



**Cochrane**  
**Library**

Cochrane Database of Systematic Reviews

## Antidepressant treatment for postnatal depression (Review)

Molyneaux E, Howard LM, McGeown HR, Karia AM, Trevillion K

Molyneaux E, Howard LM, McGeown HR, Karia AM, Trevillion K.  
Antidepressant treatment for postnatal depression.  
*Cochrane Database of Systematic Reviews* 2014, Issue 9. Art. No.: CD002018.  
DOI: [10.1002/14651858.CD002018.pub2](https://doi.org/10.1002/14651858.CD002018.pub2).

[www.cochranelibrary.com](http://www.cochranelibrary.com)

## TABLE OF CONTENTS

ABSTRACT .....	1
PLAIN LANGUAGE SUMMARY .....	2
SUMMARY OF FINDINGS .....	4
BACKGROUND .....	6
OBJECTIVES .....	7
METHODS .....	7
RESULTS .....	12
Figure 1. ....	13
Figure 2. ....	17
Figure 3. ....	18
Figure 4. ....	21
Figure 5. ....	21
DISCUSSION .....	25
AUTHORS' CONCLUSIONS .....	28
ACKNOWLEDGEMENTS .....	28
REFERENCES .....	29
CHARACTERISTICS OF STUDIES .....	33
DATA AND ANALYSES .....	47
Analysis 1.1. Comparison 1 Selective serotonin re-uptake inhibitors versus placebo, Outcome 1 Response rate at post-treatment. ....	47
Analysis 1.2. Comparison 1 Selective serotonin re-uptake inhibitors versus placebo, Outcome 2 Remission rate at post-treatment. ....	47
Analysis 2.1. Comparison 2 Antidepressants versus treatment as usual, Outcome 1 Remission rate at post-treatment. ....	48
Analysis 3.1. Comparison 3 Antidepressants versus psychosocial therapy (listening visits), Outcome 1 Remission rate at post-treatment. ....	48
Analysis 4.1. Comparison 4 Selective serotonin re-uptake inhibitors versus other pharmacological intervention (tricyclic antidepressant), Outcome 1 Response rate at post-treatment. ....	49
Analysis 4.2. Comparison 4 Selective serotonin re-uptake inhibitors versus other pharmacological intervention (tricyclic antidepressant), Outcome 2 Remission rate at post-treatment. ....	49
Analysis 5.1. Comparison 5 Sensitivity analysis: excluding trials with combined treatment, Outcome 1 SSRIs versus placebo: outcome 1.1 response rate at post-treatment. ....	49
Analysis 5.2. Comparison 5 Sensitivity analysis: excluding trials with combined treatment, Outcome 2 SSRIs versus placebo: outcome 1.2 remission rate at post-treatment. ....	50
Analysis 6.1. Comparison 6 Sensitivity analysis: removing studies with high dropout or high risk of bias in any domain, Outcome 1 SSRIs versus placebo: outcome 1.1 response rate at post-treatment. ....	50
Analysis 6.2. Comparison 6 Sensitivity analysis: removing studies with high dropout or high risk of bias in any domain, Outcome 2 SSRIs versus placebo: outcome 1.2 remission rate at post-treatment. ....	50
ADDITIONAL TABLES .....	51
WHAT'S NEW .....	52
HISTORY .....	53
CONTRIBUTIONS OF AUTHORS .....	53
DECLARATIONS OF INTEREST .....	53
DIFFERENCES BETWEEN PROTOCOL AND REVIEW .....	53
INDEX TERMS .....	54

[Intervention Review]

# Antidepressant treatment for postnatal depression

Emma Molyneaux<sup>1</sup>, Louise M Howard<sup>1</sup>, Helen R McGeown<sup>1</sup>, Amar M Karia<sup>2</sup>, Kylee Trevillion<sup>1</sup>

<sup>1</sup>Health Service and Population Research Department, The Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK. <sup>2</sup>School of Medicine, King's College London, London, UK

**Contact:** Louise M Howard, Health Service and Population Research Department, The Institute of Psychiatry, Psychology & Neuroscience, King's College London, PO31 De Crespigny Park, London, SE5 8AF, UK. [louise.howard@kcl.ac.uk](mailto:louise.howard@kcl.ac.uk).

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

**Publication status and date:** New search for studies and content updated (conclusions changed), published in Issue 9, 2014.

**Citation:** Molyneaux E, Howard LM, McGeown HR, Karia AM, Trevillion K. Antidepressant treatment for postnatal depression. *Cochrane Database of Systematic Reviews* 2014, Issue 9. Art. No.: CD002018. DOI: [10.1002/14651858.CD002018.pub2](https://doi.org/10.1002/14651858.CD002018.pub2).

Copyright © 2014 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

## ABSTRACT

### Background

Postnatal depression is a common disorder that can have adverse short- and long-term effects on maternal morbidity, the new infant and the family as a whole. Treatment is often largely by social support and psychological interventions. It is not known whether antidepressants are an effective and safe choice for treatment of this disorder. This review was undertaken to evaluate the effectiveness of different antidepressants and to compare their effectiveness with other forms of treatment, placebo or treatment as usual. It is an update of a review first published in 2001.

### Objectives

To assess the effectiveness of antidepressant drugs in comparison with any other treatment (psychological, psychosocial or pharmacological), placebo or treatment as usual for postnatal depression.

### Search methods

We searched the Cochrane Depression, Anxiety and Neurosis Group's Specialized Register (CCDANCTR) to 11 July 2014. This register contains reports of relevant randomised controlled trials (RCTs) from the following bibliographic databases: *The Cochrane Library* (all years), MEDLINE (1950 to date), EMBASE, (1974 to date) and PsycINFO (1967 to date). We also searched international trial registries and contacted pharmaceutical companies and experts in the field.

### Selection criteria

We included RCTs of women with depression with onset up to six months postpartum that compared antidepressant treatment (alone or in combination with another treatment) with any other treatment, placebo or treatment as usual.

### Data collection and analysis

Two review authors independently extracted data from the trial reports. We requested missing information from investigators wherever possible. We sought data to allow an intention-to-treat analysis. Random effects meta-analyses were conducted to pool data where sufficient comparable studies were identified.

### Main results

We included six trials with 596 participants in this review. All studies had a randomised controlled parallel group design, with two conducted in the UK, three in the US and one in Israel. Meta-analyses were performed to pool data on response and remission from studies comparing antidepressants with placebo. No meta-analyses could be conducted for other comparisons due to the small number of trials identified.

Four studies compared selective serotonin reuptake inhibitors (SSRIs) with placebo (two using sertraline, one using paroxetine and one using fluoxetine; 233 participants in total). In two of these studies both the experimental and placebo groups also received psychological therapy. Pooled risk ratios based on data from three of these studies (146 participants) showed that women randomised to SSRIs had higher rates of response and remission than those randomised to placebo (response: RR 1.43, 95% CI 1.01 to 2.03; remission: RR 1.79, 95% CI 1.08 to 2.98); the fourth study did not report data on response or remission.

One study (254 participants) compared antidepressant treatment with treatment as usual (for the first four weeks) followed by listening visits. The study found significantly higher rates of improvement in the antidepressant group than treatment-as-usual group after the first four weeks, but no difference between antidepressants and listening visits at the later follow-up. In addition, one study comparing sertraline with nortriptyline (a tricyclic antidepressant) found no difference in effectiveness (109 participants).

Side effects were experienced by a substantial proportion of women, but there was no evidence of a meaningful difference in the number of adverse effects between treatment arms in any study. There were very limited data on adverse effects experienced by breastfed infants, with no long-term follow-up. All but one of the studies were assessed as being at high or uncertain risk of attrition bias and selective outcome reporting. In particular, one of the placebo-controlled studies had over 50% drop-out.

### Authors' conclusions

The evidence base for this review was very limited, with a small number of studies and little information on a number of important outcomes, particularly regarding potential effects on the child. Risk of bias, for example from high attrition rates, as well as low representativeness of participants (e.g. exclusion of women with severe or chronic depression in several trials) also limit the conclusions that can be drawn.

Pooled estimates for response and remission found that SSRIs were significantly more effective than placebo for women with postnatal depression. However the quality of evidence contributing to this comparison was assessed as very low owing to the small sample size for this comparison (146 participants from three studies), the risk of bias in included studies and the inclusion of one study where all participants in both study arms additionally received psychological therapy. There was insufficient evidence to conclude whether, and for whom, antidepressant or psychological/psychosocial treatments are more effective, or whether some antidepressants are more effective or better tolerated than others. There is also inadequate evidence on whether the benefits of antidepressants persist beyond eight weeks or whether they have short- or long-term adverse effects on breastfeeding infants.

Professionals treating women with severe depression in the postnatal period will need to draw on other evidence, including trials among general adult populations and observational studies of antidepressant safety when breastfeeding (although the potential for confounding in non-randomised studies must be considered). More RCTs are needed with larger sample sizes and longer follow-up, including assessment of the impact on the child and safety of breastfeeding. Further larger-scale trials comparing antidepressants with alternative treatment modalities are also required.

## PLAIN LANGUAGE SUMMARY

### Antidepressants for postnatal depression

#### Why is this review important?

Postnatal depression is a common disorder that can have short- and long-term adverse effects on the mother, the new infant and the family as a whole. Antidepressants are commonly used as the first treatment option for adults with moderate to severe depression, but there is little evidence on whether antidepressants are an effective and safe choice for the treatment of this disorder in the postnatal period. This review was undertaken to evaluate the effectiveness of different antidepressants and to compare their effectiveness with other forms of treatment (e.g. psychosocial interventions such as peer support, psychological interventions such as cognitive behavioural therapy), placebo or treatment as usual.

#### Who will be interested in this review?

Parents, professionals in primary care services who work with women of reproductive age, general practitioners, professionals in adult mental health services who work with women of reproductive age and professionals working in perinatal mental health services.

#### What questions does this review aim to answer?

This review is an update of a previous Cochrane review from 2001, which found insufficient evidence to make conclusions about antidepressant treatment in postnatal depression. Therefore, this update aims to answer the following question:

What are the effects of antidepressants in comparison with other any other treatment, placebo or treatment as usual for postnatal depression?

#### Which studies were included in the review?

We searched clinical trials registries; the Cochrane Depression, Anxiety and Neurosis Group; and the Cochrane Pregnancy and Childbirth Group databases to find all high-quality studies comparing antidepressants with any other form of treatment from the upper date limit of the most recent previous searches to July 2014. We contacted drug companies and experts in the field.

To be included in the review, studies had to be randomised controlled trials (clinical studies where people were randomly put into one of two or more treatment groups) and had to include women with postnatal depression (onset of depression up to six months after giving birth) who were not taking any antidepressant medication at the start of the trial.

We included six trials of 596 women in the review. Although many of the studies were well conducted and reported, there are some areas with substantial risk of bias; for example, through incomplete follow-up (e.g. in one study over 50% of the participants dropped out prior to the primary outcome measurement).

### **What does the evidence from the review tell us?**

The quality of evidence from this review was assessed as being very low quality due to the small number of studies, risk of bias in the included studies (in particular, high proportions of participants dropped out) and the fact that many studies excluded women with chronic (i.e. long lasting) or severe depression, or both. We were able to combine data from three studies comparing a type of commonly used antidepressant called selective serotonin reuptake inhibitors (SSRIs) with placebo. The results showed that women with postnatal depression who were given SSRIs were more likely to improve or recover than those given placebo. We were unable to combine the data from studies comparing antidepressants with other treatments or treatment as usual due to the very small number of studies identified for these comparisons. There was insufficient evidence to conclude whether, and for whom, antidepressant or psychosocial/psychological treatments are more effective, or whether some antidepressants are more effective or better tolerated (or both) than others. Conclusions were also limited by the lack of data on long-term follow-up, the safety of breastfeeding or child outcomes.

### **What should happen next?**

Larger studies need to be done, and treatment decisions for women with postnatal depression will need to use evidence from other sources such as trials in general adult populations and observational studies of antidepressant safety in the postnatal period. The review authors recommend that future studies in this area should include women with severe postnatal depression, long-term follow-up on psychiatric symptoms and quality of life in mothers who have been treated for postnatal depression. In addition, more evidence is needed on outcomes for infants, particularly with regards to the safety of breastfeeding and effect of treatment for postnatal depression on the maternal-infant relationship.

## SUMMARY OF FINDINGS

### Summary of findings for the main comparison.

#### Selective serotonin reuptake inhibitors (SSRIs) compared with placebo for postnatal depression

**Patient or population:** women with postnatal depression

**Intervention:** selective serotonin reuptake inhibitors (SSRIs)

**Comparison:** placebo

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of Participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	Placebo	SSRIs				
<b>Response rate at post-treatment</b>  (as defined in individual studies)	<b>365 per 1000<sup>1</sup></b>	<b>522 per 1000</b> (369 to 741)	<b>RR 1.43</b> (1.01 to 2.03)	146 (3 studies)	⊕⊕⊕⊕ <b>very low<sup>2, 3,4</sup></b>	Yonkers 2008: response: CGI-II ≤ 2 (at 8 weeks)  Hantsoo 2013: response: < 10 HAM-D + at least 50% decrease in HAM-D score from baseline + CGI ≤ 2 (after 6 weeks of treatment)  Bloch 2012: response: > 50% reduction in MADRS or EPDS score during treatment (at 8 weeks)
<b>Remission rate at post-treatment</b>  (as defined in individual studies)	<b>257 per 1000<sup>1</sup></b>	<b>460 per 1000</b> (278 to 766)	<b>RR 1.79</b> (1.08 to 2.98)	146 (3 studies)	⊕⊕⊕⊕ <b>very low<sup>2, 3,4</sup></b>	Yonkers 2008: remission: HAM-D ≤ 8 (at 8 weeks)  Hantsoo 2013: remission: as Hantsoo response above + HAM-D < 7  Bloch 2012: remission: final score < 10 on the MADRS scale or < 7 on the EPDS (at 8 weeks)

\*The basis for the **assumed risk** is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

**CI:** Confidence interval; **RR:** Risk Ratio; **CGI:** Clinical Global Improvement; **EPDS:** Edinburgh Postnatal Depression Scale; **HAM-D:** Hamilton Rating Scale for Depression; **MADRS:** Montgomery-Åsberg Depression Rating Scale

GRADE Working Group grades of evidence

**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low quality:** We are very uncertain about the estimate.

---

- <sup>1</sup>assumed risk calculated as the proportion of women on placebo with the outcome (response or remission) in the three included studies, multiplied by 1000.
- <sup>2</sup>downgraded due to indirectness (in one of the studies included in the meta-analysis participants in both arms additionally received brief dynamic psychotherapy).
- <sup>3</sup>downgraded due to risk of bias (incomplete outcome data owing to loss to follow-up)
- <sup>4</sup>downgraded due to high imprecision (wide confidence intervals owing to the small number and small samples of included studies)

## BACKGROUND

### Description of the condition

Postnatal depression is an important and common disorder that can have short- and long-term adverse impacts on the mother, her child and the family as a whole (Letourneau 2012; Murray 1992). Postnatal depression is characterised by persistent low mood and loss of pleasure or interests, occurring with associated symptoms such as changes in appetite, psychomotor agitation or retardation, disturbed sleep and low self confidence (WHO 1992).

Both the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) (APA 2013) and the International Classification of Disease Tenth Revision (ICD-10) (WHO 2004) include postnatal depression within the standard diagnostic criteria for depression, and postnatal depression has been found to have similar phenomenology to depression in women who have not recently given birth (Nylen 2013). Somatic symptoms of depression, such as sleep disturbance and loss of libido, may occur as part of the normative postpartum experience, but evidence suggests that they are more commonly reported among depressed than non-depressed women in the postnatal period (with the exception of appetite change, Nylen 2013). An onset specifier for depression is used in both the ICD-10 (within six weeks of delivery) and DSM-5 (onset during pregnancy or in the four weeks following delivery), but many researchers use time limits between three and six months' postpartum (Munk-Olsen 2006).

One comprehensive systematic review of perinatal depression reported a prevalence of 4.7% for major depression at three months postpartum and 12.9% including minor (sub-threshold) depression (Gavin 2005), similar to estimates for adult women at non-childbearing times. However, there may be an increased risk of new episodes of depression in the period following childbirth; Cox 1993 found a three-fold increase in the incidence of depression in the first five weeks after delivery. More recent studies using medical records have supported a peak incidence in the first postpartum months (Ban 2012; Munk-Olsen 2006), although it must be noted that a substantial proportion of postnatal depression episodes begin during pregnancy or prior to conception (Wisner 2013). Most women with postpartum depression recover within a few months but about 30% of episodes last beyond the first postpartum year (Goodman 2004). Women who have had postnatal depression also have a high risk (about 40%) of both postnatal and non-postnatal relapse (Cooper 1995; Wisner 2004).

It is important that classifications distinguish postpartum depression from both the 'baby blues' and postpartum psychosis, which also occur following childbirth. The 'baby blues' are characterised by sub-threshold symptoms of depression (e.g. insomnia, fatigue, tearfulness, anxiety, irritability, impairment of concentration and mood lability) occurring soon after delivery. Prevalence estimates range from 15% to 85% among postpartum women (often around 50%, Henshaw 2003), but symptoms are usually mild and resolve within days. In contrast, postpartum psychosis is a very severe condition that affects a small proportion of postpartum women (about 2 per 1000) (Kendell 1987). Women with postpartum psychosis may present with mania, psychotic depression, schizophrenia or confusional states and in most cases, hospitalisation is indicated.

### Description of the intervention

In light of the influence of social factors, psychosocial and psychological interventions to improve outcomes for women with postnatal depression have been developed and evaluated. Reductions in depression have been identified following a range of psychosocial and psychological interventions (e.g. non-directive counselling, telephone-based peer support and cognitive behavioural therapy (CBT)) compared with usual care (Dennis 2013). However, for some women who cannot access psychosocial or psychological interventions or who have a severe depression, antidepressant drugs may be an important alternative form of treatment.

Antidepressants are drugs that treat the symptoms of depression. They are commonly used as the first treatment option for adults with moderate to severe depression, and can be classified into the following types:

- **selective serotonin re-uptake inhibitors** (SSRIs, e.g. fluoxetine) selectively block the re-uptake of serotonin. They are less dangerous in terms of overdose than most tricyclic antidepressants (TCA);
- **TCAs** (e.g. amitriptyline) are antimuscarinic drugs that block the re-uptake of both serotonin and noradrenaline (norepinephrine) and have variable sedating properties;
- **heterocyclic antidepressants** (e.g. mianserin), which block the re-uptake of noradrenaline and serotonin (5-HT);
- **monoamine oxidase inhibitors** (MAOIs, e.g. phenelzine): most drugs from this class are not commonly used due to the dangerous reactions these drugs have with various food groups and other drugs. They act by causing an accumulation of amine neurotransmitters;
- **noradrenaline re-uptake inhibitors** (NARIs, e.g. reboxetine);
- **noradrenaline-dopamine re-uptake inhibitors** (NDRIs, e.g. amineptine, bupropion);
- **serotonin-noradrenaline re-uptake inhibitors** (SNRIs, e.g. duloxetine, milnacipram, venlafaxine);
- **noradrenergic and specific serotonergic antidepressants** (NASSAs, e.g. mirtazapine);
- **serotonin antagonist and re-uptake inhibitors** (SARIs, e.g. trazodone);
- **other unclassified antidepressants** (e.g. agomelatine, vilazodone).

Antidepressants - and often their metabolites (especially if pharmacologically active) - are lipid soluble and are excreted in breast milk. These drugs are metabolised mainly in the liver and excreted via the kidneys. Exposure to antidepressants in breastfed infants is considerably lower (five- to 10-fold) than exposure in utero (Berle 2011), but immaturity or impairment of liver or kidneys (e.g. in preterm babies) may lead to higher concentrations. Breastfeeding women are advised to avoid doxepin (a TCA) as its main metabolite has been found in higher concentrations (Eberhard-Gran 2006). Some case reports and case series have described non-specific adverse events in infants exposed to other antidepressants through breastfeeding, most commonly following exposure to fluoxetine (e.g. poor feeding) and citalopram (e.g. poor sleep) (Berle 2011). There is no evidence of longer-term adverse outcomes among infants exposed to antidepressants (Berle 2011), but this could reflect a lack of studies.



Due to the limitations of the existing evidence, most manufacturers' data sheets carry warnings that antidepressants should not be given to nursing mothers. Physicians often advise women not to breastfeed when taking an antidepressant or may prescribe reduced and potentially ineffective doses or delay pharmacotherapy until after breastfeeding. However, most researchers agree that if a mother was successfully treated for depression during her pregnancy, the same medication should usually be used in the postpartum period while breastfeeding as discontinuing or switching an antidepressant treatment could lead to relapse.

### How the intervention might work

There is substantial evidence showing the effectiveness of antidepressants for depression, particularly as severity of depression increases (Fournier 2010); however, the exact mechanism by which antidepressants have their effect is unclear. One systematic review and meta-analysis of pharmacological neuroimaging studies found that, for both patients and healthy controls, repeated antidepressant administration affected activity in areas of the medial prefrontal cortex and limbic system that are associated with emotion processing (e.g. the anterior cingulate, amygdala and thalamus), with increased activity in response to positive emotions and decreased activity in response to negative emotions (Ma 2014). It appears that most antidepressants inhibit uptake of monoamine neurotransmitters (e.g. serotonin or noradrenaline (norepinephrine)) into neurons thereby increasing the concentrations of these neurotransmitters at synapses (Berton 2006). However, there is some debate over the therapeutic mechanism due to the delay before an antidepressant effect occurs (Pringle 2011).

### Why it is important to do this review

Postnatal depression is a common problem that can have adverse short- and long-term effects on the mother, her child and the wider family. Antidepressants are commonly used as the first treatment option for adults with moderate to severe depression (NICE 2007), but there are few systematic data on the effectiveness of antidepressant drugs in the postnatal period and it is important to establish the effectiveness of antidepressant drugs in comparison with other forms of treatment for postnatal depression or placebo. In addition, although antidepressants are lipid soluble and are excreted in breast milk, the safety of breastfeeding while taking these medications has not been sufficiently reviewed. There is some evidence to suggest that the benefits of breastfeeding may outweigh potential risks for healthy infants born at term (Berle 2011).

Although beyond the scope of this review, antidepressants may also be used for the treatment of pre-existing and antenatal depression during pregnancy. One forthcoming Cochrane review will complement this review by examining the effectiveness of antidepressant use, compared with placebo or psychological therapy, for the treatment of pre-existing and antenatal depression (Gordon 2013).

## OBJECTIVES

To assess the effectiveness of antidepressant drugs in comparison with any other treatment (psychological, psychosocial or

pharmacological), placebo or treatment as usual for postnatal depression.

## METHODS

### Criteria for considering studies for this review

#### Types of studies

We included all published and unpublished randomised controlled trials (RCTs) and cluster RCTs comparing antidepressant drugs with any other treatment, placebo or treatment as usual for postnatal depression. We included trials employing a cross-over design. We excluded all other study designs, including quasi-randomised studies and non-randomised studies.

#### Types of participants

##### Participant characteristics

Women of any age with postnatal depression (onset up to six months after giving birth) who were enrolled into a trial and who were not taking any antidepressant medication at the start of the trial.

##### Diagnosis

We used a broad definition of postnatal depression to include all women who were depressed during the first six months' postpartum regardless of time of onset. Thus, women were included who met criteria for depression by any of the following: use of a validated screening measure, for example, the Edinburgh Postnatal Depression Scale (EPDS) (Cox 1987), use of standard observer-rated depression diagnostic instrument, by a recognised diagnostic scheme (e.g. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) (APA 1999) or the ICD-10 (WHO 1992), or by other standardised criteria, for example, the Research Diagnostic Criteria (RDC) (Spitzer 1978). The threshold scores used for the respective scales were those used by the investigators in the trials.

