.39 , 1.26]
Wilson 1980	31	40	10	10	29.6%	0.80 [0.65 , 0.99]
Subtotal (95% CI)		**102**		**68**	**63.1%**	**0.82 [0.68 , 0.98]**
Total events:	55		35			
Heterogeneity: Tau² = 0.00; Chi² = 0.88, df = 3 (P = 0.83); I² = 0%						
Test for overall effect: Z = 2.14 (P = 0.03)						
Total (95% CI)		**197**		**169**	**100.0%**	**1.01 [0.76 , 1.35]**
Total events:	89		60			
Heterogeneity: Tau² = 0.06; Chi² = 11.95, df = 7 (P = 0.10); I² = 41%						
Test for overall effect: Z = 0.09 (P = 0.93)						
Test for subgroup differences: Chi² = 3.88, df = 1 (P = 0.05), I² = 74.2%						



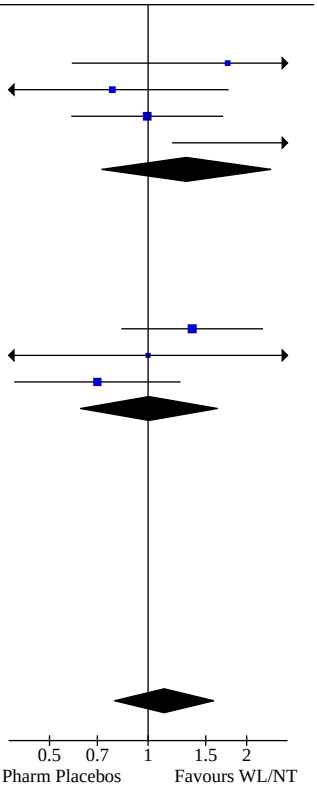
Analysis 18.15. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 15: Mental health diagnoses

Study or Subgroup	Pharmacological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
18.15.1 Formal diagnosis according to DSM/ICD								
Carlson 1993	5.2	3.3	9	5.5	3.3	9	7.5%	-0.09 [-1.01 , 0.84]
Kelley 2012	-1.64	4.52	11	0.67	4	9	7.9%	-0.51 [-1.41 , 0.38]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	9.6%	-0.81 [-1.63 , 0.00]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	10.8%	-0.12 [-0.89 , 0.65]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	27.4%	-0.08 [-0.56 , 0.40]
Subtotal (95% CI)			**65**			**123**	**63.2%**	**-0.25 [-0.57 , 0.06]**
Heterogeneity: Tau² = 0.00; Chi² = 2.87, df = 4 (P = 0.58); I² = 0%								
Test for overall effect: Z = 1.57 (P = 0.12)								
18.15.2 Fulfill symptoms of disorder ICD/DSM while not stating systems								
Borden 1986	29.7	17.5	10	32	26	10	8.3%	-0.10 [-0.98 , 0.78]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	8.0%	-0.04 [-0.93 , 0.86]
Subtotal (95% CI)			**18**			**22**	**16.3%**	**-0.07 [-0.69 , 0.56]**
Heterogeneity: Tau² = 0.00; Chi² = 0.01, df = 1 (P = 0.92); I² = 0%								
Test for overall effect: Z = 0.21 (P = 0.83)								
18.15.3 Population is classified as having a mental disorder, but full diagnostic criteria not reported								
Nandi 1976	53.2	11.2	10	57.5	13.8	8	7.3%	-0.33 [-1.27 , 0.61]
Powers 2008b	64	20	18	51	37	15	13.2%	0.44 [-0.26 , 1.13]
Subtotal (95% CI)			**28**			**23**	**20.5%**	**0.12 [-0.62 , 0.86]**
Heterogeneity: Tau² = 0.12; Chi² = 1.67, df = 1 (P = 0.20); I² = 40%								
Test for overall effect: Z = 0.32 (P = 0.75)								
Total (95% CI)			**111**			**168**	**100.0%**	**-0.14 [-0.39 , 0.11]**
Heterogeneity: Tau² = 0.00; Chi² = 6.26, df = 8 (P = 0.62); I² = 0%								
Test for overall effect: Z = 1.07 (P = 0.28)								
Test for subgroup differences: Chi² = 0.97, df = 2 (P = 0.62), I² = 0%								

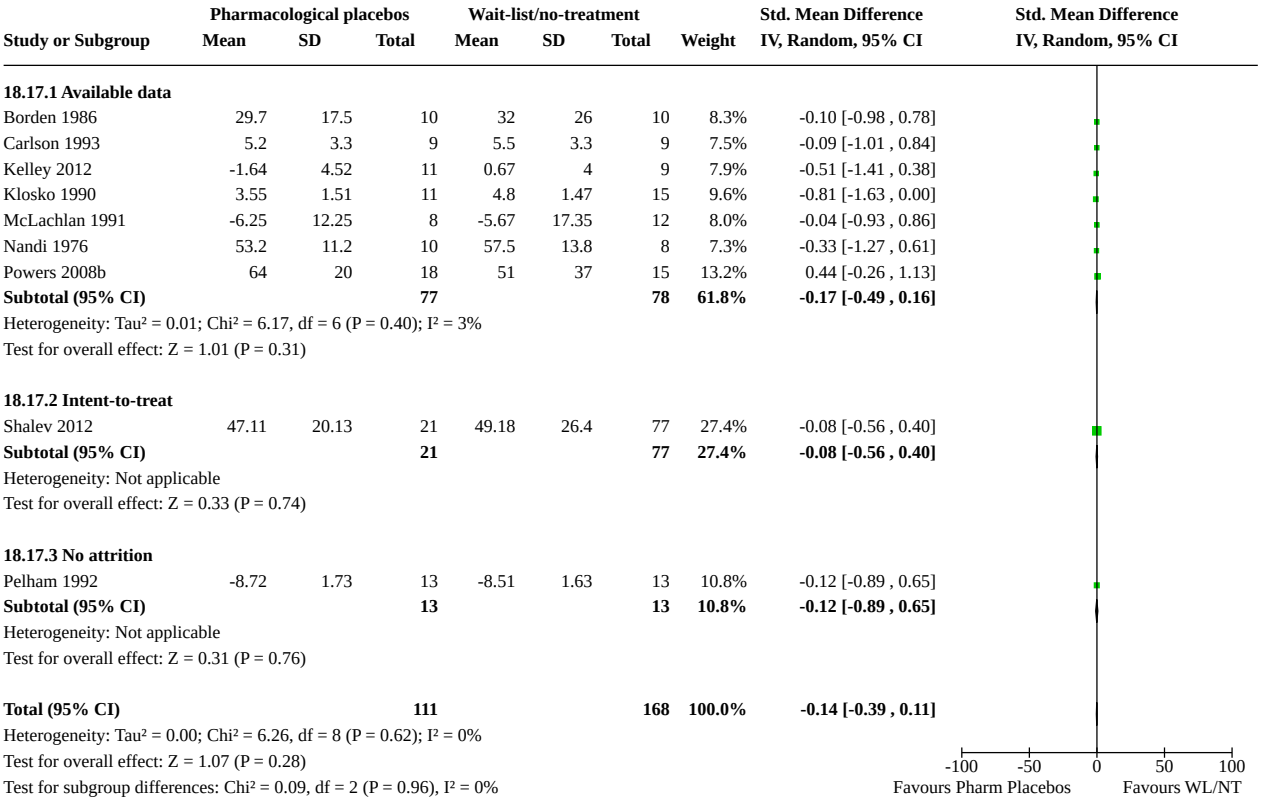


Analysis 18.16. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 16: Mental health diagnoses

Study or Subgroup	Pharmacological placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio IV, Random, 95% CI
18.16.1 Formal diagnosis according to DSM/ICD						
Klerman 1974a	7	25	4	25	8.3%	1.75 [0.58 , 5.24]
Klerman 1974b	7	25	9	25	13.1%	0.78 [0.34 , 1.76]
Rabkin 1990	14	27	12	23	22.3%	0.99 [0.58 , 1.70]
Watzl 1988	11	34	3	36	7.3%	3.88 [1.18 , 12.73]
Subtotal (95% CI)		**111**		**109**	**51.0%**	**1.31 [0.72 , 2.37]**
Total events:	39		28			
Heterogeneity: Tau² = 0.17; Chi² = 5.75, df = 3 (P = 0.12); I² = 48%						
Test for overall effect: Z = 0.89 (P = 0.38)						
18.16.2 Fulfill symptoms of disorder ICD/DSM while not stating systems						
Berg 1983	9	11	9	15	23.9%	1.36 [0.83 , 2.24]
Double 1993	3	22	3	22	4.9%	1.00 [0.23 , 4.42]
Whittaker 1963	7	13	10	13	20.2%	0.70 [0.39 , 1.26]
Subtotal (95% CI)		**46**		**50**	**49.0%**	**1.01 [0.62 , 1.63]**
Total events:	19		22			
Heterogeneity: Tau² = 0.06; Chi² = 2.90, df = 2 (P = 0.24); I² = 31%						
Test for overall effect: Z = 0.03 (P = 0.98)						
18.16.3 Population is classified as having a mental disorder, but full diagnostic criteria not reported						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
Total (95% CI)		**157**		**159**	**100.0%**	**1.12 [0.79 , 1.59]**
Total events:	58		50			
Heterogeneity: Tau² = 0.07; Chi² = 8.87, df = 6 (P = 0.18); I² = 32%						
Test for overall effect: Z = 0.64 (P = 0.52)						
Test for subgroup differences: Chi² = 0.45, df = 1 (P = 0.50), I² = 0%						



Analysis 18.17. Comparison 18: Subgroup analyses for pharmacological placebos compared with wait-list or no-treatment for continuous data, Outcome 17: Imputed data



Comparison 19. Subgroup analyses for physical placebos compared with wait-list or no-treatment for continuous data

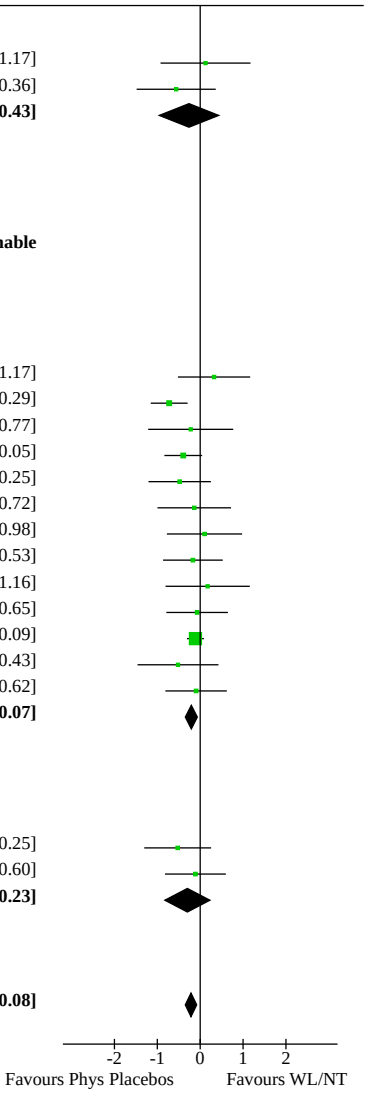
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
19.1 Type of active intervention	17	896	Std. Mean Difference (IV, Random, 95% CI)	-0.21 [-0.35, -0.08]
19.1.1 Psychological intervention	2	33	Std. Mean Difference (IV, Random, 95% CI)	-0.26 [-0.95, 0.43]
19.1.2 Pharmacological intervention	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
19.1.3 Physical intervention	13	798	Std. Mean Difference (IV, Random, 95% CI)	-0.21 [-0.35, -0.07]
19.1.4 Other or combination intervention	2	65	Std. Mean Difference (IV, Random, 95% CI)	-0.30 [-0.82, 0.23]
19.2 Risk of bias	17	896	Std. Mean Difference (IV, Random, 95% CI)	-0.21 [-0.35, -0.08]
19.2.1 Low risk of bias	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
19.2.2 High risk of bias	17	896	Std. Mean Difference (IV, Random, 95% CI)	-0.21 [-0.35, -0.08]
19.3 Type of outcome domain	13	411	Std. Mean Difference (IV, Random, 95% CI)	-0.31 [-0.51, -0.12]
19.3.1 Blinded observer-reported	8	314	Std. Mean Difference (IV, Random, 95% CI)	-0.42 [-0.64, -0.19]
19.3.2 Non-blinded observer-reported	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
19.3.3 Patient-reported	5	97	Std. Mean Difference (IV, Random, 95% CI)	-0.00 [-0.40, 0.40]
19.4 Awareness of placebo intervention	17	896	Std. Mean Difference (IV, Random, 95% CI)	-0.21 [-0.35, -0.08]
19.4.1 Open placebo	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
19.4.2 Closed placebo	17	896	Std. Mean Difference (IV, Random, 95% CI)	-0.21 [-0.35, -0.08]
19.5 The trial objective	17	896	Std. Mean Difference (IV, Random, 95% CI)	-0.21 [-0.35, -0.08]
19.5.1 A trial's objective is clearly to assess the effects of placebo	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
19.5.2 No such objectives are stated	17	896	Std. Mean Difference (IV, Random, 95% CI)	-0.21 [-0.35, -0.08]
19.6 Mean age of participants	16	503	Std. Mean Difference (IV, Random, 95% CI)	-0.30 [-0.48, -0.12]
19.6.1 Below 18 years	2	62	Std. Mean Difference (IV, Random, 95% CI)	-0.32 [-0.82, 0.19]
19.6.2 18 to 50 years	11	325	Std. Mean Difference (IV, Random, 95% CI)	-0.26 [-0.48, -0.04]
19.6.3 Above 50 years	3	116	Std. Mean Difference (IV, Random, 95% CI)	-0.40 [-0.77, -0.03]
19.7 Duration of treatment	17	926	Std. Mean Difference (IV, Random, 95% CI)	-0.22 [-0.35, -0.09]
19.7.1 Above 3 months	1	30	Std. Mean Difference (IV, Random, 95% CI)	-0.48 [-1.21, 0.25]
19.7.2 Below 3 months	17	896	Std. Mean Difference (IV, Random, 95% CI)	-0.21 [-0.35, -0.08]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
19.8 Type of physical placebo	15	473	Std. Mean Difference (IV, Random, 95% CI)	-0.31 [-0.50, -0.13]
19.8.1 Acupuncture or acupressure placebo	5	237	Std. Mean Difference (IV, Random, 95% CI)	-0.42 [-0.73, -0.11]
19.8.2 Excercise and relaxation placebo	3	67	Std. Mean Difference (IV, Random, 95% CI)	-0.21 [-0.70, 0.27]
19.8.3 Technical device placebo	5	119	Std. Mean Difference (IV, Random, 95% CI)	-0.23 [-0.60, 0.14]
19.8.4 Electromagnetic stimulation placebo	2	50	Std. Mean Difference (IV, Random, 95% CI)	-0.02 [-0.57, 0.54]
19.9 Mental health diagnoses	16	503	Std. Mean Difference (IV, Random, 95% CI)	-0.30 [-0.48, -0.12]
19.9.1 Formal diagnosis according to DSM/ICD	7	293	Std. Mean Difference (IV, Random, 95% CI)	-0.29 [-0.56, -0.02]
19.9.2 Fulfill symptoms of disorder ICD/DSM while not stating systems	3	68	Std. Mean Difference (IV, Random, 95% CI)	-0.15 [-0.62, 0.33]
19.9.3 Population is classified as having a mental disorder, but full diagnostic criteria not reported	6	142	Std. Mean Difference (IV, Random, 95% CI)	-0.32 [-0.66, 0.02]
19.10 Imputed data	16	503	Std. Mean Difference (IV, Random, 95% CI)	-0.30 [-0.48, -0.12]
19.10.1 Available data	10	231	Std. Mean Difference (IV, Random, 95% CI)	-0.13 [-0.40, 0.13]
19.10.2 Intent-to-treat	2	111	Std. Mean Difference (IV, Random, 95% CI)	-0.27 [-1.28, 0.75]
19.10.3 No attrition	4	161	Std. Mean Difference (IV, Random, 95% CI)	-0.40 [-0.71, -0.08]

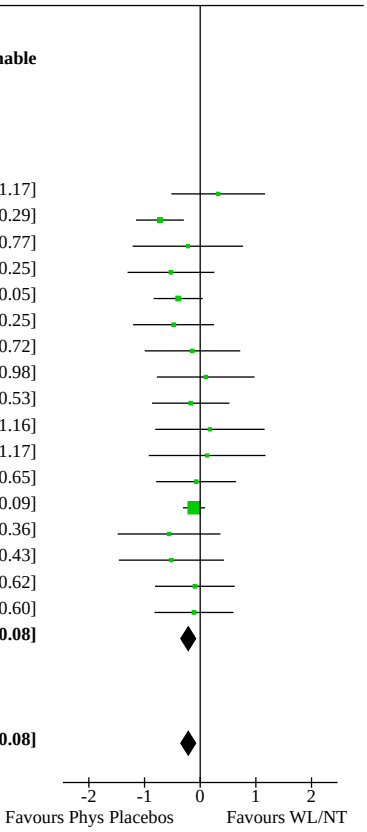
Analysis 19.1. Comparison 19: Subgroup analyses for physical placebos compared with wait-list or no-treatment for continuous data, Outcome 1: Type of active intervention

Study or Subgroup	Physical placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
19.1.1 Psychological intervention								
Pillman 2001	-85.3	14	7	-87.6	20	7	1.6%	0.12 [-0.92 , 1.17]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	2.1%	-0.56 [-1.48 , 0.36]
Subtotal (95% CI)			**16**			**17**	**3.7%**	**-0.26 [-0.95 , 0.43]**
Heterogeneity: Tau² = 0.00; Chi² = 0.92, df = 1 (P = 0.34); I² = 0%								
Test for overall effect: Z = 0.74 (P = 0.46)								
19.1.2 Pharmacological intervention								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
19.1.3 Physical intervention								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	2.5%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	9.5%	-0.72 [-1.15 , -0.29]
Freire 2007	24.6	11	7	28.2	18	9	1.8%	-0.22 [-1.21 , 0.77]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	9.1%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	3.3%	-0.48 [-1.21 , 0.25]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	2.4%	-0.14 [-1.00 , 0.72]
Lick 1977	65.86	25.45	10	62.84	31.78	10	2.3%	0.10 [-0.78 , 0.98]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	3.6%	-0.17 [-0.87 , 0.53]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.8%	0.17 [-0.81 , 1.16]
Poland 2013	22	8.2	16	22.6	8	14	3.4%	-0.07 [-0.79 , 0.65]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	44.9%	-0.11 [-0.31 , 0.09]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	2.0%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	3.4%	-0.10 [-0.81 , 0.62]
Subtotal (95% CI)			**398**			**400**	**90.0%**	**-0.21 [-0.35 , -0.07]**
Heterogeneity: Tau² = 0.00; Chi² = 10.95, df = 12 (P = 0.53); I² = 0%								
Test for overall effect: Z = 2.90 (P = 0.004)								
19.1.4 Other or combination intervention								
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	2.9%	-0.53 [-1.30 , 0.25]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	3.5%	-0.11 [-0.82 , 0.60]
Subtotal (95% CI)			**44**			**21**	**6.4%**	**-0.30 [-0.82 , 0.23]**
Heterogeneity: Tau² = 0.00; Chi² = 0.59, df = 1 (P = 0.44); I² = 0%								
Test for overall effect: Z = 1.12 (P = 0.26)								
Total (95% CI)			**458**			**438**	**100.0%**	**-0.21 [-0.35 , -0.08]**
Heterogeneity: Tau² = 0.00; Chi² = 12.59, df = 16 (P = 0.70); I² = 0%								
Test for overall effect: Z = 3.17 (P = 0.002)								
Test for subgroup differences: Chi² = 0.13, df = 2 (P = 0.94), I² = 0%								



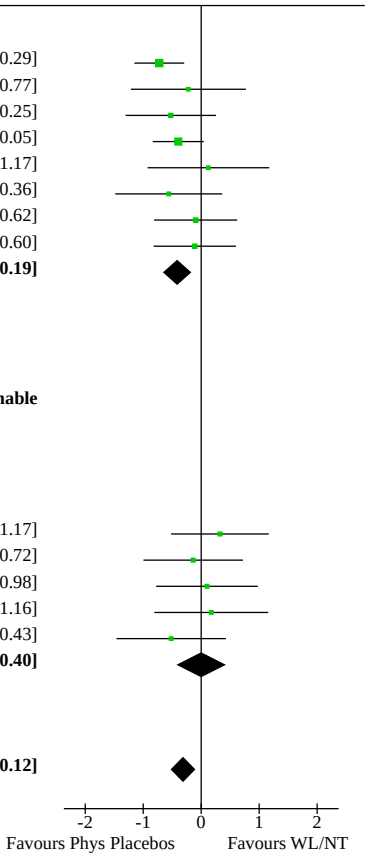
Analysis 19.2. Comparison 19: Subgroup analyses for physical placebos compared with wait-list or no-treatment for continuous data, Outcome 2: Risk of bias

Study or Subgroup	Physical placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
19.2.1 Low risk of bias								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
19.2.2 High risk of bias								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	2.5%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	9.5%	-0.72 [-1.15 , -0.29]
Freire 2007	24.6	11	7	28.2	18	9	1.8%	-0.22 [-1.21 , 0.77]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	2.9%	-0.53 [-1.30 , 0.25]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	9.1%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	3.3%	-0.48 [-1.21 , 0.25]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	2.4%	-0.14 [-1.00 , 0.72]
Lick 1977	65.86	25.45	10	62.84	31.78	10	2.3%	0.10 [-0.78 , 0.98]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	3.6%	-0.17 [-0.87 , 0.53]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.8%	0.17 [-0.81 , 1.16]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.6%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	3.4%	-0.07 [-0.79 , 0.65]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	44.9%	-0.11 [-0.31 , 0.09]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	2.1%	-0.56 [-1.48 , 0.36]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	2.0%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	3.4%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	3.5%	-0.11 [-0.82 , 0.60]
Subtotal (95% CI)			**458**			**438**	**100.0%**	**-0.21 [-0.35 , -0.08]**
Heterogeneity: Tau² = 0.00; Chi² = 12.59, df = 16 (P = 0.70); I² = 0%								
Test for overall effect: Z = 3.17 (P = 0.002)								
Total (95% CI)			**458**			**438**	**100.0%**	**-0.21 [-0.35 , -0.08]**
Heterogeneity: Tau² = 0.00; Chi² = 12.59, df = 16 (P = 0.70); I² = 0%								
Test for overall effect: Z = 3.17 (P = 0.002)								
Test for subgroup differences: Not applicable								



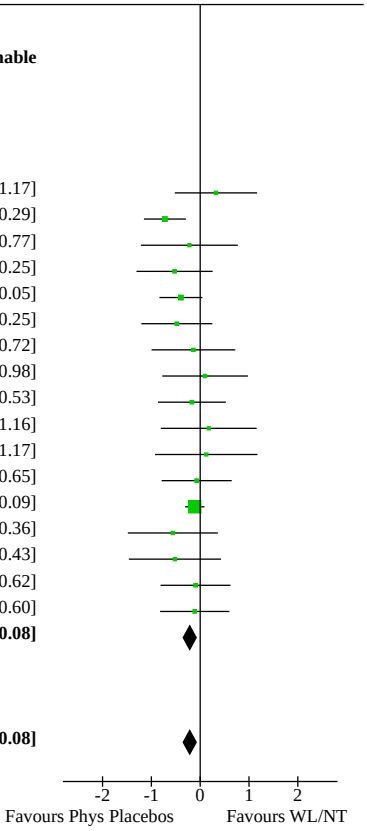
Analysis 19.3. Comparison 19: Subgroup analyses for physical placebos compared with wait-list or no-treatment for continuous data, Outcome 3: Type of outcome domain

Study or Subgroup	Physical placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
19.3.1 Blinded observer-reported								
Allen 2006	12	9.6	45	19	9.6	44	21.3%	-0.72 [-1.15 , -0.29]
Freire 2007	24.6	11	7	28.2	18	9	4.0%	-0.22 [-1.21 , 0.77]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	6.5%	-0.53 [-1.30 , 0.25]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	20.3%	-0.39 [-0.83 , 0.05]
Pillman 2001	-85.3	14	7	-87.6	20	7	3.6%	0.12 [-0.92 , 1.17]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	4.6%	-0.56 [-1.48 , 0.36]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	7.6%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	7.8%	-0.11 [-0.82 , 0.60]
Subtotal (95% CI)			**168**			**146**	**75.6%**	**-0.42 [-0.64 , -0.19]**
Heterogeneity: Tau² = 0.00; Chi² = 4.79, df = 7 (P = 0.69); I² = 0%								
Test for overall effect: Z = 3.57 (P = 0.0004)								
19.3.2 Non-blinded observer-reported								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
19.3.3 Patient-reported								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	5.5%	0.32 [-0.52 , 1.17]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	5.3%	-0.14 [-1.00 , 0.72]
Lick 1977	65.86	25.45	10	62.84	31.78	10	5.1%	0.10 [-0.78 , 0.98]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	4.1%	0.17 [-0.81 , 1.16]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	4.4%	-0.52 [-1.46 , 0.43]
Subtotal (95% CI)			**49**			**48**	**24.4%**	**-0.00 [-0.40 , 0.40]**
Heterogeneity: Tau² = 0.00; Chi² = 1.99, df = 4 (P = 0.74); I² = 0%								
Test for overall effect: Z = 0.00 (P = 1.00)								
Total (95% CI)			**217**			**194**	**100.0%**	**-0.31 [-0.51 , -0.12]**
Heterogeneity: Tau² = 0.00; Chi² = 9.89, df = 12 (P = 0.63); I² = 0%								
Test for overall effect: Z = 3.11 (P = 0.002)								
Test for subgroup differences: Chi² = 3.11, df = 1 (P = 0.08), I² = 67.8%								



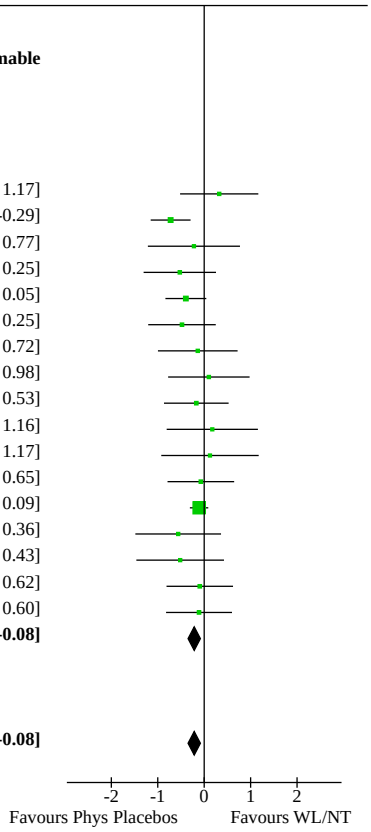
Analysis 19.4. Comparison 19: Subgroup analyses for physical placebos compared with wait-list or no-treatment for continuous data, Outcome 4: Awareness of placebo intervention

Study or Subgroup	Physical placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
19.4.1 Open placebo								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
19.4.2 Closed placebo								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	2.5%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	9.5%	-0.72 [-1.15 , -0.29]
Freire 2007	24.6	11	7	28.2	18	9	1.8%	-0.22 [-1.21 , 0.77]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	2.9%	-0.53 [-1.30 , 0.25]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	9.1%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	3.3%	-0.48 [-1.21 , 0.25]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	2.4%	-0.14 [-1.00 , 0.72]
Lick 1977	65.86	25.45	10	62.84	31.78	10	2.3%	0.10 [-0.78 , 0.98]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	3.6%	-0.17 [-0.87 , 0.53]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.8%	0.17 [-0.81 , 1.16]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.6%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	3.4%	-0.07 [-0.79 , 0.65]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	44.9%	-0.11 [-0.31 , 0.09]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	2.1%	-0.56 [-1.48 , 0.36]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	2.0%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	3.4%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	3.5%	-0.11 [-0.82 , 0.60]
Subtotal (95% CI)			**458**			**438**	**100.0%**	**-0.21 [-0.35 , -0.08]**
Heterogeneity: Tau² = 0.00; Chi² = 12.59, df = 16 (P = 0.70); I² = 0%								
Test for overall effect: Z = 3.17 (P = 0.002)								
Total (95% CI)			**458**			**438**	**100.0%**	**-0.21 [-0.35 , -0.08]**
Heterogeneity: Tau² = 0.00; Chi² = 12.59, df = 16 (P = 0.70); I² = 0%								
Test for overall effect: Z = 3.17 (P = 0.002)								
Test for subgroup differences: Not applicable								



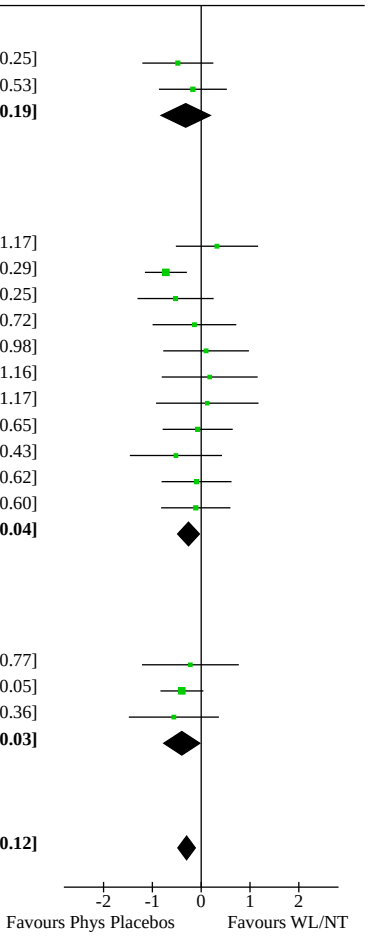
Analysis 19.5. Comparison 19: Subgroup analyses for physical placebos compared with wait-list or no-treatment for continuous data, Outcome 5: The trial objective

Study or Subgroup	Physical placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
19.5.1 A trial's objective is clearly to assess the effects of placebo								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
19.5.2 No such objectives are stated								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	2.5%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	9.5%	-0.72 [-1.15 , -0.29]
Freire 2007	24.6	11	7	28.2	18	9	1.8%	-0.22 [-1.21 , 0.77]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	2.9%	-0.53 [-1.30 , 0.25]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	9.1%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	3.3%	-0.48 [-1.21 , 0.25]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	2.4%	-0.14 [-1.00 , 0.72]
Lick 1977	65.86	25.45	10	62.84	31.78	10	2.3%	0.10 [-0.78 , 0.98]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	3.6%	-0.17 [-0.87 , 0.53]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.8%	0.17 [-0.81 , 1.16]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.6%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	3.4%	-0.07 [-0.79 , 0.65]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	44.9%	-0.11 [-0.31 , 0.09]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	2.1%	-0.56 [-1.48 , 0.36]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	2.0%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	3.4%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	3.5%	-0.11 [-0.82 , 0.60]
Subtotal (95% CI)			**458**			**438**	**100.0%**	**-0.21 [-0.35 , -0.08]**
Heterogeneity: Tau² = 0.00; Chi² = 12.59, df = 16 (P = 0.70); I² = 0%								
Test for overall effect: Z = 3.17 (P = 0.002)								
Total (95% CI)			**458**			**438**	**100.0%**	**-0.21 [-0.35 , -0.08]**
Heterogeneity: Tau² = 0.00; Chi² = 12.59, df = 16 (P = 0.70); I² = 0%								
Test for overall effect: Z = 3.17 (P = 0.002)								
Test for subgroup differences: Not applicable								



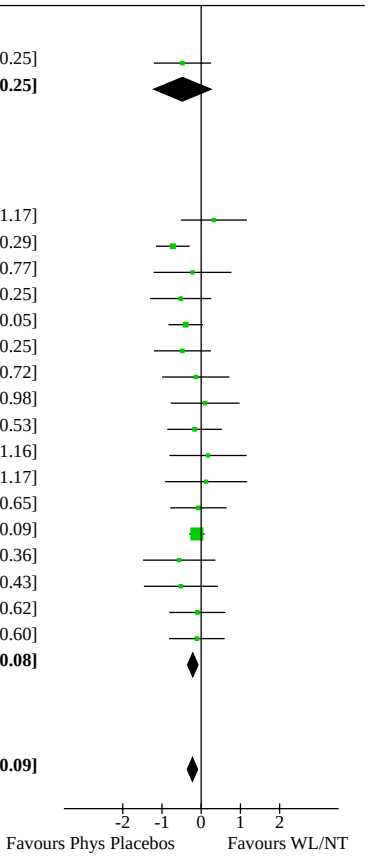
Analysis 19.6. Comparison 19: Subgroup analyses for physical placebos compared with wait-list or no-treatment for continuous data, Outcome 6: Mean age of participants

Study or Subgroup	Physical placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
19.6.1 Below 18 years								
Lacy 1990	5.71	1.95	15	6.79	2.42	15	6.0%	-0.48 [-1.21 , 0.25]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	6.6%	-0.17 [-0.87 , 0.53]
Subtotal (95% CI)			**30**			**32**	**12.6%**	**-0.32 [-0.82 , 0.19]**
Heterogeneity: Tau² = 0.00; Chi² = 0.36, df = 1 (P = 0.55); I² = 0%								
Test for overall effect: Z = 1.24 (P = 0.22)								
19.6.2 18 to 50 years								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	4.5%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	17.3%	-0.72 [-1.15 , -0.29]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	5.2%	-0.53 [-1.30 , 0.25]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	4.3%	-0.14 [-1.00 , 0.72]
Lick 1977	65.86	25.45	10	62.84	31.78	10	4.1%	0.10 [-0.78 , 0.98]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	3.3%	0.17 [-0.81 , 1.16]
Pillman 2001	-85.3	14	7	-87.6	20	7	2.9%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	6.2%	-0.07 [-0.79 , 0.65]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	3.6%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	6.2%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	6.3%	-0.11 [-0.82 , 0.60]
Subtotal (95% CI)			**176**			**149**	**64.0%**	**-0.26 [-0.48 , -0.04]**
Heterogeneity: Tau² = 0.00; Chi² = 9.66, df = 10 (P = 0.47); I² = 0%								
Test for overall effect: Z = 2.28 (P = 0.02)								
19.6.3 Above 50 years								
Freire 2007	24.6	11	7	28.2	18	9	3.2%	-0.22 [-1.21 , 0.77]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	16.5%	-0.39 [-0.83 , 0.05]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	3.7%	-0.56 [-1.48 , 0.36]
Subtotal (95% CI)			**57**			**59**	**23.4%**	**-0.40 [-0.77 , -0.03]**
Heterogeneity: Tau² = 0.00; Chi² = 0.24, df = 2 (P = 0.89); I² = 0%								
Test for overall effect: Z = 2.11 (P = 0.03)								
Total (95% CI)			**263**			**240**	**100.0%**	**-0.30 [-0.48 , -0.12]**
Heterogeneity: Tau² = 0.00; Chi² = 10.66, df = 15 (P = 0.78); I² = 0%								
Test for overall effect: Z = 3.29 (P = 0.001)								
Test for subgroup differences: Chi² = 0.40, df = 2 (P = 0.82), I² = 0%								



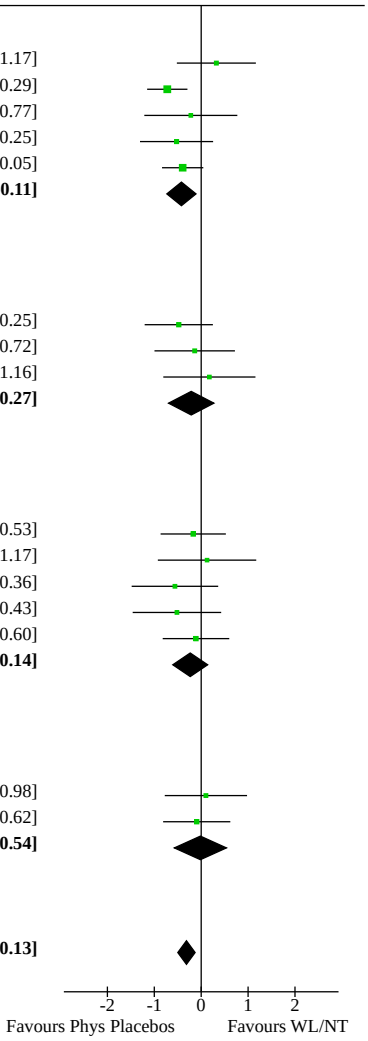
Analysis 19.7. Comparison 19: Subgroup analyses for physical placebos compared with wait-list or no-treatment for continuous data, Outcome 7: Duration of treatment

Study or Subgroup	Physical placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
19.7.1 Above 3 months								
Lacy 1990	5.71	1.95	15	6.79	2.42	15	3.2%	-0.48 [-1.21 , 0.25]
Subtotal (95% CI)			**15**			**15**	**3.2%**	**-0.48 [-1.21 , 0.25]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.29 (P = 0.20)								
19.7.2 Below 3 months								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	2.4%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	9.2%	-0.72 [-1.15 , -0.29]
Freire 2007	24.6	11	7	28.2	18	9	1.7%	-0.22 [-1.21 , 0.77]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	2.8%	-0.53 [-1.30 , 0.25]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	8.8%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	3.2%	-0.48 [-1.21 , 0.25]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	2.3%	-0.14 [-1.00 , 0.72]
Lick 1977	65.86	25.45	10	62.84	31.78	10	2.2%	0.10 [-0.78 , 0.98]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	3.5%	-0.17 [-0.87 , 0.53]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.8%	0.17 [-0.81 , 1.16]
Pillman 2001	-85.3	14	7	-87.6	20	7	1.5%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	3.3%	-0.07 [-0.79 , 0.65]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	43.4%	-0.11 [-0.31 , 0.09]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	2.0%	-0.56 [-1.48 , 0.36]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	1.9%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	3.3%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	3.4%	-0.11 [-0.82 , 0.60]
Subtotal (95% CI)			**458**			**438**	**96.8%**	**-0.21 [-0.35 , -0.08]**
Heterogeneity: Tau² = 0.00; Chi² = 12.59, df = 16 (P = 0.70); I² = 0%								
Test for overall effect: Z = 3.17 (P = 0.002)								
Total (95% CI)			**473**			**453**	**100.0%**	**-0.22 [-0.35 , -0.09]**
Heterogeneity: Tau² = 0.00; Chi² = 13.08, df = 17 (P = 0.73); I² = 0%								
Test for overall effect: Z = 3.35 (P = 0.0008)								
Test for subgroup differences: Chi² = 0.49, df = 1 (P = 0.48), I² = 0%								



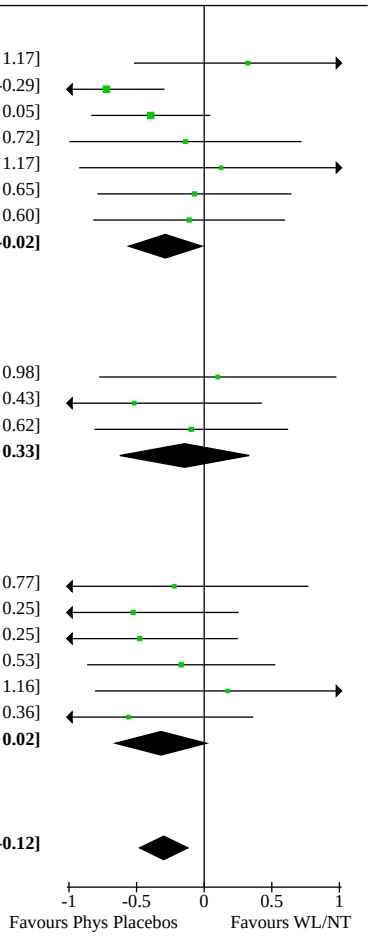
Analysis 19.8. Comparison 19: Subgroup analyses for physical placebos compared with wait-list or no-treatment for continuous data, Outcome 8: Type of physical placebo

Study or Subgroup	Physical placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
19.8.1 Acupuncture or acupressure placebo								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	4.8%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	18.4%	-0.72 [-1.15 , -0.29]
Freire 2007	24.6	11	7	28.2	18	9	3.5%	-0.22 [-1.21 , 0.77]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	5.6%	-0.53 [-1.30 , 0.25]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	17.5%	-0.39 [-0.83 , 0.05]
Subtotal (95% CI)			**123**			**114**	**49.8%**	**-0.42 [-0.73 , -0.11]**
Heterogeneity: Tau² = 0.03; Chi² = 5.09, df = 4 (P = 0.28); I² = 21%								
Test for overall effect: Z = 2.66 (P = 0.008)								
19.8.2 Excercise and relaxation placebo								
Lacy 1990	5.71	1.95	15	6.79	2.42	15	6.4%	-0.48 [-1.21 , 0.25]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	4.6%	-0.14 [-1.00 , 0.72]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	3.5%	0.17 [-0.81 , 1.16]
Subtotal (95% CI)			**34**			**33**	**14.6%**	**-0.21 [-0.70 , 0.27]**
Heterogeneity: Tau² = 0.00; Chi² = 1.14, df = 2 (P = 0.57); I² = 0%								
Test for overall effect: Z = 0.86 (P = 0.39)								
19.8.3 Technical device placebo								
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	7.0%	-0.17 [-0.87 , 0.53]
Pillman 2001	-85.3	14	7	-87.6	20	7	3.1%	0.12 [-0.92 , 1.17]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	4.0%	-0.56 [-1.48 , 0.36]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	3.8%	-0.52 [-1.46 , 0.43]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	6.7%	-0.11 [-0.82 , 0.60]
Subtotal (95% CI)			**65**			**54**	**24.6%**	**-0.23 [-0.60 , 0.14]**
Heterogeneity: Tau² = 0.00; Chi² = 1.42, df = 4 (P = 0.84); I² = 0%								
Test for overall effect: Z = 1.23 (P = 0.22)								
19.8.4 Electromagnetic stimulation placebo								
Lick 1977	65.86	25.45	10	62.84	31.78	10	4.4%	0.10 [-0.78 , 0.98]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	6.6%	-0.10 [-0.81 , 0.62]
Subtotal (95% CI)			**25**			**25**	**11.0%**	**-0.02 [-0.57 , 0.54]**
Heterogeneity: Tau² = 0.00; Chi² = 0.11, df = 1 (P = 0.73); I² = 0%								
Test for overall effect: Z = 0.06 (P = 0.95)								
Total (95% CI)			**247**			**226**	**100.0%**	**-0.31 [-0.50 , -0.13]**
Heterogeneity: Tau² = 0.00; Chi² = 10.25, df = 14 (P = 0.74); I² = 0%								
Test for overall effect: Z = 3.34 (P = 0.0008)								
Test for subgroup differences: Chi² = 1.80, df = 3 (P = 0.62), I² = 0%								



