Psychological therapies for treatment-resistant depression in adults (Review)

Ijaz S, Davies P, Williams CJ, Kessler D, Lewis G, Wiles N


www.cochranelibrary.com
Psychological therapies for treatment-resistant depression in adults (Review)

Copyright © 2018 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.
Analysis 1.17. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 17 Remission (< 7 on HAMD or < 10 on BDI) medium term (7 to 12 months). 71
Analysis 1.18. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 18 Remission (< 7 on HAMD or < 10 on BDI) long term (longer than 12 months). 71
ADDITIONAL TABLES 72
APPENDICES 75
CONTRIBUTIONS OF AUTHORS 79
DECLARATIONS OF INTEREST 79
SOURCES OF SUPPORT 79
DIFFERENCES BETWEEN PROTOCOL AND REVIEW 80
[Intervention Review]

**Psychological therapies for treatment-resistant depression in adults**

Sharea Ijaz¹, Philippa Davies², Catherine J Williams³, David Kessler³, Glyn Lewis⁴, Nicola Wiles²

¹NIHR CLAHRC West at University Hospitals Bristol NHS Foundation Trust, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. ²Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. ³School of Social and Community Medicine, University of Bristol, Bristol, UK. ⁴UCL Division of Psychiatry, UCL, London, UK

"Deceased"

Contact address: Sharea Ijaz, NIHR CLAHRC West at University Hospitals Bristol NHS Foundation Trust, Population Health Sciences, Bristol Medical School, University of Bristol, Lewins Mead, Whitefriars Building, Bristol, BS1 2NT, UK. sharcaijaz@hotmail.com, s.ijaz@bristol.ac.uk.

**Editorial group:** Cochrane Common Mental Disorders Group.

**Publication status and date:** New, published in Issue 5, 2018.


Copyright © 2018 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

**ABSTRACT**

**Background**

Antidepressants are a first-line treatment for adults with moderate to severe major depression. However, many people prescribed antidepressants for depression don’t respond fully to such medication, and little evidence is available to inform the most appropriate ‘next step’ treatment for such patients, who may be referred to as having treatment-resistant depression (TRD). National Institute for Health and Care Excellence (NICE) guidance suggests that the ‘next step’ for those who do not respond to antidepressants may include a change in the dose or type of antidepressant medication, the addition of another medication, or the start of psychotherapy. Different types of psychotherapies may be used for TRD; evidence on these treatments is available but has not been collated to date.

Along with the sister review of pharmacological therapies for TRD, this review summarises available evidence for the effectiveness of psychotherapies for adults (18 to 74 years) with TRD with the goal of establishing the best ‘next step’ for this group.

**Objectives**

To assess the effectiveness of psychotherapies for adults with TRD.

**Search methods**

We searched the Cochrane Common Mental Disorders Controlled Trials Register (until May 2016), along with CENTRAL, MEDLINE, Embase, and PsycINFO via OVID (until 16 May 2017). We also searched the World Health Organization (WHO) trials portal (ICTRP) and ClinicalTrials.gov to identify unpublished and ongoing studies. There were no date or language restrictions.

**Selection criteria**

We included randomised controlled trials (RCTs) with participants aged 18 to 74 years diagnosed with unipolar depression that had not responded to minimum four weeks of antidepressant treatment at a recommended dose. We excluded studies of drug intolerance. Acceptable diagnoses of unipolar depression were based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR) or earlier versions, International Classification of Diseases (ICD)-10, Feighner criteria, or Research Diagnostic Criteria. We included the following comparisons.
1. Any psychological therapy versus antidepressant treatment alone, or another psychological therapy.

2. Any psychological therapy given in addition to antidepressant medication versus antidepressant treatment alone, or a psychological therapy alone.

Primary outcomes required were change in depressive symptoms and number of dropouts from study or treatment (as a measure of acceptability).

**Data collection and analysis**

We extracted data, assessed risk of bias in duplicate, and resolved disagreements through discussion or consultation with a third person. We conducted random-effects meta-analyses when appropriate. We summarised continuous outcomes using mean differences (MDs) or standardised mean differences (SMDs), and dichotomous outcomes using risk ratios (RRs).

**Main results**

We included six trials (n = 698; most participants were women approximately 40 years of age). All studies evaluated psychotherapy plus usual care (with antidepressants) versus usual care (with antidepressants). Three studies addressed the addition of cognitive-behavioural therapy (CBT) to usual care (n = 522), and one each evaluated intensive short-term dynamic psychotherapy (ISTDP) (n = 60), interpersonal therapy (IPT) (n = 34), or group dialectical behavioural therapy (DBT) (n = 19) as the intervention. Most studies were small (except one trial of CBT was large), and all studies were at high risk of detection bias for the main outcome of self-reported depressive symptoms.

A random-effects meta-analysis of five trials (n = 575) showed that psychotherapy given in addition to usual care (vs usual care alone) produced improvement in self-reported depressive symptoms (MD -4.07 points, 95% confidence interval (CI) -7.07 to -1.07 on the Beck Depression Inventory (BDI) scale) over the short term (up to six months). Effects were similar when data from all six studies were combined for self-reported depressive symptoms (SMD -0.40, 95% CI -0.65 to -0.14; n = 635). The quality of this evidence was moderate. Similar moderate-quality evidence of benefit was seen on the Patient Health Questionnaire-9 Scale (PHQ-9) from two studies (MD -4.66, 95% CI 8.72 to -0.59; n = 482) and on the Hamilton Depression Rating Scale (HAMD) from four studies (MD -3.28, 95% CI -5.71 to -0.85; n = 193).

High-quality evidence shows no differential dropout (a measure of acceptability) between intervention and comparator groups over the short term (RR 0.85, 95% CI 0.58 to 1.24; six studies; n = 698).

Moderate-quality evidence for remission from six studies (RR 1.92, 95% CI 1.46 to 2.52; n = 635) and low-quality evidence for response from four studies (RR 1.80, 95% CI 1.2 to 2.7; n = 556) indicate that psychotherapy was beneficial as an adjunct to usual care over the short term.

With the addition of CBT, low-quality evidence suggests lower depression scores on the BDI scale over the medium term (12 months) (RR -3.40, 95% CI -7.21 to 0.40; two studies; n = 475) and over the long term (46 months) (RR -1.90, 95% CI -3.22 to -0.58; one study; n = 248). Moderate-quality evidence for adjunctive CBT suggests no difference in acceptability (dropout) over the medium term (RR 0.98, 95% CI 0.66 to 1.47; two studies; n = 549) and lower dropout over long term (RR 0.80, 95% CI 0.66 to 0.97; one study; n = 248).

Two studies reported serious adverse events (one suicide, two hospitalisations, and two exacerbations of depression) in 4.2% of the total sample, which occurred only in the usual care group (no events in the intervention group).

An economic analysis (conducted as part of an included study) from the UK healthcare perspective (National Health Service (NHS)) revealed that adjunctive CBT was cost-effective over nearly four years.

**Authors’ conclusions**

Moderate-quality evidence shows that psychotherapy added to usual care (with antidepressants) is beneficial for depressive symptoms and for response and remission rates over the short term for patients with TRD. Medium- and long-term effects seem similarly beneficial, although most evidence was derived from a single large trial. Psychotherapy added to usual care seems as acceptable as usual care alone.

Further evidence is needed on the effectiveness of different types of psychotherapies for patients with TRD. No evidence currently shows whether switching to a psychotherapy is more beneficial for this patient group than continuing an antidepressant medication regimen. Addressing this evidence gap is an important goal for researchers.
Are psychological therapies effective in treating depression that did not get better with previous treatment?

**Review question**

Is psychological therapy an effective treatment for adults with treatment-resistant depression (TRD)?

**Background**

Depression is a common problem often treated with antidepressant medication. However, many people do not get better with antidepressants. These patients may be said to have TRD. For these people, several different treatments can be tried - such as increasing the dose of medicine being taken, adding another medicine, or switching to a new one. Another option is to add or switch to a psychotherapy. Evidence indicates that psychotherapies can help in depression. What we don’t know is whether psychotherapies work in people with TRD. This review aimed to answer this question.

**Search date**

Searches are current up to May 2017.

**Study characteristics**

We included six randomised trials (studies in which participants are allocated at random (by chance) to receive one of the treatments being compared). These trials included 698 people and tested three different types of psychotherapy. All studies looked at whether adding psychotherapy to current medical treatment leads to improvement in depression.

**Study funding sources**

All studies were funded by public research grants.

**Key results**

We found that patients who receive psychotherapy as well as usual care with antidepressants had fewer depressive symptoms and were more often depression-free six months later compared with patients who continued with usual care alone. We are moderately confident of these findings, which means that the true effect of adding CBT may be different from what we found, although findings are likely to be close. We also found that added psychotherapy was as acceptable to patients as usual care alone. Two studies noted similar beneficial effects after 12 months, and one study at 46 months.

Two studies reported harmful effects in people receiving usual care alone (one suicide, two people hospitalised) but none in people receiving psychotherapy in addition to usual care.

**Quality of the evidence**

Because participants were aware of the treatment they had received, and because we identified only a small number of studies, we graded the evidence as moderate in quality for findings at six months and low in quality for long-term results. This assessment might change in the future, if higher-quality research results become available.
# Summary of Findings for the Main Comparison

**Psychotherapy as an adjunct to usual care compared with usual care alone for treatment-resistant depression in adults**

**Patient or population:** adults with treatment-resistant depression  
**Setting:** primary or secondary care  
**Intervention:** psychotherapy with usual care  
**Comparison:** usual care alone

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Anticipated absolute effects* (95% CI)</th>
<th>Relative effect (95% CI)</th>
<th>n° of participants (studies)</th>
<th>Quality of the evidence (GRADE)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Risk with usual care alone</td>
<td>Risk with psychotherapy as an adjunct to usual care</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Self-reported depressive symptoms short term (up to 6 months) - BDI (BDI)**
- Mean depressive symptoms short term (up to 6 months) - BDI was 21.1
  - MD 4.07 lower (7.07 lower to 1.07 lower)
  - MD -0.70 (-7.01 to -1.07)
- 575 (5 RCTs)
- ⊕⊕⊕ MODERATE
  - Third-wave cognitive/behavioural therapies given (individual CBT in 3 studies, group DBT in 1, and individual IPT in 1)

**Self-reported depressive symptoms short term (up to 6 months) - SMD (BDI & PHQ9)**
- Mean depressive symptoms short term (up to 6 months) - BDI was 21.1, and PHQ9 was 14.79
  - SMD 0.4 SD lower (0.65 lower to 0.14 lower)
  - SMD -0.4 (-0.65 to -0.14)
- 635 (6 RCTs)
- ⊕⊕⊕ MODERATE

**Observer-rated depressive symptoms short term (up to 6 months) - PHQ-9**
- Mean depressive symptoms short term (up to 6 months) - PHQ-9 was 14.8
  - MD 4.66 lower (8.72 lower to 0.59 lower)
  - MD -4.66 (-8.72 to -0.59)
- 482 (2 RCTs)
- ⊕⊕⊕ MODERATE
  - One large study from UK and one relatively small one from Canada
<table>
<thead>
<tr>
<th>Psychological therapies for treatment-resistant depression in adults (Review)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer-rated depressive symptoms short term (up to 6 months) - HAMD</td>
</tr>
<tr>
<td>Mean depressive symptoms short term (up to 6 months) - HAMD was 14.76</td>
</tr>
<tr>
<td>MD 3.28 lower (5.71 lower to 0.85 lower)</td>
</tr>
<tr>
<td>MD -3.28 (-5.71 to -0.85)</td>
</tr>
<tr>
<td>193 (4 RCTs)</td>
</tr>
<tr>
<td>⊕⊕⊕⊕ LOW c,d</td>
</tr>
<tr>
<td>Although blinded outcome assessment, 4 small studies each using a different type of psychotherapy: group DBT; ISTDP; CBT; IPT</td>
</tr>
<tr>
<td>Dropout short term (up to 6 months)</td>
</tr>
<tr>
<td>Study population</td>
</tr>
<tr>
<td>RR 0.85 (0.58 to 1.24)</td>
</tr>
<tr>
<td>698 (6 RCTs)</td>
</tr>
<tr>
<td>⊕⊕⊕⊕ HIGH</td>
</tr>
<tr>
<td>Objective outcome; data reported in all studies; Although a proxy for acceptability, it suggests that intervention may be as acceptable as usual care</td>
</tr>
<tr>
<td>Response (50% reduction in depressive symptoms from baseline) short term (up to 6 months)</td>
</tr>
<tr>
<td>Study population</td>
</tr>
<tr>
<td>RR 1.80 (1.20 to 2.69)</td>
</tr>
<tr>
<td>556 (4 RCTs)</td>
</tr>
<tr>
<td>⊕⊕⊕⊕ LOW a,e</td>
</tr>
<tr>
<td>-</td>
</tr>
<tr>
<td>Remission (&lt; 7 on HAMD or &lt; 10 on BDI) short term (up to 6 months)</td>
</tr>
<tr>
<td>Study population</td>
</tr>
<tr>
<td>RR 1.92 (1.46 to 2.52)</td>
</tr>
<tr>
<td>635 (6 RCTs)</td>
</tr>
<tr>
<td>⊕⊕⊕ MODERATEa,b</td>
</tr>
<tr>
<td>One large and 5 small studies comprising mainly women. Third-wave cognitive/behavioural therapies given (individual CBT in 3 studies; individual IPT, ISTDP, and group DBT in 1 study each)</td>
</tr>
</tbody>
</table>

* The risk in the intervention group (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; MD: mean difference; OR: odds ratio; RCT: randomised controlled trial; RR: risk ratio; SMD: standardised mean difference
### GRADE Working Group grades of evidence.

**High quality:** We are very confident that the true effect lies close to that of the estimate of the effect.  
**Moderate quality:** We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.  
**Low quality:** Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.  
**Very low quality:** We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

---

This review (Review) was last updated on [insert date]. It is intended as a summary of the evidence for the effectiveness of psychological therapies for treatment-resistant depression in adults.  

The review includes studies that compare psychological therapies to other treatments or to no treatment, or compare different psychological therapies. The focus is on the effectiveness of psychological therapies for treating depression in adults who have not responded to standard treatment.  

The study interventions included a variety of psychological therapies, such as cognitive-behavioral therapy (CBT), interpersonal therapy (IPT), and mindfulness-based therapy. The studies included in this review were conducted in different settings, such as hospitals, clinics, and community centers. The outcomes assessed included depression severity, mood, and quality of life.

---

*a* Outcome assessment not blind.  
*b* Allocation concealment unclear for one of the two smaller studies.  
*c* Risk of bias due to incomplete outcome data in two of the studies.  
*d* Studies are small. Effects not in the same direction for IPT study (n = 30).  
*e* Reporting bias likely as less frequently reported than remission or mean scores.
BACKGROUND

Description of the condition

It has been predicted that depression will be the leading cause of disability in high-income countries by the year 2030 (Mathers 2005). Severity of depression can be classified on the basis of Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV), criteria as mild (five or more symptoms with minor functional impairment), moderate (symptoms or functional impairment between 'mild' and 'severe'), or severe (most symptoms present and interfering with functioning) (NICe 2009).

Antidepressants are often prescribed as first-line treatment for adults with moderate to severe depression (APA 2010; NICe 2009). In England in 2010, 42.8 million prescriptions for antidepressants were issued at a cost of GBP220 million (The NHS Information Centre 2011). However, two-thirds of people do not respond fully to such pharmacotherapy (Trivedi 2006). Such non-response may result from intolerance to the prescribed medication or non-adherence to the treatment regimen but may also indicate treatment ‘resistance’, whereby treatment of an adequate dose and duration has been given. The World Psychiatric Association provided the earliest definition of treatment-‘resistant’ depression: “an absence of clinical response to treatment with a tricyclic antidepressant at a minimum dose of 150 mg per day of Imipramine (or equivalent drug) for 4 to 6 weeks” (WPA 1974). Subsequently, others suggested more complex classification systems based on non-response to multiple courses of treatment (Fava 2005; Fekadu 2009; Thase 1997), using terms such as ‘treatment-refractory’ depression and ‘antidepressant-resistant’ depression to describe this condition. For the purpose of this review, we will use the term ‘treatment-resistant depression’ as this is the descriptor that has generally represented the broadest definition of the condition. The burden of depression is substantial, and in the UK the average service cost to the National Health Service (NHS) has been estimated as GBP2085 per patient (McCrone 2008). Total cost of services for depression in 2007 was estimated as GBP1.7 billion, although these costs were dwarfed by the cost of lost productivity, which accounted for a further GBP5.8 billion (McCrone 2008). Similar substantial costs have been estimated for the USA, with direct treatment costs estimated at USD26.1 billion and workplace costs at a further USD51.5 billion in the year 2000 (Greenberg 2003). If up to one-third of patients have ‘treatment-resistant’ depression, it is clear that this condition represents a considerable burden to patients, the NHS, and society.

Description of the intervention

First-line treatment for adults with moderate to severe depression commonly consists of an antidepressant (APA 2010; NICe 2009). Five main types of antidepressants are available: tricyclic (TCAs) and related antidepressants; monoamine-oxidase inhibitors (MAOIs); selective serotonin reuptake inhibitors (SSRIs); serotonin and noradrenaline reuptake inhibitors (SNRIs); and noradrenergic and specific serotonin antidepressants (NaSSAs). SSRIs are safer in terms of overdose than TCAs and tend to be better tolerated than antidepressants of other classes. Hence, it is not surprising that SSRIs are the most commonly prescribed antidepressants for treating individuals with depression (Olsson 2009; The NHS Information Centre 2011).

No agreement has been reached on the standard approach for treatment of those whose depression does not respond to antidepressant medication. Guidance published by the American Psychiatric Association (APA 2010) and the National Institute for Health and Care Excellence (NICe 2009) suggests that the ‘next step’ may include increasing the dose of the antidepressant medication, switching to another antidepressant (within the same or in a different pharmacological class), or augmenting treatment via another pharmacological or psychological approach. Psychological therapies that may be given as an adjunct can be broadly categorised into four separate philosophical and theoretical schools: (1) psychodynamic/psychoanalytical (Freud 1949; Jung 1963; Klein 1960); (2) behavioural (Marks 1981; Skinner 1953; Watson 1924); (3) humanistic (Maslow 1943; May 1961; Rogers 1951); and (4) cognitive (Beck 1979; Lazarus 1971). In addition, ‘third wave’ (Hayes 2004; Hayes 2006; Hofmann 2008) and ‘integrative’ (Hollander 2007; Klerman 1984; McCullough 1984; Ryle 1990; Shapiro 1990; Weissman 2007) psychological approaches may be used. Elements of these approaches may overlap or may differ. For example, cognitive-analytical therapy (CAT) incorporates elements from several theoretical schools (Ryle 1990), whereas interpersonal therapy for depression (IPT) is disorder-specific (Klerman 1984). The most influential cognitive approaches have been merged with the behavioural approach to form cognitive-behavioural therapy (CBT) (Beck 1979; Ellis 1962), which is now viewed as a family of therapies that draw upon a common base of cognitive and behavioural models of psychological disorders (Mansell 2008).

How the intervention might work

Psychological therapies such as CBT have been shown to be effective for people with depression (Churchill 2001). When a psychological therapy is given as an adjunct to pharmacological treatment, it is hoped that the benefits gained from these different treatment approaches may be optimised. Mechanisms of action differ among psychological therapies. Cognitive-behavioural therapy targets the person’s unrealistic and unhelpful negative thoughts (“dysfunctional attitudes”) to improve outcomes, whereas behavioural therapy focuses on changing maladaptive patterns of behaviour. In contrast, humanistic therapy seeks to increase an individual’s self-awareness, and psychodynamic therapy focuses on past experiences...
and an understanding of how these events might have influenced the individual and his or her current thoughts and behaviours.

**Why it is important to do this review**

Antidepressants continue to serve as first-line treatment for many people with depression. However, only one-third of people prescribed antidepressants for depression will respond fully to such medication (Trivedi 2006). Evidence suggests that people may prefer psychotherapy to medication for depression (McHugh 2013). Therefore, summarising the evidence for effectiveness of psychological therapies for people with treatment-resistant depression (TRD) is important toward establishing the best 'next step' treatment for this patient group. Several traditional reviews have examined the evidence on treatment of people whose depression has not responded to antidepressant medication alone (e.g. Carvalho 2008; Nierenberg 2007; Papakostas 2009). Systematic reviews on the effectiveness of combination treatment for people with depression have not examined evidence for the treatment-resistant population (Friedman 2004; Pampallona 2004). Others have summarised the evidence for effectiveness of particular treatment strategies for those who have not responded to antidepressants: (1) augmentation as discussed in Carvalho 2007 with lithium - Bauer 1999 - or atypical antipsychotics - Shelton 2008; (2) within- or between-class switches (Papakostas 2008); and (3) psychological treatments (McPherson 2005). One review focused on interventions for older people (≥65 years of age) (Cooper 2011). However, several of these reviews included uncontrolled studies, non-randomised studies, or a combination of these, as well as randomised controlled trials (RCTs) (Carvalho 2007; Cooper 2011; McPherson 2005; Shelton 2008). A previous systematic review of RCTs investigating pharmacological and psychological therapies for people with TRD found no strong evidence to guide the management of such people (Stimpson 2002). However, this review, along with others (e.g. Bauer 1999, which summarised the evidence for lithium up to June 1997), is out-of-date, and several relevant RCTs were published subsequently. Another review of psychotherapies for TRD included four controlled studies of CBT (McPherson 2005); two studies showed benefit derived from CBT, and two found no difference between psychotherapy and control. No agreement has been reached on the definition of 'treatment-resistant depression'. Many studies have defined TRD as 'failure to respond to at least two previous antidepressants'. Given continued reliance upon antidepressants as first-line treatment, we have used a broader and more inclusive definition of treatment resistance - 'non-response to at least four weeks of antidepressant medication' - to help establish the best 'next step' of treatment for the significant number of people whose depression does not respond to antidepressant medication. The rise in antidepressant prescribing along with increased demand for psychotherapy in recent years (BACP 2014; McManus 2000; Middleton 2001; Pincus 1998) means that a review of the evidence for effectiveness of psychological therapies for people with TRD is timely. A connected review is examining pharmacological interventions for TRD (Williams 2013). Together, evidence from these two linked reviews will provide a comprehensive evidence base of the main interventions available for management of TRD, which will inform clinical decision-making with regards to the best 'next step' for adults whose depression has not responded to first-line treatment with medication.

**OBJECTIVES**

To assess the effectiveness of psychotherapies for adults with TRD.

**METHODS**

**Criteria for considering studies for this review**

**Types of studies**

This review includes RCTs and cluster RCTs. This review includes trials using a cross-over design but only data from the first treatment phase. Excluded from this review are trials of any other study design, including quasi-randomised studies and non-randomised studies.

**Types of participants**

**Age range**

Participants must be 18 to 74 years of age. We excluded any study that included some participants younger than 75 years and some older than 74 years if the mean age of participants was over 74 years. Similarly, we excluded any study that included some participants younger than 18 years and some older than 18 years if the mean age of participants was less than 18 years.