##### Co-morbidities

Studies involving participants with co-morbid physical conditions or other psychological disorders (e.g. anxiety) were eligible for inclusion as long as the co-morbidity was not the focus of the study.

##### Setting

We assigned no restrictions to the type of study setting.

#### Types of interventions

##### Experimental intervention

Antidepressant medication alone or in combination with another antidepressant or treatment, initiated in at least one arm of a trial.

Antidepressants were organised into classes for the purposes of this review, for example:

- **SSRIs:** citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, sertraline;
- **TCAs:** amitriptyline, clomipramine, desipramine, dothiepin, doxepin, imipramine, lofepramine, nortriptyline, protriptyline, trimipramine;
- **heterocyclic antidepressants:** mianserin;

- **MAOIs:** *irreversible*: izocarboxazid, phenelzine, tranylcipromine; *reversible*: brofaramine, moclobemide, tyrima;
- **NARIs:** reboxetine;
- **NDRIs:** amineptine, bupropion;
- **SNRIs:** duloxetine, milnacipram, venlafaxine;
- **NASSAs:** mirtazapine;
- **SARIs:** trazodone;
- **other unclassified antidepressants:** agomelatine, vilazodone.

### Comparator intervention

Any other treatment, placebo or treatment as usual. We included other treatments such as psychological interventions (e.g. CBT or interpersonal therapy), psychosocial interventions (e.g. peer support or non-directive counselling) or other pharmacological interventions (e.g. another antidepressant).

### Types of outcome measures

We included studies that met the above inclusion criteria regardless of whether they reported on the following outcomes.

#### Primary outcomes

1. Response or remission of depression, using defined dichotomous response, remission or improvement as reported in the individual studies.
2. Adverse events (or side effects) experienced by:
  - a. mother (e.g. headaches, diarrhoea, nausea);
  - b. nursing baby (e.g. respiratory depression, poor sleep, poor feeding).

We extracted all adverse events and data from side effect scales (e.g. Asberg Side Effects Rating Scale) recorded in the trial reports and summarised them narratively. We also reported overall proportions of participants experiencing adverse effects by trial arm where possible.

#### Secondary outcomes

1. Severity of depression based on rating scales (continuous data; either self reported, such as the EPDS, or clinician rated, such as the Inventory of Depression Severity (Clinician Rated Version)).
2. Acceptability of treatment both as assessed directly by questioning trial participants and indirectly by the dropout rates.
3. Cognitive development of the infant/child (e.g. assessment of the mental and psychomotor development of infants using the Mental Development Index (MDI) and Psychomotor Development Index (PDI) of the Bayley Scales of Infant Development (Bayley 2006); parent reports of developmental assessment of children aged two to three years using the Parent Report of Children's Abilities-Revised (PARCA-R) (Johnson 2008); measure of intellectual ability among children aged six years and above using the Wechsler Intelligence Scale for Children (Wechsler 1974)).
4. Overall maternal satisfaction (e.g. self report general satisfaction, satisfaction with self/baby/partner using the Mackay Childbirth Satisfaction Rating Scale (Goodman 2004); self report beliefs, values and perceived skills regarding motherhood using the Parenting Sense of Competence Scale (Gidaud-Wallston 1978)).

5. Maternal relationship with the baby (e.g. improved mother-infant interactions measured using the CARE-Index (Crittenden 1988)).
6. Ability of the mother to carry out daily activities and in her social functioning (e.g. improved score on the Global Assessment of Functioning Scale (Endicott 1976); increased social network, measured using the Social Network Index (Cohen 1997)).
7. The establishment or continuation of breastfeeding (e.g. rates of establishment, continuation or discontinuation).
8. Neglect or abuse of the baby (e.g. using the Parent-Report Multidimensional Neglectful Behavior Scale (Kaufman Kantor 2004)).
9. The effect on marital and family relationships (e.g. using the Quality of Marriage Index (Norton 1983)).
10. Quality of life (e.g. using the 36-item Short Form (SF-36) (Ware 1992)).

#### Timing of outcome assessment

- Zero to eight weeks - immediate effects.
- Nine to 16 weeks - short-term effects.
- 17 to 24 weeks - intermediate effects.
- More than 24 weeks - long-term effects.

#### Search methods for identification of studies

We identified all studies that might describe RCTs of antidepressants for postnatal depression from the Depression, Anxiety and Neurosis Cochrane Review Group Trials Registers (CCDANCTR) (most recent search, 11th July 2014).

#### The Cochrane Depression, Anxiety and Neurosis Review Group's Specialised Register (CCDANCTR)

The Cochrane Depression, Anxiety and Neurosis Group (CCDAN) maintain two clinical trials registers at their editorial base in Bristol, UK: a references register and a studies-based register. The CCDANCTR-References Register contains over 35,000 reports of RCTs in depression, anxiety and neurosis. Approximately 60% of these references have been tagged to individual, coded trials. The coded trials are held in the CCDANCTR-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, using a controlled vocabulary, please contact the CCDAN Trials Search Coordinator for further details. Reports of trials for inclusion in the Group's registers are collated from routine (weekly), generic searches of MEDLINE (1950-), EMBASE (1974-) and PsycINFO (1967-); 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 registers c/o the World Health Organization's trials portal (the International Clinical Trials Registry Platform (ICTRP)), pharmaceutical companies, the handsearching of key journals, conference proceedings and other (non-Cochrane) systematic reviews and meta-analyses.

Details of [CCDAN's generic search strategies](#) (used to identify RCTs) can be found on the Group's website.

## Electronic searches

1. **The CCDANCTR (Studies and Reference Registers)** was searched (to 11th July 2014) using the following terms, on the new Cochrane Register of Studies (CRS) platform:

#1 (antidepress\* or anti-depress\* or "anti depress\*" or MAOI\* or RIMA\* or "monoamine oxidase inhibit\*" or ((serotonin or norepinephrine or noradrenaline or neurotransmitter\* or dopamin\*) NEAR (uptake or reuptake or re-uptake or "re uptake")) or SSRI\* or SNRI\* or NARI\* or SARI\* or NDRI\* or TCA\* or tricyclic\* or tetracyclic\* or pharmacotherap\* or psychotropic\* or "drug therapy")

#2 (agomelatine or alaproclate or amoxapine or amineptine or amitriptylin\* or amitriptylinoxide or atomoxetine or bexlofaxone or benactyzine or binospirone or brofaromine or (bupropion or amfebutamone) or butriptyline or caroxazone or cianopramine or cilobamine or cimoxatone or citalopram or (chlorimipramin\* or clomipramin\* or chlomipramin\* or clomipramine) or clorgyline or clovoxamine or (CX157 or tyrima) or demexiptiline or deprenyl or (desipramine\* or pefrofrane) or desvenlafaxine or dibenzepin or diclofensine or dimetacrin\* or dosulepin or dothiepin or doxepin or duloxetine or desvenlafaxine or DVS-233 or escitalopram or etoperidone or femoxetine or fluotracen or fluoxetine or fluvoxamine or (hyperforin or hypericum or "st john\*") or imipramin\* or iprindole or iproniazid\* or ipsapirone or isocarboxazid\* or levomilnacipran or lofepramine\* or ("Lu AA21004" or vortioxetine) or "Lu AA24530" or (LY2216684 or edivoxetine) or maprotiline or melitracen or metapramine or mianserin or milnacipran or minaprine or mirtazapine or moclobemide or nefazodone or nialamide or nitroxazepine or nomifensine or norfenfluramine or nortriptylin\* or noxiptilin\* or opipramol or oxaflozane or paroxetine or phenelzine or pheniprazine or pipofezine or pirlindole or pivagabine or pizotyline or propizepine or protriptylin\* or quinupramine or reboxetine or rolipram or scopolamine or selegiline or sertraline or setiptiline or teciptiline or thozalinone or tianeptin\* or toloxatone or tranlycypromin\* or trazodone or trimipramine or venlafaxine or viloxazine or vilazodone or viqualine or zalospirone)

#3 (#1 or #2)

#4 (postpartum or post-partum or "post partum" or postnatal\* or post-natal\* or "post natal\*" or perinatal\* or peri-natal\* or "peri natal\*" or puerp\* or intrapartum or intra-partum or "intra partum" or antepartum or ante-partum or "ante partum")

#5 (pregnan\* or maternity or birth or prenatal\* or pre-natal\* or "pre natal\*" or antenatal\* or ante-natal\* or "ante natal\*") and depress\*

#6 (#4 or #5)

#7 (#3 and #6)

Records were screened by the Trials Search Co-ordinator(TSC) to remove irrelevant records (eg trials for major depression where pregnancy was an exclusion criteria).

No restriction on date, language or publication status was applied to the search. Where potentially relevant papers were identified that did not have English language full-text versions, translations were requested from contacts of the review authors or the editorial team of the Cochrane Depression, Anxiety and Neurosis group.

2. **The Cochrane Pregnancy and Childbirth Group's Specilaized Register** was also searched (25th Oct 2013) using terms for antidepressants (as listed above).

No additional studies were identified by this search, so updates were restricted to the CCDANCTR.

## 3. International Trials Registries

ClinicalTrials.gov and the WHO trials portal were searched on 11th July 2014 to identify ongoing and/or unpublished studies.

## Searching other resources

### Reference lists

Forward and backward citation tracking of all included studies was carried out to identify additional studies missed from the original electronic searches (for example unpublished or in-press citations).

### Personal communication

The following Pharmaceutical companies were contacted directly for any relevant unpublished data: Pfizer, Roche, Astrazeneca, Abbott, Lilly, Bayer, GSK, Sanofi, Rosemont pharma, Johnson & Johnson, Merck, Novartis, Teva, Alliance, Amdipharm, Dallas Burston Ashbourne, Lundbeck, Abbvie, Alcon, Britannia Pharmaceuticals Lts, Cox Pharma, Crawford Pharmaceuticals, De Novo Pharmaceuticals, ECRON, Valeant, Viastris, BHR Pharma, Actavis, Forest Pharmaceuticals, Mitsubishi Pharmaceuticals, Ranbaxy, Bristol Myers-Squibb (responses received from: Lilly, Sanofi, Johnson & Johnson, Merck, Teva, Lundbeck, Mylan, Actavis and Bristol Myers-Squibb).

Contact was made with authors of identified trials and with experts in the field (Professor Lee Cohen, Dr Kimberly A. Yonkers, Professor Philip Boyce, Professor Katherine Wisner, Professor Ian Jones, Professor Salvatore Gentile).

The International Marcé Society was also contacted.

## Data collection and analysis

### Selection of studies

Two of three review authors (KT, HM or AK) independently inspected abstracts retrieved from the search. We obtained the full-text articles for any publication that was potentially relevant. Two authors independently assessed the full articles for inclusion based on the previously defined inclusion criteria. We resolved any disagreements by consensus discussions with an additional review author (EM). If it was impossible to resolve disagreements, we contacted the authors of the papers for clarification.

The review authors excluded duplicate records and recorded reasons for exclusion of ineligible studies (see [Characteristics of excluded studies](#) table). We collated multiple reports that related to the same study so that each study rather than each report was the unit of interest in the review. We recorded the selection processes in sufficient detail to complete a PRISMA flow diagram and [Characteristics of included studies](#) table.

We processed included trial data as described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011), and the guidelines issued by the National Health Service (NHS) Centre for Reviews and Dissemination ([Centre for Research and Dissemination 2009](#)).

## Data extraction and management

The review authors designed and piloted a data extraction form, based on the following study characteristics.

1. Methods: date of study, study design, study setting, details of blinding/allocation concealment, total duration of study, details of any 'run-in' period, number of study centres and location, and withdrawals.
2. Participants: total number and number of each group, inclusion and exclusion criteria, mean age, age range, severity of condition and diagnostic criteria.
3. Interventions: number of intervention groups, type of interventions and comparisons, duration of intervention and key details (e.g. dosage, adherence, quality of delivery), concomitant medications and excluded medications.
4. Outcomes: details of measures used to assess outcomes (e.g. details of validation), primary and secondary outcomes specified and collected, time points reported and adverse events.
5. Analysis: statistical techniques used, unit of analysis for each outcome, subgroup analyses, number of participants followed up from each condition.
6. Notes: publication type, funding for trial and notable conflicts of interest of trial authors.

Two review authors (HM and AK) independently extracted data from included studies into standard paper or electronic forms. We checked all data for consistency and resolved any disagreements by going back to the original papers, and by discussion with a third review author (EM or KT) where necessary. If necessary, we contacted authors of the studies for clarification or when inadequate details of randomisation and other characteristics of trials were provided.

One review author (KT) transferred data into Review Manager 5 ([RevMan 2012](#)), which was then double-checked by comparing the data presented in the systematic review with the study reports. A second review author (EM) also spot-checked study characteristics for accuracy against the trial report.

### Main comparisons

The main planned comparisons were as follows:

1. Antidepressants versus placebo;
2. Antidepressants versus treatment as usual;
3. Antidepressants versus psychological intervention;
4. Antidepressants versus psychosocial intervention;
5. Antidepressants versus other pharmacological intervention.

We had planned to include antidepressants versus psychological intervention, but no included studies provided data for this. We also planned to present comparisons on a drug level; however, due to the amount of data available, comparisons were combined by class of drug (see [Types of interventions](#)).

### Assessment of risk of bias in included studies

Three review authors (HM, AK and EM) independently assessed risk of bias for each study using the criteria outlined in the *Cochrane Handbook for Systematic Reviews of Interventions* ([Higgins 2011](#)). The three review authors resolved any disagreements by discussion

or by involving another review author (KT). We assessed risk of bias according to the following domains.

1. Random sequence generation.
2. Allocation concealment.
3. Blinding of participants and personnel.
4. Blinding of outcome assessment.
5. Incomplete outcome data.
6. Selective outcome reporting.
7. Other bias (adherence to medication).

We judged each potential source of bias as high, low or unclear and included a supporting quotation from the study report together with a justification for the review authors judgement in the 'Risk of bias' table. We summarised the risk of bias judgements across different studies for each of the domains listed. Where information on risk of bias related to unpublished data or correspondence with a trialist, we noted this in the 'Risk of bias' table.

### Measures of treatment effect

We presented the primary outcome of depression response or remission using risk ratios (RR) for all studies. We summarised other outcomes using the data as quoted in the original papers (e.g. odds ratio (OR), RR, mean difference (MD)). If there were sufficient data for meta-analyses to be performed on any outcomes, we calculated RRs for dichotomous outcomes and MDs or standardised mean difference (SMD) for continuous data.

### Dichotomous data

We calculated the RR and its 95% confidence interval (CI) for primary outcome dichotomous data. It has been shown that RR is more intuitive than ORs and that OR tend to be interpreted as RR by clinicians ([Bland 2000](#)). This misinterpretation then leads to an overestimate of the impression of the effect.

Where possible, we attempted to convert outcome measures to dichotomous data using cut-off points on rating scales to identify those who did and did not fulfil the criteria for depression.

### Continuous data

If a meta-analysis was conducted for continuous data, we would analyse this by calculating the MD between groups, if studies use the same outcome measure for comparison. If studies used different outcome measures to assess the same outcome, we would calculate SMD and 95% CIs.

When standard errors instead of standard deviations (SD) were presented, we converted the former to SDs. If SDs were not reported and could not be calculated from available data, we asked authors to supply the data. In the absence of data from authors, we used the mean SD from other studies.

### Unit of analysis issues

#### Cluster-randomised trials

It is important to ensure that the data analysed from cluster RCTs takes into account the clustered nature of the data. No cluster-RCTs met the inclusion criteria for this review, but if any are included in future updates we will deal with them as follows. We will extract the intra-cluster correlation coefficient (ICC) for each trial; where no such data are reported, we will request the information from

study authors. If this information is not available, in line with the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011), we will use estimates from similar studies in order to 'correct' data for clustering, where this had not been done. We will use generic inverse variance methods to meta-analyse results from cluster RCTs (Higgins 2011).

### Cross-over trials

A major concern of cross-over trials is the carry-over effect. It occurs if an effect (e.g. pharmacological or psychological) of the treatment in the first phase is carried over to the second phase. As a consequence, on entry to the second phase the participants can differ systematically from their initial state despite a wash-out phase. For the same reason, cross-over trials are not appropriate if the condition of interest is unstable (Elbourne 2002). Both of these effects are very likely in postnatal depression; although we identified no cross-over trials in the review, if any are identified for inclusion in future updates we will only use data from the first randomised treatment period.

### Studies with multiple treatment groups

Trials that have more than two arms (e.g. pharmacological intervention (A); psychological intervention (B); and control (C)) can cause issues with regards to pair-wise meta-analysis. In line with the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011), if we identified any studies with two or more active treatment arms, then we took the following approach, dependent on whether the outcome was dichotomous or continuous:

For a dichotomous outcome: we combined active treatment groups into a single arm for comparison against the control group (in relation to the number of people with events and sample sizes), or the control group was split equally.

For a continuous outcome: we pooled means, 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 to be compared against the control group.

### Dealing with missing data

At some degree of loss of follow-up, data must lose credibility (Xia 2009); therefore, in the protocol, we determined that we would exclude studies with more than 50% loss to follow-up.

In the case where attrition for a binary outcome was between 0% and 50% and outcomes for these people were presented, we reported the data. We presented data on a 'once-randomised always-analyse' basis, assuming an intention-to-treat (ITT) analysis. We assumed that women lost to follow-up had a negative outcome, with the exception of the outcome of death. For example, for the outcome of remission of depression, we assumed that this had not occurred for any of the women lost to follow-up.

In the case where attrition for a continuous outcome was between 0% and 50% and completer-only data were reported, we reproduced these.

We used ITT analysis when available. It was anticipated that some studies would have used the method of last observation carried forward (LOCF) to do an ITT analysis. As with all methods of imputation to deal with missing data, LOCF introduces uncertainty about the reliability of the results. Therefore, where we have

reported LOCF data in this review it is indicated. We presented ITT analysis for all primary outcomes. Where ITT analyses were not available for secondary outcomes, we reported this in the relevant section of the results.

### Assessment of heterogeneity

If there were sufficient data for a meta-analysis, we assessed statistical heterogeneity visually by studying the degree of overlap of the CIs for individual studies in a forest plot. We also carried out more formal assessments using a Chi<sup>2</sup> test with the P value set at 0.1 and the I<sup>2</sup> statistic, as the Chi<sup>2</sup> test has low power to detect diversity when the number of studies is low or sample size is small. The I<sup>2</sup> statistic only provides an approximate estimate of the variability due to heterogeneity so the following overlapping bands would be used to guide our interpretation of the I<sup>2</sup> statistic, as suggested in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011):

- 0% to 40% might not be important;
- 30% to 60% may represent moderate heterogeneity;
- 50% to 90% may represent substantial heterogeneity;
- 75% to 100% represents considerable heterogeneity;

We interpreted the I<sup>2</sup> value using the results of the Chi<sup>2</sup> test as well as the magnitude of the pooled effect size.

### Assessment of reporting biases

Had there been more than 10 included studies, we would have generated a funnel plot and visually inspected it for asymmetry. Asymmetry in the plot could be attributed to publication bias; however, there are other causes of funnel plot asymmetry that we would have also considered.

### Data synthesis

We planned a random-effects meta-analysis to synthesise data from studies with comparable methods (using the same class of antidepressants and the same comparison group, e.g. placebo, listening visits) if three or more studies were identified for each comparison.

### Subgroup analysis and investigation of heterogeneity

We planned subgroup analyses to assess the effectiveness of the intervention in the following groups:

1. women with mild to moderate depressive disorder (as defined by diagnostic interview or a validated scale) versus women with severe depressive disorder (as defined by diagnostic interview or a validated scale);
2. women with chronic depression (onset pre-pregnancy) versus women with onset in pregnancy versus new-onset postpartum depression;
3. interventions lasting eight weeks or less versus interventions lasting more than eight weeks.

### Sensitivity analysis

We planned a priori sensitivity analyses (if sufficient data were identified) to explore the robustness of pooled estimates to decisions made in the systematic review. The effect of excluding studies with the following characteristics was assessed:

1. study quality: excluding studies that had a high risk of bias in any domain;
2. blinding: excluding antidepressant versus placebo trial studies that were unblinded;
3. attrition: excluding studies with more than 20% drop-out. Based on the change to the protocol (see below and [Differences between protocol and review](#)), we also planned a second sensitivity analysis for attrition excluding studies with greater than 50% attrition;
4. validation: excluding outcomes based on non-validated scales from the analyses.

For outcomes with both skewed data and non-skewed data, we investigated the effect of combining all data together and if there was no substantive difference then we left the potentially skewed data in the analyses.

### Summary tables

We produced summary tables for the key findings of the review for all main comparisons. The tables present the findings for remission and response with outcomes for individual trials and pooled estimates, where calculated.

## RESULTS

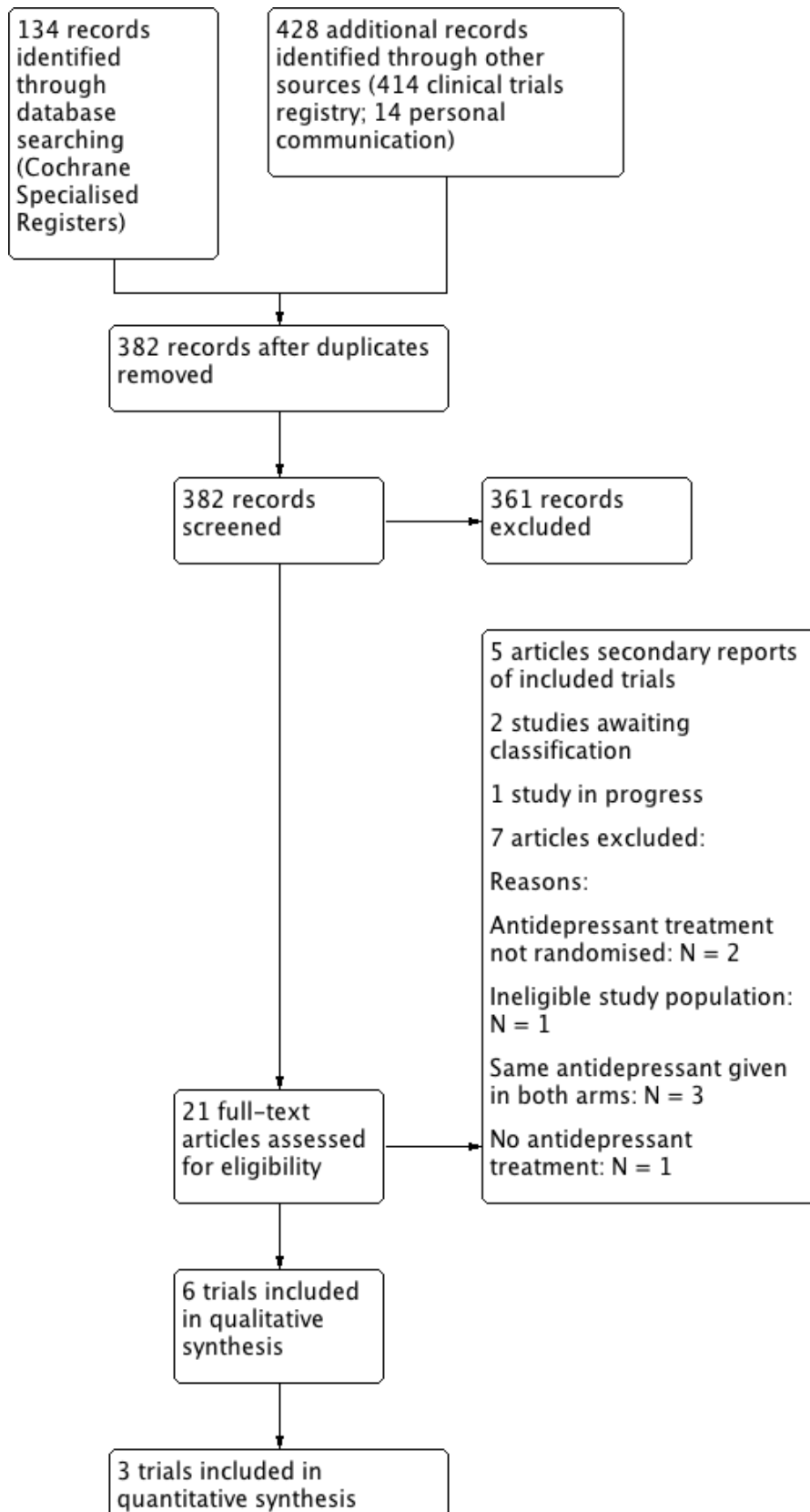
### Description of studies

#### Results of the search

We conducted searches to July 2014 retrieving 134 references from the specialised registers of the two Cochrane Review Groups (CCDAN and PCG). We retrieved an additional 428 records from other sources, including nine studies suggested by pharmaceutical companies. After de-duplication, two review authors (HM, AK or KT) independently screened 382 records and excluded 361 records (on title and abstract) as they did not meet the inclusion criteria. We retrieved the full-text papers for the remaining 21 reports and assessed them for eligibility. After discussion, the review authors decided that the protocol should be altered to allow studies with more than 50% attrition rate to be included in the review due to the small number of relevant RCTs. Therefore, we included [Yonkers 2008](#) (56% dropout).