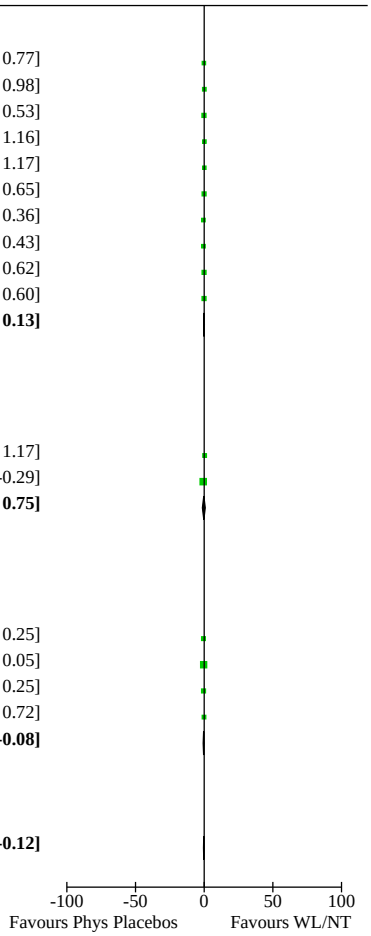
Analysis 19.9. Comparison 19: Subgroup analyses for physical placebos compared with wait-list or no-treatment for continuous data, Outcome 9: Mental health diagnoses

Study or Subgroup	Physical placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
19.9.1 Formal diagnosis according to DSM/ICD								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	4.5%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	17.3%	-0.72 [-1.15 , -0.29]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	16.5%	-0.39 [-0.83 , 0.05]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	4.3%	-0.14 [-1.00 , 0.72]
Pillman 2001	-85.3	14	7	-87.6	20	7	2.9%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	6.2%	-0.07 [-0.79 , 0.65]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	6.3%	-0.11 [-0.82 , 0.60]
Subtotal (95% CI)			**156**			**137**	**57.9%**	**-0.29 [-0.56 , -0.02]**
Heterogeneity: Tau² = 0.02; Chi² = 7.38, df = 6 (P = 0.29); I² = 19%								
Test for overall effect: Z = 2.08 (P = 0.04)								
19.9.2 Fulfill symptoms of disorder ICD/DSM while not stating systems								
Lick 1977	65.86	25.45	10	62.84	31.78	10	4.1%	0.10 [-0.78 , 0.98]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	3.6%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	6.2%	-0.10 [-0.81 , 0.62]
Subtotal (95% CI)			**34**			**34**	**13.9%**	**-0.15 [-0.62 , 0.33]**
Heterogeneity: Tau² = 0.00; Chi² = 0.92, df = 2 (P = 0.63); I² = 0%								
Test for overall effect: Z = 0.60 (P = 0.55)								
19.9.3 Population is classified as having a mental disorder, but full diagnostic criteria not reported								
Freire 2007	24.6	11	7	28.2	18	9	3.2%	-0.22 [-1.21 , 0.77]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	5.2%	-0.53 [-1.30 , 0.25]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	6.0%	-0.48 [-1.21 , 0.25]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	6.6%	-0.17 [-0.87 , 0.53]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	3.3%	0.17 [-0.81 , 1.16]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	3.7%	-0.56 [-1.48 , 0.36]
Subtotal (95% CI)			**73**			**69**	**28.1%**	**-0.32 [-0.66 , 0.02]**
Heterogeneity: Tau² = 0.00; Chi² = 1.90, df = 5 (P = 0.86); I² = 0%								
Test for overall effect: Z = 1.86 (P = 0.06)								
Total (95% CI)			**263**			**240**	**100.0%**	**-0.30 [-0.48 , -0.12]**
Heterogeneity: Tau² = 0.00; Chi² = 10.66, df = 15 (P = 0.78); I² = 0%								
Test for overall effect: Z = 3.29 (P = 0.001)								
Test for subgroup differences: Chi² = 0.36, df = 2 (P = 0.84), I² = 0%								



Analysis 19.10. Comparison 19: Subgroup analyses for physical placebos compared with wait-list or no-treatment for continuous data, Outcome 10: Imputed data

Study or Subgroup	Physical placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
19.10.1 Available data								
Freire 2007	24.6	11	7	28.2	18	9	3.2%	-0.22 [-1.21 , 0.77]
Lick 1977	65.86	25.45	10	62.84	31.78	10	4.1%	0.10 [-0.78 , 0.98]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	6.6%	-0.17 [-0.87 , 0.53]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	3.3%	0.17 [-0.81 , 1.16]
Pillman 2001	-85.3	14	7	-87.6	20	7	2.9%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	6.2%	-0.07 [-0.79 , 0.65]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	3.7%	-0.56 [-1.48 , 0.36]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	3.6%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	6.2%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	6.3%	-0.11 [-0.82 , 0.60]
Subtotal (95% CI)			**121**			**110**	**46.2%**	**-0.13 [-0.40 , 0.13]**
Heterogeneity: Tau² = 0.00; Chi² = 2.42, df = 9 (P = 0.98); I² = 0%								
Test for overall effect: Z = 1.00 (P = 0.32)								
19.10.2 Intent-to-treat								
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	4.5%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	17.3%	-0.72 [-1.15 , -0.29]
Subtotal (95% CI)			**56**			**55**	**21.8%**	**-0.27 [-1.28 , 0.75]**
Heterogeneity: Tau² = 0.43; Chi² = 4.70, df = 1 (P = 0.03); I² = 79%								
Test for overall effect: Z = 0.51 (P = 0.61)								
19.10.3 No attrition								
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	5.2%	-0.53 [-1.30 , 0.25]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	16.5%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	6.0%	-0.48 [-1.21 , 0.25]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	4.3%	-0.14 [-1.00 , 0.72]
Subtotal (95% CI)			**86**			**75**	**32.0%**	**-0.40 [-0.71 , -0.08]**
Heterogeneity: Tau² = 0.00; Chi² = 0.50, df = 3 (P = 0.92); I² = 0%								
Test for overall effect: Z = 2.47 (P = 0.01)								
Total (95% CI)			**263**			**240**	**100.0%**	**-0.30 [-0.48 , -0.12]**
Heterogeneity: Tau² = 0.00; Chi² = 10.66, df = 15 (P = 0.78); I² = 0%								
Test for overall effect: Z = 3.29 (P = 0.001)								
Test for subgroup differences: Chi² = 1.59, df = 2 (P = 0.45), I² = 0%								



Comparison 20. Subgroup analyses for usual care compared with wait-list or no-treatment for continuous data

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
20.1 Type of active intervention	5	523	Std. Mean Difference (IV, Random, 95% CI)	-0.33 [-0.83, 0.16]
20.1.1 Psychological intervention	4	346	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.39, 0.10]
20.1.2 Pharmacological intervention	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
20.1.3 Physical intervention	1	177	Std. Mean Difference (IV, Random, 95% CI)	-1.10 [-1.42, -0.79]
20.1.4 Other or combination intervention	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable

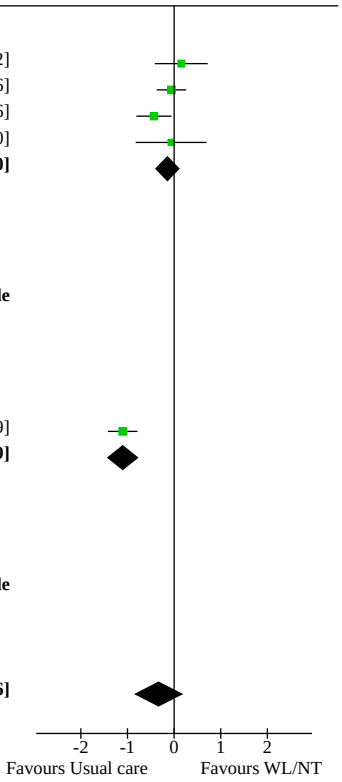
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
20.2 Type of active intervention	2	260	Risk Ratio (IV, Random, 95% CI)	0.93 [0.34, 2.55]
20.2.1 Psychological intervention	2	260	Risk Ratio (IV, Random, 95% CI)	0.93 [0.34, 2.55]
20.2.2 Pharmacological intervention	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
20.2.3 Physical intervention	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
20.2.4 Other intervention	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
20.3 Risk of bias	5	523	Std. Mean Difference (IV, Random, 95% CI)	-0.33 [-0.83, 0.16]
20.3.1 Low risk of bias	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
20.3.2 High risk of bias	5	523	Std. Mean Difference (IV, Random, 95% CI)	-0.33 [-0.83, 0.16]
20.4 Risk of bias	2	260	Risk Ratio (IV, Random, 95% CI)	0.93 [0.34, 2.55]
20.4.1 Low risk of bias	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
20.4.2 High risk of bias	2	260	Risk Ratio (IV, Random, 95% CI)	0.93 [0.34, 2.55]
20.5 Type of outcome domain	5	523	Std. Mean Difference (IV, Random, 95% CI)	-0.33 [-0.83, 0.16]
20.5.1 Blinded observer-reported	4	473	Std. Mean Difference (IV, Random, 95% CI)	-0.44 [-0.99, 0.10]
20.5.2 Non-blinded observer-reported	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
20.5.3 Patient-reported	1	50	Std. Mean Difference (IV, Random, 95% CI)	0.15 [-0.41, 0.72]
20.6 Mean age of participants	5	523	Std. Mean Difference (IV, Random, 95% CI)	-0.33 [-0.83, 0.16]
20.6.1 Below 18 years	1	177	Std. Mean Difference (IV, Random, 95% CI)	-1.10 [-1.42, -0.79]
20.6.2 18 to 50 years	3	316	Std. Mean Difference (IV, Random, 95% CI)	-0.15 [-0.46, 0.17]
20.6.3 Above 50 years	1	30	Std. Mean Difference (IV, Random, 95% CI)	-0.06 [-0.82, 0.70]
20.7 Duration of treatment	5	523	Std. Mean Difference (IV, Random, 95% CI)	-0.36 [-0.84, 0.12]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
20.7.1 Above 3 months	3	443	Std. Mean Difference (IV, Random, 95% CI)	-0.53 [-1.17, 0.10]
20.7.2 Below 3 months	2	80	Std. Mean Difference (IV, Random, 95% CI)	-0.00 [-0.46, 0.45]
20.8 Duration of treatment	2	260	Risk Ratio (IV, Random, 95% CI)	0.93 [0.34, 2.55]
20.8.1 Above 3 months	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
20.8.2 Below 3 months	2	260	Risk Ratio (IV, Random, 95% CI)	0.93 [0.34, 2.55]
20.9 Type of usual care	5	523	Std. Mean Difference (IV, Random, 95% CI)	-0.33 [-0.83, 0.16]
20.9.1 Psychological usual care	4	473	Std. Mean Difference (IV, Random, 95% CI)	-0.44 [-0.99, 0.10]
20.9.2 Pharmacological usual care	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
20.9.3 Physical usual care	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
20.9.4 Other or a combination usual care	1	50	Std. Mean Difference (IV, Random, 95% CI)	0.15 [-0.41, 0.72]
20.10 Type of usual care	2	260	Risk Ratio (IV, Random, 95% CI)	0.93 [0.34, 2.55]
20.10.1 Psychological usual care	1	26	Risk Ratio (IV, Random, 95% CI)	0.50 [0.20, 1.26]
20.10.2 Pharmacological usual care	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
20.10.3 Physical usual care	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable
20.10.4 Other or a combination usual care	1	234	Risk Ratio (IV, Random, 95% CI)	1.42 [1.17, 1.73]
20.11 Standarised usual care	5	523	Std. Mean Difference (IV, Random, 95% CI)	-0.33 [-0.83, 0.16]
20.11.1 The usual care intervention was intentionally standardised or manualised	4	493	Std. Mean Difference (IV, Random, 95% CI)	-0.38 [-0.94, 0.18]
20.11.2 No standardisation or manualisation	1	30	Std. Mean Difference (IV, Random, 95% CI)	-0.06 [-0.82, 0.70]
20.12 Standarised usual care	2	260	Risk Ratio (IV, Random, 95% CI)	0.93 [0.34, 2.55]
20.12.1 The usual care intervention was intentionally standardised or manualised	0	0	Risk Ratio (IV, Random, 95% CI)	Not estimable

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
20.12.2 No standardisation or manualisation	2	260	Risk Ratio (IV, Random, 95% CI)	0.93 [0.34, 2.55]
20.13 Mode of psychological treatment in usual care	4	473	Std. Mean Difference (IV, Random, 95% CI)	-0.44 [-0.99, 0.10]
20.13.1 Individual treatment	2	185	Std. Mean Difference (IV, Random, 95% CI)	-0.06 [-0.35, 0.23]
20.13.2 Group treatment	2	288	Std. Mean Difference (IV, Random, 95% CI)	-0.78 [-1.43, -0.12]
20.13.3 Combination of individual and group	0	0	Std. Mean Difference (IV, Random, 95% CI)	Not estimable
20.14 Imputed data	5	523	Std. Mean Difference (IV, Random, 95% CI)	-0.36 [-0.84, 0.12]
20.14.1 Available data	2	80	Std. Mean Difference (IV, Random, 95% CI)	-0.00 [-0.46, 0.45]
20.14.2 Intent-to-treat	3	443	Std. Mean Difference (IV, Random, 95% CI)	-0.53 [-1.17, 0.10]

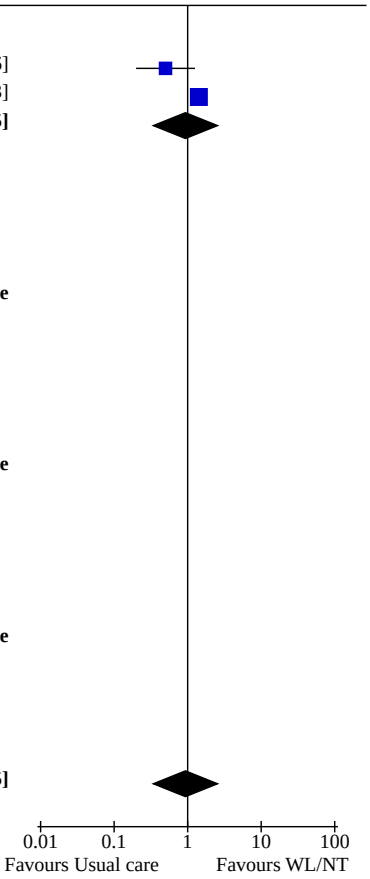
Analysis 20.1. Comparison 20: Subgroup analyses for usual care compared with wait-list or no-treatment for continuous data, Outcome 1: Type of active intervention

Study or Subgroup	Usual care Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
20.1.1 Psychological intervention								
Crisp 1991	-5.5	0.6	30	-5.6	0.7	20	18.5%	0.15 [-0.41 , 0.72]
Glogowska 2000	-4.9	1.7	71	-4.8	1.8	84	22.3%	-0.06 [-0.37 , 0.26]
Rapee 2007	28.63	15.01	59	35.337	15.83	52	21.4%	-0.43 [-0.81 , -0.06]
Teri 1997	13.9	4.14	10	14.1	2.32	20	15.5%	-0.06 [-0.82 , 0.70]
Subtotal (95% CI)			**170**			**176**	**77.7%**	**-0.14 [-0.39 , 0.10]**
Heterogeneity: Tau² = 0.01; Chi² = 3.64, df = 3 (P = 0.30); I² = 18%								
Test for overall effect: Z = 1.15 (P = 0.25)								
20.1.2 Pharmacological intervention								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
20.1.3 Physical intervention								
Rapee 2006	3.4	2.6	90	5.8	1.6	87	22.3%	-1.10 [-1.42 , -0.79]
Subtotal (95% CI)			**90**			**87**	**22.3%**	**-1.10 [-1.42 , -0.79]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 6.82 (P < 0.00001)								
20.1.4 Other or combination intervention								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
Total (95% CI)			**260**			**263**	**100.0%**	**-0.33 [-0.83 , 0.16]**
Heterogeneity: Tau² = 0.26; Chi² = 27.62, df = 4 (P < 0.0001); I² = 86%								
Test for overall effect: Z = 1.32 (P = 0.19)								
Test for subgroup differences: Chi² = 22.02, df = 1 (P < 0.00001), I² = 95.5%								



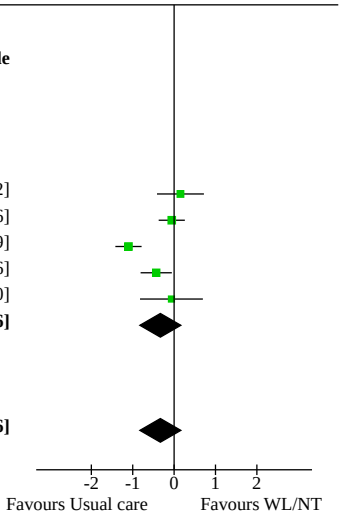
Analysis 20.2. Comparison 20: Subgroup analyses for usual care compared with wait-list or no-treatment for continuous data, Outcome 2: Type of active intervention

Study or Subgroup	Usual care Events	Usual care Total	Wait-list/no-treatment Events	Wait-list/no-treatment Total	Weight	Risk Ratio IV, Random, 95% CI
20.2.1 Psychological intervention						
Milby 1980	4	13	8	13	40.3%	0.50 [0.20 , 1.26]
Robin 1976	88	116	63	118	59.7%	1.42 [1.17 , 1.73]
Subtotal (95% CI)		**129**		**131**	**100.0%**	**0.93 [0.34 , 2.55]**
Total events:	92		71			
Heterogeneity: Tau² = 0.43; Chi² = 4.72, df = 1 (P = 0.03); I² = 79%						
Test for overall effect: Z = 0.14 (P = 0.89)						
20.2.2 Pharmacological intervention						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
20.2.3 Physical intervention						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
20.2.4 Other intervention						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
Total (95% CI)		**129**		**131**	**100.0%**	**0.93 [0.34 , 2.55]**
Total events:	92		71			
Heterogeneity: Tau² = 0.43; Chi² = 4.72, df = 1 (P = 0.03); I² = 79%						
Test for overall effect: Z = 0.14 (P = 0.89)						
Test for subgroup differences: Not applicable						



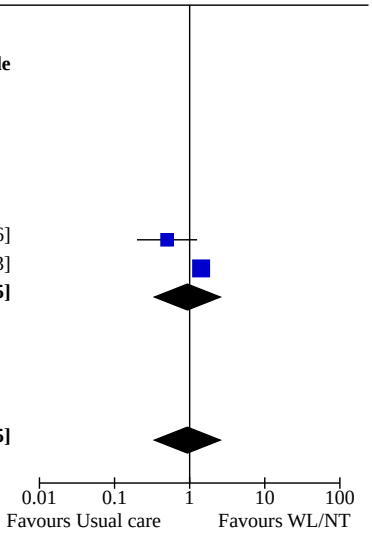
Analysis 20.3. Comparison 20: Subgroup analyses for usual care compared with wait-list or no-treatment for continuous data, Outcome 3: Risk of bias

Study or Subgroup	Usual care Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
20.3.1 Low risk of bias								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
20.3.2 High risk of bias								
Crisp 1991	-5.5	0.6	30	-5.6	0.7	20	18.5%	0.15 [-0.41 , 0.72]
Glogowska 2000	-4.9	1.7	71	-4.8	1.8	84	22.3%	-0.06 [-0.37 , 0.26]
Rapee 2006	3.4	2.6	90	5.8	1.6	87	22.3%	-1.10 [-1.42 , -0.79]
Rapee 2007	28.63	15.01	59	35.337	15.83	52	21.4%	-0.43 [-0.81 , -0.06]
Teri 1997	13.9	4.14	10	14.1	2.32	20	15.5%	-0.06 [-0.82 , 0.70]
Subtotal (95% CI)			**260**			**263**	**100.0%**	**-0.33 [-0.83 , 0.16]**
Heterogeneity: Tau² = 0.26; Chi² = 27.62, df = 4 (P < 0.0001); I² = 86%								
Test for overall effect: Z = 1.32 (P = 0.19)								
Total (95% CI)			**260**			**263**	**100.0%**	**-0.33 [-0.83 , 0.16]**
Heterogeneity: Tau² = 0.26; Chi² = 27.62, df = 4 (P < 0.0001); I² = 86%								
Test for overall effect: Z = 1.32 (P = 0.19)								
Test for subgroup differences: Not applicable								



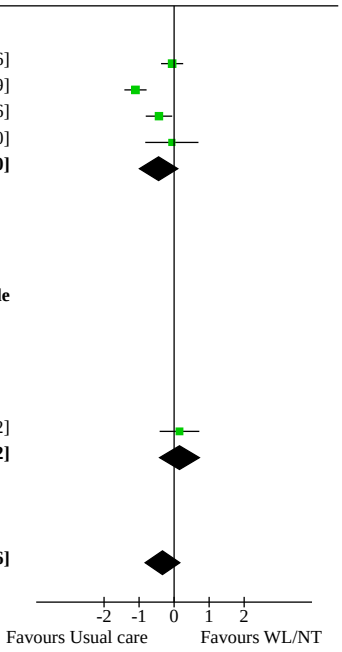
Analysis 20.4. Comparison 20: Subgroup analyses for usual care compared with wait-list or no-treatment for continuous data, Outcome 4: Risk of bias

Study or Subgroup	Usual care Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio IV, Random, 95% CI
20.4.1 Low risk of bias						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
20.4.2 High risk of bias						
Milby 1980	4	13	8	13	40.3%	0.50 [0.20 , 1.26]
Robin 1976	88	116	63	118	59.7%	1.42 [1.17 , 1.73]
Subtotal (95% CI)		**129**		**131**	**100.0%**	**0.93 [0.34 , 2.55]**
Total events:	92		71			
Heterogeneity: Tau² = 0.43; Chi² = 4.72, df = 1 (P = 0.03); I² = 79%						
Test for overall effect: Z = 0.14 (P = 0.89)						
Total (95% CI)		**129**		**131**	**100.0%**	**0.93 [0.34 , 2.55]**
Total events:	92		71			
Heterogeneity: Tau² = 0.43; Chi² = 4.72, df = 1 (P = 0.03); I² = 79%						
Test for overall effect: Z = 0.14 (P = 0.89)						
Test for subgroup differences: Not applicable						



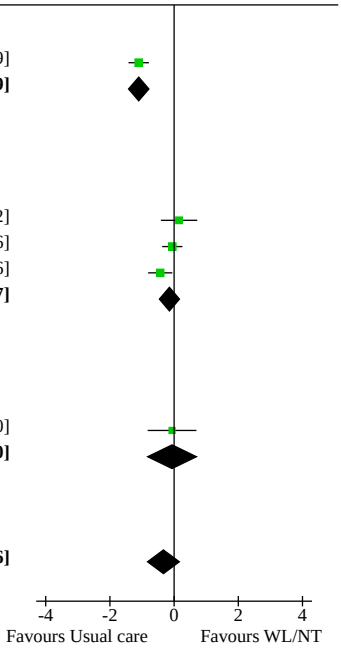
Analysis 20.5. Comparison 20: Subgroup analyses for usual care compared with wait-list or no-treatment for continuous data, Outcome 5: Type of outcome domain

Study or Subgroup	Usual care Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
20.5.1 Blinded observer-reported								
Glogowska 2000	-4.9	1.7	71	-4.8	1.8	84	22.3%	-0.06 [-0.37 , 0.26]
Rapee 2006	3.4	2.6	90	5.8	1.6	87	22.3%	-1.10 [-1.42 , -0.79]
Rapee 2007	28.63	15.01	59	35.337	15.83	52	21.4%	-0.43 [-0.81 , -0.06]
Teri 1997	13.9	4.14	10	14.1	2.32	20	15.5%	-0.06 [-0.82 , 0.70]
Subtotal (95% CI)			**230**			**243**	**81.5%**	**-0.44 [-0.99 , 0.10]**
Heterogeneity: Tau² = 0.26; Chi² = 22.84, df = 3 (P < 0.0001); I² = 87%								
Test for overall effect: Z = 1.59 (P = 0.11)								
20.5.2 Non-blinded observer-reported								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
20.5.3 Patient-reported								
Crisp 1991	-5.5	0.6	30	-5.6	0.7	20	18.5%	0.15 [-0.41 , 0.72]
Subtotal (95% CI)			**30**			**20**	**18.5%**	**0.15 [-0.41 , 0.72]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.53 (P = 0.60)								
Total (95% CI)			**260**			**263**	**100.0%**	**-0.33 [-0.83 , 0.16]**
Heterogeneity: Tau² = 0.26; Chi² = 27.62, df = 4 (P < 0.0001); I² = 86%								
Test for overall effect: Z = 1.32 (P = 0.19)								
Test for subgroup differences: Chi² = 2.20, df = 1 (P = 0.14), I² = 54.6%								



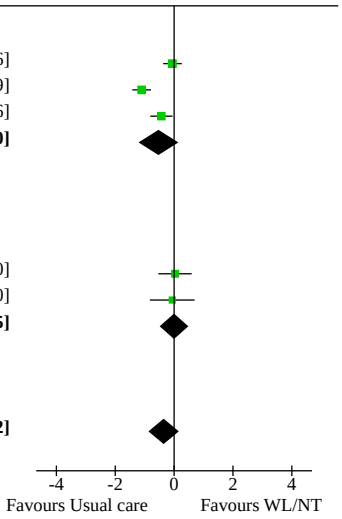
Analysis 20.6. Comparison 20: Subgroup analyses for usual care compared with wait-list or no-treatment for continuous data, Outcome 6: Mean age of participants

Study or Subgroup	Usual care Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
20.6.1 Below 18 years								
Rapee 2006	3.4	2.6	90	5.8	1.6	87	22.3%	-1.10 [-1.42 , -0.79]
Subtotal (95% CI)			**90**			**87**	**22.3%**	**-1.10 [-1.42 , -0.79]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 6.82 (P < 0.00001)								
20.6.2 18 to 50 years								
Crisp 1991	-5.5	0.6	30	-5.6	0.7	20	18.5%	0.15 [-0.41 , 0.72]
Glogowska 2000	-4.9	1.7	71	-4.8	1.8	84	22.3%	-0.06 [-0.37 , 0.26]
Rapee 2007	28.63	15.01	59	35.337	15.83	52	21.4%	-0.43 [-0.81 , -0.06]
Subtotal (95% CI)			**160**			**156**	**62.2%**	**-0.15 [-0.46 , 0.17]**
Heterogeneity: Tau² = 0.03; Chi² = 3.59, df = 2 (P = 0.17); I² = 44%								
Test for overall effect: Z = 0.92 (P = 0.36)								
20.6.3 Above 50 years								
Teri 1997	13.9	4.14	10	14.1	2.32	20	15.5%	-0.06 [-0.82 , 0.70]
Subtotal (95% CI)			**10**			**20**	**15.5%**	**-0.06 [-0.82 , 0.70]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.17 (P = 0.87)								
Total (95% CI)			**260**			**263**	**100.0%**	**-0.33 [-0.83 , 0.16]**
Heterogeneity: Tau² = 0.26; Chi² = 27.62, df = 4 (P < 0.0001); I² = 86%								
Test for overall effect: Z = 1.32 (P = 0.19)								
Test for subgroup differences: Chi² = 19.65, df = 2 (P < 0.0001), I² = 89.8%								



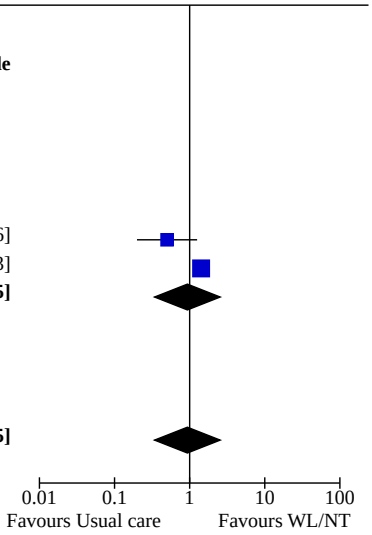
Analysis 20.7. Comparison 20: Subgroup analyses for usual care compared with wait-list or no-treatment for continuous data, Outcome 7: Duration of treatment

Study or Subgroup	Usual care Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
20.7.1 Above 3 months								
Glogowska 2000	-4.9	1.7	71	-4.8	1.8	84	22.4%	-0.06 [-0.37 , 0.26]
Rapee 2006	3.4	2.6	90	5.8	1.6	87	22.4%	-1.10 [-1.42 , -0.79]
Rapee 2007	28.63	15.01	59	35.337	15.83	52	21.5%	-0.43 [-0.81 , -0.06]
Subtotal (95% CI)			**220**			**223**	**66.3%**	**-0.53 [-1.17 , 0.10]**
Heterogeneity: Tau² = 0.28; Chi² = 21.42, df = 2 (P < 0.0001); I² = 91%								
Test for overall effect: Z = 1.64 (P = 0.10)								
20.7.2 Below 3 months								
Crisp 1991	-5.5	3.2863	30	-5.6	3.1305	20	18.4%	0.03 [-0.54 , 0.60]
Teri 1997	13.9	4.14	10	14.1	2.32	20	15.3%	-0.06 [-0.82 , 0.70]
Subtotal (95% CI)			**40**			**40**	**33.7%**	**-0.00 [-0.46 , 0.45]**
Heterogeneity: Tau² = 0.00; Chi² = 0.04, df = 1 (P = 0.84); I² = 0%								
Test for overall effect: Z = 0.01 (P = 0.99)								
Total (95% CI)			**260**			**263**	**100.0%**	**-0.36 [-0.84 , 0.12]**
Heterogeneity: Tau² = 0.24; Chi² = 26.02, df = 4 (P < 0.0001); I² = 85%								
Test for overall effect: Z = 1.45 (P = 0.15)								
Test for subgroup differences: Chi² = 1.76, df = 1 (P = 0.18), I² = 43.3%								



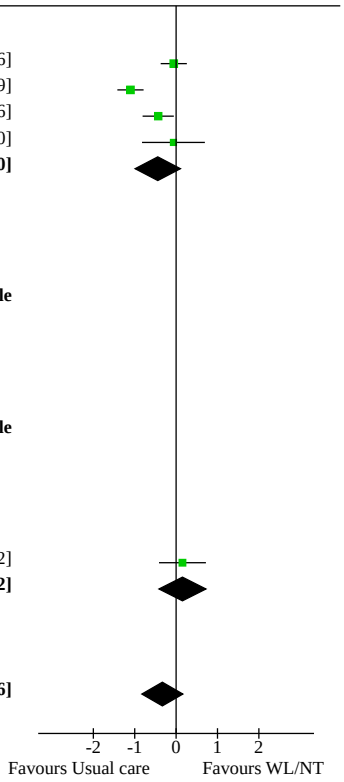
Analysis 20.8. Comparison 20: Subgroup analyses for usual care compared with wait-list or no-treatment for continuous data, Outcome 8: Duration of treatment

Study or Subgroup	Usual care Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio IV, Random, 95% CI
20.8.1 Above 3 months						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
20.8.2 Below 3 months						
Milby 1980	4	13	8	13	40.3%	0.50 [0.20 , 1.26]
Robin 1976	88	116	63	118	59.7%	1.42 [1.17 , 1.73]
Subtotal (95% CI)		**129**		**131**	**100.0%**	**0.93 [0.34 , 2.55]**
Total events:	92		71			
Heterogeneity: Tau² = 0.43; Chi² = 4.72, df = 1 (P = 0.03); I² = 79%						
Test for overall effect: Z = 0.14 (P = 0.89)						
Total (95% CI)		**129**		**131**	**100.0%**	**0.93 [0.34 , 2.55]**
Total events:	92		71			
Heterogeneity: Tau² = 0.43; Chi² = 4.72, df = 1 (P = 0.03); I² = 79%						
Test for overall effect: Z = 0.14 (P = 0.89)						
Test for subgroup differences: Not applicable						



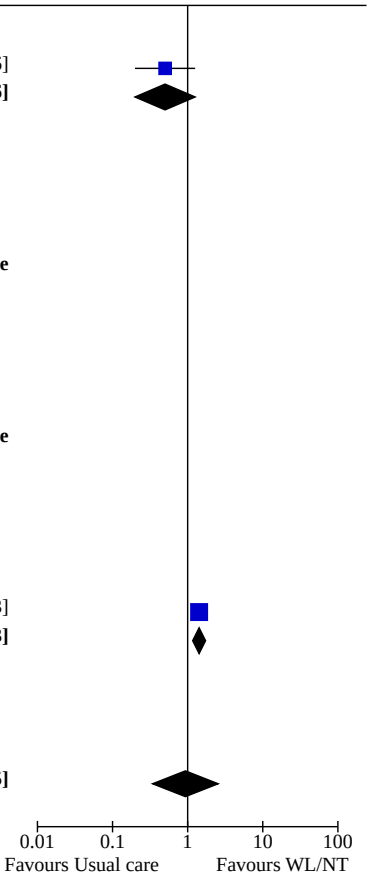
Analysis 20.9. Comparison 20: Subgroup analyses for usual care compared with wait-list or no-treatment for continuous data, Outcome 9: Type of usual care

Study or Subgroup	Usual care Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
20.9.1 Psychological usual care								
Glogowska 2000	-4.9	1.7	71	-4.8	1.8	84	22.3%	-0.06 [-0.37 , 0.26]
Rapee 2006	3.4	2.6	90	5.8	1.6	87	22.3%	-1.10 [-1.42 , -0.79]
Rapee 2007	28.63	15.01	59	35.337	15.83	52	21.4%	-0.43 [-0.81 , -0.06]
Teri 1997	13.9	4.14	10	14.1	2.32	20	15.5%	-0.06 [-0.82 , 0.70]
Subtotal (95% CI)			**230**			**243**	**81.5%**	**-0.44 [-0.99 , 0.10]**
Heterogeneity: Tau² = 0.26; Chi² = 22.84, df = 3 (P < 0.0001); I² = 87%								
Test for overall effect: Z = 1.59 (P = 0.11)								
20.9.2 Pharmacological usual care								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
20.9.3 Physical usual care								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
20.9.4 Other or a combination usual care								
Crisp 1991	-5.5	0.6	30	-5.6	0.7	20	18.5%	0.15 [-0.41 , 0.72]
Subtotal (95% CI)			**30**			**20**	**18.5%**	**0.15 [-0.41 , 0.72]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.53 (P = 0.60)								
Total (95% CI)			**260**			**263**	**100.0%**	**-0.33 [-0.83 , 0.16]**
Heterogeneity: Tau² = 0.26; Chi² = 27.62, df = 4 (P < 0.0001); I² = 86%								
Test for overall effect: Z = 1.32 (P = 0.19)								
Test for subgroup differences: Chi² = 2.20, df = 1 (P = 0.14), I² = 54.6%								



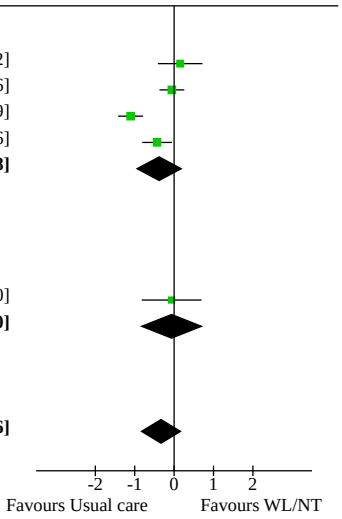
Analysis 20.10. Comparison 20: Subgroup analyses for usual care compared with wait-list or no-treatment for continuous data, Outcome 10: Type of usual care

Study or Subgroup	Usual care Events	Usual care Total	Wait-list/no-treatment Events	Wait-list/no-treatment Total	Weight	Risk Ratio IV, Random, 95% CI
20.10.1 Psychological usual care						
Milby 1980	4	13	8	13	40.3%	0.50 [0.20 , 1.26]
Subtotal (95% CI)		**13**		**13**	**40.3%**	**0.50 [0.20 , 1.26]**
Total events:	4		8			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.47 (P = 0.14)						
20.10.2 Pharmacological usual care						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
20.10.3 Physical usual care						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
20.10.4 Other or a combination usual care						
Robin 1976	88	116	63	118	59.7%	1.42 [1.17 , 1.73]
Subtotal (95% CI)		**116**		**118**	**59.7%**	**1.42 [1.17 , 1.73]**
Total events:	88		63			
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.49 (P = 0.0005)						
Total (95% CI)		**129**		**131**	**100.0%**	**0.93 [0.34 , 2.55]**
Total events:	92		71			
Heterogeneity: Tau² = 0.43; Chi² = 4.72, df = 1 (P = 0.03); I² = 79%						
Test for overall effect: Z = 0.14 (P = 0.89)						
Test for subgroup differences: Chi² = 4.72, df = 1 (P = 0.03), I² = 78.8%						



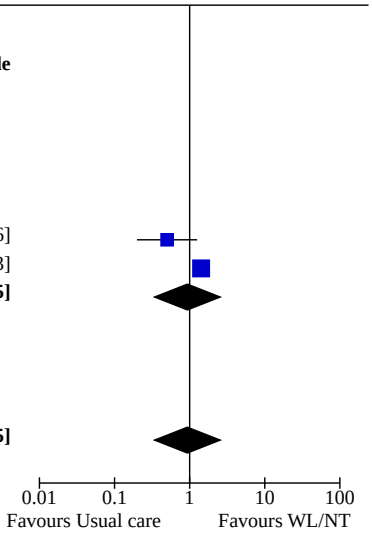
Analysis 20.11. Comparison 20: Subgroup analyses for usual care compared with wait-list or no-treatment for continuous data, Outcome 11: Standarised usual care

Study or Subgroup	Usual care Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
20.11.1 The usual care intervention was intentionally standardised or manualised								
Crisp 1991	-5.5	0.6	30	-5.6	0.7	20	18.5%	0.15 [-0.41 , 0.72]
Glogowska 2000	-4.9	1.7	71	-4.8	1.8	84	22.3%	-0.06 [-0.37 , 0.26]
Rapee 2006	3.4	2.6	90	5.8	1.6	87	22.3%	-1.10 [-1.42 , -0.79]
Rapee 2007	28.63	15.01	59	35.337	15.83	52	21.4%	-0.43 [-0.81 , -0.06]
Subtotal (95% CI)			**250**			**243**	**84.5%**	**-0.38 [-0.94 , 0.18]**
Heterogeneity: Tau² = 0.29; Chi² = 26.59, df = 3 (P < 0.00001); I² = 89%								
Test for overall effect: Z = 1.33 (P = 0.18)								
20.11.2 No standardisation or manualisation								
Teri 1997	13.9	4.14	10	14.1	2.32	20	15.5%	-0.06 [-0.82 , 0.70]
Subtotal (95% CI)			**10**			**20**	**15.5%**	**-0.06 [-0.82 , 0.70]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.17 (P = 0.87)								
Total (95% CI)			**260**			**263**	**100.0%**	**-0.33 [-0.83 , 0.16]**
Heterogeneity: Tau² = 0.26; Chi² = 27.62, df = 4 (P < 0.0001); I² = 86%								
Test for overall effect: Z = 1.32 (P = 0.19)								
Test for subgroup differences: Chi² = 0.43, df = 1 (P = 0.51), I² = 0%								



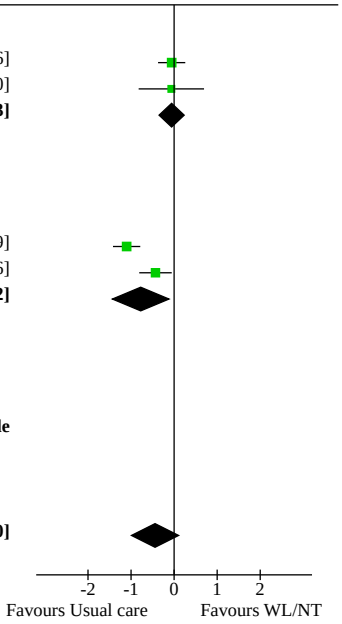
Analysis 20.12. Comparison 20: Subgroup analyses for usual care compared with wait-list or no-treatment for continuous data, Outcome 12: Standarised usual care

Study or Subgroup	Usual care Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio IV, Random, 95% CI
20.12.1 The usual care intervention was intentionally standardised or manualised						
Subtotal (95% CI)		**0**		**0**		**Not estimable**
Total events:	0		0			
Heterogeneity: Not applicable						
Test for overall effect: Not applicable						
20.12.2 No standardisation or manualisation						
Milby 1980	4	13	8	13	40.3%	0.50 [0.20 , 1.26]
Robin 1976	88	116	63	118	59.7%	1.42 [1.17 , 1.73]
Subtotal (95% CI)		**129**		**131**	**100.0%**	**0.93 [0.34 , 2.55]**
Total events:	92		71			
Heterogeneity: Tau² = 0.43; Chi² = 4.72, df = 1 (P = 0.03); I² = 79%						
Test for overall effect: Z = 0.14 (P = 0.89)						
Total (95% CI)		**129**		**131**	**100.0%**	**0.93 [0.34 , 2.55]**
Total events:	92		71			
Heterogeneity: Tau² = 0.43; Chi² = 4.72, df = 1 (P = 0.03); I² = 79%						
Test for overall effect: Z = 0.14 (P = 0.89)						
Test for subgroup differences: Not applicable						



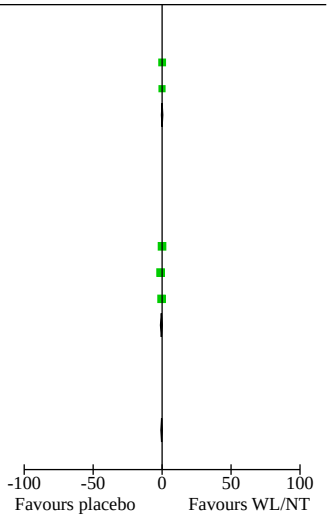
Analysis 20.13. Comparison 20: Subgroup analyses for usual care compared with wait-list or no-treatment for continuous data, Outcome 13: Mode of psychological treatment in usual care

Study or Subgroup	Usual care Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
20.13.1 Individual treatment								
Glogowska 2000	-4.9	1.7	71	-4.8	1.8	84	27.3%	-0.06 [-0.37 , 0.26]
Teri 1997	13.9	4.14	10	14.1	2.32	20	19.0%	-0.06 [-0.82 , 0.70]
Subtotal (95% CI)			**81**			**104**	**46.4%**	**-0.06 [-0.35 , 0.23]**
Heterogeneity: Tau² = 0.00; Chi² = 0.00, df = 1 (P = 0.99); I² = 0%								
Test for overall effect: Z = 0.39 (P = 0.70)								
20.13.2 Group treatment								
Rapee 2006	3.4	2.6	90	5.8	1.6	87	27.3%	-1.10 [-1.42 , -0.79]
Rapee 2007	28.63	15.01	59	35.337	15.83	52	26.3%	-0.43 [-0.81 , -0.06]
Subtotal (95% CI)			**149**			**139**	**53.6%**	**-0.78 [-1.43 , -0.12]**
Heterogeneity: Tau² = 0.19; Chi² = 7.11, df = 1 (P = 0.008); I² = 86%								
Test for overall effect: Z = 2.32 (P = 0.02)								
20.13.3 Combination of individual and group								
Subtotal (95% CI)			**0**			**0**		**Not estimable**
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
Total (95% CI)			**230**			**243**	**100.0%**	**-0.44 [-0.99 , 0.10]**
Heterogeneity: Tau² = 0.26; Chi² = 22.84, df = 3 (P < 0.0001); I² = 87%								
Test for overall effect: Z = 1.59 (P = 0.11)								
Test for subgroup differences: Chi² = 3.84, df = 1 (P = 0.05), I² = 73.9%								



Analysis 20.14. Comparison 20: Subgroup analyses for usual care compared with wait-list or no-treatment for continuous data, Outcome 14: Imputed data