**Definition of treatment-resistant depression**

We defined treatment-resistant depression as 'A primary diagnosis of unipolar depression that has not responded (or has only partially responded) to a minimum of four weeks of antidepressant treatment at a recommended dose (at least 150 mg/d imipramine or equivalent antidepressant (e.g. 20 mg/d citalopram)).' We excluded studies that included people who had not responded because of intolerance of antidepressant medication. Although initiatives have sought to improve access to psychological therapies in England and elsewhere, access to psychological...
treatment remains limited and antidepressants are often given as first-line treatment for adults with depression. Therefore, this review does not include studies of interventions intended for those who have not responded to psychological treatment.

**Diagnosis**

Acceptable diagnoses of unipolar depression include those based on criteria from DSM-IV-TR or earlier versions of this publication (APA 2000), International Classification of Diseases (ICD)-10 (WHO 1992), Feighner criteria (Feighner 1972), or Research Diagnostic Criteria (Spitzer 1978). We excluded studies that did not use standardised diagnostic criteria.

**Comorbidities**

Excluded from this review are studies of participants with comorbid schizophrenia or bipolar disorder. Also excluded are studies including participants with both unipolar and bipolar depression unless data are available for the subgroup of unipolar participants. This review includes studies involving participants with comorbid physical conditions or other psychological disorders (e.g. anxiety) for whom psychological therapy was not being primarily used to manage the physical illness, in other words, the focus of treatment was TRD - not the comorbidity.

**Types of interventions**

**Experimental interventions**

1. Any psychological therapy provided as monotherapy, that is, the intervention comprised only a psychological therapy.
2. Any psychological therapy provided as an adjunct to antidepressant therapy, that is, the intervention was given in addition to an antidepressant.

We grouped psychological therapies into (1) psychodynamic/psychoanalytical; (2) cognitive-behavioural; (3) humanistic; and (4) integrated therapies. The ‘integrated therapies’ category includes integrative therapies such as IPT and CAT, which involve components of different psychological therapy models. Group 2 includes ‘third wave’ cognitive-behavioural therapy-based approaches.

**Comparator interventions**

1. An antidepressant that is included in one of five main types: TCAs, MAOIs, SSRIs, SNRIs, and NaSSAs.
2. Another psychological therapy - grouped as above.
3. An attentional control providing the same level of support and attention from a practitioner (as is received by those in the experimental intervention arm) but not containing any of the key ‘active’ ingredients of the experimental intervention.

The authors of another review have included studies examining pharmacological interventions for individuals with TRD (Williams 2013).

**Types of outcome measures**

**Primary outcomes**

1. Change in depressive symptoms as measured on rating scales for depression, either
   1. Clinician-rated depressive symptoms (e.g. Hamilton Rating Scale for Depression (HAMD) - Hamilton 1960; Montgomery-Asberg Depression Rating Scale (MADRS) - Montgomery 1979), or
   2. Self-reported depressive symptoms (e.g. Beck Depression Inventory (BDI) - Beck 1961; Beck 1996; other validated measures). We analysed data on observer-rated and self-reported outcomes separately.
2. Number of dropouts from study or treatment (all-cause dropout) within trials.

When available, we collected data on reasons for dropout and summarised them in narrative form.

**Secondary outcomes**

3. Response or remission rates, or both, based on changes in depression measures - either clinician-rated (e.g. HAMD - Hamilton 1960) or self-report (e.g. BDI - Beck 1961; Beck 1996) or other validated measures. Response is frequently quantified as at least a 50% reduction in symptoms on HAMD or BDI, but we accepted the study’s original definition. Remission is based on the absolute score on the depression measure. Examples of definitions of remission include scores of 7 or less on the HAMD and 10 or less on the BDI. Again, we accepted the study authors’ original definition.
4. When available, we summarised in narrative form data on improvements in social adjustment and social functioning including Global Assessment of Function scores, as provided in Luborsky 1962.
5. When available, we summarised in narrative form data on improvement in quality of life as measured on Short Form (SF)-36 (Ware 1993), Health of the Nation Outcome Scales (HoNOS) (Wing 1994), or World Health Organization Quality of Life (WHOQOL - WHOQOL 1998) or similar scales.
6. When reported, we summarised in narrative form data on economic outcomes, for example, days of work absence/ability to return to work, number of appointments with primary care physician, number of referrals to secondary services, and use of additional treatments.
7. When reported, we summarised in narrative form data on adverse effects, for example, completed/attempted suicides.
Timing of outcome assessment
We summarised outcomes at each reported follow-up point. When appropriate, and when the data allowed, we categorised outcomes as short term (up to six months), medium term (seven to 12 months post treatment), and long term (longer than 12 months).

Search methods for identification of studies

Cochrane Common Mental Disorders Controlled Trials Register (CCMDCTR)
The Cochrane Common Mental Disorders Group (CCMD) maintains two archived clinical trials registers at its editorial base in York, UK: a references register and a studies-based register. The CCMDCTR-References Register contains over 40,000 reports of RCTs examining depression, anxiety, and neurosis. Approximately 50% of these references have been tagged to individual coded trials. Coded trials are held in the CCMDCTR-Studies Register, and records are linked between the two registers through the use of unique Study ID tags. Coding of trials is based on the EU-Psi coding manual, which uses a controlled vocabulary (please contact the CCMD Information Specialist for further details). Reports of trials for inclusion in the Group’s registers are collated from routine (weekly) generic searches of MEDLINE (1950 to 2016), Embase (1974 to 2016), and PsycINFO (1967 to 2016); quarterly searches of the Cochrane Central Register of Controlled Trials (CENTRAL); and review-specific searches of additional databases. Reports of trials are also sourced from international trials registries via the World Health Organization trials portal (International Clinical Trials Registry Platform (ICTRP)), pharmaceutical companies, handsearching of key journals, conference proceedings, and other (non-Cochrane) systematic reviews and meta-analyses. Details of CCMD’s generic search strategies (used to identify RCTs) can be found on the Group’s website. The Group’s Specialised Register had fallen out of date with the Editorial Group’s move from Bristol to York in the summer of 2016.

Electronic searches
We searched the CCMDCTR-Studies Register using the following terms:
Condition = ((depressi* or "affective disorder" or "mood disorder") and (“treatment-resistant” or recurrent))
We searched the CCMDCTR-References Register using a more sensitive set of terms (keywords and subject headings) to identify additional untagged/uncoded references:
1. depressi* [Ti, Ab, KW]
2. (*refractory* or *resistan* or *recurren*) [Ti, Ab]
3. (augment* or potentiat*) [Ti, Ab]
4. (chronicity or "chronic depress*" or "chronically depress*" or "depressed chronic*" or "chronic major depressi*" or "chronic affective disorder*" or "chronic mood disorder*" or (chronic* and (relaps* or recur*))]) [Ti, Ab, KW]
5. (“persistent depress*” or "persistently depress*” or "depression persist*” or "persistent major depress*” or “persistance of depress*” or “persistence of major depress*”) [Ti, Ab]
6. (nonrespon* or non-respon* or "non respon*" or "not respon*” or "no respon*” or "partial respon*” or "incomplete respon*” or "incompletely respon*” or unrespon*) [Ti, Ab]
7. (“failed to respond” or "failed to improve” or “failure to respon*” or “failure to improve” or "failed medication*” or "antidepressant fail*” or "treatment fail*”) [Ti, Ab]
8. (inadequate* and respon*) [Ti, Ab]
9. "treatment-resistant depression " [KW]
10. (recurrence or “recurrent depression” or “recurrent disease”) [KW]
11. "drug resistance” [KW]
12. "treatment failure” [KW]
13. "drug potentiation” [KW]
14. augmentation [KW]
15. or/2-14
16. (1 and 15)
We applied no date or language restrictions to our search. Our search of the CCMDCTR was up-to-date as of 18 March 2016. We ran additional searches via the following biomedical databases (1 January 2016 to 16 May 2017) (Appendix 1):
1. Medline/Premedline = 553
2. Embase = 546
3. CENTRAL = 477
4. Psychinfo = 246
5. Web of Science = 673
We used the term ‘treatment-resistant’ or ‘treatment refractory’ depression to search international trials registries, including the WHO trials portal (ICTRP) and ClinicalTrials.gov (to 30 June 2017) (Appendix 1), to identify any additional ongoing and unpublished studies. We contacted Principal Investigators, when necessary, to request further details of ongoing/unpublished studies or trials reported as conference abstracts only. These searches are up-to-date until 30 June 2017.

Searching other resources
We searched the reference lists of all included studies and other relevant systematic reviews for studies that may meet review inclusion criteria. We contacted subject experts to ensure that we had considered for inclusion all relevant published and unpublished studies.

Data collection and analysis

Selection of studies
One review author (NW or PD or SI) examined titles and abstracts and removed obviously irrelevant reports, then screened study abstracts against inclusion criteria using a standardised abstract screening form. In any case of uncertainty, an over-inclusive approach was taken and the full paper was obtained, along with full papers for studies assessed as meeting the inclusion criteria. Two review authors screened each paper for inclusion or exclusion from the review. If any disagreements arose, these were discussed with a third review author. If it was not possible to determine eligibility for a study, review authors added that study to the list of those awaiting assessment and contacted trial authors to request further information or clarification. Review authors documented the study selection process using a PRISMA study selection flow diagram.

Data extraction and management

Two review authors used a standardised data extraction form to independently extract data regarding participants, interventions and their comparators, methodological details, treatment effects including dropouts, and possible biases. If any disagreements arose, they discussed these with a third review author. The data extraction form was piloted during the first phase of data extraction. Review authors abstracted information related to study populations, definition of TRD, sample size, interventions, comparators, potential biases in conduct of the trial, outcomes, follow-up, and methods of statistical analysis.

Main planned comparisons

- Any psychological therapy versus antidepressant treatment alone.
- Any psychological therapy versus another psychological therapy.
- Any psychological therapy given in addition to antidepressant medication versus antidepressant treatment alone.
- Any psychological therapy given in addition to antidepressant medication versus a psychological therapy alone.
- Any psychological therapy versus an attention control.

For comparison 2, review authors grouped the different types of psychological therapies according to the list given earlier. If we identified enough studies, we planned to pool the evidence for CBT, IPT, CAT, etc., individually within various categories for comparisons 1, 3, and 4.

Assessment of risk of bias in included studies

Two review authors independently assessed risk of bias for each included study using the 'Risk of bias' tool of the Cochrane Collaboration (Higgins 2017). We discussed any disagreements with a third review author. We assessed the following criteria.

1. Sequence generation: Was the allocation sequence adequately generated?
2. Allocation concealment: Was allocation adequately concealed?
3. Blinding of participants, study personnel, and outcome assessors for each outcome: Was knowledge of the allocated treatment adequately prevented during the study?
4. Incomplete outcome data for each main outcome or class of outcomes: Were incomplete outcome data adequately addressed?
5. Selective outcome reporting: Were reports of the study free of the suggestion of selective outcome reporting?
6. Other sources of bias: Was the study apparently free of other problems that could put it at high risk of bias? For example, not reporting baseline numbers, describing differential attrition, following up only on people who continued taking medication. Review authors extracted a description of what was reported to have happened in each study and judged risk of bias for each domain within and across studies, based on the following three categories: low risk of bias; unclear risk of bias; and high risk of bias. When studies provided few or no details about the process of randomisation, review authors contacted trial authors to seek clarification.

Measures of treatment effect

We analysed continuous outcomes by calculating the mean difference (MD) between groups if studies used the same outcome measure for comparison. If studies used different outcome measures to assess the same outcome, we calculated the standardised mean difference (SMD) and 95% confidence intervals (CIs). The SMD can be interpreted as follows: 0.2 represents a small effect, 0.5 a moderate effect, and 0.8 a large effect (Cohen 1988). We calculated risk ratios (RRs) for dichotomous outcomes. When overall risks were significant, we planned to calculate the number needed to treat for an additional beneficial outcome (NNTB) or the number needed to treat for an additional harmful outcome (NNTH) to produce one outcome by combining the overall RR with an estimate of prevalence of the event in the control groups of trials.

Unit of analysis issues

Cluster-randomised trials

We planned to incorporate results from cluster RCTs into the review using generic inverse variance methods (Higgins 2011). With cluster RCTs, it is important to ensure that data were analysed with consideration of their clustered nature. The intracluster correlation coefficient (ICC) for each trial was to be extracted. When no such data were reported, we planned to request them from study authors. If these data were not available, in line with the Cochrane
Handbook for Systematic Reviews of Interventions (Higgins 2011), we planned to use estimates from similar studies to 'correct' data for clustering when this had not been done.

Cross-over trials
For cross-over trials, we planned to include in the analysis only results from the first randomised treatment period.

Studies with multiple treatment groups
Studies that include more than two arms (e.g. psychological intervention (A); psychological intervention (B); and control) can cause problems in pair-wise meta-analysis. For studies with two or more active treatment arms, we undertook the following approach according to whether the outcome was continuous or dichotomous.

For a continuous outcome: We pooled means, standard deviations (SDs), and the number of participants for each active treatment group across treatment arms as a function of the number of participants in each arm for comparison against the control group (Higgins 2011).

For a dichotomous outcome: We planned to combine active treatment groups into a single arm for comparison against the control group (in terms of numbers of people with events and sample sizes) or to split the control group equally (Higgins 2011).

Dealing with missing data
We contacted study authors to request data when missing. If an outcome was missing for more than 50% of participants, we excluded this study from the analysis. When available, we used intention-to-treat (ITT) analyses from the study reports and wrote to study authors to request relevant unreported analyses.

Assessment of heterogeneity
We assessed heterogeneity using the Chi² test, which provides evidence of variation in effect estimates beyond that of chance. The Chi² test has low power to assess heterogeneity when included studies are few or numbers of participants small; so we set the P value conservatively at 0.1. We also quantified heterogeneity using the I² statistic, which calculates the percentage of variability due to heterogeneity rather than chance. We expected, a priori, that clinical heterogeneity between studies would be considerable; therefore we considered I² values between 50% and 90% to represent substantial statistical heterogeneity that would need to be explored further.

Assessment of reporting biases
We managed reporting bias by undertaking comprehensive searches for papers in all languages and studies outside the peer-reviewed domain. We determined outcome reporting bias for all included studies and sought trial protocols whenever possible. If outcome data were missing, we requested these from trial authors. We had planned to use funnel plots to help detect reporting biases and to conduct formal testing for small-study effects using the Egger test (Egger 1997) if 10 or more studies were included in the review (Higgins 2011).

Data synthesis
Given the potential for heterogeneity in the included interventions, we used a random-effects model for all analyses. This approach incorporates the assumption that different studies are estimating different, yet related, intervention effects and takes into account differences between studies even if no statistically significant heterogeneity is found. We tested heterogeneity formally using both the Chi² test and the I² statistic (as outlined above).

We sought clinical advice regarding combining treatment groups to ensure that findings were clinically meaningful. When a meta-analysis was not possible (e.g. owing to insufficient data or substantial heterogeneity), we provided a narrative assessment of the evidence in which we summarised the evidence according to intervention type.

Subgroup analysis and investigation of heterogeneity
A priori, we considered the degree of treatment resistance recorded at the point of entry to the trial a potential effect modifier. Therefore we planned the following subgroup analyses (based on two variables).

1. Severity of depression: classifying participants as 'non-responders' or 'partial responders' at baseline
2. Length of acute treatment phase (before trial entry): four weeks or longer, 12 weeks or longer, or six months or longer

We planned to conduct such subgroup analyses when we had obtained data from at least 10 included studies (Higgins 2011).

Sensitivity analysis
We planned to conduct sensitivity analyses to explore how much of the variation between studies comparing psychological therapies for TRD was accounted for by between-study differences in:

1. study quality: allocation concealment used as a marker of trial quality; studies that have not used allocation concealment were excluded;
2. attrition: studies with more than 20% dropout excluded;
3. missing data: studies that have imputed missing data excluded;
4. treatment fidelity: studies that have not measured treatment fidelity of the psychological model excluded; or
5. publication type: studies that have not been published in full (conference abstracts/proceedings, doctoral dissertations) excluded.
In the original protocol, we stated that we would produce 'Summary of findings' (SoF) tables for all relevant comparisons. However, the current recommendation to Cochrane review authors is that they select one 'primary' time point that they will report for all outcomes in the SoF tables. Following this, we present the short-term outcome in the SoF table.

**Results**

**Description of studies**

**Results of the search**

We found 4705 records via our electronic searches. We located four further papers through complementary searches of references and study author contacts. After removing duplicates, we screened 4515 titles and abstracts, of which we excluded 4408. However, four studies (five references) are still awaiting full-text assessment (Characteristics of studies awaiting classification) as, to date, we could not obtain full-text papers and we identified one ongoing study (Characteristics of ongoing studies), for which a full paper is not yet available. We therefore screened 102 full-text articles. We excluded 82 articles (pertaining to 63 studies) and provided reasons for exclusion in Figure 1; we presented additional details under Characteristics of excluded studies. We included in this review 20 full-text articles pertaining to six studies. All six studies contributed data to meta-analyses. We contacted the authors of all included studies with regards to points of clarification and received a response from five of the six. We also contacted two of the authors of excluded studies to request clarification on methods and received a response from one.
Figure 1. Study flow diagram.

4705 records identified through database searching

4515 records after duplicates removed

2515 records screened

4408 records excluded

63 studies (82 full-text articles) excluded:
- not RCT 9
- intervention criteria not met 11
- age criteria not met 5
- diagnostic criteria for depression not met 9
- criteria for TRD not met 28
- irrelevant comparison 1

102 full-text papers assessed for eligibility

6 studies (20 full-text) included in qualitative synthesis

6 studies included in quantitative synthesis (meta-analysis)
We have presented details of study flow in a PRISMA flow diagram in Figure 1.

**Included studies**
Six studies met all of our inclusion criteria, and we included them in this review (see Characteristics of included studies) (Harley 2008; Nakagawa 2017; Souza 2016; Town 2017; Wiles 2007; Wiles 2016).

**Design**
All six studies were parallel-group randomised trials conducted to compare the effectiveness of psychotherapy as an adjunct to usual care that included antidepressant medication versus usual care alone.

**Sample size**
Three of the six studies were small, recruiting fewer than 50 participants in total (Harley 2008; Souza 2016; Wiles 2007). Only one study was a large multi-centre RCT with a total of 469 participants randomised between two groups (Wiles 2016).

**Setting**
Two studies were reported from the same UK research group, which recruited participants from general practices (primary care) (Wiles 2007; Wiles 2016). The other four studies recruited participants from the psychiatric outpatient departments of hospitals (secondary care) and were conducted in the USA (Harley 2008), Canada (Town 2017), Japan (Nakagawa 2017), and Brazil (Souza 2016).

**Participants**
The mean age of participants in these studies ranged from 40.6 years - in Nakagawa 2017 - to 49.3 years - in Souza 2016 - and most participants were women (63.3% in Town 2017 to 85% in Souza 2016); Nakagawa 2017 was the only study that recruited more male than female participants (36%).

**Interventions**
All included studies addressed the same comparison: psychotherapy as adjunct to usual care (including antidepressants) compared with usual care alone. Trialists studied four types of psychotherapies: cognitive-behavioural therapy (CBT); dialectical behavioural therapy (DBT); interpersonal therapy (IPT); and Intensive short-term dynamic psychotherapy (ISTDP).

Three studies - one from Japan - Nakagawa 2017 - and two from the UK - Wiles 2007 and Wiles 2016 - evaluated individual cognitive-behavioural therapy (CBT) for depression using the model proposed by Beck et al (Beck 1979). The number of sessions was similar across these studies: 16 to 20 sessions in Nakagawa 2017, 12 to 20 sessions in Wiles 2007, and 12 to 18 sessions in Wiles 2016. In the two UK studies, sessions lasted up to an hour and were provided by trained and supervised therapists representative of the NHS psychological therapy services (two therapists delivered CBT in Wiles 2007, and 11 part-time therapists delivered treatment in Wiles 2016). In the Japanese study, individual sessions were 50 minutes in duration and were provided by four trained and supervised psychiatrists, one clinical psychologist, and one psychiatric nurse. In all three studies, patients in both groups continued to receive usual care from their treating doctors as needed during the study. Participants were expected to continue taking antidepressant medication as part of usual care.

Harley 2008 studied group dialectical behaviour therapy (DBT), which shares key elements of CBT, namely, change-oriented cognitive-behavioural strategies. Participants received 16 weekly sessions, each lasting 1.5 hours, with weekly between-session homework assignments. The group was run by two clinical psychologists, both of whom had received DBT training and had at least 7 years experience of leading DBT skills groups.

Souza 2016 evaluated interpersonal psychotherapy (IPT) as an adjunct to usual care. Participants received 16 individual weekly sessions, each 40 minutes in duration. One psychiatrist and one third-year psychiatry resident delivered therapy sessions. A senior IPT therapist supervised the sessions weekly.

Town 2017 studied Intensive short-term dynamic psychotherapy (ISTDP) - a brief psychotherapy format tailored to the patient’s anxiety tolerance that helps the patient identify and address emotional factors that culminate into, exacerbate, and perpetuate depression.

**Intervention engagement**
In the pilot study of CBT (Wiles 2007), participants attended, on average (median), 9.5 sessions (interquartile range (IQR) 2, 12]. In the US study of DBT (Harley 2008), participants did not attend, on average, 1.8 out of 16 sessions (range 0 to 3). In the large UK multi-centre trial (Wiles 2016), participants received an average (median) of 12 sessions of CBT (IQR 6 to 17) by 12-month follow-up. In total, 141 participants (60.3%) received at least 12 sessions of CBT, and the average duration of therapy was 6.3 months (SD 3.0). In Nakagawa 2017, the mean number of sessions completed was 15 (SD 3), with 97.5% of participants completing the CBT course.

Participants in the Brazilian study received on average 11 sessions...
of IPT, with 70% participants receiving at least eight sessions (Souza 2016). Harley 2008 reported no information on the number of sessions. The mean number of sessions completed in Town 2017 was 16.1 (SD 6.68), and 72% (n = 24) of participants received at least 15 sessions of ISTDP.

### Intervention fidelity

Four studies also assessed the fidelity of the intervention to the respective psychotherapy models (CBT and ISTDP) using a validated rating scale completed by independent raters (Nakagawa 2017; Town 2017; Wiles 2007; Wiles 2016).

### Primary outcomes

#### Change in depressive symptoms

All included studies reported the main outcome of change in depressive symptoms. All six studies reported short-term follow-up. Two studies reported medium-term follow-up (7 to 12 months) (Nakagawa 2017; Wiles 2016), and one study reported long-term (longer than 12 months) follow-up (Wiles 2016).

Five recorded outcomes using the self-report BDI scale. Three studies used BDI version II (Beck 1996) (Nakagawa 2017; Wiles 2007; Wiles 2016), and two studies used BDI version I (Beck 1961) (Harley 2008; Souza 2016).

Two studies also reported change in depressive symptoms on the HAMD scale (Harley 2008; Souza 2016), and Town 2017 reported HAMD-GRID change scores.

Two studies also reported change in depressive symptoms on the Patient Health Questionnaire (PHQ)-9 scale (Town 2017; Wiles 2016).

#### Number of dropouts

All trials reported the number of dropouts by group. When reported, the most common reason for dropout was inability to contact the participant, followed by withdrawal from treatment. All reported reasons are listed in Table 1 by study and group allocation.

### Secondary outcomes

#### Response or remission rates

All six studies measured remission. Three studies defined remission as participants scoring less than 7 on the HAMD scale (Harley 2008; Souza 2016; Town 2017); Nakagawa 2017 defined it as scoring less than 7 on the HAMD-GRID; and Wiles 2007 and Wiles 2016 defined remission as scoring less than 10 on the BDI-II.