We required further information to determine the eligibility of one study ([Wisner 2006](#)); the trial investigator provided this (Katherine Wisner). Information on antidepressant prescriptions in [Sharp 2010](#) was also provided following contact with the study author (Debbie Sharp). We translated the two Chinese papers, but neither was eligible for the review. Forward and backward citation tracking of included articles yielded no further relevant trials. The PRISMA flow diagram details the study selection process (see [Figure 1](#)).

**Figure 1. Study selection flow diagram.**



**Figure 1. (Continued)**

3 trials included in quantitative synthesis (meta-analysis for SSRIs versus placebo comparison)

Six trials met the inclusion criteria and were included in the review. We also identified one potentially eligible on-going study (NCT00602355) and two studies awaiting classification with no data available at the time of production of this review (NCT00744328; NCT02122393). More detail on these studies is given in the [Characteristics of ongoing studies](#) table and [Characteristics of studies awaiting classification](#) tables.

**Included studies**

We included six trials in this updated review (see [Characteristics of included studies](#) table).

**Design**

All included studies used a randomised controlled parallel groups design. We identified no eligible cluster-randomised or cross-over trials.

**Sample sizes**

The study by [Sharp 2010](#) had the largest study population with 254 participants randomised. A total of 109 women participated in [Wisner 2006](#), 87 in [Appleby 1997](#), 70 in [Yonkers 2008](#), and 36 in [Hantsoo 2013](#). We included 40 participants from the [Bloch 2012](#) study (42 were randomised but two withdrew immediately following randomisation and were not included in the ITT analysis). The total number of participants included in the review was 596.

**Setting**

Two studies took place in the UK ([Appleby 1997](#); [Sharp 2010](#)), three in the US ([Hantsoo 2013](#); [Wisner 2006](#); [Yonkers 2008](#)), and one in Israel ([Bloch 2012](#)).

**Participants**

The studies had broadly similar inclusion and exclusion criteria.

All required women to meet criteria for depression in the postpartum period, although different criteria were used between studies. [Bloch 2012](#) and [Hantsoo 2013](#) both assessed depression using the Structured Clinical Interview for DSM-IV; [Hantsoo 2013](#) also requiring participants to score 18 or greater on the Hamilton Rating Scale for Depression (HAM-D) at study entry and to have symptoms rated as at least 'moderate' of the Global Clinical Impressions (CGI) severity of illness scale. [Wisner 2006](#) and [Yonkers 2008](#) both required participants to meet the DSM-IV criteria for major depressive disorder and score above a cut-off on the HAM-D (18 or greater for [Wisner 2006](#) and 16 or greater for [Yonkers 2008](#)). [Sharp 2010](#) assessed depression using the Revised Clinical Interview Schedule (CIS-R) for ICD-10 and the EPDS (participants had to score 13 or greater at entry to the study). [Appleby 1997](#) required women to score 10 or greater on the EPDS and 12 or greater on the CIS-R, as well as satisfying researcher diagnostic criteria for major or minor depression (see [Characteristics of included studies](#) for more details on these measures).

Enrolment times varied between six to eight weeks' postpartum ([Appleby 1997](#)), and within 12 months' postpartum ([Hantsoo 2013](#)). In all studies, onset of the depressive episode had to be before six months' postpartum (ranging from four weeks' postpartum ([Wisner 2006](#)), to 26 weeks' postpartum ([Sharp 2010](#))). In all included trials, participants were not taking any antidepressant medication at the commencement of the study. In three trials, participants were also not eligible for the study if they were receiving psychological therapy ([Sharp 2010](#); [Wisner 2006](#); [Yonkers 2008](#)).

All studies restricted the population of depressed women with further exclusion criteria. In order to restrict the participants to women with moderate depression, women were excluded from the [Hantsoo 2013](#) study if they scored 32 or greater on the HAM-D and from [Bloch 2012](#) if they scored 30 or greater on the Montgomery-Åsberg Depression Rating Scale (MADRS). Four studies excluded women with suicidal ideation ([Bloch 2012](#); [Hantsoo 2013](#); [Sharp 2010](#); [Yonkers 2008](#)). Four studies also excluded women based on the duration of existing symptoms of depression (over two years: [Appleby 1997](#); over six months: [Bloch 2012](#); onset of major depressive disorder during pregnancy or before: [Hantsoo 2013](#); [Yonkers 2008](#)). [Sharp 2010](#) did not exclude women based on length of depressive episode and [Wisner 2006](#) also included women with chronic depression (defined as an episode of major depression that began before the index pregnancy), but only after additional funding was obtained part way through the trial.

Three studies excluded women with treatment-resistant depression ([Appleby 1997](#); [Bloch 2012](#); [Hantsoo 2013](#)), defined as two failed trials of antidepressants by [Bloch 2012](#), and past failed trial of sertraline by [Hantsoo 2013](#). Five studies excluded women with current alcohol or drug misuse ([Appleby 1997](#); [Bloch 2012](#); [Hantsoo 2013](#); [Sharp 2010](#); [Yonkers 2008](#)), and five studies excluded women with current or past psychotic symptoms or disorders (such as bipolar disorder, schizophrenia or schizoaffective disorder) ([Bloch 2012](#); [Hantsoo 2013](#); [Sharp 2010](#); [Wisner 2006](#); [Yonkers 2008](#)). [Appleby 1997](#) excluded any women with severe illness. Three studies stated that they excluded women with major physical illness ([Appleby 1997](#); [Bloch 2012](#); [Hantsoo 2013](#)), and one study excluded mothers who were breastfeeding ([Appleby 1997](#)).

Where age inclusion criteria were stated, these were largely 18+ or 18 to 45 years ([Bloch 2012](#); [Hantsoo 2013](#); [Sharp 2010](#)); however, [Yonkers 2008](#) included women from 16 years of age and [Wisner 2006](#) included women from 15 years. Where mean age was reported in studies, this ranged from 23.1 years ([Appleby 1997](#), in the placebo plus one session of counselling group) to 30.8 ± 4.0 years ([Hantsoo 2013](#)).

The predominant ethnicity was reported was white, ranging from 48.6% of participants in the study by [Yonkers 2008](#) to 94.4% of participants in the [Hantsoo 2013](#) study. In one study, there was a significant minority of Hispanic participants (35.7%) ([Yonkers 2008](#)), with a small minority of participants in the study by [Hantsoo 2013](#) being Hispanic (5.6%). Two studies had a minority of black



participants (12.9%: [Yonkers 2008](#); 11.5%: [Sharp 2010](#)), and the study by [Sharp 2010](#) also had 13 Asian participants (5.2% of those randomised). In the study by [Wisner 2006](#), 40% of women randomised to sertraline and 19% of women randomised to nortriptyline had non-white ethnicity. Two studies provided no data on ethnicity ([Bloch 2012](#), [Appleby 1997](#)). Information provided on socioeconomic status was highly varied, making any comparisons of socioeconomic status across trials difficult.

All studies assessed severity of depression at baseline. [Wisner 2006](#) reported that baseline severity was assessed using several scales including the HAM-D and CGI with no difference between the two study groups, but did not report scale scores. [Appleby 1997](#) reported geometric means scores on the EPDS and the HAM-D for all women randomised to take fluoxetine (all also receiving either one or six sessions of counselling) and all women randomised to placebo (again all also receiving either one or six sessions of counselling). Geometric mean HAM-D scores were 14.2 (95% CI 13.0 to 15.5) for the fluoxetine group and 13.9 (95% CI 12.5 to 15.4) for the placebo group; on the HAM-D scores in the range 8 to 16 indicate mild depression ([Zimmerman 2013](#)). Geometric mean EPDS scores were 17.2 (95% CI 16.2 to 18.2) for the fluoxetine group and 16.9 (95% CI 15.8 to 18.1) for the placebo group. In [Bloch 2012](#), baseline EPDS scores showed similar means to [Appleby 1997](#) with 16.05 (SD 4.84) in the brief dynamic psychotherapy group plus placebo group and 18.40 (SD 4.83) in the brief dynamic psychotherapy group plus antidepressant group. Similar baseline severity was also found in the [Sharp 2010](#) study based on EPDS scores (mean  $\pm$  SD: 17.3  $\pm$  3.3 for the antidepressant group and 17.7  $\pm$  3.5 for the treatment as usual followed by listening visits group).

Higher baseline severity was found in two placebo-controlled studies (mean HAM-D scores in the range 17 to 23 indicating 'moderate depression' and  $\geq 24$  indicating 'severe depression'; [Zimmerman 2013](#)). [Hantsoo 2013](#) measured baseline severity with both the EPDS and HAM-D; on the EPDS the mean score for women randomised to the antidepressants (sertraline) was 18.8 (SD 2.6) and 20.8 (SD 5.7) for women randomised to placebo. On the HAM-D, these scores were 20.6 (SD 2.8) for the antidepressant group and 23.2 (SD 3.9) for the placebo group. Similar but slightly higher scores were recorded at baseline by [Yonkers 2008](#); in this study women randomised to antidepressants had a mean HAM-D score of 23.6 (SD 4.7) and women randomised to placebo had a mean HAM-D score of 24.7 (SD 5.0). Further details on these measures are given in the [Characteristics of included studies](#) table.

Women were recruited from a variety of settings, including general practice, postnatal wards, obstetric care settings and general advertising. [Sharp 2010](#) sent an information pack containing an EPDS questionnaire to all new mothers within the catchment area (data obtained from birth registry office and general practitioner (GP) records). The length of the recruitment period ranged from 20 months ([Appleby 1997](#)) to 10 years ([Hantsoo 2013](#)) (not described in [Wisner 2006](#)).

### Interventions

Antidepressant prescriptions varied between studies with three prescribing sertraline ([Bloch 2012](#); [Hantsoo 2013](#); [Wisner 2006](#)), one fluoxetine ([Appleby 1997](#)), one paroxetine ([Yonkers 2008](#)), and one nortriptyline ([Wisner 2006](#); used as a comparison with sertraline). [Sharp 2010](#) allowed choice of antidepressants based on physician and participant preference. Although GPs were given prescribing

guidelines in this study (with SSRIs recommended as the first-line therapy in keeping with national guidelines), there were no set drugs for the trial. Information on the antidepressants prescribed was obtained through participant self report at all follow-up points and by recording prescribing information from medical notes. Most participants were prescribed citalopram, fluoxetine or sertraline; full details of the antidepressants prescribed and the number of participants prescribed each antidepressant are given in the [Characteristics of included studies](#) table.

One study ([Hantsoo 2013](#)) had a one-week run-in period to the trial during which all participants took placebo only, followed by participants in the antidepressant group being given sertraline 50 mg per day. In the other two trials prescribing sertraline, both [Bloch 2012](#) and [Wisner 2006](#) had initial doses of 25 mg per day, increasing to 50 mg after two ([Wisner 2006](#)) or seven days ([Bloch 2012](#)). In [Hantsoo 2013](#) and [Wisner 2006](#), the maximum dose allowed was 200 mg per day, in [Bloch 2012](#), it was 100 mg per day. Participants randomised to nortriptyline in the [Wisner 2006](#) study started on 10 mg per day, increasing to 25 mg per day up to a maximum of 150 mg per day. The prescription of paroxetine in the [Yonkers 2008](#) study began with 10 mg per day, increasing to a maximum of 40 mg per day. Initial dosage of antidepressants were described as increasing at regular intervals in all of the trials except [Appleby 1997](#) (where no data on prescribing patterns were given), as guided by tolerability of treatment and effect on symptoms. Where specified, dosage was once daily. Data on dosage was collected in [Sharp 2010](#), but is not reported here owing to the heterogeneity of treatments given. Adherence was monitored with pill counts in two trials ([Bloch 2012](#); [Yonkers 2008](#)), self report plus review of prescription data in one trial ([Sharp 2010](#)), and serum drug level monitoring in one trial ([Wisner 2006](#)).

Four studies had a placebo control with study personnel and participants blinded to group allocation ([Appleby 1997](#); [Bloch 2012](#); [Hantsoo 2013](#); [Yonkers 2008](#)). [Bloch 2012](#) and [Appleby 1997](#) also included psychological therapy in both the placebo and the active treatment arms (brief dynamic psychotherapy ([Bloch 2012](#)) and CBT-based counselling ([Appleby 1997](#))). In [Appleby 1997](#), participants were randomly assigned to receive either one or six sessions of the CBT-based counselling. [Wisner 2006](#) compared the efficacy of two pharmacological treatments (nortriptyline and sertraline) in a two-arm blinded RCT. [Sharp 2010](#) conducted an unblinded pragmatic RCT in which participants were randomised to antidepressants or four weeks of treatment as usual followed by listening visits. Antidepressants were prescribed by the participant's GP, who was requested to provide no other counselling or psychological intervention for women in this arm of the trial. However, the participants receiving antidepressants also received usual care and had several GP appointments for antidepressant monitoring. The comparison group received treatment as usual (general supportive care from GPs) for the first four weeks to allow the effectiveness of antidepressants to be compared with treatment as usual and to replicate the waiting period that would likely occur prior to a woman beginning counselling for postnatal depression. The GPs were requested not to prescribe antidepressants or additional psychological interventions unless clinically necessary. The listening visits (non-directive counselling) began after this four-week period and were delivered by trained health visitors (up to eight sessions). The primary aim of this study was to evaluate the clinical effectiveness of antidepressants for mothers with postnatal

depression compared with treatment as usual (i.e. outcomes at four weeks prior to the commencement of listening visits). The secondary aim of this study was to compare outcomes in the two groups at 18 weeks (i.e. women randomised to antidepressants compared with women randomised to listening visits following treatment as usual). In this trial, women were also able to change to (or add in) the alternative intervention (i.e. antidepressants or listening visits) at any point after four weeks.

Follow-up intervals ranged from seven weeks ([Hantsoo 2013](#)) to 24 weeks ([Wisner 2006](#); main outcomes as eight weeks followed by a 16-week continuation phase). In the study by [Hantsoo 2013](#), follow-up at seven weeks included the one-week run-in placebo period; therefore, outcomes after six weeks of active treatment versus placebo were assessed. One study followed up participants at four and 18 weeks (with the four-week follow-up comparing antidepressants with treatment as usual, and the 18-week follow-up comparing antidepressants with listening visits) ([Sharp 2010](#)). In another study, outcome assessments took place at eight weeks ([Yonkers 2008](#)). [Appleby 1997](#) and [Bloch 2012](#) had a 12-week follow-up period. In the final four weeks of the [Bloch 2012](#) trial (after the main outcomes at eight weeks), the trial was converted to an open trial for the continuation phase.

## Outcomes

### Primary outcome assessment

The primary outcome in this review was a dichotomous measure of depression response or remission, which was assessed in five of the six included studies (data on response and remission were not available from [Appleby 1997](#)). This was defined in the following ways:

- [Bloch 2012](#): response: greater than 50% reduction in MADRS or EPDS score during treatment; remission: final score less than 10 on the MADRS scale or less than 7 on the EPDS (outcomes at eight weeks);
- [Sharp 2010](#): remission (termed 'improvement' in the original trial): less than 13 on the EPDS (outcomes at four and 18 weeks);
- [Wisner 2006](#): response: 50% reduction in HAM-D from baseline; remission: less than 7 on the HAM-D (outcomes at eight weeks);
- [Yonkers 2008](#): response: CGI scale score of 1 or 2; remission: HAM-D score 8 or less (outcomes at eight weeks);
- [Hantsoo 2013](#): response: 10 or less on HAM-D plus at least 50% decrease in HAM-D score from baseline plus CGI (improvement scale) 2 or less; remission: as 'response' plus HAM-D score less than 7 (outcomes after six weeks of treatment (study week seven, including the one-week run-in period)).

Further details on these scales are given in the [Characteristics of included studies](#) tables.

### Adverse effects

In two studies, specific side effect rating scales were used (Asberg Side Effects Rating Scale ([Wisner 2006](#)), and the UKU Side Effect

Rating Scale ([Bloch 2012](#))). Other trials reported adverse outcomes but the method of the assessment was not specified.

### Secondary outcomes

Secondary outcomes were severity of depression, acceptability, cognitive development of the infant, maternal satisfaction, maternal relationship with the baby, social functioning, establishment or continuation of breastfeeding, neglect or abuse of the baby, effect on marital or family relationships and quality of life, although there were no data from the included studies for several of the outcomes.

### Excluded studies

We excluded studies for the following reasons: antidepressant treatment not randomised (two studies: [Rojas 2007](#); [Suri 2005](#)), same antidepressants given in both arms (three studies: [Misri 2004](#); [Yu 2006](#); [Zhao 2006](#)), no antidepressant treatment (one study: [Bennett 2001](#)) and ineligible study population ([Stein 2012](#)). See [Characteristics of excluded studies](#) for further details.

### Ongoing studies

We identified one ongoing RCT comparing sertraline with interpersonal psychotherapy and with placebo ([NCT00602355](#)). Based on the limited available information, we believe that this study will be eligible for inclusion when completed (see [Characteristics of ongoing studies](#)).

### Studies awaiting classification

We identified two RCTs awaiting classification: one comparing sertraline with transdermal oestradiol and with placebo ([NCT00744328](#)) and one comparing sertraline with CBT and with combined therapy (sertraline and CBT) ([NCT02122393](#)). From the available evidence it appears that both studies would be eligible for the review, but no data is currently available for either study (see [Characteristics of studies awaiting classification](#) for more details).

### New studies found at this update

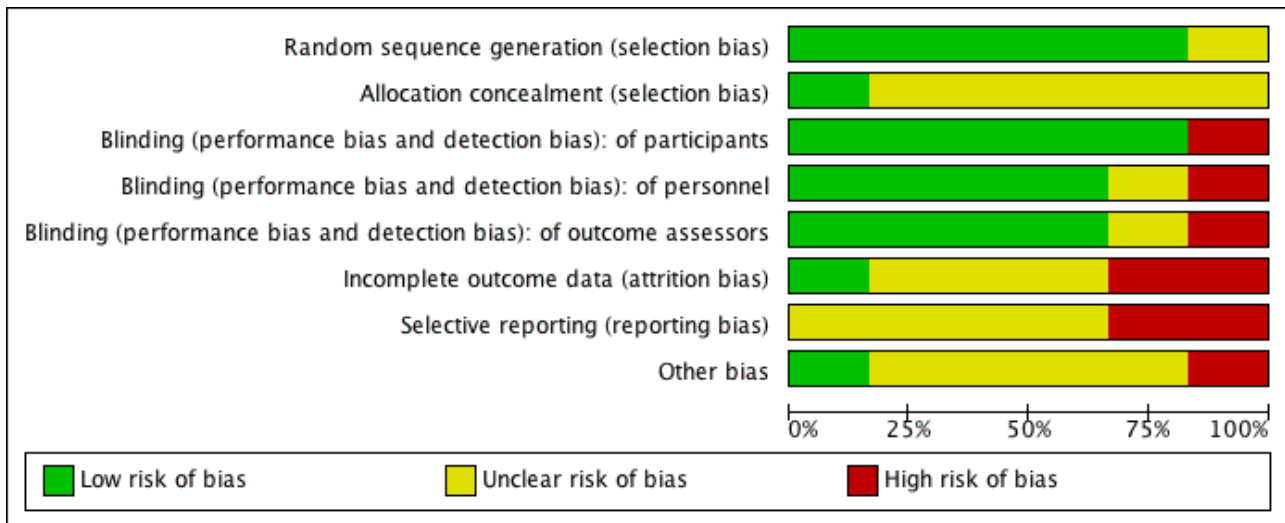
We included five new studies in this update ([Bloch 2012](#); [Hantsoo 2013](#); [Sharp 2010](#); [Wisner 2006](#); [Yonkers 2008](#)).

### Risk of bias in included studies

We used the Cochrane 'Risk of bias' assessment tool to evaluate each study in five domains of potential bias ([Higgins 2011](#)): random sequence generation, allocation concealment, blinding, incomplete outcome data and selective outcome reporting. We also assessed adherence to medication as an additional potential source of bias.

See [Characteristics of included studies](#) table for full details of risk of bias judgements for each study. Graphical representations of the overall risk of bias in included studies are presented in [Figure 2](#) and [Figure 3](#).

**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.**

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding (performance bias and detection bias): of participants	Blinding (performance bias and detection bias): of personnel	Blinding (performance bias and detection bias): of outcome assessors	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Appleby 1997	+	?	+	+	+	?	?	?
Bloch 2012	+	?	+	?	?	+	?	?
Hantsoo 2013	?	?	+	+	+	?	?	?
Sharp 2010	+	+	-	-	-	?	-	-
Wisner 2006	+	?	+	+	+	-	?	+
Yonkers 2008	+	?	+	+	+	-	-	?

## Allocation

### Sequence generation

Five of the six included studies described methods of random sequence generation with low risk of bias (e.g. using computer- or pharmacy-generated random numbers). For one study, the risk of bias in this domain was unclear; the study was described as randomised but there were insufficient details to assess whether appropriate methods of randomisation were used ([Hantsoo 2013](#)).

### Allocation concealment

Only one included study gave sufficient information on allocation concealment to ensure low risk of bias in this domain (a remote computerised randomisation service was used and the methods of sequence generation were concealed from those involved in the enrolment and randomisation of participants) ([Sharp 2010](#)). None of the other studies provided details on allocation concealment.

## Blinding

### Blinding of participants and personnel

Four studies had low risk of bias with study personnel and participants blinded to treatment allocation. One study compared the antidepressant intervention with listening visits so blinding of study personnel and participants was not possible, leading to high risk of bias in this domain ([Sharp 2010](#)). The risk of bias was unclear in [Bloch 2012](#). Participants and the managing psychiatrist were blinded to treatment condition, but when the blind was assessed at the end of the study the psychiatrist guessed group assignment incorrectly in every case. This suggests that there may have been some differences between the intervention groups; incorrect assignment of every participant when there are only two treatment options implies that the psychiatrist correctly grouped the participants with others who had received the same treatment, although incorrectly guessed treatment status of these groups. It should be noted that only [Bloch 2012](#) reported assessment of the success of blinding, although [Hantsoo 2013](#) described the withdrawal of one participant following accidental unblinding in the penultimate week of the study.

### Blinding of outcome assessment

Four studies had low risk of bias with outcome assessors blinded to treatment allocation. [Sharp 2010](#) did not blind outcome assessors, leading to high risk of bias. No details were provided on who performed the outcome assessments in the [Bloch 2012](#) study so the risk of bias for blinding of outcome assessment was unclear.