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
20.14.1 Available data								
Crisp 1991	-5.5	3.2863	30	-5.6	3.1305	20	18.4%	0.03 [-0.54 , 0.60]
Teri 1997	13.9	4.14	10	14.1	2.32	20	15.3%	-0.06 [-0.82 , 0.70]
Subtotal (95% CI)			**40**			**40**	**33.7%**	**-0.00 [-0.46 , 0.45]**
Heterogeneity: Tau² = 0.00; Chi² = 0.04, df = 1 (P = 0.84); I² = 0%								
Test for overall effect: Z = 0.01 (P = 0.99)								
20.14.2 Intent-to-treat								
Glogowska 2000	-4.9	1.7	71	-4.8	1.8	84	22.4%	-0.06 [-0.37 , 0.26]
Rapee 2006	3.4	2.6	90	5.8	1.6	87	22.4%	-1.10 [-1.42 , -0.79]
Rapee 2007	28.63	15.01	59	35.337	15.83	52	21.5%	-0.43 [-0.81 , -0.06]
Subtotal (95% CI)			**220**			**223**	**66.3%**	**-0.53 [-1.17 , 0.10]**
Heterogeneity: Tau² = 0.28; Chi² = 21.42, df = 2 (P < 0.0001); I² = 91%								
Test for overall effect: Z = 1.64 (P = 0.10)								
Total (95% CI)			**260**			**263**	**100.0%**	**-0.36 [-0.84 , 0.12]**
Heterogeneity: Tau² = 0.24; Chi² = 26.02, df = 4 (P < 0.0001); I² = 85%								
Test for overall effect: Z = 1.45 (P = 0.15)								
Test for subgroup differences: Chi² = 1.76, df = 1 (P = 0.18), I² = 43.3%								



Comparison 21. Sensitivity analyses for all placebos compared with wait-list or no-treatment

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
21.1 Fixed effects	65	2446	Std. Mean Difference (IV, Fixed, 95% CI)	-0.33 [-0.42, -0.25]
21.2 Fixed effects	9	385	Risk Ratio (IV, Fixed, 95% CI)	0.91 [0.77, 1.07]
21.3 End of intervention (post-treatment scores, changes scores excluded)	65	2446	Std. Mean Difference (IV, Random, 95% CI)	-0.37 [-0.49, -0.25]
21.4 Type of data collection	14		Risk Difference (IV, Random, 95% CI)	0.03 [-0.02, 0.08]
21.4.1 Structured list of reporting	1		Risk Difference (IV, Random, 95% CI)	-1.54 [-2.97, -0.11]
21.4.2 Spontaneous reporting	13		Risk Difference (IV, Random, 95% CI)	0.03 [-0.02, 0.07]
21.5 Transformed data for TSA analysis	65	2456	Std. Mean Difference (IV, Random, 95% CI)	-0.34 [-0.44, -0.23]

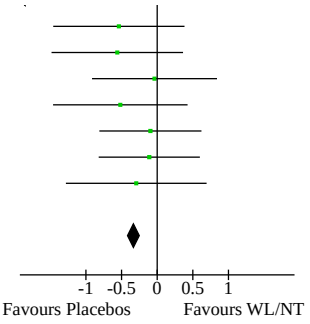
Analysis 21.1. Comparison 21: Sensitivity analyses for all placebos compared with wait-list or no-treatment, Outcome 1: Fixed effects

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	3.0%	-0.45 [-0.93 , 0.03]
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	1.0%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	3.7%	-0.72 [-1.15 , -0.29]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	2.1%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	0.7%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	0.9%	-1.71 [-2.60 , -0.82]
Borden 1986	29.7	17.5	10	32	26	10	0.9%	-0.10 [-0.98 , 0.78]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	0.9%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	1.0%	-0.25 [-1.09 , 0.59]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	1.8%	-0.11 [-0.72 , 0.49]
Bramston 1985	-78	14.21	12	-72	13.82	12	1.0%	-0.41 [-1.22 , 0.40]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	0.8%	-0.09 [-1.01 , 0.84]
Carter 2003	3.8	1.2	28	3.8	1.3	29	2.5%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	2.5%	-0.64 [-1.16 , -0.12]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	1.1%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	1.1%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	0.9%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	4.6%	-1.10 [-1.48 , -0.71]
Freire 2007	24.6	11	7	28.2	18	9	0.7%	-0.22 [-1.21 , 0.77]
Fuchs 1977	14.3	7	10	21.4	7	10	0.8%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	1.1%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	3.5%	-0.08 [-0.51 , 0.36]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	1.1%	-0.53 [-1.30 , 0.25]
Kelley 2012	-1.64	4.52	11	0.67	4	9	0.8%	-0.51 [-1.41 , 0.38]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	0.8%	-0.59 [-1.49 , 0.31]
Kilmann 1987	69	27	4	90	14	4	0.3%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	0.5%	-0.69 [-1.87 , 0.49]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	1.0%	-0.81 [-1.63 , 0.00]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	0.9%	-0.35 [-1.24 , 0.53]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	3.5%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	1.3%	-0.48 [-1.21 , 0.25]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	2.8%	0.03 [-0.46 , 0.52]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	0.9%	-0.14 [-1.00 , 0.72]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	0.7%	-0.71 [-1.67 , 0.25]
Lick 1977	65.86	25.45	10	62.84	31.78	10	0.9%	0.10 [-0.78 , 0.98]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	0.9%	0.34 [-0.55 , 1.22]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	0.8%	-0.04 [-0.93 , 0.86]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	0.9%	-0.45 [-1.32 , 0.42]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	1.4%	-0.17 [-0.87 , 0.53]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	0.8%	-0.33 [-1.27 , 0.61]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	0.7%	0.17 [-0.81 , 1.16]
Peck 1976	-12.5	4.25	4	-7	3.27	4	0.3%	-1.26 [-2.90 , 0.37]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	1.1%	-0.12 [-0.89 , 0.65]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	1.0%	-1.33 [-2.16 , -0.49]
Pillman 2001	-85.3	14	7	-87.6	20	7	0.6%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	1.3%	-0.07 [-0.79 , 0.65]
Powers 2004	46	20.09	12	41.6	13.15	15	1.2%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	1.1%	-0.10 [-0.90 , 0.70]
Powers 2008b	64	20	18	51	37	15	1.4%	0.44 [-0.26 , 1.13]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	17.3%	-0.11 [-0.31 , 0.09]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	2.3%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	1.3%	-0.43 [-1.15 , 0.28]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	0.8%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	0.7%	-0.39 [-1.37 , 0.58]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	2.9%	-0.08 [-0.56 , 0.40]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	1.2%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	0.9%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	0.5%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	0.8%	-0.54 [-1.46 , 0.38]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	0.8%	-0.56 [-1.48 , 0.36]

Analysis 21.1. (Continued)

Tan 1986	9.5	7	10	13.44	7	9	0.8%	-0.54 [-1.46 , 0.38]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	0.8%	-0.56 [-1.48 , 0.36]
Turner 1979	-4.1	2.644	10	-4	2.449	10	0.9%	-0.04 [-0.91 , 0.84]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	0.8%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	1.3%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	1.3%	-0.11 [-0.82 , 0.60]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	0.7%	-0.29 [-1.28 , 0.69]
Total (95% CI)			**1246**			**1200**	**100.0%**	**-0.33 [-0.42 , -0.25]**

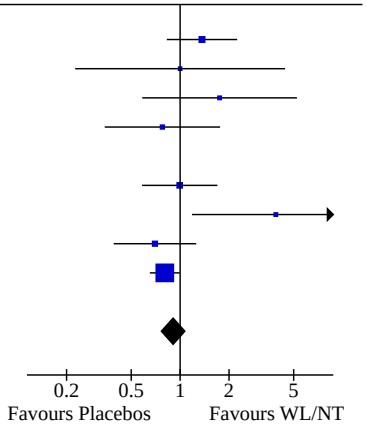
Heterogeneity: Chi² = 108.28, df = 64 (P = 0.0005); I² = 41%
Test for overall effect: Z = 7.95 (P < 0.00001)
Test for subgroup differences: Not applicable



Analysis 21.2. Comparison 21: Sensitivity analyses for all placebos compared with wait-list or no-treatment, Outcome 2: Fixed effects

	All placebos		Wait-list/no-treatment			Risk Ratio	Risk Ratio
Study or Subgroup	**Events**	**Total**	**Events**	**Total**	**Weight**	**IV, Fixed, 95% CI**	**IV, Fixed, 95% CI**
Berg 1983	9	11	9	15	11.1%	1.36 [0.83 , 2.24]	
Double 1993	3	22	3	22	1.2%	1.00 [0.23 , 4.42]	
Klerman 1974a	7	25	4	25	2.3%	1.75 [0.58 , 5.24]	
Klerman 1974b	7	25	9	25	4.1%	0.78 [0.34 , 1.76]	
Mealiea 1971	0	9	0	10		Not estimable	
Rabkin 1990	14	27	12	23	9.7%	0.99 [0.58 , 1.70]	
Watzl 1988	11	34	3	36	2.0%	3.88 [1.18 , 12.73]	
Whittaker 1963	7	13	10	13	8.1%	0.70 [0.39 , 1.26]	
Wilson 1980	31	40	10	10	61.5%	0.80 [0.65 , 0.99]	
Total (95% CI)		**206**		**179**	**100.0%**	**0.91 [0.77 , 1.07]**	
Total events:	89		60				

Heterogeneity: Chi² = 11.95, df = 7 (P = 0.10); I² = 41%
Test for overall effect: Z = 1.17 (P = 0.24)
Test for subgroup differences: Not applicable



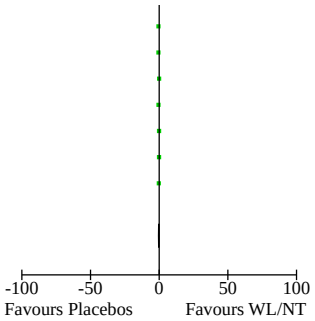
Analysis 21.3. Comparison 21: Sensitivity analyses for all placebos compared with wait-list or no-treatment, Outcome 3: End of intervention (post-treatment scores, changes scores excluded)

	All placebos			Wait-list/no-treatment				Std. Mean Difference	Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI	IV, Random, 95% CI
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	2.5%	-0.45 [-0.93 , 0.03]	
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	1.3%	0.32 [-0.52 , 1.17]	
Allen 2006	12	9.6	45	19	9.6	44	2.8%	-0.72 [-1.15 , -0.29]	
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	2.2%	-0.09 [-0.66 , 0.48]	
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	1.1%	-0.10 [-1.05 , 0.85]	
Ayen 2004	14.6	4.5	20	22.1	3.7	10	1.2%	-1.71 [-2.60 , -0.82]	
Borden 1986	29.7	17.5	10	32	26	10	1.3%	-0.10 [-0.98 , 0.78]	
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	1.2%	-0.71 [-1.60 , 0.18]	
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	1.3%	-0.25 [-1.09 , 0.59]	
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	2.0%	-0.11 [-0.72 , 0.49]	
Bramston 1985	-78	14.21	12	-72	13.82	12	1.4%	-0.41 [-1.22 , 0.40]	
Carlson 1993	5.2	3.3	9	5.5	3.3	9	1.2%	-0.09 [-1.01 , 0.84]	
Carter 2003	3.8	1.2	28	3.8	1.3	29	2.3%	0.00 [-0.52 , 0.52]	
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	2.3%	-0.64 [-1.16 , -0.12]	
Espie 1989a	63.5	30.6	14	96.5	63.5	13	1.5%	-0.65 [-1.43 , 0.13]	
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	1.5%	-0.33 [-1.12 , 0.46]	
Foa 1991	18.09	7.13	11	19.5	7.18	10	1.3%	-0.19 [-1.05 , 0.67]	
Foa 2018	18.65	4.28	110	23.06	3.09	40	3.0%	-1.10 [-1.48 , -0.71]	
Freire 2007	24.6	11	7	28.2	18	9	1.1%	-0.22 [-1.21 , 0.77]	
Fuchs 1977	14.3	7	10	21.4	7	10	1.1%	-0.97 [-1.91 , -0.03]	
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	1.5%	-0.80 [-1.59 , -0.01]	
Hippman 2016	61.4	15.9	40	62.7	17.8	40	2.7%	-0.08 [-0.51 , 0.36]	
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	1.5%	-0.53 [-1.30 , 0.25]	
Kelley 2012	-1.64	4.52	11	0.67	4	9	1.2%	-0.51 [-1.41 , 0.38]	
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	1.2%	-0.59 [-1.49 , 0.31]	
Kilmann 1987	69	27	4	90	14	4	0.5%	-0.85 [-2.35 , 0.65]	
Klein 1977	10.17	5.49	6	15	7.32	6	0.8%	-0.69 [-1.87 , 0.49]	
Klosko 1990	3.55	1.51	11	4.8	1.47	15	1.4%	-0.81 [-1.63 , 0.00]	
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	1.3%	-0.35 [-1.24 , 0.53]	
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	2.7%	-0.39 [-0.83 , 0.05]	
Lacy 1990	5.71	1.95	15	6.79	2.42	15	1.6%	-0.48 [-1.21 , 0.25]	
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	2.5%	0.03 [-0.46 , 0.52]	
Legrand 2016	28.43	7.46	11	29.9	12.57	10	1.3%	-0.14 [-1.00 , 0.72]	
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	1.1%	-0.71 [-1.67 , 0.25]	
Lick 1977	65.86	25.45	10	62.84	31.78	10	1.3%	0.10 [-0.78 , 0.98]	
Liddle 1990	19.3	6.93	10	16.9	6.79	10	1.3%	0.34 [-0.55 , 1.22]	
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	1.2%	-0.04 [-0.93 , 0.86]	
Miranda 1997	5.2	3.6	11	6.9	3.7	10	1.3%	-0.45 [-1.32 , 0.42]	
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	1.7%	-0.17 [-0.87 , 0.53]	
Nandi 1976	53.2	11.2	10	57.5	13.8	8	1.1%	-0.33 [-1.27 , 0.61]	
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	1.1%	0.17 [-0.81 , 1.16]	
Peck 1976	-12.5	4.25	4	-7	3.27	4	0.5%	-1.26 [-2.90 , 0.37]	
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	1.5%	-0.12 [-0.89 , 0.65]	
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	1.4%	-1.33 [-2.16 , -0.49]	
Pillman 2001	-85.3	14	7	-87.6	20	7	1.0%	0.12 [-0.92 , 1.17]	
Poland 2013	22	8.2	16	22.6	8	14	1.7%	-0.07 [-0.79 , 0.65]	
Powers 2004	46	20.09	12	41.6	13.15	15	1.5%	0.26 [-0.51 , 1.02]	
Powers 2008a	66	36	15	69	16	10	1.4%	-0.10 [-0.90 , 0.70]	
Powers 2008b	64	20	18	51	37	15	1.7%	0.44 [-0.26 , 1.13]	
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	3.9%	-0.11 [-0.31 , 0.09]	
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	2.3%	-0.24 [-0.78 , 0.30]	
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	1.7%	-0.43 [-1.15 , 0.28]	
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.1%	-2.91 [-3.85 , -1.97]	
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.1%	-0.39 [-1.37 , 0.58]	
Shalev 2012	47.11	20.13	21	49.18	26.4	77	2.5%	-0.08 [-0.56 , 0.40]	
Shealy 1979	29.6	30.82	14	35.78	24.4	14	1.6%	-0.22 [-0.96 , 0.53]	
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	1.3%	-1.10 [-1.97 , -0.23]	
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	0.8%	-0.88 [-2.11 , 0.35]	
Tan 1986	9.5	7	10	13.44	7	9	1.2%	-0.54 [-1.46 , 0.38]	
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	1.2%	-0.56 [-1.48 , 0.36]	

Analysis 21.3. (Continued)

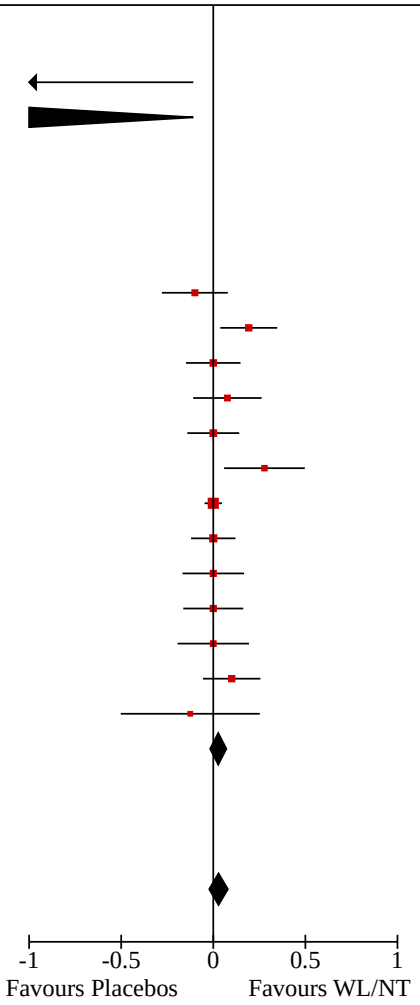
Tan 1986	9.5	7	10	13.44	7	9	1.2%	-0.54 [-1.46 , 0.38]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	1.2%	-0.56 [-1.48 , 0.36]
Turner 1979	-4.1	2.644	10	-4	2.449	10	1.3%	-0.04 [-0.91 , 0.84]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	1.1%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	1.7%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	1.7%	-0.11 [-0.82 , 0.60]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	1.1%	-0.29 [-1.28 , 0.69]
Total (95% CI)			**1246**			**1200**	**100.0%**	**-0.37 [-0.49 , -0.25]**

Heterogeneity: Tau² = 0.08; Chi² = 108.28, df = 64 (P = 0.0005); I² = 41%
Test for overall effect: Z = 6.16 (P < 0.00001)
Test for subgroup differences: Not applicable



Analysis 21.4. Comparison 21: Sensitivity analyses for all placebos compared with wait-list or no-treatment, Outcome 4: Type of data collection

Study or Subgroup	RD	SE	Weight	Risk Difference IV, Random, 95% CI
21.4.1 Structured list of reporting				
Ayen 2004	-1.5385	0.7296	0.1%	-1.54 [-2.97 , -0.11]
Subtotal (95% CI)			**0.1%**	**-1.54 [-2.97 , -0.11]**
Heterogeneity: Not applicable				
Test for overall effect: Z = 2.11 (P = 0.03)				
21.4.2 Spontaneous reporting				
Ehlers 2014	-0.1	0.0913	6.1%	-0.10 [-0.28 , 0.08]
Foa 2018	0.1925	0.0789	7.5%	0.19 [0.04 , 0.35]
Freire 2007	0	0.0754	8.0%	0.00 [-0.15 , 0.15]
Goldstein 2000	0.0769	0.0948	5.7%	0.08 [-0.11 , 0.26]
Karst 2007	0	0.0719	8.6%	0.00 [-0.14 , 0.14]
Klosko 1990	0.2778	0.1118	4.4%	0.28 [0.06 , 0.50]
Kwan 2017	0	0.024	20.9%	0.00 [-0.05 , 0.05]
Lacy 1990	0	0.0615	10.4%	0.00 [-0.12 , 0.12]
Legrand 2016	0	0.0853	6.7%	0.00 [-0.17 , 0.17]
McLachlan 1991	0	0.0831	7.0%	0.00 [-0.16 , 0.16]
Nandi 1976	0	0.0989	5.4%	0.00 [-0.19 , 0.19]
Wilson 1980	0.1	0.0795	7.5%	0.10 [-0.06 , 0.26]
Wollersheim 1991	-0.125	0.1926	1.7%	-0.13 [-0.50 , 0.25]
Subtotal (95% CI)			**99.9%**	**0.03 [-0.02 , 0.07]**
Heterogeneity: Tau² = 0.00; Chi² = 14.86, df = 12 (P = 0.25); I² = 19%				
Test for overall effect: Z = 1.20 (P = 0.23)				
Total (95% CI)			**100.0%**	**0.03 [-0.02 , 0.08]**
Heterogeneity: Tau² = 0.00; Chi² = 19.41, df = 13 (P = 0.11); I² = 33%				
Test for overall effect: Z = 1.11 (P = 0.27)				
Test for subgroup differences: Chi² = 4.60, df = 1 (P = 0.03), I² = 78.3%				



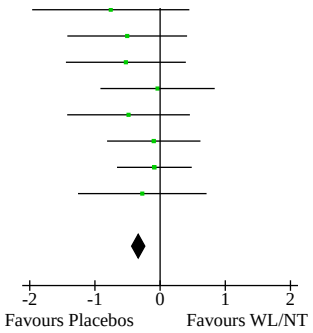
Analysis 21.5. Comparison 21: Sensitivity analyses for all placebos compared with wait-list or no-treatment, Outcome 5: Transformed data for TSA analysis

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Abikoff 2004	15	3.7376	35	16.66	3.7376	34	2.8%	-0.44 [-0.92 , 0.04]
Allen 1998	12	3.7289	11	10.82	3.7289	11	1.2%	0.30 [-0.54 , 1.15]
Allen 2006	10	3.8205	45	12.66	3.8205	44	3.2%	-0.69 [-1.12 , -0.26]
Alvarez 1997	14.67	3.7034	23	15	3.7034	25	2.2%	-0.09 [-0.65 , 0.48]
Ascher 1979	11	3.6838	8	11.37	3.6838	9	1.0%	-0.10 [-1.05 , 0.86]
Ayen 2004	9	4.3377	20	15.3	4.3377	10	1.2%	-1.41 [-2.26 , -0.56]
Borden 1986	12	3.7119	10	12.37	3.7119	10	1.1%	-0.10 [-0.97 , 0.78]
Borkovec 1975	12	3.845	11	14.62	3.845	10	1.1%	-0.65 [-1.54 , 0.23]
Borkovec 1976	8	3.7054	11	8.92	3.7054	11	1.2%	-0.24 [-1.08 , 0.60]
Bornovalova 2008	10	3.6879	18	10.41	3.6879	25	2.0%	-0.11 [-0.72 , 0.50]
Bramston 1985	11	3.7232	12	12.51	3.7232	12	1.3%	-0.39 [-1.20 , 0.42]
Carlson 1993	15	3.6911	9	15.33	3.6911	9	1.0%	-0.09 [-1.01 , 0.84]
Carter 2003	7	3.6668	28	7	3.6668	28	2.5%	0.00 [-0.52 , 0.52]
Ehlers 2014	10	3.7955	30	12.36	3.7955	30	2.5%	-0.61 [-1.13 , -0.10]
Espie 1989a	11	3.8166	14	13.4	3.8166	13	1.4%	-0.61 [-1.38 , 0.17]
Etringer 1982	14	3.722	13	15.22	3.722	12	1.3%	-0.32 [-1.11 , 0.47]
Foa 1991	6	3.7077	11	6.7	3.7077	10	1.2%	-0.18 [-1.04 , 0.68]
Foa 2018	10	3.8995	110	14.06	3.8995	40	3.6%	-1.04 [-1.42 , -0.65]
Freire 2007	8	3.6908	7	8.81	3.6908	9	0.9%	-0.21 [-1.20 , 0.78]
Fuchs 1977	13	3.9578	10	16.58	3.9578	10	1.0%	-0.87 [-1.79 , 0.06]
Goldstein 2000	10	3.8685	13	12.95	3.8685	14	1.4%	-0.74 [-1.52 , 0.05]
Hippman 2016	14	3.6672	40	14.3	3.6672	40	3.1%	-0.08 [-0.52 , 0.36]
Karst 2007	12	3.7371	19	13.96	3.7371	10	1.4%	-0.51 [-1.29 , 0.27]
Kelley 2012	5	3.7378	11	6.88	3.7378	9	1.1%	-0.48 [-1.38 , 0.41]
Kennedy 1974	10	3.779	10	12.18	3.779	10	1.1%	-0.55 [-1.45 , 0.34]
Kilmann 1987	6	3.9881	4	9.14	3.9881	4	0.4%	-0.68 [-2.15 , 0.78]
Klein 1977	10	3.8452	6	12.55	3.8452	6	0.7%	-0.61 [-1.78 , 0.56]
Klosko 1990	11	5.8948	11	13.99	5.8948	15	1.3%	-0.49 [-1.28 , 0.30]
Krapfl 1970	10	1.4982	10	11.29	1.4982	10	1.0%	-0.82 [-1.75 , 0.10]
Kwan 2017	9	3.7347	41	10.44	3.7347	40	3.1%	-0.38 [-0.82 , 0.06]
Lacy 1990	7	3.7519	15	8.77	3.7519	15	1.5%	-0.46 [-1.19 , 0.27]
Lai 2004	6	3.6976	35	5.89	3.6976	30	2.7%	0.03 [-0.46 , 0.52]
Legrand 2016	11	3.7077	11	11.52	3.7077	10	1.2%	-0.13 [-0.99 , 0.72]
Lick 1975	9	3.8396	9	11.62	3.8396	9	1.0%	-0.65 [-1.60 , 0.30]
Lick 1977	12	3.7119	10	11.63	3.7119	10	1.1%	0.10 [-0.78 , 0.97]
Liddle 1990	12	3.7342	10	10.75	3.7342	10	1.1%	0.32 [-0.56 , 1.20]
McLachlan 1991	5	3.6807	8	5.15	3.6807	12	1.1%	-0.04 [-0.93 , 0.86]
Miranda 1997	10	3.7534	11	11.66	3.7534	10	1.1%	-0.42 [-1.29 , 0.44]
Mitchell 2008	13	3.7262	15	13.63	3.7262	17	1.6%	-0.16 [-0.86 , 0.53]
Nandi 1976	6	3.7315	10	7.22	3.7315	8	1.0%	-0.31 [-1.25 , 0.63]
Nicassio 1974	14	3.7	8	13.37	3.7	8	0.9%	0.16 [-0.82 , 1.14]
Peck 1976	10	4.3558	4	14.65	4.3558	4	0.4%	-0.93 [-2.45 , 0.60]
Pelham 1992	14	3.6968	13	14.44	3.6968	13	1.4%	-0.12 [-0.88 , 0.65]
Pendleton 1983	15	4.1432	13	19.91	4.1432	15	1.3%	-1.15 [-1.96 , -0.34]
Pillman 2001	16	3.6855	7	15.56	3.6855	7	0.8%	0.11 [-0.94 , 1.16]
Poland 2013	16	3.7162	16	16.26	3.7162	14	1.6%	-0.07 [-0.79 , 0.65]
Powers 2004	6	3.7181	12	5.04	3.7181	15	1.4%	0.25 [-0.51 , 1.01]
Powers 2008a	13	3.6987	15	13.37	3.6987	10	1.3%	-0.10 [-0.90 , 0.70]
Powers 2008b	13	3.7471	18	11.38	3.7471	15	1.7%	0.42 [-0.27 , 1.12]
Proudfoot 2013	10	3.7665	195	10.41	3.7665	198	5.8%	-0.11 [-0.31 , 0.09]
Quayhagen 1995	7	3.7069	28	7.89	3.7069	25	2.4%	-0.24 [-0.78 , 0.30]
Roehrich 1993	15	3.7563	16	16.59	3.7563	15	1.6%	-0.41 [-1.13 , 0.30]
Rosa-Alcatraz 2009	11	5.4479	18	21.74	5.4479	20	1.4%	-1.93 [-2.71 , -1.15]
Rosen 1976	8	3.7337	10	9.44	3.7337	7	0.9%	-0.37 [-1.34 , 0.61]
Shalev 2012	10	3.6558	21	10.3	3.6558	77	2.8%	-0.08 [-0.56 , 0.40]
Shealy 1979	9	3.7041	14	9.81	3.7041	14	1.5%	-0.21 [-0.96 , 0.53]
Steinmark 1974	15	4.0172	12	19.06	4.0172	12	1.2%	-0.98 [-1.83 , -0.12]
Szymanski 1995	6	3.9622	5	9.25	3.9622	7	0.6%	-0.76 [-1.96 , 0.45]
Tan 1986	6	3.7652	10	7.99	3.7652	9	1.0%	-0.50 [-1.42 , 0.41]
Tori 1973	12	3.7652	9	14.07	3.7652	10	1.0%	-0.53 [-1.45 , 0.39]

Analysis 21.5. (Continued)

Tan 1986	6	3.7652	10	7.99	3.7652	9	1.0%	-0.50 [-1.42 , 0.41]
Tori 1973	12	3.7652	9	14.07	3.7652	10	1.0%	-0.53 [-1.45 , 0.39]
Turner 1979	14	3.6895	10	14.15	3.6895	10	1.1%	-0.04 [-0.92 , 0.84]
Vanderplate 1983	16	3.776	9	17.92	3.776	9	1.0%	-0.48 [-1.43 , 0.46]
Weingaertner 1971	16	3.6971	15	16.37	3.6971	15	1.6%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	13	4.5839	25	13.41	4.5839	22	2.2%	-0.09 [-0.66 , 0.49]
Wollersheim 1991	11	3.7	8	12.07	3.7	8	0.9%	-0.27 [-1.26 , 0.71]
Total (95% CI)			**1246**			**1210**	**100.0%**	**-0.34 [-0.44 , -0.23]**

Heterogeneity: Tau² = 0.04; Chi² = 83.42, df = 64 (P = 0.05); I² = 23%
Test for overall effect: Z = 6.51 (P < 0.00001)
Test for subgroup differences: Not applicable



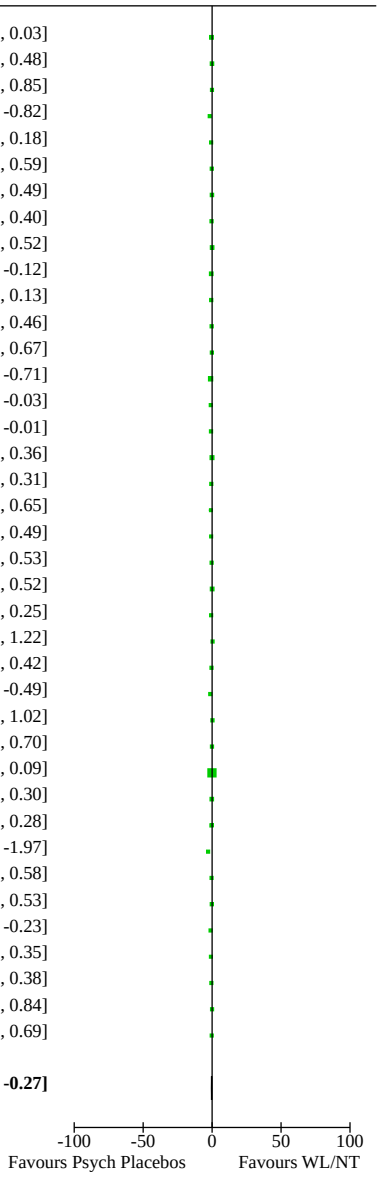
Comparison 22. Sensitivity analyses for psychological placebo compared with wait-list or no-treatment

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
22.1 Fixed effects	39	1656	Std. Mean Difference (IV, Fixed, 95% CI)	-0.37 [-0.47, -0.27]
22.2 End of intervention (post-treatment scores)	39	1656	Std. Mean Difference (IV, Random, 95% CI)	-0.47 [-0.64, -0.30]
22.3 Transformed data for TSA analysis	38	1263	Std. Mean Difference (IV, Random, 95% CI)	-0.44 [-0.59, -0.29]

Analysis 22.1. Comparison 22: Sensitivity analyses for psychological placebo compared with wait-list or no-treatment, Outcome 1: Fixed effects

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	4.4%	-0.45 [-0.93 , 0.03]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	3.1%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	1.1%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	1.3%	-1.71 [-2.60 , -0.82]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	1.3%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	1.4%	-0.25 [-1.09 , 0.59]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	2.7%	-0.11 [-0.72 , 0.49]
Bramston 1985	-78	14.21	12	-72	13.82	12	1.5%	-0.41 [-1.22 , 0.40]
Carter 2003	3.8	1.2	28	3.8	1.3	29	3.7%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	3.7%	-0.64 [-1.16 , -0.12]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	1.6%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	1.6%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	1.4%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	6.8%	-1.10 [-1.48 , -0.71]
Fuchs 1977	14.3	7	10	21.4	7	10	1.1%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	1.6%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	5.2%	-0.08 [-0.51 , 0.36]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	1.2%	-0.59 [-1.49 , 0.31]
Kilmann 1987	69	27	4	90	14	4	0.4%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	0.7%	-0.69 [-1.87 , 0.49]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	1.3%	-0.35 [-1.24 , 0.53]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	4.2%	0.03 [-0.46 , 0.52]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	1.1%	-0.71 [-1.67 , 0.25]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	1.3%	0.34 [-0.55 , 1.22]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	1.3%	-0.45 [-1.32 , 0.42]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	1.4%	-1.33 [-2.16 , -0.49]
Powers 2004	46	20.09	12	41.6	13.15	15	1.7%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	1.6%	-0.10 [-0.90 , 0.70]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	25.5%	-0.11 [-0.31 , 0.09]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	3.4%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	2.0%	-0.43 [-1.15 , 0.28]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	1.1%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.0%	-0.39 [-1.37 , 0.58]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	1.8%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	1.3%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	0.7%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	1.2%	-0.54 [-1.46 , 0.38]
Turner 1979	-4.1	2.644	10	-4	2.449	10	1.3%	-0.04 [-0.91 , 0.84]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	1.0%	-0.29 [-1.28 , 0.69]
Total (95% CI)			**868**			**788**	**100.0%**	**-0.37 [-0.47 , -0.27]**

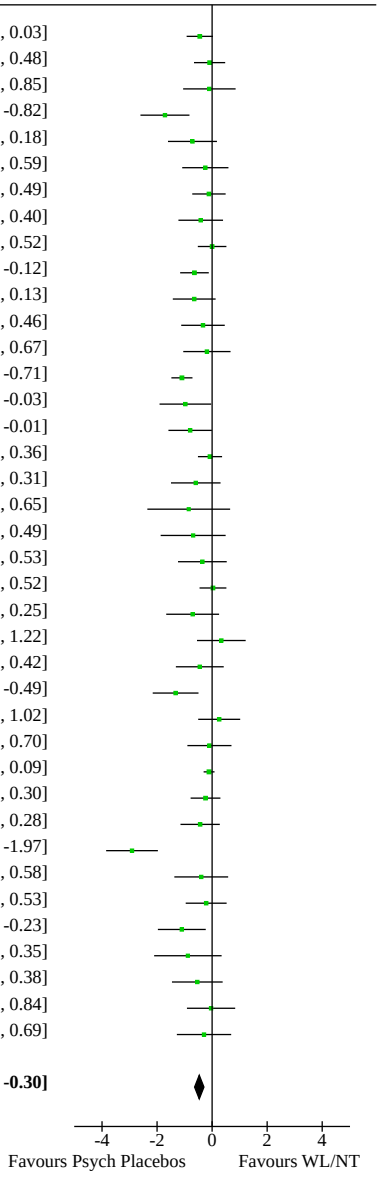
Heterogeneity: Chi² = 87.11, df = 38 (P < 0.0001); I² = 56%
Test for overall effect: Z = 7.30 (P < 0.00001)
Test for subgroup differences: Not applicable



Analysis 22.2. Comparison 22: Sensitivity analyses for psychological placebo compared with wait-list or no-treatment, Outcome 2: End of intervention (post-treatment scores)

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Abikoff 2004	-88	18.1	35	-78.5	23.6	34	3.7%	-0.45 [-0.93 , 0.03]
Alvarez 1997	20.96	6.84	23	21.56	6.09	25	3.3%	-0.09 [-0.66 , 0.48]
Ascher 1979	-3.25	1.669	8	-3.111	0.928	9	2.0%	-0.10 [-1.05 , 0.85]
Ayen 2004	14.6	4.5	20	22.1	3.7	10	2.2%	-1.71 [-2.60 , -0.82]
Borkovec 1975	2.15	0.51	11	2.98	1.53	10	2.2%	-0.71 [-1.60 , 0.18]
Borkovec 1976	20.9	29.3	11	27.6	22.6	11	2.3%	-0.25 [-1.09 , 0.59]
Bornovalova 2008	68.67	13.84	18	69.94	8.08	25	3.2%	-0.11 [-0.72 , 0.49]
Bramston 1985	-78	14.21	12	-72	13.82	12	2.4%	-0.41 [-1.22 , 0.40]
Carter 2003	3.8	1.2	28	3.8	1.3	29	3.6%	0.00 [-0.52 , 0.52]
Ehlers 2014	47.88	31.77	30	65.28	20.64	30	3.6%	-0.64 [-1.16 , -0.12]
Espie 1989a	63.5	30.6	14	96.5	63.5	13	2.5%	-0.65 [-1.43 , 0.13]
Etringer 1982	-11.39	3.95	13	-10.08	3.68	12	2.5%	-0.33 [-1.12 , 0.46]
Foa 1991	18.09	7.13	11	19.5	7.18	10	2.3%	-0.19 [-1.05 , 0.67]
Foa 2018	18.65	4.28	110	23.06	3.09	40	4.2%	-1.10 [-1.48 , -0.71]
Fuchs 1977	14.3	7	10	21.4	7	10	2.0%	-0.97 [-1.91 , -0.03]
Goldstein 2000	1.54	0.52	13	2.05	0.7	14	2.5%	-0.80 [-1.59 , -0.01]
Hippman 2016	61.4	15.9	40	62.7	17.8	40	3.9%	-0.08 [-0.51 , 0.36]
Kennedy 1974	-9.3	4.25	10	-6.95	3.27	10	2.1%	-0.59 [-1.49 , 0.31]
Kilmann 1987	69	27	4	90	14	4	1.0%	-0.85 [-2.35 , 0.65]
Klein 1977	10.17	5.49	6	15	7.32	6	1.5%	-0.69 [-1.87 , 0.49]
Krapfl 1970	-11.8	4.25	10	-10.4	3.27	10	2.2%	-0.35 [-1.24 , 0.53]
Lai 2004	-1.41	0.26	35	-1.42	0.37	30	3.7%	0.03 [-0.46 , 0.52]
Lick 1975	-6.33	2.74	9	-4.33	2.65	9	2.0%	-0.71 [-1.67 , 0.25]
Liddle 1990	19.3	6.93	10	16.9	6.79	10	2.2%	0.34 [-0.55 , 1.22]
Miranda 1997	5.2	3.6	11	6.9	3.7	10	2.2%	-0.45 [-1.32 , 0.42]
Pendleton 1983	-50.81	20.82	13	-29.73	8.4	15	2.3%	-1.33 [-2.16 , -0.49]
Powers 2004	46	20.09	12	41.6	13.15	15	2.6%	0.26 [-0.51 , 1.02]
Powers 2008a	66	36	15	69	16	10	2.4%	-0.10 [-0.90 , 0.70]
Proudfoot 2013	39.01	23.94	195	41.56	22.11	198	5.0%	-0.11 [-0.31 , 0.09]
Quayhagen 1995	-108.3	14.8	28	-104.8	13.9	25	3.5%	-0.24 [-0.78 , 0.30]
Roehrich 1993	-53	10.18	16	-48.93	7.81	15	2.7%	-0.43 [-1.15 , 0.28]
Rosa-Alcatraz 2009	3.05	2.04	18	8.25	1.44	20	2.0%	-2.91 [-3.85 , -1.97]
Rosen 1976	-4.5	4.55	10	-2.86	2.8	7	1.9%	-0.39 [-1.37 , 0.58]
Shealy 1979	29.6	30.82	14	35.78	24.4	14	2.6%	-0.22 [-0.96 , 0.53]
Steinmark 1974	2.17	0.51	12	3.47	1.53	12	2.2%	-1.10 [-1.97 , -0.23]
Szymanski 1995	-5.4	2.7	5	-3.57	1.13	7	1.4%	-0.88 [-2.11 , 0.35]
Tan 1986	9.5	7	10	13.44	7	9	2.1%	-0.54 [-1.46 , 0.38]
Turner 1979	-4.1	2.644	10	-4	2.449	10	2.2%	-0.04 [-0.91 , 0.84]
Wollersheim 1991	9.13	5.62	8	10.88	5.64	8	1.9%	-0.29 [-1.28 , 0.69]
Total (95% CI)			**868**			**788**	**100.0%**	**-0.47 [-0.64 , -0.30]**

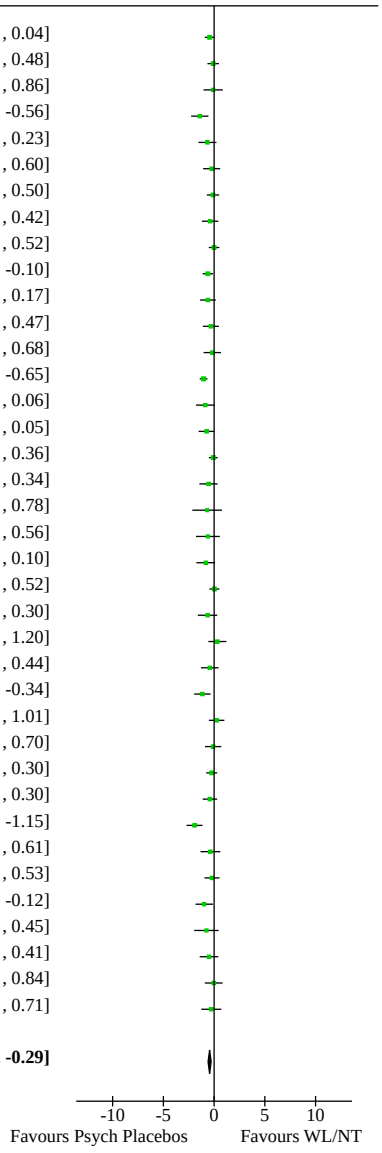
Heterogeneity: Tau² = 0.14; Chi² = 87.11, df = 38 (P < 0.0001); I² = 56%
Test for overall effect: Z = 5.40 (P < 0.00001)
Test for subgroup differences: Not applicable



Analysis 22.3. Comparison 22: Sensitivity analyses for psychological placebo compared with wait-list or no-treatment, Outcome 3: Transformed data for TSA analysis