Four studies measured response (Nakagawa 2017; Souza 2016; Wiles 2007; Wiles 2016). Both UK studies defined response as a reduction in BDI-II score of at least 50% compared with baseline (Wiles 2007; Wiles 2016). Nakagawa 2017 and Souza 2016 defined response as a 50% reduction on the HAMD and HAMD-GRID scales, respectively.

#### Social adjustment and social functioning

Only one trial reported social functioning, which was measured on SAS work and LIFE work scales (Harley 2008). See Table 2.

#### Quality of life

Five trials measured quality of life (Nakagawa 2017; Souza 2016; Town 2017; Wiles 2007; Wiles 2016). However, results for this outcome from the Town 2017 study are not yet available. Wiles 2007 used an unpublished tool to measure quality of life, Town 2017 and Wiles 2016 used the SF-12 (mental and physical sub-scales); Souza 2016 used the WHOQOL scale; and Nakagawa 2017 used the SF-36 scale. Data are reported in Table 3.

#### Economic outcomes

Four studies collected economic data (Nakagawa 2017; Town 2017; Wiles 2007; Wiles 2016); however results from the recent studies are not yet available (Nakagawa 2017; Town 2017). An analysis of cost-effectiveness was conducted for the large-scale multi-centre trial (Wiles 2016), whereas in the pilot study (Wiles 2007), trial authors piloted the method of data collection and reported costs per patient for the entire sample (intervention and control groups combined).

#### Adverse effects

Two of the included studies reported adverse effect data (Nakagawa 2017; Town 2017). We have presented these in Table 4.

#### Follow-up times reported

Harley 2008 and Wiles 2007 reported all outcome data at four months post randomisation, and Wiles 2016 at six, 12, and (on average) 46 months post randomisation. Souza 2016 reported all outcomes at two, four, five, and six months of follow-up. Nakagawa 2017 reported six and 12 months' follow-up for all outcomes (except economic outcomes), and Town 2017 reported six months' follow-up for only the main outcomes of depression score and dropout.
Excluded studies

We have listed excluded studies with reasons for exclusion in the Characteristics of excluded studies table.
In total, we excluded 68 full-text articles referring to 55 studies. Primary reasons for exclusion were as follows: study was not an RCT (n = 8); study did not meet intervention criteria (n = 9); age criteria were not met (n = 5); diagnostic criteria for depression were not met (n = 7); comparison was irrelevant (n = 1); and criteria for TRD were not met (n = 25). Those not meeting TRD criteria included: relapse prevention and/or recurrent depression (n = 13), other mood or depressive disorders (n = 7), and psychotic disorders (n = 2) (see Characteristics of excluded studies for detail for each study). We excluded some studies (n = 7) for more than one primary reason.
We did not include three large and well-known trials (STAR*D, REVAMP, and TADS) in this review as they did not meet inclusion criteria.
The STAR*D trial did not apply diagnostic criteria at the stage of randomisation to psychotherapy (Thase, 2007- STAR*D). This study also originally included those who could not tolerate antidepressant medication as well as those who had not responded to medication.
For the REVAMP trial (Kocsis, 2009 - REVAMP), not all participants met the DSM diagnosis at the stage of randomisation to the psychotherapy phase.
The TADS study used a definition of TRD that did not fit our review: two failed attempts with treatments - one with an antidepressant medication, and one with either an antidepressant medication or a psychological treatment (McPherson 2003 -TADS). No criterion pertained to the dose/duration of treatment in defining a 'failed attempt,' whereas our definition included a minimum of four weeks’ treatment at an adequate dose. Further, studies of interventions for those who have not responded to psychological treatments were outside the scope of our review.

Ongoing studies

We will add one ongoing trial to the update of this review (Lynch 2015 - RO-DBT (REFRAMED)).

Studies awaiting classification

Four studies (five references) await assessment and classification as we have found no full texts to date via interlibrary loans or contact with study authors (Checkley, 1999; Moras, 1999; Spooner 1999; Strauss, 2002).

Risk of bias in included studies

We have given full details of the risk of bias for included studies under Characteristics of included studies. We have provided graphical representations of the overall risk of bias in included studies for each risk of bias item in Figure 2, and for each study in Figure 3. Given the small number of studies included, we undertook no formal comparison of reporting bias based on a funnel plot.

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.

<table>
<thead>
<tr>
<th></th>
<th>Random sequence allocation (selection bias)</th>
<th>Allocation concealment (selection bias)</th>
<th>Blinding of participants and personnel (performance bias)</th>
<th>Blinding of outcome assessment (detection bias)</th>
<th>Incomplete outcome data (attrition bias)</th>
<th>Selective reporting (reporting bias)</th>
<th>Other bias</th>
</tr>
</thead>
</table>
**Allocation**

**Random sequence generation (selection bias)**

All six studies stated that they were randomised and reported adequate random sequence generation; therefore they were at low risk of bias for this item.

**Allocation concealment (selection bias)**

Five of the six studies were at low risk of bias in this domain (Nakagawa 2017; Souza 2016; Town 2017; Wiles 2007; Wiles 2016), and one was at unclear risk (Harley 2008).

In terms of how allocation was concealed from individual recruiting participants, Harley 2008 did not provide details of the method of allocation concealment used; therefore we marked this one as having unclear risk of bias for this domain. Wiles 2007 and Town 2017 used an individual independent of the recruiting researchers. Wiles 2016 used a telephone randomisation service to conceal allocation from those recruiting participants. Souza 2016 used sequentially numbered brown sealed envelopes containing the randomisation sequence. Nakagawa 2017 used an automated computer system for allocation.

**Blinding**

For psychological interventions, it is very difficult to blind patients and therapists to the intervention being provided. Participants and personnel providing treatment were not blind to treatment allocation in any of the studies. Therefore for the domain of performance bias, we considered all six studies to be at high risk of bias due to lack of blinding.

For the same reason, all self-completed BDI outcome assessments were unblinded and at high risk of detection bias. Harley 2008, Nakagawa 2017, and Souza 2016 minimised the likelihood of observer bias by blinding outcome assessors administering the observer-rated (HAMD) scale. Hence, for the HAMD depression outcome, risk of detection bias was low.

We considered risk of bias due to lack of blinding for the second primary outcome (dropout from the study) as low for all studies, as this is not likely to be affected by observer bias.

**Incomplete outcome data**

Nakagawa 2017, Wiles 2007, and Wiles 2016 were at low risk of bias in this domain, having conducted their analyses on an intention-to-treat (ITT) basis. Nakagawa 2017 and Wiles 2016 analysed outcomes without imputation for missing data first, then evaluated the robustness of their findings (assumed missing at random) by conducting sensitivity analyses imputing missing data. Wiles 2007 used last observation carried forward in the ITT analysis. Souza 2016 indicated that researchers used an ITT approach; however this likely referred to people who received the allocated intervention, as the CONSORT diagram reported including only completers in final analysis and excluding those who did not receive an allocated intervention or were lost to follow-up. Hence, with 27% dropout, we considered this study to be at high risk for bias for this domain. We also marked Harley 2008 as having high risk of bias due to high dropout without an ITT analysis to account for missing data.

**Selective reporting**

Three studies were at low risk of bias in this domain, as contact with study authors provided additional or missing information regarding planned and additional analyses (Harley 2008; Wiles 2007; Wiles 2016). Souza 2016 discussed in the study report two new outcomes that were not listed in the protocol. We have not yet received clarification from study authors on this; therefore we have marked risk of bias for this study as unclear. We also considered the two recent studies to be at unclear risk in this domain, as papers (and communication from study authors) stated that remaining outcomes will be reported in future papers (Nakagawa 2017; Town 2017).

**Other potential sources of bias**

We considered studies at high risk of bias from other sources if we noted any inconsistencies between papers reporting findings that could not be explained by study protocols or study authors in terms of differential attrition, selective follow-up, or baseline numbers not reported. We observed no such anomalies and therefore considered all six studies to be at low risk of bias from other sources.

**Effects of interventions**

See: **Summary of findings for the main comparison**

Psychotherapy as an adjunct to usual care compared with usual care alone for treatment-resistant depression in adults - short-term effects; **Summary of findings 2** Psychotherapy as an adjunct to usual care compared with usual care alone for treatment-resistant depression in adults - medium- to long-term effects

**Comparison 1. Psychotherapy + usual care (including antidepressant medication) versus usual care (including antidepressant medication)**


Primary outcomes

1.1 Depressive symptoms

Up to six months (short term)

Self-reported depressive symptoms

The pooled mean difference on the self-reported BDI scale (mean difference (MD) -4.07, 95% confidence interval (CI) -7.01 to -1.07; five trials, n = 575) favoured the addition of psychotherapy to usual care with antidepressant medication compared with usual care alone. Data show little evidence of heterogeneity (I² = 27%). Stratified by therapy type, analyses showed a similar effect for trials focusing on CBT (MD -4.56, 95% CI -7.49 to -1.63; three trials, n = 522; Analysis 1.1). The size of this effect was similar to the overall pooled estimate because a substantial proportion of weight (46.3%) in the combined analysis came from the Wiles 2016 study included in this group. For group-based DBT given in addition to usual care compared with usual care (Harley 2008, data show a larger difference in mean BDI score but wide confidence intervals and the null value of zero (MD -10.79, 95% CI -23.83 to 2.25; one study; n = 19; Analysis 1.1). The addition of Interpersonal psychotherapy (IPT) (Souza 2016) to usual care showed no difference between addition of IPT to usual care compared with usual care alone (MD 0.80, 95% CI -6.70 to 8.30; one trial, n = 34; Analysis 1.1).

Two studies reported results at six months for self-reported depressive symptoms on the PHQ-9 scale (Analysis 1.2) (Town 2017; Wiles 2016). This analysis also provided evidence of the benefit of adding psychotherapy to usual care (MD -4.66, 95% CI -8.72 to -0.59; two trials, n = 482; moderate-quality evidence). However, heterogeneity was substantial (I² = 73%). Combining self-reported depressive scales across all included studies produced very similar results (SMD -0.40, 95% CI -0.65 to -0.14; six trials, n = 635, moderate-quality evidence; Analysis 1.3). Heterogeneity was also similar (I² = 37%).

Clinician-rated depressive symptoms

Four studies used an observer-rated instrument, the HAMD, to measure depressive symptoms (Harley 2008; Nakagawa 2017; Souza 2016; Town 2017). The pooled effect showed a small between-group difference favouring psychotherapy as an adjunct to usual care (MD -3.28, 95% CI -5.71 to -0.85; four trials, n = 193; low-quality evidence; Analysis 1.4). We noted some evidence of heterogeneity (I² = 30%). Stratified by type of therapy, analyses showed that effects were similar for CBT (MD -3.20, 95% CI -5.75 to -0.65; one trial; n = 80), ISTDP (MD -5.84, 95% CI -11.22 to -0.46; one trial, n = 60), and DBT (MD -5.81, 95% CI -11.04 to -0.58; one trial, n = 19) in single studies but not for IPT (MD 0.10, 95% CI -4.05 to 4.25; one trial, n = 34; Analysis 1.4).

7 to 12 months (medium term)

Self-reported depressive symptoms

Two studies examined outcomes over the medium term (Nakagawa 2017; Wiles 2016). Those who received CBT in addition to usual care, had a BDI-II score that was lower compared than the score for those who continued with usual care (MD -3.40, 95% CI -7.21 to 0.40; two trials, n = 475; low-quality evidence; Analysis 1.5); however, the effect included the null value of zero. Heterogeneity was moderate (I² = 41%).

Researchers found beneficial effects of adjunctive therapy in terms of depressive symptoms measured on the PHQ-9 (Wiles 2016), with a small difference at one year (MD -1.90 points, 95% CI -3.2 to -0.58; one trial, n = 395; low-quality evidence; Analysis 1.6).

Clinician-rated depressive symptoms

Nakagawa also reported 12-month results on the HAMD-GRID scale. These results were similar to those for self-reported symptoms for this length of follow-up (MD -4.70, 95% CI -7.88 to -1.52; one trial, n = 80; Analysis 1.7).

Longer than 12 months (long term)

Self-reported depressive symptoms

Only one trial reported long-term outcomes (Wiles 2016). This long-term follow-up took place, on average, 46 months after randomisation. At 46 months, those who had received CBT in addition to usual care had fewer symptoms of depression on the BDI scale compared with those given usual care alone (MD -4.2, 95% CI -7.57 to -0.83; one trial, n = 248; low-quality evidence; Analysis 1.8).

Benefit was also evident in terms of depressive symptoms on PHQ-9 scores (MD -1.6, 95% CI -3.26 to -0.06; one study, n = 252; low-quality evidence; Analysis 1.9).

Clinician-rated depressive symptoms

No study reported this outcome over the long term.

1.2 Dropout

Up to six months (short term)

Random-effects meta analysis combining all six studies showed that dropout did not differ between adjunct psychotherapy and usual care groups (RR 0.85, 95% CI 0.58 to 1.24; six trials, n = 698; high-quality evidence) (Harley 2008; Nakagawa 2017; Souza 2016; Town 2017; Wiles 2007; Wiles 2016). Data show no evidence of heterogeneity (I² = 0%; Analysis 1.10).

When analysing stratification by therapy type, studies that focused on CBT findings in terms of dropout were similar (RR 0.74, 95% CI 0.48 to 1.16; three trials, n = 574; Analysis 1.10). For DBT, whilst the risk ratio for dropout favoured the control group, confidence intervals were wide and included the null value of one.
(RR 1.27, 95% CI 0.26 to 6.28; one trial, n = 24; Analysis 1.10). Similarly, in the single study on ISTDP (Town 2017), even though dropout was more than twice that of the usual care group, confidence intervals were wide and included the null value of one (RR 2.33, 95% CI 0.67 to 8.18; one trial, n = 60; Analysis 1.10). For IPT psychotherapy as adjunct to usual care (Souza 2016), dropout was lower than for usual care but confidence intervals were wide again and included the null value of one (RR 0.68, 95% CI 0.20 to 2.33; one trial, n = 40; Analysis 1.10).

7 to 12 months (medium term)

At 12 months’ follow-up, combined results from two studies showed no difference in dropout between intervention and usual care groups (RR 0.98, 95% CI 0.66 to 1.47; two trials, n = 549; moderate-quality evidence; Analysis 1.11) (Nakagawa 2017; Wiles 2016).

Longer than 12 months (long term)

At long-term follow-up (average of 46 months), Wiles 2016 found that those who received CBT as an adjunct to usual care were less likely to drop out of the study than those randomised to usual care (RR 0.80, 95% CI 0.66 to 0.97; one trial, n = 469; low-quality evidence; Analysis 1.12).

Secondary outcomes

1.3 Response (50% reduction from baseline) and remission (< 7 on HAMD or < 10 on BDI)

Up to six months (short term)

Response

Four studies reported the outcome of response over the short term, indicating clear benefit of psychotherapy as an adjunct to usual care when compared with usual care alone (RR 1.80, 95% CI 1.20 to 2.69; four trials, n = 556; low-quality evidence; Analysis 1.13) (Nakagawa 2017; Souza 2016; Wiles 2007; Wiles 2016). We found no evidence of heterogeneity (I² = 0%).

The number needed to treat for an additional beneficial outcome (NNTB) for an RR of 1.8 was 4.7 (for control group, risk of 0.264), meaning that one person on average would show response for every 4.7 treated with added psychotherapy.

Remission

All six included studies provided data for this outcome in the short term (Harley 2008; Nakagawa 2017; Souza 2016; Town 2017; Wiles 2007; Wiles 2016). Two studies defined remission on the BDI scale (less than 10) (Wiles 2007; Wiles 2016), and four studies on the HAMD scale (less than 7) (Harley 2008; Nakagawa 2017; Souza 2016; Town 2017). A random-effects meta-analysis showed that those who received psychotherapy in addition to usual care had a two-fold higher likelihood of remission over the short term compared with those given usual care alone (RR 1.92, 95% CI 1.46 to 2.52; six trials, n = 635; moderate-quality evidence; Analysis 1.16). Data show no evidence of heterogeneity (I² = 0%). The NNTB for an RR of 1.92 was 6.5 (for control group risk of 0.16), meaning that one person on average would reach remission for every 6.5 treated with added psychotherapy.

7 to 12 months (medium term)

Response

In terms of the outcome of response measured over the medium term (7 to 12 months), data from two studies show that those who received CBT in addition to usual care were more likely to meet criteria for response compared with those randomised to continue with usual care (RR 1.73, 95% CI 1.42 to 2.10; two trials, n = 475; low-quality evidence; Analysis 1.14) (Nakagawa 2017; Wiles 2016).

Remission

Data on remission were available over the medium term (7 to 12 months) for two studies (Nakagawa 2017; Wiles 2016). We found that those who received CBT in addition to usual care were twice as likely to meet criteria for remission at 12 months’ follow-up compared with those who were randomised to continue with usual care (RR 1.97, 95% CI 1.51 to 2.56; two trials, n = 475; moderate-quality evidence; Analysis 1.17).

Longer than 12 months (long term)

Response

Wiles 2016 was the only study that reported outcomes over the long term (on average, 46 months). Again, those who received CBT in addition to usual care were more likely to meet criteria for response compared with those randomised to continue with usual care (RR 1.62, 95% CI 1.13 to 2.32; one trial, n = 248; low-quality evidence; Analysis 1.15).

Remission

The same trial reported remission outcomes over the long term (an average of 46 months) (Wiles 2016). Those who had received CBT in addition to usual care had a 50% increased risk of meeting criteria for remission compared with those randomised to continue with usual care, although the 95% CI included just the null value of one (RR 1.56, 95% CI 0.97 to 2.53; one trial, n = 248; low-quality evidence; Analysis 1.18).
1.4 Social functioning

Harley 2008 was the only study reporting this outcome. Trialists used selected domains of a clinician-rated (LIFE RIFT) scale to measure social functioning in the short term (at four months) by assessors who were blinded to treatment allocation. In addition, researchers assessed social functioning using two self-reported scales - Social Adjustment Scale-Self-Report (SAS-SR) and Schwartz Outcome Scale-10 (SOS-10). Data from 19 participants revealed that all scales showed improved social functioning in the group given psychotherapy as an adjunct; however effects were not consistent across observer-rated and self-rated scales for the same domains (Table 2).

1.5 Quality of life

Five studies measured quality of life outcomes (Nakagawa 2017; Souza 2016; Town 2017; Wiles 2007; Wiles 2016), but only four reported these data. Town 2017 indicated that trialists expect to report these outcomes in a later publication. We have presented the key results in Table 3. The QOL measure used in Wiles 2007 was an unpublished six-item instrument on which scores could range from zero to 12, with lower scores denoting poorer quality of life. Wiles 2007 found little evidence to suggest that quality of life differed between intervention and usual care groups (MD 1.20, 95% CI -1.61 to 4.01). Wiles 2016 reported six-, 12-, and 46-month data for quality of life on the SF-12 mental and physical subscales. Results indicated improved QOL on the SF mental subscale for all follow-ups but no differences on the physical subscale. Souza 2016 used the WHOQOL scale to assess quality of life. At six months, data showed no differences in quality of life scores between the compared groups in physical, psychological, social, or overall quality of life. Nakagawa 2017 reported six- and 12-month results on the SF-36 mental and physical subscales. Study authors found no significant differences between groups at either follow-up for any of the subscales.

1.6 Economic outcomes

Although Wiles 2007 reported economic data, researchers did not compare data between the two groups. This pilot study focused on evaluating the feasibility of collecting such data. Wiles 2016 performed a cost utility analysis and presented an incremental cost-effectiveness ratio (ICER) from the societal perspective over a 12-month horizon (costs in GBP for 2010 without discounting). Analysis showed that for the base case, the addition of CBT to usual care would be cost-effective (cost per quality-adjusted life-year (QALY) gain = £14,911) at the currently accepted range for the National Health Service (NHS) (£20,000 to 30,000 per QALY). Sensitivity analyses showed that the ICER could range from £13,006 to £29,626. When assessed at 46-month follow-up, data show that the ICER for the base case was £5,374 (ranging from £4,622 to £6,890 in sensitivity analyses). At a societal willingness to pay of £20,000 per QALY, the net monetary benefit per patient per year was £782, which had a probability of 0.92 of being cost-effective.

The two recently published reports of trials did not present cost data (Nakagawa 2017; Town 2017); however both protocols stated that a cost-effectiveness evaluation would be performed as part of the study, so future reports are awaited for these results.

1.7 Adverse events

Two studies reported adverse events at six and 12 months, respectively (Nakagawa 2017; Town 2017). We have presented these in Table 4. Both studies reported few adverse events only for control group participants.

Other comparisons

No included studies addressed any of the following three comparisons.
1. Any psychological therapy versus antidepressant treatment alone.
2. Any psychological therapy versus another psychological therapy.
3. Any psychological therapy given in addition to antidepressant medication versus a psychological therapy alone.
4. Any psychological therapy versus an attention control.

Subgroup analyses

We did not conduct planned subgroup analysis on the following aspects because fewer than 10 studies were available for any comparison:
1. Severity of depression (non-responders vs partial responders).
2. Length of acute treatment phase (12 weeks or longer vs six months or longer).

Sensitivity analyses

We conducted preplanned sensitivity analyses for our primary outcome of depressive symptoms as follows.

Study quality: allocation concealment

When limited to studies that adequately concealed allocation, the pooled effect size (MD -4.66, 95% CI -7.94 to -1.37) for depressive outcomes on the BDI was not materially different from that in the main analysis (MD -5.10, 95% CI -7.89 to -2.30; Analysis 1.1).
Attrition: more than 20% at six months
When we removed from analysis the two studies with more than 20% dropout (Harley 2008; Souza 2016), the pooled effect size (MD -5.67, 95% CI -8.13 to -3.21) was consistent with the figure presented earlier (MD -5.10, 95% CI -7.89 to -2.30).

Missing data: imputed missing data
No trial imputed missing data for the primary analysis, so this analysis was not necessary.

Treatment fidelity: not measuring treatment fidelity of the psychological model
Four studies measured treatment fidelity (Nakagawa 2017; Town 2017; Wiles 2007; Wiles 2016). Exclusion of studies that did not measure fidelity to the psychological model yielded an estimate of treatment effect (MD -5.67, 95% CI -8.13 to -3.21) that was consistent with the main findings (MD -5.10, 95% CI -7.89 to -2.30) (Harley 2008; Souza 2016).

Publication type: studies that have not been published in full (conference abstract/proceedings, doctoral dissertation)
All studies included were available in the peer-reviewed domain; hence this was not conducted.