### Incomplete outcome data

The greatest risk of bias in the studies included in this review came from incomplete outcome data (attrition bias), with only one study having low risk of bias in this domain ([Bloch 2012](#)). [Yonkers 2008](#) had very high risk of attrition bias, with 39 women withdrawing from the study out of the 70 women randomised (56%). This means that the study findings must be interpreted extremely cautiously. However, dropout reasons and numbers were similar between treatment groups and sensitivity analyses assuming that all drop-outs had either positive or negative outcomes in the trial found that antidepressants remained associated with significantly higher remission rates than placebo in both scenarios. This suggests that the primary finding was robust to a range of outcomes for drop-outs.

[Wisner 2006](#) had high risk of bias from incomplete outcome data. Significantly more participants withdrew from the sertraline than the nortriptyline group in the first eight weeks of the study and there was high attrition from both groups (withdrawal rates: 23/55 (42%) with sertraline and 13/54 (24%) with nortriptyline; Wilcoxon P value = 0.02). Reasons for withdrawal were assessed and although "side effects" and "clinical deterioration" did not differ in frequency between the two groups, significantly more women in the sertraline group withdrew by personal choice or were lost to follow-up without reasons given, which may reflect factors associated with clinical outcomes or side effects.

In four studies, there was some evidence of potential risk of bias, but we rated this as 'unclear' owing to insufficient details in reporting. In [Appleby 1997](#), 26 of the 87 women dropped out over the course of the study, although with relatively similar rates across treatment groups. Timing and reasons for drop-out were reported but in many cases this was "no reason given", meaning that it is difficult to assess whether reasons for drop-out varied between groups. In [Hantsoo 2013](#), seven of the 36 women dropped out over the course of the trial. Again, there were similar numbers of drop-outs in the two treatment groups but all women dropping out due to clinical deterioration had received the placebo, which may have led to an underestimation of the intervention effect. [Sharp 2010](#) reported some differential drop-out between the antidepressant group and the treatment as usual followed by listening visit group (higher in the antidepressant group), which was not statistically significant at four weeks (antidepressant group 18% drop-out, 23 women; treatment as usual group 10% drop-out, 13 women; P value = 0.09) but was significant at 18 weeks (antidepressant group 25% drop-out, 32 women; listening visits 13% drop-out, 16 women; P value = 0.015). However, sensitivity analyses examining the impact of attrition (including multiple imputation) found that study findings were robust to a range of outcomes for study drop-outs. [Sharp 2010](#) did not give characteristics of the drop-outs separately by intervention group and reasons for withdrawal were not described so it is not possible to assess whether these differ between groups.

All studies used LOCF in cases of missing data, except for [Sharp 2010](#) who used multiple imputation. All studies conducted ITT analyses for their primary outcomes and all results for the primary outcome of this review are reported using ITT, although some secondary analyses were reported using complete case analysis only (and are indicated as such).

### Selective reporting

Bias from selective outcome reporting was unclear in four of the studies: in most studies, the protocols were unavailable and, in one study, the protocol had insufficient detail on outcomes to assess selective reporting ([Bloch 2012](#)). In [Sharp 2010](#), the protocol was available and all pre-specified primary outcomes were reported, but outcomes from assessments of participants' partners, the HOME measure and Bayley Scale of Infant Development were not reported. Authors stated that these outcomes would be reported in a subsequent paper but we could not find this. In [Yonkers 2008](#), the Social Adjustment Scale and SF-36 were included in the methods but not reported in the results. We believe that the general absence of data on child outcomes and breastfeeding safety in the six studies reflects the fact that these data were not collected, rather than selective outcome reporting; however, this cannot be assessed without access to the study protocols.

## Other potential sources of bias

Adherence to antidepressants is often low, which could bias study findings, so we assessed risk of bias related to low adherence in the included studies. [Wisner 2006](#) assessed serum levels as a measure of compliance and found that 14 women (out of 95 study participants) had minimal levels of the antidepressant in their blood, despite claiming compliance. These women were evenly distributed between the two antidepressant groups and results did not alter when these women were removed from analysis. Therefore, the risk of bias related to adherence in this study was low. [Sharp 2010](#) collected self reported data on adherence to medication and had high risk of bias in this area; only 56% (59 women) of the 106 women who were randomised to antidepressant treatment and followed up reported taking any antidepressants in the first four weeks after randomisation. Of the women followed up at 18 weeks, 64% of participants randomised to antidepressants reported taking antidepressants in the previous four weeks (62/97), and 34% of women randomised to listening visits reported taking antidepressants in the previous four weeks (37/109). Two studies provided no details of adherence to antidepressant medication so risk of bias in these studies was unclear ([Appleby 1997](#); [Hantsoo 2013](#)). One study stated that pill counts were used to monitor compliance but it is unclear from the results how many women were non-compliant ([Bloch 2012](#)). [Yonkers 2008](#) used pill counts to assess adherence and found that seven of the 35 women randomised to antidepressant treatment were non-compliant (took less than 80% of pills) at one visit, four were non-compliant at two visits and one was consistently non-compliant and consequently removed from active treatment in the study. In the placebo group, 10 of the 35 women were non-compliant at one visit, three were non-compliant at two or more visits and one was non-compliant at four visits. It is unclear to what extent this may have biased study findings, as we do not know how much medication the non-compliant women were taking (this could range between 0% and 79% based on reported data) or whether adherence was only reported for women who completed the study.

## Effects of interventions

See: [Summary of findings for the main comparison](#)

All included studies reported effect on depressive symptoms as their primary outcome. Five studies included response or remission rates (the primary outcome for this review). The way these were characterised for each study is listed in the [Included studies](#) section

above and the summary of findings tables. Further details on the scales are given in the [Characteristics of included studies](#) tables.

Due to the small number of studies and heterogeneity between papers it was only possible to conduct a meta-analysis for the comparison of antidepressant vs placebo. The meta-analysis included two studies where women were randomised to SSRIs or placebo, and one where women were randomised to SSRIs and psychological therapy or placebo and psychological therapy. Other findings are discussed narratively for the following comparison groups based on class of drug, where possible: antidepressants vs treatment as usual, antidepressants vs psychosocial interventions; antidepressants vs other pharmacological interventions. No data were available for our planned comparison of antidepressants versus psychological interventions.

## Comparison 1: antidepressants versus placebo

### *Selective serotonin re-uptake inhibitors versus placebo*

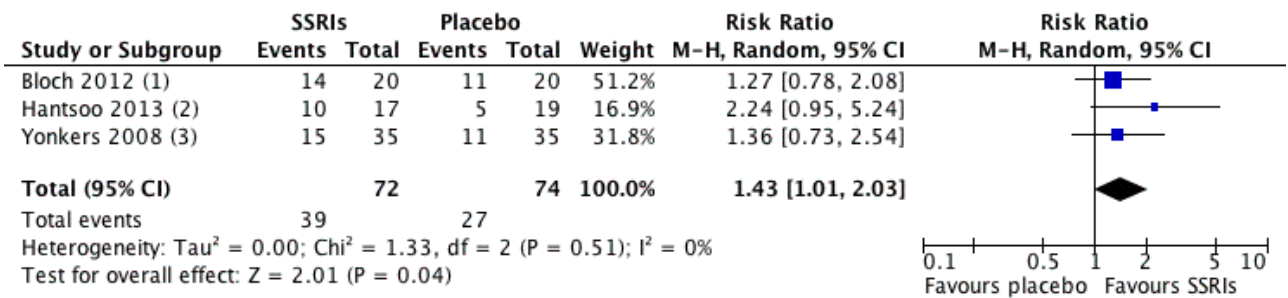
#### *Primary outcomes*

##### **1.1 Remission/response of depression**

Four studies investigated the effect of SSRIs versus placebo ([Appleby 1997](#); [Bloch 2012](#); [Hantsoo 2013](#); [Yonkers 2008](#)). Two studies (106 women) compared SSRIs (sertraline ([Hantsoo 2013](#)); paroxetine ([Yonkers 2008](#))) with placebo, with both studies assessing response and remission. The other two studies (127 women) compared SSRIs (fluoxetine ([Appleby 1997](#)); sertraline ([Bloch 2012](#))) with placebo, with both study arms additionally receiving psychological therapy (CBT-based counselling ([Appleby 1997](#)); brief dynamic psychotherapy ([Bloch 2012](#))). Of these two, only [Bloch 2012](#) assessed response or remission.

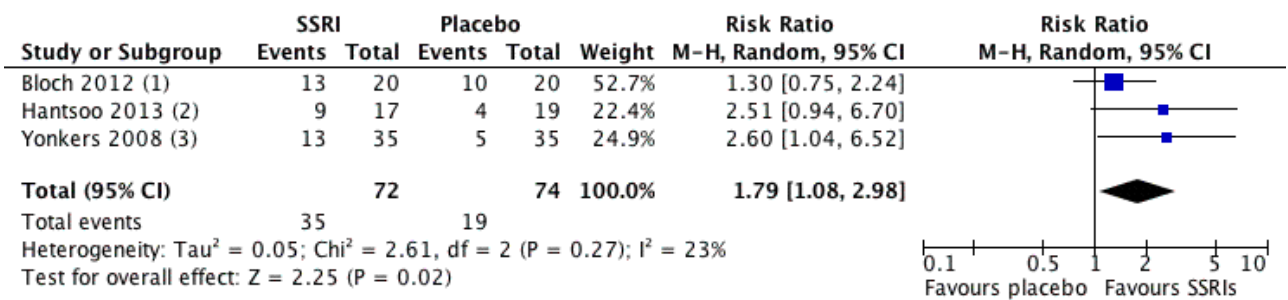
Random effects meta-analyses were conducted to pool data on response and remission from the three studies with this data (146 participants). The pooled estimate showed a 43% greater chance of responding for those randomised to SSRIs compared with those randomised to placebo (RR 1.43, 95% CI 1.01 to 2.03; [Analysis 1.1, Figure 4](#)). There was no evidence of meaningful heterogeneity in this meta-analysis:  $I^2$  was 0%, the  $\chi^2$  test for heterogeneity was not significant and the confidence intervals from individual studies overlapped. The pooled estimate for remission found a 79% greater chance of remission for those randomised to SSRIs compared with those randomised to placebo (RR 1.79, 95% CI 1.08 to 2.98; [Analysis 1.2, Figure 5](#)). Again, heterogeneity was found to be low:  $I^2$  was 23%, the  $\chi^2$  test was not significant and confidence intervals from individual studies overlapped.

**Figure 4. Forest plot of comparison: 1 Selective serotonin re-uptake inhibitors versus placebo, outcome: 1.1 Response rate at post-treatment.**



- (1) Sertraline and psychological therapy (brief dynamic psychotherapy) versus placebo and psychological therapy (brief dynar)
- (2) Sertraline versus placebo
- (3) Paroxetine versus placebo

**Figure 5. Forest plot of comparison: 1 Selective serotonin re-uptake inhibitors versus placebo, outcome: 1.2 Remission rate at post-treatment.**



- (1) Sertraline and psychological therapy (brief dynamic psychotherapy) versus placebo and psychological therapy (brief dynar)
- (2) Sertraline versus placebo
- (3) Paroxetine versus placebo

**1.2 Adverse effects**

All studies comparing SSRI versus placebo examined side effects, with none describing a significant difference between groups (Appleby 1997; Bloch 2012; Hantsoo 2013; Yonkers 2008).

Participants in the Yonkers 2008 trial reported decreased appetite (antidepressant group 3/35 women, 9%; placebo group 2/35 women, 6%), diarrhoea (antidepressant group 4/35, 11%; placebo group 4/35, 11%), dizziness (antidepressant group 6/35, 17%; placebo group 3/35, 9%), dry mouth (antidepressant group 4/35, 11%; placebo group 0/35), headache (antidepressant group 9/35, 26%; placebo group 13/35, 37%), nausea (antidepressant group 5/35, 14%; placebo group 6/35, 17%), somnolence and drowsiness (antidepressant group 5/35, 14%; placebo group 5/35, 14%). Although some side effects appeared more common in the antidepressant group (e.g. dizziness, dry mouth), no significant differences were found in symptoms experienced by participants in the paroxetine as compared with the placebo group (P values ranging from 0.11 to greater than 0.99). The overall proportion of women experiencing side effects was not reported.

In Hantsoo 2013, 3/17 women from the sertraline group reported side effects: nausea (3/17 women; 17.6%), headache (1/17 women; 5.8%) and diarrhoea (1/17 women; 5.8%). Frequent diarrhoea was

reported by one participant in the placebo group (1/19; 5.3%). No participants dropped out due to side effects and no adverse events were reported for any of the participants or their breastfeeding infants (number breastfeeding: 6/17 in the sertraline group, 5/19 in the placebo group).

Bloch 2012 reported a hypomanic switch in two women from the brief dynamic psychotherapy plus sertraline group at week eight (one woman on sertraline 50 mg and one woman on sertraline 100 mg). UKU Side Effect Rating scores showed no significant difference between treatment groups at week eight (P value = 0.456), or at 12 weeks (P value = 0.937), although the overall proportion of women experiencing side effects in each group was not given, neither were the details of types of side effects experienced.

In the Appleby 1997 study, one woman dropped out of the fluoxetine group and three women dropped out of the placebo group due to side effects, but the nature of these side effects was not reported. Side effects were only reported among women who dropped out of the study.

## Secondary outcomes

### 1.3 Severity of depression

In the trial by [Yonkers 2008](#), change in severity of depression was assessed as the difference in mean score on repeated measures of the HAM-D, Inventory of Depressive Symptomatology, Self Report (IDS-SR), and CGI-S (Clinical Global Impressions, Severity of Illness scale) between the antidepressant and placebo groups. There was no significant difference between groups on the HAM-D (-1.62; P value = 0.22). There was a significant main effect of group on the IDS-SR scores (-4.98; P value = 0.019); however, IDS-SR scores were significantly different between the antidepressant and placebo groups at baseline and the authors concluded that it was this baseline difference carried over at later time points as the group by time interaction effect was not significant. The authors concluded that there was significantly greater improvement in the antidepressant than placebo group based on CGI-S scores (main effect -0.48; P value = 0.047, groups did not differ at baseline).

In the trial by [Hantsoo 2013](#), severity of depression was defined as a treatment group by time interaction with the baseline score as a covariate on the HAM-D and EPDS. Analysis of change in scores over time in the ITT group showed that there was not a significant group by time effect for the HAM-D ( $F(1,145) = 2.05$ ; P value = 0.15) or EPDS ( $F(1,137) = 0.43$ ; P value = 0.51).

In the study by [Bloch 2012](#), ITT analyses using LOCF indicated no group by time interaction effect for depression scores on either the MADRS ( $F[4,35] = 0.97$ ; P value = 0.44) or EPDS ( $F[4,35] = 0.62$ ; P value = 0.65) across the eight-week intervention period.

[Appleby 1997](#) assessed reduction in depression using the CIS-R, HAM-D and EPDS, with ITT analysis (LOCF). The geometric mean scores at 12 weeks on the CIS-R were 11.1 (95% CI 6.9 to 17.6) for fluoxetine plus *one* session of counselling and 19.1 (95% CI 15.4 to 23.5) for placebo plus *one* session of counselling. Geometric mean scores at 12 weeks on the CIS-R were 10.5 (95% CI 6.6 to 16.6) for fluoxetine plus *six* sessions and 13.0 (95% CI 9.2 to 18.1) for placebo plus *six* sessions of counselling. The calculated percentage difference in geometric mean scores between fluoxetine and placebo at 12 weeks was 40.7% (95% CI 10.9% to 60.6%). Similar findings were observed using the HAM-D with the following geometric mean scores reported at 12 weeks: 4.4 (95% CI 2.4 to 7.4) for fluoxetine plus *one* session of counselling and 8.1 (95% CI 6.1 to 10.7) for placebo plus *one* session of counselling; 5.1 (95% CI 2.6 to 9.2) for fluoxetine plus *six* sessions of counselling and 4.9 (95% CI 3.0 to 8.9) for placebo plus *six* sessions of counselling. Finally, the geometric mean scores on the EPDS at 12 weeks were reported as follows: 7.1 (95% CI 5.0 to 10.1) for fluoxetine plus *one* session of counselling and 10.3 (95% CI 8.1 to 13.2) for placebo plus *one* session of counselling; 7.5 (95% CI 4.6 to 11.8) for fluoxetine plus *six* sessions of counselling and 9.5 (95% CI 7.2 to 12.5) for placebo plus *six* sessions of counselling. Authors of this study concluded that fluoxetine was significantly more effective than placebo and, after an initial session of counselling, was as effective as a full course of CBT counselling in the treatment of postnatal depression. However, the 95% CIs of geometric mean scores for all treatment groups overlap.

### 1.4 Acceptability

While no direct assessments of treatment acceptability were made, treatment adherence and withdrawal rate may be indicative of

the level of treatment acceptability. In the paroxetine arm of [Yonkers 2008](#), one woman withdrew due to nausea, six due to lack of efficacy, five who felt well and no longer desired treatment, one who became pregnant and one who was not adherent to medication. A further six women were lost to follow-up. Among the women on placebo, four left due to perceived adverse effects, seven discontinued due to lack of efficacy, two felt better and no longer desired treatment and one moved out of the area. A further nine women in this arm were lost to follow-up. In terms of treatment compliance, among the women on paroxetine, seven were taking less than 80% of medication at one visit, four were non-adherent on a second visit and one was excluded from the trial for ongoing non-adherence. Among the women on placebo, 10 were non-compliant at one visit, three were non-compliant at two visits and one was non-compliant on a fourth visit.

In the trial comparing sertraline and placebo, [Hantsoo 2013](#) reported that 36/36 (100%) women remained in the trial at week two, 33/36 (92%) completed through week four and 29/36 (81%) completed through week seven. Three of 19 women on placebo left the trial due to clinical worsening. Among the women randomised to antidepressants, 3/17 (17.6%) dropped out (e.g. due to death in the family or unable to contact) and 1/17 (5.9%) was excluded after the blind was broken due to an administrative error. The authors did not report on adherence to medication in either treatment arm.

[Appleby 1997](#) reported a similar dropout rate in all treatment arms. Fourteen of 43 (32.6%) women on fluoxetine dropped out of the trial. Of these, two reported that they 'disliked the drug', three reported a lack of improvement, one reported side effects (type not specified) and eight gave no reason for dropping out. Twelve of 44 (27.3%) women taking placebo dropped out of the trial, of whom three disliked the drug, three reported side effects (type not specified) and six gave no reason.

[Bloch 2012](#) reported that seven women (3/20 from the placebo group and 4/20 from the antidepressant group) discontinued the trial between week four and week eight, at which point primary outcomes were reported. Among drop-outs from the placebo group, two reported lack of motivation as their reason for drop-out and one reported clinical deterioration; among drop-outs from the antidepressant group, two reported lack of motivation as their reason for drop-out and two reported clinical deterioration). No details on adherence to medications in either treatment arm were provided by either [Appleby 1997](#) or [Bloch 2012](#).

### 1.5 Cognitive development of the infant

No data available.

### 1.6 Maternal satisfaction

No data available.

### 1.7 Maternal relationship with the baby

No data available.

### 1.8 Social functioning

No data available.

### 1.9 Establishment or continuation of breastfeeding

No data available.



### 1.10 Neglect or abuse of the baby

No data available.

### 1.11 Effect on marital and family relationships

No data available.

### 1.12 Quality of life

No data available.

## Comparison 2: antidepressants versus treatment as usual

### Primary outcomes

#### 2.1 Remission/response of depression

One study (254 women) compared antidepressants with treatment as usual (outcomes assessed after four weeks) (Sharp 2010). There were no set antidepressants (prescriptions based on GP/participant choice) although most participants were prescribed citalopram, fluoxetine or sertraline (see [Characteristics of included studies](#) table for details). Results showed higher remission (EPDS less than 13) for women in the antidepressant group compared with treatment as usual after four weeks (improvement: 37.2% (48/129) women with antidepressant group versus 17.6% (22/125) women with treatment as usual). The RR was significant (RR 2.11, 95% CI 1.36 to 3.28; Analysis 2.1). This RR was based on the assumption that all the women who were lost to follow-up did not have remission of depression (calculated for this review on a 'once-randomised always-analyse' basis). In the Sharp 2010 study, multiple imputation was also used to impute missing data and ORs calculated. At four weeks, the imputed OR for remission in the antidepressant group compared with treatment as usual was 3.2 (95% CI 1.7 to 6.1).

#### 2.2 Adverse effects

No adverse events or serious side effects of treatment were reported in the Sharp 2010 trial. No data were reported on adverse effects related to infants or the safety of breastfeeding.

### Secondary outcomes

#### 2.3 Severity of depression

When outcomes were assessed in terms of continuous measurement of EPDS (for those followed up only; adjusted for baseline EPDS score and centre), this resulted in a two-point difference in means in favour of the antidepressant group when compared with treatment as usual at four weeks (MD -2.1, 95% CI -3.3 to -0.9; P value < 0.001).

#### 2.4 Acceptability

In the Sharp 2010 trial, more women in the antidepressant group withdrew or were lost to follow-up than women randomised to treatment as usual (withdrawal at four weeks: antidepressant group 23/129 (17.8%), treatment as usual group 13/125 (10.4%); P value = 0.090). From the reported data, it was not possible to determine whether this difference in withdrawal was due to a lack of acceptability of treatment with antidepressants or to other factors. Adherence was assessed in this trial using the Morisky Adherence Scale and four items adapted from a scale reported by Schroeder 2006. Authors reported low adherence to treatment: the percentage of women randomised to antidepressants who reported actually taking antidepressants in the previous four weeks

was 56% at four weeks (59/106; only reported for the women followed up).

#### 2.5 Cognitive development of the infant

No data available.

#### 2.6 Maternal satisfaction

No data available.

#### 2.7 Maternal relationship with the baby

Sharp 2010 examined the effect on maternal functioning using the Maternal Adjustment and Maternal Attitudes (MAMA) Attitudes Towards Pregnancy and Baby subscale (postpartum version). The study found weak evidence of benefit to the antidepressant group compared with treatment as usual at four weeks (adjusted MD 1.1, 95% CI -0.02 to 2.2 P value = 0.05).

#### 2.8 Social functioning

No data available.

#### 2.9 Establishment or continuation of breastfeeding

No data available.

#### 2.10 Neglect or abuse of the baby

No data available.

#### 2.11 Effect on marital and family relationships

Sharp 2010 examined the effect of treatment on marital relationships using the Golombok Rust Inventory of Marital State (GRIMS) scale. There was no evidence of any difference in marital relationship between treatment groups (4 weeks: -0.6, 95% CI -1.9 to 0.7; P value = 0.39).

#### 2.12 Quality of life

Sharp 2010 investigated the effect of treatment with antidepressants compared with treatment as usual on health-related quality of life, measured with the 12-item Short Form (SF-12) Mental Health and Physical Health components and the EQ-5D. They found a significant difference in favour of antidepressants on the SF-12 Mental Health component at four weeks (MD 0.36, 95% CI 0.14 to 0.57; P value = 0.001) but no difference on the SF-12 physical component score (MD -0.002, 95% CI -0.24 to 0.23; P value = 0.98). The EQ-5D utility score showed marginal, non-significant evidence in favour of the antidepressant group at four weeks with a reported adjusted MD of 0.05 (95% CI -0.002 to 0.11; P value = 0.059) and using the EQ-5D visual analogue scale there was no significant difference (MD 3.5, 95% CI -1.8 to 8.8; P value = 0.20).