Study or Subgroup	Psychological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Abikoff 2004	15	3.7376	35	16.66	3.7376	34	4.4%	-0.44 [-0.92 , 0.04]
Alvarez 1997	14.67	3.7034	23	15	3.7034	25	3.7%	-0.09 [-0.65 , 0.48]
Ascher 1979	11	3.6838	8	11.37	3.6838	9	1.9%	-0.10 [-1.05 , 0.86]
Ayen 2004	9	4.3377	20	15.3	4.3377	10	2.2%	-1.41 [-2.26 , -0.56]
Borkovec 1975	12	3.845	11	14.62	3.845	10	2.1%	-0.65 [-1.54 , 0.23]
Borkovec 1976	8	3.7054	11	8.92	3.7054	11	2.3%	-0.24 [-1.08 , 0.60]
Bornovalova 2008	10	3.6879	18	10.41	3.6879	25	3.5%	-0.11 [-0.72 , 0.50]
Bramston 1985	11	3.7232	12	12.51	3.7232	12	2.4%	-0.39 [-1.20 , 0.42]
Carter 2003	7	3.6668	28	7	3.6668	29	4.1%	0.00 [-0.52 , 0.52]
Ehlers 2014	10	3.7955	30	12.36	3.7955	30	4.1%	-0.61 [-1.13 , -0.10]
Espie 1989a	11	3.8166	14	13.4	3.8166	13	2.6%	-0.61 [-1.38 , 0.17]
Etringer 1982	14	3.722	13	15.22	3.722	12	2.5%	-0.32 [-1.11 , 0.47]
Foa 1991	6	3.7077	11	6.7	3.7707	10	2.2%	-0.18 [-1.04 , 0.68]
Foa 2018	10	3.8995	110	14.06	3.8995	40	5.3%	-1.04 [-1.42 , -0.65]
Fuchs 1977	13	3.9578	10	16.58	3.9578	10	2.0%	-0.87 [-1.79 , 0.06]
Goldstein 2000	10	3.8685	13	12.95	3.8685	14	2.5%	-0.74 [-1.52 , 0.05]
Hippman 2016	14	3.6672	40	14.3	3.6672	40	4.7%	-0.08 [-0.52 , 0.36]
Kennedy 1974	10	3.779	10	12.18	3.779	10	2.1%	-0.55 [-1.45 , 0.34]
Kilmann 1987	6	3.9881	4	9.14	3.9881	4	0.9%	-0.68 [-2.15 , 0.78]
Klein 1977	10	3.8452	6	12.55	3.8452	6	1.4%	-0.61 [-1.78 , 0.56]
Krapfl 1970	10	1.4982	10	11.29	1.4982	10	2.0%	-0.82 [-1.75 , 0.10]
Lai 2004	6	3.6976	35	5.89	3.6976	30	4.3%	0.03 [-0.46 , 0.52]
Lick 1975	9	3.8396	9	11.62	3.8396	9	1.9%	-0.65 [-1.60 , 0.30]
Liddle 1990	12	3.7342	10	10.75	3.7342	10	2.1%	0.32 [-0.56 , 1.20]
Miranda 1997	10	3.7534	11	11.66	3.7534	10	2.2%	-0.42 [-1.29 , 0.44]
Pendleton 1983	15	4.1432	13	19.91	4.1432	15	2.4%	-1.15 [-1.96 , -0.34]
Powers 2004	6	3.7181	12	5.04	3.7182	15	2.6%	0.25 [-0.51 , 1.01]
Powers 2008a	13	3.6987	15	13.37	3.6987	10	2.4%	-0.10 [-0.90 , 0.70]
Quayhagen 1995	7	3.7069	28	7.89	3.7069	25	3.9%	-0.24 [-0.78 , 0.30]
Roehrich 1993	15	3.7563	16	16.59	3.7563	15	2.8%	-0.41 [-1.13 , 0.30]
Rosa-Alcatraz 2009	11	5.4479	18	21.74	5.4479	20	2.5%	-1.93 [-2.71 , -1.15]
Rosen 1976	8	3.7337	10	9.44	3.7337	7	1.8%	-0.37 [-1.34 , 0.61]
Shealy 1979	9	3.7041	14	9.81	3.7041	14	2.7%	-0.21 [-0.96 , 0.53]
Steinmark 1974	15	4.0172	12	19.06	4.0172	12	2.2%	-0.98 [-1.83 , -0.12]
Szymanski 1995	6	3.9622	5	9.25	3.9622	7	1.3%	-0.76 [-1.96 , 0.45]
Tan 1986	6	3.7652	10	7.99	3.7652	9	2.0%	-0.50 [-1.42 , 0.41]
Turner 1979	14	3.6895	10	14.15	3.6895	10	2.1%	-0.04 [-0.92 , 0.84]
Wollersheim 1991	11	3.7	8	12.07	3.7	8	1.8%	-0.27 [-1.26 , 0.71]
Total (95% CI)			**673**			**590**	**100.0%**	**-0.44 [-0.59 , -0.29]**

Heterogeneity: Tau² = 0.07; Chi² = 57.60, df = 37 (P = 0.02); I² = 36%
Test for overall effect: Z = 5.79 (P < 0.00001)
Test for subgroup differences: Not applicable



Comparison 23. Sensitivity analyses for pharmacological placebo compared with wait-list or no-treatment

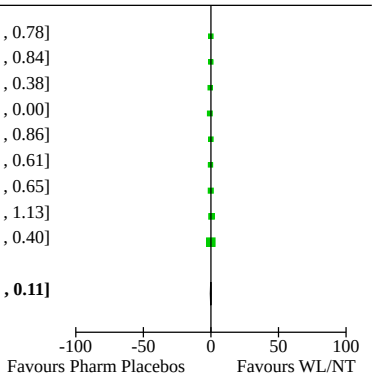
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
23.1 Fixed effects	9	279	Std. Mean Difference (IV, Fixed, 95% CI)	-0.14 [-0.39, 0.11]
23.2 Fixed effects	8	366	Risk Ratio (IV, Fixed, 95% CI)	0.91 [0.77, 1.07]
23.3 End of intervention (post-treatment scores)	9	279	Std. Mean Difference (IV, Random, 95% CI)	-0.14 [-0.39, 0.11]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
23.4 Transformed data for TSA analysis	9	279	Std. Mean Difference (IV, Random, 95% CI)	-0.08 [-0.33, 0.17]

Analysis 23.1. Comparison 23: Sensitivity analyses for pharmacological placebo compared with wait-list or no-treatment, Outcome 1: Fixed effects

Study or Subgroup	Pharmacological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Borden 1986	29.7	17.5	10	32	26	10	8.3%	-0.10 [-0.98 , 0.78]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	7.5%	-0.09 [-1.01 , 0.84]
Kelley 2012	-1.64	4.52	11	0.67	4	9	7.9%	-0.51 [-1.41 , 0.38]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	9.6%	-0.81 [-1.63 , 0.00]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	8.0%	-0.04 [-0.93 , 0.86]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	7.3%	-0.33 [-1.27 , 0.61]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	10.8%	-0.12 [-0.89 , 0.65]
Powers 2008b	64	20	18	51	37	15	13.2%	0.44 [-0.26 , 1.13]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	27.4%	-0.08 [-0.56 , 0.40]
Total (95% CI)			**111**			**168**	**100.0%**	**-0.14 [-0.39 , 0.11]**

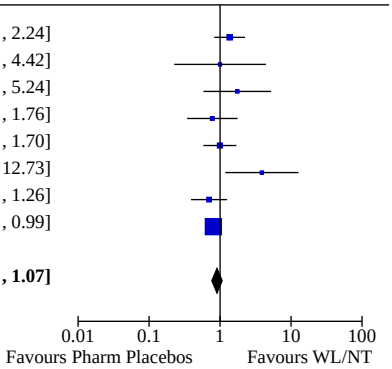
Heterogeneity: Chi² = 6.26, df = 8 (P = 0.62); I² = 0%
Test for overall effect: Z = 1.07 (P = 0.28)
Test for subgroup differences: Not applicable



Analysis 23.2. Comparison 23: Sensitivity analyses for pharmacological placebo compared with wait-list or no-treatment, Outcome 2: Fixed effects

Study or Subgroup	Pharmacological placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio IV, Fixed, 95% CI
Berg 1983	9	11	9	15	11.1%	1.36 [0.83 , 2.24]
Double 1993	3	22	3	22	1.2%	1.00 [0.23 , 4.42]
Klerman 1974a	7	25	4	25	2.3%	1.75 [0.58 , 5.24]
Klerman 1974b	7	25	9	25	4.1%	0.78 [0.34 , 1.76]
Rabkin 1990	14	27	12	23	9.7%	0.99 [0.58 , 1.70]
Watzl 1988	11	34	3	36	2.0%	3.88 [1.18 , 12.73]
Whittaker 1963	7	13	10	13	8.1%	0.70 [0.39 , 1.26]
Wilson 1980	31	40	10	10	61.5%	0.80 [0.65 , 0.99]
Total (95% CI)		**197**		**169**	**100.0%**	**0.91 [0.77 , 1.07]**
Total events:	89		60			

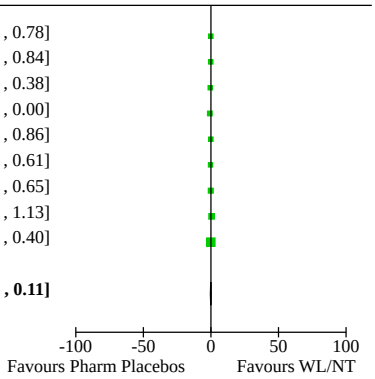
Heterogeneity: Chi² = 11.95, df = 7 (P = 0.10); I² = 41%
Test for overall effect: Z = 1.17 (P = 0.24)
Test for subgroup differences: Not applicable



Analysis 23.3. Comparison 23: Sensitivity analyses for pharmacological placebo compared with wait-list or no-treatment, Outcome 3: End of intervention (post-treatment scores)

Study or Subgroup	Pharmacological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Borden 1986	29.7	17.5	10	32	26	10	8.3%	-0.10 [-0.98 , 0.78]
Carlson 1993	5.2	3.3	9	5.5	3.3	9	7.5%	-0.09 [-1.01 , 0.84]
Kelley 2012	-1.64	4.52	11	0.67	4	9	7.9%	-0.51 [-1.41 , 0.38]
Klosko 1990	3.55	1.51	11	4.8	1.47	15	9.6%	-0.81 [-1.63 , 0.00]
McLachlan 1991	-6.25	12.25	8	-5.67	17.35	12	8.0%	-0.04 [-0.93 , 0.86]
Nandi 1976	53.2	11.2	10	57.5	13.8	8	7.3%	-0.33 [-1.27 , 0.61]
Pelham 1992	-8.72	1.73	13	-8.51	1.63	13	10.8%	-0.12 [-0.89 , 0.65]
Powers 2008b	64	20	18	51	37	15	13.2%	0.44 [-0.26 , 1.13]
Shalev 2012	47.11	20.13	21	49.18	26.4	77	27.4%	-0.08 [-0.56 , 0.40]
Total (95% CI)			**111**			**168**	**100.0%**	**-0.14 [-0.39 , 0.11]**

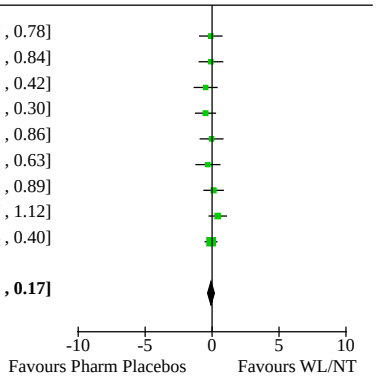
Heterogeneity: Tau² = 0.00; Chi² = 6.26, df = 8 (P = 0.62); I² = 0%
Test for overall effect: Z = 1.07 (P = 0.28)
Test for subgroup differences: Not applicable



Analysis 23.4. Comparison 23: Sensitivity analyses for pharmacological placebo compared with wait-list or no-treatment, Outcome 4: Transformed data for TSA analysis

Study or Subgroup	Pharmacological placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Borden 1986	10	23.3669	10	12.33	23.3669	10	8.2%	-0.10 [-0.97 , 0.78]
Carlson 1993	7	23.2921	9	9.09	23.2921	9	7.4%	-0.09 [-1.01 , 0.84]
Kelley 2012	9	23.6503	11	20.87	23.6503	9	7.9%	-0.48 [-1.38 , 0.42]
Klosko 1990	14	37.2331	11	32.85	37.2331	15	10.1%	-0.49 [-1.28 , 0.30]
McLachlan 1991	11	23.2892	8	11.93	23.2892	12	7.9%	-0.04 [-0.93 , 0.86]
Nandi 1976	9	23.5273	10	16.67	23.5273	8	7.2%	-0.31 [-1.25 , 0.63]
Pelham 1992	13	23.3025	13	10.21	23.3025	13	10.7%	0.12 [-0.65 , 0.89]
Powers 2008b	14	23.5696	18	3.76	23.5696	15	13.2%	0.42 [-0.27 , 1.12]
Shalev 2012	10	23.1535	21	11.86	23.1535	77	27.2%	-0.08 [-0.56 , 0.40]
Total (95% CI)			**111**			**168**	**100.0%**	**-0.08 [-0.33 , 0.17]**

Heterogeneity: Tau² = 0.00; Chi² = 4.32, df = 8 (P = 0.83); I² = 0%
Test for overall effect: Z = 0.63 (P = 0.53)
Test for subgroup differences: Not applicable



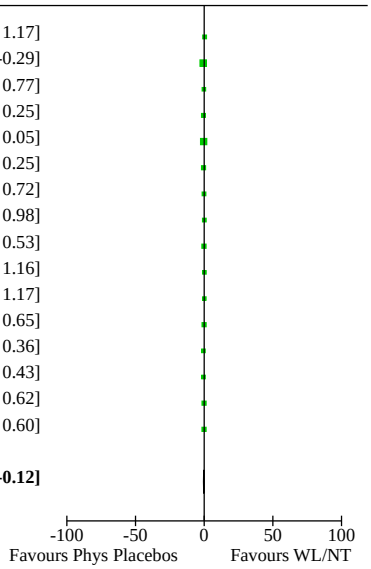
Comparison 24. Sensitivity analyses for physical placebo compared with wait-list or no-treatment

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
24.1 Fixed effects	16	503	Std. Mean Difference (IV, Fixed, 95% CI)	-0.30 [-0.48, -0.12]
24.2 End of intervention (post-treatment scores)	16	503	Std. Mean Difference (IV, Random, 95% CI)	-0.30 [-0.48, -0.12]
24.3 Transformed data for TSA analysis	17	896	Std. Mean Difference (IV, Random, 95% CI)	-0.21 [-0.34, -0.07]

Analysis 24.1. Comparison 24: Sensitivity analyses for physical placebo compared with wait-list or no-treatment, Outcome 1: Fixed effects

Study or Subgroup	Physical placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	4.5%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	17.3%	-0.72 [-1.15 , -0.29]
Freire 2007	24.6	11	7	28.2	18	9	3.2%	-0.22 [-1.21 , 0.77]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	5.2%	-0.53 [-1.30 , 0.25]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	16.5%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	6.0%	-0.48 [-1.21 , 0.25]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	4.3%	-0.14 [-1.00 , 0.72]
Lick 1977	65.86	25.45	10	62.84	31.78	10	4.1%	0.10 [-0.78 , 0.98]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	6.6%	-0.17 [-0.87 , 0.53]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	3.3%	0.17 [-0.81 , 1.16]
Pillman 2001	-85.3	14	7	-87.6	20	7	2.9%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	6.2%	-0.07 [-0.79 , 0.65]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	3.7%	-0.56 [-1.48 , 0.36]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	3.6%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	6.2%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	6.3%	-0.11 [-0.82 , 0.60]
Total (95% CI)			**263**			**240**	**100.0%**	**-0.30 [-0.48 , -0.12]**

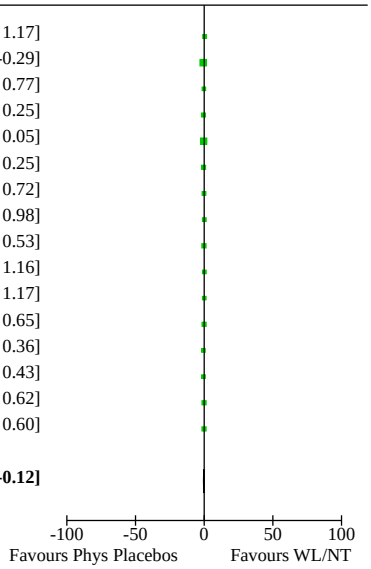
Heterogeneity: Chi² = 10.66, df = 15 (P = 0.78); I² = 0%
Test for overall effect: Z = 3.29 (P = 0.001)
Test for subgroup differences: Not applicable



Analysis 24.2. Comparison 24: Sensitivity analyses for physical placebo compared with wait-list or no-treatment, Outcome 2: End of intervention (post-treatment scores)

Study or Subgroup	Physical placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Allen 1998	-2.9	7.9	11	-6.1	10.9	11	4.5%	0.32 [-0.52 , 1.17]
Allen 2006	12	9.6	45	19	9.6	44	17.3%	-0.72 [-1.15 , -0.29]
Freire 2007	24.6	11	7	28.2	18	9	3.2%	-0.22 [-1.21 , 0.77]
Karst 2007	-2.79	0.54	19	-2.5	0.53	10	5.2%	-0.53 [-1.30 , 0.25]
Kwan 2017	39.96	16.9683	41	46.43	15.4319	40	16.5%	-0.39 [-0.83 , 0.05]
Lacy 1990	5.71	1.95	15	6.79	2.42	15	6.0%	-0.48 [-1.21 , 0.25]
Legrand 2016	28.43	7.46	11	29.9	12.57	10	4.3%	-0.14 [-1.00 , 0.72]
Lick 1977	65.86	25.45	10	62.84	31.78	10	4.1%	0.10 [-0.78 , 0.98]
Mitchell 2008	-79.2	6.5	15	-77.5	11.9	17	6.6%	-0.17 [-0.87 , 0.53]
Nicassio 1974	117.29	133.7	8	99.25	35.3	8	3.3%	0.17 [-0.81 , 1.16]
Pillman 2001	-85.3	14	7	-87.6	20	7	2.9%	0.12 [-0.92 , 1.17]
Poland 2013	22	8.2	16	22.6	8	14	6.2%	-0.07 [-0.79 , 0.65]
Tori 1973	-6.22	2.22	9	-4.9	2.28	10	3.7%	-0.56 [-1.48 , 0.36]
Vanderplate 1983	2	0.6	9	2.3	0.5	9	3.6%	-0.52 [-1.46 , 0.43]
Weingaertner 1971	1.06	1.29	15	1.2	1.56	15	6.2%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	136.87	16.98	25	138.91	20.14	11	6.3%	-0.11 [-0.82 , 0.60]
Total (95% CI)			**263**			**240**	**100.0%**	**-0.30 [-0.48 , -0.12]**

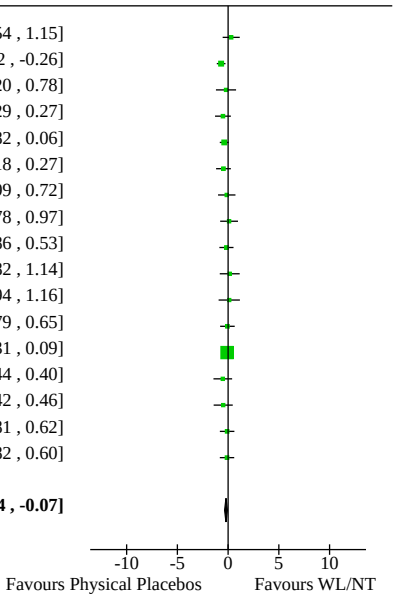
Heterogeneity: Tau² = 0.00; Chi² = 10.66, df = 15 (P = 0.78); I² = 0%
Test for overall effect: Z = 3.29 (P = 0.001)
Test for subgroup differences: Not applicable



Analysis 24.3. Comparison 24: Sensitivity analyses for physical placebo compared with wait-list or no-treatment, Outcome 3: Transformed data for TSA analysis

Study or Subgroup	Physical placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Allen 1998	8	9.7092	11	4.93	9.7092	11	2.5%	0.30 [-0.54 , 1.15]
Allen 2006	15	9.9522	45	21.91	9.9522	44	9.6%	-0.69 [-1.12 , -0.26]
Freire 2007	11	9.6239	7	13.11	9.6239	9	1.8%	-0.21 [-1.20 , 0.78]
Karst 2007	6	9.7266	19	11.09	9.7266	10	2.9%	-0.51 [-1.29 , 0.27]
Kwan 2017	12	9.7193	41	15.74	9.7193	40	9.1%	-0.38 [-0.82 , 0.06]
Lacy 1990	7	9.8042	15	11.61	9.8042	15	3.3%	-0.46 [-1.18 , 0.27]
Legrand 2016	12	9.6354	11	13.34	9.6354	10	2.4%	-0.13 [-0.99 , 0.72]
Lick 1977	10	9.6375	10	9.04	9.6375	10	2.3%	0.10 [-0.78 , 0.97]
Mitchell 2008	8	9.6825	15	9.63	9.6825	17	3.6%	-0.16 [-0.86 , 0.53]
Nicassio 1974	13	9.64	8	11.37	9.64	8	1.8%	0.16 [-0.82 , 1.14]
Pillman 2001	6	9.5786	7	4.85	9.5786	7	1.6%	0.11 [-0.94 , 1.16]
Poland 2013	14	9.6458	16	14.67	9.6458	14	3.4%	-0.07 [-0.79 , 0.65]
Proudfoot 2013	9	9.7136	195	10.06	9.7136	198	44.8%	-0.11 [-0.31 , 0.09]
Tori 1973	11	9.8157	9	16.38	9.8157	10	2.1%	-0.52 [-1.44 , 0.40]
Vanderplate 1983	11	9.8217	9	15.99	9.8217	9	2.0%	-0.48 [-1.42 , 0.46]
Weingaertner 1971	8	9.5851	15	8.96	9.5851	15	3.4%	-0.10 [-0.81 , 0.62]
Wolitzky 2009	15	9.6182	25	16.06	9.6182	11	3.5%	-0.11 [-0.82 , 0.60]
Total (95% CI)			**458**			**438**	**100.0%**	**-0.21 [-0.34 , -0.07]**

Heterogeneity: Tau² = 0.00; Chi² = 11.34, df = 16 (P = 0.79); I² = 0%
Test for overall effect: Z = 3.07 (P = 0.002)
Test for subgroup differences: Not applicable



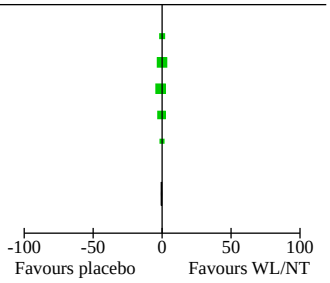
Comparison 25. Sensitivity analyses for usual care compared with wait-list or no-treatment

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
25.1 Fixed effects	5	523	Std. Mean Difference (IV, Fixed, 95% CI)	-0.46 [-0.64, -0.28]
25.2 Transformed data for TSA analysis	5	523	Std. Mean Difference (IV, Random, 95% CI)	-0.32 [-0.78, 0.14]

Analysis 25.1. Comparison 25: Sensitivity analyses for usual care compared with wait-list or no-treatment, Outcome 1: Fixed effects

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Fixed, 95% CI
Crisp 1991	-5.5	3.2863	30	-5.6	3.1305	20	9.8%	0.03 [-0.54 , 0.60]
Glogowska 2000	-4.9	1.7	71	-4.8	1.8	84	31.4%	-0.06 [-0.37 , 0.26]
Rapee 2006	3.4	2.6	90	5.8	1.6	87	31.3%	-1.10 [-1.42 , -0.79]
Rapee 2007	28.63	15.01	59	35.337	15.83	52	22.0%	-0.43 [-0.81 , -0.06]
Teri 1997	13.9	4.14	10	14.1	2.32	20	5.4%	-0.06 [-0.82 , 0.70]
Total (95% CI)			**260**			**263**	**100.0%**	**-0.46 [-0.64 , -0.28]**

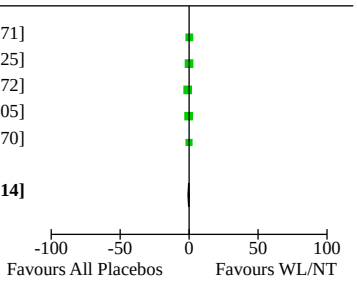
Heterogeneity: Chi² = 26.02, df = 4 (P < 0.0001); I² = 85%
Test for overall effect: Z = 5.07 (P < 0.00001)
Test for subgroup differences: Not applicable



Analysis 25.2. Comparison 25: Sensitivity analyses for usual care compared with wait-list or no-treatment, Outcome 2: Transformed data for TSA analysis

Study or Subgroup	All placebos Mean	SD	Total	Wait-list/no-treatment Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Crisp 1991	10	1.7321	30	9.74	1.7321	20	18.3%	0.15 [-0.42 , 0.71]
Glogowska 2000	12	1.7368	71	12.11	1.7368	84	22.5%	-0.06 [-0.38 , 0.25]
Rapee 2006	8	1.8623	90	9.93	1.8623	87	22.6%	-1.03 [-1.35 , -0.72]
Rapee 2007	11	1.7349	59	11.75	1.7349	52	21.6%	-0.43 [-0.81 , -0.05]
Teri 1997	12	1.7558	10	12.11	1.7558	20	15.0%	-0.06 [-0.82 , 0.70]
Total (95% CI)			**260**			**263**	**100.0%**	**-0.32 [-0.78 , 0.14]**

Heterogeneity: Tau² = 0.22; Chi² = 24.19, df = 4 (P < 0.0001); I² = 83%
Test for overall effect: Z = 1.36 (P = 0.17)
Test for subgroup differences: Not applicable



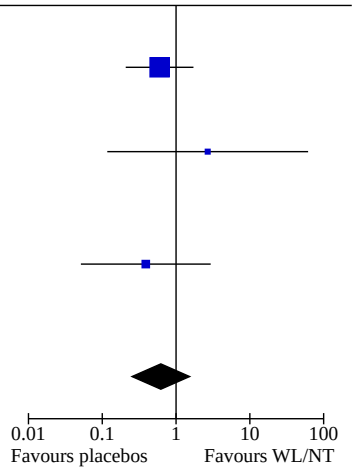
Comparison 26. Sensitivity analysis for serious adverse events of placebos

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
26.1 Data in RR for TSA analysis	11	517	Risk Ratio (M-H, Random, 95% CI)	0.62 [0.25, 1.53]

Analysis 26.1. Comparison 26: Sensitivity analysis for serious adverse events of placebos, Outcome 1: Data in RR for TSA analysis

Study or Subgroup	Placebos Events	Total	Wait-list/no-treatment Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Ehlers 2014	0	30	0	30		Not estimable
Foa 2018	8	107	5	40	72.1%	0.60 [0.21 , 1.72]
Freire 2007	0	12	0	12		Not estimable
Karst 2007	0	19	0	10		Not estimable
Klosko 1990	1	18	0	16	8.2%	2.68 [0.12 , 61.58]
Kwan 2017	0	41	0	40		Not estimable
Lacy 1990	0	15	0	15		Not estimable
Legrand 2016	0	11	0	10		Not estimable
McLachlan 1991	1	9	4	14	19.6%	0.39 [0.05 , 2.95]
Nandi 1976	0	10	0	8		Not estimable
Wilson 1980	0	40	0	10		Not estimable
Total (95% CI)		**312**		**205**	**100.0%**	**0.62 [0.25 , 1.53]**
Total events:	10		9			

Heterogeneity: Tau² = 0.00; Chi² = 1.05, df = 2 (P = 0.59); I² = 0%
Test for overall effect: Z = 1.04 (P = 0.30)
Test for subgroup differences: Not applicable



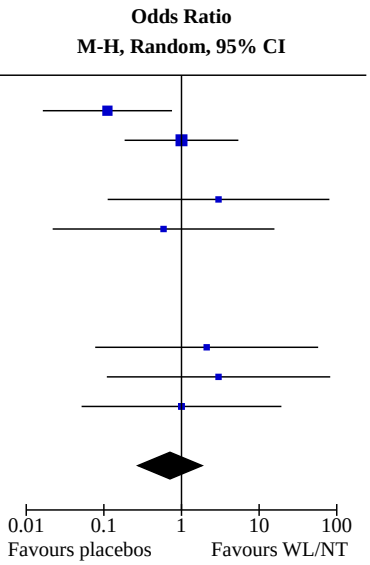
Comparison 27. Sensitivity analysis for non-serious adverse events of placebos

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
27.1 Data in RR for TSA analysis	11	362	Odds Ratio (M-H, Random, 95% CI)	0.71 [0.27, 1.84]

Analysis 27.1. Comparison 27: Sensitivity analysis for non-serious adverse events of placebos, Outcome 1: Data in RR for TSA analysis

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Odds Ratio M-H, Random, 95% CI
Ayen 2004	2	20	5	10	24.6%	0.11 [0.02 , 0.75]
Ehlers 2014	27	30	27	30	31.8%	1.00 [0.19 , 5.40]
Freire 2007	12	12	12	12		Not estimable
Goldstein 2000	13	13	13	14	8.4%	3.00 [0.11 , 80.39]
Karst 2007	18	19	10	10	8.4%	0.59 [0.02 , 15.75]
Kwan 2017	41	41	40	40		Not estimable
Lacy 1990	15	15	15	15		Not estimable
Legrand 2016	11	11	10	10		Not estimable
McLachlan 1991	9	9	13	14	8.3%	2.11 [0.08 , 57.61]
Nandi 1976	10	10	10	11	8.2%	3.00 [0.11 , 82.40]
Wollersheim 1991	7	8	7	8	10.3%	1.00 [0.05 , 19.36]
Total (95% CI)		**188**		**174**	**100.0%**	**0.71 [0.27 , 1.84]**
Total events:	165		162			

Heterogeneity: Tau² = 0.00; Chi² = 5.71, df = 6 (P = 0.46); I² = 0%
Test for overall effect: Z = 0.70 (P = 0.48)
Test for subgroup differences: Not applicable



ADDITIONAL TABLES

Table 1. Description of experimental interventions

Name of intervention	Other common names reported in the literature	Definition	Examples
Psychological placebo	Attention placebo, credible placebo, common-factor treatment control, sham intervention, pseudo control	Psychological placebos target the non-specific or shared components of psychological treatments, such as human interaction variables, attending sessions, and patient expectations	In an example from Tan 1986, the psychological placebo participants were exposed to sessions of supportive group counselling, which were thought to represent the non-specific component of the active intervention, which was cognitive behavioural therapy
Pharmacological placebo	Pill placebo, placebo tablet, dummy pill	Pharmacological placebos are inert substances, typically in pill or liquid form, which do not contain the active ingredients of a given drug treatment	Participants typically receive a pill containing starch, sugar, or lactose (Double 1993; Meissner 2011b).

Table 1. Description of experimental interventions *(Continued)*

Physical placebo	Sham intervention, credible placebo, pseudo control	Physical placebos target the inert components of a given physical treatment (such as acupuncture, needle injection, exercise regimens, surgery, or electromagnetic stimulation)	This could be sham acupuncture where the needles are blunted (Tough 2009), or sham electromagnetic stimulation, where the machine is not turned on or electrodes are attached to inactive sites (Sommer 2006)
Usual care	Treatment as usual (TAU), standard care, outpatient care, standard practice, support as usual, clinical care, routine care, existing-practice control	Usual care reflects locally accepted treatment practices for a given mental health disorder. It is provided either by private or public practitioners and may involve both pharmacological and psychological treatment.	Patients allocated to usual care might receive a large variety of therapies with a theoretical blend of psychodynamic, humanistic and behavioural approaches (Borduin 2009)
Wait-list	Minimal contact control, delayed treatment	Wait-list participants are assessed on repeated occasions, but are promised the "active" intervention after the trial has ended	In Ertl 2011, participants allocated to the wait-list group were reassessed at baseline and follow-up, and subsequently offered the active treatment, which was narrative exposure therapy
No-treatment	Minimal contact control	No-treatment participants are assessed on repeated occasions without receiving the active treatment intervention. Unlike wait-list interventions, no-treatment participants are not promised the "active" intervention after trial completion	No-treatment participants in Miranda 2003 did not receive any mental-health related treatment, and were not promised the active intervention (antidepressants or psychotherapy) after trial completion

Table 2. Description of trials with missing data

Trial ID	Missing data	Methods of data generation
Borden 1986	Missing information regarding how many participants randomised to each group	Since 30 patients was included in total, we assumed due to randomisation that there was an equal distribution of patients in each group.
Borkovec 1975	The standard deviation (SD) was not reported on the Daily Sleep Questionnaire (DSQ) – Subscale difficulty experienced in falling asleep	An artificial SD was imported from Steinmark 1974 due to similar outcome, population, and control interventions
Brill 1964a ; Brill 1964b	No usable data reported	Not possible to generate data
Carlson 1993	Missing information regarding how many participants randomised to each group	Since 28 patients was included in total, we assumed due to randomisation and ethical principles that the active arm (in this case methylphenidate) included an additional patient compared with the placebo and no-treatment group
Crouch 1988	No usable data reported	Not possible to generate data, and could not get in contact with authors
Doty 1975	No usable data reported	Not possible to generate data

Table 2. Description of trials with missing data *(Continued)*

Fuchs 1977	The standard deviation (SD) on the improvement of the mean on Beck Depression Inventory was not reported.	The SD was calculated from a F-test statistic
Goldwasser 1987	The SD was not reported on the Mini-Mental State (MMS)	Not possible to generate data
Hekmat 1984	The SD was not reported on the Timed Behavior Checklist (BCL)	Not possible to generate data
Howlin 2007	No usable data reported	Not possible to generate data, and could not get in contact with authors
Krapfl 1970	Mean was only reported in a figure, and the SD on the Behavioral Avoidance Test (BAT) was not reported	Mean value was generated from figure 1, and the SD was generated from Etringer 1982 ; Rosen 1976 , due to that both trials included similar outcome, population, and control interventions
Lang 1965	No usable data reported	Not possible to generate data
Pelham 1992	Missing information regarding how many participants randomised to each group	Since 38 patients was included in total, we assumed due to randomisation and ethical principles that the active arm (methylphenidate) and placebo included an additional patient each
Pendleton 1983	Missing information regarding how many participants randomised to each group	Since 62 patients was included in total, we assumed due to randomisation and ethical principles that the two active arms (Negative practice and Desensitization) included an additional patient each
Roth 1964	The standard deviation (SD) was not reported on either Global estimate of the severity of the patient's illness and a global measure of over-all improvement	Not possible to generate data
Rupert 1978	No usable data reported	Not possible to generate data
Shealy 1979	Separate means missing mild to moderate group. The standard deviation (SD) was not reported on the Daily Sleep and Relaxation Practice Questionnaire (DSRQ)	We pooled the means from mild and moderate due to no information of patients in each group. A mean score was generated from the two groups. SDs was generated from Ascher 1979 ; Steinmark 1974 due to that these trials hade the same outcomes and population
Sibilio 1957	No usable data reported	Not possible to generate data
Sommerness 1955	No usable data reported	Not possible to generate data
Tan 1986	The standard deviation (SD) on the improvement of the mean on Beck Depression Inventory was not reported	The SD was imported from Fuchs 1977 due to similar outcome and population
Trexler 1972	The standard deviation (SD) was not reported on the Behavioral Checklist	Not possible to generate data
Matson 1980	No usable data reported	Not possible to generate data, and could not get in contact with authors

Table 2. Description of trials with missing data *(Continued)*

Pearl 1956	No usable data reported	Not possible to generate data, and could not get in contact with authors

Table 3. Key demographic characteristics of the included studies

Category	Study frequency	Study ID
Author correspondence		
Contacted	35	Abikoff 2004 ; Allen 1998 ; Allen 2006 ; Ayen 2004 ; Bornovalova 2008 ; Carlson 1993 ; Carter 2003 ; Crouch 1988 ; Foa 1991 ; Foa 2018 ; Freire 2007 ; Glogowska 2000 ; Goldstein 2000 ; Hippman 2016 ; Karst 2007 ; Kelley 2012 ; Kwan 2017 ; Lai 2004 ; Legrand 2016 ; Miranda 1997 ; Mitchell 2008 ; Pelham 1992 ; Poland 2013 ; Powers 2004 ; Powers 2008a ; Powers 2008b ; Proudfoot 2013 ; Rabkin 1990 ; Rapee 2006 ; Rapee 2007 ; Rosa-Alcatraz 2009 ; Shalev 2012 ; Szymanski 1995 ; Teri 1997 ; Wolitzky 2009
Authors responded	16	Abikoff 2004 ; Abikoff 2004 (pers comm) ; Allen 1998 ; Allen 1998 (pers comm) ; Allen 2006 ; Allen 2006 (pers comm) ; Ayen 2004 ; Ayen 2004 (pers comm) ; Bornovalova 2008 ; Bornovalova 2008 (pers comm) ; Karst 2007 ; Karst 2007 (pers comm) ; Kelley 2012 ; Kelley 2012 (pers comm) ; Kwan 2017 ; Kwan 2017 (pers comm) ; Lai 2004 ; Lai 2004 (pers comm) ; Legrand 2016 ; Legrand 2016 (pers comm) ; Miranda 1997 ; Miranda 1997 (pers comm) ; Proudfoot 2013 ; Proudfoot 2013 (pers comm) ; Rapee 2006 ; Rapee 2006 (pers comm) ; Rapee 2007 ; Rapee 2007 (pers comm) ; Szymanski 1995 ; Szymanski 1995 (pers comm) ; Teri 1997 ; Teri 1997 (pers comm)
Setting		
Studies with outpatient settings	74	Abikoff 2004 ; Allen 1998 ; Allen 2006 ; Alvarez 1997 ; Ascher 1979 ; Ayen 2004 ; Berg 1983 ; Borden 1986 ; Borkovec 1975 ; Borkovec 1976 ; Brill 1964a ; Brill 1964b ; Carlson 1993 ; Crouch 1988 ; Ehlers 2014 ; Espie 1989a ; Etringer 1982 ; Foa 1991 ; Foa 2018 ; Freire 2007 ; Fuchs 1977 ; Glogowska 2000 ; Goldstein 2000 ; Hekmat 1984 ; Howlin 2007 ; Karst 2007 ; Kelley 2012 ; Kennedy 1974 ; Kilmann 1987 ; Klein 1977 ; Klerman 1974a ; Klerman 1974b ; Klosko 1990 ; Krapfl 1970 ; Kwan 2017 ; Lang 1965 ; Lick 1975 ; Lick 1977 ; Liddle 1990 ; Matson 1980 ; Mealiea 1971 ; Milby 1980 ; Miranda 1997 ; Mitchell 2008 ; Nandi 1976 ; Nicassio 1974 ; Pelham 1992 ; Pendleton 1983 ; Poland 2013 ; Powers 2004 ; Powers 2008a ; Powers 2008b ; Proudfoot 2013 ; Quayhagen 1995 ; Rabkin 1990 ; Rapee 2006 ; Rapee 2007 ; Robin 1976 ; Rosa-Alcatraz 2009 ; Rosen 1976 ; Roth 1964 ; Shalev 2012 ; Shealy 1979 ; Steinmark 1974 ; Szymanski 1995 ; Tan 1986 ; Teri 1997 ; Tori 1973 ; Trexler 1972 ; Turner 1979 ; Vanderplate 1983 ; Wilson 1980 ; Wolitzky 2009 ; Wollersheim 1991
Studies with inpatient settings	20	Bornovalova 2008 ; Bramston 1985 ; Carter 2003 ; Doty 1975 ; Double 1993 ; Goldwasser 1987 ; Lacy 1990 ; Lai 2004 ; Legrand 2016 ; McLachlan 1991 ; Peck 1976 ; Pearl 1956 ; Pillman 2001 ; Roehrich 1993 ; Rupert 1978 ; Sibilio 1957 ; Sommerness 1955 ; Watzl 1988 ; Weingaertner 1971 ; Whittaker 1963
Studies with both inpatient and outpatient settings	2	Crisp 1991 ; Hippman 2016
Sample size		

Table 3. Key demographic characteristics of the included studies *(Continued)*

Sample size above 100 participants	4	Foa 2018 ; Glogowska 2000 ; Proudfoot 2013 ; Rapee 2006
Gender		
Only females included	19	Allen 1998 ; Ayen 2004 ; Brill 1964a ; Brill 1964b ; Carter 2003 ; Crisp 1991 ; Ehlers 2014 ; Foa 1991 ; Fuchs 1977 ; Kennedy 1974 ; Klerman 1974a ; Klerman 1974b ; Krapfl 1970 ; Lick 1975 ; Mealiea 1971 ; Shealy 1979 ; Sibilio 1957 ; Vanderplate 1983 ; Watzl 1988
Only males included	14	Carlson 1993 ; Doty 1975 ; Kilmann 1987 ; Klein 1977 ; Pearl 1956 ; Pelham 1992 ; Pillman 2001 ; Poland 2013 ; Roehrich 1993 ; Roth 1964 ; Rupert 1978 ; Sommerness 1955 ; Weingaertner 1971 ; Whittaker 1963
Diagnostic classification		
DSM-II diagnosis	2	Klerman 1974a ; Klerman 1974b
DSM-III diagnosis	4	Klosko 1990 ; Liddle 1990 ; Rabkin 1990 ; Roehrich 1993
DSM-III-R diagnosis	8	Abikoff 2004 ; Carlson 1993 ; Crisp 1991 ; Foa 1991 ; Pelham 1992 ; Szymanski 1995 ; Teri 1997 ; Wollersheim 1991
DSM-IV diagnosis	15	Allen 1998 ; Allen 2006 ; Alvarez 1997 ; Ayen 2004 ; Carter 2003 ; Goldstein 2000 ; Hippman 2016 ; Kelley 2012 ; Lai 2004 ; Pillman 2001 ; Poland 2013 ; Rapee 2006 ; Rapee 2007 ; Shalev 2012 ; Wolitzky 2009
DSM-IV-TR diagnosis	3	Foa 1991 ; Kwan 2017 ; Rosa-Alcatraz 2009
ICD-9 diagnosis	1	Watzl 1988
The most commonly assessment instrument was Structured Clinical Interview for DSM (SCID; Spitzer 1989)	8	Allen 1998 ; Allen 2006 ; Ayen 2004 ; Ehlers 2014 ; Goldstein 2000 ; Hippman 2016 ; Kelley 2012 ; Poland 2013
Trials fulfilled the symptoms of a mental health disorder from the available diagnostic classifications system at the time of the trial, but did not report classification system	24	Ascher 1979 ; Berg 1983 ; Borden 1986 ; Bramston 1985 ; Brill 1964a ; Brill 1964b ; Double 1993 ; Espie 1989a ; Goldwasser 1987 ; Howlin 2007 ; Kilmann 1987 ; Lick 1977 ; Matson 1980 ; McLachlan 1991 ; Quayhagen 1995 ; Robin 1976 ; Rosen 1976 ; Roth 1964 ; Shealy 1979 ; Steinmark 1974 ; Turner 1979 ; Vanderplate 1983 ; Weingaertner 1971 ; Whittaker 1963
Trials reported a population classified as having a mental health disorder, but full diagnostic criteria was not reported	35	Borkovec 1975 ; Borkovec 1976 ; Bornovalova 2008 ; Crouch 1988 ; Doty 1975 ; Etringer 1982 ; Freire 2007 ; Fuchs 1977 ; Glogowska 2000 ; Hekmat 1984 ; Karst 2007 ; Kennedy 1974 ; Klein 1977 ; Krapfl 1970 ; Lacy 1990 ; Lang 1965 ; Lick 1975 ; Mealiea 1971 ; Milby 1980 ; Miranda 1997 ; Mitchell 2008 ; Nandi 1976 ; Nicassio 1974 ; Pendleton 1983 ; Powers 2004 ; Powers 2008a ; Powers 2008b ; Proudfoot 2013 ; Rupert 1978 ; Sibilio 1957 ; Sommerness 1955 ; Tan 1986 ; Tori 1973 ; Trexler 1972 ; Wilson 1980
Diagnoses		