Reporting bias
We minimised reporting biases in the review by including unpublished information and studies, as planned. However, with few included studies, we could not test for small-study effect and therefore cannot be certain that bias due to selective reporting can be ruled out. We considered bias due to selective reporting when grading the evidence (Summary of findings for the main comparison; Summary of findings 2) and assigned findings a 'low' rating for response outcome.
### ADDITIONAL SUMMARY OF FINDINGS

**Psychotherapy as an adjunct to usual care compared with usual care alone for treatment-resistant depression in adults**

**Patient or population:** adults with treatment-resistant depression  
**Setting:** outpatient primary or secondary care  
**Intervention:** psychotherapy as an adjunct to usual care  
**Comparison:** usual care alone

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Anticipated absolute effects* (95% CI)</th>
<th>Relative effect (95% CI)</th>
<th>n of participants (studies)</th>
<th>Quality of the evidence (GRADE)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Risk with usual care</td>
<td>Risk with psychotherapy as an adjunct to usual care</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Self-reported depressive symptoms medium term (7 to 12 months) - BDI</strong></td>
<td>Mean depressive symptoms score at medium term - BDI was 17.5 (7.21 lower to 0.4 higher)</td>
<td>-</td>
<td>475 (2 RCTs)</td>
<td>⊕⊕⊕⃝⃝ LOW&lt;sup&gt;a,b&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td><strong>Observer-rated depressive symptoms medium term (7 to 12 months) - PHQ-9</strong></td>
<td>Mean depressive symptoms score at medium term - BDI was 13 (3.22 lower to 0.58 lower)</td>
<td>-</td>
<td>395 (1 RCT)</td>
<td>⊕⊕⊕⃝⃝ LOW&lt;sup&gt;a,c&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td><strong>Self (patient)-reported depressive symptoms long term (longer than 12 months) - BDI</strong></td>
<td>Mean depressive symptoms score at long term - BDI was 23.4 (7.57 lower to 0.83 lower)</td>
<td>-</td>
<td>248 (1 RCT)</td>
<td>⊕⊕⊕⃝⃝ LOW&lt;sup&gt;a,c&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td><strong>Observer-rated depressive symptoms long term (longer than 12 months) - PHQ-9</strong></td>
<td>Mean depressive symptoms score at long term - BDI was 11.1 (3.26 lower to 0.06 higher)</td>
<td>-</td>
<td>252 (1 RCT)</td>
<td>⊕⊕⊕⃝⃝ LOW&lt;sup&gt;a,b,c&lt;/sup&gt;</td>
</tr>
</tbody>
</table>
### Dropout medium term (7 to 12 months)*

| Study population | RR 0.98  
| (0.66 to 1.47) | 549 (2 RCTs) |
| 149 per 1000 | 146 per 1000 (98 to 219) |

* The risk in the intervention group (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).

---

### Dropout long term (longer than 12 months)*

| Study population | RR 0.80  
| (0.66 to 0.97) | 469 (1 RCT) |
| 523 per 1000 | 419 per 1000 (345 to 508) |

---

### Response (50% reduction in BDI) medium term (7 to 12 months)*

| Study population | RR 1.73  
| (1.42 to 2.10) | 475 (2 RCTs) |
| 345 per 1000 | 434 per 1000 (489 to 724) |

---

### Response (50% reduction in BDI) long term (longer than 12 months)*

| Study population | RR 1.62  
| (1.13 to 2.32) | 248 (1 RCT) |
| 268 per 1000 | 434 per 1000 (303 to 621) |

---

### Remission (< 7 on HAMD or < 10 on BDI) medium term (7 to 12 months)*

| Study population | RR 1.97  
| (1.51 to 2.56) | 475 (2 RCTs) |
| 223 per 1000 | 439 per 1000 (336 to 570) |

---

### Remission (< 7 on HAMD or < 10 on BDI) long term (longer than 12 months)*

| Study population | RR 1.56  
| (0.97 to 2.53) | 248 (1 RCT) |
| 179 per 1000 | 279 per 1000 (173 to 452) |

CI: confidence interval; RCT: randomised controlled trial; RR: risk ratio
GRADE Working Group grades of evidence.
High quality: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of the effect.

Outcome assessment not blind.
Wide confidence intervals.
Single study data.
Results at 46 months favour psychotherapy intervention when earlier results (6-month and 12-month) showed no difference.
Reporting bias likely as less frequently reported than remission or mean scores.
**DISCUSSION**

This review focused on psychological interventions for adults with treatment-resistant depression (TRD) when psychotherapy was the sole treatment (monotherapy) or was provided as an adjunct to antidepressant medication. We identified no studies that examined psychotherapy as monotherapy for patients with TRD. However, we found six studies that focused on psychotherapy as an adjunct to usual care (which included antidepressant medication) versus usual care alone. In three of these studies, those randomised to intervention received cognitive-behavioural therapy (CBT); in one study each, the intervention comprised interpersonal therapy (IPT), intensive short-term dynamic psychotherapy (ISTDP), and group dialectical behaviour therapy (DBT).

**Summary of main results**

We have summarised the findings of this review in two key tables for short-term (Summary of findings for the main comparison) and medium to long-term (Summary of findings 2) outcomes. The pooled mean difference for self-reported depressive symptoms (as measured on the Beck Depression Inventory (BDI)) on data from five of the six included studies shows that any psychotherapy given in addition to usual care (vs usual care alone) was beneficial over the short term (up to six months). Results were similar for the pooled standardised mean difference for six studies. The National Institute for Health and Care Excellence (NICE) guidelines group previously suggested that the minimum clinically important difference (MCID) equates to at least 3 points on the BDI (NICE 2004), and hence the benefits observed here (mean difference -4.7 points on BDI) may be clinically relevant. However, more recent evidence suggests that MCID values should be considered in terms of relative rather than absolute changes, and that patients with TRD need larger improvements on the BDI to report feeling better (Button 2015), although this requires further investigation.

Quality of evidence was moderate for the short-term outcome (Summary of findings for the main comparison). Trials provided low- to moderate-quality evidence to support the beneficial effects of adjunctive psychotherapy in terms of dichotomous outcomes of response and remission over the short term (Summary of findings for the main comparison). High-quality evidence from six pooled studies show that dropout was not differential between intervention and comparator groups. We found no evidence to support a difference in clinician/observer-rated depressive symptoms (on the Hamilton Depression Rating Scale (HAMDI)) between treatment groups, but we graded the quality of this evidence as low (Summary of findings for the main comparison).

Only two studies reported outcomes over the medium term (up to 12 months). These two studies provided low-quality evidence suggesting a similarly beneficial effect for depression outcomes at medium-term follow-up (7 to 12 months) when psychotherapy was added to usual care. Acceptability measured by dropout also was not different at this follow-up, and we considered the evidence to be of moderate quality because studies were few. Only one trial reported outcomes over the longer term. This large study based in UK primary care provided moderate-quality evidence for longer-term benefits of CBT as an adjunct to antidepressant medication (Summary of findings 2). Beneficial effects, in terms of self-reported depressive symptoms (on the BDI) and the dichotomous outcome of response, were maintained over the long term (46 months). Those randomised to receive the intervention were more likely to achieve remission, although confidence intervals at 46 months included the null value of one. At 46 months, those randomised to receive the intervention in addition to standard care were less likely to have dropped out of the study, with the confidence interval just excluding the null.

Data for outcomes of quality of life, social functioning, and resource use were limited. Based on a single study from the UK, addition of CBT to usual care appears cost-effective from the perspective of the healthcare system. For the outcome of quality of life, we noted mixed findings, with some studies providing some evidence for beneficial effects in terms of quality of life, and others finding no differences between groups. We could draw no conclusions in terms of social functioning, as these findings relate to only a single small study.

**Overall completeness and applicability of evidence**

Our search was comprehensive; we utilised the CMDCTR’s trial reference and studies registers, which were collated from searches (from inception) of multiple databases, and also included assessment of unpublished literature accessed by contacting study authors. In addition, we screened reference lists of included studies and contacted study authors for any unpublished or ongoing studies.

Findings from this review are applicable to adults with TRD defined as “depression (meeting diagnostic criteria) that has not responded to at least 4 weeks treatment with therapeutic dose of antidepressant medication”. We identified six randomised controlled trials (RCTs) as providing evidence relevant to our review in terms of effectiveness, acceptability, and safety of four different types of psychotherapy. Evidence was stronger for CBT because of the large UK trial with long-term follow-up but was based on only three studies (Nakagawa 2017; Wiles 2007; Wiles 2016). The findings for groups given DBT, ISTDP, and IPT need to be interpreted with caution, as each set of findings is based on a single small study (Harley 2008; Souza 2016; Town 2017, respectively). In terms of development of policy and services for adults with TRD, it will be important that these findings are incorporated into future revisions of treatment guidelines, both in the UK (NICE) and internationally (e.g. Canadian Network for Mood and Anxiety Treatments (CANMAT)).
The age and gender distribution of participants recruited in the included studies reflect the expected average age/gender profile of depressed patients - women in their 30s and 40s - with the exception of Nakagawa 2017, for which most participants were men. In terms of severity of depression, mean depression scores (BDI) in the included studies ranged from 26 to 31, which refers to moderate (17 to 29 on BDI; 20 to 28 on BDI-II) and severe (30 to 63 on BDI; 29 to 63 on BDI-II) depression categories for this scale (Beck 1961; Beck 1996). The two UK trials recruited participants from primary care, and the other four trials recruited outpatients from hospital (secondary care) outpatient clinics (Harley 2008; Nakagawa 2017; Souza 2016; Town 2017). Thus inclusion of data from trials in different practice settings increases the generalisability of review findings.

The six included studies were conducted in four different countries - two in the UK (n = 2) and one each in the USA, Canada, Brazil, and Japan. Nonetheless, study findings may have limited generalisability to other countries where the systems for delivering mental health care are substantially different from those described in the included studies.

We identified several gaps in the literature. We found no studies that provided evidence on the effectiveness of a psychological therapy as monotherapy for this patient group (TRD), or that compared different types of psychotherapy, or active treatment versus an attentional control.

**Quality of the evidence**

See Summary of findings for the main comparison and Summary of findings 2 for details.

**Study limitations**

In all included studies, participants could not be blinded to treatment allocation; therefore it is possible that detection bias affected measurements. For the observer-rated measure of depression HAMD, reported in four studies, investigators carried out blinded outcome assessments even though participants and providers were not blind. For the other main outcome of dropout, which was taken to indicate acceptability, awareness of treatment status is not expected to affect this more objective outcome as much. For the secondary clinical outcome of response, we considered that some reporting bias was likely.

Three studies (n = 522, including one trial with 469 participants) examined CBT as an adjunct to antidepressant medication, and one study each looked at IPT, ISTDP, and group DBT. All studies were relatively small, except the large UK multi-centre CBT trial; therefore it is not possible to draw robust conclusions about the effectiveness of different types of psychotherapy.

**Consistency of effect**

Data show little statistical heterogeneity for the main outcome of depressive symptom on BDI (27%) and HAMD (30%) scores, but high heterogeneity (75%) in the meta-analysis of two studies on the Patient Health Questionnaire (PHQ)-9 scale. One of six included studies showed no benefit of psychotherapy given as an adjunct to usual care on both BDI and HAMD scale scores (Souza 2016). This may be due to the fact that this was a relatively small study (n = 40) that would have been underpowered to detect a difference between groups or differences in the effectiveness of the different types of psychotherapy provided in the included studies. Differences in the study setting may also be relevant. As more evidence becomes available, it may be possible to draw more robust conclusions about the comparative effectiveness of different types of psychotherapy.

**Imprecision**

Confidence intervals surrounding the pooled effect for both main outcomes - depressive symptoms and dropout - were relatively narrow. The estimate of effect for dropout favoured the intervention group; however, the confidence interval included the null value of one, suggesting that psychotherapy as an adjunct to usual care was as acceptable as usual care alone.

**Indirectness**

The studies included in this review were free from indirectness - in terms of the comparison of interest (all direct comparisons) as well as the population and interventions examined.

**Publication bias**

We minimised the likelihood of publication bias by searching for and including unpublished and ongoing studies. However, we did not examine funnel plots and did not perform a formal test of small-study bias (Eggers test) given the small number of studies available. This precludes any formal conclusion regarding the absence of publication bias for this review question.

**Potential biases in the review process**

We attempted to contact study authors to resolve queries related to published and unpublished studies, and to obtain missing data relevant to this review. This attempt addressed any risk of selective bias or non-reporting of outcomes in this review. One study reported additional analyses beyond those specified in the protocol for secondary outcomes (response, remission) (Souza 2016). These analyses were based on accepted definitions of response and remission and were conducted under the assumption that all lost to follow-up had a poor outcome, which should move the effect estimate towards null. However, we cannot say for sure what effect
Agreements and disagreements with other studies or reviews

Four previous systematic reviews have addressed the question examined by our review (Carvalho 2014; McIntyre 2014; McPherson 2005; Trivedi 2011).

Of these four previous reviews, McPherson 2005 was the only review that was focused on the effectiveness of psychotherapy for TRD, similar to our review. This review used broader inclusion criteria for TRD than we did. Review authors in the McPherson review accepted any ‘author’s description of non-response as valid’ to maximise the number of included studies because no study fulfilled the four-week criterion for prior treatment duration that was used in the earlier review by Stimpson 2002. McPherson 2005 also included a wider range of interventions and comparators (any ‘talking’ intervention versus any other intervention) and study designs (including both controlled and uncontrolled studies) compared with our review, which was restricted to RCTs. McPherson 2005 found four controlled studies of CBT. These review authors reported that two studies showed benefit and two reported no data on effect but stated that there was no difference between psychotherapy and control. This method of synthesis (vote counting) is not recommended (Deeks 2011). Thus the review did not come to a clear conclusion regarding the benefit of psychotherapy for TRD but did suggest improved methodological rigour and the need to focus in future studies on measuring outcomes important to patients.

McIntyre 2014 included only one study (Wiles 2016); review authors considered results from this study as providing ‘compelling evidence’ of effectiveness of psychotherapy of CBT type for patients with TRD. The Trivedi 2011 review included two of our included studies (Harley 2008; Wiles 2007), along with three of our excluded studies (Kennedy, 2003; Scott, 2000; Thase, 2007– STAR*D). It also included one study that we evaluated at the abstract stage because it did not include patients with TRD (Blackburn 1997). The Trivedi 2011 review thus included six trials, although it counted two stages of STAR*D as two trials and therefore refers to seven trials. It concluded that psychotherapy is useful in TRD but effects are variable and the evidence base is small. Carvalho 2014 included one systematic review (Trivedi 2011) and commented on the six trials included in this review, concluding that cognitive therapies are effective in treating patients with TRD.

Two other systematic reviews that addressed the question of the effectiveness of psychological therapy as treatment for patients with TRD did not identify any relevant studies (Cooper 2011; Stimpson 2002). Given that Stimpson 2002 searched the literature up to January 2001, it is not surprising that review authors found no studies, because all studies included in the current review were published from 2007 onwards. The Cooper review reported not finding any studies conducted to assess psychotherapy for TRD, although review authors mention STAR*D under lithium augmentation. The Cooper 2011 review did not present the search strategy nor the excluded studies list, making it difficult to comment on review findings.

Of these earlier reviews, only two assessed the quality of included studies. McPherson 2005 assessed study quality using a scoring checklist, and Cooper 2011 reported study findings using randomisation as the indicator of high study quality, but neither incorporated study quality in the discussion nor in conclusions on the strength of evidence.

We used preplanned random-effects meta-analyses to pool the evidence on the effectiveness of psychotherapy not only in terms of the primary outcome of depressive symptoms but also as related to other patient important outcomes of remission and acceptability. Hence our review provides a more comprehensive and up-to-date assessment of relevant outcomes and provides a quantitative summary of the magnitude of the treatment effects that can be expected. This goes further than previous reviews, which, whilst concluding that psychotherapy may be useful in treating patients with TRD, did not synthesise the results quantitatively and only advocated that better quality studies assessing the effectiveness of psychotherapy for patients with TRD are needed.

Authors’ conclusions

Implications for practice

This review found moderate-quality evidence to show that psychotherapy given in addition to usual care (which includes antidepressant medication) for individuals with treatment-resistant depression (TRD) has beneficial effects in terms of improvement in depressive symptoms and response and remission rates over the short term. Evidence of medium- and long-term beneficial effects is similar but has been derived from two studies and one study, respectively. High-quality evidence shows that adjunct psychotherapy is as acceptable (measured with all-cause dropout) as usual...
care with antidepressant medication. In addition, cognitive-behavioural therapy as an adjunct to usual care appears to be cost-effective over almost four years from the perspective of the UK health care system (National Health Service (NHS)), although this evidence relates to a single multi-centre study.

No direct evidence on the comparative effectiveness of different types of psychotherapies is currently available.

Implications for research

Evidence is needed on the relative effectiveness of different types of psychological interventions for the large group of patients who do not respond to antidepressant medication. In addition, currently no evidence is available to answer the question of whether switching to a psychological treatment is more beneficial for this patient group compared with continuing on existing antidepressant medication. Given that many depressed patients may express a preference for psychological treatment (McHugh 2013), it is important that this evidence gap is addressed. Collaborative care has been found to improve depression outcomes, and this may be due to better monitoring resulting in more active management of treatment, but such a model has not been evaluated in patients with TRD. As the evidence base increases in this area, it will be important that future researchers compare the effectiveness of different psychological approaches with each other, and with pharmacological interventions. Conducting a network meta-analysis that enables multiple treatment options to be compared simultaneously would inform policy and practice.

Careful consideration needs to be given to the design of future trials on this topic. Many of the trials considered for inclusion in this review did not apply diagnostic criteria at the point of randomisation; therefore we excluded them. Other trialists randomised those who had not responded to treatment to two or more different interventions. Such trials can only answer the question of whether there is a difference in outcomes for those randomised to two different treatment strategies rather than informing the key question of whether it is beneficial to try an alternative strategy (be that switching to psychotherapy or augmenting with psychotherapy) compared with continuing on the existing treatment.

Future studies need to incorporate additional measures of outcomes across a range of domains that are important to patients such as quality of life, long-term remission, and adverse effects, as well as measuring resource use to inform discussions of cost-effectiveness that are key in the context of limited healthcare resources. Whilst the use of different scales to measure a single outcome presents challenges for combining data, extensions to recent work that has mapped outcomes measured on different scales - Kounali 2016 - may be useful in enabling future studies to better compare outcomes. In addition, it is important that due consideration is given to what patients with depression regard as the minimal clinically important difference in various outcome measures - as in Button 2015 - for powering future trials.

Finally, it is widely acknowledged that identification of adverse effects represents a challenge in psychotherapy trials; indeed, only two of the studies included in this review reported adverse effects. Therefore, it is important that future studies incorporate recent recommendations from Duggan 2014 to ensure that any negative effects of psychotherapy are identified and reported, and that acceptability of treatment is evaluated.

Acknowledgements

CRG funding acknowledgement

The National Institute for Health Research (NIHR) is the largest single funder of the Cochrane Common Mental Disorders Group.

Disclaimer

The views and opinions expressed therein are those of the review authors and do not necessarily reflect those of the NIHR, the NHS, or the Department of Health.

References

References to studies included in this review

Harley 2008 [published data only]

Psychological therapies for treatment-resistant depression in adults (Review)

Copyright © 2018 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

Souza 2016 [published data only]
Fleck MPA. Efficacy of interpersonal psychotherapy in treatment resistant depression. clinicaltrials.gov/show/NCT01141426 2013.

Town 2017 [published data only]
NCT. Halifax treatment refractory depression trial: a randomized controlled trial of intensive short-term dynamic psychotherapy (ISTDP) compared to secondary care treatment as usual. clinicaltrials.gov/show/NCT01141426 2014.

Wiles 2007 [published data only]

Wiles 2016 [published data only]

References to studies excluded from this review

Arnow, 2013 [published data only]

Asarnow, 2011 [published data only]

Barnhofer, 2015 [published data only]

Berner, 1974 [published data only]
Beutel 2016 {published data only}

Biesheuvel-Leliefeld, 2012 {published data only}

Bowie, 2013 {published data only}

Britton, 2010 {published data only}

Carty 2001 {published data only}

Chaput, 2008 {published data only}


Cladder-Micus, 2015 {published data only}

Crane, 2012 {published data only}

Davidson, 2005 {published data only}

Dekker 2013 {published data only}

Douglas, 2015 {published data only}

Ducasse 2016 {published data only}

Eisendrath 2016 {published data only}


Farrand, 2014 {published data only}

Frank, 1987 {published data only}

Gois, 2014 {published data only}

Greenlee, 2010 {published data only}
Hides, 2006 [published data only]

Hollandare, 2011 [published data only]

Hollon, 2014 [published data only]

Huijbers, 2012 [published data only]

Jha, 2014 [published data only]

Kearn, 2016 [published data only]

Kennard, 2006 [published data only]

Kennedy, 2003 [published data only]

Kocsis, 2009 - REVAMP [published data only]
Kocsis JH. The REVAMP (Research Evaluating the Value of Augmenting Medication with Psychotherapy) study for chronic depression. 45th Annual NCDEU (New Clinical Drug Evaluation Unit) Meeting; 2005 June 6 - 9; Boca Raton, FL: 86. 2005.

Koenig 2015 [published data only]

Kuyken, 2014 [published data only]

Ludman, 2016 [published data only]

Luyten, 2013 [published data only]

Lynch, 2007 [published data only]

Maddux, 2015 [published data only]
Maddux R. Clinical depression: is treatment outcome impacted by comorbid depressive personality?. Conference: 23rd European Congress of Psychiatry, EPA 2015 Vienna


**Moore, 1997** *published data only*

**Morris, 2010** *published data only*

**Mota, 2014** *published data only*

**Omid, 2013** *published data only*

**Otto, 2013** *published data only*

**Papadopoulos, 2014** *published data only*

**Paykel, 1999** *published data only*
Paykel ES, Scott J, Teasdale JD, Johnson AL, Garland A, Moore R, et al. Prevention of relapse in residual depression...

Pearce 2016 (published data only)


Reynolds, 2010 (published data only)

Schramm, 2011 (published data only)

Schramm, 2011a (published data only)

Schramm, 2015 (published data only)


Schroer, 2012 (published data only)

Schuling 2016 (published data only)

Scott, 2000 (published data only)

Scott, 2001 (published data only)

Shallcross 2016 (published data only)

Teismann, 2014 (published data only)

Thase, 2007- STAR*D (published data only)


Watkins, 2009 (published data only)

Watkins, 2011 (published data only)
References to studies awaiting assessment

Checkley, 1999 *(published data only)*

Moras, 1999 *(published data only)*

Spooner 1999 *(published data only)*

Strauss, 2002 *(published data only)*

References to ongoing studies

Lynch 2015 - RO-DBT (REFRAMED) *(published data only)*

Additional references

APA 2000

APA 2010

BACP, 2014
BACP. Attitudes to counselling and psychotherapy: key findings of our 2014 survey. www.bacp.co.uk 2014:www.bacp.co.uk/docs/pdf/13381_attitudes%20survey%202014%20key%20findings.pdf.