## Comparison 3: antidepressants versus psychosocial interventions

### Antidepressants versus listening visits

#### Primary outcomes

#### 3.1 Remission/response of depression

Sharp 2010 compared antidepressants with listening visits at the 18-week outcome assessment. Antidepressants were not significantly more effective than listening visits, with remission of depression (EPDS less than 13) occurring in 46.5% (60/129)

of women randomised to antidepressants compared with 44.8% (56/125) of women randomised to listening visits (RR 1.04, 95% CI 0.79 to 1.36; see Analysis 3.1). [Sharp 2010](#) also performed multiple imputation to address missing data and calculated an OR for remission of depression at 18 weeks, which was not significant (OR 1.4, 95% CI 0.8 to 2.4).

### 3.2 Adverse effects

[Sharp 2010](#) reported no adverse events or serious side effects of treatment at 18 weeks.

#### Secondary outcomes

### 3.3 Severity of depression

When antidepressants and listening visits were compared at 18 weeks in [Sharp 2010](#) (for those followed-up only), there was no evidence of a significant difference between the groups in severity of depression (MD in EPDS scores: -0.7, 95% C.I. -2.1 to 0.8; P value = 0.37).

### 3.4 Acceptability

Significantly more women withdrew or were lost to follow-up in the antidepressant group (32/129 (24.8%)) than the listening visits group (16/125 (12.8%)) (P value = 0.015). In terms of adherence at 18 weeks, 64% of women randomised to antidepressants reported taking antidepressants in the previous four weeks (62/97 women followed up) and 34% of women randomised to listening visits reported taking antidepressants in the past four weeks (37/109 followed up).

### 3.5 Cognitive development of the infant

No data available.

### 3.6 Maternal satisfaction

No data available.

### 3.7 Maternal relationship with the baby

There was no difference in maternal functioning between the antidepressant and listening visits groups at 18 weeks in [Sharp 2010](#) (adjusted differences 1.0, 95% CI -0.3 to 2.2; P value = 0.14), assessed using the MAMA Attitudes Towards Pregnancy and Baby subscale (postpartum version).

### 3.8 Social functioning

No data available.

### 3.9 Establishment or continuation of breastfeeding

No data available.

### 3.10 Neglect or abuse of the baby

No data available.

### 3.11 Effect on marital and family relationships

There was no evidence of any difference between the antidepressant and listening visits group for marital relationships (assessed using the GRIMS scale; -1.2 (95% CI -2.8 to 0.4; P value = 0.14) in [Sharp 2010](#).

### 3.12 Quality of life

At 18 weeks, [Sharp 2010](#) compared antidepressants with listening visits and found no difference on the SF-12 Mental Health component score (MD 0.09, 95% CI -0.19 to 0.37; P value = 0.53) or Physical Health component score (MD 0.12, 95% CI -0.12 to 0.36; P value = 0.34). No significant differences were identified on the EQ-5D utility score (adjusted difference -0.01, 95% CI -0.08 to 0.05; P value = 0.68) or on the EQ-5D visual analogue scale (MD -1.6, 95% CI -8.1 to 4.9; P value = 0.63).

## Comparison 4: antidepressants versus other pharmacological interventions

### Selective serotonin re-uptake inhibitors versus tricyclic antidepressants

#### Primary outcomes

#### 4.1 Remission/response of depression

One study (109 women) compared an SSRI (sertraline) with a TCA (nortriptyline) and found no significant difference in effectiveness for treating postnatal depression, with primary outcome measures at week eight of treatment ([Wisner 2006](#)).

No differences between drug groups were observed using ITT analysis of the proportion of women who responded (50% reduction in HAM-D from baseline to week eight), 56% (31/55) of women randomised to sertraline and 69% (37/54) of women randomised to nortriptyline (RR 0.82, 95% CI 0.61 to 1.10; Analysis 4.1). There was also no difference in the proportion of women who remitted (HAM-D less than 7 at week eight): 46% (25/55) of women randomised to sertraline and 48% (26/54) of women randomised to nortriptyline (RR 0.94, 95% CI 0.63 to 1.41; Analysis 4.2).

#### 4.2 Adverse effects

In the [Wisner 2006](#) study, there was no difference between sertraline and nortriptyline in the overall number of side effects reported (using the Asberg Side Effects Rating Scale;  $\chi^2_1 = 0.00$ ; P value = 1.00). However, some side effects were more common among women who took nortriptyline than women taking sertraline: cholinergic symptoms such as moderate to severe thirst (after week three: 19% to 23% with nortriptyline versus 3% to 4% with sertraline; P value = 0.02), dry mouth (20% to 40% with nortriptyline versus 2% to 11% with sertraline for weeks two to eight; P value = 0.001) and constipation (23% to 25% with nortriptyline versus 7% to 12% with sertraline; P value = 0.05). Other side effects were more common in the sertraline than nortriptyline group: constant or severe headaches (10% to 15% with sertraline versus 1% to 2% with nortriptyline; P value = 0.05 at weeks two and three), slight to moderate increased perspiration (35% to 40% with sertraline versus 15% to 20% with nortriptyline for weeks one to three; P value = 0.04) and hot flushes interrupting sleep (4% to 10% with sertraline versus 0% to 2% with nortriptyline for weeks one to three; P value = 0.04).

[Wisner 2006](#) reported that babies of breastfeeding mothers in the trial had no adverse effects.

#### Secondary outcomes

#### 4.3 Severity of depression

For the 83 women providing a minimum of three weeks' follow-up data, there was no difference between drug groups for depression symptoms at four and eight weeks or across eight to 24 weeks ( $\chi^2$

$I^2 = 0.08$ ;  $P$  value = 0.77). The interaction of time by drug group on depressive symptoms was also not significant ( $\text{Chi}^2_8 = 3.64$ ;  $P$  value = 0.89).

#### 4.4 Acceptability

Significantly more women randomised to sertraline than women randomised to nortriptyline withdrew from the study in the first eight weeks (23/55 (42%) with sertraline versus 13/54 (24%) with nortriptyline; Wilcoxon;  $P$  value = 0.02), with a significantly higher proportion of women lost to follow-up or withdrawing by personal choice in the sertraline group (20% with sertraline versus 6% with nortriptyline; Wilcoxon  $\text{Chi}^2_1 = 4.86$ ;  $P$  value = 0.03). The proportion of women withdrawing for other reasons (side effects, hypomania occurrence or clinical deterioration) did not differ significantly between the two drug groups. There were no significant differences in rates of withdrawal after week eight (entering the continuation phase of the trial). Of those eligible to enter the continuation phase, 24/32 (75%) of those randomised to sertraline and 25/40 (63%) of those randomised to nortriptyline chose to do so ( $\text{Chi}^2_1 = 1.28$ ;  $P$  value = 0.26). Adherence (assessed through serum levels) found that 14 women had minimal drug levels in their blood despite claims of compliance. There was no significant difference found in lack of compliance between women assigned to nortriptyline (9/51, 18%) and women assigned to sertraline (5/44, 11%; Fisher exact test;  $P$  value = 0.29).

#### 4.5 Cognitive development of the infant

No data available.

#### 4.6 Maternal satisfaction

No data available.

#### 4.7 Maternal relationship with the baby

No data available.

#### 4.8 Social functioning

[Wisner 2006](#) investigated the effect of treatment with nortriptyline versus sertraline on social functioning as assessed by the Social Problems Questionnaire. No significant effect of treatment modality was found at week eight (change in log-likelihood  $\text{Chi}^2_1 = 0.25$ ;  $P$  value = 0.62) or when the interaction of time by drug group was examined (change in log-likelihood  $\text{Chi}^2_2 = 2.22$ ;  $P$  value = 0.33). There were also no significant differences between antidepressant groups at week 24.

#### 4.9 Establishment or continuation of breastfeeding

No data available.

#### 4.10 Neglect or abuse of the baby

No data available.

#### 4.11 Effect on marital and family relationships

No data available.

#### 4.12 Quality of life

No data available.

## Subgroup analyses

We were unable to conduct subgroup analyses due to the small number of studies included in the meta-analysis for the SSRIs versus placebo comparison and the lack of data for other meta-analyses.

## Sensitivity analyses

### *Comparison: selective serotonin reuptake inhibitors versus placebo*

Sensitivity analyses were conducted to examine the effect of removing studies with combined treatment (i.e. [Bloch 2012](#) in which all participants received brief dynamic psychotherapy as well as sertraline or placebo). After removing [Bloch 2012](#) the pooled risk ratio for response was 1.62 (95% CI 0.98 to 2.67; [Analysis 5.1](#)) and the pooled risk ratio for remission was 2.56 (95% CI 1.31 to 5.00; [Analysis 5.2](#)). In both cases this reflects an increase in risk ratios (from RR 1.43 for response and RR 1.79 for remission in the main analyses); however confidence intervals also increased as the pooled estimates were now based on just two studies and only the effect on remission remained statistically significant.

Additional sensitivity analyses were conducted to examine the effects of removing studies with high dropout or high risk of bias in any domain; both of these required [Yonkers 2008](#) to be removed from the pooled estimates. After removing [Yonkers 2008](#) from the analyses, the pooled risk ratio for response was 1.52 (95% CI 0.89 to 2.58; [Analysis 6.1](#)) and the pooled risk ratio for remission was 1.60 (95% CI 0.86 to 2.97; [Analysis 6.2](#)). The effect size for response in this sensitivity analyses was slightly larger than in the main analyses (from RR 1.43 to RR 1.52), but the effect size for remission was reduced (from RR 1.79 to RR 1.60). Again, these pooled estimates were now based on two studies only so confidence intervals were extremely wide and neither effect was statistically significant.

## Reporting bias

We were unable to assess reporting bias due to the small number of studies included in the meta-analysis for the SSRIs versus placebo comparison and the lack of data for other meta-analyses.

## DISCUSSION

### Summary of main results

We identified six RCTs (596 women) examining the effectiveness of antidepressants (predominantly SSRIs: sertraline, paroxetine or fluoxetine, and a TCA: nortriptyline) for postnatal depression. Four studies comparing SSRIs with placebo were identified; in two of these trials all participants also received psychological therapy. Meta-analyses of the three studies (146 participants) for this comparison with relevant data found that participants randomised to SSRIs were significantly more likely to show response or remission of depression at follow-up compared with participants randomised to placebo. However, these findings must be interpreted with caution; the quality of evidence was graded as 'very low' owing the small number of included studies, high risk of bias in some trials (including over 50% drop-out in one study) and the pooling of results from one study which provided psychological therapy in both the SSRI or placebo arms ([Bloch 2012](#)) with two studies that compared SSRIs and placebo only (see [Summary of findings for the main comparison](#)).

It was not possible to conduct meta-analyses for the other comparisons and there is insufficient evidence to conclude whether, and for whom, antidepressant or psychosocial treatments are more effective, or whether some antidepressants are more effective or better tolerated than others. One study showed a significant benefit of antidepressants compared with treatment as usual (Table 1), but there was no evidence of a benefit of antidepressants compared with listening visits after these had been introduced at later follow-up (see Table 2). No difference in effectiveness was demonstrated in the one study comparing sertraline with nortriptyline (see Table 3), although a significantly higher proportion of women randomised to sertraline withdrew in the first eight weeks of the study compared with those randomised to nortriptyline which may suggest a difference in the acceptability of the treatments. The current evidence on antidepressant treatment of postnatal depression is limited by the small number of RCTs, underpowered samples, lack of long-term follow-up or child outcomes and other study limitations such as risk of bias. There were few data on the safety of breastfeeding, adverse effects for the infants or long-term outcomes for the mother and child. Two studies reported that there were no adverse effects for breastfed infants (Hantsoo 2013; Wisner 2006), but this was based on small numbers (e.g. six mothers randomised to sertraline were breastfeeding in Hantsoo 2013) and limited assessment.

Side effects were reported by a substantial proportion of women and were mainly characteristic of the type of antidepressant used with nausea, diarrhoea and headaches reported with SSRIs and constipation with nortriptyline. It was often difficult to interpret the severity of side effects and several studies were limited in their assessment and reporting of side effects and adverse events; Appleby 1997 only reported adverse events in reasons for withdrawal so we do not know whether there were side effects among women who remained in trial and Sharp 2010 reported only 'serious side effects' (of which there were none). High attrition rates with limited reporting of reasons for withdrawal and the overall lack of child outcomes make it difficult to draw conclusions about adverse outcomes in the included studies, particularly any adverse outcomes related to breastfed infants. It was also difficult to make any conclusions on most of secondary outcomes, as very few of these were addressed in any included studies. Although acceptability was generally not specifically assessed, high drop-out, low adherence and low recruitment may reflect limited acceptability of antidepressants or RCTs in the postnatal period.

### Overall completeness and applicability of evidence

This evidence base is limited by the fact that there have been few trials of antidepressants in the postnatal period, and most of these were underpowered to detect significant differences in treatment effect. Data could only be pooled for the SSRI versus placebo comparison and this meta-analysis included only three studies. Sensitivity analyses were conducted to examine the effect of removing Bloch 2012 (all participants received psychological therapy as well as SSRI or placebo) or Yonkers 2008 (high drop-out and high risk of bias) from the pooled estimates but are difficult to interpret as each was based on two studies only and had extremely wide confidence intervals. There were insufficient data to conduct meta-analyses for the other comparisons so these conclusions are drawn from individual studies.

Studies recruited women presenting with a major depressive disorder (except for Appleby 1997, which also included women with

minor depressive disorder) in the first few months postpartum. However, many studies had restrictive inclusion criteria and excluded women with the most severe disorders, which will affect the generalisability of findings. This is a particular limitation given the need to establish an evidence base for the treatment of severe depression in the postnatal period to enable informed decision making for women who may be reluctant to take antidepressants. For example, four studies excluded women with chronic depression (variously defined), four studies excluded women with suicidal ideation and three studies excluded women with treatment-resistant depression (usually defined as past failed trial or trials of antidepressant treatment). Women with psychotic disorders or drug or alcohol use were also excluded from the majority of studies. The findings of these studies may therefore not be generalisable to these women, who may be those in most need of pharmacological treatment. Chronic depression or depression with other co-morbid mental disorders may be more resistant to treatment; however, there is evidence from non-pregnant adults that antidepressants are more effective compared with placebo for individuals with severe depression (Kirsch 2008). The three studies that excluded women who had previously been resistant to treatment for depression may have biased findings, particularly as Hantsoo 2013 excluded women who had previously not responded to sertraline from their trial comparing sertraline with placebo. One study allowed a choice of antidepressant medication with 14 different antidepressants prescribed over the course of the study (majority citalopram, fluoxetine, sertraline; see Characteristics of included studies for full details) (Sharp 2010). However, as the results were not reported separately for each antidepressant, conclusions about the effectiveness and safety of individual antidepressants cannot be drawn from this study. The fact that there was choice of antidepressant in the Sharp 2010 study means that this (more pragmatic) trial addresses a slightly different question from those that randomised participants to a particular antidepressant. It should also be noted that the antidepressant groups in all studies received usual care alongside their prescriptions, including appointments with the study doctors (to monitor prescriptions). This needs to be taken into account when interpreting findings, particularly for the antidepressants versus treatment as usual comparison in Sharp 2010, where both groups received the same usual care (supportive visits with their GP).

The studies in this review are all from high-income countries, with five of the six studies from UK or US settings. Conclusions therefore cannot be generalised to medium- and low- income settings, a limitation given the elevated prevalence of perinatal mental disorders observed in these settings (Fisher 2012). Most of the women in the included studies were also white Caucasian, which may prevent results being applicable to ethnic minority populations. Teenage mothers were also often excluded (Wisner 2006 included women over 15 years of age and Yonkers 2008 included women over 16 years of age; other studies were limited to women over 18 years). Again, this limits generalisation to a group particularly vulnerable to postnatal depression (Figueiredo 2007).

The evidence is also limited by the small number of antidepressants individually assessed (four antidepressants only, of which three were SSRIs) and the extremely limited data on longer-term, infant and secondary outcomes. Response and remission rates varied substantially between studies, with relatively low rates even in the antidepressant groups in some studies (e.g. Sharp

2010). This is likely to reflect differences in study methodology (particularly participant inclusion and exclusion criteria and length of follow up) and adherence but should be considered when interpreting findings. Studies of other antidepressants, including evaluation of the impact of antidepressants on outcomes for breastfeeding mothers and infants, would increase the evidence base on treatment options for postnatal women.

### Quality of the evidence

The risk of bias varied between domains in the six trials included in the review. Random sequence generation and blinding were generally performed and reported adequately (with some exceptions), but assessment of allocation concealment and selective outcome reporting was problematic due to insufficient detail in most included studies and lack of availability of study protocols. The greatest limitations to the quality of evidence were issues with recruitment and attrition. Most studies experienced difficulties in recruitment, leading to studies being underpowered to detect a significant treatment effect, and many also had high rates of drop-out (the highest being 56% drop-out in [Yonkers 2008](#)). Although ITT analysis is reported for all primary outcomes in this review, drop-out may have introduced bias particularly if reasons for drop-out, such as clinical worsening, differed between groups. In addition, some secondary outcomes were not reported as ITT, which is likely to have introduced bias for these data. The amount of attrition was generally well described, but the reasons for and timing of drop-out were often insufficiently detailed to assess the likelihood of meaningful bias. In addition, adherence was low in many studies (particularly [Sharp 2010](#), where there was also substantial cross-over between the study groups). Future studies should report more clearly on all areas of potential bias, with particular focus on reducing and describing attrition. Assessing studies using the GRADEpro criteria (see [Summary of findings for the main comparison](#); [Table 1](#); [Table 2](#); [Table 3](#)) showed that the quality of the evidence for all comparisons was low or very low (low: antidepressant versus antidepressant; very low: antidepressant versus placebo, antidepressants versus treatment as usual, antidepressants versus listening visits). This demonstrates the uncertainty in the estimates and the likelihood that further research will have an important impact on our confidence in the estimates of the effect sizes or the effects themselves, or both. The evidence was limited by the risk of bias and imprecision in the included study estimates.

As RRs were used for the primary outcome in this review but not in any original papers, the conclusions here do not always match those of the primary papers. For example, [Hantsoo 2013](#) reported significant differences in response and remission between the antidepressant and placebo group based on Chi<sup>2</sup> tests; however, the RRs for these data were not significant. This is due to the small sample size and low precision in the estimate; the RRs themselves suggest substantial benefit to the antidepressant group but the 95% CIs were very wide.

### Potential biases in the review process

We believe that all relevant RCTs were identified by the systematic review process. Extensive efforts were made to identify relevant papers through database searches, citation tracking, and contacting pharmaceutical companies and experts in the field for knowledge of relevant papers or unpublished data. Our searches of clinical trial registries identified one ongoing study that may meet

criteria for the review; this is described in the [Characteristics of ongoing studies](#) table. We also identified two studies where data collection has been completed but data are not yet available, these are listed in the [Characteristics of studies awaiting classification](#) tables. We also contacted authors of conference presentations identified by the initial search results when there was no record of an associated publication. Despite these efforts, it is possible that publication bias may have influenced the review findings. This could not be assessed (e.g. through funnel plots) owing to the small number of studies.

Two review authors independently performed study screening, data extraction and risk of bias assessment with a third review author resolving any discrepancies remaining after discussion. We included one study with a dropout rate above 50% ([Yonkers 2008](#)). This may have introduced bias; however, drop-out in this study was balanced between the antidepressant and placebo groups, and the findings of the paper were robust to sensitivity analyses assuming that all drop-outs were remitters or non-remitters. Owing to the small number of relevant studies, we believe that the data from this trial were valuable, although its conclusions must be interpreted with caution. A number of other updates to the protocol were made (e.g. the specification of response/remission as the primary depression outcomes, the inclusion of quality of life as a secondary outcome and the increased sub-categorisation of antidepressant types). We also updated the protocol so that non-validated scales could be included as outcome measures (with planned sensitivity analyses to examine the effect of non-validated scales on findings). This change was made to reduce the potential for bias or limitations to the evidence base from excluding these measures, balanced against the need to assess the potential for bias from including non-validated measures. While changes to the protocol may introduce bias in the review process, most changes made here had no impact on the current review as they related to items that were not present in included studies (e.g. certain types of antidepressants, non-validated scales). However, these changes should benefit future updates of this review.

### Agreements and disagreements with other studies or reviews

Although the effectiveness of antidepressants as a treatment for depression has been established in the general adult population ([Arroll 2005](#)), there is a paucity of studies, particularly RCTs, examining this in the postnatal period. Our review includes similar studies to those reported in a systematic review focused on SSRIs for postnatal depression ([De Crescenzo 2014](#)), although an additional published study is reported here ([Hantsoo 2013](#)), and [De Crescenzo 2014](#) included a study comparing paroxetine only with paroxetine plus CBT ([Misri 2004](#)), which did not meet our inclusion criteria. De Crescenzo et al concluded that antidepressants appear to reduce postnatal depression effectively without severe adverse effects but emphasised limitations including the small number of studies, high dropout rates, unrepresentative samples and lack of long-term follow up or assessment of acceptability. They also highlighted that there is insufficient evidence to demonstrate a superiority of SSRIs over other treatments. The addition of [Hantsoo 2013](#) in this review allowed us to conduct a small meta-analysis comparing SSRIs with placebo, however our overall conclusions are similar and also emphasise the limitations of the evidence, particularly regarding potential adverse outcomes for the mother and infant.

## AUTHORS' CONCLUSIONS

### Implications for practice

The evidence base reported in this review is of very low quality and includes only a small number of studies, which imposes significant limitations for conclusions on both efficacy and potential adverse outcomes for the mother and baby. It is difficult to draw implications for practice from the findings reported here, particularly due to the lack of evidence on child outcomes, which may be particularly important for women with postnatal depression. The trials included here focused on mild to moderate depression and suggested that, while SSRIs were found to be significantly more effective than placebo, there was little difference in effectiveness when comparing antidepressants with psychological/psychosocial interventions.

Women with mild to moderate depression should be informed that there is no current evidence to suggest a clear difference between the effectiveness of antidepressants and psychological/psychosocial treatments; shared decision-making to weigh up the potential for benefits and harms for both the mother and child needs to be implemented in treatment decisions. Women with severe or chronic depression or suicidal ideation were excluded from several of the included studies, which is also a major limitation for making clinical recommendations. Clinicians treating women with severe depression in the postnatal period will need to draw on the evidence base for severe depression from outside of the postnatal period as well as observational studies examining the impact of antidepressant medication on breastfeeding infants (taking into account the potential for confounding in non-randomised studies).

### Implications for research

Postnatal depression is the most common complication of childbearing and is associated with substantial morbidity for the mother and child. However, there are few randomised controlled trials examining the effectiveness of antidepressant treatments in the postnatal period and more trials in this period are needed. Further research is particularly required to demonstrate the impact and safety of antidepressant treatments in the postnatal period regarding child outcomes and long-term follow-up. This is of particular importance for breastfeeding women and their infants, as positive or negative effects related to antidepressant treatment may not emerge immediately. The safety of antidepressants for the baby and in breastfeeding are often critical concerns for women in the postnatal period, so these outcomes must be assessed and reported by future research. Other outcomes that are meaningful for women, such as maternal satisfaction, mother-child

interactions and quality of life, also need to be evaluated, as does the severity of side-effects experienced.

Future trials should try to address other limitations of the current evidence base, including small sample sizes and high attrition. Reasons for non-participation and drop-out should be assessed in detail in future studies and substantial efforts should be made to reduce their impact, including realistic estimates of recruitment and dropout rates in sample size calculations and study protocols. Broader inclusion criteria (e.g. not excluding women with chronic depression) could improve recruitment and also increase the generalisability of findings. It is particularly important to develop an evidence base for women with severe/chronic depression or suicidal ideation in the postnatal period.