Table 3. Key demographic characteristics of the included studies *(Continued)*

Attention Deficit Hyperactivity Disorder (ADHD) or Attention Deficit Disorder (ADD)	5	Abikoff 2004 ; Borden 1986 ; Carlson 1993 ; Klein 1977 ; Pelham 1992
Anorexia	1	Crisp 1991
Anxiety disorders of different kind (e.g. specific anxiety, social anxiety, panic disorder)	25	Etringer 1982 ; Goldstein 2000 ; Hekmat 1984 ; Karst 2007 ; Kennedy 1974 ; Klosko 1990 ; Krapfl 1970 ; Lang 1965 ; Lick 1975 ; Mealiea 1971 ; Peck 1976 ; Pendleton 1983 ; Powers 2004 ; Powers 2008a ; Powers 2008b ; Proudfoot 2013 ; Rapee 2006 ; Rapee 2007 ; Rosa-Alcatraz 2009 ; Rosen 1976 ; Rupert 1978 ; Szymanski 1995 ; Tori 1973 ; Trexler 1972 ; Wolitzky 2009
Autism spectrum disorder	1	Howlin 2007
Bulimia	1	Carter 2003
Depression	16	Allen 1998 ; Allen 2006 ; Ayen 2004 ; Fuchs 1977 ; Kelley 2012 ; Klerman 1974a ; Klerman 1974b ; Legrand 2016 ; Liddle 1990 ; Nandi 1976 ; Poland 2013 ; Proudfoot 2013 ; Rabkin 1990 ; Tan 1986 ; Teri 1997 ; Wollersheim 1991
Encopresis	1	Berg 1983
Erectile dysfunction	1	Kilmann 1987
Intellectual disability	2	Bramston 1985 ; Matson 1980
Learning disability	3	Glogowska 2000 ; Miranda 1997 ; Mitchell 2008)
Neurodegenerative disease	6	Goldwasser 1987 ; Kwan 2017 ; Lai 2004 ; McLachlan 1991 ; Quayhagen 1995 ; Teri 1997
Post-Traumatic Stress-Disorder (PTSD)	4	Ehlers 2014 ; Foa 1991 ; Foa 2018 ; Shalev 2012
Schizophrenia	5	Hippman 2016 ; Pearl 1956 ; Sibilio 1957 ; Weingaertner 1971 ; Whittaker 1963
Sleep-wake disorders (e.g. insomnia, sleep disturbance)	11	Ascher 1979 ; Borkovec 1975 ; Borkovec 1976 ; Espie 1989a ; Freire 2007 ; Lick 1977 ; Nicassio 1974 ; Shealy 1979 ; Steinmark 1974 ; Turner 1979
Substance use disorders of different kind (e.g. alcohol, cocaine)	8	Alvarez 1997 ; Bornovalova 2008 ; Crouch 1988 ; Milby 1980 ; Pillman 2001 ; Roehrich 1993 ; Watzl 1988 ; Wilson 1980
Other unspecified disorders	8	Brill 1964a ; Brill 1964b ; Doty 1975 ; Double 1993 ; Lacy 1990 ; Robin 1976 ; Roth 1964 ; Sommerness 1955
Duration of interventions		
Less than three months	74	Allen 1998 ; Allen 2006 ; Alvarez 1997 ; Ascher 1979 ; Borkovec 1975 ; Borkovec 1976 ; Bornovalova 2008 ; Bramston 1985 ; Carlson 1993 ; Carter 2003 ; Crisp 1991 ; Crouch 1988 ; Doty 1975 ; Double 1993 ; Espie 1989a ; Etringer 1982 ; Foa 1991 ; Foa 2018 ; Freire 2007 ; Fuchs 1977 ; Goldstein 2000 ; Goldwasser 1987 ; Hekmat 1984 ; Howlin 2007 ; Karst 2007 ; Kelley 2012 ; Kilmann 1987 ; Klein 1977 ; Krapfl 1970 ; Kwan 2017 ; Lai 2004 ; Lang 1965 ; Legrand 2016 ; Lick 1975

Table 3. Key demographic characteristics of the included studies *(Continued)*

		; Lick 1977 ; Liddle 1990 ; Matson 1980 ; Mealiea 1971 ; Milby 1980 ; Miranda 1997 ; Mitchell 2008 ; Nandi 1976 ; Nicassio 1974 ; Pearl 1956 ; Peck 1976 ; Pelham 1992 ; Pendleton 1983 ; Pillman 2001 ; Poland 2013 ; Powers 2004 ; Powers 2008a ; Powers 2008b ; Proudfoot 2013 ; Rabkin 1990 ; Robin 1976 ; Roehrich 1993 ; Rosen 1976 ; Roth 1964 ; Rupert 1978 ; Shealy 1979 ; Sibilio 1957 ; Steinmark 1974 ; Szymanski 1995 ; Tan 1986 ; Teri 1997 ; Tori 1973 ; Trexler 1972 ; Turner 1979 ; Vanderplate 1983 ; Weingaertner 1971 ; Whittaker 1963 ; Wilson 1980 ; Wolitzky 2009 ; Wollersheim 1991
More than three months	21	Abikoff 2004 ; Ayen 2004 ; Berg 1983 ; Borden 1986 ; Brill 1964a ; Brill 1964b ; Ehlers 2014 ; Glogowska 2000 ; Kennedy 1974 ; Klerman 1974a ; Klerman 1974b ; Klosko 1990 ; Lacy 1990 ; McLachlan 1991 ; Quayhagen 1995 ; Rapee 2006 ; Rapee 2007 ; Rosa-Alcatraz 2009 ; Shalev 2012 ; Sommerness 1955 ; Watzl 1988
Not reported	1	Hippman 2016
Formats of interventions		
Individual treatment	25	Ascher 1979 ; Bornovalova 2008 ; Brill 1964a ; Carter 2003 ; Ehlers 2014 ; Espie 1989a ; Foa 1991 ; Foa 2018 ; Glogowska 2000 ; Goldstein 2000 ; Hekmat 1984 ; Hippman 2016 ; Kennedy 1974 ; Lai 2004 ; Lang 1965 ; Lick 1975 ; Matson 1980 ; Powers 2004 ; Powers 2008a ; Proudfoot 2013 ; Quayhagen 1995 ; Roehrich 1993 ; Rosen 1976 ; Teri 1997 ; Turner 1979
Group treatment	18	Alvarez 1997 ; Ayen 2004 ; Bramston 1985 ; Doty 1975 ; Fuchs 1977 ; Goldwasser 1987 ; Kilmann 1987 ; Liddle 1990 ; Miranda 1997 ; Pendleton 1983 ; Rapee 2006 ; Rapee 2007 ; Rosa-Alcatraz 2009 ; Steinmark 1974 ; Szymanski 1995 ; Tan 1986 ; Trexler 1972 ; Wollersheim 1991
Combined treatment	3	Abikoff 2004 ; Borkovec 1975 ; Borkovec 1976
Types of interventions		
Types of psychological placebos		
Interactional component in psychological placebo	19	Abikoff 2004 ; Ayen 2004 ; Bramston 1985 ; Doty 1975 ; Ehlers 2014 ; Foa 1991 ; Foa 2018 ; Fuchs 1977 ; Goldwasser 1987 ; Hekmat 1984 ; Kilmann 1987 ; Klein 1977 ; Lai 2004 ; Lang 1965 ; Proudfoot 2013 ; Quayhagen 1995 ; Rosa-Alcatraz 2009 ; Shealy 1979 ; Tan 1986 ; Wollersheim 1991
Psychoeducational component in psychological placebo	9	Alvarez 1997 ; Bornovalova 2008 ; Carter 2003 ; Hippman 2016 ; Liddle 1990 ; Miranda 1997 ; Roehrich 1993 ; Rosen 1976 ; Szymanski 1995
Exposure component in psychological placebo	16	Ascher 1979 ; Borkovec 1975 ; Borkovec 1976 ; Espie 1989a ; Etringer 1982 ; Goldstein 2000 ; Kennedy 1974 ; Krapfl 1970 ; Lick 1975 ; Mealiea 1971 ; Pendleton 1983 ; Powers 2004 ; Powers 2008a ; Steinmark 1974 ; Trexler 1972 ; Turner 1979
Labels for psychological placebo		
Attention placebo control	12	Abikoff 2004 ; Alvarez 1997 ; Bramston 1985 ; Goldstein 2000 ; Goldwasser 1987 ; Hekmat 1984 ; Kilmann 1987 ; Klein 1977 ; Liddle 1990 ; Peck 1976 ; Tan 1986 ; Trexler 1972

Table 3. Key demographic characteristics of the included studies *(Continued)*

Non-specific placebo counselling or treatment	11	Ayen 2004 ; Bornovalova 2008 ; Doty 1975 ; Foa 1991 ; Fuchs 1977 ; Lai 2004 ; Lang 1965 ; Mealiea 1971 ; Pendleton 1983 ; Rosen 1976 ; Wollersheim 1991
Non-specific educational placebo	4	Carter 2003 ; Hippman 2016 ; Roehrich 1993 ; Szymanski 1995
Quasi-desensitization placebo	7	Ascher 1979 ; Borkovec 1975 ; Borkovec 1976 ; Kennedy 1974 ; Krapfl 1970 ; Steinmark 1974 ; Turner 1979
Standarised psychological placebo (e.g. present-centred therapy or emotion-focused supportive therapy)	2	Ehlers 2014 ; Foa 2018
Credible placebos or imagery relief placebo	8	Etringer 1982 ; Lick 1975 ; Miranda 1997 ; Powers 2004 ; Powers 2008a ; Quayhagen 1995 ; Rosa-Alcatraz 2009 ; Shealy 1979
Labels for pharmacological placebos		
Placebo pill	16	Borden 1986 ; Brill 1964a ; Carlson 1993 ; Crouch 1988 ; Double 1993 ; Kelley 2012 ; Klosko 1990 ; McLachlan 1991 ; Nandi 1976 ; Pelham 1992 ; Rabkin 1990 ; Roth 1964 ; Shalev 2012 ; Sibilio 1957 ; Sommerness 1955 ; Whittaker 1963
Implants	1	Wilson 1980
Injection	1	Watzl 1988
Psychological treatment and pharmacological placebo	4	Berg 1983 ; Klerman 1974a ; Klerman 1974b ; Powers 2008b
Labels for physical placebos		
Acupuncture or acupressure	5	Allen 1998 ; Allen 2006 ; Freire 2007 ; Karst 2007 ; Kwan 2017
Electromagnetic stimulation	2	Lick 1975 ; Weingaertner 1971
Exercise and relaxation	3	Lacy 1990 ; Legrand 2016 ; Nicassio 1974
Technical device	6	Mitchell 2008 ; Pillman 2001 ; Rupert 1978 ; Tori 1973 ; Vanderplate 1983 ; Wolitzky 2009
Labels for usual care		
Standard treatment	3	Matson 1980 ; Rapee 2006 ; Rapee 2007
Community based therapy	1	Glogowska 2000
Typical care control	1	Teri 1997

Table 3. Key demographic characteristics of the included studies *(Continued)*

Outpatient psychotherapy	3	Brill 1964b ; Crisp 1991 ; Milby 1980
Other	1	Robin 1976
Concomitant treatment		
Allowed the participant to receive a concomitant psychotherapy	18	Borden 1986 ; Bornovalova 2008 ; Carlson 1993 ; Crouch 1988 ; Ehlers 2014 ; Kwan 2017 ; Lacy 1990 ; Lai 2004 ; McLachlan 1991 ; Pelham 1992 ; Rapee 2006 ; Rapee 2007 ; Roth 1964 ; Sommerness 1955 ; Tan 1986 ; Teri 1997 ; Watzl 1988 ; Whittaker 1963
Did not allow the participant to receive a concomitant psychotherapy	20	Allen 1998 ; Allen 2006 ; Borkovec 1976 ; Doty 1975 ; Fuchs 1977 ; Goldstein 2000 ; Kelley 2012 ; Klerman 1974a ; Klerman 1974b ; Klosko 1990 ; Krapfl 1970 ; Lang 1965 ; Legrand 2016 ; Milby 1980 ; Rosen 1976 ; Shealy 1979 ; Sibilio 1957 ; Steinmark 1974 ; Szymanski 1995 ; Vanderplate 1983
Allowed the participants to receive a concomitant pharmacotherapy	29	Abikoff 2004 ; Alvarez 1997 ; Ayen 2004 ; Bornovalova 2008 ; Bramston 1985 ; Carter 2003 ; Crouch 1988 ; Doty 1975 ; Double 1993 ; Ehlers 2014 ; Espie 1989a ; Foa 2018 ; Kilmann 1987 ; Klosko 1990 ; Kwan 2017 ; Lai 2004 ; Legrand 2016 ; Lick 1977 ; Peck 1976 ; Pelham 1992 ; Poland 2013 ; Proudfoot 2013 ; Rapee 2006 ; Rapee 2007 ; Rupert 1978 ; Sommerness 1955 ; Tan 1986 ; Teri 1997 ; Whittaker 1963
Did not allow the participant to receive a concomitant pharmacotherapy	19	Allen 1998 ; Allen 2006 ; Borkovec 1976 ; Crisp 1991 ; Freire 2007 ; Goldstein 2000 ; Kelley 2012 ; Lacy 1990 ; Nicassio 1974 ; Quayhagen 1995 ; Roehrich 1993 ; Rosen 1976 ; Roth 1964 ; Shealy 1979 ; Steinmark 1974 ; Szymanski 1995 ; Vanderplate 1983 ; Weingaertner 1971 ; Wollersheim 1991
Comparators		
Wait-list control interventions	39	Allen 1998 ; Allen 2006 ; Ascher 1979 ; Ayen 2004 ; Borkovec 1975 ; Borkovec 1976 ; Bramston 1985 ; Brill 1964a ; Brill 1964b ; Ehlers 2014 ; Espie 1989a ; Foa 1991 ; Foa 2018 ; Fuchs 1977 ; Glogowska 2000 ; Goldstein 2000 ; Kelley 2012 ; Kilmann 1987 ; Kwan 2017 ; Lacy 1990 ; Lick 1975 ; Milby 1980 ; Pendleton 1983 ; Powers 2004 ; Powers 2008a ; Proudfoot 2013 ; Quayhagen 1995 ; Rapee 2006 ; Rapee 2007 ; Robin 1976 ; Rosa-Alcatraz 2009 ; Rosen 1976 ; Shealy 1979 ; Steinmark 1974 ; Trexler 1972 ; Turner 1979 ; Vanderplate 1983 ; Wolitzky 2009 ; Wollersheim 1991
No-treatment control interventions	56	Abikoff 2004 ; Alvarez 1997 ; Berg 1983 ; Borden 1986 ; Bornovalova 2008 ; Carlson 1993 ; Carter 2003 ; Crisp 1991 ; Crouch 1988 ; Doty 1975 ; Double 1993 ; Etringer 1982 ; Freire 2007 ; Goldwasser 1987 ; Hekmat 1984 ; Hippman 2016 ; Karst 2007 ; Kennedy 1974 ; Klein 1977 ; Klerman 1974a ; Klerman 1974b ; Klosko 1990 ; Krapfl 1970 ; Lai 2004 ; Lang 1965 ; Legrand 2016 ; Lick 1977 ; Liddle 1990 ; Matson 1980 ; McLachlan 1991 ; Mealiea 1971 ; Miranda 1997 ; Mitchell 2008 ; Nandi 1976 ; Nicassio 1974 ; Pearl 1956 ; Peck 1976 ; Pelham 1992 ; Pillman 2001 ; Poland 2013 ; Powers 2008b ; Rabkin 1990 ; Roehrich 1993 ; Roth 1964 ; Rupert 1978 ; Shalev 2012 ; Sibilio 1957 ; Sommerness 1955 ; Szymanski 1995 ; Tan 1986 ; Teri 1997 ; Tori 1973 ; Watzl 1988 ; Weingaertner 1971 ; Whittaker 1963 ; Wilson 1980
Labelled wait-list control intervention but the description is a no-treatment control intervention	10	Ascher 1979 ; Borkovec 1975 ; Borkovec 1976 ; Bramston 1985 ; Espie 1989a ; Kilmann 1987 ; Lick 1975 ; Rosen 1976 ; Trexler 1972 ; Turner 1979

Table 3. Key demographic characteristics of the included studies *(Continued)*

Labelled no-treatment control intervention but the description is a wait-list control intervention	4	Carter 2003 ; Hekmat 1984 ; Hippman 2016 ; Teri 1997
Labelled as 'Minimal contact group' but description is a wait-list control intervention	1	Foa 2018
Labelled as 'Delayed treatment group' but description is a wait-list control intervention	3	Howlin 2007 ; Robin 1976 ; Wollersheim 1991
Add-on drug treatment to each group	1	Abikoff 2004
Add-on psychotherapeutic treatment on each group	4	Berg 1983 ; Klerman 1974a ; Klerman 1974b ; Powers 2008b
Outcomes (four or more trials)		
Behavioral Avoidance Test (BAT)	8	Etringer 1982 ; Kennedy 1974 ; Krapfl 1970 ; Lang 1965 ; Peck 1976 ; Rosen 1976 ; Szymanski 1995 ; Wolitzky 2009
Daily Sleep Questionnaire (DSQ)	6	Ascher 1979 ; Borkovec 1975 ; Espie 1989a ; Steinmark 1974 ; Trexler 1972 ; Vanderplate 1983
Beck Depression Inventory (BDI)	4	Ayen 2004 ; Fuchs 1977 ; Legrand 2016 ; Tan 1986
Adverse events		
Serious adverse events	11	Ehlers 2014 ; Foa 2018 ; Freire 2007 ; Karst 2007 ; Klosko 1990 ; Kwan 2017 ; Lacy 1990 ; Legrand 2016 ; McLachlan 1991 ; Nandi 1976 ; Wilson 1980
Non-serious adverse events	14	Ayen 2004 ; Ehlers 2014 ; Foa 2018 ; Freire 2007 ; Goldstein 2000 ; Karst 2007 ; Klosko 1990 ; Kwan 2017 ; Lacy 1990 ; Legrand 2016 ; McLachlan 1991 ; Nandi 1976 ; Wilson 1980 ; Wollersheim 1991
Type of adverse events		
Complaint list	1	Ayen 2004
Clinician-administered posttraumatic stress disorder (PTSD) Scale for DSM-5 (CAPS-5)	1	Ehlers 2014
Disulfiram-ethanol reaction (DER)	1	Wilson 1980

Table 3. Key demographic characteristics of the included studies *(Continued)*

Spontaneous reporting	11	Foa 2018 ; Freire 2007 ; Goldstein 2000 ; Karst 2007 ; Klosko 1990 ; Kwan 2017 ; Lacy 1990 ; Legrand 2016 ; McLachlan 1991 ; Nandi 1976 ; Wollersheim 1991

Table 4. Risk of bias in included studies

Category	Study frequency	Study ID
Random sequence generation		
Low risk of bias	27	Allen 1998 ; Allen 2006 ; Bornovalova 2008 ; Carter 2003 ; Doty 1975 ; Ehlers 2014 ; Foa 2018 ; Freire 2007 ; Hippman 2016 ; Howlin 2007 ; Karst 2007 ; Kelley 2012 ; Kilmann 1987 ; Kwan 2017 ; Legrand 2016 ; McLachlan 1991 ; Poland 2013 ; Powers 2008a ; Powers 2008b ; Proudfoot 2013 ; Rabkin 1990 ; Rapee 2006 ; Rapee 2007 ; Rosen 1976 ; Sommerness 1955 ; Szymanski 1995 ; Teri 1997
High risk of bias	4	Lai 2004 ; Pillman 2001 ; Robin 1976 ; Shalev 2012
Unclear risk of bias	65	Abikoff 2004 ; Alvarez 1997 ; Ascher 1979 ; Ayen 2004 ; Berg 1983 ; Borden 1986 ; Borkovec 1975 ; Borkovec 1976 ; Bramston 1985 ; Brill 1964a ; Brill 1964b ; Carlson 1993 ; Crisp 1991 ; Crouch 1988 ; Double 1993 ; Espie 1989a ; Etringer 1982 ; Foa 1991 ; Fuchs 1977 ; Glogowska 2000 ; Goldstein 2000 ; Goldwasser 1987 ; Hekmat 1984 ; Kennedy 1974 ; Klein 1977 ; Klerman 1974a ; Klerman 1974b ; Klosko 1990 ; Krapfl 1970 ; Lacy 1990 ; Lang 1965 ; Lick 1975 ; Lick 1977 ; Liddle 1990 ; Matson 1980 ; Mealiea 1971 ; Milby 1980 ; Miranda 1997 ; Mitchell 2008 ; Nandi 1976 ; Nicassio 1974 ; Peck 1976 ; Pearl 1956 ; Pelham 1992 ; Pendleton 1983 ; Powers 2004 ; Quayhagen 1995 ; Roehrich 1993 ; Rosa-Alcatraz 2009 ; Roth 1964 ; Rupert 1978 ; Shealy 1979 ; Sibilio 1957 ; Steinmark 1974 ; Tan 1986 ; Tori 1973 ; Trexler 1972 ; Turner 1979 ; Vanderplate 1983 ; Watzl 1988 ; Weingaertner 1971 ; Whittaker 1963 ; Wilson 1980 ; Wolitzky 2009 ; Wollersheim 1991
Allocation concealment		
Low risk of bias	18	Allen 1998 ; Allen 2006 ; Bornovalova 2008 ; Ehlers 2014 ; Freire 2007 ; Glogowska 2000 ; Karst 2007 ; Kelley 2012 ; Kwan 2017 ; Legrand 2016 ; Poland 2013 ; Proudfoot 2013 ; Rapee 2006 ; Roth 1964 ; Sibilio 1957 ; Sommerness 1955 ; Teri 1997 ; Whittaker 1963
Unclear risk of bias	78	Abikoff 2004 ; Alvarez 1997 ; Ascher 1979 ; Ayen 2004 ; Berg 1983 ; Borden 1986 ; Borkovec 1975 ; Borkovec 1976 ; Bramston 1985 ; Brill 1964a ; Brill 1964b ; Carlson 1993 ; Carter 2003 ; Crisp 1991 ; Crouch 1988 ; Doty 1975 ; Double 1993 ; Espie 1989a ; Etringer 1982 ; Foa 1991 ; Foa 2018 ; Fuchs 1977 ; Goldstein 2000 ; Goldwasser 1987 ; Hekmat 1984 ; Hippman 2016 ; Howlin 2007 ; Kennedy 1974 ; Kilmann 1987 ; Klein 1977 ; Klerman 1974a ; Klerman 1974b ; Klosko 1990 ; Krapfl 1970 ; Lacy 1990 ; Lai 2004 ; Lang 1965 ; Lick 1975 ; Lick 1977 ; Liddle 1990 ; Matson 1980 ; McLachlan 1991 ; Mealiea 1971 ; Milby 1980 ; Miranda 1997 ; Mitchell 2008 ; Nandi 1976 ; Nicassio 1974 ; Pearl 1956 ; Peck 1976 ; Pelham 1992 ; Pendleton 1983 ; Pillman 2001 ; Powers 2004 ; Powers 2008a ; Powers 2008b ; Quayhagen 1995 ; Rabkin 1990 ; Rapee 2007 ; Robin 1976 ; Roehrich 1993 ; Rosa-Alcatraz 2009 ; Rosen 1976 ; Rupert 1978 ; Shalev 2012 ; Shealy 1979 ; Steinmark 1974 ; Szymanski 1995 ; Tan 1986 ; Tori 1973 ; Trexler 1972 ; Turner 1979 ; Vanderplate 1983 ; Watzl 1988 ; Weingaertner 1971 ; Wilson 1980 ; Wolitzky 2009 ; Wollersheim 1991

Table 4. Risk of bias in included studies *(Continued)*

Blinding of outcome assessors		
Low risk of bias	45	Allen 1998 ; Allen 2006 ; Borden 1986 ; Borkovec 1976 ; Bramston 1985 ; Carlson 1993 ; Doty 1975 ; Double 1993 ; Ehlers 2014 ; Etringer 1982 ; Foa 1991 ; Foa 2018 ; Freire 2007 ; Glogowska 2000 ; Hekmat 1984 ; Karst 2007 ; Kelley 2012 ; Kennedy 1974 ; Klein 1977 ; Klosko 1990 ; Kwan 2017 ; Lai 2004 ; Lick 1975 ; Matson 1980 ; McLachlan 1991 ; Mealiea 1971 ; Nandi 1976 ; Pendleton 1983 ; Pillman 2001 ; Rabkin 1990 ; Rapee 2006 ; Rapee 2007 ; Robin 1976 ; Rosa-Alcatraz 2009 ; Rosen 1976 ; Roth 1964 ; Shalev 2012 ; Sibilio 1957 ; Sommerness 1955 ; Szymanski 1995 ; Teri 1997 ; Tori 1973 ; Trexler 1972 ; Weingaertner 1971 ; Wolitzky 2009
High risk of bias	32	Abikoff 2004 ; Alvarez 1997 ; Ascher 1979 ; Ayen 2004 ; Berg 1983 ; Borkovec 1975 ; Bornovalova 2008 ; Carter 2003 ; Crisp 1991 ; Crouch 1988 ; Espie 1989a ; Fuchs 1977 ; Goldstein 2000 ; Hippman 2016 ; Howlin 2007 ; Kilmann 1987 ; Legrand 2016 ; Lick 1977 ; Liddle 1990 ; Nicassio 1974 ; Pelham 1992 ; Powers 2004 ; Powers 2008a ; Powers 2008b ; Proudfoot 2013 ; Shealy 1979 ; Steinmark 1974 ; Tan 1986 ; Turner 1979 ; Vanderplate 1983 ; Whittaker 1963 ; Wilson 1980
Unclear risk of bias	19	Brill 1964a ; Brill 1964b ; Goldwasser 1987 ; Klerman 1974a ; Klerman 1974b ; Krapfl 1970 ; Lacy 1990 ; Lang 1965 ; Milby 1980 ; Miranda 1997 ; Mitchell 2008 ; Pearl 1956 ; Peck 1976 ; Poland 2013 ; Quayhagen 1995 ; Roehrich 1993 ; Rupert 1978 ; Watzl 1988 ; Wollersheim 1991
Blinding of participants and personnel		
High risk of bias	96	Abikoff 2004 ; Allen 1998 ; Allen 2006 ; Alvarez 1997 ; Ascher 1979 ; Ayen 2004 ; Berg 1983 ; Borden 1986 ; Borkovec 1975 ; Borkovec 1976 ; Bornovalova 2008 ; Bramston 1985 ; Brill 1964a ; Brill 1964b ; Carlson 1993 ; Carter 2003 ; Crisp 1991 ; Crouch 1988 ; Doty 1975 ; Double 1993 ; Ehlers 2014 ; Espie 1989a ; Etringer 1982 ; Foa 1991 ; Foa 2018 ; Freire 2007 ; Fuchs 1977 ; Glogowska 2000 ; Goldstein 2000 ; Goldwasser 1987 ; Hekmat 1984 ; Hippman 2016 ; Howlin 2007 ; Karst 2007 ; Kelley 2012 ; Kennedy 1974 ; Kilmann 1987 ; Klein 1977 ; Klerman 1974a ; Klerman 1974b ; Klosko 1990 ; Krapfl 1970 ; Kwan 2017 ; Lacy 1990 ; Lai 2004 ; Lang 1965 ; Legrand 2016 ; Lick 1975 ; Lick 1977 ; Liddle 1990 ; Matson 1980 ; McLachlan 1991 ; Mealiea 1971 ; Milby 1980 ; Miranda 1997 ; Mitchell 2008 ; Nandi 1976 ; Nicassio 1974 ; Pearl 1956 ; Peck 1976 ; Pelham 1992 ; Pendleton 1983 ; Pillman 2001 ; Poland 2013 ; Powers 2004 ; Powers 2008a ; Powers 2008b ; Proudfoot 2013 ; Quayhagen 1995 ; Rabkin 1990 ; Rapee 2006 ; Rapee 2007 ; Robin 1976 ; Roehrich 1993 ; Rosa-Alcatraz 2009 ; Rosen 1976 ; Roth 1964 ; Rupert 1978 ; Shalev 2012 ; Shealy 1979 ; Sibilio 1957 ; Sommerness 1955 ; Steinmark 1974 ; Szymanski 1995 ; Tan 1986 ; Teri 1997 ; Tori 1973 ; Trexler 1972 ; Turner 1979 ; Vanderplate 1983 ; Watzl 1988 ; Weingaertner 1971 ; Whittaker 1963 ; Wilson 1980 ; Wolitzky 2009 ; Wollersheim 1991
Incomplete outcome data		
Low risk of bias	41	Allen 1998 ; Allen 2006 ; Ayen 2004 ; Berg 1983 ; Borkovec 1976 ; Bramston 1985 ; Ehlers 2014 ; Glogowska 2000 ; Goldstein 2000 ; Hippman 2016 ; Howlin 2007 ; Karst 2007 ; Kilmann 1987 ; Klerman 1974a ; Klerman 1974b ; Kwan 2017 ; Lacy 1990 ; Lai 2004 ; Lang 1965 ; Legrand 2016 ; Lick 1977 ; Liddle 1990 ; Matson 1980 ; Mealiea 1971 ; Miranda 1997 ; Mitchell 2008 ; Nicassio 1974 ; Pelham 1992 ; Pendleton 1983 ; Powers 2004 ; Powers 2008a ; Powers 2008b ; Rabkin 1990 ; Robin 1976 ; Rosa-Alcatraz 2009 ; Shalev 2012 ; Steinmark 1974 ; Tan 1986 ; Turner 1979 ; Watzl 1988 ; Wolitzky 2009

Table 4. Risk of bias in included studies *(Continued)*

High risk of bias	29	Abikoff 2004 ; Alvarez 1997 ; Borkovec 1975 ; Bornovalova 2008 ; Brill 1964a ; Brill 1964b ; Crisp 1991 ; Crouch 1988 ; Doty 1975 ; Double 1993 ; Espie 1989a ; Foa 1991 ; Freire 2007 ; Fuchs 1977 ; Kennedy 1974 ; Klosko 1990 ; McLachlan 1991 ; Nandi 1976 ; Pillman 2001 ; Poland 2013 ; Quayhagen 1995 ; Rapee 2006 ; Roehrich 1993 ; Rosen 1976 ; Roth 1964 ; Szymanski 1995 ; Teri 1997 ; Vanderplate 1983 ; Wollersheim 1991
Unclear risk of bias	26	Ascher 1979 ; Borden 1986 ; Carlson 1993 ; Carter 2003 ; Etringer 1982 ; Foa 2018 ; Goldwasser 1987 ; Hekmat 1984 ; Kelley 2012 ; Klein 1977 ; Krapfl 1970 ; Lick 1975 ; Milby 1980 ; Peck 1976 ; Proudfoot 2013 ; Rapee 2007 ; Rupert 1978 ; Pearl 1956 ; Shealy 1979 ; Sibilio 1957 ; Sommerness 1955 ; Tori 1973 ; Trexler 1972 ; Weingaertner 1971 ; Whittaker 1963 ; Wilson 1980
Selective reporting		
Low risk of bias	6	Allen 2006 ; Ehlers 2014 ; Hippman 2016 ; Kwan 2017 ; Legrand 2016 ; Shalev 2012
High risk of bias	5	Abikoff 2004 ; Bornovalova 2008 ; Foa 2018 ; Kelley 2012 ; Proudfoot 2013
Unclear risk of bias	85	Allen 1998 ; Alvarez 1997 ; Ascher 1979 ; Ayen 2004 ; Berg 1983 ; Borden 1986 ; Borkovec 1975 ; Borkovec 1976 ; Bramston 1985 ; Brill 1964a ; Brill 1964b ; Carlson 1993 ; Carter 2003 ; Crisp 1991 ; Crouch 1988 ; Doty 1975 ; Double 1993 ; Espie 1989a ; Etringer 1982 ; Foa 1991 ; Freire 2007 ; Fuchs 1977 ; Glogowska 2000 ; Goldstein 2000 ; Goldwasser 1987 ; Hekmat 1984 ; Howlin 2007 ; Karst 2007 ; Kennedy 1974 ; Kilmann 1987 ; Klein 1977 ; Klerman 1974a ; Klerman 1974b ; Klosko 1990 ; Krapfl 1970 ; Lacy 1990 ; Lai 2004 ; Lang 1965 ; Lick 1975 ; Lick 1977 ; Liddle 1990 ; Matson 1980 ; McLachlan 1991 ; Mealiea 1971 ; Milby 1980 ; Miranda 1997 ; Mitchell 2008 ; Nandi 1976 ; Nicassio 1974 ; Pearl 1956 ; Peck 1976 ; Pelham 1992 ; Pendleton 1983 ; Pillman 2001 ; Poland 2013 ; Powers 2004 ; Powers 2008a ; Powers 2008b ; Quayhagen 1995 ; Rabkin 1990 ; Rapee 2006 ; Rapee 2007 ; Robin 1976 ; Roehrich 1993 ; Rosa-Alcatraz 2009 ; Rosen 1976 ; Roth 1964 ; Rupert 1978 ; Shealy 1979 ; Sibilio 1957 ; Sommerness 1955 ; Steinmark 1974 ; Szymanski 1995 ; Tan 1986 ; Teri 1997 ; Tori 1973 ; Trexler 1972 ; Turner 1979 ; Vanderplate 1983 ; Watzl 1988 ; Weingaertner 1971 ; Whittaker 1963 ; Wilson 1980 ; Wolitzky 2009 ; Wollersheim 1991
Other potential sources of bias		
Low risk of bias	87	Abikoff 2004 ; Allen 1998 ; Allen 2006 ; Alvarez 1997 ; Ascher 1979 ; Ayen 2004 ; Berg 1983 ; Borden 1986 ; Borkovec 1975 ; Borkovec 1976 ; Bornovalova 2008 ; Bramston 1985 ; Brill 1964a ; Brill 1964b ; Carlson 1993 ; Carter 2003 ; Crisp 1991 ; Crouch 1988 ; Doty 1975 ; Double 1993 ; Ehlers 2014 ; Espie 1989a ; Etringer 1982 ; Foa 1991 ; Foa 2018 ; Freire 2007 ; Fuchs 1977 ; Glogowska 2000 ; Goldstein 2000 ; Goldwasser 1987 ; Hekmat 1984 ; Hippman 2016 ; Howlin 2007 ; Karst 2007 ; Kelley 2012 ; Kennedy 1974 ; Kilmann 1987 ; Klein 1977 ; Klerman 1974a ; Klerman 1974b ; Klosko 1990 ; Krapfl 1970 ; Kwan 2017 ; Lacy 1990 ; Lang 1965 ; Legrand 2016 ; Lick 1975 ; Liddle 1990 ; Matson 1980 ; McLachlan 1991 ; Mealiea 1971 ; Milby 1980 ; Miranda 1997 ; Nandi 1976 ; Nicassio 1974 ; Pearl 1956 ; Peck 1976 ; Pendleton 1983 ; Pillman 2001 ; Poland 2013 ; Powers 2004 ; Powers 2008a ; Powers 2008b ; Proudfoot 2013 ; Rabkin 1990 ; Rapee 2006 ; Rapee 2007 ; Roehrich 1993 ; Rosa-Alcatraz 2009 ; Rosen 1976 ; Roth 1964 ; Rupert 1978 ; Shalev 2012 ; Shealy 1979 ; Sibilio 1957 ; Sommerness 1955 ; Szymanski 1995 ; Tan 1986 ; Teri 1997 ; Tori 1973 ; Vanderplate 1983 ; Watzl 1988 ; Weingaertner 1971 ; Whittaker 1963 ; Wilson 1980 ; Wolitzky 2009 ; Wollersheim 1991

Table 4. Risk of bias in included studies *(Continued)*

High risk of bias	8	Lick 1975 ; Mitchell 2008 ; Pelham 1992 ; Quayhagen 1995 ; Robin 1976 ; Steinmark 1974 ; Trexler 1972 ; Turner 1979
Unclear risk of bias	1	Lai 2004
Researchers or authors provided the treatment	4	Lick 1975 ; Steinmark 1974 ; Trexler 1972 ; Turner 1979
Attention bias or differences in duration of treatment	2	Mitchell 2008 ; Robin 1976
Carry-over effects	1	Pelham 1992
Exceed the passivity of placebo	1	Quayhagen 1995
Time bias or assessment at different point for the groups	1	Robin 1976
Affiliation bias (allegiance or industry bias)		
Low risk of bias	90	Allen 1998 ; Allen 2006 ; Alvarez 1997 ; Ascher 1979 ; Ayen 2004 ; Berg 1983 ; Borden 1986 ; Borkovec 1975 ; Borkovec 1976 ; Bramston 1985 ; Brill 1964a ; Brill 1964b ; Carlson 1993 ; Carter 2003 ; Crisp 1991 ; Crouch 1988 ; Doty 1975 ; Double 1993 ; Espie 1989a ; Etringer 1982 ; Freire 2007 ; Glogowska 2000 ; Goldstein 2000 ; Goldwasser 1987 ; Hekmat 1984 ; Hippman 2016 ; Howlin 2007 ; Karst 2007 ; Kelley 2012 ; Kennedy 1974 ; Kilmann 1987 ; Klein 1977 ; Klerman 1974a ; Klerman 1974b ; Klosko 1990 ; Krapfl 1970 ; Kwan 2017 ; Lacy 1990 ; Lai 2004 ; Lang 1965 ; Legrand 2016 ; Lick 1975 ; Lick 1977 ; Liddle 1990 ; Matson 1980 ; McLachlan 1991 ; Mealiea 1971 ; Milby 1980 ; Miranda 1997 ; Mitchell 2008 ; Nandi 1976 ; Nicassio 1974 ; Pearl 1956 ; Peck 1976 ; Pelham 1992 ; Pendleton 1983 ; Pillman 2001 ; Poland 2013 ; Powers 2004 ; Powers 2008a ; Powers 2008b ; Proudfoot 2013 ; Quayhagen 1995 ; Rabkin 1990 ; Rapee 2006 ; Rapee 2007 ; Robin 1976 ; Roehrich 1993 ; Rosa-Alcatraz 2009 ; Rosen 1976 ; Roth 1964 ; Rupert 1978 ; Shalev 2012 ; Shealy 1979 ; Sibilio 1957 ; Sommerness 1955 ; Steinmark 1974 ; Szymanski 1995 ; Tan 1986 ; Teri 1997 ; Tori 1973 ; Trexler 1972 ; Turner 1979 ; Vanderplate 1983 ; Watzl 1988 ; Weingaertner 1971 ; Whittaker 1963 ; Wilson 1980 ; Wolitzky 2009 ; Wollersheim 1991
High risk of bias	6	Abikoff 2004 ; Bornovalova 2008 ; Ehlers 2014 ; Foa 1991 ; Foa 2018 ; Fuchs 1977

Table 5. Effects of methods

Category	Study frequency	Study ID
Beneficial outcome data		
Dichotomous data	11	Berg 1983 ; Double 1993 ; Klerman 1974a ; Klerman 1974b ; Mealiea 1971 ; Milby 1980 ; Rabkin 1990 ; Robin 1976 ; Watzl 1988 ; Whittaker 1963 ; Wollersheim 1991

Table 5. Effects of methods *(Continued)*

Continuous data	71	Abikoff 2004 ; Allen 1998 ; Allen 2006 ; Alvarez 1997 ; Ascher 1979 ; Ayen 2004 ; Borden 1986 ; Borkovec 1975 ; Borkovec 1976 ; Bornovalova 2008 ; Bramston 1985 ; Carlson 1993 ; Carter 2003 ; Crisp 1991 ; Ehlers 2014 ; Espie 1989a ; Etringer 1982 ; Foa 1991 ; Foa 2018 ; Freire 2007 ; Fuchs 1977 ; Glogowska 2000 ; Goldstein 2000 ; Hippman 2016 ; Karst 2007 ; Kelley 2012 ; Kennedy 1974 ; Kilmann 1987 ; Klein 1977 ; Klosko 1990 ; Krapfl 1970 ; Kwan 2017 ; Lacy 1990 ; Lai 2004 ; Legrand 2016 ; Lick 1975 ; Lick 1977 ; Liddle 1990 ; Matson 1980 ; McLachlan 1991 ; Miranda 1997 ; Mitchell 2008 ; Nandi 1976 ; Nicassio 1974 ; Peck 1976 ; Pelham 1992 ; Pendleton 1983 ; Pillman 2001 ; Poland 2013 ; Powers 2004 ; Powers 2008a ; Powers 2008b ; Proudfoot 2013 ; Quayhagen 1995 ; Rapee 2006 ; Rapee 2007 ; Roehrich 1993 ; Rosa-Alcatraz 2009 ; Rosen 1976 ; Shalev 2012 ; Shealy 1979 ; Steinmark 1974 ; Szymanski 1995 ; Tan 1986 ; Teri 1997 ; Tori 1973 ; Turner 1979 ; Vanderplate 1983 ; Weingaertner 1971 ; Wilson 1980 ; Wolitzky 2009
Did not report usable data	14(13)	Brill 1964a ; Brill 1964b ; Crouch 1988 ; Doty 1975 ; Goldwasser 1987 ; Hekmat 1984 ; Lang 1965 ; Pearl 1956 ; Roth 1964 ; Rupert 1978 ; Sibilio 1957 ; Sommerness 1955 ; Trexler 1972 ; Matson 1980
Efficacy of all placebos		
Reported dichotomous data on a outcome for efficacy of placebo	9	Berg 1983 ; Double 1993 ; Klerman 1974a ; Klerman 1974b ; Mealiea 1971 ; Rabkin 1990 ; Watzl 1988 ; Whittaker 1963 ; Wilson 1980
Reported continuous data on a outcome for efficacy of placebo	65	Abikoff 2004 ; Allen 1998 ; Allen 2006 ; Alvarez 1997 ; Ascher 1979 ; Ayen 2004 ; Borden 1986 ; Borkovec 1975 ; Borkovec 1976 ; Bornovalova 2008 ; Bramston 1985 ; Carlson 1993 ; Carter 2003 ; Crouch 1988 ; Ehlers 2014 ; Espie 1989a ; Etringer 1982 ; Foa 1991 ; Foa 2018 ; Freire 2007 ; Fuchs 1977 ; Glogowska 2000 ; Hippman 2016 ; Karst 2007 ; Kelley 2012 ; Kennedy 1974 ; Kilmann 1987 ; Klein 1977 ; Klosko 1990 ; Krapfl 1970 ; Kwan 2017 ; Lacy 1990 ; Lai 2004 ; Legrand 2016 ; Lick 1975 ; Lick 1977 ; Liddle 1990 ; McLachlan 1991 ; Miranda 1997 ; Mitchell 2008 ; Nandi 1976 ; Nicassio 1974 ; Peck 1976 ; Pelham 1992 ; Pendleton 1983 ; Pillman 2001 ; Poland 2013 ; Powers 2004 ; Powers 2008a ; Powers 2008b ; Proudfoot 2013 ; Quayhagen 1995 ; Rosa-Alcatraz 2009 ; Rosen 1976 ; Shalev 2012 ; Shealy 1979 ; Steinmark 1974 ; Szymanski 1995 ; Tan 1986 ; Tori 1973 ; Turner 1979 ; Vanderplate 1983 ; Weingaertner 1971 ; Wolitzky 2009 ; Wollersheim 1991
Did not report usable data	12	Brill 1964a ; Crouch 1988 ; Doty 1975 ; Goldwasser 1987 ; Hekmat 1984 ; Lang 1965 ; Roth 1964 ; Rupert 1978 ; Pearl 1956 ; Sibilio 1957 ; Sommerness 1955 ; Trexler 1972 .
Efficacy of psychological placebos		
Reported dichotomous data on a outcome for efficacy of placebo	1	Mealiea 1971
Reported continuous data on a outcome for efficacy of placebo	38	Abikoff 2004 ; Alvarez 1997 ; Ascher 1979 ; Ayen 2004 ; Borkovec 1975 ; Borkovec 1976 ; Bornovalova 2008 ; Bramston 1985 ; Carter 2003 ; Ehlers 2014 ; Espie 1989a ; Etringer 1982 ; Foa 1991 ; Foa 2018 ; Fuchs 1977 ; Goldstein 2000 ; Hippman 2016 ; Kennedy 1974 ; Kilmann 1987 ; Klein 1977 ; Krapfl 1970 ; Lai 2004 ; Lick 1975 ; Liddle 1990 ; Miranda 1997 ; Pendleton 1983 ; Powers 2004 ; Powers 2008a ; Quayhagen 1995 ; Roehrich 1993 ; Rosa-Alcatraz 2009 ; Rosen 1976 ; Shealy 1979 ; Stein-