Bauer 1999

Beck 1961

Beck 1979

Beck 1996

Blackburn 1997

Button 2015

Carvalho 2007

Carvalho 2008

Carvalho 2014

Churchill 2001

Cohen 1988

Cooper 2011

Deeks 2011
Duggan 2014

Egger 1997

Ellis 1962

Fava 2005

Feighner 1972

Fekadu 2009

Freud 1949

Friedman 2004

Greenberg 2003

Hamilton 1960

Hayes 2004

Hayes 2006

Higgins 2011

Higgins 2017

Hofmann 2008

Hollander 2007

Jung 1963

Klein 1960

Klerman 1984

Kounali 2016

Lazarus 1971

Luborsky 1962

Mansell 2008

Marks 1981

Maslow 1943

Mathers 2005
Psychological therapies for treatment-resistant depression in adults (Review)

Copyright © 2018 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

May 1961

McCrone 2008

McCullough 1984

McHugh 2013

McIntyre 2014

McManus 2000

McPherson 2005

Middleton 2001

Montgomery 1979

NICE 2004

NICE 2009

Nierenberg 2007

Olfson 2009

Pampallona 2004

Papakostas 2008

Papakostas 2009

Pincus 1998

Rogers 1951

Ryle 1990

Shapiro 1990

Shelton 2008

Skinner 1953

Spitzer 1978

Stimpson 2002
Thase 1997

The NHS Information Centre 2011

Trivedi 2006

Trivedi 2011

Ware 1993

Watson 1924

Weissman 2007

WHO 1992

WHOQOL 1998

Williams 2013

Wing 1994

WPA 1974

References to other published versions of this review

Wiles 2013

* Indicates the major publication for the study
## Characteristics of included studies  
*ordered by study ID*

### Harley 2008

<table>
<thead>
<tr>
<th>Methods</th>
<th>RCT, parallel groups</th>
</tr>
</thead>
</table>
| Participants | Recruited from outpatient clinic of a hospital  
Location: USA  
Criteria for depression: Structured Clinical Interview for DSM-IV diagnoses (SCID-I)  
Age: range 18-65 years (mean 41.8 years)  
24 participants in total (n = 18 (75%) female)  
Baseline HAMD score: DBT + TAU = 16.9; waitlist/TAU = 18.9  
Baseline BDI score: DBT + TAU = 27.9; waitlist/TAU = 27.4 |
| Interventions | Group I: dialectical behavioural therapy plus usual care  
DBT: 16 weekly sessions, each lasting 1 hour 30 minutes, with weekly between-session homework assignments. The group was co-led by 2 clinical psychologists trained and experienced in group DBT  
Group II: waitlist and usual care  
Both groups continued antidepressant treatment as part of usual care |
| Outcomes | Continuous measure of depressive symptoms (Hamilton Depression Scale (HAMD) and Beck Depression Inventory (BDI) scores)  
Follow-up at 16 weeks |
| Definition of TRD | Non-response to at least 6 weeks of an adequate dose of an antidepressant  
Standard effective doses were predefined by consensus of 2 senior psychiatrists with expertise in depression |
| Notes |  |

### Risk of bias

<table>
<thead>
<tr>
<th>Bias</th>
<th>Authors’ judgement</th>
<th>Support for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation (selection bias)</td>
<td>Low risk</td>
<td>Eligible participants were “block randomised” by gender and age</td>
</tr>
<tr>
<td>Allocation concealment (selection bias)</td>
<td>Unclear risk</td>
<td>No information was given regarding concealment.</td>
</tr>
</tbody>
</table>
| Blinding of participants and personnel (performance bias)  
All outcomes | High risk | Participants and therapists were aware of the treatment allocation |
| Blinding of outcome assessment (detection bias)  
Patient reported depressive symptoms | High risk | Participants and therapists were aware of the treatment allocation |
Blinding of outcome assessment (detection bias) 
Observer rated depressive symptoms | Low risk | Quote: “The independent assessors were blind to the condition each patient had been assigned when they conducted week 0 and week 16 assessments.”

Incomplete outcome data (attrition bias) 
All outcomes | High risk | 23% and 18% dropout in intervention and control groups, respectively. Study authors did not report any approach to dealing with missing data and, when contacted, confirmed that no such analyses were undertaken.

Selective reporting (reporting bias) | Low risk | No protocol is available. Outcomes listed in methods were all reported in results. Study author was contacted for any additional outcomes or analyses conducted, and they provided a later publication of the study with additional (post hoc) analyses on 1 of the outcomes.

Other bias | Low risk | No differential attrition, baseline differences, or selective follow-up is apparent.

Nakagawa 2017

Methods | RCT, parallel group

Participants | Recruited from outpatient departments of 2 hospitals
Location: Tokyo, Japan
Criteria for depression: DMS-IV and HAMD ≥ 16
Age: mean (SD) 39.5 (9.2) years for CBT group, 41.7 (10.7) for usual care group; range 20 to 65 years
80 participants (n = 29 (36%) female)
Baseline HAMD-GRID score: CBT + TAU = 20.9; TAU = 20.8
Baseline BDI score: CBT + TAU = 27; TAU = 27.2

Interventions | Group I: cognitive-behavioural therapy in addition to usual care
CBT was delivered by 4 psychiatrists, 1 clinical psychologist, and 1 psychiatric nurse, all of whom were trained in delivering CBT and were supervised and independently rated by a specialist on adherence to CBT protocols. Participants could receive between 16 and 20 sessions
Group II: usual care
Investigators imposed no restrictions on the treatment that participants in the usual care group could receive, except CBT
Both groups continued antidepressant therapy as part of usual care

Outcomes | Change in clinician-rated 17-item GRID-Hamilton Depression Rating Scale (GRID-HAMD) score at 16 weeks
Severity and change in scores of subjective depression symptoms (BDI)-II
Dropout
Proportions of responders (> 50% reduction in GRID-HAMD from baseline) and remitters (< 7 GRID-HAMD)
Safety (numbers of serious adverse events)
<table>
<thead>
<tr>
<th>Risk of bias</th>
<th>Authors' judgement</th>
<th>Support for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation (selection bias)</td>
<td>Low risk</td>
<td>Quote: “Central computerised registration system designed for this study (which) automatically randomises patients and generates a message noting their assigned treatment.”</td>
</tr>
<tr>
<td>Allocation concealment (selection bias)</td>
<td>Low risk</td>
<td>Quote: “Allocation will be concealed and stratified by site (n = 2) with the minimisation method to balance the age and baseline HAMD score. The cutoff for age and depression level used for minimisation will not be disclosed until study completion.”</td>
</tr>
<tr>
<td>Blinding of participants and personnel (performance bias)</td>
<td>High risk</td>
<td>Quote: “Due to the nature of the intervention, neither the patients, the treating psychiatrists, [n]or the study therapists can be completely blinded, but the two latter groups are strongly instructed not to disclose the randomisation status to patients at assessments.”</td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias)</td>
<td>High risk</td>
<td>Quote: “Due to the nature of the intervention, neither the patients, the treating psychiatrists, [n]or the study therapists can be completely blinded, but the two latter groups are strongly instructed not to disclose the randomisation status to patients at assessments.”</td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias)</td>
<td>Low risk</td>
<td>Quote: “Study design uses a standardised psychiatric interview to assess depression symptomatology by blind raters; the assessors were not involved with treatment delivery and study coordination and were prohibited from accessing any information that could confer participant allocation. The success of assessor masking was tested...percent agreement was 52 and kappa coefficient was 0.00, indicating masking was successful.”</td>
</tr>
<tr>
<td>Incomplete outcome data (attrition bias)</td>
<td>Low risk</td>
<td>Primary analysis was ITT with all randomised participants included. Missing values were not imputed but were tested in sensitivity analyses using imputations for departures from missing at random</td>
</tr>
<tr>
<td>Selective reporting (reporting bias)</td>
<td>Unclear risk</td>
<td>Comment: Some outcomes are expected in future publications, so we cannot judge yet</td>
</tr>
</tbody>
</table>
### Nakagawa 2017 (Continued)

<table>
<thead>
<tr>
<th>Bias</th>
<th>Authors' judgement</th>
<th>Support for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation (selection bias)</td>
<td>Low risk</td>
<td>Quote: &quot;randomisation sequence generated by computer prior to the recruitment&quot;</td>
</tr>
<tr>
<td>Allocation concealment (selection bias)</td>
<td>Low risk</td>
<td>Quote: &quot;Single randomisation was carried out by means of sequentially numbered brown sealed envelopes containing the randomisation sequence.&quot;</td>
</tr>
</tbody>
</table>

Other bias | Low risk | No differential attrition, baseline differences, [n]or selective follow-up apparent |

### Souza 2016

<table>
<thead>
<tr>
<th>Methods</th>
<th>RCT, parallel group</th>
</tr>
</thead>
</table>
| Participants | Recruited from outpatient clinic of a hospital  
Location: Brazil  
Criteria for depression: DMS-IV  
Age: mean (SD) 49.3 (12.3) years for IPT group, 49.18 (12.5) for usual care group  
40 participants (n = 34 (85%) female)  
Baseline BDI score: CBT + TAU = 31.4; TAU = 28.8  
Baseline HAMD score: CBT + TAU = 19.8; TAU = 18.4 |
| Interventions | Group I: interpersonal psychotherapy (IPT) plus usual care  
IPT performed according to treatment guidelines; 16 individual 40-minute weekly sessions administered by third year psychiatric residents and 1 psychiatrist  
Group II: Usual care - pharmacotherapy and clinical management. Clinicians were free to choose medication(s) plus other treatments that followed standard clinical guidelines  
Both groups continued antidepressant therapy as part of usual care |
| Outcomes | Depressive symptoms: HAMD 17 as continuous score and dichotomous outcome (response - defined as 50% reduction; remission < 7); BDI as continuous score; CGI-S as continuous score  
QOL: WHOQOL continuous score  
Dropout |
| Definition of TRD | Non-response to at least 1 trial of antidepressant medication in adequate dose and duration. Adequate dose was defined as the equivalent of at least 75 mg of amitriptyline. Adequate duration - at least 4 weeks |
| Notes | |

**Risk of bias**

<table>
<thead>
<tr>
<th>Bias</th>
<th>Authors' judgement</th>
<th>Support for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation (selection bias)</td>
<td>Low risk</td>
<td>Quote: &quot;randomisation sequence generated by computer prior to the recruitment&quot;</td>
</tr>
<tr>
<td>Allocation concealment (selection bias)</td>
<td>Low risk</td>
<td>Quote: &quot;Single randomisation was carried out by means of sequentially numbered brown sealed envelopes containing the randomisation sequence.&quot;</td>
</tr>
</tbody>
</table>
| Blinding of participants and personnel (performance bias)  
All outcomes | High risk | No information is available, but it is likely that participants and personnel provided/received psychotherapy and knew about it; no placebo was used. It is possible that the clinicians who deliv- |
### Souza 2016  (Continued)

<table>
<thead>
<tr>
<th>Bias Type</th>
<th>Risk</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blinding of outcome assessment (detection bias)</td>
<td>High</td>
<td>No information is available, but it is unlikely that participants were blind to treatment; no placebo was used</td>
</tr>
<tr>
<td>Patient reported depressive symptoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias)</td>
<td>Low</td>
<td>Quote: “Investigators responsible for the outcome assessments were blinded to the treatment assignment.”</td>
</tr>
<tr>
<td>Observer rated depressive symptoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incomplete outcome data (attrition bias)</td>
<td>High</td>
<td>27.5% dropout; IPT group (n = 6), usual care (n = 5); although authors stated using ITT, not all randomised participants were analysed as per their CONSORT figure</td>
</tr>
<tr>
<td>All outcomes</td>
<td></td>
<td>Quote: “We performed analyses using the full data set including all patients randomly assigned to any of the two interventions; Considering the intention-to-treat sample the differences in response rates and remission were not significant.” Comment: awaiting trial author response on ITT queries</td>
</tr>
<tr>
<td>Selective reporting (reporting bias)</td>
<td>Unclear</td>
<td>The paper additionally reports response and remission, which were not stated in the protocol. Unclear how this may affect trial findings</td>
</tr>
<tr>
<td>Other bias</td>
<td>Low</td>
<td>No differential attrition, baseline differences, nor selective follow-up apparent</td>
</tr>
<tr>
<td>Other bias</td>
<td>Low</td>
<td>No differential attrition, baseline differences, nor selective follow-up apparent</td>
</tr>
</tbody>
</table>

### Town 2017

<table>
<thead>
<tr>
<th>Section</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methods</td>
<td>RCT, parallel group</td>
</tr>
<tr>
<td>Participants</td>
<td>Recruited from secondary care outpatient departments</td>
</tr>
<tr>
<td></td>
<td>Location: Halifax, Canada</td>
</tr>
<tr>
<td></td>
<td>Criteria for depression: DMS-IV and HAMD ≥ 16</td>
</tr>
<tr>
<td></td>
<td>Age: range 18 to 65 years (mean age 38.9 years for ISTDP group, 44.2 years for usual care group)</td>
</tr>
<tr>
<td></td>
<td>60 participants in total (n = 38 (63.3%) female)</td>
</tr>
<tr>
<td></td>
<td>Baseline score HAMD: ISTDP = 23.5; TAU = 24.03</td>
</tr>
<tr>
<td>Interventions</td>
<td>Intensive short-term dynamic psychotherapy with treatment as usual vs standard care/treatment as usual</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Change in depression severity (HAMD; PHQ-9)</td>
</tr>
<tr>
<td></td>
<td>Dropout</td>
</tr>
<tr>
<td></td>
<td>Quality of life (SF-12)</td>
</tr>
<tr>
<td></td>
<td>Cost-effectiveness</td>
</tr>
</tbody>
</table>
**Definition of TRD**

Participants with non-remitting depression following at least 1 course of antidepressants. Participants will have had at least 1 treatment trial of antidepressants at an acceptable therapeutic dose (length \( \geq 6 \) weeks) for the current depressive episode without adequate response (score on the Hamilton Rating Scale for Depression \( \geq 16 \)) at the time of screening interview.

**Notes**

The published paper presented only the depression outcome and dropout at 6 months and stated that other outcomes will be presented in a following paper. Request made to study author for unpublished data; no response yet.

---

**Risk of bias**

<table>
<thead>
<tr>
<th>Bias</th>
<th>Authors’ judgement</th>
<th>Support for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation (selection bias)</td>
<td>Low risk</td>
<td>Quote: “For purposes of randomisation, a researcher external to the study team generated a permuted block randomisation sequence using a digital random number generator.”</td>
</tr>
<tr>
<td>Allocation concealment (selection bias)</td>
<td>Low risk</td>
<td>Quote: “Screening assessments and enrolment were conducted by the study research assistant who remained blind throughout the randomisation and allocation process. Allocation was conducted at the end of enrolment by an administrative assistant.”</td>
</tr>
<tr>
<td>Blinding of participants and personnel (performance bias)</td>
<td>High risk</td>
<td>Quote: “Therapists and patients could not be blinded to treatment allocation.”</td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias) Patient reported depressive symptoms</td>
<td>High risk</td>
<td>Quote: “Therapists and patients could not be blinded to treatment allocation.”</td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias) Observer rated depressive symptoms</td>
<td>Low risk</td>
<td>Quote: “The study research assistant was blind to treatment condition; and to maintain concealment, patients were instructed to refrain from discussing their treatment during assessments; the study (included) use of blinded outcome ratings.”</td>
</tr>
<tr>
<td>Incomplete outcome data (attrition bias) All outcomes</td>
<td>Low risk</td>
<td>Quote: “The 60 randomised participants were included in our primary 'intention to treat' analysis sample; the missing data [were] distributed equally between the two groups.”</td>
</tr>
<tr>
<td>Selective reporting (reporting bias)</td>
<td>Unclear risk</td>
<td>Comment: some outcomes expected in future publications, so we cannot judge yet</td>
</tr>
<tr>
<td>Other bias</td>
<td>Low risk</td>
<td>No differential attrition, baseline differences, nor selective follow-up apparent</td>
</tr>
</tbody>
</table>
**Wiles 2007**

<table>
<thead>
<tr>
<th>Methods</th>
<th>RCT, parallel groups</th>
</tr>
</thead>
</table>
| Participants | Recruited from GP practices  
Location: UK  
Criteria for depression: ICD-10 and at least 15 on the Beck Depression Inventory (BDI-II)  
Age: range 18 to 65 years (mean age 45.5 years for CBT group, 45.1 years for usual care group)  
25 participants in total (n = 21 (84%) female)  
Baseline BDI score: CBT + TAU = 31.3; TAU = 26.6 |
| Interventions | Group I: cognitive-behavioural therapy in addition to usual care  
CBT was delivered by 2 therapists who received weekly supervision of a specialist. Participants could receive between 12 and 20 sessions  
Group II: usual care  
Investigators applied no restrictions on the treatment that participants in the usual care group could receive  
Both groups continued antidepressant therapy as part of usual care |
| Outcomes | Depressive symptoms (BDI-II) as continuous score and dichotomous outcome (response - defined as at least a 50% reduction in BDI score) |
| Definition of TRD | Non-response to at least 6 weeks of antidepressant treatment at British National Formulary (BNF) recommended doses |

**Notes**

**Risk of bias**

<table>
<thead>
<tr>
<th>Bias</th>
<th>Authors’ judgement</th>
<th>Support for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation (selection bias)</td>
<td>Low risk</td>
<td>Computer-generated randomisation</td>
</tr>
<tr>
<td>Allocation concealment (selection bias)</td>
<td>Low risk</td>
<td>Randomisation conducted by an individual independent of the recruitment process</td>
</tr>
<tr>
<td>Blinding of participants and personnel (performance bias)</td>
<td>High risk</td>
<td>Participants and therapists were aware of the treatment allocation</td>
</tr>
<tr>
<td>All outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias)</td>
<td>High risk</td>
<td>Participants were aware of the treatment allocation.</td>
</tr>
<tr>
<td>Patient reported depressive symptoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias)</td>
<td>Low risk</td>
<td>Not applicable - no observer-rated scales</td>
</tr>
<tr>
<td>Observer rated depressive symptoms</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Incomplete outcome data (attrition bias)

| All outcomes | Low risk | Last observation carried forward approach used for participants with missing data (n = 2) |

### Selective reporting (reporting bias)

| Low risk | Trial protocol available. Only BDI results were reported in the 2007 publication, but data on unpublished outcomes (QOL, final mean, SD values for BDI for each group, reasons for dropout by group) were obtained from the study author |

### Other bias

| Low risk | No differential attrition, baseline differences, nor selective follow-up apparent |

### Wiles 2016

**Methods**

RCT, parallel group (CoBalT trial)

**Participants**

Recruited from GP practices

Location: UK

Criteria for depression: ICD-10 criteria for depression and at least 14 on the Beck Depression Inventory (BDI-II)

Age: range 18 to 75 years (mean age 49.2 years for CBT group, 50.0 years for usual care group)

469 participants in total (n = 339 (72%) female)

Baseline BDI score: CBT + TAU = 31.8; TAU = 31.8

**Interventions**

Group I: cognitive-behavioural therapy plus usual care

Participants received a course of 12 sessions of CBT, with (up to) 6 additional sessions if deemed necessary by the therapist. Eleven trained and supervised therapists were representative of NHS psychological therapy services.

Sessions typically lasted 50 to 60 minutes.

At 12 months, the median number of sessions received was 12 (IQR 6 to 17)

Group II: usual care

Investigators applied no restrictions on treatment options for participants randomised to be managed as usual by their GP. Participants could be referred for counselling or for secondary care (including for CBT)

Both groups continued antidepressant medication as part of usual care

**Outcomes**

Depressive symptoms were measured by (1) Beck Depression Inventory (BDI-II) - mean scores, response (reduction in BDI of at least 50% compared with baseline), and remission (BDI-II score < 10); and (2) Patient Health Questionnaire (PHQ-9)

Measures of anxiety (Generalised Anxiety Disorder Assessment (GAD-7)) and panic (Brief PHQ)

Quality of life (QOL): SF mental and SF physical scales of the SF-12 version 2 and the EuroQOL (EQ-5D)

Economic outcomes: primary and secondary care resource use, direct costs to NHS and Personal Social Services, and participants’ out-of-pocket personal expenses and indirect costs such as travel
Definition of TRD | Non-response to at least 6 weeks of antidepressant treatment at British National Formulary (BNF) recommended doses
---|---
Notes | A predefined analysis plan was agreed upon with the Trial Steering Committee. The primary outcome for the main trial (measured at 6 months post randomisation) was a dichotomous outcome of response (defined as at least a 50% reduction in depressive symptoms compared with baseline), but for long-term follow-up (on average, 46 months post randomisation), the primary outcome was specified as a continuous outcome (BDI-II score) to maximise power. This change was made at the time the request for additional funding was submitted to the funder (6 November 2012)

### Risk of bias

<table>
<thead>
<tr>
<th>Bias</th>
<th>Authors’ judgement</th>
<th>Support for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation (selection bias)</td>
<td>Low risk</td>
<td>Quote: “Randomisation was by means of a computer-generated code....Allocation was stratified by centre and minimised (with a probability weighting of 0.80) according to baseline BDI score (14-19, 20-28, ≥ 29); whether the general practice had a counsellor (yes or no); previous treatment with antidepressants (yes or no); and duration of present episode of depression (&lt; 1 year, 1-2 years, ≥ 2 years).”</td>
</tr>
<tr>
<td>Allocation concealment (selection bias)</td>
<td>Low risk</td>
<td>Quote: “randomisation...from a remote automated telephone randomisation service, which thus ensured that the treatment allocation was concealed from the recruiting researcher”</td>
</tr>
<tr>
<td>Blinding of participants and personnel (performance bias)</td>
<td>High risk</td>
<td>Quote: “Because of the nature of the intervention, it was not possible to mask participants, general practitioners, CBT therapists, or researchers to the treatment allocation.”</td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias)</td>
<td>High risk</td>
<td>Quote: “Because of the nature of the intervention, it was not possible to mask participants, general practitioners, CBT therapists, or researchers to the treatment allocation.”</td>
</tr>
<tr>
<td>Incomplete outcome data (attrition bias)</td>
<td>Low risk</td>
<td>ITT done with and without imputation for missing data with similar results Quote: “Trial dealt with any missing data at an individual item level by adopting the following rule.”</td>
</tr>
</tbody>
</table>

### Notes

A predefined analysis plan was agreed upon with the Trial Steering Committee. The primary outcome for the main trial (measured at 6 months post randomisation) was a dichotomous outcome of response (defined as at least a 50% reduction in depressive symptoms compared with baseline), but for long-term follow-up (on average, 46 months post randomisation), the primary outcome was specified as a continuous outcome (BDI-II score) to maximise power. This change was made at the time the request for additional funding was submitted to the funder (6 November 2012).
If > 10% of the items were incomplete, then the data collected on that measure for that participant were disregarded. However, if < 10% of items on a particular measure were missing, missing item(s) were imputed using the mean of the remaining items (rounded to an integer). Sensitivity analyses were conducted using the method of multiple imputation by chained equation (MICE) to examine the impact of missing data on the main findings.