One study in this review allowed a choice of antidepressants (decision made by study GPs and participants) (Sharp 2010), and future studies could compare treatment effects in a group randomised to a choice of antidepressant medications, compared with women randomised to a specific antidepressant with no choice. Fidelity to antidepressant and comparison treatments among trial participants should also be assessed in all future trials. Further research could also examine combination therapy in comparison with either antidepressants or psychological therapy. Finally, future studies should examine whether psychological interventions are more likely to prevent relapse than antidepressants; cognitive and behavioural strategies acquired through psychological interventions may help prevent relapse when pharmacological treatment has been stopped but there has been little research in this area.

## ACKNOWLEDGEMENTS

We would like to thank the Trials Search Co-ordinator of the Cochrane Collaboration Depression, Anxiety and Neurosis Group, and Reinhard Wentz of the Institute of Child Health, for assistance with developing the search strategy for this review.

We would also like to thank the authors of the previous version of this review (Sara Hoffbrand, Louise Howard and Helen Crawley) and Wei Dai, Hui Li and Jin Huajie for their assistance with translation.

### CRG funding acknowledgement:

The National Institute for Health Research (NIHR) is the largest single funder of the Cochrane Depression, Anxiety and Neurosis Group.

### Disclaimer:

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

## REFERENCES

### References to studies included in this review

#### Appleby 1997 {published data only}

Appleby L. A controlled study of fluoxetine and cognitive-behavioural counselling in the treatment of postnatal depression. Proceedings of the 9th Congress of the Association of European Psychiatrists; 1998 Sept 20-24; Copenhagen, Denmark.

\* Appleby L, Warner R, Whitton A, Faragher B. A controlled study of fluoxetine and cognitive-behavioural counselling in the treatment of postnatal depression. *BMJ* 1997;**314**(7085):932-6.

Warner R, Appleby L, Whitton A, Faragher B. A controlled study of fluoxetine and cognitive behavioural counselling in the treatment of post-natal depression. Proceedings of the Royal College of Psychiatrists Winter Meeting; 1997 Jan 21-24; Cardiff, UK.

#### Bloch 2012 {published data only}

Bloch M, Meiboom H, Lorberblat M, Aharonov I, Bluvstein I, Ben-Itzhak S, et al. Treatment of postpartum depression with psychotherapy and add-on sertraline: a double-blind, randomised, placebo-controlled study. *European Neuropsychopharmacology* 2011;**21**(Suppl 3):S359-60.

\* Bloch M, Meiboom H, Lorberblatt M, Bluvstein I, Aharonov I, Schreiber S. The effect of sertraline add-on to brief dynamic psychotherapy for the treatment of postpartum depression: a randomized, double-blind, placebo-controlled study. *Journal of Clinical Psychiatry* 2012;**73**(2):235-41.

#### Hantsoo 2013 {published data only}

Epperson CN, McDougle CJ, Ward-O'Brien D, Price LH. A controlled study of antidepressant treatment of postpartum depression. Proceedings of the 149th Annual Meeting of the American Psychiatric Association; 1996 May 4-9; New York.

\* Hantsoo L, Ward-O'Brien D, Czarkowski KA, Gueorguieva R, Price LH, Epperson CN. A randomized, placebo-controlled, double-blind trial of sertraline for postpartum depression. *Psychopharmacology* 2014;**231**(5):939-48. [DOI: [10.1007/s00213-013-3316-1](https://doi.org/10.1007/s00213-013-3316-1)]

#### Sharp 2010 {published data only}

Chew-Graham C, Chamberlain E, Turner K, Folkes L, Caulfield L, Sharp D. GPs' and health visitors' views on the diagnosis and management of postnatal depression: a qualitative study. *British Journal of General Practice* 2008;**58**(548):169-76.

Chew-Graham CA, Sharp D, Chamberlain E, Folkes L, Turner KM. Disclosure of symptoms of postnatal depression, the perspectives of health professionals and women: a qualitative study. *BMC family practice* 2009;**10**:7.

Howard L, Chew-Graham CA, Tylee A, Lewis G, Anderson I, Abel K, et al. The RESPOND trial: a randomized evaluation of antidepressants and support for women with postnatal depression. *Archives of Women's Mental Health* 2011;**14**(Suppl 1).

Howard LM, Flach C, Mehay A, Sharp D, Tylee A. The prevalence of suicidal ideation identified by the Edinburgh Postnatal Depression Scale in postpartum women in primary care: findings from the RESPOND trial. *BMC Pregnancy and Childbirth* 2011;**11**:57.

\* Sharp D J, Chew-Graham C A, Tylee A, Lewis G, Howard L, Anderson I, et al. A pragmatic randomised controlled trial to compare antidepressants with a community-based psychosocial intervention for the treatment of women with postnatal depression: the RESPOND trial. *Health Technology Assessment* 2010;**14**(43):1-181.

Turner KM, Chew-Graham C, Folkes L, Sharp D. Women's experiences of health visitor delivered listening visits as a treatment for postnatal depression: a qualitative study. *Patient Education and Counseling* 2010;**78**(2):234-9.

Turner KM, Sharp D, Folkes L, Chew GC. Women's views and experiences of antidepressants as a treatment for postnatal depression: a qualitative study. *Family Practice* 2008;**25**(6):450-5.

Wan MW, Sharp DJ, Howard LM, Abel KM. Attitudes and adjustment to the parental role in mothers following treatment for postnatal depression. *Journal of Affective Disorders* 2011;**131**(1-3):284-29.

#### Wisner 2006 {published data only}

Lanza di Scalea T, Hanusa BH, Wisner KL. Sexual function in postpartum women treated for depression: results from a randomized trial of nortriptyline versus sertraline. *Journal of Clinical Psychiatry* 2009;**70**(3):423-8.

Logsdon MC, Wisner K, Hanusa BH. Does maternal role functioning improve with antidepressant treatment in women with postpartum depression?. *Journal of Women's Health* 2009;**18**(1):85-90.

Logsdon MC, Wisner K, Hanusa BH, Philips A. Role functioning and symptom remission in women with postpartum depression after antidepressant treatment. *Archives of Psychiatric Nursing* 2003;**17**(6):276-83.

\* Wisner KL, Hanusa BH, Perel JM, Peindl KS, Piontek CM, Sit DK, et al. Postpartum depression: a randomized trial of sertraline versus nortriptyline. *Journal of Clinical Psychopharmacology* 2006;**26**(4):353-60.

#### Yonkers 2008 {published data only}

Heath AC, Yonkers KA, Rush AJ. Paroxetine in the treatment of postpartum depression. Proceedings of the 153rd Annual Meeting of the American Psychiatric Association; 2000 May 13-8; Chicago, IL.

Heath AC, Yonkers KA, Rush AJ. Paroxetine in the treatment of postpartum depression. Proceedings of the 155th Annual Meeting of the American Psychiatric Association; 2002 May 18-23; Philadelphia, PA.

\* Yonkers KA, Lin H, Howell HB, Heath AC, Cohen LS. Pharmacologic treatment of postpartum women with new-onset major depressive disorder: a randomized controlled trial with paroxetine. *Journal of Clinical Psychiatry* 2008;**69**(4):659-65.

## References to studies excluded from this review

### Bennett 2001 {published data only}

Bennett P. The costs and benefits of treatment of post-natal depression in the community. National Research Register 2001.

Honey KL, Bennett P, Morgan M. A brief psycho-educational group intervention for postnatal depression. *British Journal of Clinical Psychology* 2002;**41**:405-9.

### Misri 2004 {published data only}

Misri S, Reebye P, Corral M, Milis L. The use of paroxetine and cognitive-behavioral therapy in postpartum depression and anxiety: a randomized controlled trial. *Journal of Clinical Psychiatry* 2004;**65**(9):1236-41.

Misri S, Reebye P, Milis L, Shah S. The impact of treatment intervention on parenting stress in postpartum depressed mothers: a prospective study. *American Journal of Orthopsychiatry* 2006;**76**(1):115-9.

### Rojas 2007 {published data only}

Rojas G, Fritsch R, Solis J, Jadresic E, Castillo C, González M, et al. Treatment of postnatal depression in low-income mothers in primary-care clinics in Santiago, Chile: a randomised controlled trial. *Lancet* 2007;**370**(9599):1629-37.

Zayas LH. Six-month multicomponent intervention improves postnatal depression in low-income settings. *Evidence-Based Mental Health* 2008;**11**(3):80.

### Stein 2012 {published data only}

Stein A, Murphy S, Artech A, Lehtonen A, Harvey A, Craske MG, et al. Effects of reboxetine and citalopram on appraisal of infant facial expressions and attentional biases. *Journal of Psychopharmacology* 2012;**26**(5):670-6.

### Suri 2005 {published data only}

Suri R, Burt VK, Altshuler LL. Nefazodone for the treatment of postpartum depression. *Archives of Women's Mental Health* 2005;**8**(1):55-6.

### Yu 2006 {published data only}

Yu H, Liu C, Wang X. The therapeutic effects of paroxetine addition of psychological intervention on post-partum depression. *Progress in Obstetrics and Gynecology* 2006;**15**(4):289-91.

### Zhao 2006 {published data only}

Zhao X, Lin H. Study on postpartum depression treated with combination of fluoxetine and Chaihu Shugan powder. *Liaoning Journal of Traditional Chinese Medicine* 2006;**33**(5):586-7.

## References to studies awaiting assessment

### NCT00744328 {unpublished data only}

Wisner KL. Postpartum Depression: Transdermal Estradiol Versus Sertraline (E2 SERT). <http://clinicaltrials.gov/show/NCT00744328> 2014 (accessed 04 August 2014).

Wisner KL. Should estradiol be prescribed for postpartum depression? [conference abstract]. *Biological Psychiatry* 2014;**9**(Suppl 1):299S.

### NCT02122393 {unpublished data only}

Milgrom J. A Randomised Trial of Sertraline, Cognitive Behaviour Therapy & Combined Therapy for Postnatal Depression. <http://www.clinicaltrials.gov/ct2/show/NCT02122393> 2014 (accessed 04 August 2014).

## References to ongoing studies

### NCT00602355 {unpublished data only}

Zlotnick C, Stuart S. Placebo controlled trial of sertraline and interpersonal psychotherapy for postpartum depression. <http://clinicaltrials.gov/ct2/show/NCT00602355> 2008 (accessed 27 July 2014).

## Additional references

### APA 1999

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th Edition. Washington, DC: American Psychiatric Association, 1999.

### APA 2013

American Psychiatric Association. DSM 5. American Psychiatric Association, 2013.

### Arroll 2005

Arroll B, Macgillivray S, Ogston S, Reid I, Sullivan F, Williams B, et al. Efficacy and tolerability of tricyclic antidepressants and SSRIs compared with placebo for treatment of depression in primary care: a meta-analysis. *Annals of Family Medicine* 2005;**3**(5):449-56.

### Ban 2012

Ban L, Gibson JE, West J, Fiaschi L, Oates MR, Tata LJ. Impact of socioeconomic deprivation on maternal perinatal mental illnesses presenting to UK general practice. *British Journal of General Practice* 2012;**62**(603):e671-8.

### Bayley 2006

Bayley N. Bayley Scales of Infant and Toddler Development® 3rd Edition (Bayley-III®). San Antonio, TX: The Psychological Corporation, 2006.

### Berle 2011

Berle JØ, Spigset O. Antidepressant use during breastfeeding. *Current Women's Health Reviews* 2011;**7**(1):28.



**Berton 2006**

Berton O, Nestler EJ. New approaches to antidepressant drug discovery: beyond monoamines. *Nature Reviews Neuroscience* 2006;**7**:137-51.

**Bland 2000**

Bland JM, Altman DG. Statistics notes: the odds ratio. *BMJ* 2000;**320**(7247):1468.

**Centre for Research and Dissemination 2009**

Centre for Research and Dissemination. Systematic reviews: CRD's guidance for undertaking systematic reviews in healthcare. York, UK: University of York, 2009.

**Cohen 1997**

Cohen S, Doyle WJ, Skoner DP, Rabin BS, Gwaltney JM Jr. Social ties and susceptibility to the common cold. *JAMA* 1997;**277**:1940-4.

**Cooper 1995**

Cooper PJ, Murray L. Course and recurrence of postnatal depression. Evidence for the specificity of the diagnostic concept. *British Journal of Psychiatry* 1995;**166**:191-5.

**Cox 1987**

Cox J, Holden JM, Sagovsky R. Detection of postnatal depression: development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry* 1987;**150**:782-6.

**Cox 1993**

Cox JL, Murray D, Chapman G. A controlled study of the onset, duration and prevalence of postnatal depression. *British Journal of Psychiatry* 1993;**163**:27-31.

**Crittenden 1988**

Crittenden PM. Relationships at Risk. In: Belsky J, Nezworski T editor(s). Clinical implications of attachment. Hillsdale, NJ: Lawrence Erlbaum Association, 1988:136-74.

**De Crescenzo 2014**

De Crescenzo F, Perelli F, Armando M, Vicatri S. Selective serotonin reuptake inhibitors (SSRIs) for post-partum depression (PPD): a systematic review of randomized clinical trials. *Journal of Affective Disorders* 2014;**152**:39-44.

**Dennis 2013**

Dennis C-L, Dowswell T. Interventions (other than pharmacological, psychosocial or psychological) for treating antenatal depression. *Cochrane Database of Systematic Reviews* 2013, Issue 7. [DOI: [10.1002/14651858.CD006795.pub3](https://doi.org/10.1002/14651858.CD006795.pub3)]

**Eberhard-Gran 2006**

Eberhard-Gran M, Esklid A, Opjordsmoen S. Use of psychotropic medications in treating mood disorders during lactation. Practical recommendations. *CNS Drugs* 2006;**20**:187-98.

**Elbourne 2002**

Elbourne DR, Altman DG, Higgins JPT, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: methodological issues. *International Journal of Epidemiology* 2002;**31**(1):140-9.

**Endicott 1976**

Endicott J, Spitzer RL, Fleiss JL, Cohen J. The Global Assessment Scale: a procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry* 1976;**33**(6):766.

**Figueiredo 2007**

Figueiredo B, Pacheco A, Costa R. Depression during pregnancy and the postpartum period in adolescent and adult Portuguese mothers. *Archives of Women's Mental Health* 2007;**10**(3):103-9.

**Fisher 2012**

Fisher J, de Mello MC, Patel V, Rahman A, Tran T, Holton S, et al. Prevalence and determinants of common perinatal mental disorders in women in low-and lower-middle-income countries: a systematic review. *Bulletin of the World Health Organization* 2012;**90**(2):139-49.

**Fournier 2010**

Fournier JC, DeRubeis RJ, Hollon SD, Dimidjian S, Amsterdam JD, Shelton RC, et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA* 2010;**303**(1):47-53.

**Gavin 2005**

Gavin NI, Gaynes BN, Lohr KN, Meltzer-Brody S, Gartlehner G, Swinson T. Perinatal depression: a systematic review of prevalence and incidence. *Obstetrics & Gynecology* 2005;**106**(5, Part 1):1071-83.

**Gidaud-Wallston 1978**

Gibaud-Wallston J, Wandersman LP. Development and utility of the Parenting Sense of Competence Scale. Proceedings of the American Psychological Association; 1978 Aug 28 - Sept 1; Toronto, Canada.

**Goodman 2004**

Goodman JH. Postpartum depression beyond the early postpartum period. *Journal of Obstetric, Gynecologic, & Neonatal Nursing* 2004;**33**(4):410-20.

**Gordon 2013**

Gordon A, Mikocka-Walus A, Grzeskowiak LE, Jayaasekara RS. Antidepressants for depression during pregnancy. *Cochrane Database of Systematic Reviews* 2013, Issue 8. [DOI: [10.1002/14651858.CD010710](https://doi.org/10.1002/14651858.CD010710)]

**Henshaw 2003**

Henshaw C. Mood disturbance in the early puerperium: a review. *Archives of Women's Mental Health* 2003;**6**(2):s33-42.

**Higgins 2011**

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from [www.cochrane-handbook.org](http://www.cochrane-handbook.org).

**Johnson 2008**

Johnson S, Dieter W, Marlow N. Developmental assessment of preterm infants at 2 years: validity of parent reports. *Developmental Medicine and Child Neurology* 2008;**50**(1):58-62.

**Kaufman Kantor 2004**

Kaufman Kantor G, Holt MK, Mebert CJ, Straus MA, Drach KM, Ricci LR, et al. Development and preliminary psychometric properties of the Multidimensional Neglectful Behavior Scale-Child Report. *Child Maltreatment* 2004;**9**(5):409-28.

**Kendell 1987**

Kendell RE, Chalmers L, Platz C. The epidemiology of puerperal psychoses. *British Journal of Psychiatry* 1987;**150**:662-73.

**Kirsch 2008**

Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Medicine* 2008;**5**(2):e45.

**Letourneau 2012**

Letourneau NL, Dennis C-L, Benzies K, Duffett-Leger L, Stewart M, Tryphonopoulos PD, et al. Postpartum depression is a family affair: addressing the impact on mothers, fathers, and children. *Issues in Mental Health Nursing* 2012;**33**(7):445-57.

**Ma 2014**

Ma Y. Neuropsychological mechanism underlying antidepressant effect: a systematic meta-analysis. *Molecular Psychiatry* 2014 Mar 25 [Epub ahead of print].

**Munk-Olsen 2006**

Munk-Olsen T, Laursen TM, Pedersen CB, Mors O, Mortensen PB. New parents and mental disorders. *JAMA* 2006;**296**(21):2582-9.

**Murray 1992**

Murray L. The impact of postnatal depression on infant development. *Journal of Child Psychology and Psychiatry* 1992;**33**(3):543-61.

**NICE 2007**

National Institute for Health and Clinical Excellence. Antenatal and Postnatal Mental Health: Clinical Management and Service Guidance. CG45. Vol. **45**, National Institute for Health and Clinical Excellence, 2007.

**Norton 1983**

Norton R. Measuring marital quality: a critical look at the dependent variable. *Journal of Marriage and Family* 1983;**45**:141-51.

**Nylen 2013**

Nylen KJ, Williamson JA, O'Hara MW, Watson D, Engeldinger J. Validity of somatic symptoms as indicators of depression in pregnancy. *Archives of Women's Mental Health* 2013;**16**(3):1-8.

**Pringle 2011**

Pringle A, Browning M, Cowen PJ, Harmer CJ. A cognitive neuropsychological model of antidepressant drug action. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 2011;**35**(7):1586-92.

**RevMan 2012 [Computer program]**

The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.2. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2012.

**Schroeder 2006**

Schroeder K, Fahey T, Hay AD, Montgomery A, Peters TJ. Adherence to antihypertensive medication assessed by self-report was associated with electronic monitoring compliance. *Journal of Clinical Epidemiology* 2006;**59**:650-1.

**Spitzer 1978**

Spitzer RL, Endicott J, Robins E. Research diagnostic criteria: rationale and reliability. *Archives of General Psychiatry* 1978;**35**(6):773-82.

**Ware 1992**

Ware JE, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): I. Conceptual framework and item selection. *Medical Care* 1992;**30**:473-83.

**Wechsler 1974**

Wechsler D. Wechsler Intelligence Scale for Children - Revised. New York, NY: Psychological Corp, 1974.

**WHO 1992**

World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders. Geneva: World Health Organization, 1992.

**WHO 2004**

World Health Organization. ICD-10: International Statistical Classification of Diseases and Related Health Problems. World Health Organization, 2004.

**Wisner 2004**

Wisner KL, Perel JM, Peindl KS, Hanusa BH. Timing of depression recurrence in the first year after birth. *Journal of Affective Disorders* 2004;**78**(3):249-52.

**Wisner 2013**

Wisner KL, Sit DK, McShea MC, Rizzo DM, Zoretich RA, Hughes CL, et al. Onset timing, thoughts of self-harm, and diagnoses in postpartum women with screen-positive depression findings. *JAMA Psychiatry* 2013;**70**(5):490-8.

**Xia 2009**

Xia J, Adams C, Bhagat N, Bhagat V, Bhoopathi P, El-Sayeh H, et al. Losing participants before the trial ends erodes credibility of findings. *Psychiatric Bulletin* 2009;**33**(7):254-7.

**Zimmerman 2013**

Zimmerman M, Martinez JH, Young D, Chelminski I, Dalrymple K. Severity classification on the Hamilton depression rating scale. *Journal of Affective Disorders* 2013;**150**:384-388.

## References to other published versions of this review

### Hoffbrand 2001

Hoffbrand SE, Howard L, Crawley H. Antidepressant treatment for post-natal depression. *Cochrane Database of Systematic Reviews* 2001, Issue 1. [DOI: [10.1002/14651858.CD002018](https://doi.org/10.1002/14651858.CD002018)]

\* Indicates the major publication for the study

## CHARACTERISTICS OF STUDIES

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

#### Appleby 1997

Methods	<p>Randomisation method: computer-generated random numbers</p> <p>Analysis by ITT: yes (LOCF), in addition to analysis by completion</p> <p>Power calculation: none stated</p>
Participants	<p>Setting: community-based: women on maternity wards were asked to allow assessment of their mood in their homes 6-8 weeks later</p> <p>Country: UK</p> <p>Inclusion criteria: women who scored <math>\geq 10</math> on the EPDS at the screening visit were assessed with the CIS-R and eligible to participate if they scored <math>\geq 12</math>, as well as satisfying research diagnostic criteria for major or minor depressive disorder</p> <p>Exclusion criteria: chronic (<math>&gt; 2</math> years) or resistant depression, current drug or alcohol misuse, severe illness requiring close monitoring or hospital admission, breastfeeding</p> <p>Number recruited: 87</p> <p>Number dropped out: 26</p> <p>Number analysed: 87 (additional completers analysis with 61 participants)</p> <p>Age (mean): fluoxetine + 1 counselling session 25.7 years; fluoxetine + 6 counselling sessions 26.6 years; placebo + 1 counselling session 23.1 years; placebo + 6 counselling sessions 26.0 years</p> <p>Ethnicity: no details</p> <p>Socioeconomic status: no details</p>
Interventions	<p>Women were randomly assigned to 1 of 4 groups:</p> <ul style="list-style-type: none"> <li>Fluoxetine + 1 session of counselling (22 women)</li> <li>Fluoxetine + 6 sessions of counselling (21 women)</li> <li>Placebo + 1 session of counselling (23 women)</li> <li>Placebo + 6 sessions of counselling (21 women)</li> </ul> <p>Counselling was derived from CBT and structured to offer reassurance and practical advice on areas of concern to depressed mothers</p>
Outcomes	<p>Assessments were carried out at week 1, 4 and 12</p> <p>Outcome was effect on depressive symptoms as measured by mean scores on the CIS-R, the HAM-D (week 1 and 12 only) and the EPDS</p>
Notes	<p>Funding source not given.</p> <p>See footnote for abbreviations and description of outcome measures.</p>

**Appleby 1997** (Continued)

**Risk of bias**

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Subjects were allocated to one of four treatment groups by using computer generated random numbers"
Allocation concealment (selection bias)	Unclear risk	No details on allocation concealment given
Blinding (performance bias and detection bias) of participants	Low risk	"The counselling was delivered by a psychologist... supervised by a second psychiatrist, both were blind to drug treatment, as were trial subjects"
Blinding (performance bias and detection bias) of personnel	Low risk	"The counselling was delivered by a psychologist... supervised by a second psychiatrist, both were blind to drug treatment, as were trial subjects"
Blinding (performance bias and detection bias) of outcome assessors	Low risk	"The assessment interviews were conducted by a psychiatrist blind to subject treatment group"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	<p>"Drop-out rates were similar in the four groups. Drop outs were younger than subjects who completed the study and more likely to have an unemployed partner and to have a planned pregnancy, but the groups did not differ on initial psychiatric morbidity scores, employment, obstetric complications, parity, family history, or personal history of depression, including postnatal depression"</p> <p>Of 87 total participants, 14/43 from the fluoxetine plus counselling group dropped out and 12/44 of the placebo plus counselling group dropped out</p> <p>Details of dropout timings and reasons were reported, but mainly "no reason given". Lack of improvement was the reason for 3 drop-outs in the fluoxetine group but 0 in the placebo group. In contrast, 3 women in the placebo group but only 1 woman in the intervention group dropped out due to side effects</p>
Selective reporting (reporting bias)	Unclear risk	Protocol unavailable
Other bias	Unclear risk	<p>No details given on adherence to medication</p> <p>See footnote for abbreviations and description of outcome measures</p>

**Bloch 2012**

Methods	Randomisation method: pharmacy-generated random serial numbers  Analysis by ITT: yes (LOCF)  Power calculation: yes
Participants	Setting: maternity ward and baby care centre  Country: Israel

**Bloch 2012** (Continued)

Inclusion criteria: aged 18-45 years; met criteria for current MDD during screening and baseline visit according to DSM-IV (Structured Clinical Interview for DSM-IV Axis I disorders), onset of depression within 2 months of delivery

Exclusion criteria: MADRS score  $\geq 30$ , suicidal ideation (MADRS item 10 score  $\geq 5$ ), psychotic symptoms or bipolar disorder, current depressive episode  $> 6$  months, current treatment with antidepressants, 2 failed adequate trials of antidepressants, major physical illness, alcoholism or drug use

Number recruited: 42

Number dropped after baseline assessment: 2 (both from placebo + BDP group)

Number dropped out by week 8: 4 from sertraline + BDP group; 3 from placebo + BDP group (not including the 2 dropped out after baseline assessment)

Number analysed: 40 (2 participants who dropped out after baseline excluded)

Age: no data

Ethnicity: no data

Socioeconomic status: sertraline + BDP group: high income: 7/20 (35%), middle income: 10/20 (50%), low income: 3/20 (15%); placebo + BD group: high income: 4/20 (20%), middle income: 7/20 (35%), low income 9/20 (45%)

**Interventions**

Women were randomly assigned to 1 of 2 groups:

- Sertraline + BDP (20 women): sertraline dosage: week 1 25 mg once daily, week 2 50 mg once daily, week 4 increase to 100 mg if  $< 20\%$  improvement in MADRS or no improvement in CGI. Blinded Psychiatrist decision on whether to increase dose
- Placebo + BDP (22 women). Dummy pills identical to sertraline were delivered to women according to the same protocol as the sertraline group along with BDP

BDP is a time-limited psychotherapeutic intervention that aims to enhance the patient's insights about repetitive circumstances.