Table 5. Effects of methods *(Continued)*

		mark 1974 ; Szymanski 1995 ; Tan 1986 ; Turner 1979 ; Wollersheim 1991
Did not report usable data	5	Doty 1975 ; Goldwasser 1987 ; Hekmat 1984 ; Lang 1965 ; Trexler 1972
Efficacy of pharmacological placebos		
Reported dichotomous data on a outcome for efficacy of placebo	8	Berg 1983 ; Double 1993 ; Klerman 1974a ; Klerman 1974b ; Rabkin 1990 ; Watzl 1988 ; Whittaker 1963 ; Wilson 1980
Reported continuous data on a outcome for efficacy of placebo	9	Borden 1986 ; Carlson 1993 ; Kelley 2012 ; Klosko 1990 ; McLachlan 1991 ; Nandi 1976 ; Pelham 1992 ; Powers 2008b ; Shalev 2012
Did not report usable data	6	Brill 1964a ; Crouch 1988 ; Roth 1964 ; Pearl 1956 ; Sibilio 1957 ; Sommerness 1955
Efficacy of physical placebos		
Reported dichotomous data on a outcome for efficacy of placebo	0	
Reported continuous data on a outcome for efficacy of placebo	16	Allen 1998 ; Allen 2006 ; Freire 2007 ; Karst 2007 ; Kwan 2017 ; Lacy 1990 ; Legrand 2016 ; Lick 1977 ; Mitchell 2008 ; Nicassio 1974 ; Pillman 2001 ; Poland 2013 ; Tori 1973 ; Vanderplate 1983 ; Weingaertner 1971 ; Wolitzky 2009
Did not report usable data	1	Rupert 1978
Efficacy of Usual care		
Reported dichotomous data on a outcome for efficacy of usual care	2	Milby 1980 ; Robin 1976
Reported continuous data on a outcome for efficacy of usual care	5	Crisp 1991 ; Glogowska 2000 ; Rapee 2006 ; Rapee 2007 ; Teri 1997
Did not report usable data	2	Brill 1964b ; Matson 1980
Reported serious adverse events		
All placebos	11	Ehlers 2014 ; Foa 2018 ; Freire 2007 ; Karst 2007 ; Klosko 1990 ; Kwan 2017 ; Lacy 1990 ; Legrand 2016 ; McLachlan 1991 ; Nandi 1976 ; Wilson 1980
Psychological placebos	2	Ehlers 2014 ; Foa 2018
Pharmacologica placebos	4	Klosko 1990 ; McLachlan 1991 ; Nandi 1976 ; Wilson 1980
Physical placebos	5	Freire 2007 ; Karst 2007 ; Kwan 2017 ; Lacy 1990 ; Legrand 2016
Usual care	0	

Table 5. Effects of methods *(Continued)*

Reported data on specific mental health diagnoses for all placebos (three or more trials)		
Attention Deficit Hyperactivity Disorder (ADHD) or Attention Deficit Disorder (ADD)	5	Abikoff 2004 ; Borden 1986 ; Carlson 1993 ; Klein 1977 ; Pelham 1992
Anxiety disorders of different kind (e.g. specific anxiety, social anxiety, panic disorder)	16	Etringer 1982 ; Goldstein 2000 ; Karst 2007 ; Klosko 1990 ; Krapfl 1970 ; Lick 1975 ; Miranda 1997 ; Pendleton 1983 ; Powers 2004 ; Powers 2008a ; Powers 2008b ; Rosa-Alcatraz 2009 ; Rosen 1976 ; Szymanski 1995 ; Tori 1973 ; Wolitzky 2009
Depression	10	Allen 1998 ; Allen 2006 ; Ayen 2004 ; Fuchs 1977 ; Kelley 2012 ; Legrand 2016 ; Liddle 1990 ; Nandi 1976 ; Poland 2013 ; Wollersheim 1991
Neurodegenerative disease	4	Kwan 2017 ; Lai 2004 ; Mitchell 2008 ; Quayhagen 1995
Post-Traumatic Stress-Disorder (PTSD)	4	Ehlers 2014 ; Foa 1991 ; Foa 2018 ; Shalev 2012
Sleep-wake disorders (e.g. insomnia, sleep disturbance)	11	Ascher 1979 ; Borkovec 1975 ; Borkovec 1976 ; Espie 1989a ; Freire 2007 ; Lick 1977 ; Nicassio 1974 ; Shealy 1979 ; Steinmark 1974 ; Turner 1979 ; Vanderplate 1983
Substance use disorders of different kind (e.g. alcohol, cocaine)	4	Alvarez 1997 ; Bornovalova 2008 ; Pillman 2001 ; Roehrich 1993
Reported data on specific mental health diagnoses for psychological placebos (three or more trials)		
Anxiety disorders of different kind (e.g. specific anxiety, social anxiety, panic disorder)	11	Etringer 1982 ; Goldstein 2000 ; Krapfl 1970 ; Lick 1975 ; Miranda 1997 ; Pendleton 1983 ; Powers 2004 ; Powers 2008a ; Rosa-Alcatraz 2009 ; Rosen 1976 ; Szymanski 1995
Depression	4	Ayen 2004 ; Fuchs 1977 ; Liddle 1990 ; Wollersheim 1991
Post-Traumatic Stress-Disorder (PTSD)	3	Ehlers 2014 ; Foa 1991 ; Foa 2018
Sleep-wake disorders (e.g. insomnia, sleep disturbance)	7	Ascher 1979 ; Borkovec 1975 ; Borkovec 1976 ; Espie 1989a ; Shealy 1979 ; Steinmark 1974 ; Turner 1979
Substance use disorders of different kind (e.g. alcohol, cocaine)	3	Alvarez 1997 ; Bornovalova 2008 ; Roehrich 1993
Reported data on specific mental health diagnoses for pharmacological placebos (three or more trials)		

Table 5. Effects of methods *(Continued)*

Attention Deficit Hyperactivity Disorder (ADHD) or Attention Deficit Disorder (ADD)	3	Borden 1986 ; Carlson 1993 ; Pelham 1992
Reported data on specific mental health diagnoses for physical placebos (three or more trials)		
Anxiety disorders of different kind (e.g. specific anxiety, social anxiety, panic disorder)	3	Karst 2007 ; Tori 1973 ; Wolitzky 2009
Depression	4	Allen 1998 ; Allen 2006 ; Legrand 2016 ; Poland 2013
Sleep-wake disorders (e.g. insomnia, sleep disturbance)	4	Freire 2007 ; Lick 1977 ; Nicassio 1974 ; Vanderplate 1983
Reported non-serious adverse events		
All placebos	14	Ayen 2004 ; Ehlers 2014 ; Foa 2018 ; Freire 2007 ; Goldstein 2000 ; Karst 2007 ; Klosko 1990 ; Kwan 2017 ; Lacy 1990 ; Legrand 2016 ; McLachlan 1991 ; Nandi 1976 ; Wilson 1980 ; Wollersheim 1991
Psychological placebos	5	Ayen 2004 ; Ehlers 2014 ; Foa 2018 ; Goldstein 2000 ; Wollersheim 1991
Pharmacologica placebos	5	Klosko 1990 ; McLachlan 1991 ; Nandi 1976 ; Wilson 1980
Physical placebos	5	Freire 2007 ; Karst 2007 ; Kwan 2017 ; Lacy 1990 ; Legrand 2016
Usual care	0	

Table 6. Differences from protocol and review

Section	Protocol	Review
Pooling of placebos	We planned to pool psychological, pharmacological, and physical placebos into one placebo for the first primary outcome.	We pooled psychological, pharmacological, and physical placebos into a group called 'all placebos' for all outcomes. This was done to increase the power of the analysis regarding adverse events and specific mental health diagnosis.
Dealing with missing data	We planned to contact study authors for relevant missing data on our prImary and secondary outcomes. If the authors dId not respond after two attempts to contact them, we planned to stop communications. If we are not able to obtain missing data, we will use the available data (incomplete data) in the analyses. If data are not reported in a usable way, we will consult a statistician to explore its transformation.	We contacted study authors for relevant missing data. However, we made a pragmatic decision not to contact authors from studies before 1990. This was due to lack of probability that these data would have been preserved. In trials that did not report standard deviation (SD), we created an artificially SD from trials with the same population, outcome, experimental and control intervention.

Table 6. Differences from protocol and review *(Continued)*

		Furthermore, if trials did not report the amount of participants included in each group, and the total amount of participant could be equally divided, we expected the groups to have the same amount of participants in each group. However, if the number could not be equally divided, we anticipated that the active interventions and active control interventions included more.
Subgroup analysis and investigation of heterogeneity	We planned to conduct subgroup analyses to make hypotheses about the subgroups mentioned below. 1. Type of active intervention: i) psychological intervention, ii) pharmacological intervention, iii) physical intervention, or iv) other intervention. 2. Overall risk of bias: i) high risk of bias or ii) low risk of bias. 3. Type of outcome domain: i) blinded observer-reported, ii) non-blinded observer-reported, or iii) patient-reported. 4. Type of comparator intervention: i) wait-list or ii) no-treatment. 5. Awareness of placebo intervention: i) participants were aware that they might receive a placebo, or ii) participants were not aware of such. 6. The trial objective: i) a trial's objective is clearly to assess the effects of placebo, usual care, or wait-list interventions, or ii) no such objectives are stated. 7. Mean age of participants: i) < 18 years, ii) 18 to 50 years, or iii) > 50 years. 8. Duration of intervention: i) three months or above or ii) below three months. 9. Type of usual care: i) pharmacological, ii) psychological, iii) physical, or iv) other. 10. Standardised usual care: i) the usual care intervention was intentionally standardised or manualised, or ii) no standardisation or manualisation. 11. Mode of psychological treatment in usual care and psychological placebo: i) individual psychological treatment, or ii) group psychological treatment.	We added four post-hoc subgroup analyses: 1. Mental health diagnoses: i) formal diagnosis according to DSM/ICD, ii) fulfil symptoms of disorder ICD/DSM while not stating classifications systems, and iii) population is classified as having a mental disorder, but full diagnostic criteria not reported. 2. Type of psychological placebo: i) interaction placebo, ii) educational placebo, and iii) exposure placebo. 3. Type of physical placebo: i) acupuncture or acupressure placebo, ii) exercise and relaxation placebo, iii) technical device placebo, and iv) electromagnetic stimulation placebo. 4. Affiliation bias: i) Risk of affiliation, industry, allegiance bias, and ii) no risk of affiliation, industry, allegiance bias (Leichsenring 2019)
Sensitivity analysis	Studies contributing to heterogeneity ('outliers') was planned to be removed to evaluate the impact of their statistical heterogeneity on the overall pooled effect estimate. We will remove outliers one by one and assess the impact on the overall outcome. We will conduct sensitivity analyses to determine whether findings are sensitive to the following decisions made during the review process. 1. Our assessment of the level of clinical heterogeneity. 2. Analytical technique (e.g. fixed-effect and random-effects models). 3. Type of data collection (e.g. different ways to measure adverse events). 4. Imputed data (comparing the analyses with available outcome data with those following the ITT principle). 5. Combination of data in continuous outcomes (end of intervention or change scores). 6. Use of cluster-randomised trials.	We were only able to perform the analyses 2, 3, 4, and 5 due to a lack of sufficient data. Furthermore, we included a test for imprecision, as assessed by GRADE, by conducting Trial Sequential Analysis (TSA) on our primary outcomes included in the Summary of findings 3 . We also included these two sensitivity analyses: 1. The impact of including wait-list interventions described as no-intervention 2. The impact of including no-interventions described as wait-list interventions

Table 6. Differences from protocol and review *(Continued)*

7. Impact of non-normally distributed data.

Meta-regression	It was planned that we would conduct supplementary meta-regression analyses on continuous outcome, based on the findings from the subgroup analyses. We planned to choose covariates based on relevant subgroup analyses, such as type of intervention, risk of bias, type of outcome domain, mean participant age and duration of interventions.	Due to a lack of data, meta-regression analyses were not possible to conduct.
Selections of studies (participants)	All patients in each included trial was required to have a formal diagnosis of a mental health disorder, according to the Diagnostic and Statistical Manual of Mental Disorders (DSM), First Edition (DSM-I;APA 1952), SecondEdition (DSM-II;APA 1968),Third Edition (DSM-III; APA 1980), Third Edition Revised (DSM-IIIR; APA 1987), Fourth Edition (DSM-IV; APA 1994), Fourth Edition Text Revision (DSM-IV-TR; APA 2000), and Fifth Edition (DSM-5; APA 2013), or according to the International Classification ofDiseases andRelated-Health Problems (ICD), Sixth Edition (ICD-6; WHO 1949), Seventh Edition (ICD-7; WHO 1955), Eighth Edition (ICD-8; WHO 1967), Ninth Edition (ICD-9; WHO 1975), 10th Edition (ICD-10; WHO 1993), or 11th Edition (ICD-11; WHO 2018).	During the inclusion phase we identified trials, where patients fulfilled the symptoms of the disorder according to ICD or DSM, but where the trial did not explicitly state a classification system. Moreover, trials were also identified where the populations was classified as having a mental health disorder, but full diagnostic symptoms were not reported. We chose to include these in our analyses, and included a post-hoc subgroup analysis testing the difference
Decision hierarchy of outcomes	We used the following decision hierarchy to select the outcomes measuring effect. 1. We first included the outcome indicated as the primary outcome in the trial report. This could for instance be the one used for the sample size calculation. We preferred data from end of treatment over follow-up data. 2. If the trial did not differentiate between primary and secondary outcomes or if more than one primary outcome was stated, we preferred continuous to dichotomous outcomes. 3. If there were multiple continuous outcomes, we preferred observer-reported to patient-reported outcomes, and blinded to non-blinded outcomes.	During the extraction phase, we had to extend our decision hierarchy due to the fact that the pre-defined hierarchy proved not to be sufficient. We added: 1. If trials reported several observer-reported outcomes, we included the outcomes that best captured the core symptoms of the mental health population being treated. Here, we preferred global scores over sub-scores. 2. We then identified the outcome measure with the best psychometric properties (e.g. validity and reliability). 3. If still undecided, the final outcome measure was randomly decided.

APPENDICES

Appendix 1. Search strategy

Database: Ovid MEDLINE(R) <1946 to March Week 5 2018 > Search Strategy:

1. exp mental disorders/ or exp anxiety disorders/ or exp agoraphobia/ or exp anxiety, separation/ or exp neurocirculatory asthenia/ or exp neurotic disorders/ or exp obsessive-compulsive disorder/ or exp hoarding disorder/ or exp panic disorder/ or exp phobic disorders/ or exp phobia, social/ or exp "bipolar and related disorders"/ or exp bipolar disorder/ or exp "disruptive, impulse control, and conduct disorders"/ or exp firesetting behavior/ or exp gambling/ or exp trichotillomania/ or exp dissociative disorders/ or exp multiple personality disorder/ or exp elimination disorders/ or exp encopresis/ or exp enuresis/ or exp "feeding and eating disorders"/ or exp anorexia nervosa/ or exp binge-eating disorder/ or exp bulimia nervosa/ or exp "feeding and eating disorders of childhood"/ or exp female athlete triad syndrome/ or exp food addiction/ or exp night eating syndrome/ or exp pica/ or exp mood disorders/ or exp depressive disorder/ or exp depression, postpartum/ or exp depressive disorder, major/ or exp depressive disorder, treatment-resistant/ or exp dysthymic disorder/ or exp premenstrual dysphoric disorder/ or exp seasonal affective disorder/ or exp cyclothymic disorder/ or exp motor disorders/ or exp neurocognitive disorders/ or exp amnesia/ or exp alcoholic korsakoff syndrome/ or exp amnesia,

anterograde/ or exp amnesia, retrograde/ or exp amnesia, transient global/ or exp cognition disorders/ or exp auditory perceptual disorders/ or exp huntington disease/ or exp cognitive dysfunction/ or exp consciousness disorders/ or exp delirium/ or exp emergence delirium/ or exp dementia/ or exp aids dementia complex/ or exp alzheimer disease/ or exp aphasia, primary progressive/ or exp primary progressive nonfluent aphasia/ or exp creutzfeldt-jakob syndrome/ or exp dementia, vascular/ or exp dementia, multi-infarct/ or exp diffuse neurofibrillary tangles with calcification/ or exp frontotemporal lobar degeneration/ or exp frontotemporal dementia/ or exp "pick disease of the brain"/ or exp kluver-bucy syndrome/ or exp lewy body disease/ or exp dyslexia, acquired/ or exp alexia, pure/ or exp neurodevelopmental disorders/ or exp "attention deficit and disruptive behavior disorders"/ or exp attention deficit disorder with hyperactivity/ or exp conduct disorder/ or exp child behavior disorders/ or exp child development disorders, pervasive/ or exp autism spectrum disorder/ or exp asperger syndrome/ or exp autistic disorder/ or exp communication disorders/ or exp childhood-onset fluency disorder/ or exp social communication disorder/ or exp speech sound disorder/ or exp developmental disabilities/ or exp intellectual disability/ or exp learning disorders/ or exp dyscalculia/ or exp dyslexia/ or exp specific learning disorder/ or exp motor skills disorders/ or exp mutism/ or exp reactive attachment disorder/ or exp schizophrenia, childhood/ or exp stereotypic movement disorder/ or exp tic disorders/ or exp paraphilic disorders/ or exp exhibitionism/ or exp "fetishism (psychiatric)"/ or exp masochism/ or exp pedophilia/ or exp sadism/ or exp transvestism/ or exp voyeurism/ or exp personality disorders/ or exp antisocial personality disorder/ or exp borderline personality disorder/ or exp compulsive personality disorder/ or exp dependent personality disorder/ or exp histrionic personality disorder/ or exp hysteria/ or exp paranoid personality disorder/ or exp passive-aggressive personality disorder/ or exp schizoid personality disorder/ or exp schizotypal personality disorder/ or exp "schizophrenia spectrum and other psychotic disorders"/ or exp affective disorders, psychotic/ or exp capgras syndrome/ or exp delusional parasitosis/ or exp morgellons disease/ or exp paranoid disorders/ or exp psychotic disorders/ or exp psychoses, substance-induced/ or exp schizophrenia/ or exp schizophrenia, catatonic/ or exp schizophrenia, disorganized/ or exp schizophrenia, paranoid/ or exp shared paranoid disorder/ or exp sexual dysfunctions, psychological/ or exp dyspareunia/ or exp erectile dysfunction/ or exp gender dysphoria/ or exp premature ejaculation/ or exp "sexual and gender disorders"/ or exp vaginismus/ or exp sleep wake disorders/ or exp dyssomnias/ or exp sleep deprivation/ or exp sleep disorders, circadian rhythm/ or exp sleep disorders, intrinsic/ or exp parasomnias/ or exp nocturnal paroxysmal dystonia/ or exp rem sleep parasomnias/ or exp restless legs syndrome/ or exp sleep arousal disorders/ or exp sleep bruxism/ or exp sleep-wake transition disorders/ or exp somatoform disorders/ or exp body dysmorphic disorders/ or exp conversion disorder/ or exp factitious disorders/ or exp munchausen syndrome/ or exp munchausen syndrome by proxy/ or exp hypochondriasis/ or exp neurasthenia/ or exp substance-related disorders/ or exp alcohol-related disorders/ or exp alcohol amnestic disorder/ or exp alcohol withdrawal delirium/ or exp alcoholic intoxication/ or exp alcoholism/ or exp binge drinking/ or exp psychoses, alcoholic/ or exp wernicke encephalopathy/ or exp amphetamine-related disorders/ or exp cocaine-related disorders/ or exp inhalant abuse/ or exp marijuana abuse/ or exp "marijuana use"/ or exp neonatal abstinence syndrome/ or exp opioid-related disorders/ or exp morphine dependence/ or exp opium dependence/ or exp phencyclidine abuse/ or exp substance abuse, intravenous/ or exp substance abuse, oral/ or exp substance withdrawal syndrome/ or exp "tobacco use disorder"/ or exp "trauma and stressor related disorders"/ or exp adjustment disorders/ or exp stress disorders, traumatic/ or exp battered child syndrome/ or exp combat disorders/ or exp psychological trauma/ or exp stress disorders, post-traumatic/ or exp stress disorders, traumatic, acute/

2. exp PLACEBO EFFECT/ or exp Placebos/
3. (control* or compar* or nonspecific or non-specific or un-specific or unspecific or vehicle* or placebo* or credible or pseudo or sham or mock or fake or dumm* or attention or "common factor*").ab,hw,kf,ti.
4. (usual or clinic* or standard* or enhanc* or routine or outpatient* or convention* or gener* or local* or structur* or manual* or optim*).ab,hw,kf,ti.
5. (care or practi* or management* or treat* or intervention* or contact* or pill* or tablet* or medic* or therap* or surger* or operat*).ab,hw,kf,ti.
6. 4 and 5
7. TAU.ab,hw,kf,ti.
8. exp Waiting Lists/
9. ("no* care" or "no* practi*" or "no* management*" or "no* treat*" or "no* intervention*" or "no* contact*" or "no* pill*" or "no* tablet*" or "no* medic*" or "no* therap*" or "no* surger*" or "no* operat*" or "no* active*" or "no* experimental*").ab,hw,kf,ti.
10. (no*care or no*practi* or no*management* or no*treat* or no*intervention* or no*contact* or no*pill* or no*tablet* or no*medic* or no*therap* or no*surger* or no*operat* or no*active* or no*experimental*).ab,hw,kf,ti.
11. (uncare or unpracti* or unmanagement* or untreat* or unintervention* or uncontact* or unmedic* or untherap* or unsurger* or unoperat*or unactive* or unexperimental*).ab,hw,kf,ti.
12. ("un care" or "un practi*" or "un management*" or "un treat*" or "un intervention*" or "un contact*" or "un medic*" or "un therap*" or "un surger*" or "un operat*" or "un active*" or "un experimental*").ab,hw,kf,ti.
13. ("minim* care" or "minim* practi*" or "minim* management*" or "minim* treat*" or "minim* intervention*" or "minim* contact*" or "minim* medic*" or "minim* therap*" or "minim* surger*" or "minim* operat*" or "minim* active*" or "minim* experimental*" or "minim* period*" or "minim* time*").ab,hw,kf,ti.
14. ("without care" or "without practi*" or "without management*" or "without treat*" or "without intervention*" or "without contact*" or "without pill*" or "without tablet*" or "without medic*" or "without therap*" or "without surger*" or "without operat*" or "without active*" or "without experimental*").ab,hw,kf,ti.

15. ("delay* care" or "delay* practi*" or "delay* management*" or "delay* treat*" or "delay* intervention*" or "delay* contact*" or "delay* pill*" or "delay* tablet*" or "delay* medic*" or "delay* therap*" or "delay* surger*" or "delay* operat*" or "delay* active*" or "delay* experimental*" or "delay* list*" or "delay* period*" or "delay* time").ab,hw,kf,ti.
16. (await* or wait*).ab,hw,kf,ti.
17. randomi#ed controlled trial.pt.
18. controlled clinical trial.pt.
19. randomi#ed.ab.
20. "placebo* ".ab.
21. drug therapy.fs
22. randomly.ab.
23. trial.ab.
24. groups.ab.
25. exp Animals/
26. Humans/
27. 25 not 26
28. 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24
29. 28 not 27
30. 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16
31. 2 or 3 or 6
32. 1 and 29 and 30 and 31

Database: PsycINFO <1806 to April Week 1 2018> Search Strategy:

1. exp mental disorders/ or exp adjustment disorders/ or exp affective disorders/ or exp alexithymia/ or exp anxiety disorders/ or exp autism spectrum disorders/ or exp chronic mental illness/ or exp dementia/ or exp dissociative disorders/ or exp eating disorders/ or exp elective mutism/ or exp factitious disorders/ or exp gender identity disorder/ or exp hoarding disorder/ or exp hysteria/ or exp impulse control disorders/ or exp neurosis/ or exp paraphilias/ or exp personality disorders/ or exp pseudodementia/ or exp psychosis/ or exp schizoaffective disorder/ or exp abnormal psychology/ or exp adaptive behavior/ or exp attention deficit disorder/ or exp attention deficit disorder with hyperactivity/ or exp behavior disorders/ or exp borderline states/ or exp brain disorders/ or exp communication disorders/ or exp conduct disorder/ or exp consciousness disturbances/ or exp emotional disturbances/ or exp infantilism/ or exp intellectual development disorder/ or exp learning disorders/ or exp narcissism/ or exp personality processes/ or exp psychiatric patients/ or exp psychiatric symptoms/ or exp psychodiagnosis/ or exp psychopathology/ or exp sexual function disturbances/ or exp sleep disorders/
2. exp PLACEBO/
3. exp OUTPATIENTS/
4. (control* or compar* or nonspecific or non-specific or un-specific or unspecific or vehicle* or placebo* or fals* or credible or pseudo or sham or mock or fake or dumm* or neutral or attention or "common factor*").ab,hw,id,ti.
5. (usual or clinic* or standard* or enhanc* or routine or outpatient* or convention* or gener* or local* or structur* or manual* or optim*).ab,hw,id,ti.
6. (care or practi* or management* or treat* or intervention* or contact* or pill* or tablet* or medic* or therap* or surger* or operat* or period* or time).ab,hw,id,ti.
7. 5 and 6
8. TAU.ab,hw,id,ti.
9. exp Experiment Controls/
10. ("no* care" or "no* practi*" or "no* management*" or "no* treat*" or "no* intervention*" or "no* contact*" or "no* pill*" or "no* tablet*" or "no* medic*" or "no* therap*" or "no* surger*" or "no* operat*" or "no* active*" or "no* experimental*").ab,hw,id,ti.
11. (no*care or no*practi* or no*management* or no*treat* or no*intervention* or no*contact* or no*pill* or no*tablet* or no*medic* or no*therap* or no*surger* or no*operat* or no*active* or no*experimental*).ab,hw,id,ti.
12. (uncare or unpracti* or unmanagement* or untreat* or unintervention* or uncontact* or unmedic* or untherap* or unsurger* or unoperat*or unactive* or unexperimental*).ab,hw,id,ti.
13. ("un care" or "un practi*" or "un management*" or "un treat*" or "un intervention*" or "un contact*" or "un medic*" or "un therap*" or "un surger*" or "un operat*" or "un active*" or "un experimental*").ab,hw,id,ti.
14. ("minim* care" or "minim* practi*" or "minim* management*" or "minim* treat*" or "minim* intervention*" or "minim* contact*" or "minim* medic*" or "minim* therap*" or "minim* surger*" or "minim* operat*" or "minim* active*" or"minim* experimental*" or "minim* period*" or "minim* time*").ab,hw,id,ti.

15. ("without care" or "without practi*" or "without management*" or "without treat*" or "without intervention*" or "without contact*" or "without pill*" or "without tablet*" or "without medic*" or "without therap*" or "without surger*" or "without operat*" or "without active*" or "without experimental*").ab,hw,id,ti.
16. ("delay* care" or "delay* practi*" or "delay* management*" or "delay* treat*" or "delay* intervention*" or "delay* contact*" or "delay* pill*" or "delay* tablet*" or "delay* medic*" or "delay* therap*" or "delay* surger*" or "delay* operat*" or "delay* active*" or "delay* experimental*" or "delay* list*" or "delay* period*" or "delay* time").ab,hw,id,ti.
17. (await* or wait*).ab,hw,id,ti.
18. exp Clinical Trials/
19. (random* adj allocat*).ab.
20. randomi?ed.ab.
21. placebo.ab.
22. "random* ".ab.
23. "trial* ".ab.
24. "group* ".ab.
25. drug therapy.sh.
26. exp Animals/ not Humans/
27. 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25
28. 27 not 26
29. 2 or 3 or 4 or 7 or 8
30. 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17
31. 1 and 28 and 29 and 30

Database: Embase <1974 to 2018 April 09> Search Strategy:

1. exp mental disease/ or exp addiction/ or exp adjustment disorder/ or exp alexithymia/or exp anxiety disorder/ or exp autism/ or exp behavior disorder/ or exp delirium/ or exp dissociative disorder/ or exp emotional disorder/ or exp learning disorder/ or exp memory disorder/ or exp mental deficiency/ or exp mental infantilism/ or exp mental instability/ or exp mood disorder/ or exp neurosis/ or exp organic brain syndrome/ or exp personality disorder/ or exp psychosexual disorder/ or exp psychosis/ or exp psychosomatic disorder/ or exp psychotrauma/ or exp schizophrenia spectrum disorder/ or exp stupor/ or exp thought disorder/
2. exp placebo effect/ or exp placebo/
3. exp outpatient care/
4. (control* or compar* or nonspecific or non-specific or un-specific or unspecific or vehicle* or placebo* or credible or pseudo or sham or mock or fake or dumm* or attention or "common factor*").ab,hw,kw,ti.
5. (usual or clinic* or standard* or enhanc* or routine or outpatient* or convention* or gener* or local* or structur* or manual* or optim*).ab,hw,kw,ti.
6. (care or practi* or management* or treat* or intervention* or contact* or pill* or tablet* or medic* or therap* or surger* or operat* or period* or time).ab,hw,kw,ti.
7. 5 and 6
8. TAU.ab,hw,kw,ti.
9. exp control group/
10. ("no* care" or "no* practi*" or "no* management*" or "no* treat*" or "no* intervention*" or "no* contact*" or "no* pill*" or "no* tablet*" or "no* medic*" or "no* therap*" or "no* surger*" or "no* operat*" or "no* active*" or "no* experimental*").ab,hw,kw,ti.
11. (no*care or no*practi* or no*management* or no*treat* or no*intervention* or no*contact* or no*pill* or no*tablet* or no*medic* or no*therap* or no*surger* or no*operat* or no*active* or no*experimental*).ab,hw,kw,ti.
12. (uncare or unpracti* or unmanagement* or untreat* or unintervention* or uncontact* or unmedic* or untherap* or unsurger* or unoperat*or unactive* or unexperimental*).ab,hw,kw,ti.
13. ("un care" or "un practi*" or "un management*" or "un treat*" or "un intervention*" or "un contact*" or "un medic*" or "un therap*" or "un surger*" or "un operat*" or "un active*" or "un experimental*").ab,hw,kw,ti.
14. ("minim* care" or "minim* practi*" or "minim* management*" or "minim* treat*" or "minim* intervention*" or "minim* contact*" or "minim* medic*" or "minim* therap*" or "minim* surger*" or "minim* operat*" or "minim* active*" or "minim* experimental*" or "minim* period*" or "minim* time*").ab,hw,kw,ti.
15. ("without care" or "without practi*" or "without management*" or "without treat*" or "without intervention*" or "without contact*" or "without pill*" or "without tablet*" or "without medic*" or "without therap*" or "without surger*"or "without operat*" or "without active*" or "without experimental*").ab,hw,kw,ti.

16. ("delay* care" or "delay* practi*" or "delay* management*" or "delay* treat*" or "delay* intervention*" or "delay* contact*" or "delay* pill*" or "delay* tablet*" or "delay* medic*" or "delay* therap*" or "delay* surger*" or "delay* operat*" or "delay* active*" or "delay* experimental*" or "delay* list*" or "delay* period*" or "delay* time").ab,hw,kw,ti.
17. (await* or wait*).ab,hw,kw,ti.
18. controlled clinical trial/ or exp clinical trial/ or exp controlled study/ or exp randomized controlled trial/
19. (random* adj allocat*).ab.
20. randomi?ed.ab.
21. placebo.ab.
22. "random* ".ab.
23. "trial* ".ab.
24. drug therapy.fs.
25. exp Animals/ not Humans/
26. 18 or 19 or 20 or 21 or 22 or 23 or 24
27. 26 not 25
28. 2 or 3 or 4 or 7 or 8
29. 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17
30. 1 and 27 and 28 and 29

Database: Cochrane Central Register of Controlled Trials; current issue:

1. MeSH descriptor: [Mental Disorders] explode all trees
2. MeSH descriptor: [Placebos] explode all trees
3. MeSH descriptor: [Placebo Effect] explode all trees
4. (control* or compar* or nonspecific or non-specific or un-specific or unspecific or vehicle* or placebo* or credible or pseudo or sham or mock or fake or dumm* or attention or "common factor*"):ti,ab,kw (Word variations have been searched)
5. (usual or clinic* or standard* or enhanc* or routine or outpatient* or convention* or gener* or local* or structur* or manual* or optim*) next (care or practi* or management* or treat* or intervention* or contact* or pill* or tablet* or medic* or therap* or surger* or operat* or period* or time):ti,ab,kw (Word variations have been searched)
6. TAU:ti,ab,kw (Word variations have been searched)
7. MeSH descriptor: [Waiting Lists] explode all trees
8. "no care" or "non care" or "no treat*" or "non treat*" or "no pract*" or "non pract*" or "no intervention" or "non intervention" or "no management" or "non management" or "no pill*" or "non pill" or "no contact" or "non contact" or "no tablet" or "non tablet" or "no medication*" or "non medication" or "no therap*" or "non therap*" or "no surger*" or "non surger*" or "no operat*" or "non operat*" or "no active" or "non active" or "no experimental" or "non experimental":ti,ab,kw (Word variations have been searched)
9. un near (care or practi* or management* or treat* or intervention* or contact* or pill* or tablet* or medic* or therap* or surger* or operat* or active* or experimental*):ti,ab,kw (Word variations have been searched)
10. minim* near (care or practi* or management* or treat* or intervention* or contact* or pill* or tablet* or medic* or therap* or surger* or operat* or active* or experimental*):ti,ab,kw (Word variations have been searched)
11. without near (care or practi* or management* or treat* or intervention* or contact* or pill* or tablet* or medic* or therap* or surger* or operat* or active* or experimental*):ti,ab,kw (Word variations have been searched)
12. delay near (care or practi* or management* or treat* or intervention* or contact* or pill* or tablet* or medic* or therap* or surger* or operat* or active* or experimental*):ti,ab,kw (Word variations have been searched)
13. (await* or wait*):ti,ab,kw (Word variations have been searched)
14. #2 or #3 or #4 or #5 or #6
15. #7 or #8 or #9 or #10 or #11 or #12 or #13
16. #1 and #14 and #15 in Trials

Database - AMED - The Allied and Complementary Medicine Database:

Interface - EBSCOhost Research Databases

Search Screen - Advanced Search

Search modes - Boolean/Phrase

1. TX Mental* OR psych* OR Anxi* OR Bipolar OR Conduct disorder* OR Dissociative OR Elimination Disorder* OR Eat* OR Mood* OR Motor Disorder* OR Neuro* OR Paraphilic OR Personality OR Schizophren* OR Sexual Dys* OR Sleep* OR Somatoform* OR Substance* OR Trauma

2. placebo
3. (control* or compar* or nonspecific or non-specific or un-specific or unspecific or vehicle* or placebo* or credible or pseudo or sham or mock or fake or dumm* or attention or "common factor*")
4. (usual or clinic* or standard* or enhanc* or routine or outpatient* or convention* or gener* or local* or structur* or manual* or optim*)
5. (care or practi* or management* or treat* or intervention* or contact* or pill* or tablet* or medic* or therap* or surger* or operat*)
6. (S4 AND S5)
7. TAU
8. ("no* care" or "no* practi*" or "no* management*" or "no* treat*" or "no* intervention*" or "no* contact*" or "no* pill*" or "no* tablet*" or "no* medic*" or "no* therap*" or "no* surger*" or "no* operat*" or "no* active*" or "no* experimental*")
9. (no*care or no*practi* or no*management* or no*treat* or no*intervention* or no*contact* or no*pill* or no*tablet* or no*medic* or no*therap* or no*surger* or no*operat* or no*active* or no*experimental*)
10. (uncare or unpracti* or unmanagement* or untreat* or unintervention* or uncontact* or unmedic* or untherap* or unsurger* or unoperat*or unactive* or unexperimental*)
11. ("un care" or "un practi*" or "un management*" or "un treat*" or "un intervention*" or "un contact*" or "un medic*" or "un therap*" or "un surger*" or "un operat*" or "un active*" or "un experimental*")
12. ("minim* care" or "minim* practi*" or "minim* management*" or "minim* treat*" or "minim* intervention*" or "minim* contact*" or "minim* medic*" or "minim* therap*" or "minim* surger*" or "minim* operat*" or "minim* active*" or "minim* experimental*" or "minim* period*" or "minim* time*")
13. ("without care" or "without practi*" or "without management*" or "without treat*" or "without intervention*" or "without contact*" or "without pill*" or "without tablet*" or "without medic*" or "without therap*" or "without surger*" or "without operat*" or "without active*" or "without experimental*")
14. ("delay* care" or "delay* practi*" or "delay* management*" or "delay* treat*" or "delay* intervention*" or "delay* contact*" or "delay* pill*" or "delay* tablet*" or "delay* medic*" or "delay* therap*" or "delay* surger*" or "delay* operat*" or "delay* active*" or "delay* experimental*" or "delay* list*" or "delay* period*" or "delay* time")
15. (await* or wait*)
16. Randomi#ed or controlled trial* or clinical trial* or placebo* or random* or trial or groups
17. (S2 OR S6 OR S7)
18. (S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15)
19. S1 AND S16 AND S17 AND S18

Database: Web of Science; 1900 to current:

1. TOPIC: (Mental* OR psych* OR Anxi* OR Bipolar OR Conduct disorder* OR Dissociative OR Elimination Disorder* OR Eat* OR Mood* OR Motor Disorder* OR Neuro* OR Paraphilic OR Personality OR Schizophren* OR Sexual Dys* OR Sleep* OR Somatoform* OR Substance* OR Trauma*) *Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI Timespan=All years*
2. TS=(Placebo or control or comparison or vehicle or false or credible or pseudo or sham or mock or dummy or neutral or "standard care" or "usual intervention" or "routine care" or TAU or "treatment as usual" or "usual care" or "standard care" or "standard intervention" or "enhanced care" or "convention* care") *Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI Timespan=All years*
3. TS=(("waiting list" or wait* or await* or "no*intervention" or "no therapy" or "no*treatment" or "no*care" or "no care" or "no treatment" or "minim*treatment" or "minim*care" or "minim* therapy" or "without care" or "without treatment" or "without intervention" or "without therapy" or "delayed care" or "delayed treatment" or "delayed therapy" or "delayed intervention" *indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI Timespan=All years*
4. TI=(randomized or randomised or controlled trial* or clinical trial* or placebo* or drug therapy or random* or trial or groups *Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI Timespan=All years*
5. #4 AND #3 AND #2 AND #1 *Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, ESCI Timespan=All years*

Database: ProQuest Dissertations and Theses A&I; 1743 to current:

- ((Mental* OR psych* OR Anxi* OR Bipolar OR Conduct disorder* OR Dissociative OR Elimination Disorder* OR Eat* OR Mood* OR Motor Disorder* OR Neuro* OR paedophilic OR Personality OR Schizophren* OR Sexual Dys* OR Sleep* OR Somatoform* OR Substance* OR Trauma*)
- AND (placebo OR control OR comparison OR vehicle OR false OR credible OR pseudo OR sham OR mock OR dummy OR neutral OR "standard care" OR "usual intervention" OR "routine care" OR TAU OR "treatment as usual" OR "usual care" OR "standard care" OR "standard intervention" OR "enhanced care" OR "convention* care")
- AND ("waiting list" OR wait* OR await* OR "no*intervention" OR "no therapy" OR "no*treatment" OR "no*care" OR "no care" OR "no treatment" OR "minim*treatment" OR "minim*care" OR "minim* therapy" OR "without care" OR "without treatment" OR "without intervention" OR "without therapy" OR "delayed care" OR "delayed treatment" OR "delayed therapy" OR "delayed intervention"))
- AND diskw.Exact("PLACEBO" OR "Random" OR "Clinical Trial" OR "Controlled

- trials" OR "randomised controlled trial" OR "Randomized Controlled Trial")

Database: Sociological Abstracts ProQuest; 1952 to current:

- (Mental* OR psych* OR Anxi* OR Bipolar OR Conduct disorder* OR Dissociative OR Elimination Disorder* OR Eat* OR Mood* OR Motor Disorder* OR Neuro* OR paedophilic OR Personality OR Schizophren* OR Sexual Dys* OR Sleep* OR Somatoform* OR Substance* OR Trauma*)
- AND ab(placebo OR control OR comparison OR vehicle OR false OR credible OR pseudo OR sham OR mock OR dummy OR neutral OR "standard care" OR "usual intervention" OR "routine care" OR TAU OR "treatment as usual" OR "usual care" OR "standard care" OR "standard intervention" OR "enhanced care" OR "convention* care")
- AND ab("waiting list" OR wait* OR await* OR "no*intervention" OR "no therapy" OR"no*treatment" OR "no*care" OR "no care" OR "no treatment" OR "minim*treatment" OR "minim*care" OR "minim* therapy" OR "without care" OR "without treatment" OR "without intervention" OR "without therapy" OR "delayed care" OR "delayed treatment" OR "delayed therapy" OR "delayed intervention")
- AND ab("PLACEBO" OR "Random" OR "Clinical Trial" OR "Controlled trials" OR "randomised controlled trial" OR "Randomized Controlled Trial")

Database: Google Scholar; top 200 of relevance according to Bramer 2017 **:**

- Mental | psychiatric | psychological | no treatment | waitlist | placebo | usual care | random | clinical trials

Database: BIOSIS Previews; 1969 to current:

1. TOPIC: (mental disorder*) *DocType=All document types; Language=All languages;*
2. *TOPIC: ("waiting list" or wait* or await* or "no*intervention" or "no therapy" or "no*treatment" or "no*care" or "no care" or "no treatment" or "minim*treatment" or "minim*care" or "minim* therapy" or "without care" or "without treatment" or "without intervention" or "without therapy" or "delayed care" or "delayed treatment" or "delayed therapy" or "delayed intervention") DocType=All document types; Language=All languages;*
3. *TOPIC: (Placebo or control or comparison or vehicle or false or credible or pseudo or sham or mock or dummy or neutral or "standard care" or "usual intervention" or "routine care" or TAU or "treatment as usual" or "usual care" or "standard care" or "standard intervention" or "enhanced care" or "convention* care")DocType=All document types; Language=All languages;*
4. *#3 AND #2 AND #1DocType=All document types; Language=All languages;*

Database: Open Grey;1997 to current

- (mental* OR psych*) AND (placebo OR usual care OR "treatment as usual" OR wait-list OR wait list OR await* OR wait*) AND (no treatment OR wait-list OR wait list OR await* OR wait*) AND (random*)

Trial registry: Australian New Zealand Clinical Trials Registry (ANZCTR); www.anzctr.org.au/BasicSearch.aspx

- Search terms: (mental OR psychiatric) AND (placebo OR usual care OR waitlist) AND (no treatment OR waitlist)
- Allocation to treatment: Randomised
- Condition category: Mental health
- Healthy Volunteers: No

Trial registry: Clinical Trials; clinicaltrials.gov

- Condition category: Mental Disorder
- Other terms: (placebo OR usual care OR wait-list)
- Intervention/treatment: No treatment

Trial registry: EU Clinical Trials Register; www.clinicaltrialsregister.eu/ctr-search/search

- (Mental disorder OR psychiatric) AND (placebo OR usual care OR wait-list) AND (no treatment OR wait-list)

Trial registry: ISRCTN; www.isrctn.com

Search 1:

- Condition category: Mental and behavioural disorders
- Interventions: No treatment

Search 2:

- Condition category: Mental and behavioural disorders

- Interventions: Waitlist

Trial registry: UK Clinical Trials Gateway; www.ukctg.nihr.ac.uk/#popoverSearchDivId

- (Mental disorder OR psychiatric) AND (placebo OR usual care OR wait-list) AND (no treatment OR wait-list)

Trial registry: World Health Organization International Clinical Trials Registry Platform (WHO ICTRP); http://apps.who.int/trialsearch/

- Condition: (Mental disorder OR psychiatric)
- Intervention: No treatment OR wait-list

Appendix 2. Template for extraction sheet

Source	Trial ID (Original trial, e.g. Plizska 2000) Full citationOther publications on same studyAuthor contact informationPublication type (e.g. full report, abstract, letter) Form filled by (date, name)Ethical approval
Eligibility	Confirm eligibility: yes/no/awaiting
Correspondence	Correspondence required: yes/no
Methods	Design (number of arms): Sample calculation: Cluster randomised: (yes/no): Duration of trial (baseline to post): Duration of participation (trial + follow-up): Setting: Purpose of trial: Open or closed placebo:
Participants	Number of participants screened: Number of participants included: Number of participants followed-up: Number of participants randomly assigned to: • Control 1: • Control 2: • Active treatment: Number of withdrawals: • Control 1: • Control 2: • Active treatment: Diagnosis: Diagnostic manual (DSM/ICD) Means of assessment:

(Continued)

	Comorbidity:
	Age: mean years (range)
	IQ:
	Sex: (male/female)
	Ethnicity:
	Country:
	Country of treatment (where did the treatment take place):
	Inclusion criteria
	Exclusion criteria
Interventions Remember to state (if applicable): 1. whether pharmacological, psychological other not specified 2. whether individual or group psychological treat	**Control intervention** Treatment name (type): Description of intervention: Individual or group treatment: Exposure/intensity to treatment: Duration of treatment: Concomitant psychotherapy: Concomitant pharmacotherapy: **Comparator intervention** Comparison name (type): Description of intervention: Exposure/intensity to treatment: Duration treatment: Concomitant psychotherapy: Concomitant pharmacotherapy:
Outcomes Remember to state: 1. whether self-rated or observer-rated 2. Hierarchy	Relevant outcomes for effect: Relevant outcomes for adverse events:
Notes	Key conclusion from study authors Key limitations from study authors Other notes from review authors
Risk of bias 1. Support for judgement	Random sequence generation Allocation concealment

(Continued)

2. Risk of bias (low, unclear, high)	Blinding of outcome assessment Blinding of participants and personnel Incomplete outcome data Selective outcome reporting(Trial registry ID (search: clinicaltrials.gov (from 2008) and who.int/ictrp/en (from 2004) Other sources of bias: • Allegiance bias/industry has vested interest in one of the interventions • Attention bias (differences duration) • Baseline differences between groups

Appendix 3. Cochrane's Risk of bias tool 1.0

Random sequence generation

1. Low risk of bias: an adequate method for randomisation sequence generation was used (e.g. computer-generated random numbers or a table of random numbers), or the method was unlikely to introduce selection bias.
2. Unclear risk of bias: there was insufficient information to determine whether the applied randomisation method could introduce selection bias.
3. High risk of bias: the method applied was likely to introduce selection bias.