<table>
<thead>
<tr>
<th>Selective reporting (reporting bias)</th>
<th>Low risk</th>
<th>Protocol is available. Depression outcomes and main QOL measures (SF-12) are reported fully for all time points described. Additional QOL measures collected for the economic analyses (EQ-5D-3L, SF-6D) are not reported separately but were used to derive QALYs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other bias</td>
<td>Low risk</td>
<td>No differential attrition, baseline differences, nor selective follow-up apparent</td>
</tr>
</tbody>
</table>

BDI: Beck Depression Inventory.
BNF: British National Formulary.
CBT: cognitive-behavioural therapy.
CGI-S: Clinical Global Impressions Scale.
DBT: dialectical behaviour therapy.
DSM: *Diagnostic and Statistical Manual of Mental Disorders*.
EQ-5D: EuroQOL Group Quality of Life Questionnaire based on five dimensions.
EQ-5D-3L: EuroQOL Group Quality of Life Questionnaire based on a three-level scale.
EuroQOL: EuroQOL Group Quality of Life Questionnaire.
GAD: generalised anxiety disorder.
GP: general practice.
HAMD: Hamilton Depression Rating Scale.
ICD: International Classification of Diseases.
IPT: interpersonal therapy.
IQR: interquartile range.
ISTDP: intensive short-term dynamic psychotherapy.
ITT: intention-to-treat.
MICE: multiple imputation by chained equation.
mg: milligram.
NHS: National Health Service.
PHQ: Patient Health Questionnaire.
QALY: quality-adjusted life-year.
QOL: quality of life.
RCT: randomised controlled trial.
SCID: Structured Clinical Interview for DSM-IV diagnoses.
SD: standard deviation.
SF: Short Form.
TAU: treatment as usual.
### Characteristics of excluded studies  [ordered by study ID]

<table>
<thead>
<tr>
<th>Study</th>
<th>Reason for exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arnow, 2013</td>
<td>All participants did not meet diagnostic criteria at randomisation</td>
</tr>
<tr>
<td>Asarnow, 2011</td>
<td>Participants did not meet age criteria</td>
</tr>
<tr>
<td>Barnhofer, 2015</td>
<td>All participants did not meet diagnostic criteria at randomisation</td>
</tr>
<tr>
<td>Berner, 1974</td>
<td>Not an RCT</td>
</tr>
<tr>
<td>Beutel 2016</td>
<td>All participants did not meet diagnostic criteria at randomisation</td>
</tr>
<tr>
<td>Biesheuvel-Leliefeld, 2012</td>
<td>Did not meet intervention criteria; not TRD: participants currently recovered from depression</td>
</tr>
<tr>
<td>Bowie, 2013</td>
<td>Not TRD: no dose or duration of prior treatment provided as criteria for TRD</td>
</tr>
<tr>
<td>Britton, 2010</td>
<td>Not TRD: no prior antidepressant treatment</td>
</tr>
<tr>
<td>Carty 2001</td>
<td>Not an RCT</td>
</tr>
<tr>
<td>Cladder-Micus, 2015</td>
<td>Not TRD: not all participants on prior antidepressant treatment</td>
</tr>
<tr>
<td>Crane, 2012</td>
<td>Not TRD: major depressive disorder</td>
</tr>
<tr>
<td>Davidson, 2005</td>
<td>All participants did not meet diagnostic criteria at randomisation</td>
</tr>
<tr>
<td>Dekker 2013</td>
<td>Not TRD: not all participants on prior antidepressant treatment</td>
</tr>
<tr>
<td>Douglas, 2015</td>
<td>Participants did not meet diagnostic criteria at randomisation; not TRD: recurrent mood disorder</td>
</tr>
<tr>
<td>Ducasse 2016</td>
<td>Participants did not meet diagnostic criteria at randomisation</td>
</tr>
<tr>
<td>Eisendrath 2016</td>
<td>Irrelevant comparison</td>
</tr>
<tr>
<td>Farrand, 2014</td>
<td>Did not meet intervention criteria; not TRD: no prior antidepressant treatment</td>
</tr>
<tr>
<td>Frank, 1987</td>
<td>Not an RCT</td>
</tr>
<tr>
<td>Gois, 2014</td>
<td>Not TRD: depression secondary to type 2 diabetes</td>
</tr>
<tr>
<td>Greenlee, 2010</td>
<td>Participants did not meet age criteria and did not meet diagnostic criteria at randomisation</td>
</tr>
<tr>
<td>Hides, 2006</td>
<td>Did not meet intervention criteria; not TRD: comorbid depression with substance abuse</td>
</tr>
<tr>
<td>Study</td>
<td>TRD Description</td>
</tr>
<tr>
<td>------------------</td>
<td>---------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Hollandare, 2011</td>
<td>Did not meet intervention criteria. All participants did not meet diagnostic criteria at point of randomisation; not TRD: 'partly remitted depression' - no further detail</td>
</tr>
<tr>
<td>Hollon, 2014</td>
<td>Not TRD: recurrent or chronic depression</td>
</tr>
<tr>
<td>Huijbers, 2012</td>
<td>Not TRD: remitted people only</td>
</tr>
<tr>
<td>Jha, 2014</td>
<td>Not TRD: recurrent depression</td>
</tr>
<tr>
<td>Kearns, 2016 - DARE</td>
<td>Not TRD: MDD (recurrent) or bipolar disorder; participants in remission</td>
</tr>
<tr>
<td>Kennard, 2006</td>
<td>Participants did not meet age criteria; not TRD: relapse prevention/currently remitted</td>
</tr>
<tr>
<td>Kennedy, 2003</td>
<td>All participants did not meet diagnostic criteria at the point of randomisation</td>
</tr>
<tr>
<td>Kocsis, 2009 - REV AMP</td>
<td>All participants did not meet diagnostic criteria at randomisation: REVAMP</td>
</tr>
<tr>
<td>Koenig 2015</td>
<td>Did not meet intervention criteria; not TRD: MDD comorbid with chronic medical illness</td>
</tr>
<tr>
<td>Kuyken, 2014</td>
<td>Not TRD: relapse prevention</td>
</tr>
<tr>
<td>Ludman, 2016</td>
<td>Did not meet intervention criteria</td>
</tr>
<tr>
<td>Luyten, 2013</td>
<td>Not RCT; not TRD: MDD</td>
</tr>
<tr>
<td>Lynch, 2007</td>
<td>Not TRD: MDD or MDD with personality disorder</td>
</tr>
<tr>
<td>Maddux, 2015</td>
<td>Not TRD: chronic depression with comorbid personality disorder</td>
</tr>
<tr>
<td>Markowitz, 2012</td>
<td>Not an RCT</td>
</tr>
<tr>
<td>Martin, 2006</td>
<td>Not TRD: included participants with history of antidepressant medication</td>
</tr>
<tr>
<td>Martire, 2010</td>
<td>Participants did not meet age criteria</td>
</tr>
<tr>
<td>McPherson 2003 - TADS</td>
<td>Not TRD: included participants without antidepressant treatment; dose and duration of treatment not a consideration for inclusion</td>
</tr>
<tr>
<td>Melyani 2015</td>
<td>Not TRD: relapsing depression; no other details</td>
</tr>
<tr>
<td>Michalak 2015</td>
<td>Not TRD: included participants without antidepressant treatment</td>
</tr>
<tr>
<td>Moore, 1997</td>
<td>Not TRD: recurrent major depression</td>
</tr>
<tr>
<td>Morriss, 2010</td>
<td>Did not meet intervention criteria</td>
</tr>
<tr>
<td>Mota, 2014</td>
<td>Did not meet intervention criteria</td>
</tr>
<tr>
<td>Reference</td>
<td>Reason for Exclusion</td>
</tr>
<tr>
<td>--------------------</td>
<td>--------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Omidi, 2013</td>
<td>Did not meet intervention criteria; all patients did not meet diagnostic criteria at randomisation</td>
</tr>
<tr>
<td>Otto, 2013</td>
<td>Not an RCT</td>
</tr>
<tr>
<td>Papadopoulos, 2014</td>
<td>Not an RCT</td>
</tr>
<tr>
<td>Paykel, 1999</td>
<td>Not TRD: prevention of relapse in partly remitted participants with residual depressive symptoms</td>
</tr>
<tr>
<td>Pearce 2016</td>
<td>Did not meet intervention criteria</td>
</tr>
<tr>
<td>Reynolds, 2010</td>
<td>Participants did not meet age criteria.</td>
</tr>
<tr>
<td>Schramm, 2011</td>
<td>Not TRD: chronic depression; no inclusion criteria for dose and duration of antidepressant treatment</td>
</tr>
<tr>
<td>Schramm, 2011a</td>
<td>Not TRD: chronic depression; no inclusion criteria for dose and duration of antidepressant treatment</td>
</tr>
<tr>
<td>Schramm, 2015</td>
<td>Not TRD: chronic depression; no inclusion criteria for dose and duration of antidepressant treatment</td>
</tr>
<tr>
<td>Schroer, 2012</td>
<td>No TRD: chronic depression comorbid with chronic medical illness</td>
</tr>
<tr>
<td>Schuling 2016</td>
<td>Participants did not meet diagnostic criteria; not TRD: recurrent depressive disorder with or without a current episode</td>
</tr>
<tr>
<td>Scott, 2000</td>
<td>Not TRD: included patients with residual depressive symptoms and with psychotic depression</td>
</tr>
<tr>
<td>Scott, 2001</td>
<td>All patients did not meet diagnostic criteria at randomisation; not TRD: partly remitted recent MDD</td>
</tr>
<tr>
<td>Shallcross 2016</td>
<td>Participants did not meet diagnostic criteria.</td>
</tr>
<tr>
<td>Teismann, 2014</td>
<td>Not TRD: included patients without antidepressant treatment</td>
</tr>
<tr>
<td>Thase, 2007- STAR*D</td>
<td>Not TRD: included those who could not tolerate antidepressant medication; diagnostic criteria not applied at randomisation</td>
</tr>
<tr>
<td>Watkins, 2009</td>
<td>Not an RCT</td>
</tr>
<tr>
<td>Watkins, 2011</td>
<td>Not TRD: MDD within past 18 months, but not in previous two months; duration of antidepressant less than 4 weeks</td>
</tr>
</tbody>
</table>

MDD: major depressive disorder.
RCT: randomized controlled trial.
TRD: treatment-resistant depression.
<table>
<thead>
<tr>
<th>Study</th>
<th>Methods</th>
<th>Participants</th>
<th>Interventions</th>
<th>Outcomes</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Checkley, 1999</td>
<td>RCT - parallel group</td>
<td>Unclear if TRD</td>
<td>Psychotherapy added to usual care with AD</td>
<td>Depression</td>
<td>No full text</td>
</tr>
<tr>
<td>Moras, 1999</td>
<td>RCT - parallel group</td>
<td>Unclear if TRD</td>
<td>Psychotherapy added to usual care with AD</td>
<td>Depression</td>
<td>No full text found</td>
</tr>
<tr>
<td>Spooner 1999</td>
<td>RCT - parallel group</td>
<td>Unclear if TRD</td>
<td>Unclear</td>
<td>Depression</td>
<td>No full text found</td>
</tr>
<tr>
<td>Strauss, 2002</td>
<td>RCT - parallel group</td>
<td>Unclear if TRD</td>
<td>Psychotherapy added to usual care with AD</td>
<td>Depression</td>
<td></td>
</tr>
</tbody>
</table>
AD: antidepressant.
RCT: randomised controlled trial.
TRD: treatment-resistant depression.

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

**Lynch 2015 - RO-DBT (REFRAMED)**

<table>
<thead>
<tr>
<th>Trial name or title</th>
<th>ISRCTN85784627 (REFRAMED)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methods</td>
<td>RCT</td>
</tr>
<tr>
<td>Participants</td>
<td>18 years or older; TRD defined as 2 or more previous episodes of depression or chronic depression; in current episode, participants must have taken an adequate dose of antidepressant medication for at least 6 weeks without relief</td>
</tr>
<tr>
<td>Interventions</td>
<td>Radically Open Dialectical Behaviour Therapy (RO-DBT) vs Treatment as Usual</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Primary outcomes: Depression (HAMD, LIFE-RIFT) at 6 and 12 month after treatment</td>
</tr>
<tr>
<td></td>
<td>Health-related quality of life (EQ-5D-3 L)</td>
</tr>
<tr>
<td></td>
<td>Health services use/costs (AD-SUS)</td>
</tr>
<tr>
<td></td>
<td>Secondary outcomes: Suicide (MSSI, SBQ)</td>
</tr>
<tr>
<td></td>
<td>Depression and affect (PHQ-9, PANAS)</td>
</tr>
<tr>
<td>Starting date</td>
<td>01/01/2012</td>
</tr>
<tr>
<td>Contact information</td>
<td>Dr Roelie Hempel</td>
</tr>
<tr>
<td></td>
<td>University Road</td>
</tr>
<tr>
<td></td>
<td>Southampton</td>
</tr>
<tr>
<td></td>
<td>SO17 1BJ</td>
</tr>
<tr>
<td></td>
<td>United Kingdom</td>
</tr>
<tr>
<td>Notes</td>
<td><a href="http://www.isrctn.com/ISRCTN85784627">http://www.isrctn.com/ISRCTN85784627</a></td>
</tr>
</tbody>
</table>

AD-SUS: Adult Service Use Schedule.
EQ: EuroQOL.
EQ-5D-3L: EuroQol Group Quality of Life Questionnaire based on three-level scale.
HAMD: Hamilton Depression Scale.
LIFE-RIFT: Range of Impaired Functioning Tool.
MSSI: Modified Scale for Suicide Ideation.
PANAS: Positive and Negative Affect Schedule.
Comparison 1. Psychotherapy with usual care versus usual care alone

<table>
<thead>
<tr>
<th>Outcome or subgroup title</th>
<th>No. of studies</th>
<th>No. of participants</th>
<th>Statistical method</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Self-reported depressive symptoms short term (up to 6 months) - BDI</td>
<td>5</td>
<td>575</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-4.07 [-7.07, -1.07]</td>
</tr>
<tr>
<td>1.1 CBT with usual care vs usual care alone</td>
<td>3</td>
<td>522</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-4.56 [-7.49, -1.63]</td>
</tr>
<tr>
<td>1.2 DBT with usual care vs usual care alone</td>
<td>1</td>
<td>19</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-10.79 [-23.83, 2.25]</td>
</tr>
<tr>
<td>1.3 IPT with usual care vs usual care alone</td>
<td>1</td>
<td>34</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>0.80 [-6.70, 8.30]</td>
</tr>
<tr>
<td>2 Self-reported depressive symptoms short term (up to 6 months) - PHQ-9</td>
<td>2</td>
<td>482</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-4.66 [-8.72, -0.59]</td>
</tr>
<tr>
<td>2.1 ISTDP with usual care vs usual care alone</td>
<td>1</td>
<td>60</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-7.25 [-11.37, -3.13]</td>
</tr>
<tr>
<td>2.2 CBT with usual care vs usual care alone</td>
<td>1</td>
<td>422</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-3.0 [-4.27, -1.73]</td>
</tr>
<tr>
<td>3 Self-reported depressive symptoms short term (up to 6 months) - SMD (BDI &amp; PHQ-9)</td>
<td>6</td>
<td>635</td>
<td>Std. Mean Difference (IV, Random, 95% CI)</td>
<td>-0.40 [-0.65, -0.14]</td>
</tr>
<tr>
<td>3.1 CBT with usual care vs usual care alone</td>
<td>3</td>
<td>522</td>
<td>Std. Mean Difference (IV, Random, 95% CI)</td>
<td>-0.35 [-0.56, -0.13]</td>
</tr>
<tr>
<td>3.2 DBT with usual care vs usual care alone</td>
<td>1</td>
<td>19</td>
<td>Std. Mean Difference (IV, Random, 95% CI)</td>
<td>-0.72 [-1.66, 0.21]</td>
</tr>
<tr>
<td>3.3 IPT with usual care vs usual care alone</td>
<td>1</td>
<td>34</td>
<td>Std. Mean Difference (IV, Random, 95% CI)</td>
<td>0.07 [-0.60, 0.74]</td>
</tr>
<tr>
<td>3.4 ISTDP with usual care vs usual care alone</td>
<td>1</td>
<td>60</td>
<td>Std. Mean Difference (IV, Random, 95% CI)</td>
<td>-0.88 [-1.41, -0.35]</td>
</tr>
<tr>
<td>4 Clinician-rated depressive symptoms short term (up to 6 months) - HAMD</td>
<td>4</td>
<td>193</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-3.28 [-5.71, -0.85]</td>
</tr>
<tr>
<td>4.1 DBT with usual care vs usual care alone</td>
<td>1</td>
<td>19</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-5.81 [-11.04, -0.58]</td>
</tr>
<tr>
<td>4.2 IPT with usual care vs usual care alone</td>
<td>1</td>
<td>34</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>0.10 [-4.05, 4.25]</td>
</tr>
<tr>
<td>4.3 ISTDP with usual care vs usual care</td>
<td>1</td>
<td>60</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-5.84 [-11.22, -0.46]</td>
</tr>
<tr>
<td>4.4 CBT with usual care vs usual care alone</td>
<td>1</td>
<td>80</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-3.20 [-5.75, -0.65]</td>
</tr>
<tr>
<td>5 Self-reported depressive symptoms medium term (7 to 12 months) - BDI</td>
<td>2</td>
<td>475</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-3.40 [-7.21, 0.40]</td>
</tr>
<tr>
<td>Study</td>
<td>Comparison</td>
<td>Statistic</td>
<td>Effect Size (95% CI)</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------------</td>
<td>-----------</td>
<td>---------------------</td>
<td></td>
</tr>
<tr>
<td>5.1</td>
<td>CBT with usual care vs usual care alone</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-3.40 [-7.21, 0.40]</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Self-reported depressive symptoms medium term (7 to 12 months) - PHQ-9</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-1.90 [-3.22, -0.58]</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Clinician-rated depressive symptoms medium term (7 to 12 months) - HAMD</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-4.70 [-7.88, -1.52]</td>
<td></td>
</tr>
<tr>
<td>7.1</td>
<td>CBT with usual care vs usual care alone</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-4.70 [-7.88, -1.52]</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Self-reported depressive symptoms long term (longer than 12 months) - BDI</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-4.20 [-7.57, -0.83]</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Self-reported depressive symptoms long term (longer than 12 months) - PHQ-9</td>
<td>Mean Difference (IV, Random, 95% CI)</td>
<td>-1.60 [-3.26, 0.06]</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Dropout short term (up to 6 months)</td>
<td>Risk Ratio (M-H, Random, 95% CI)</td>
<td>0.85 [0.58, 1.24]</td>
<td></td>
</tr>
<tr>
<td>10.1</td>
<td>CBT with usual care vs usual care alone</td>
<td>Risk Ratio (M-H, Random, 95% CI)</td>
<td>0.74 [0.48, 1.16]</td>
<td></td>
</tr>
<tr>
<td>10.2</td>
<td>IPT with usual care vs usual care alone</td>
<td>Risk Ratio (M-H, Random, 95% CI)</td>
<td>0.68 [0.20, 2.33]</td>
<td></td>
</tr>
<tr>
<td>10.3</td>
<td>DBT with usual care vs usual care alone</td>
<td>Risk Ratio (M-H, Random, 95% CI)</td>
<td>1.27 [0.26, 6.28]</td>
<td></td>
</tr>
<tr>
<td>10.4</td>
<td>ISTDP with usual care vs usual care alone</td>
<td>Risk Ratio (M-H, Random, 95% CI)</td>
<td>2.33 [0.67, 8.18]</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Dropout medium term (7 to 12 months)</td>
<td>Risk Ratio (M-H, Random, 95% CI)</td>
<td>0.98 [0.66, 1.47]</td>
<td></td>
</tr>
<tr>
<td>11.1</td>
<td>CBT with usual care vs usual care alone</td>
<td>Risk Ratio (M-H, Random, 95% CI)</td>
<td>0.98 [0.66, 1.47]</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Dropout long term (longer than 12 months)</td>
<td>Risk Ratio (M-H, Random, 95% CI)</td>
<td>0.80 [0.66, 0.97]</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Response (50% reduction in depressive symptoms from baseline) short term (up to 6 months)</td>
<td>Risk Ratio (M-H, Random, 95% CI)</td>
<td>1.80 [1.20, 2.69]</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Response (50% reduction in depressive symptoms from baseline) medium term (7 to 12 months)</td>
<td>Risk Ratio (M-H, Random, 95% CI)</td>
<td>1.73 [1.42, 2.10]</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Response (50% reduction in depressive symptoms from baseline) long term (longer than 12 months)</td>
<td>Risk Ratio (M-H, Random, 95% CI)</td>
<td>1.62 [1.13, 2.32]</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Remission (&lt; 7 on HAMD or &lt; 10 on BDI) short term (up to 6 months)</td>
<td>Risk Ratio (IV, Random, 95% CI)</td>
<td>1.92 [1.46, 2.52]</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Remission (&lt; 7 on HAMD or &lt; 10 on BDI) medium term (7 to 12 months)</td>
<td>Risk Ratio (M-H, Random, 95% CI)</td>
<td>1.97 [1.51, 2.56]</td>
<td></td>
</tr>
</tbody>
</table>
Analysis 1.1. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 1 Self-reported depressive symptoms short term (up to 6 months) - BDI.

Review: Psychological therapies for treatment-resistant depression in adults

Comparison: 1 Psychotherapy with usual care versus usual care alone

Outcome: 1 Self-reported depressive symptoms short term (up to 6 months) - BDI

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean(SD)</td>
<td>N</td>
<td>Mean(SD)</td>
<td>IV ,Random,95% CI</td>
</tr>
<tr>
<td>1 CBT with usual care vs usual care alone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nakagawa 2017</td>
<td>40</td>
<td>12.5 (12.4)</td>
<td>40</td>
<td>13.3 (12.3)</td>
<td>21.5 %</td>
</tr>
<tr>
<td>Wiles 2007</td>
<td>14</td>
<td>13.1 (11.9)</td>
<td>9</td>
<td>19.3 (5.3)</td>
<td>14.2 %</td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>206</td>
<td>18.9 (14.2)</td>
<td>213</td>
<td>24.5 (13.1)</td>
<td>46.3 %</td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td>260</td>
<td>82.0 %</td>
<td>-4.56 [-7.49, -1.63]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: $\tau^2 = 1.84$; $\chi^2 = 2.60$, df = 2 ($P = 0.27$); $I^2 = 23%$

Test for overall effect: $Z = 3.05$ ($P = 0.0023$)

2 DBT with usual care vs usual care alone

Harley 2008

<table>
<thead>
<tr>
<th>N</th>
<th>Mean(SD)</th>
<th>N</th>
<th>Mean(SD)</th>
<th>IV ,Random,95% CI</th>
<th>IV ,Random,95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>15.1 (12.13)</td>
<td>9</td>
<td>25.89 (16.3)</td>
<td>4.9 %</td>
<td>-10.79 [-23.83, 2.25]</td>
</tr>
</tbody>
</table>

**Subtotal (95% CI)**

<table>
<thead>
<tr>
<th>N</th>
<th>Mean(SD)</th>
<th>N</th>
<th>Mean(SD)</th>
<th>IV ,Random,95% CI</th>
<th>IV ,Random,95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>9</td>
<td>4.9 %</td>
<td>-10.79 [-23.83, 2.25]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: not applicable

Test for overall effect: $Z = 1.62$ ($P = 0.10$)

3 IPT with usual care vs usual care alone

Souza 2016

<table>
<thead>
<tr>
<th>N</th>
<th>Mean(SD)</th>
<th>N</th>
<th>Mean(SD)</th>
<th>IV ,Random,95% CI</th>
<th>IV ,Random,95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>23.4 (11.6)</td>
<td>18</td>
<td>22.6 (10.6)</td>
<td>13.1 %</td>
<td>0.80 [-6.70, 8.30]</td>
</tr>
</tbody>
</table>

**Subtotal (95% CI)**

<table>
<thead>
<tr>
<th>N</th>
<th>Mean(SD)</th>
<th>N</th>
<th>Mean(SD)</th>
<th>IV ,Random,95% CI</th>
<th>IV ,Random,95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>18</td>
<td>13.1 %</td>
<td>0.80 [-6.70, 8.30]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: not applicable

Test for overall effect: $Z = 0.21$ ($P = 0.83$)

**Total (95% CI)**

<table>
<thead>
<tr>
<th>N</th>
<th>Mean(SD)</th>
<th>N</th>
<th>Mean(SD)</th>
<th>IV ,Random,95% CI</th>
<th>IV ,Random,95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>286</td>
<td>289</td>
<td>100.0 %</td>
<td>-4.07 [-7.07, -1.07]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: $\tau^2 = 3.27$; $\chi^2 = 5.50$, df = 4 ($P = 0.24$); $I^2 = 27%$

Test for overall effect: $Z = 2.66$ ($P = 0.0078$)

Test for subgroup differences: $\chi^2 = 2.74$, df = 2 ($P = 0.25$), $I^2 = 27%$
### Analysis 1.2. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 2 Self-reported depressive symptoms short term (up to 6 months) - PHQ-9.