**Outcomes**

Outcome measures carried out at weeks 0, 2, 4, 6, 8, 12

Primary outcome: continuous change in depressive symptoms as measured by the MADRS and EPDS during 8-week randomisation phase

Secondary outcomes: continuous change in MADRS and EPDS during open phase of the study (weeks 8-12), proportion of women meeting response and remission status at week 8 (response defined as  $> 50\%$  reduction in MADRS or EPDS scores during treatment and remission as a final score of  $< 10$  on the MADRS or  $< 7$  on the EPDS)

Other secondary ratings: measurements of symptom severity using the Clinical Global Impression scale (CGI-I, CGI-S), assessment of global mental health with the MHI and assessment of adverse effects using the UKU Side Effect Rating Scale

**Notes**

This study was funded by an Independent Investigator Award to Dr Bloch from the National Alliance for Research on Schizophrenia and Depression, Great Neck, New York.

See footnote for abbreviations and description of outcome measures.

**Risk of bias**

**Bias**

**Authors' judgement**

**Support for judgement**

Random sequence generation (selection bias)

Low risk

"The institution's pharmacy-generated random patient serial numbers with active versus placebo ratio 1:1 were issued to the researchers and randomly

**Bloch 2012** (Continued)

		assigned to eligible patients by the psychiatrist after the informed consent was signed".
Allocation concealment (selection bias)	Unclear risk	Insufficient information given to be certain of allocation concealment
Blinding (performance bias and detection bias) of participants	Low risk	"The second group received dummy pills daily, identical in appearance to the active pills, according to the same protocol as the active group"
Blinding (performance bias and detection bias) of personnel	Unclear risk	"The managing psychiatrist was blinded to treatment condition... The managing psychiatrist was also asked at the end of the full protocol to document her assessment of whether the patient received active or placebo pills, and indeed, was unable to correctly guess this factor in every instance"
Blinding (performance bias and detection bias) of outcome assessors	Unclear risk	No details given on who assessed outcomes so unclear whether outcome assessors were blinded
Incomplete outcome data (attrition bias) All outcomes	Low risk	<p>"Seven patients discontinued medication between weeks 4 and 8, three from the placebo group and four from the active group. Discontinuation was due to lack of motivation (n=4: placebo group, n=2; sertraline group, n=2) and clinical deterioration (n=3: placebo group, n=1; sertraline group, n=2)."</p> <p>42 participants were originally in the study, 2 participants dropped out of the placebo group immediately after the baseline. 40 participants are included in the ITT</p>
Selective reporting (reporting bias)	Unclear risk	Insufficient detail in protocol
Other bias	Unclear risk	<p>"A pill count was conducted to monitor compliance. Protocol violation was defined as &lt;80% compliance by pill count"</p> <p>It is unclear whether the 7 patients who discontinued medication are the only participants with low compliance</p> <p>"The compliance for psychotherapy was good: in the sertraline group, 92% of the psychotherapy sessions were attended compared to 87% in the placebo group (P=NS) [not significant]"</p>

**Hantsoo 2013**

Methods	Randomisation method: unclear  Analysis by ITT: yes (LOCF for response and remission analyses) and by evaluable group  Power calculation: yes
Participants	Setting: mixed setting - recruitment via local obstetrician-gynaecologists, paediatricians, mental health professionals, postnatal depression support groups, and advertisements in local newspapers  Country: USA  Inclusion criteria: aged 18-45 years, depression onset reported within 3 months after delivery, no psychotropic medication for 5 or more weeks, and given birth within the last 12 months to an infant without serious medical issues. Participants were required to have a diagnosis of postnatal depression based on the SCID, to score $\geq 18$ and $< 32$ on the 19-item HAM-D and to have at least "moderate" symptoms on the severity of illness rating of the CGI scale. Only English speaking women were eligible.

**Antidepressant treatment for postnatal depression (Review)**

**Hantsoo 2013** (Continued)

Exclusion criteria: onset of MDD during pregnancy (indicated on the SCID), screened positive for thyroid disease (unless thyroid condition stable), drug or alcohol dependence in the last 6 months or positive urine drug test during screening, current or history of psychotic disorder (Axis I, including bipolar type I), active suicidal ideation, any significant medical conditions, planning to become pregnant or past failed trial of sertraline.

Number recruited: 38 (36 randomised after the placebo run-in week: 2 participants had > 30% decline in HAM-D scores during the run-in week and were removed from the study as per protocol)

Number dropped out: 7 dropped out by week 7 (final week)

Number analysed: 36 analysed on an ITT basis. Repeated analyses with evaluable group had at least 3 post-randomisation assessments (33 women)

Age (mean ± SD): 30.8 ± 4.0 years, with no between-group differences; sertraline: 29.6 ± 4.0 years; placebo: 31.7 ± 3.7 years

Ethnicity: sertraline group: 16 Caucasian, 1 Hispanic; placebo group: 18 Caucasian, 1 Hispanic

Years of education (mean ± SD): sertraline group: 14.4 ± 2.0 years; placebo group: 14.0 ± 1.2 years

**Interventions**

All participants underwent a 1-week single-blind placebo lead-in. Participants who still met the inclusion criteria and had had a less than 30% reduction in HAM-D scores were randomly assigned to 1 of 2 groups:

- Sertraline: treatment commenced with 50 mg daily. Dosage was then increased as tolerated by 1 capsule (50 mg) every 1-2 weeks until clinical remission was obtained or up to a maximum of 4 capsules (200 mg) per day. The mean daily dose (± SD) at week 7 was 100.0 ± 54.0 mg
- Placebo: dosage followed the pattern described above. The mean dose for the placebo group at week 7 was 119.4 ± 51.8 mg

**Outcomes**

Primary outcomes: response in psychiatric symptoms: treatment response was defined as a score of ≤ 10 on the HAM-D, at least a 50% decrease in HAM-D score from baseline, and a score of "much improved" or "very much improved" on the improvement scale of the CGI (after 6 weeks of treatment); remission defined as per criteria for response to treatment in addition to a HAM-D score ≤ 7 (after 6 weeks of treatment)

Secondary outcomes: trends over time in depressive symptoms as rated by the HAM-D and the EPDS, and in anxiety symptoms as rated by the HAM-A. The predominant interest was the treatment group by linear time interaction

**Notes**

This study was funded by Pfizer (New York, NY, USA), the National Institute of Mental Health (P50 MH099910 and K23 MH01830) and the National Institute of Drug Abuse (K24 DA03031).

See footnote for abbreviations and description of outcome measures.

**Risk of bias**

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	"After the lead-in, all the subjects.. were randomised to a 6-week, double-blind trial"
Allocation concealment (selection bias)	Unclear risk	No details given on allocation concealment
Blinding (performance bias and detection bias) of participants	Low risk	1 participant was excluded after the study began due to accidental unblinding

**Hantsoo 2013** (Continued)

Blinding (performance bias and detection bias) of personnel	Low risk	"A research pharmacist was responsible for creating a blinding table and distributing the study drug; all other study personnel remained blind to subject treatment status"
Blinding (performance bias and detection bias) of outcome assessors	Low risk	"All other study personnel remained blind to subject treatment status"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	"A total of 17 women were randomised to the sertraline group and 19 were randomised to placebo, for a total of 36 women in the intent-to-treat group. The reasons for failure to the full 7 weeks included clinical deteriorating (n=3, all in the placebo group), loss to follow-up (n=3), and accidental unblinding (n=1)"  There could have been an underestimation of treatment effect as women dropping out for clinical deterioration were all in the control group
Selective reporting (reporting bias)	Unclear risk	Protocol unavailable
Other bias	Unclear risk	No details given on adherence to medication

**Sharp 2010**

Methods	Randomisation method: web-based randomisation programme  Analysis by ITT: ITT, multiple imputation and complete-case analysis all employed
Participants	Setting: community-based: recruitment was from 77 general practices based in Bristol, South London, and Manchester  Country: UK  Inclusion criteria: women aged $\geq 18$ years who had a recent live birth and were living with their baby were eligible for screening phase. After screening, deemed to be eligible if: score of $\geq 13$ on baseline EPDS, ICD-10 primary diagnosis of depression on the CIS-R, proficient in English at a level to complete all research assessments and recently delivered baby was $< 26$ weeks old  Exclusion criteria: stillbirth or neonatal death, baby $> 26$ weeks old, baby fostered or adopted. Women were also not eligible if they had psychosis, alcohol or drug abuse, were already receiving treatment for depression or were actively suicidal  Number recruited: 254  Number dropped out by week 4: antidepressants: 23/129, treatment as usual: 13/125  Number dropped out by week 18: antidepressants: 32/129, listening visits: 16/125  Number analysed: 218 primary analysis on an ITT basis at 4 weeks, 206 primary analysis on an ITT basis at 18 weeks, also analysed as all 254 randomised.  Age (mean $\pm$ SD): 29.3 $\pm$ 6.3 years  Ethnicity: white 196 (77.8%), black 29 (11.5%), Asian 13 (5.2%), other 14 (5.6%)  Socioeconomic status: highest educational qualification: none: 36 (14.8%), GCSE (school exams taken at 16) 67 (27.5%), A level (school exams taken at 18) 32 (13.1%), NVQ (National Vocational Qualification) 48 (19.7%), degree 61 (25.0%)



**Sharp 2010** (Continued)

Interventions	<p>Women were randomly assigned to 1 of 2 groups:</p> <ul style="list-style-type: none"> <li>• Antidepressants (129 women). SSRI recommended as a first-line treatment; however, a pragmatic approach whereby the GP and the woman agreed which antidepressant medication should be prescribed was employed. Most women were prescribed citalopram (68 women), fluoxetine (49 women) or sertraline (22 women). Other antidepressants prescribed were amitriptyline (4 women), cipramil (1 woman), clomipramine (1 woman), dosulepin (5 women), escitalopram (6 women), imipramine (1 woman), iofepamine (1 woman), mirtazapine (4 women), paroxetine (7 women), prothiaden (1 woman) and venlafaxine (2 women). Trial design allowed women to receive the alternative intervention at any time after four weeks. 68 women in the antidepressant arm requested listening visits after the 4-week follow-up. Of these, 64 had at least one visit. Adherence to treatment: at 4 weeks 56% of the women randomised to antidepressants reported taking any antidepressants (59/106, only calculated for those followed-up).</li> <li>• Treatment as usual and listening visits (125 women). Listening visits commenced about 4 weeks after randomisation to mimic waiting list times (4 weeks of treatment as usual). Listening visits were delivered in a series of up to 8 sessions by trained research health visitors. Women allocated listening visits were able to visit their GP for antidepressants at any time during the study, but GPs could not prescribe antidepressants until 4 weeks unless absolutely necessary</li> </ul>
Outcomes	<p>Timing of each outcome and relevant domain</p> <p>All assessments carried out at week 0, 4 and 18</p> <p>Primary outcome: assessment of remission of postnatal depression using EPDS &lt; 13 at follow-up</p> <p>Secondary outcomes: change in depressive symptoms (EPDS) as continuous variable, physical and mental health assessment (SF-12), assessment of maternal functioning (MAMA), health-related quality of life (EQ-5D), quality of marital relationships (GRIMS). If women had a male partner he was asked to complete the following: assessment of relationship with partner (GRIMS), assessment of paternal functioning (PAPA), general health assessment (GHQ and SF-12)</p>
Notes	<p>This study was funded by the National Institute for Health Research Health Technology Assessment programme.</p> <p>See footnote for abbreviations and description of outcome measures.</p>

**Risk of bias**

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	<p>"Before the baseline home visit, the women's trial identification number, date of birth and trial centre were entered into a web-based randomisation program"</p> <p>"The randomisation sequences was generated using a computer program with block sizes of six, eight and ten, varied randomly"</p>
Allocation concealment (selection bias)	Low risk	<p>"After eligibility had been determined and consent had been obtained at the home visit, the researcher telephoned the remote computerised randomisation service and responded to a series of questions by keying numbers (e.g. patient identification number, baseline EPDS score) of the telephone keypad"</p> <p>"The methods of sequence generation were concealed from the researchers involved in enrolling and randomising the women into the trial"</p>
Blinding (performance bias and detection bias) of participants	High risk	<p>"Participants, researchers and those delivering the interventions were not blinded to the treatment allocation"</p>

**Sharp 2010** (Continued)

Blinding (performance bias and detection bias) of personnel	High risk	"Participants, researchers and those delivering the interventions were not blinded to the treatment allocation"
Blinding (performance bias and detection bias) of outcome assessors	High risk	"Participants, researchers and those delivering the interventions were not blinded to the treatment allocation"
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	<p>"More women in the antidepressant group withdrew or were lost to follow-up [4 weeks: antidepressants 23 (18%), listening visits 13 (10%) p = 0.090; 18 weeks: antidepressants 32 (25%), listening visits 16 (13%) p = 0.015]"</p> <p>Reasons for drop-out are not given and characteristics of drop-outs are not given separately by intervention group</p> <p>Sensitivity analyses (including multiple imputation) were performed to examine the impact of missing data. The imputation of missing data had no material effect on the results</p>
Selective reporting (reporting bias)	High risk	All primary outcomes reported. Some evidence of selective outcome reporting from the protocol where the HOME measure and Bayley Scale of Infant Development are pre-specified but not reported in the main paper. Paternal measures are also detailed in the protocol; these are detailed in the methods of the main paper and it is stated that they will be discussed in a separate report, but this could not be identified
Other bias	High risk	"At 4 weeks only 59 (56%) of the 106 women followed up among those randomised to the antidepressants and who completed the [adherence] questionnaire reported taking any antidepressants. In the listening visits groups seven (6%) of the 112 women followed up also reported taking antidepressants... At the 18-week time point, the numbers in each group who reported taking antidepressants during the previous 4 weeks were 62 (64% of the 97 followed up) and 37 (34% of 109 followed up) in the antidepressants and listening visits groups, respectively"

**Wisner 2006**

Methods	Randomisation method: block randomisation with a sequence generated in SPSS  Analysis by ITT: yes for primary outcomes (response and remission), LOCF  Power calculation: yes
Participants	Setting: no details  Country: USA  Inclusion criteria: women aged 15-45 years with major depression within 4 weeks of birth. Women with chronic depression (an episode on major depression beginning before the index pregnancy) were also included after additional funding was obtained part-way through the trial. Mothers had to present for treatment within 3 months of delivery and score $\geq 18$ on the HAM-D  Exclusion criteria: presence of any other Axis I disorder except generalised anxiety disorder or panic disorder, contraindications to TCA treatment, and concurrent psychiatric treatment  Number recruited: 109  Number dropped out: 23 from the sertraline group (42%), 13 from the nortriptyline group (24%)

**Wisner 2006** (Continued)

Number analysed: ITT and analyses presented for 95 women who took the assigned medication for at least 1 week and provided at least 1 week of follow-up data and 83 women who provided at least 3 weeks of follow-up data

Age: no data

Ethnicity: significantly more non-white women were randomly assigned to sertraline (40%) than nortriptyline (19%) (Fisher exact test; P value = 0.02). There were no other demographic differences between the 2 drug groups at baseline and no other details on ethnicity of socio-economic status were given

Interventions	<p>The aim was to compare the effect on postnatal depression symptoms of treatment with sertraline compared with nortriptyline</p> <p>Women were randomly assigned to 1 of 2 groups:</p> <ul style="list-style-type: none"> <li>• Sertraline: the dosing began with 25 mg/day for 2 days. Thereafter, the doses was increased to 50 mg/day and further increased until either response or side effects prohibited further dose escalation. The maximum dose was 200 mg/day</li> <li>• Nortriptyline: initial dose of 10 mg/day. This was then increased to 25 mg/day and then further increased until either response or side effects prohibited further dose escalation. Maximum dose was 150 mg/day</li> </ul>
Outcomes	<p>Followed up at weekly intervals for weeks 1-8, then again at week 24</p> <p>Primary outcomes: response to treatment at 8 weeks (50% reduction in HAM-D from baseline); remission of depression (HAM-D &lt; 7 at week 8); continuous change in HAM-D; severity of symptoms of depression (CGI scale at week 8); overall functioning as measured by the GAS; issues in income, housing, relationships and work (SPQ)</p> <p>Secondary outcome: side effects on the Asberg Side Effects Rating Scale in addition to time to withdrawal due to side effects, obsessions and compulsions measured with the YBOCS, emergence of mania was screening for safety reasons using the Mania Rating Scale (derived from the Schedule for Affective Disorders and Schizophrenia)</p>
Notes	<p>This study was funded by the National Institute of Mental Health.</p> <p>See footnote for abbreviations and description of outcome measures.</p>

**Risk of bias**

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Subjects were randomised 1:1 to either nortriptyline or sertraline in block of 8 to 12 with a sequence generated by SPSS"
Allocation concealment (selection bias)	Unclear risk	No details given on allocation concealment
Blinding (performance bias and detection bias) of participants	Low risk	"Prescriptions were assembled by the research pharmacist. The nortriptyline and sertraline were delivered in 2 doses, with breakfast and at bedtime. The opaque, inert gelatine capsules contained either sertraline (AM)/placebo(HS) or placebo(AM)/nortriptyline (HS)"
Blinding (performance bias and detection bias) of personnel	Low risk	"The primary staff (side effects monitor, mood symptom rater, and study psychiatrist) were blind to drug assignment until project completion. The medication monitoring function (nurse) was separate from (and blind to) the mood monitoring (interviewer)"
Blinding (performance bias and detection bias)	Low risk	"The primary staff (side effects monitor, mood symptom rater, and study psychiatrist) were blind to drug assignment until project completion"

**Antidepressant treatment for postnatal depression (Review)**

**Wisner 2006** (Continued)  
 of outcome assessors

Incomplete outcome data (attrition bias) All outcomes	High risk	"Significantly more women who took sertraline compared with nortriptyline withdrew from the study in the first 8 weeks (23/55 [42%] versus 13/54 [24%], respectively [P = 0.02]). The proportion of women who were lost to follow-up or withdrew by personal choice differed significantly (sertraline, 20%, vs. nortriptyline, 6%; Wilcoxon $\chi^2_1 = 4.86$ ; P = 0.03). Other reasons for withdrawal (side effects, hypomania occurrence, or clinical deterioration) did not differ between the 2 drug groups"  It is unclear why the difference in withdrawal between study groups was so high - but likely to cause bias in results
Selective reporting (reporting bias)	Unclear risk	Protocol unavailable
Other bias	Low risk	"Fourteen women had minimal drug in their blood despite claims of compliance. The results remained the same when data from these 14 women were removed. Drug assignment in the 14 women was distributed similarly between nortriptyline (n = 9/51, 18%) and sertraline (n = 5/44, 11%; Fisher exact test, P = 0.29)"

**Yonkers 2008**

Methods	<p>Randomisation method: pre-determined with a computer-generated schedule in blocked sets of 4 and was stratified by site</p> <p>Analysis by ITT: yes (LOCF for response and remission analyses)</p>
Participants	<p>Setting: community/secondary care. Women were recruited by advertisement or referral from obstetric care providers</p> <p>Country: USA</p> <p>Inclusion criteria: aged <math>\geq 16</math> years, met diagnostic criteria for MDD with an onset in the 3 months post-delivery, had given birth within the previous 9 months and had a score on the 17-item HAM-D of at least 16 at the initial visit. Women who were breastfeeding were allowed to participate</p> <p>Exclusion criteria: onset of MDD prior to delivery, current suicidal ideation with intent, current (within the last 6 months) alcohol or drug abuse or dependence, current psychotic symptoms, lifetime diagnosis of schizophrenia, bipolar disorder or schizoaffective disorder, currently receiving treatment (pharmacotherapy or psychotherapy) for a psychiatric disorder, currently pregnant, unwilling to be randomised or unable to attend treatment visits at a participating site</p> <p>Number recruited: 70 women (35 active treatment, 35 placebo)</p> <p>Number dropped out by final week (week 8 <math>\pm</math> 7 days): paroxetine group: 20/35 (57%); placebo group: 23/35 (66%)</p> <p>Number analysed: ITT analysis and evaluation at week 8 for results from 17 women in paroxetine group and 14 women in the placebo group</p> <p>Age (mean <math>\pm</math> SD): paroxetine: mean 26.1 <math>\pm</math> 6.5; placebo: 25.9 <math>\pm</math> 6.5</p> <p>Ethnicity: paroxetine: white: 18 (51.4%), black: 5 (14.3%), Hispanic: 11 (31.4%), other 1 (2.9%); placebo: white: 16 (45.7%), black: 4 (11.4%), Hispanic: 14 (40.0%), other 1 (2.9%)</p> <p>Socioeconomic status: paroxetine: &lt; 12 years of education: 11 (37.9%), &gt; 12 years of education: 18 (62.1%); placebo: &lt; 12 years of education: 15 (53.6%), &gt; 12 years of education: 13 (46.4%)</p>