Allocation concealment

1. Low risk of bias: the method to conceal intervention allocations (e.g. central allocation) was unlikely to bias the results.
2. Unclear risk of bias: there was insufficient information to determine whether the applied method could bias allocation to interventions.
3. High risk of bias: the method applied (e.g. open random allocation schedule) could have biased the allocations to interventions.

Blinding of participants and personnel

1. Low risk of bias: the method of blinding was sufficiently described and blinding was conducted in a satisfactory way.
2. Unclear risk of bias: there was insufficient information to determine whether adequate blinding was used and whether it was likely to bias the effect estimates.
3. High risk of bias: no blinding procedures were used or the blinding procedures were incomplete.

It is important to highlight that blinding of participants and personnel were not possible in the included trials, since participants would be aware if they received any kind of care or no care. Therefore, all of the trials would be rated as high risk of bias.

Blinding of outcome assessment

1. Low risk of bias: the method of blinding was described and blinding was conducted in a satisfactory way.
2. Unclear risk of bias: there was insufficient information to determine whether the type of blinding was likely to bias the effect estimates.
3. High risk of bias: no blinding used or incomplete blinding was used.

Incomplete outcome data

1. Low risk of bias: missing data did probably not affect the outcome measures, as all missing data can be considered as missing at random or all data were reported.
2. Unclear risk of bias: there was insufficient information to determine whether missing data, or the method used to handle missing data was likely to bias the effect estimates.
3. High risk of bias: the crude estimate of effects could have been biased given the attrition rates, the reasons for the missing data, or the insufficient methods used to handle missing data.

Selective outcome reporting

1. Low risk of bias: the trial protocol was available and all pre-specified outcomes of interest were reported.
2. Unclear risk of bias: there was insufficient information to determine whether selective outcome reporting could have occurred.
3. High risk of bias: not all of the primary outcomes specified beforehand were reported or participants were excluded after randomisation.

Other sources of bias

1. Low risk of bias: the trial appeared to be free of other sources of bias.
2. Unclear risk of bias: there was insufficient information to determine the extent of other possible sources of bias.
3. High risk of bias: other sources of bias were identified.

Appendix 4. Trial Sequential Analysis (TSA) and Funnel Plot figures

The analysis on usual care compared with wait-list and no-treatment showed that the required information size was not reached. See Summary of findings 2. See Figure 4. We were not able to draw funnel plots for Usual care due to lack of data.

We performed a TSA on the primary outcomes all placebos, psychological placebos, pharmacological placebos, and pharmacological placebos included in the Summary of findings 3.

The analysis on all placebos compared with wait-list and no-treatment showed that the required information size was reached. See Figure 7. We drew a funnel plot for the comparison between all placebos and wait-list and no-treatment. The funnel plot shows no signs of asymmetry. See Figure 8.

Figure 7. When comparing all placebos with wait-list and no-treatment on beneficial effects, we performed a trial sequential analysis on the primary outcome. The analysis shows that the required information size was reached. See Figure 5 above. MIREDIF: Minimum relevant difference

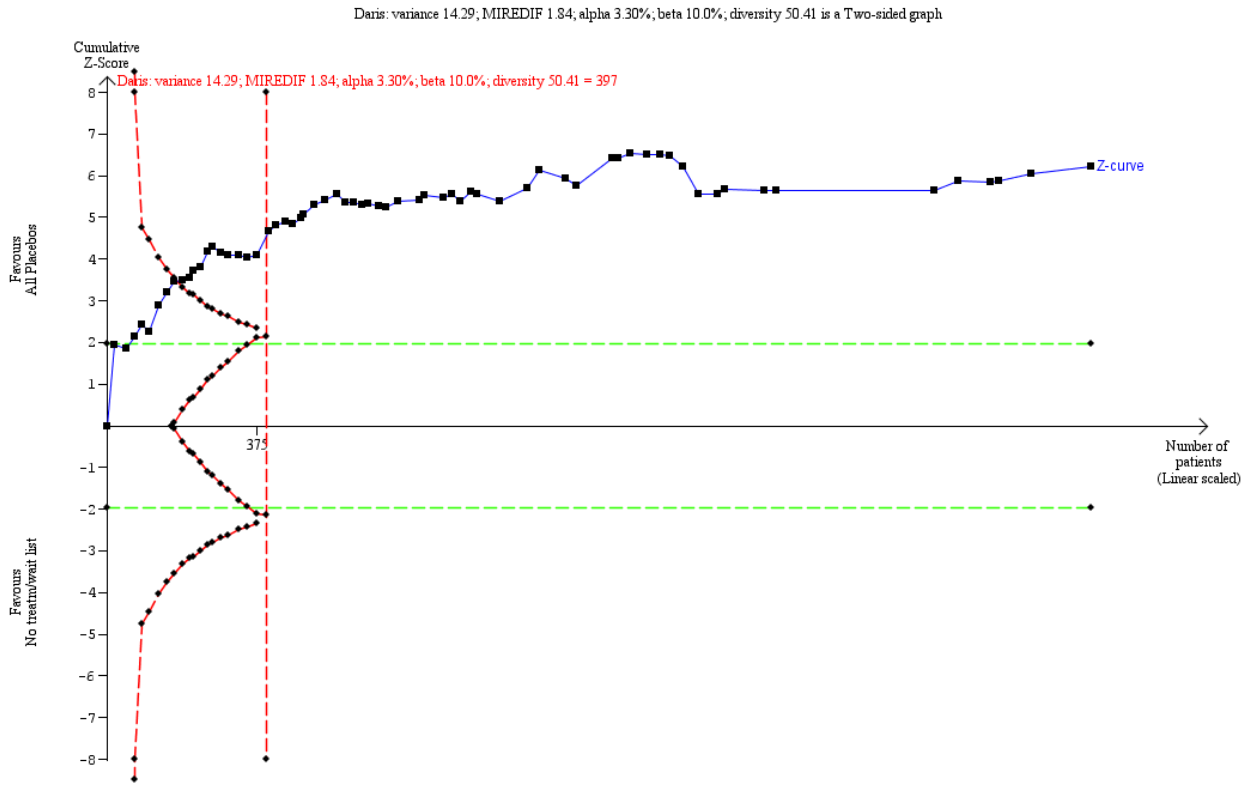
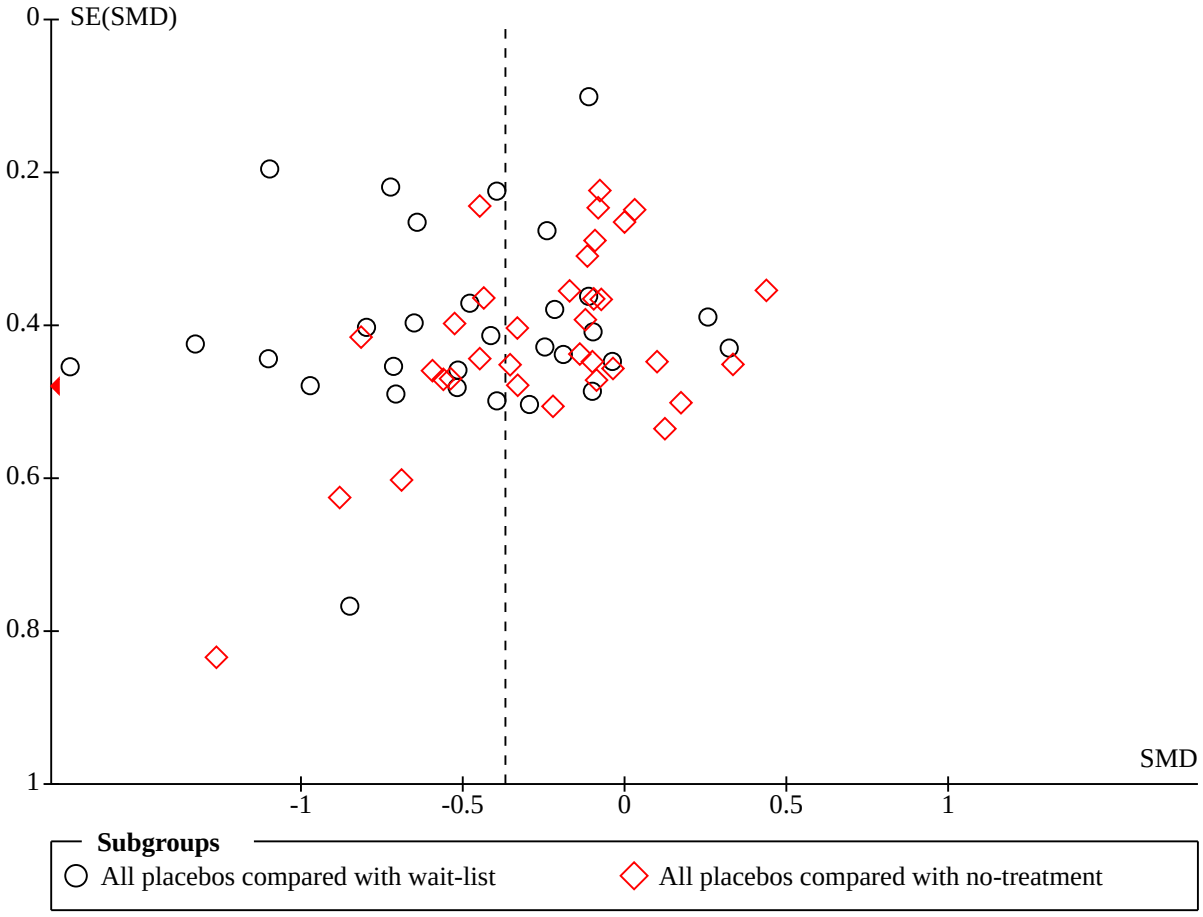
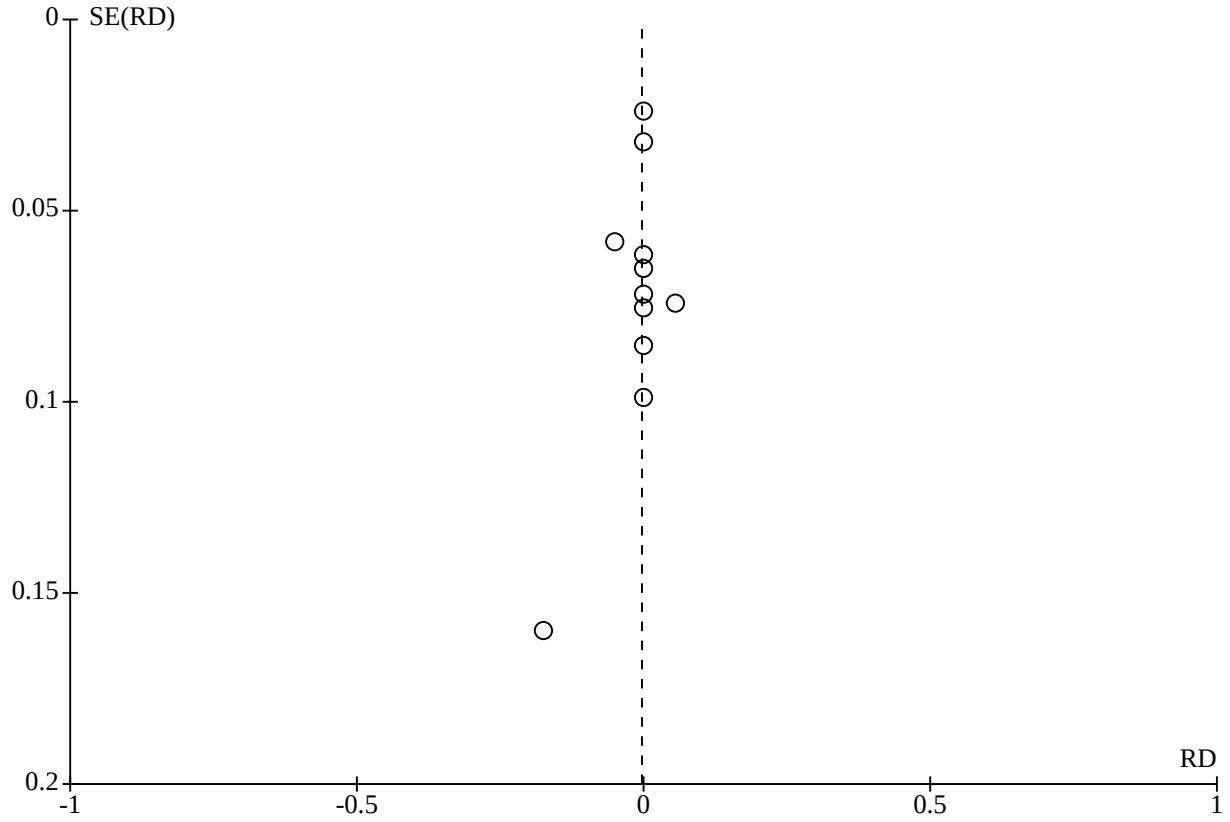


Figure 8. Funnel plot of comparison: 6.1. Efficacy of all placebos compared with wait-list/no-treatment for continuous data.



We drew a funnel plot for the serious adverse events between all placebos and wait-list and no-treatment. It was not possible to perform a TSA for serious adverse events. The funnel plot shows no signs of asymmetry. See Figure 9.

Figure 9. Funnel plot of comparison: 7.1 Serious adverse events of all placebos compared with wait-list/no-treatment for dichotomous data.



The analysis on psychological placebos compared with wait-list and no-treatment showed that the required information size was reached. See Figure 10. We drew a funnel plot for the comparison between all placebos and wait-list and no-treatment. The funnel plot shows no signs of asymmetry. See Figure 11.

Figure 10. We performed a trial sequential analysis on the primary outcome for efficacy of psychological placebos compared with wait-list and no-treatment. The analysis shows that the required information size was reached. See Figure 8 above. MIREDIF: Minimum relevant difference

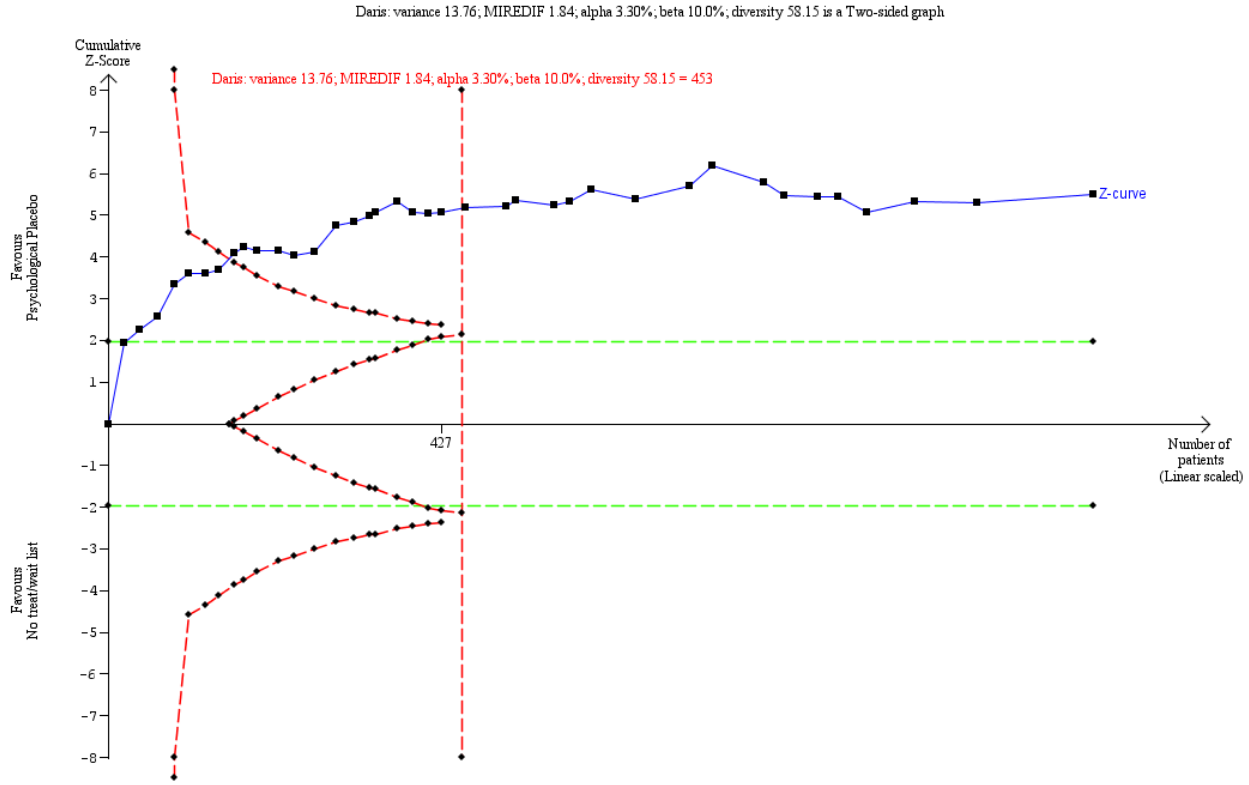
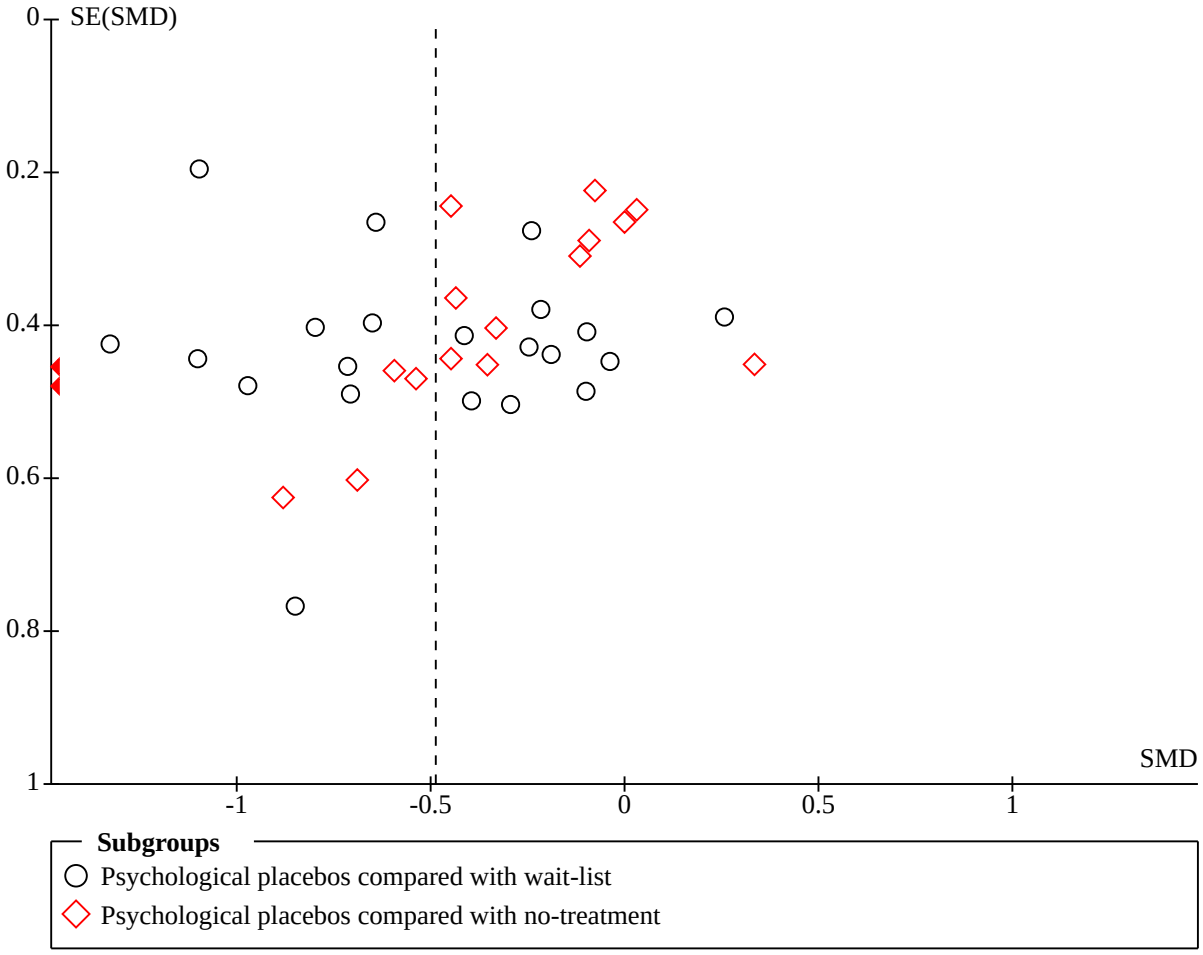
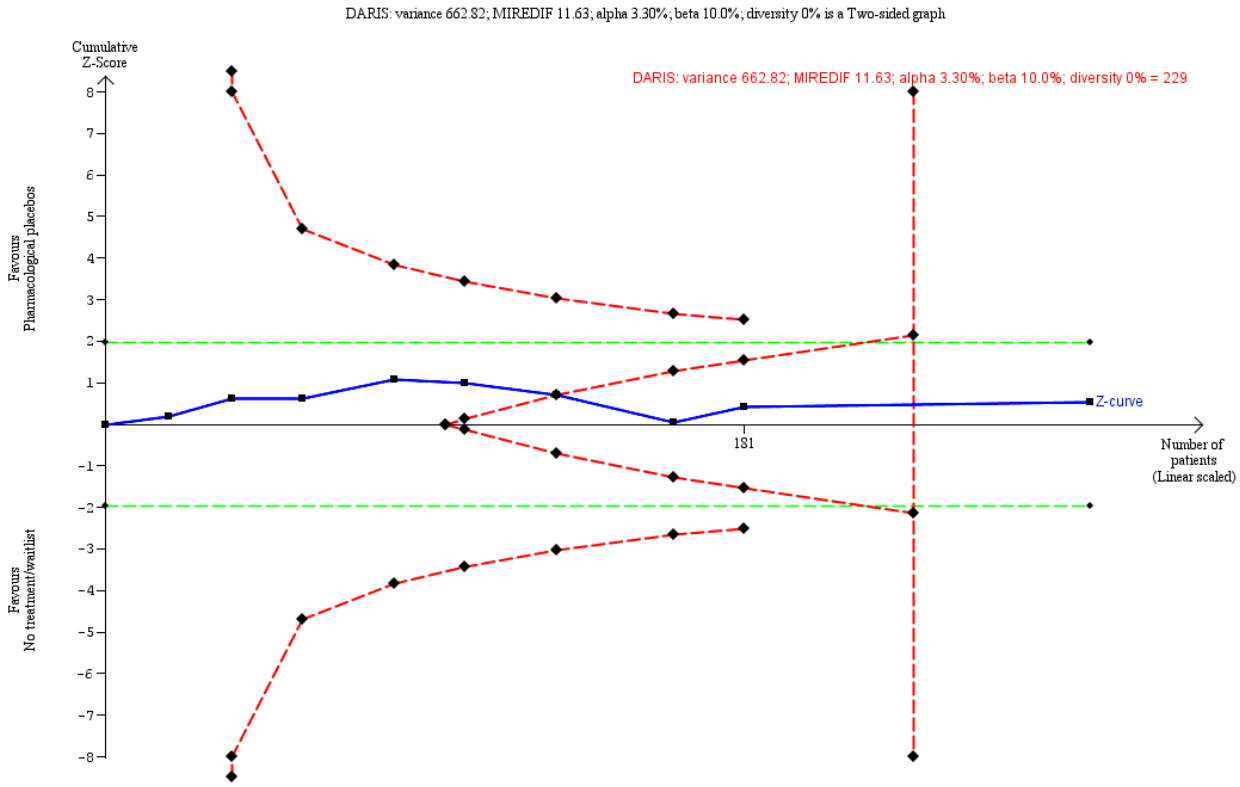


Figure 11. Funnel plot of comparison: 9.1. Efficacy psychological placebos compared with wait-list/no-treatment for continuous data.



The analysis on pharmacological placebos compared with wait-list and no-treatment showed that the required information size was not reached. See Figure 12. A funnel plot for the comparison between all placebos and wait-list and no-treatment were not possible due to lack of data.

Figure 12. We performed a trial sequential analysis on the primary outcome for efficacy of pharmacological placebos compared with wait-list and no-treatment. The TSA showed the cumulated Z curve enters the futility area. See Figure 10 above. MIREDIF: Minimum relevant difference



The analysis on physical placebos compared with wait-list and no-treatment showed that the required information size was reached. See Figure 13. We drew a funnel plot for the comparison between all placebos and wait-list and no-treatment. The funnel plot shows no signs of asymmetry. See Figure 14.

Figure 13. We performed a trial sequential analysis on the primary outcome for efficacy of physical placebos compared with wait-list and no-treatment. The analysis shows that the required information size was reached. See Figure 11 above. MIREDIF: Minimum relevant difference

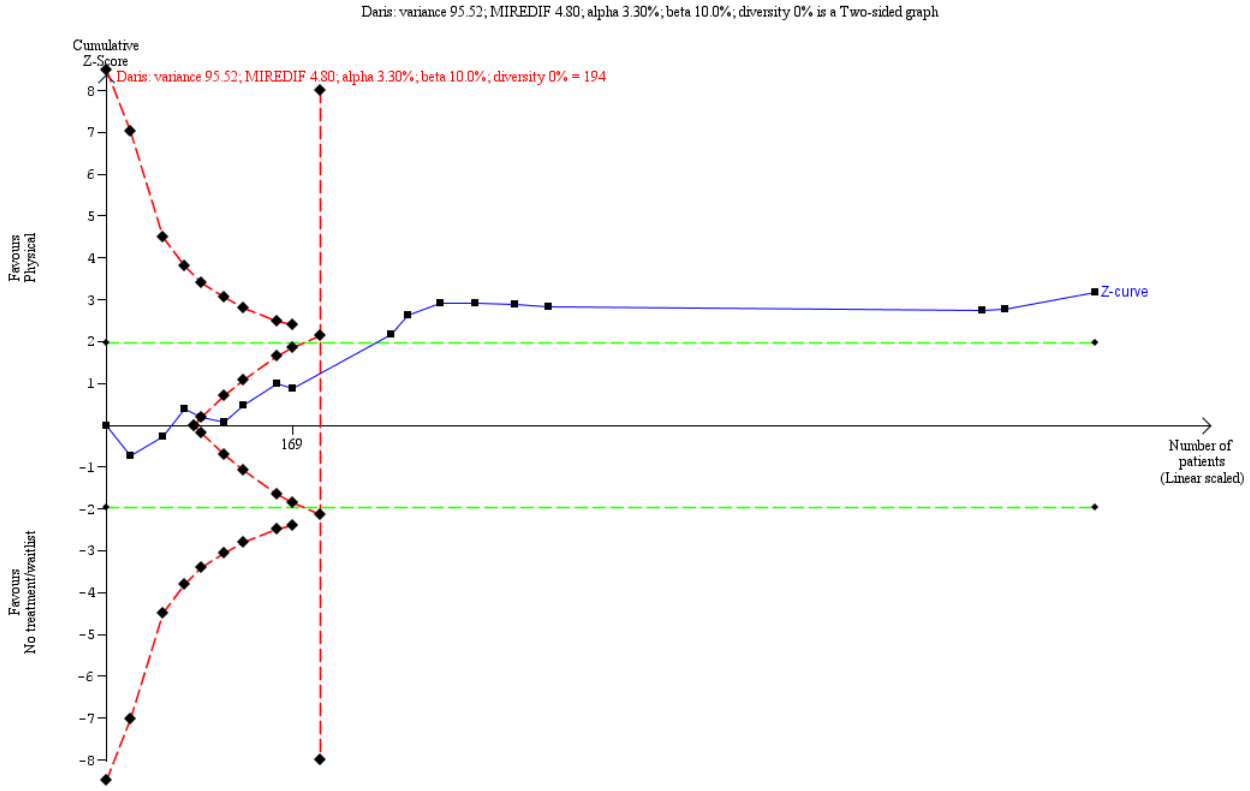
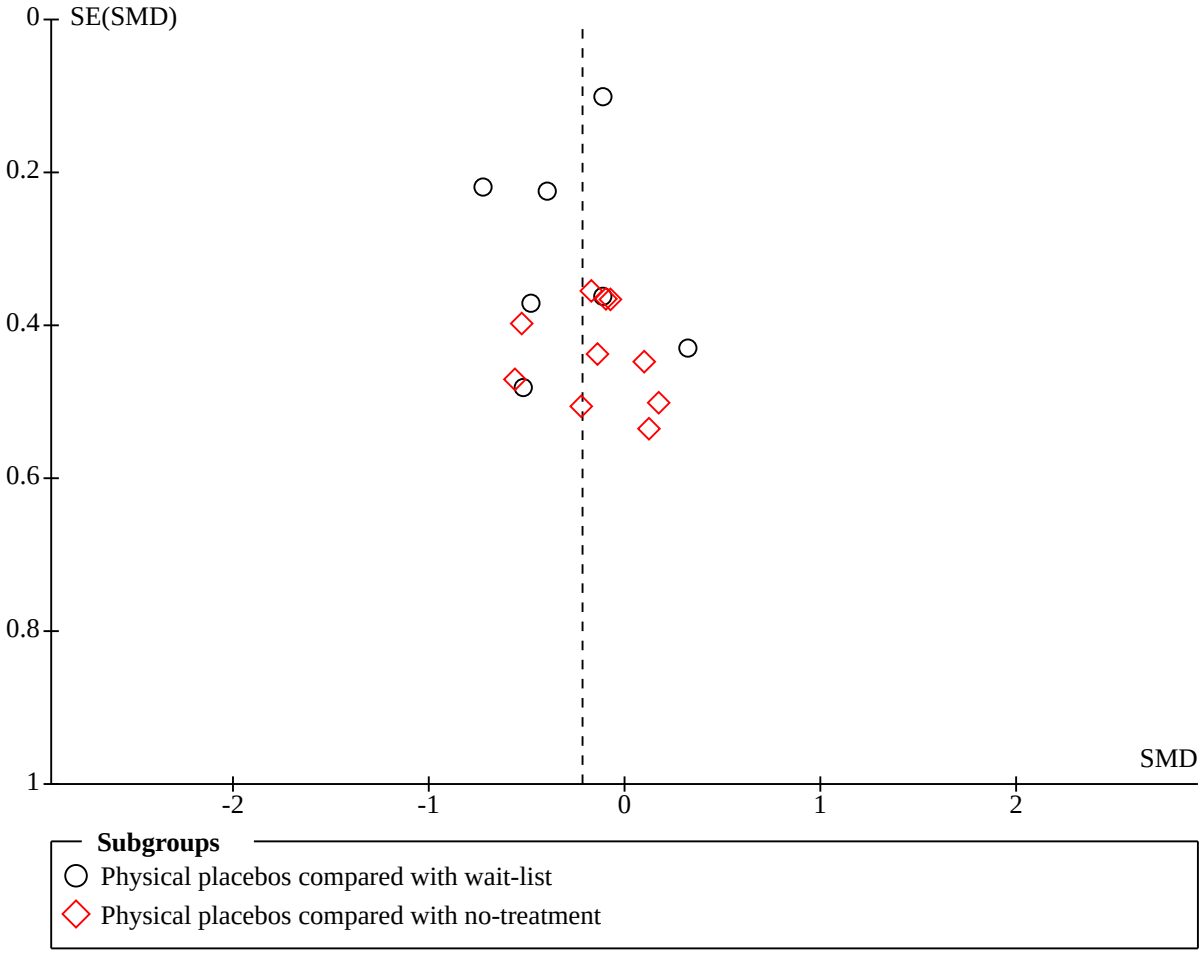


Figure 14. Funnel plot of comparison: 11.1 Efficacy of physical placebos compared with wait-list/no-treatment for continuous data.



Appendix 5. Subgroup analyses

Analyses for specific mental health disorders

Differences in outcomes measuring benefits between usual care versus wait-list or no-treatment for specific mental health disorders

There were not at least three trials for a specific mental health disorders regarding usual care.

Differences in outcomes measuring benefits between all placebos versus wait-list or no treatment for specific mental health disorders

All placebos compared with wait-list or no-treatment for for people with substance use disorders

Four trials compared all placebos with no-treatment (see Table 5).

All placebos compared with wait-list or no-treatment for continuous data: no differences were found for efficacy for patients with substance use disorders comparing all placebos with wait-list or no-treatment (SMD -0.15, 95% CI -0.49 to 0.19; 4 trials, 136 participants; P = 0.37; I^2 = 0%; low-quality evidence; Analysis 12.1). Test for subgroup differences: not applicable.

All placebos compared with wait-list or no-treatment for for people with sleep-wake disorders

11 trials compared all placebos with wait-list or no-treatment (see Table 5).

All placebos compared with wait-list or no-treatment for continuous data:all placebos had a beneficial effect for patients with sleep-wake disorders compared with wait-list or no-treatment (SMD -0.34, 95% CI -0.60 to -0.07; 11 trials, 229 participants; P = 0.01; I^2 = 0%; very low-quality evidence; Analysis 12.2). No differences were found between subgroups: all placebos compared with wait-list (SMD -0.58, 95%

CI -1.10 to -0.05; 3 trials, 70 participants; P = 0.03; I^2 = 14%), and all placebos compared with no-treatment (SMD -0.24, 95% CI -0.55 to 0.08; 8 trials, 159 participants;P = 0.14; I^2 = 0%). Test for subgroup differences: Chi ^{2} = 1.19, df = 1 (P = 0.28; I^2 = 15.7%)

12.3. All placebos compared with wait-list or no-treatment for for people with depression

10 trials compared all placebos with wait-list or no-treatment (Table 5).

All placebos compared with wait-list or no-treatment for continuous data: all placebos had a beneficial effect for patients with depression compared with wait-list or no-treatment (SMD -0.42, 95% CI -0.78 to -0.05; 10 trials, 286 participants; P = 0.03; I^2 = 51%; low-quality evidence; Analysis 12.3) No differences were found between subgroups: all placebos compared with wait-list (SMD -0.65, 95% CI -1.15 to -0.15; 6 trials, 197 participants; P = 0.01; I^2 = 57%), and all placebos compared with no-treatment (SMD -0.05, 95% CI -0.47 to 0.37; 4 trials, 89 participants; P = 0.82; I^2 = 0%). Test for subgroup differences: Chi ^{2} = 3.31, df = 1 (P = 0.07: I^2 = 69.8%)

All placebos compared with wait-list or no-treatment for for people with post-traumatic stress disorder (PTSD)

Four trials comparing all placebos with wait-list or no-treatment (Table 5).

All placebos compared with wait-list or no-treatment for continuous data: all placebos had a beneficial effect for patients with PTSD compared with wait-list or no-treatment (SMD -0.54, 95% CI -1.06 to -0.02; 4 trials, 329 participants; P = 0.04; I^2 = 74%; low-quality evidence; Analysis 12.4). No differences were found between subgroups: all placebos compared with wait-list (SMD -0.75, 95% CI -1.23 to -0.27; 3 trials, 231 participants; P = 0.002; I^2 = 55%), and all placebos compared with no-treatment (SMD -0.08, 95% CI -0.56 to 0.40; 1 trial, 98 participants; I^2 = not applicable; P = 0.74). Test for subgroup differences: Chi ^{2} = 3.70, df = 1 (P = 0.05; I^2 = 73%).

All placebos compared with wait-list or no-treatment for for people with anxiety disorder

Sixteen trials compared all placebos with wait-list or no-treatment (Table 5).

All placebos compared with wait-list or no-treatment for continuous data : All placebos had a beneficial effect for patients with anxiety disorders compared with wait-list or no-treatment (SMD -0.57, 95% CI -0.93 to -0.21; 16 trials, 401 participants; P = 0.002; I^2 = 66%; very low-quality evidence; Analysis 12.5). No differences were found between subgroups: all placebos compared with wait-list (SMD -0.81, 95% CI -1.66 to 0.05; 6 trials, 181 participants; I^2 = 85%; P = 0.07), and all placebos compared with no-treatment (SMD -0.39, 95% CI -0.66 to -0.11; 10 trials, 220 participants; P = 0.006; I^2 = 0%). Test for subgroup differences: Chi ^{2} = 0.83, df = 1 (P = 0.36; I^2 = 0%).

12.6. All placebos compared with wait-list or no-treatment for for people with attention deficit hyperactivity disorder (ADHD) or attention deficit disorder (ADD)

Five trials compared all placebos with no-treatment (Table 5)

All placebos compared with wait-list or no-treatment for continuous data: no differences were found for beneficial effect for patients with ADHD or ADD comparing all placebos with wait-list or no-treatment (SMD -0.31, 95% CI -0.64 to 0.02; 5 trials, 145 participants; P = 0.06; I^2 = 0%; very low-quality evidence; Analysis 12.6). Test for subgroup differences: not applicable.

All placebos compared with wait-list or no-treatment forfor people with neurodegenerative disorders

Four trials compared all placebos with no-treatment (see Table 5).

All placebos compared with wait-list or no-treatment for continuous data: no differences were found for beneficial effect for patients with neurodegenerative disorders comparing all placebos with wait-list or no-treatment (SMD -0.21, 95% CI -0.47 to 0.05; 4 trials, 231 participants; P = 0.12; I^2 = 0%; very low-quality evidence; Analysis 12.7). No differences were found between subgroups: all placebos compared with wait-list (SMD -0.33, 95% CI -0.67 to 0.01; 2 trials, 124 participants; P = 0.06; I^2 = 0%), and all placebos compared with no-treatment (SMD -0.03, 95% CI -0.43 to 0.36; 2 trials, 97 participants; P = 0.86; I^2 = 0%). Test for subgroup differences: Chi ^{2} = 1.24, df = 1 (P = 0.27; I^2 = 19.2%).

Differences in outcomes measuring benefits between psychological placebos versus wait-list or no treatment for specific mental health disorders

Psychological placebos compared with wait-list or no-treatment for people with substance use disorders

Three trials compared psychological placebos with no-treatment (Table 5).

Psychological placebos compared with wait-list or no-treatment for continuous data: no differences were found for beneficial effect for patients with substance use disorders comparing psychological placebos with wait-list or no-treatment (SMD -0.19, 95% CI -0.54 to 0.17; 3 trials, 122 participants; P = 0.31; I^2 = 0%; very low-quality evidence; Analysis 13.1). Test for subgroup differences: not applicable.

Psychological placebos compared with wait-list or no-treatment for people with sleep-wake disorders

Seven trials compared psychological placebos with wait-list or no-treatment (Table 5).

Psychological placebos compared with wait-list or no-treatment for continuous data: psychological placebos had a beneficial effect for patients with sleep-wake disorders compared with wait-list or no-treatment (SMD -0.44, 95% CI -0.76 to -0.12; 7 trials, 159 participants; P = 0.007; I^2 = 0%; low-quality evidence; Analysis 13.2). No differences were found between subgroups: psychological placebos compared with wait-list (SMD -0.63, 95% CI -1.49 to 0.24; 2 trials, 52 participants; P = 0.15; I^2 = 56%), and psychological placebos compared with no-treatment (SMD -0.37, 95% CI -0.75 to 0.02; 5 trials, 107 participants; P = 0.06; I^2 = 0%). Test for subgroup differences: Chi ^{2} = 0.29, df = 1 (P = 0.59; I^2 = 0%).

13.3. Psychological placebos compared with wait-list or no-treatment for people with depression

Four trials compared psychological placebos with wait-list or no-treatment (Table 5).

Psychological placebos compared with wait-list or no-treatment for continuous data: no differences were found for beneficial effect for patients with depression comparing psychological placebos with wait-list or no-treatment (SMD -0.66, 95% CI -1.56 to 0.23; 4 trials, 86 participants; P = 0.15; I^2 = 73%; very low-quality evidence; Analysis 13.3). There were differences between subgroups: psychological placebos compared with wait-list (SMD -1.01, 95% CI -1.82 to -0.21; 3 trials, 66 participants; P = 0.01; I^2 = 55%), and psychological placebos compared with no-treatment (SMD 0.34, 95% CI -0.55 to 1.22; 1 trial, 20 participants; P = 0.46; I^2 = not applicable). Test for subgroup differences: Chi ^{2} = 4.90, df = 1 (P = 0.03, I^2 = 79.6%).

Psychological placebos compared with wait-list or no-treatment for people with post-traumatic stress disorder (PTSD)

Three trials compared psychological placebos with wait-list or no-treatment (Table 5).