Review: Psychological therapies for treatment-resistant depression in adults

Comparison: 1 Psychotherapy with usual care versus usual care alone

Outcome: 2 Self-reported depressive symptoms short term (up to 6 months) - PHQ-9

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean(SD)</td>
<td>N</td>
<td>Mean(SD)</td>
<td>IV</td>
</tr>
<tr>
<td>1 ISTDP with usual care vs usual care alone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Town 2017</td>
<td>30</td>
<td>9.84 (8.4626)</td>
<td>30</td>
<td>17.09 (7.7931)</td>
<td>-38.9</td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td>30</td>
<td>30</td>
<td>-38.9</td>
<td>%</td>
<td>-7.25 [-11.37, -3.13]</td>
</tr>
<tr>
<td>Heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 3.45 (P = 0.00056)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 CBT with usual care vs usual care alone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>209</td>
<td>9.5 (6.7)</td>
<td>213</td>
<td>12.5 (6.6)</td>
<td>61.1</td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td>209</td>
<td>213</td>
<td>61.1</td>
<td>%</td>
<td>-3.00 [-4.27, -1.73]</td>
</tr>
<tr>
<td>Heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 4.63 (P &lt; 0.00001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td>239</td>
<td>243</td>
<td>100.0</td>
<td>%</td>
<td>-4.66 [-8.72, -0.59]</td>
</tr>
<tr>
<td>Heterogeneity: Tau² = 6.62; Chi² = 3.74, df = 1 (P = 0.05); I² =73%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 2.25 (P = 0.025)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for subgroup differences: Chi² = 3.74, df = 1 (P = 0.05), I² =73%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---
## Analysis 1.3. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 3 Self-reported depressive symptoms short term (up to 6 months) - SMD (BDI & PHQ-9).

Review: Psychological therapies for treatment-resistant depression in adults

Comparison: 1 Psychotherapy with usual care versus usual care alone

Outcome: 3 Self-reported depressive symptoms short term (up to 6 months) - SMD (BDI & PHQ-9)

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Std. Mean Difference</th>
<th>Weight</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N  Mean(SD)</td>
<td>N  Mean(SD)</td>
<td>IV,Random,95% CI</td>
<td></td>
<td>IV,Random,95% CI</td>
</tr>
<tr>
<td>1 CBT with usual care vs usual care alone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nakagawa 2017</td>
<td>40 12.5 (12.4)</td>
<td>40 13.3 (12.3)</td>
<td>20.3 % -0.06 [-0.50, 0.37]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiles 2007</td>
<td>14 13.1 (11.9)</td>
<td>9 19.3 (5.3)</td>
<td>7.7 % -0.60 [-1.46, 0.26]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>206 18.9 (14.2)</td>
<td>213 24.5 (13.1)</td>
<td>38.2 % -0.41 [-0.60, -0.22]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>260</td>
<td>262</td>
<td>66.1 % -0.35 [-0.56, -0.13]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: Tau² = 0.01; Chi² = 2.31, df = 2 (P = 0.32); I² =13%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 3.13 (P = 0.0017)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 DBT with usual care vs usual care alone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harley 2008</td>
<td>10 15.1 (12.13)</td>
<td>9 25.89 (16.3)</td>
<td>6.6 % -0.72 [-1.66, 0.21]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>10</td>
<td>9</td>
<td>6.6 % -0.72 [-1.66, 0.21]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 1.51 (P = 0.13)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 IPT with usual care vs usual care alone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Souza 2016</td>
<td>16 23.4 (11.6)</td>
<td>18 22.6 (10.6)</td>
<td>11.3 % 0.07 [-0.60, 0.74]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>16</td>
<td>18</td>
<td>11.3 % 0.07 [-0.60, 0.74]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 0.21 (P = 0.84)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 ISTDP with usual care vs usual care alone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Town 2017</td>
<td>30 9.84 (8.4626)</td>
<td>30 17.09 (7.7931)</td>
<td>15.9 % -0.88 [-1.41, -0.35]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>30</td>
<td>30</td>
<td>15.9 % -0.88 [-1.41, -0.35]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 3.24 (P = 0.0012)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>316</td>
<td>319</td>
<td>100.0 % -0.40 [-0.65, -0.14]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: Tau² = 0.04; Chi² = 7.93, df = 5 (P = 0.16); I² =37%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 2.99 (P = 0.0028)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for subgroup differences: Chi² = 5.68, df = 3 (P = 0.13), I² =47%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

![Favours psychotherapy vs. usual care](image-url)
Analysis 1.4. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 4 Clinician-rated depressive symptoms short term (up to 6 months) - HAMD.

Review: Psychological therapies for treatment-resistant depression in adults

Comparison: 1 Psychotherapy with usual care versus usual care alone

Outcome: 4 Clinician-rated depressive symptoms short term (up to 6 months) - HAMD

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Mean(SD)</td>
<td>N Mean(SD)</td>
<td>N/Random,95% CI</td>
<td></td>
<td>N/Random,95% CI</td>
</tr>
<tr>
<td>1 DBT with usual care vs usual care alone</td>
<td>10 11.3 (5.31)</td>
<td>9 17.11 (6.23)</td>
<td>170 % -5.81 [ -11.04, -0.58 ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td>10 9</td>
<td></td>
<td>17.0 % -5.81 [ -11.04, -0.58 ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 2.18 (P = 0.030)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 IPT with usual care vs usual care alone</td>
<td>16 14 (6.4)</td>
<td>18 13.9 (5.9)</td>
<td>240 % 0.10 [ -4.05, 4.25 ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td>16 18</td>
<td></td>
<td>24.0 % 0.10 [ -4.05, 4.25 ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 0.05 (P = 0.96)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 ISTDP with usual care vs usual care</td>
<td>30 12.9 (13.124)</td>
<td>30 18.74 (7.3379)</td>
<td>163 % -5.84 [ -11.22, -0.46 ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td>30 30</td>
<td></td>
<td>163 % -5.84 [ -11.22, -0.46 ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 2.13 (P = 0.033)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 CBT with usual care vs usual care alone</td>
<td>40 6.1 (4.5)</td>
<td>40 9.3 (6.9)</td>
<td>427 % -3.20 [ -5.75, -0.65 ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td>40 40</td>
<td></td>
<td>427 % -3.20 [ -5.75, -0.65 ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 2.46 (P = 0.014)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td>96 97</td>
<td></td>
<td>100.0 % -3.28 [ -5.71, -0.85 ]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: Tau^2 = 1.90; Chi^2 = 4.30, df = 3 (P = 0.23); I^2 =30%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 2.65 (P = 0.0081)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for subgroup differences: Chi^2 = 4.30, df = 3 (P = 0.23), I^2 =30%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Analysis 1.5. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 5 Self-reported depressive symptoms medium term (7 to 12 months) - BDI.

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean(SD)</td>
<td>N</td>
<td>Mean(SD)</td>
<td>IVRandom,95% CI</td>
</tr>
<tr>
<td>1 CBT with usual care vs usual care alone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nakagawa 2017</td>
<td>40</td>
<td>12.5 (12.1)</td>
<td>40</td>
<td>13 (14)</td>
<td>30.9 %</td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>197</td>
<td>17 (14)</td>
<td>198</td>
<td>21.7 (12.9)</td>
<td>69.1 %</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>237</strong></td>
<td><strong>238</strong></td>
<td></td>
<td></td>
<td><strong>100.0 %</strong></td>
</tr>
</tbody>
</table>

Heterogeneity: Tau² = 3.62; Ch² = 1.70, df = 1 (P = 0.19); I² =41%
Test for overall effect: Z = 1.75 (P = 0.080)
Test for subgroup differences: Not applicable

### Analysis 1.6. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 6 Self-reported depressive symptoms medium term (7 to 12 months) - PHQ-9.

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean(SD)</td>
<td>N</td>
<td>Mean(SD)</td>
<td>IVRandom,95% CI</td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>197</td>
<td>9 (7)</td>
<td>198</td>
<td>10.9 (6.4)</td>
<td>100.0 %</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>197</strong></td>
<td><strong>198</strong></td>
<td></td>
<td></td>
<td><strong>100.0 %</strong></td>
</tr>
</tbody>
</table>

Heterogeneity: not applicable
Test for overall effect: Z = 2.81 (P = 0.0049)
Test for subgroup differences: Not applicable
Analysis 1.7. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 7 Clinician-rated depressive symptoms medium term (7 to 12 months) - HAMD.

Review: Psychological therapies for treatment-resistant depression in adults

Comparison: 1 Psychotherapy with usual care versus usual care alone

Outcome: 7 Clinician-rated depressive symptoms medium term (7 to 12 months) - HAMD

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Mean(SD)</td>
<td>N Mean(SD)</td>
<td>IV,Random,95% CI</td>
<td></td>
<td>IV,Random,95% CI</td>
</tr>
<tr>
<td>CBT with usual care vs usual care alone</td>
<td>40 5.4 (5.9)</td>
<td>40 10.1 (8.4)</td>
<td>-4.70 [-7.88, -1.52]</td>
<td>100.0 %</td>
<td>-4.70 [-7.88, -1.52]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>40</td>
<td>40</td>
<td>100.0 % -4.70 [-7.88, -1.52]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Heterogeneity: not applicable

Test for overall effect: Z = 2.90 (P = 0.0038)

Test for subgroup differences: Not applicable

---

Psychological therapies for treatment-resistant depression in adults (Review)
Copyright © 2018 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.
### Analysis 1.8. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 8 Self-reported depressive symptoms long term (longer than 12 months) - BDI.

**Review:** Psychological therapies for treatment-resistant depression in adults

**Comparison:** 1 Psychotherapy with usual care versus usual care alone

**Outcome:** 8 Self-reported depressive symptoms long term (longer than 12 months) - BDI

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wiles 2016</td>
<td>136 19.2 (13.8)</td>
<td>112 23.4 (13.2)</td>
<td>-4.20 [-7.57, -0.83]</td>
<td>100.0 %</td>
<td>100.0 %</td>
</tr>
</tbody>
</table>

Total (95% CI) 136 112 100.0 % -4.20 [-7.57, -0.83]

Heterogeneity: not applicable

Test for overall effect: Z = 2.44 (P = 0.015)

Test for subgroup differences: Not applicable

### Analysis 1.9. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 9 Self-reported depressive symptoms long term (longer than 12 months) - PHQ-9.

**Review:** Psychological therapies for treatment-resistant depression in adults

**Comparison:** 1 Psychotherapy with usual care versus usual care alone

**Outcome:** 9 Self-reported depressive symptoms long term (longer than 12 months) - PHQ-9

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Mean Difference</th>
<th>Weight</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wiles 2016</td>
<td>126 9.5 (7.1)</td>
<td>126 11.1 (6.3)</td>
<td>-1.60 [-3.26, 0.06]</td>
<td>100.0 %</td>
<td>100.0 %</td>
</tr>
</tbody>
</table>

Total (95% CI) 126 126 100.0 % -1.60 [-3.26, 0.06]

Heterogeneity: not applicable

Test for overall effect: Z = 1.89 (P = 0.058)

Test for subgroup differences: Not applicable
### Analysis 1.10. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 10 Dropout short term (up to 6 months).

**Review:** Psychological therapies for treatment-resistant depression in adults

**Comparison:** 1 Psychotherapy with usual care versus usual care alone

**Outcome:** 10 Dropout short term (up to 6 months)

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Risk Ratio (M-H,Random,95% CI)</th>
<th>Weight</th>
<th>Risk Ratio (M-H,Random,95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n/N</td>
<td>n/N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 CBT with usual care vs usual care alone</td>
<td>2/40</td>
<td>2/40</td>
<td>4.0% 1.00 [0.15, 6.76]</td>
<td>4.0% 1.00 [0.15, 6.76]</td>
<td></td>
</tr>
<tr>
<td>Nakagawa 2017</td>
<td>2/40</td>
<td>2/40</td>
<td>4.0% 1.00 [0.15, 6.76]</td>
<td>4.0% 1.00 [0.15, 6.76]</td>
<td></td>
</tr>
<tr>
<td>Wiles 2007</td>
<td>0/14</td>
<td>2/11</td>
<td>1.7% 0.16 [0.01, 3.03]</td>
<td>1.7% 0.16 [0.01, 3.03]</td>
<td></td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>28/234</td>
<td>37/235</td>
<td>69.9% 0.76 [0.48, 1.20]</td>
<td>69.9% 0.76 [0.48, 1.20]</td>
<td></td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td><strong>288</strong></td>
<td><strong>286</strong></td>
<td><strong>75.5% 0.74 [0.48, 1.16]</strong></td>
<td><strong>75.5% 0.74 [0.48, 1.16]</strong></td>
<td></td>
</tr>
<tr>
<td>Total events: 30 (Psychotherapy), 41 (Usual care)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: Tau² = 0.0; Chi² = 1.16, df = 2 (P = 0.56); I² = 0.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 1.32 (P = 0.19)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 IPT with usual care vs usual care alone</td>
<td>3/17</td>
<td>6/23</td>
<td>9.5% 0.68 [0.20, 2.33]</td>
<td>9.5% 0.68 [0.20, 2.33]</td>
<td></td>
</tr>
<tr>
<td>Souza 2016</td>
<td>3/17</td>
<td>6/23</td>
<td>9.5% 0.68 [0.20, 2.33]</td>
<td>9.5% 0.68 [0.20, 2.33]</td>
<td></td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td><strong>17</strong></td>
<td><strong>23</strong></td>
<td><strong>9.5% 0.68 [0.20, 2.33]</strong></td>
<td><strong>9.5% 0.68 [0.20, 2.33]</strong></td>
<td></td>
</tr>
<tr>
<td>Total events: 3 (Psychotherapy), 6 (Usual care)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 0.62 (P = 0.54)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 DBT with usual care vs usual care alone</td>
<td>3/13</td>
<td>2/11</td>
<td>5.7% 1.27 [0.26, 6.28]</td>
<td>5.7% 1.27 [0.26, 6.28]</td>
<td></td>
</tr>
<tr>
<td>Harley 2008</td>
<td>3/13</td>
<td>2/11</td>
<td>5.7% 1.27 [0.26, 6.28]</td>
<td>5.7% 1.27 [0.26, 6.28]</td>
<td></td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td><strong>13</strong></td>
<td><strong>11</strong></td>
<td><strong>5.7% 1.27 [0.26, 6.28]</strong></td>
<td><strong>5.7% 1.27 [0.26, 6.28]</strong></td>
<td></td>
</tr>
<tr>
<td>Total events: 3 (Psychotherapy), 2 (Usual care)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 0.29 (P = 0.77)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 ISTDP with usual care vs usual care alone</td>
<td>7/30</td>
<td>3/30</td>
<td>9.2% 2.33 [0.67, 8.18]</td>
<td>9.2% 2.33 [0.67, 8.18]</td>
<td></td>
</tr>
<tr>
<td>Town 2017</td>
<td>7/30</td>
<td>3/30</td>
<td>9.2% 2.33 [0.67, 8.18]</td>
<td>9.2% 2.33 [0.67, 8.18]</td>
<td></td>
</tr>
<tr>
<td><strong>Subtotal (95% CI)</strong></td>
<td><strong>30</strong></td>
<td><strong>30</strong></td>
<td><strong>9.2% 2.33 [0.67, 8.18]</strong></td>
<td><strong>9.2% 2.33 [0.67, 8.18]</strong></td>
<td></td>
</tr>
<tr>
<td>Total events: 7 (Psychotherapy), 3 (Usual care)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: not applicable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 1.32 (P = 0.19)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>348</strong></td>
<td><strong>350</strong></td>
<td><strong>100.0% 0.85 [0.58, 1.24]</strong></td>
<td><strong>100.0% 0.85 [0.58, 1.24]</strong></td>
<td></td>
</tr>
<tr>
<td>Total events: 43 (Psychotherapy), 52 (Usual care)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity: Tau² = 0.0; Chi² = 4.36, df = 5 (P = 0.50); I² = 0.0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect: Z = 0.86 (P = 0.39)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for subgroup differences: Chi² = 3.21, df = 3 (P = 0.36), I² = 7%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Psychological therapies for treatment-resistant depression in adults (Review)**

**Copyright © 2018 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.**
**Analysis 1.11. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 11 Dropout medium term (7 to 12 months).**

Review: Psychological therapies for treatment-resistant depression in adults

Comparison: 1 Psychotherapy with usual care versus usual care alone

Outcome: 11 Dropout medium term (7 to 12 months)

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Risk Ratio M HRandom</th>
<th>95% CI M HRandom</th>
<th>Weight %</th>
<th>Risk Ratio M HRandom</th>
<th>95% CI M HRandom</th>
</tr>
</thead>
<tbody>
<tr>
<td>n/N</td>
<td>n/N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 CBT with usual care vs usual care alone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nakagawa 2017</td>
<td>3/40</td>
<td>4/40</td>
<td>7.9 %</td>
<td>0.75 [ 0.18, 3.14 ]</td>
<td>7.9 %</td>
<td>0.75 [ 0.18, 3.14 ]</td>
<td></td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>37/234</td>
<td>37/235</td>
<td>92.1 %</td>
<td>1.00 [ 0.66, 1.53 ]</td>
<td>92.1 %</td>
<td>1.00 [ 0.66, 1.53 ]</td>
<td></td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>274</strong></td>
<td><strong>275</strong></td>
<td><strong>100.0 %</strong></td>
<td><strong>0.98 [ 0.66, 1.47 ]</strong></td>
<td><strong>100.0 %</strong></td>
<td><strong>0.98 [ 0.66, 1.47 ]</strong></td>
<td></td>
</tr>
</tbody>
</table>

Total events: 40 (Psychotherapy), 41 (Usual care)

Heterogeneity: Tau² = 0.0; Chi² = 1.5, df = 1 (P = 0.70); I² = 0.0%

Test for overall effect: Z = 0.09 (P = 0.93)

Test for subgroup differences: Not applicable

---

**Psychological therapies for treatment-resistant depression in adults (Review)**

Copyright © 2018 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.
Analysis 1.12. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 12 Dropout long term (longer than 12 months).

Review: Psychological therapies for treatment-resistant depression in adults

Comparison: 1 Psychotherapy with usual care versus usual care alone

Outcome: 12 Dropout long term (longer than 12 months)

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Risk Ratio</th>
<th>Weight</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n/N</td>
<td>n/N</td>
<td>M-H, Random, 95% CI</td>
<td></td>
<td>M-H, Random, 95% CI</td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>98/234</td>
<td>123/235</td>
<td>100.0 %</td>
<td>0.80</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[0.66, 0.97]</td>
<td></td>
<td>[0.66, 0.97]</td>
</tr>
</tbody>
</table>

Total (95% CI) 234 235 100.0 % 0.80 [0.66, 0.97]

Total events: 98 (Psychotherapy), 123 (Usual care)
Heterogeneity: not applicable
Test for overall effect: Z = 2.25 (P = 0.024)
Test for subgroup differences: Not applicable

Favours psychotherapy Favours usual care
### Analysis 1.13. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 13 Response (50% reduction in depressive symptoms from baseline) short term (up to 6 months).

Review: Psychological therapies for treatment-resistant depression in adults

Comparison: 1 Psychotherapy with usual care versus usual care alone

Outcome: 13 Response (50% reduction in depressive symptoms from baseline) short term (up to 6 months)

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Risk Ratio M-H,Random,95% CI</th>
<th>Weight</th>
<th>Risk Ratio M-H,Random,95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n/N</td>
<td>n/N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nakagawa 2017</td>
<td>34/40</td>
<td>24/40</td>
<td>43.7 % 1.42 [1.07, 1.88]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Souza 2016</td>
<td>6/16</td>
<td>4/18</td>
<td>11.3 % 1.69 [0.58, 4.92]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiles 2007</td>
<td>8/14</td>
<td>0/9</td>
<td>2.1 % 11.33 [0.73, 175.10]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>95/206</td>
<td>46/213</td>
<td>42.9 % 2.14 [1.59, 2.87]</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td>276</td>
<td>280</td>
<td>100.0 % 1.80 [1.20, 2.69]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total events: 143 (Psychotherapy), 74 (Usual care)

Heterogeneity: Tau² = 0.08; Chi² = 6.90, df = 3 (P = 0.08); I² = 57%

Test for overall effect: Z = 2.85 (P = 0.0043)

Test for subgroup differences: Not applicable

---

Psychological therapies for treatment-resistant depression in adults (Review) © 2018 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.
## Analysis 1.14. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 14 Response (50% reduction in depressive symptoms from baseline) medium term (7 to 12 months).

Review: Psychological therapies for treatment-resistant depression in adults

Comparison: Psychotherapy with usual care versus usual care alone

Outcome: Response (50% reduction in depressive symptoms from baseline) medium term (7 to 12 months)

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Risk Ratio M H,Random,95% CI</th>
<th>Weight</th>
<th>Risk Ratio M H,Random,95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n/N</td>
<td>n/N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nakagawa 2017</td>
<td>33/40</td>
<td>20/40</td>
<td>+</td>
<td>33.4 %</td>
<td>1.65 [1.17, 2.32]</td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>109/197</td>
<td>62/198</td>
<td>-</td>
<td>66.6 %</td>
<td>1.77 [1.39, 2.25]</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>237</strong></td>
<td><strong>238</strong></td>
<td>+</td>
<td><strong>100.0 %</strong></td>
<td><strong>1.73 [1.42, 2.10]</strong></td>
</tr>
</tbody>
</table>

- Total events: 142 (Psychotherapy), 82 (Usual care)
- Heterogeneity: Tau² = 0.0; Chi² = 0.11, df = 1 (P = 0.74); I² = 0.0%
- Test for overall effect: Z = 5.43 (P < 0.00001)
- Test for subgroup differences: Not applicable

## Analysis 1.15. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 15 Response (50% reduction in depressive symptoms from baseline) long term (longer than 12 months).

Review: Psychological therapies for treatment-resistant depression in adults

Comparison: Psychotherapy with usual care versus usual care alone

Outcome: Response (50% reduction in depressive symptoms from baseline) long term (longer than 12 months)

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Risk Ratio M H,Random,95% CI</th>
<th>Weight</th>
<th>Risk Ratio M H,Random,95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n/N</td>
<td>n/N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>59/136</td>
<td>30/112</td>
<td>+</td>
<td>100.0 %</td>
<td>1.62 [1.13, 2.32]</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>136</strong></td>
<td><strong>112</strong></td>
<td>+</td>
<td><strong>100.0 %</strong></td>
<td><strong>1.62 [1.13, 2.32]</strong></td>
</tr>
</tbody>
</table>

- Total events: 59 (Psychotherapy), 30 (Usual care)
- Heterogeneity: not applicable
- Test for overall effect: Z = 2.61 (P = 0.0089)
- Test for subgroup differences: Not applicable
### Analysis 1.16. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 16 Remission (< 7 on HAMD or < 10 on BDI) short term (up to 6 months).

**Review:** Psychological therapies for treatment-resistant depression in adults

**Comparison:** 1 Psychotherapy with usual care versus usual care alone

**Outcome:** 16 Remission (< 7 on HAMD or < 10 on BDI) short term (up to 6 months)

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy n/N</th>
<th>Usual care n/N</th>
<th>Risk Ratio IVRandom,95% CI</th>
<th>Weight IVRandom,95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harley 2008</td>
<td>3/10</td>
<td>0/9</td>
<td></td>
<td>0.9 % 6.36 [0.37, 108.56]</td>
</tr>
<tr>
<td>Nakagawa 2017</td>
<td>28/40</td>
<td>16/40</td>
<td>*</td>
<td>40.6 % 1.75 [1.14, 2.69]</td>
</tr>
<tr>
<td>Souza 2016</td>
<td>5/16</td>
<td>3/18</td>
<td></td>
<td>4.7 % 1.88 [0.53, 6.63]</td>
</tr>
<tr>
<td>Town 2017</td>
<td>11/30</td>
<td>1/30</td>
<td></td>
<td>19 % 1.10 [1.51, 79.96]</td>
</tr>
<tr>
<td>Wiles 2007</td>
<td>6/14</td>
<td>1/9</td>
<td></td>
<td>2 % 3.86 [0.55, 26.6]</td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>57/206</td>
<td>32/213</td>
<td>*</td>
<td>49.9 % 1.84 [1.25, 2.72]</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>316</strong></td>
<td><strong>319</strong></td>
<td><strong>100.0 %</strong> 1.92 [1.46, 2.52]</td>
<td></td>
</tr>
</tbody>
</table>

Total events: 110 (Psychotherapy), 53 (Usual care)

Heterogeneity: \( \tau^2 = 0.0; \ \chi^2 = 4.38, \ df = 5 (P = 0.50); \ I^2 = 0.0\%

Test for overall effect: \( Z = 4.66 (P < 0.00001) \)

Test for subgroup differences: Not applicable

---

*Psychological therapies for treatment-resistant depression in adults (Review)*

Copyright © 2018 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.
Analysis 1.17. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 17 Remission (< 7 on HAMD or < 10 on BDI) medium term (7 to 12 months).

Review: Psychological therapies for treatment-resistant depression in adults

Comparison: 1 Psychotherapy with usual care versus usual care alone

Outcome: 17 Remission (< 7 on HAMD or < 10 on BDI) medium term (7 to 12 months)

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Risk Ratio M H, Random, 95% CI</th>
<th>Weight</th>
<th>Risk Ratio M H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nakagawa 2017</td>
<td>29/40</td>
<td>17/40</td>
<td></td>
<td>41.3%</td>
<td>1.71 [1.13, 2.56]</td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>78/197</td>
<td>36/198</td>
<td></td>
<td>58.7%</td>
<td>2.18 [1.55, 3.07]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>237</td>
<td>238</td>
<td></td>
<td>100.0%</td>
<td>1.97 [1.51, 2.56]</td>
</tr>
</tbody>
</table>

Total events: 107 (Psychotherapy), 53 (Usual care)
Heterogeneity: Tau^2 = 0.0; Chi^2 = 0.85, df = 1 (P = 0.36); I^2 = 0.0%
Test for overall effect: Z = 5.07 (P < 0.00001)
Test for subgroup differences: Not applicable

0.01 0.1 1 10 100
Favours psychotherapy Favours usual care

Analysis 1.18. Comparison 1 Psychotherapy with usual care versus usual care alone, Outcome 18 Remission (< 7 on HAMD or < 10 on BDI) long term (longer than 12 months).

Review: Psychological therapies for treatment-resistant depression in adults

Comparison: 1 Psychotherapy with usual care versus usual care alone

Outcome: 18 Remission (< 7 on HAMD or < 10 on BDI) long term (longer than 12 months)

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Psychotherapy</th>
<th>Usual care</th>
<th>Risk Ratio M H, Random, 95% CI</th>
<th>Weight</th>
<th>Risk Ratio M H, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wiles 2016</td>
<td>38/136</td>
<td>20/112</td>
<td></td>
<td>100.0%</td>
<td>1.56 [0.97, 2.53]</td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td>136</td>
<td>112</td>
<td></td>
<td>100.0%</td>
<td>1.56 [0.97, 2.53]</td>
</tr>
</tbody>
</table>

Total events: 38 (Psychotherapy), 20 (Usual care)
Heterogeneity: not applicable
Test for overall effect: Z = 1.83 (P = 0.068)
Test for subgroup differences: Not applicable

0.01 0.1 1 10 100
Favours psychotherapy Favours usual care

Psychological therapies for treatment-resistant depression in adults (Review)

Copyright © 2018 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.
### ADDITIONAL TABLES

#### Table 1. Psychotherapy with usual care versus usual care alone - reasons for dropout

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Total N randomised</th>
<th>Follow-up time point, months</th>
<th>Reason for dropout given in Intervention group (psychotherapy as an adjunct to usual care)</th>
<th>Reason for dropout given in control group (usual care alone)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harley 2008</td>
<td>24</td>
<td>4</td>
<td>1 difficulty finding childcare; 1 work schedule conflict; 1 decided group was not a good fit</td>
<td>1 moved; 1 medical problem</td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>469</td>
<td>6</td>
<td>25 not followed up 14 withdrew from study 6 lost to follow-up 4 unable to contact 1 died</td>
<td>22 not followed up 13 withdrew from study 6 lost to follow-up 3 unable to contact</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>36 not followed up 17 withdrew from study 17 lost to follow-up 2 died</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>37 not followed up 15 withdrew from study 22 lost to follow-up</td>
</tr>
<tr>
<td>Wiles 2007</td>
<td>25</td>
<td>4</td>
<td>NA</td>
<td>2 lost to follow-up</td>
</tr>
<tr>
<td>Nakagawa 2017</td>
<td>80</td>
<td>6</td>
<td>1 not contactable; 1 patient discontinued because of lumbago</td>
<td>1 not contactable; patient discontinued owing to family health problem</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>1 not contactable</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1 not contactable; 1 died</td>
</tr>
<tr>
<td>Town 2017</td>
<td>60</td>
<td>6</td>
<td>2 did not start therapy; 3 not contactable; 2 withdrew</td>
<td>3 not contactable</td>
</tr>
</tbody>
</table>

N: number
NA: not available
Table 2. Psychotherapy with usual care versus usual care alone for social functioning

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Measure</th>
<th>Time point (months)</th>
<th>N psychotherapy + usual care</th>
<th>N usual care</th>
<th>Mean difference</th>
<th>95% CI lower</th>
<th>95% CI upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harley 2008</td>
<td>SAS work&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4</td>
<td>14</td>
<td>9</td>
<td>1.20</td>
<td>-1.61</td>
<td>4.01</td>
</tr>
<tr>
<td>Harley 2008</td>
<td>LIFE work&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6</td>
<td>201</td>
<td>209</td>
<td>6</td>
<td>3.5</td>
<td>8.2</td>
</tr>
<tr>
<td>Harley 2008</td>
<td>LIFE recreation&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6</td>
<td>194</td>
<td>195</td>
<td>4.1</td>
<td>1.6</td>
<td>6.7</td>
</tr>
</tbody>
</table>

N: number
P: P value
SD: standard deviation

<sup>a</sup>Cohen's D > 0.5 is moderate effect and > 0.8 is large effect.
<sup>b</sup>SAS-SR and LIFE-RIFT (SAS work/social recreational, LIFE work/recreation/satisfaction): Lower scores are healthier.
<sup>c</sup>Schwartz Outcome Scale-10 (SOS-10): Higher scores are healthier.

Table 3. Psychotherapy with usual care versus usual care alone for quality of life

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Measure</th>
<th>Time point (months)</th>
<th>N psychotherapy + usual care</th>
<th>N usual care</th>
<th>Mean difference</th>
<th>95% CI lower</th>
<th>95% CI upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wiles 2007</td>
<td>Unpublished tool&lt;sup&gt;a&lt;/sup&gt;</td>
<td>4</td>
<td>14</td>
<td>9</td>
<td>1.20</td>
<td>-1.61</td>
<td>4.01</td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>SF-12 mental&lt;sup&gt;b&lt;/sup&gt;</td>
<td>6</td>
<td>201</td>
<td>209</td>
<td>6</td>
<td>3.5</td>
<td>8.2</td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>SF-12 mental&lt;sup&gt;b&lt;/sup&gt;</td>
<td>12</td>
<td>194</td>
<td>195</td>
<td>4.1</td>
<td>1.6</td>
<td>6.7</td>
</tr>
<tr>
<td>Source</td>
<td>Measure</td>
<td>N</td>
<td>Mean</td>
<td>CI</td>
<td>SD</td>
<td>SE</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------</td>
<td>----</td>
<td>------</td>
<td>-------</td>
<td>------</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>SF-12 mental b</td>
<td>46</td>
<td>132</td>
<td>110</td>
<td>3.5</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SF-12 physical b</td>
<td>6</td>
<td>201</td>
<td>209</td>
<td>−1.7</td>
<td>−3.4</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SF-12 physical b</td>
<td>12</td>
<td>194</td>
<td>195</td>
<td>0.3</td>
<td>−1.4</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiles 2016</td>
<td>SF-12 physical b</td>
<td>46</td>
<td>132</td>
<td>110</td>
<td>0.9</td>
<td>−2</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Souza 2016</td>
<td>WHOQOL overall QOL c</td>
<td>6</td>
<td>16</td>
<td>18</td>
<td>0.80</td>
<td>−2.67</td>
<td>4.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Souza 2016</td>
<td>WHOQOL physical c</td>
<td>6</td>
<td>16</td>
<td>18</td>
<td>7.10</td>
<td>−3.04</td>
<td>17.24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Souza 2016</td>
<td>WHOQOL psychological c</td>
<td>6</td>
<td>16</td>
<td>18</td>
<td>3.00</td>
<td>−8.51</td>
<td>14.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Souza 2016</td>
<td>WHOQOL social c</td>
<td>6</td>
<td>16</td>
<td>18</td>
<td>6.50</td>
<td>−6.71</td>
<td>19.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nakagawa 2017</td>
<td>SF-36 mental b</td>
<td>6</td>
<td>40</td>
<td>40</td>
<td>−2.32</td>
<td>−7.25</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nakagawa 2017</td>
<td>SF-36 mental b</td>
<td>12</td>
<td>40</td>
<td>40</td>
<td>−1.27</td>
<td>−6.26</td>
<td>3.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nakagawa 2017</td>
<td>SF-36 physical b</td>
<td>6</td>
<td>40</td>
<td>40</td>
<td>−1.17</td>
<td>−6.46</td>
<td>3.81</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nakagawa 2017</td>
<td>SF-36 physical b</td>
<td>12</td>
<td>40</td>
<td>40</td>
<td>0.95</td>
<td>−4.4</td>
<td>6.82</td>
</tr>
</tbody>
</table>

CI: confidence interval
N: number
A 6-item instrument (unpublished) on which one could score between zero and 12: lower scores denote poorer QOL
SF physical/mental: Higher score denotes better quality of life
WHOQOL: Higher scores denote higher quality of life.
<table>
<thead>
<tr>
<th>Study ID</th>
<th>Outcome</th>
<th>Measure</th>
<th>Time point, months</th>
<th>N psychotherapy + usual care</th>
<th>N usual care</th>
<th>N usual care with outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nakagawa 2017</td>
<td>Serious adverse event</td>
<td>Hospitalisation due to depression exacerbation</td>
<td>12</td>
<td>40</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nakagawa 2017</td>
<td>Serious adverse event</td>
<td>Suicide</td>
<td>6</td>
<td>40</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Town 2017</td>
<td>Adverse event</td>
<td>Increases in depressive symptoms</td>
<td>6</td>
<td>30</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

N: number

**APPENDICES**

**Appendix 1. Search strategy**

**1 MEDLINE search strategy**

Database: Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) <1946 to Present>

1 Depressive Disorder, Treatment-Resistant/ (726)
2 (depress* and ((antidepress* or SSRI* or SNRI* or (serotonin adj3 (uptake or reuptake or re-uptake)) or medication* or psychotropic or treatment* or respon*) adj2 fail*)).ti,ab,kf. (1488)
3 (depress* and ((antidepress* or SSRI* or SNRI* or (serotonin adj3 (uptake or reuptake or re-uptake)) or psychotropic medication* or treatment*) adj2 ("no respon*" or "not respon*" or nonrespon* or non-respon* or unrespon*)).ti,ab,kf. (561)
4 (depress* adj3 (refractor* or resistan* or chronic* or persist*)).ti,ab,kf. (10275)
5 (depress* adj3 (relaps* or recur*)).ti,kf. (1201)
6 (depress* and (augment* or potentiat*)).mp. (14786)
7 or/1-6 (26920)
8 randomized controlled trial.pt. or exp randomized controlled trial/ or exp Randomized Controlled Trials as Topic/ (570582)
9 controlled clinical trial.pt. (94063)
10 (RCT or randomi#ed or at random or (random* adj3 (assign* or allocat* or divide* or division or number*)]).ti,ab,kf. (645282)
11 ((placebo or sham or mock or fake or dummy) and (control* or group?)).ti,ab,kf. (211689)
12 double-blind*.ti,ab,kf,hw. (185041)
13 trial.ti. (181076)
14 ((cluster or crossover* or cross-over*) adj3 (random* or trial or study or control* or group))).ti,ab,kf. (62904)
15 or/8-14 (1056236)
16 7 and 15 (3636)
17 letter/ (970850)
18 editorial/ (439091)
19 news/ (183195)
20 exp historical article/ (382466)
21 Anecdotes as topic/ (4929)
22 comment/ (690728)
23 case report/ (1883468)
24 (letter or comment*).ti. (123606)
25 exp animals/ not humans/ (4399234)
26 exp Animals, Laboratory/ (798339)
27 exp Animal Experimentation/ not (exp human experimentation/ or humans/) (4881)
28 exp Models, Animal/ (492528)
29 exp rodentia/ (2975987)
30 (rat or rats or mouse or mice or rodent*).ti. (1264597)
31 or/17-30 (8908789)
32 16 not 31 (1264597)
33 (2016* or 2017*).yr,dc,ed,ep. (2581993)
34 (in-data-review or in-process or publisher).st. (1331518)
35 33 or 34 (2894741)
36 32 and 35 (553)

2 Embase search strategy

1 treatment resistant depression/
2 (depress* and ((antidepress* or SSRI* or SNRI* or (serotonin adj3 (uptake or reuptake or re-uptake)) or medication* or psychotropic or treatment* or respon*) adj2 fail*)).ti,ab,kf.
3 (depress* and ((antidepress* or SSRI* or SNRI* or (serotonin adj3 (uptake or reuptake or re-uptake)) or psychotropic medication* or treatment*) adj2 ("no respon*" or "not respon*" or nonrespon* or non-respon* or unrespon*))).ti,ab,kf.
4 (depress* adj3 (refractor* or resistant* or chronic* or persist*)).ti,ab,kf.
5 (depress* adj3 (relaps* or recurr*)).ti,kf.
6 (depress* and (augment* or potentiat*)).mp.
7 or/1-6
8 randomized controlled trial/ or "randomized controlled trial (topic)="/n
10 "double blind procedure="/n
11 "single-blind procedure="/n
12 (RCT or randomi#ed or at random or (random* adj3 (assign* or allocat* or divide* or division or number*))).ti,ab,kf.
13 trial.ti.
14 ((cluster or crossover* or cross-over*) adj3 (random* or trial or study or control* or group))).ti,ab,kf.
15 double-blind*.ti,ab.
16 ((placebo or sham or mock or fake or dummy) and (control* or group?)).ti,ab,kf.
17 or/8-16
18 7 and 17
19 letter.pt. or letter/
20 note.pt.
21 editorial.pt.
22 case report/ or case study/
23 (letter or comment*).ti.
24 exp animal/ not human/
25 nonhuman/
26 exp experimental animal/
27 exp animal experiment/
28 exp animal model/
29 exp rodent/
30 (rat or rats or mouse or mice or rodent*).ti.
31 or/19-30
32 18 not 31

3 PsycINFO search strategy
1 (depress* and ((antidepress* or SSRI* or SNRI* or (serotonin adj3 (uptake or reuptake or re-uptake)) or medication* or psychotropic or treatment* or respon*) adj2 fail*)).ti,ab,id.
2 (depress* and ((antidepress* or SSRI* or SNRI* or (serotonin adj3 (uptake or reuptake or re-uptake)) or psychotropic medication* or treatment*) adj2 ("no respon*" or "not respon*" or nonrespon* or non-respon* or unrespon*))).ti,ab,id.
3 (depress* adj3 (refractor* or resistan* or chronic* or persist*))).ti,ab,id.
4 (depress* and (augment* or potentiat*)).mp.
5 treatment resistant depression/
6 (depress* adj3 (relaps* or recurr*)).ti,id.
7 or/1-6
8 clinical trials/
9 (RCT or randomi#ed or at random or (random* adj3 (assign* or allocat* or divide* or division or number*))).ti,ab,id.
10 double-blind*.ti,ab,id,hw.
11 ((placebo or sham or mock or fake or dummy) and (control* or group?)).ti,ab,id.
12 trial.ti.
13 ((cluster or crossover* or cross-over*) adj3 (random* or trial or study or control* or group?)).ti,ab,id.
14 or/8-13
15 7 and 14
17 scientific communication/
18 case report/
19 (letter or comment*).ti.
20 exp animals/ or animal models/
21 (rat or rats or mouse or mice or rodent*).ti.
22 or/16-21
23 15 not 22

4 CENTRAL search strategy
#1MeSH descriptor: [Depressive Disorder, Treatment-Resistant] explode all trees
#2(depress* and ((antidepress* or SSRI* or SNRI* or (serotonin near/3 (uptake or reuptake or re-uptake)) or medication* or psychotropic or treatment* or respon*) near/2 fail*)):ti,ab,kw
#3(depress* and ((antidepress* or SSRI* or SNRI* or (serotonin near/3 (uptake or reuptake or re-uptake)) or "psychotropic medication" or "psychotropic medications" or treatment*) near/2 ("no respon*" or "not respon*" or nonrespon* or non-respon* or unrespon*)):ti,ab,kw
#4(depress* near/3 (refractor* or resistan* or chronic* or persist*)):ti,ab,kw
#5(depress* near/3 (relaps* or recurr*)):ti,kw
#6(depress* and (augment* or potentiat*)):ti,ab,kw
#7#1 or #2 or #3 or #4 or #5 or #6
5 Web of Science search strategy

# 15 #11 not #14
# 14 #13 OR #12
# 13 TS=((animal* near/2 experiment*) or (animal* near/2 model*) or (animal* near/2 laborator*))
# 12 TI= (rat or rats or mouse or mice or rodent* or animal* or comment* or letter or “case study” or “case report” or anecdote* or editorial* or news )
# 11 #10 AND #6
# 10 #9 OR #8 OR #7
# 9 TI= trial
# 8 TS= (RCT or randomized or randomised or "at random" or (random* near/3 (assign* or allocat* or divide* or division or number*))
# 7 TS= ((controlled near/2 "clinical trial") or double-blind* or ((placebo or sham or mock or fake or dummy) and (control* or group?)) or ((cluster or crossover* or cross-over*) near/3 (random* or trial or study or control* or group?)))
# 6 #5 OR #4 OR #3 OR #2 OR #1
# 5 TS=(depress* and (augment* or potentiat*))
# 4 TS=(depress* near/3 (relaps* or recur*))
# 3 TS=(depress* near/3 (refractor* or resistan* or chronic* or persist*))
# 2 TS=(depress* and ((antidepress* or SSRI* or SNRI* or (serotonin near/3 (uptake or reuptake or re-uptake)) or "psychotropic medication" or "psychotropic medications" or treatment*) near/2 ("no respon*” or "not respon*” or nonrespon* or non-respon* or unrespon*))
# 1 TS=((depress* and ((antidepress* or SSRI* or SNRI* or (serotonin near/3 (uptake or reuptake or re-uptake)) or medication* or psychotropic or treatment* or respon*) near/2 fail*)))

6 Trial registry search strategy

Types of Study=Interventional
Condition 1= treatment resistant depression
Condition 2= refractory depression
Condition 3= recurrent depression
Condition 4= chronic depression

Search details for 2017 May updates

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review area</td>
<td>Treating treatment-resistant depression</td>
</tr>
<tr>
<td>Objectives</td>
<td>To identify which pharmacological and/or psychological therapies are effective for treatment-resistant depression</td>
</tr>
<tr>
<td>Populations/aspect</td>
<td>Adults with treatment-resistant depression</td>
</tr>
<tr>
<td>Interventions</td>
<td>Pharmacological and/or psychological therapies</td>
</tr>
<tr>
<td>Study design</td>
<td>RCT/cluster/cross-over</td>
</tr>
<tr>
<td>Exclusions</td>
<td>Animal studies/editorials/anecdotes/case reports/letters</td>
</tr>
<tr>
<td>How the information was searched</td>
<td>Databases: MEDLINE, Premedline, Embase, Cochrane Library, PsycInfo, Web of Science Language: all Dates: 2016 to date</td>
</tr>
</tbody>
</table>
Date searched | 16 May 2017
---|---

Search results | MEDLINE/Premedline = 553  
Embase = 546  
Cochrane = 477  
PsycInfo = 246  
Web of Science = 673  
Total = 2495  
Total de-duplicated = 1309  
With previously seen references removed = 1193  
Trial registers search update June 2017 - 678 (clinicaltrials.gov), 67 (WHO ICTRP)

**Contributions of Authors**

NW drafted the protocol, which was finalised following comments from all protocol authors. NW undertook abstract screening for the primary search, and PD and SI undertook update screening. All review authors contributed to extraction of data from papers included in the review. SI wrote the first draft of the review, which was commented upon by all review authors. NW is the guarantor of the review.

CW died in Spring 2018 and the living authors did not make substantive changes to the review beyond this author’s contribution.

**Declarations of Interest**

NW was the Chief Investigator of the National Institute for Health Research HTA-funded CoBalT trial (CBT as an adjunct to pharmacotherapy for TRD in primary care: ISRCTN38231611), and DK and GL were Principal Investigators. NW, DK, and GL are authors on two of the studies included in this review (*Wiles 2007*; *Wiles 2016*).

PD and SI have no conflicts to declare.

CW is deceased; declarations of interest published in the protocol: “CW has no conflicts to declare”.

**Sources of Support**

**Internal sources**
- Cochrane Common Mental Disorders Review Group, UK.
External sources

- NIHR CLAHRC West, UK. This research was supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care West (NIHR CLAHRC West). The views expressed in this article are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.
- NIHR Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust and the University of Bristol, UK. This research was also supported by the NIHR Biomedical Research Centre at University Hospitals Bristol NHS Foundation Trust and the University of Bristol. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health and Social Care.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

In the protocol, we had planned that missing data would be addressed in additional analyses assuming best (all who dropped out had positive outcomes) and worst (all who dropped out had negative outcomes) case scenarios. However, we later agreed that this was not necessary because study level data were more robust than participant level data imputations determined by review authors.

We have reported comparison 5 (any psychological therapy vs an attention control) in the methods of the review; this was not stated in the protocol. This occurred because attention control was listed as a comparator intervention in the section Types of interventions but was missed in error under planned comparisons in the protocol stage. We found no studies for this comparison.

In the protocol, we had said we were going to calculate and convert the odds ratio (OR) for each study to risk ratio (RR). However, we found that this was not needed because studies reported event data for dichotomous outcomes in full (only two studies additionally reported OR). We did calculate OR for each study but found that these values were the same as RR figures, and since the final presentations were to include RRs, we have chosen to forego presenting ORs.