Psychological placebos compared with wait-list or no-treatment for continuous data: psychological placebos had a beneficial effect for patients with PTSD compared with wait-list or no-treatment (SMD -0.75, 95% CI -1.23 to -0.27; 3 trials, 231 participants; P = 0.002; I^2 = 55%; very low-quality evidence; Analysis 13.4). Test for subgroup differences: not applicable.

Psychological placebos compared with wait-list or no-treatment for people with anxiety disorders

Eleven trials compared psychological placebos with wait-list or no-treatment (Table 5).

Psychological placebos compared with wait-list or no-treatment for continuous data: psychological placebos had a beneficial effect for patients with anxiety disorders compared with wait-list or no-treatment (SMD -0.71, 95% CI -1.19 to -0.22; 11 trials, 258 participants; P = 0.005; I^2 = 70%; very low-quality evidence; Analysis 13.5). No differences were found between subgroups: psychological placebos compared with wait-list (SMD -0.94, 95% CI -1.83 to -0.05; 6 trials, 157 participants; P = 0.04); I^2 = 84%), and psychological placebos compared with no-treatment (SMD -0.43, 95% CI -0.83 to -0.04; 5 trials, 101 participants; P = 0.03; I^2 = 0%). Test for subgroup differences: Chi ^{2} = 1.04, df = 1 (P = 0.31; I^2 = 3.7%).

Differences in outcomes measuring benefits between pharmacological placebos versus wait-list or no treatment for specific mental health disorders

Pharmacological placebos compared with wait-list or no-treatment for people with attention deficit hyperactivity disorder (ADHD) or attention deficit disorder (ADD)

Three trials compared pharmacological placebos with no-treatment(Table 5).

Pharmacological placebos compared with wait-list or no-treatment for continuous data: no differences were found for beneficial effect for patients with ADHD or ADD comparing pharmacological placebos with wait-list or no-treatment (SMD -0.10, 95% CI -0.59 to 0.39; 3 trials, 64 participants; P = 0.68; I^2 = 0%; low-quality evidence; Analysis 14.1). Test for subgroup differences: not applicable.

Differences in outcomes measuring benefits between physical placebos versus wait-list or no treatment for specific mental health disorders

Physical placebos compared with wait-list or no-treatment for people with sleep-wake disorders

Four trials comparing all placebos with wait-list or no-treatment (Table 5).

Physical placebos compared with wait-list or no-treatment for continuous data: no differences were found for beneficial effect for patients with sleep-wake disorders comparing physical placebos with wait-list or no-treatment (SMD -0.11, 95% CI -0.58 to 0.36; 4 trials, 70 participants; P = 0.65; I^2 = 0%; very low-quality evidence; Analysis 15.1). No differences were found between subgroups: physical placebos compared with wait-list (SMD -0.52, 95% CI -1.46 to 0.43; 1 trial, 18 participants; P = 0.28; I^2 = not applicable), and physical placebos compared with no-treatment (SMD 0.03, 95% CI -0.52 to 0.57; 3 trials, 52 participants; P = 0.93; I^2 = 0%). Test for subgroup differences: Chi ^{2} = 0.95, df = 1 (P = 0.33; I^2 = 0%).

Physical placebos compared with wait-list or no-treatment for people with depression

Four trials compared physical placebos with wait-list or no-treatment (Table 5).

Physical placebos compared with wait-list or no-treatment for continuous data: no differences were found for beneficial effect for patients with depression comparing physical placebos with wait-list or no-treatment (SMD -0.24, 95% CI -0.73 to 0.25; 4 trials, 162 participants; P = 0.12; I^2 = 0%; very low-quality evidence; Analysis 15.2). No differences were found between subgroups: physical placebos compared with wait-list (SMD -0.27, 95% CI -1.28 to 0.75; 2 trials, 111 participants; P = 0.61; I^2 = 79%), and physical placebos compared with no-treatment (SMD -0.10, 95% CI -0.65 to 0.45; 2 trials, 51 participants; P = 0.72; I^2 = 0%). Test for subgroup differences: Chi ^{2} = 0.08, df = 1 (P = 0.78; I^2 = 0%).

Physical placebos compared with wait-list or no-treatment for people with anxiety disorders

Three trials compared physical placebos with wait-list or no-treatment (Table 5).

Physical placebos compared with wait-list or no-treatment for continuous data: no differences were found for beneficial effect for patients with anxiety disorders comparing physical placebos with wait-list or no-treatment (SMD -0.36, 95% CI -0.82 to 0.09; 3 trials, 84 participants; P = 0.002; I^2 = 66%; very low-quality evidence; Analysis 15.3). No differences were found between subgroups: physical placebos compared with wait-list (SMD -0.11, 95% CI -0.82 to 0.60; 1 trial, 36 participants; P = 0.76; I^2 = not applicable), and physical placebos compared with no-treatment (SMD -0.54, 95% CI -1.14 to 0.06; 2 trials, 48 participants; P = 0.08; I^2 = 0%). Test for subgroup differences: Chi ^{2} = 0.82, df = 1 (P = 0.36, I^2 = 0%).

Ordinary Subgroup analyses

Subgroup analyses were conducted if two or more trials were included in each group. Therefore, some of the predefined subgroup analyses were not possible (see Table 6).

Subgroup analyses for all placebos

Type of active intervention (continuous data). Comparing psychological interventions to pharmacological, physical, and other or combination interventions

No differences were found between subgroups: psychological interventions (SMD -0.48, 95% CI -0.66 to -0.30; 38 trials, 1215 participants; I^2 = 54%; P < 0.0001), pharmacological interventions (SMD -0.29, 95% CI -0.62 to 0.04; 7 trials, 148 participants; I^2 = 0%; P = 0.08), physical interventions (SMD -0.21, 95% CI -0.35 to -0.07; 13 trials, 798 participants; I^2 = 0%; P = 0.004) and other or combination interventions (SMD -0.32, 95% CI -0.59 to -0.04; 6 trials, 252 participants; I^2 = 0%; P = 0.02). Test for subgroup differences: Chi ^{2} = 5.46, df = 3 (P = 0.14), I^2 = 45.1%; Analysis 16.1)

Type of active intervention (dichotomous data). Comparing psychological interventions to pharmacological, physical, and other or combination interventions

No differences were found between subgroups: psychological interventions (RR not estimable; 1 trial, 19 participants; I^2 = not applicable; P = not applicable), pharmacological interventions (RR 1.04, 95% CI 0.74 to 1.48; 7 trials, 316 participants; I^2 = 49%; P = 0.82), and other or combination interventions (RR 0.99, 95% CI 0.58 to 1.70; 1 trial, 50 participants; I^2 = not applicable; P = 0.98). Test for subgroup differences: Chi ^{2} = 0.02, df = 1 (P = 0.89), I^2 = 0%; Analysis 16.2)

Risk of bias (continuous data). Comparing trials of low risk of bias to trial of high risk of bias

Test for subgroup differences: not applicable (Analysis 16.3)

Risk of bias (dichotomous data). Comparing trials of low risk of bias to trial of high risk of bias

Test for subgroup differences: not applicable (Analysis 16.4)

Type of outcome domain (continuous data). Comparing blinded observer-reported to non-blinded observer-reported and patient-reported outcomes

There were differences between subgroups: blinded observer-reported (SMD -0.50, 95% CI -0.70 to -0.31; 29 trials, 1046 participants; I^2 = 50%; P < 0.00001), non-blinded observer-reported (SMD -0.54, 95% CI -0.95 to -0.13; 2 trials, 96 participants; I^2 = 0%; P = 0.009), and patient-reported (SMD -0.19, 95% CI -0.35 to -0.03; 25 trials, 1063 participants; I^2 = 26%; P = 0.02). Test for subgroup differences: Chi ^{2} = 6.73, df = 2 (P = 0.03), I^2 = 70.3%; Analysis 16.5)

Type of outcome domain (dichotomous data). Comparing blinded observer-reported to non-blinded observer-reported and patient-reported outcomes

No differences were found between subgroups: blinded observer-reported (RR 0.99, 95% CI 0.60 to 1.64; 3 trials, 113 participants; I^2 = 0%; P = 0.98), and patient-reported (RR 1.00, 95% CI 0.60 to 1.65; 2 trials, 76 participants; I^2 = 73%; P = 0.99). Test for subgroup differences: Chi ^{2} = 0.00, df = 1 (P = 1.00), I^2 = 0%; Analysis 16.6)

Awareness of placebo intervention (continuous data). Comparing open placebos to closed placebos

No differences were found between subgroups: open placebo (SMD -0.51, 95% CI -1.41 to 0.38; 1 trial, 20 participants; I^2 = not applicable; P = 0.26), and closed placebos (SMD -0.37, 95% CI -0.49 to -0.25; 64 trials, 2426 participants; I^2 = 42%; P < 0.0001). Test for subgroup differences: Chi ^{2} = 0.10, df = 1 (P = 0.75), I^2 = 0%; Analysis 16.7)

Awareness of placebo intervention (dichotomous data). Comparing open placebos to closed placebos

Test for subgroup differences: not applicable (Analysis 16.8)

The trial objective (continuous data). Comparing a trial's objective is clearly to assess the effects of placebo to trials without no such objectives

Test for subgroup differences: not applicable (Analysis 16.9)

The trial objective (dichotomous data). Comparing a trial's objective is clearly to assess the effects of placebo to trials without no such objectives

Test for subgroup differences: not applicable (Analysis 16.10)

Mean age of participants (continuous data). Comparing below 18 years of age to 18 to 50 years of age and to above 50 years of age

No differences were found between subgroups: below 18 years of age (SMD -0.47, 95% CI -0.98 to 0.04; 9 trials, 274 participants; I^2 = 75%; P = 0.07), 18 to 50 years of age (SMD -0.31, 95% CI -0.44 to -0.18; 40 trials, 1735 participants; I^2 = 31%; P < 0.00001), and above 50 years of age (SMD -0.41, 95% CI -0.75 to -0.06; 8 trials, 292 participants; I^2 = 45%; P = 0.02). Test for subgroup differences: Chi ^{2} = 0.54, df = 2 (P = 0.76), I^2 = 0%; Analysis 16.11)

Mean age of participants (dichotomous data). Comparing below 18 years of age to 18 to 50 years of age and to above 50 years of age

No differences were found between subgroups: 18 to 50 years of age (RR 1.07, 95% CI 0.70 to 1.62; 5 trials, 270 participants; I^2 = 53%; P = 0.76), and above 50 years of age (RR 0.73, 95% CI 0.43 to 1.27; 2 trials, 70 participants; I^2 = 0%; P = 0.27). Test for subgroup differences: Chi ^{2} = 1.14, df = 1 (P = 0.29), I^2 = 12.2%; Analysis 16.12)

Duration of treatment (continuous data). Comparing above three months of duration to below three months of duration

No differences were found between subgroups: above three months (SMD -0.68, 95% CI -1.09 to -0.27; 11 trials, 464 participants; I^2 = 74%; P = 0.001), and below three months (SMD -0.32, 95% CI -0.43 to -0.20; 53 trials, 1902 participants; I^2 = 18%; P < 0.00001). Test for subgroup differences: Chi ^{2} = 2.86, df = 1 (P = 0.09), I^2 = 65.1%; Analysis 16.13)

Duration of treatment (dichotomous data). Comparing above three months of duration to below three months of duration

No differences were found between subgroups: above three months (RR 1.45, 95% CI 0.84 to 2.49; 4 trials, 196 participants; I^2 = 40%; P = 0.18), and below three months (RR 0.82, 95% CI 0.68 to 0.98; 4 trials, 170 participants; I^2 = 0%; P < 0.03). Test for subgroup differences: Chi ^{2} = 3.88, df = 1 (P = 0.05), I^2 = 74.2%; Analysis 16.14)

Mental health diagnoses (continuous data). Comparing formal diagnosis according to DSM or ICD to fulfilment of symptoms according to DSM or ICD and population being classified as having a mental health disorder, but full diagnostic criteria not reported

No differences were found between subgroups: formal diagnosis according to DSM/ICD (SMD -0.42, 95% CI -0.62 to -0.21;29 trials, 1256 participants; I^2 = 63%; P < 0.001), fulfil symptoms of disorder according to DSM/ICD while not stating classifications systems (SMD -0.30, 95% CI -0.52 to -0.08; 14 trials, 326 participants; I^2 = 0%; P = 0.008), and population is classified as having a mental health disorder, but full diagnostic criteria not reported (SMD -0.27, 95% CI -0.42 to -0.12; 21 trials, 864 participants; I^2 = 4%; P = 0.0004). Test for subgroup differences: Chi ^{2} = 1.27, df = 2 (P = 0.53), I^2 = 0%; Analysis 16.15)

Mental health diagnoses (dichotomous data). Comparing formal diagnosis according to DSM or ICD to fulfilment of symptoms according to DSM or ICD, while not stating classifications system, and population being classified as having a mental health disorder, but full diagnostic criteria not reported

No differences were found between subgroups: formal diagnosis according to DSM/ICD (RR 1.31, 95% CI 0.72 to 2.37; 4 trials, 220 participants; I^2 = 48%; P = 0.38), fulfil symptoms of disorder according to ICD/DSM while not stating classifications systems (RR 1.01, 95% CI 0.62 to 1.63; 3 trials, 96 participants; I^2 = 31%; P = 0.98), and population is classified as having a mental health disorder, but full diagnostic criteria not provided (RR 0.80, 95% CI 0.65 to 0.99; 2 trials, 69 participants; I^2 = NA; P = 0.04). Test for subgroup differences: Chi ^{2} = 2.67, df = 2 (P = 0.26), I^2 = 25.2%; Analysis 16.16)

Affiliation bias (continuous data). Comparing risk of affiliation, industry and allegiance bias to no risk found of affiliation, industry and allegiance bias

No differences were found between subgroups: risk of affiliation bias (SMD -0.61, 95% CI -0.95 to -0.26; 6 trials, 363 participants; I^2 = 52%; P = 0.005), and no risk of affiliation bias (SMD -0.33, 95% CI -0.45 to -0.21; 59 trials, 2083 participants; I^2 = 34%; P < 0.00001). Test for subgroup differences: Chi ^{2} = 2.22, df = 1 (P = 0.14), I^2 = 55%; Analysis 16.17)

Risk of bias (participants and personnel excluded). Comparing low risk of bias with unclear risk of bias and high risk of bias.

No differences were found between subgroups: Low risk of bias (SMD -0.58, 95% CI -0.85 to -0.32; 3 trials, 230 participants; I^2 = 0%; P < 0.001), unclear risk of bias (SMD -0.12, 95% CI -0.95 to 0.71; 2 trials, 51 participants; I^2 = 52%; P = 0.58) and high risk of bias (SMD -0.36, 95% CI -0.48 to -0.23, 60 trials, 2165 participants; I^2 = 42%; P < 0.00001). Test for subgroup differences: Chi^2 = 2.70, df = 2 (P = 0.26), I^2 = 26.0%; Analysis 16.18)

Imputed data. Comparing available data with intention-to-treat and no-attrition data

No differences were found between subgroups: available data (SMD -0.36, 95% CI -0.48 to -0.24, 63 trials, 2416 participants, I^2 = 19%; P < 0.00001), intention-to-treat (SMD -0.22, 95% CI -0.44 to -0.01, 9 trials, 891 participants, I^2 = 46%; P = 0.04), and no attrition (SMD -0.50, 95% CI -1.00 to 0.00, 9 trials, 289 participants, I^2 = 75%; P = 0.05 Test for subgroup differences: Chi^2 = 5.65, df = 2 (P = 0.06), I^2 = 64%; Analysis 16.19

Subgroup analyses for psychological placebos

Type of active intervention (continuous data). Comparing psychological interventions to pharmacological, physical, and other or combination interventions

No differences were found between subgroups: psychological interventions (SMD -0.48, 95% CI -0.67 to -0.29; 35 trials, 1174 participants; I^2 = 56%; P < 0.0001), and other or combination interventions (SMD -0.51, 95% CI -0.94 to -0.09; 3 trials, 89 participants; I^2 = 0%; P = 0.02). Test for subgroup differences: Chi ^{2} = 0.01, df = 1 (P = 0.91), I^2 = 0%; Analysis 17.1)

Risk of bias (continuous data). Comparing trials of low risk of bias to trial of high risk of bias

Test for subgroup differences: not applicable (Analysis 17.2)

Type of outcome domain (continuous data). Comparing blinded observer-reported to non-blinded observer-reported and patient-reported outcomes

No differences were found between subgroups: blinded observer-reported (SMD -0.72, 95% CI -1.08 to -0.36; 14 trials, 512 participants; I^2 = 68%; P < 0.00001), non-blinded observer-reported (SMD -0.54, 95% CI -0.95 to -0.13; 2 trials, 96 participants; I^2 = 0%; P = 0.009), and patient-reported (SMD -0.27, 95% CI -0.47 to -0.08; 18 trials, 907 participants; I^2 = 36%; P = 0.007). Test for subgroup differences: Chi ^{2} = 5.02, df = 2 (P = 0.08), I^2 = 60.2%; Analysis 17.3)

Awareness of placebo intervention (continuous data). Comparing open placebos to closed placebos

Test for subgroup differences: not applicable (Analysis 17.4)

The trial objective (dichotomous data). Comparing a trial's objective is clearly to assess the effects of placebo to trials without no such objectives

Test for subgroup differences: not applicable (Analysis 17.5)

Mean age of participants (continuous data). Comparing below 18 years of age to 18 to 50 years of age and above 50 years of age

No differences were found between subgroups: below 18 years of age (SMD -0.85, 95% CI -2.02 to -0.32; 4 trials, 148 participants; I^2 = 89%; P = 0.16), 18 to 50 years of age (SMD -0.38, 95% CI -0.55 to -0.21; 25 trials, 1233 participants; I^2 = 40%; P < 0.00001), and above 50 years of age (SMD -0.59, 95% CI -1.33 to 0.14; 4 trials, 156 participants; I^2 = 75%; P = 0.11). Test for subgroup differences: Chi ^{2} = 0.88, df = 2 (P = 0.64), I^2 = 0%; Analysis 17.6)

Duration of treatment (continuous data). Comparing above three months of duration to below three months of duration

No differences were found between subgroups: above three months (SMD -1.02, 95% CI -1.68 to -0.36; 6 trials, 270 participants; I^2 = 83%; P = 0.003), and below three months (SMD -0.38, 95% CI -0.53 to -0.22; 32 trials, 1306 participants; I^2 = 33%; P < 0.00001). Test for subgroup differences: Chi ^{2} = 3.42, df = 1 (P = 0.06), I^2 = 70.8%; Analysis 17.7)

Type of psychological placebo (continuous data). Comparing interactive placebo to psychoeducational and exposure placebo

There were differences between subgroups: interaction placebo (SMD -0.70, 95% CI -1.04 to -0.36; 15 trials, 613 participants; I^2 = 70%; P < 0.0001), psychoeducational placebo (SMD -0.14, 95% CI -0.36 to 0.07; 9 trials, 329 participants; I^2 = 0%; P = 0.20), and exposure placebo (SMD -0.48, 95% CI -0.71 to -0.24; 14 trials, 321 participants; I^2 = 6%; P < 0.0001). Test for subgroup differences: Chi ^{2} = 8.68, df = 2 (P = 0.01), I^2 = 76.9%; Analysis 17.8)

Mode of psychological placebo (continuous data). Comparing individual treatment to group and a combination of both

No differences were found between subgroups: individual treatment (SMD -0.31, 95% CI -0.49 to -0.12; 19 trials, 1151 participants; I^2 = 45%; P = 0.001), group treatment (SMD -0.85, 95% CI -1.30 to -0.39; 13 trials, 308 participants; I^2 = 70%; P = 0.0003), and combination of individual and group (SMD -0.45, 95% CI -0.83 to -0.08; 3 trials, 112 participants; I^2 = 0%; P = 0.02). Test for subgroup differences: Chi ^{2} = 4.62, df = 2 (P = 0.10), I^2 = 56.7%; Analysis 17.9)

Mental health diagnoses (continuous data). Comparing formal diagnosis according to DSM or ICD to fulfilment of symptoms according to DSM or ICD, while not stating classifications system, and population being classified as having a mental health disorder, but full diagnostic criteria not reported

No differences were found between subgroups: formal diagnosis according to DSM/ICD (SMD -0.53, 95% CI -0.86 to -0.21; 16 trials, 767 participants; I^2 = 76%; P = 0.001), fulfil symptoms of a disorder according to DSM/ICD while not stating classifications systems (SMD -0.39, 95% CI -0.66 to -0.12; 9 trials, 218 participants; I^2 = 0%; P = 0.005), and population is classified as having a mental health disorder, but full diagnostic criteria not reported (SMD -0.37, 95% CI -0.59 to -0.15; 14 trials, 671 participants; I^2 = 19%; P = 0.0008). Test for subgroup differences: Chi ^{2} = 0.70, df = 2 (P = 0.70), I^2 = 0%; Analysis 17.10)

Affiliation bias (continuous data). Comparing risk of affiliation, industry and allegiance bias to no risk found of affiliation, industry and allegiance bias

No differences were found between subgroups: risk of affiliation bias (SMD -0.61, 95% CI -0.95 to -0.26; 6 trials, 363 participants; I^2 = 52%; P = 0.005), and no risk of affiliation bias (SMD -0.33, 95% CI -0.45 to -0.21; 59 trials, 2083 participants; I^2 = 34%; P < 0.00001). Test for subgroup differences: Chi ^{2} = 2.22, df = 1 (P = 0.14), I^2 = 55%; Analysis 17.11)

Imputed data. Comparing available data with intention-to-treat and no-attrition data.

No differences were found between subgroups: available data (SMD -0.50, 95% CI -0.68 to -0.33, 27 trials, 842 participants, I^2 = 27%; P < 0.00001), intention-to-treat (SMD -0.18, 95% CI -0.39 to -0.03, 6 trials, 682 participants, I^2 = 30%; P = 0.09), and no attrition (SMD -0.75, 95% CI -2.11 to 0.62, 4 trials, 102 participants, I^2 = 90%; P = 0.28. Test for subgroup differences: Chi^2 = 5.65, df = 2 (P = 0.28), I^2 = 64.6%; Analysis 17.12

Subgroup analyses for pharmacological placebos

Type of active intervention (continuous data). Comparing psychological interventions to pharmacological, physical, and other or combination interventions

No differences were found between subgroups: psychological interventions (SMD 0.44, 95% CI -0.26 to 1.13; 1 trial, 33 participants; I^2 = not applicable; P = 0.22), pharmacological interventions (SMD -0.29, 95% CI -0.62 to 0.04; 7 trials, 148 participants; I^2 = 0%; P = 0.08), and other or combination interventions (SMD -0.08, 95% CI -0.56 to 0.40; 1 trial, 98 participants; I^2 = not applicable; P = 0.74). Test for subgroup differences: Chi ^{2} = 3.55, df = 2 (P = 0.17), I^2 = 43.7%; Analysis 18.1)

Type of active intervention (dichotomous data). Comparing psychological interventions to pharmacological, physical, and other or combination interventions

No differences were found between subgroups: pharmacological interventions (RR 1.09, 95% CI 0.71 to 1.67; 7 trials, 316 participants; I^2 = 66%; P = 0.71), and other or combination interventions (RR 0.99, 95% CI 0.58 to 1.70; 1 trial, 50 participants; I^2 = not applicable; P = 0.98). Test for subgroup differences: Chi ^{2} = 0.06, df = 1 (P = 0.80), I^2 = 0%; Analysis 18.2)

Risk of bias (continuous data). Comparing trials of low risk of bias to trial of high risk of bias

Test for subgroup differences: not applicable (Analysis 18.3)

Risk of bias (dichotomous data). Comparing trials of low risk of bias to trial of high risk of bias

Test for subgroup differences: not applicable (Analysis 18.4)

Type of outcome domain (continuous data). Comparing blinded observer-reported to non-blinded observer-reported and patient-reported outcomes

No differences were found between subgroups: blinded observer-reported (SMD -0.24, 95% CI -0.53 to 0.05; 7 trials, 220 participants; I^2 = 0%; P = 0.10), and patient-reported (SMD 0.18, 95% CI -0.36 to 0.73; 2 trials, 59 participants; I^2 = 11%; P = 0.51). Test for subgroup differences: Chi ^{2} = 1.82, df = 1 (P = 0.18), I^2 = 45.0%; Analysis 18.5)

Type of outcome domain (dichotomous data). Comparing blinded observer-reported to non-blinded observer-reported and patient-reported outcomes

No differences were found between subgroups: blinded observer-reported (RR 0.99, 95% CI 0.60 to 1.64; 2 trials, 94 participants; I^2 = 0%; P = 0.98), and patient-reported (RR 1.00, 95% CI 0.60 to 1.65; 2 trials, 76 participants; I^2 = 73%; P = 0.99). Test for subgroup differences: Chi ^{2} = 0.00, df = 1 (P = 1.00), I^2 = 0%; Analysis 18.6)

Awareness of placebo intervention (continuous data). Comparing open placebos to closed placebos

No differences were found between subgroups: open placebo (SMD -0.51, 95% CI -1.41 to 0.38; 1 trial, 20 participants; I^2 = not applicable; P = 0.26), and closed placebos (SMD -0.11, 95% CI -0.37 to 0.16; 8 trials, 259 participants; I^2 = 0%; P = 0.43). Test for subgroup differences: Chi ^{2} = 0.73, df = 1 (P = 0.39), I^2 = 0%; Analysis 18.7)

Awareness of placebo intervention (dichotomous data). Comparing open placebos to closed placebos

Test for subgroup differences: not applicable (Analysis 18.8)

The trial objective (dichotomous data). Comparing a trial's objective is clearly to assess the effects of placebo to trials without no such objectives

Test for subgroup differences: not applicable (Analysis 18.9)

The trial objective (dichotomous data). Comparing a trial's objective is clearly to assess the effects of placebo to trials withoutno such objectives

Test for subgroup differences: not applicable (Analysis 18.10)

Mean age of participants (continuous data). Comparing below 18 years of age to 18 to 50 years of age and above 50 years of age

No differences were found between subgroups: below 18 years of age (SMD -0.10, 95% CI -0.59 to 0.39; 3 trials, 64 participants; I^2 = 0%; P = 0.68), and 18 to 50 years of age (SMD -0.19, 95% CI -0.68 to 0.31; 4 trials, 177 participants; I^2 = 50%; P = 0.46). Test for subgroup differences: Chi ^{2} = 0.88, df = 2 (P = 0.64), I^2 = 0%; Analysis 18.11)

Mean age of participants (dichotomous data). Comparing below 18 years of age to 18 to 50 years of age and above 50 years of age

No differences were found between subgroups: 18 to 50 years of age (RR 1.07, 95% CI 0.70 to 1.62; 5 trials, 270 participants; I^2 = 53%; P = 0.76), and above 50 years of age (RR 0.73, 95% CI 0.43 to 1.27; 2 trials, 70 participants; I^2 = 0%; P = 0.27), Test for subgroup differences: Chi ^{2} = 1.14, df = 1 (P = 0.29), I^2 = 12.2%; Analysis 18.12)

Duration of treatment (continuous data). Comparing above three months of duration to below three months of duration

No differences were found between subgroups: above three months (SMD -0.21, 95% CI -0.56 to 0.14; 4 trials, 164 participants; I^2 = 0%; P = 0.23), and below three months (SMD -0.06, 95% CI -0.43 to 0.31; 5 trials, 115 participants; I^2 = 0%; P = 0.77). Test for subgroup differences: Chi ^{2} = 0.35, df = 1 (P = 0.55), I^2 = 0%; Analysis 18.13)

Duration of treatment (dichotomous data). Comparing above three months of duration to below three months of duration

No differences were found between subgroups: above three months (RR 1.45, 95% CI 0.84 to 2.49; 4 trials, 196 participants; I^2 = 40%; P = 0.18), and below three months (RR 0.82, 95% CI 0.68 to 0.98; 4 trials, 170 participants; I^2 = 0%; P = 0.03). Test for subgroup differences: Chi ^{2} = 3.88, df = 1 (P = 0.05), I^2 = 74.2%; Analysis 18.14)

Mental health diagnoses (continuous data). Comparing formal diagnosis according to DSM or ICD to fulfilment of symptoms according to DSM or ICD, while not stating classifications system, and population being classified as having a mental health disorder, but full diagnostic criteria not reported

No differences were found between subgroups: formal diagnosis according to DSM/ICD (SMD -0.25, 95% CI -0.57 to 0.06; 5 trials, 188 participants; I^2 = 0%; P = 0.12), fulfil symptoms of disorder ICD/DSM while not stating classifications systems (SMD -0.07, 95% CI -0.69 to 0.56; 2 trials, 40 participants; I^2 = 0%; P = 0.83), and population is classified as having a mental health disorder, but full diagnostic criteria not reported (SMD 0.12, 95% CI -0.62 to 0.86; 2 trials, 51 participants; I^2 = 40%; P = 0.75). Test for subgroup differences: Chi ^{2} = 0.97, df = 2 (P = 0.62), I^2 = 0%; Analysis 18.15)

Mental health diagnoses (dichotomous data). Comparing formal diagnosis according to DSM or ICD to fulfilment of symptoms according to DSM or ICD, while not stating classifications system, and population being classified as having a mental health disorder, but full diagnostic criteria not reported

No differences were found between subgroups: formal diagnosis according to DSM/ICD (RR 1.31, 95% CI 0.72 to 2.37; 4 trials, 220 participants; I^2 = 48%; P = 0.38), and fulfil symptoms of disorder ICD/DSM while not stating classifications systems (RR 1.01, 95% CI 0.62 to 1.63; 3 trials, 96 participants; I^2 = 31%; P = 0.98). Test for subgroup differences: Chi ^{2} = 0.45, df = 1 (P = 0.50), I^2 = 0%; Analysis 18.16)

Imputed data. Comparing available data with intention-to-treat and no-attrition data. +

No differences were found between subgroups: available able (SMD -0.17, 95% CI -0.49 to 0.16, 7 trials, 155 participants, I^2 = 3%; P = 0.31), intention-to-treat (SMD -0.08, 95% CI -0.56 to 0.40, 1 trial, 98 participants, I^2 = not applicable; P = 0.74), and no attrition (SMD -0.12, 95% CI -0.89 to 0.65, 1 trial, 26 participants, I^2 = not applicable; P = 0.76. Test for subgroup differences: Chi^2 = 0.09, df = 2 (P = 0.96), I^2 = 0%; Analysis 18.17)

Subgroup analyses for physical placebos

Type of active intervention (continuous data). Comparing psychological interventions to pharmacological, physical, and other or combination interventions

No differences were found between subgroups: psychological interventions (SMD -0.26, 95% CI -0.95 to 0.43; 2 trials, 33 participants; I^2 = 0%; P = 0.46), physical interventions (SMD -0.21, 95% CI -0.35 to -0.07; 13 trials, 798 participants; I^2 = 0%; P = 0.004), and other or combination interventions (SMD -0.30, 95% CI -0.82 to 0.23; 2 trials, 65 participants; I^2 = 0%; P = 0.26). Test for subgroup differences: Chi ^{2} = 0.13, df = 2 (P = 0.26), I^2 = 0%; Analysis 19.1)

Risk of bias (continuous data). Comparing trials of low risk of bias to trial of high risk of bias

Test for subgroup differences: not applicable (Analysis 19.2)

Type of outcome domain (continuous data). Comparing blinded observer-reported to non-blinded observer-reported and patient-reported outcomes

No differences were found between subgroups: blinded observer-reported (SMD -0.42, 95% CI -0.64 to -0.19; 8 trials, 314 participants; I^2 = 0%; P = 0.0004), and patient-reported (SMD -0.00, 95% CI -0.40 to 0.40; 5 trials, 97 participants; I^2 = 0%; P = 1.00). Test for subgroup differences: Chi ^{2} = 3.11, df = 1 (P = 0.08), I^2 = 67.8%; Analysis 19.3)

Awareness of placebo intervention (continuous data). Comparing open placebos to closed placebos

Test for subgroup differences: not applicable (Analysis 19.4)

The trial objective (continuous data). Comparing a trial's objective is clearly to assess the effects of placebo to trials without no such objectives

Test for subgroup differences: not applicable (Analysis 19.5)

Mean age of participants (continuous data). Comparing below 18 years of age to 18 to 50 years of age and above 50 years of age

No differences were found between subgroups: below 18 years of age (SMD -0.32, 95% CI -0.82 to 0.19; 2 trials, 62 participants; I^2 = 0%; P = 0.22), 18 to 50 years of age (SMD -0.26, 95% CI -0.48 to -0.04; 11 trials, 325 participants; I^2 = 0%; P = 0.02), and above 50 years of age (SMD -0.40, 95% CI -0.77 to -0.03; 3 trials, 116 participants; I^2 = 0%; P = 0.03). Test for subgroup differences: Chi ^{2} = 0.40, df = 2 (P = 0.82), I^2 = 0%; Analysis 19.6)

Duration of treatment (continuous data). Comparing above three months of duration to below three months of duration

No differences were found between subgroups: above three months (SMD -0.48, 95% CI -1.21 to 0.25; 1 trial, 30 participants; I^2 = not applicable; P = 0.20), and below three months (SMD -0.21, 95% CI -0.35 to -0.08; 17 trials, 896 participants; I^2 = 0%; P = 0.002). Test for subgroup differences: Chi ^{2} = 0.49, df = 1 (P = 0.48), I^2 = 0%; Analysis 19.7)

Type of physical placebo (continuous data). Comparing acupuncture or acupressure placebo to exercise and relaxation, technical device and electromagnetic stimulation placebo

No differences were found between subgroups: acupuncture or acupressure placebo (SMD -0.42, 95% CI -0.73 to -0.11; 5 trials, 237 participants; I^2 = 21%; P = 0.008), exercise and relaxation placebo (SMD -0.21, 95% CI -0.70 to 0.27; 3 trials, 67 participants; I^2 = 0%; P = 0.39), technical device placebo (SMD -0.23, 95% CI -0.60 to 0.14; 5 trials, 119 participants; I^2 = 0%; P = 0.22), and electromagnetic stimulation

placebo (SMD -0.02, 95% CI -0.57 to 0.54; 2 trials, 50 participants; I^2 = 0%; P = 0.95). Test for subgroup differences: Chi ^{2} = 1.80, df = 3 (P = 0.62), I^2 = 0%; Analysis 19.8)

Mental health diagnoses (continuous data). Comparing formal diagnosis according to DSM or ICD to fulfilment of symptoms according to DSM or ICD, while not stating classifications system, and population being classified as having a mental health disorder, but full diagnostic criteria not reported

No differences were found between subgroups: formal diagnosis according to DSM/ICD (SMD -0.29, 95% CI -0.56 to -0.02; 7 trials, 293 participants; I^2 = 19%; P = 0.04), fulfil symptoms of disorder ICD/DSM while not stating classifications systems (SMD -0.15, 95% CI -0.62 to 0.33; 3 trials, 68 participants; I^2 = 0%; P = 0.55), and population is classified as having a mental health disorder, but full diagnostic criteria not reported (SMD -0.32, 95% CI -0.66 to 0.02; 6 trials, 142 participants; I^2 = 0%; P = 0.06). Test for subgroup differences: Chi ^{2} = 0.36, df = 2 (P = 0.84), I^2 = 0%; Analysis 19.9)

Imputed data. Comparing available data with intention-to-treat and no-attrition data

No differences were found between subgroups: available data (SMD -0.13, 95% CI -0.40 to 0.13; 10 trials, 231 participants, I^2 = 0 %; P = 0.32), intention-to-treat (SMD -0.27, 95% CI -1.28 to 0.75;2 trials, 111 participants, I^2 = 79; P = 0.6), and no attrition (SMD -0.40, -0.71 to -0.08; 4 trials, 161 participants, I^2 = 0%; P = 0.01. Test for subgroup differences: Chi^2 = 1.59, df = 2 (P = 0.45), I^2 = 0%; Analysis 19.10)

Subgroup analyses for usual care

Type of active intervention (continuous data). Comparing psychological interventions to pharmacological, physical, and other or combination interventions

No differences were found between subgroups: psychological interventions (SMD -0.14, 95% CI -0.39 to 0.10; 4 trials, 346 participants; I^2 = 18%; P 0 0.25), physical interventions (SMD -1.10, 95% CI -1.42 to -0.79; 1 trial, 177 participants; I^2 = not applicable; P < 0.00001). Test for subgroup differences: Chi ^{2} = 22.02, df = 1 (P < 0.00001), I^2 = 95.5%; Analysis 20.1)

Type of active intervention (dichotomous data). Comparing psychological interventions to pharmacological, physical, and other or combination interventions

Test for subgroup differences: not applicable (Analysis 20.2)

Risk of bias (continuous data). Comparing trials of low risk of bias to trial of high risk of bias

Test for subgroup differences: not applicable (Analysis 20.3)

Risk of bias (dichotomous data). Comparing trials of low risk of bias to trial of high risk of bias

Test for subgroup differences: not applicable (Analysis 20.4)

Type of outcome domain (continuous data). Comparing blinded observer-reported to non-blinded observer-reported and patient-reported outcomes

No differences were found between subgroups: wait-list (SMD -0.53, 95% CI -1.17 to 0.10; 3 trials, 443 participants; I^2 = 91%; P = 0.10), and no-treatment (SMD 0.08, 95% CI -0.38 to 0.53; 2 trials, 80 participants; I^2 = 0%; P = 0.74). Test for subgroup differences: Chi ^{2} = 2.33, df = 1 (P = 0.13), I^2 = 57.1%; Analysis 20.6)

Mean age of participants (continuous data). Comparing below 18 years of age to 18 to 50 years of age and above 50 years of age

No differences were found between subgroups: below 18 years of age (SMD -1.10, 95% CI -1.42 to -0.79; 1 trial, 177 participants; I^2 = not applicable; P < 0.00001), 18 to 50 years of age (SMD -0.15, 95% CI -0.46 to 0.17; 3 trials, 316 participants; I^2 = 44%; P = 0.36), and above 50 years of age (SMD -0.06, 95% CI -0.82 to 0.70; 1 trial, 30 participants; I^2 = not applicable; P = 0.87). Test for subgroup differences: Chi ^{2} = 19.65, df = 2 (P < 0.00001), I^2 = 89.8%; Analysis 20.6)

Duration of treatment (continuous data). Comparing above three months of duration to below three months of duration

No differences were found between subgroups: above three months (SMD -0.53, 95% CI -1.17 to 0.10; 3 trials, 443 participants; I^2 = 91%; P = 0.10), and below three months (SMD 0.08, 95% CI -0.38 to 0.53; 2 trials, 80 participants; I^2 = 0%; P = 0.74). Test for subgroup differences: Chi ^{2} = 2.33, df = 1 (P = 0.13), I^2 = 57.1%;Analysis 20.7)

Duration of treatment (dichotomous data). Comparing above three months of duration to below three months of duration

Test for subgroup differences: not applicable (Analysis 20.8)

Type of usual care (continuous data). Comparing psychological usual care to pharmacological, physical, and other or combination usual care

No differences were found between subgroups: psychological usual care (SMD -0.44, 95% CI -0.99 to 0.10; 4 trials, 473 participants; I^2 = 87%; P = 0.11), and other or combination of usual care (SMD 0.15, 95% CI -0.41 to 0.72; 1 trial, 50 participants; I^2 = not applicable; P = 0.60). Test for subgroup differences: Chi ^{2} = 2.20, df = 1 (P = 0.14), I^2 = 54.6%; Analysis 20.9)

Type of usual care (dichotomous data). Comparing psychological usual care to pharmacological, physical, and other or combination usual care

No differences were found between subgroups: psychological usual care (RR 0.50, 95% CI 0.20 to 1.26; 1 trial, 26 participants; I^2 = not applicable; P = 0.14), and other or combination of usual care (RR 1.42, 95% CI 1.17 to 1.73; 1 trial, 234 participants; I^2 = not applicable; P = 0.0005). Test for subgroup differences: Chi ^{2} = 4.72, df = 1 (P = 0.03), I^2 = 78.8%; Analysis 20.10)

Standarised usual care (continuous data). Comparing trials of usual care intervention was intentionally standardised or manualised to trials of no standardisation or manualisation

No differences were found between subgroups: standardised or manualised (SMD -0.38, 95% CI -0.94 to 0.18; 4 trials, 493 participants; I^2 = 89%; P = 0.18), and no standardisation or manualisation (SMD -0.06, 95% CI -0.82 to 0.70; 1 trial, 30 participants; I^2 = not applicable; P = 0.87). Test for subgroup differences: Chi ^{2} = 0.43, df = 1 (P = 0.51), I^2 = 0%; Analysis 20.11)

Standarised usual care (dichotomous data). Comparing trials of usual care intervention was intentionally standardised or manualised to trials of no standardisation or manualisation

Test for subgroup differences: not applicable (Analysis 20.12)

Mode of psychological treatment in usual care (continuous data). Comparing individual treatment to group and a combination of both

No differences were found between subgroups: individual treatment (SMD -0.06, 95% CI -0.35 to 0.23; 2 trials, 185 participants; I^2 = 0%; P = 0.70), and group treatment (SMD -0.78, 95% CI -1.43 to -0.12; 2 trials, 288 participants; I^2 = 86%; P = 0.02). Test for subgroup differences: Chi ^{2} = 3.84, df = 2 (P = 0.05), I^2 = 73.9%; Analysis 20.13)

Imputed data. Comparing available data with intention-to-treat and no-attrition data

No differences were found between subgroups: available data (SMD -0.36, 95% CI -0.84 to 0.12; 2 trials, 80 participants, I^2 = 0%; P = 0.84), intention-to-treat (SMD -0.53, 95% CI -1.17 to 0.10; 3 trials, 443 participants, I^2 =91%; P = 0.10). Test for subgroup differences: Chi^2 = 1.76, df = 1 (P = 0.18), I^2 = 43.3%

HISTORY

Protocol first published: Issue 1, 2019

CONTRIBUTIONS OF AUTHORS

EF: protocol development, screening, data extraction, statistical analyses, writing, revisions

AT: protocol development, screening, data extraction, statistical analyses, writing, revisions

LB: screening, data extraction, statistical analyses, revisions

AH: protocol development, writing, revisions

CG: protocol development, statistical analyses, writing, revisions

MK: protocol development, writing, revisions

ES: protocol development, writing, revisions

OJ: protocol development, statistical analyses, writing, revisions

DECLARATIONS OF INTEREST

The review authors have no relevant interests to declare.

SOURCES OF SUPPORT

Internal sources

- Psychiatric Research Unit, Region Zealand Psychiatry, Slagelse, Denmark

  EF, AT, MTK, ES, and OJS we re personally salaried by the institution during the period of this review.
- Center for Evidence-Based Medicine, Odense University Hospital and University of Southern Denmark, Denmark

  AH wa s personally salaried by the institution during the period of this review.
- The Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen, Denmark

  CG wa s personally salaried by the institution during the period of this review.

External sources

- No sources of support provided

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

There were some differences between the protocol and review. These included i) pooling of placebos, ii) dealing with missing data, iii) subgroup analyses, iv) sensitivity analyses, v) selection of studies, vi) an extended decision hierarchy, and vii) no meta-regression analyses. For more information, see Table 6.

INDEX TERMS

Medical Subject Headings (MeSH)

Anxiety; Anxiety Disorders; *Depressive Disorder, Major; *Mental Health; Psychotherapy; Randomized Controlled Trials as Topic

MeSH check words

Humans