**Yonkers 2008** (Continued)

Interventions	<p>Women were randomly assigned to 1 of 2 groups:</p> <ul style="list-style-type: none"> <li>• Paroxetine: week 1 and 2: 1 capsule (10 mg) of immediate release paroxetine daily; week 3 and 4: 2 capsules (20 mg) of immediate release paroxetine daily unless side effects limited an increase. Further increments to 30 mg by week 4 and then 40 mg by week 6 were encouraged if improvement was assessed as &lt; 30% compared with baseline</li> <li>• Placebo: identical placebo administered according to same protocol as paroxetine</li> </ul>
Outcomes	<p>All primary outcomes listed were assessed at weeks 1, 2, 3, 4, 6 and for a final visit, at week 8 (<math>\pm</math> 7 days)</p> <p>Primary outcome: change in depressive symptoms measured by the HAM-D, CGI and the Inventory of Depressive Symptomatology - Self-report scale</p> <p>Secondary outcomes: rates of remission, defined as a HAM-D score of <math>\leq</math> 8, and response, defined as a CGI-Improvement scale score of 1 or 2; predictors of remission defined as above; Social Adjustment as measured by the SAS; SF-36</p>
Notes	<p>This study was supported by a Collaborative Research Trial, Investigator-Initiated grant from GlaxoSmithKline to Drs Yonkers and Cohen and by National Institute of Mental Health grant MH01648 to Dr Yonkers.</p> <p>See footnote for abbreviations and description of outcome measures.</p>

**Risk of bias**

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Subjects were randomly assigned to take identical capsules of either paroxetine or placebo. Random assignment was predetermined with a computer-generated schedule in blocked sets of 4 and was stratified by site. A study statistician was responsible for random assignment"
Allocation concealment (selection bias)	Unclear risk	Insufficient details provided to be sure of allocation concealment
Blinding (performance bias and detection bias) of participants	Low risk	"Subjects were instructed to take 1 capsule (10mg of immediate-release paroxetine or identical placebo)"
Blinding (performance bias and detection bias) of personnel	Low risk	"A study statistician was responsible for random assignment, and remaining study staff were blind to group assignment."
Blinding (performance bias and detection bias) of outcome assessors	Low risk	"..remaining study staff were blind to group assignment"
Incomplete outcome data (attrition bias) All outcomes	High risk	<p>"Seventy women qualified for the study, and 31 completed study treatment... Subjects withdrew from the active treatment for the following reasons: 1 due to an adverse event (nausea), 6 due to lack of efficacy, including 1 subject who was psychiatrically hospitalised, 6 who were lost to follow-up, 5 who felt well and no longer desired treatment, 1 who became pregnant and 1 who was non-compliant</p> <p>In subjects randomly assigned to placebo, 4 left the study because of perceived adverse events (rash, nausea, diarrhoea, headache), 7 discontinued because of lack of efficacy, including 1 subject who required hospitalisation, 9 were lost to follow-up, 2 improved and no longer desired treatment, and 1 subject moved"</p>

**Yonkers 2008** (Continued)

"Given the high rate of dropout, we explored additional models to assess the robustness of remission results. These models first assumed that all dropouts were remitters and then that they were all nonremitters. In both models, treatment with paroxetine remained significantly better than treatment with placebo"

Drop out numbers are similar in the 2 groups and some reasons account for similar numbers across the 2 groups but for a substantial proportion "lost to follow up" the reason for drop-out is unknown. Sensitivity analyses only performed for the primary outcome

Selective reporting (reporting bias)	High risk	The Social Adjustment Scale and SF-36 were included in the methods but not reported in the results
Other bias	Unclear risk	<p>"Pill counts revealed that, among women assigned to paroxetine, 7 were non-compliant (took less than 80% of prescribed pills at 1 visit, and 4 were non-compliant at 2 visits. One subject assigned to active treatment was discontinued due to on-going lack of compliance; of the remaining subject, no others fell below the 80% compliance rate at more than 2 visits. Among subjects assigned to placebo, 10 were noncompliant at 1 visit, 3 were noncompliant during at least 2 visits, and 1 was noncompliant on 4 occasions"</p> <p>The potential bias was unclear as we do not know whether non-compliant women were taking 0% or 79% of their medication. It is also not clear whether the numbers of non-compliant participants were reported for the study as a whole (26/70 women) or only for those who did not drop out (26/31 women)</p>

**Abbreviations:** BPD: brief dynamic psychotherapy; CBT: cognitive behavioural therapy; CGI: Clinical Global Improvement; CIS-R: Revised Clinical Interview Schedule; DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; EPDS: Edinburgh Postnatal Depression Scale; GAS: Global Assessment Scale; GHQ: General Health Questionnaire; GRIMS: Golombok Rust Inventory of Marital State; HAM-A: Hamilton Rating Scale for Anxiety; HAM-D: Hamilton Rating Scale for Depression; ICD-10: International Classification of Disease Tenth Revision; ITT: intention to treat; LOCF: last observation carried forward; MADRS: Montgomery-Åsberg Depression Rating Scale; MAMA: Maternal Adjustment and Maternal Attitudes; MDD: major depressive disorder; MIH: Mental Health Index; PAPA: Preschool Age Psychiatric Assessment; SAS: Social Adjustment Scale; SCID: Structured Clinical Interview for DSM-IV; SD: standard deviation; SF-12: 12-item Short Form; SF-36: 36-item Short Form; SPQ: Social Problems Questionnaire; YBOCS: Yale-Brown Obsessive Compulsive Scale.

**CIS-R** is a structured diagnostic interview schedule for the diagnosis of common mental disorders. The CIS-R is widely used in population and primary care surveys to provide estimates of depression.

**CGI-Improvement Scale** is a clinician-rated scale that assesses changes in symptoms. The scales are rated on a scale of 1 = very much improved; 2 = much improved; 3 = minimally improved; 4 = no change; 5 = minimally worse; 6 = much worse or 7 = very much worse. Each component of the CGI is rated separately and the scales do not yield a global score.

**CGI-Severity of Illness** measure is a clinician-rated scale that assess the severity of symptoms. The CGI-Severity of Illness is rated on a scale of 1 = not at all ill; 2 = borderline mentally ill; 3 = mildly ill; 4 = moderately ill; 5 = markedly ill; 6 = severely ill or 7 = extremely ill. The CGI-Improvement scale is a clinician-rated scale that assesses changes in symptoms. The scales are rated on a scale of 1 = very much improved; 2 = much improved; 3 = minimally improved; 4 = no change; 5 = minimally worse; 6 = much worse or 7 = very much worse. Each component of the CGI is rated separately and the scales do not yield a global score.

**EPDS** is a 10-item self administered screen for perinatal depression, validated in 20 languages. For each item, women are asked to select 1 of 4 responses that most closely describe how they have felt over the past 7 days. Each response has a value of 0-3; scores for the 10 items are summed to give a total score between 0 and 30. The EPDS is the most widely used screening instrument for postpartum depression and has a positive predictive value for postnatal major depression of 9-64% (with a cut-off score of 9/10) or 17-100% (with a cut-off of 12/13). A cut-off score of 12/13 is used in most studies to indicate postpartum depression. The EPDS does not discriminate levels of depression and additional information is required to meet diagnostic criteria for depression.

**EQ-5D** is a preference-based measure of health-related quality of life measured on 5 dimensions (i.e. mobility, self care, usual activities, pain/discomfort and anxiety/depression), each rated on 3 levels (i.e. no problems, some problems and severe problems). Participants are classified into 1 of 243 health states, each associated with a score that can be used to calculate quality-adjusted life years. The measure has been extensively used in health economic evaluations and its psychometric properties are adequate.

**GAS** is a rating scale for evaluating the overall functioning of a person during a specified time period on a continuum from psychological or psychiatric sickness to health.

**GRIMS** is a 28-item self complete questionnaire that assesses the quality of the relationship between a married or co-habiting couple.

**HAM-A** is a clinician-rated screening instrument that assesses the presence and severity of anxiety. Total scores are obtained by summing the score of each item, 0-4 (symptom is absent, mild, moderate or severe). For the 14-item HAM-A version total scores range from 0 to 56.

A score of 0-13 is indicative of no anxiety; 14-17 is indicative of mild anxiety; 18-24 is indicative of moderate anxiety and 25-30 is indicative of severe anxiety.

**HAM-D** is a clinician rated screening instrument that assesses the presence and severity of depression. Total scores are obtained by summing the score of each item, 0-4 (symptom is absent, mild, moderate or severe) or 0-2 (absent, slight or trivial, or clearly present). For the 17-item HAM-D version, total scores range from 0 to 54. A score of 0-6 is indicative of no depression, 7-17 is indicative of mild depression, 18-24 is indicative of moderate depression and  $\geq 25$  is indicative of severe depression. For most raters, a total score of  $\leq 7$  after treatment is a typical indicator of remission and a decrease of 50% or more from baseline is considered an indicator of a clinically significant change.

**MADRS** is a diagnostic instrument that measures the severity of depressive episodes. Each response has a value of 0-6; scores for the 10 items are summed to give a total score between 0 and 60. A score of 0-6 is indicative of no depression, 7-19 is indicative of mild depression; 20-34 is indicative of moderate depression and  $\geq 35$  is indicative of severe depression.

**MAMA** is a self administered questionnaire that examines perceptions of maternal adjustment and attitudes towards marital relationships and the baby. The postnatal sub-scale of the MAMA questionnaire comprises 12 items rated on a 4-point scale from 1 = "not at all" to 4 = "very much".

**SF-12** is a 12-item self-complete questionnaire that measures functional health and well-being. The measure is a widely used and well-validated generic measure of functional quality of life.

**SPQ** is a 33-item self report questionnaire that covers 10 areas or domains, including housing conditions; occupation; financial status; social and leisure activities; contacts with relatives, friends and neighbours; family functioning; child-parent interaction; relationship with spouse or partner and legal matters. The individual items are rated on a 4-point scale ranging from 0 (no social difficulties/satisfactory adjustment) to 3 (severe social difficulties/very poor adjustment).

### Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
<a href="#">Bennett 2001</a>	No antidepressant treatment
<a href="#">Misri 2004</a>	Not comparing antidepressants with another intervention - both arms had same antidepressant (paroxetine vs. paroxetine + cognitive behavioural therapy)
<a href="#">Rojas 2007</a>	No consistent randomised comparison of antidepressants to another intervention (multicomponent intervention vs. usual care)
<a href="#">Stein 2012</a>	Ineligible study population
<a href="#">Suri 2005</a>	Antidepressant treatment not randomised
<a href="#">Yu 2006</a>	Not comparing antidepressants with another intervention - both arms had same antidepressant (paroxetine vs. paroxetine + psychological intervention)
<a href="#">Zhao 2006</a>	Not comparing antidepressants with another intervention - both arms had same antidepressant (fluoxetine vs. fluoxetine + shugan powder)

### Characteristics of studies awaiting assessment [ordered by study ID]

#### [NCT00744328](#)

Methods	8-week, double-blind, placebo-controlled randomised controlled trial
Participants	85 women
Interventions	Transdermal oestradiol (50-200 µg/day) compared with sertraline (25-200 mg/day) compared with placebo
Outcomes	Primary outcome measures include assessing the efficacy of oestradiol as a treatment for postpartum depression, and efficacy in comparison to placebo and sertraline.

**NCT00744328** (Continued)

Secondary outcome measures will include data on infant development using Bayley Scales of Infant Development, mother-infant serum oestradiol and sertraline levels, quality of mother-infant interactions.

Notes The study (NCT00744328) is led by Professor Katherine Wisner in the USA, data collection has finished (terminated early due to recruitment issues) and analysis is in progress.

Contact information: Emily A. Pinheiro: [emily.pinheiro@northwestern.edu](mailto:emily.pinheiro@northwestern.edu)

**NCT02122393**

Methods	24-week, single-blinded (outcome assessor) randomised controlled trial
Participants	45 women
Interventions	Sertraline compared with cognitive behavioural therapy compared with combined therapy
Outcomes	Beck Depression Inventory (primary outcome), Beck Anxiety Inventory, Parenting Stress Index (secondary outcome)

Notes This study is led by Jeannette Milgrom and Alan W Gemmill in Melbourne, Australia. The clinicaltrials.gov record states that data collection was completed in April 2005 (study retrospectively registered in April 2014) but correspondence with study authors indicated that the report is currently in progress and data are not yet available.

Contact information: [jeannette.milgrom@austin.org.au](mailto:jeannette.milgrom@austin.org.au)

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

**NCT00602355**

Trial name or title	Effectiveness of Sertraline Alone and Interpersonal Psychotherapy Alone in Treating Women with Postpartum Depression
Methods	13-week, double-blind, placebo-controlled randomised controlled trial
Participants	The study is expected to enrol 100 women
Interventions	Sertraline 25-200 mg/day alone compared with interpersonal psychotherapy, administered as 50-minute sessions every week for 13 weeks, compared with placebo.
Outcomes	Primary outcome measure is monitoring of depressive symptoms severity using Hamilton Depression Rating Scale.  Secondary outcomes include: monitoring of depressive symptoms using the Beck Depression Inventory and Edinburgh Postnatal Depression Scale, general illness severity using Clinical Global Impression scale and social functioning assessed with Postpartum Adjustment Questionnaire and anxiety assessed by Hamilton Rating Scale for Anxiety. Follow-up assessments are due to take place 3 and 6 months post intervention.
Starting date	February 2008
Contact information	Jennifer Bowman-Reif: <a href="mailto:jennifer-bowman-reif@uiowa.edu">jennifer-bowman-reif@uiowa.edu</a>
Notes	The study is led by Dr Caron Zlotnick in the USA and is due to be completed in 2014

**Antidepressant treatment for postnatal depression (Review)**

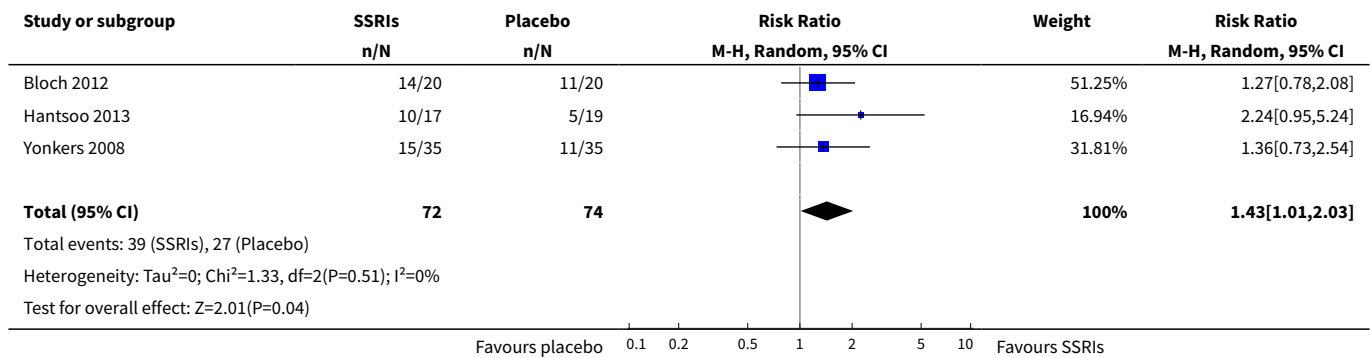


**DATA AND ANALYSES**

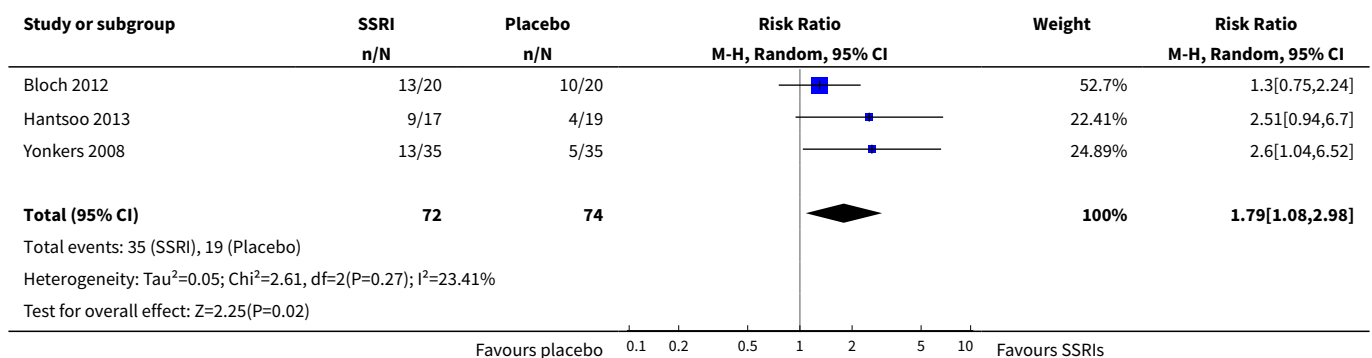
**Comparison 1. Selective serotonin re-uptake inhibitors versus placebo**

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Response rate at post-treatment	3	146	Risk Ratio (M-H, Random, 95% CI)	1.43 [1.01, 2.03]
2 Remission rate at post-treatment	3	146	Risk Ratio (M-H, Random, 95% CI)	1.79 [1.08, 2.98]

**Analysis 1.1. Comparison 1 Selective serotonin re-uptake inhibitors versus placebo, Outcome 1 Response rate at post-treatment.**



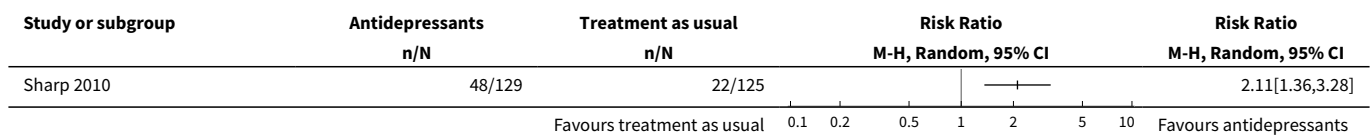
**Analysis 1.2. Comparison 1 Selective serotonin re-uptake inhibitors versus placebo, Outcome 2 Remission rate at post-treatment.**



**Comparison 2. Antidepressants versus treatment as usual**

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Remission rate at post-treatment	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected

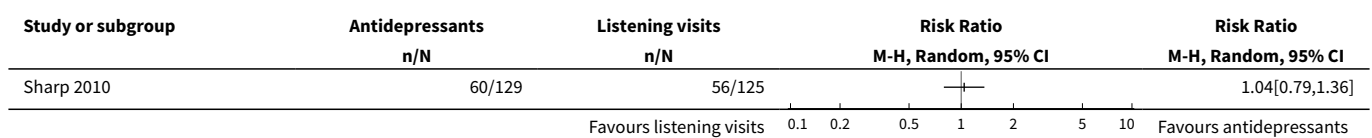
**Analysis 2.1. Comparison 2 Antidepressants versus treatment as usual, Outcome 1 Remission rate at post-treatment.**



**Comparison 3. Antidepressants versus psychosocial therapy (listening visits)**

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Remission rate at post-treatment	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected

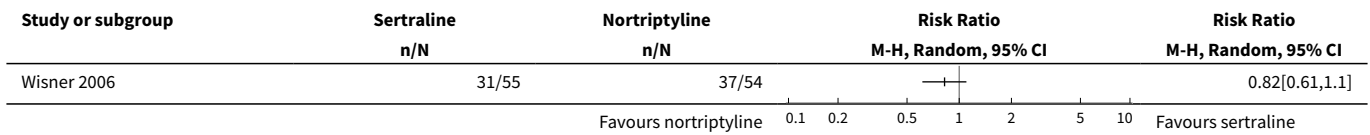
**Analysis 3.1. Comparison 3 Antidepressants versus psychosocial therapy (listening visits), Outcome 1 Remission rate at post-treatment.**



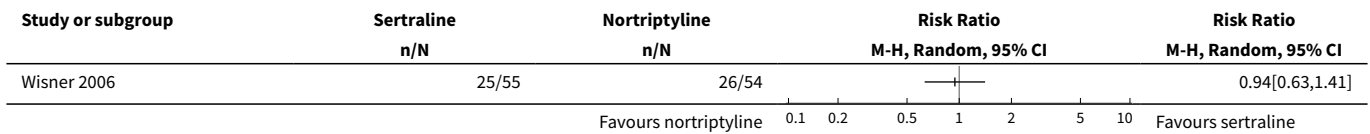
**Comparison 4. Selective serotonin re-uptake inhibitors versus other pharmacological intervention (tricyclic antidepressant)**

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Response rate at post-treatment	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected
2 Remission rate at post-treatment	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected

**Analysis 4.1. Comparison 4 Selective serotonin re-uptake inhibitors versus other pharmacological intervention (tricyclic antidepressant), Outcome 1 Response rate at post-treatment.**



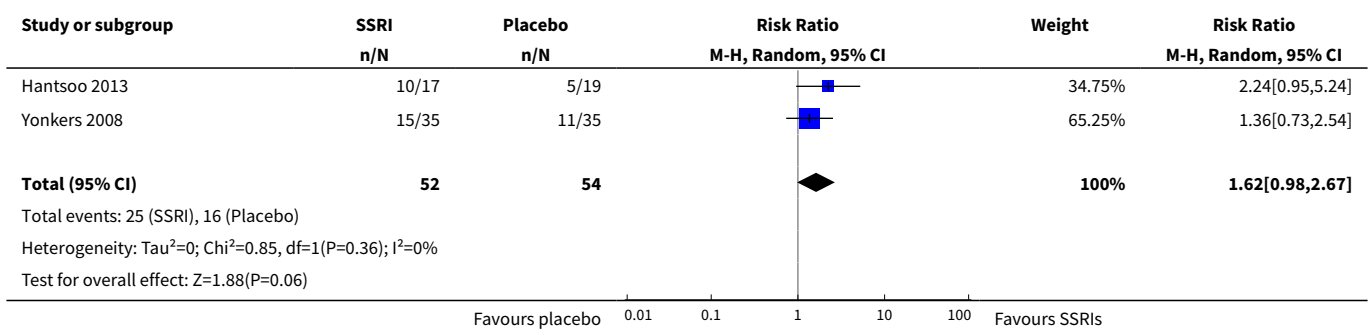
**Analysis 4.2. Comparison 4 Selective serotonin re-uptake inhibitors versus other pharmacological intervention (tricyclic antidepressant), Outcome 2 Remission rate at post-treatment.**



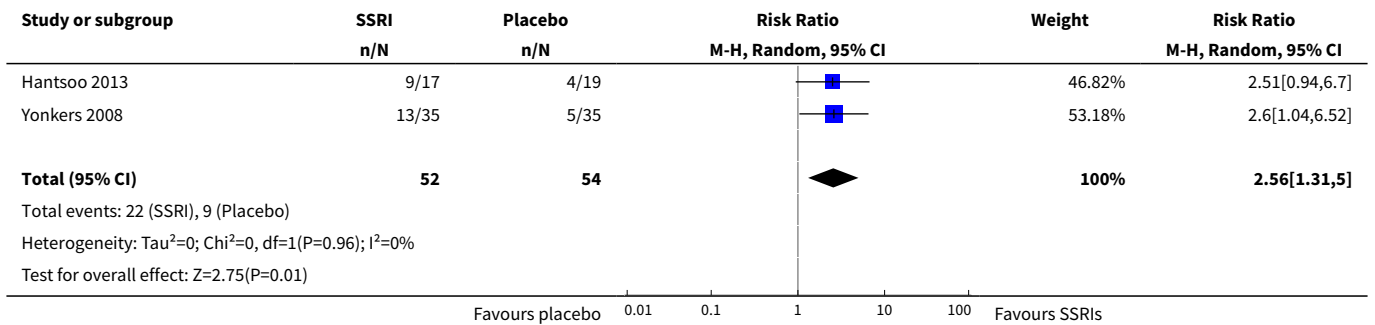
**Comparison 5. Sensitivity analysis: excluding trials with combined treatment**

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 SSRIs versus placebo: outcome 1.1 response rate at post-treatment	2	106	Risk Ratio (M-H, Random, 95% CI)	1.62 [0.98, 2.67]
2 SSRIs versus placebo: outcome 1.2 remission rate at post-treatment	2	106	Risk Ratio (M-H, Random, 95% CI)	2.56 [1.31, 5.00]

**Analysis 5.1. Comparison 5 Sensitivity analysis: excluding trials with combined treatment, Outcome 1 SSRIs versus placebo: outcome 1.1 response rate at post-treatment.**



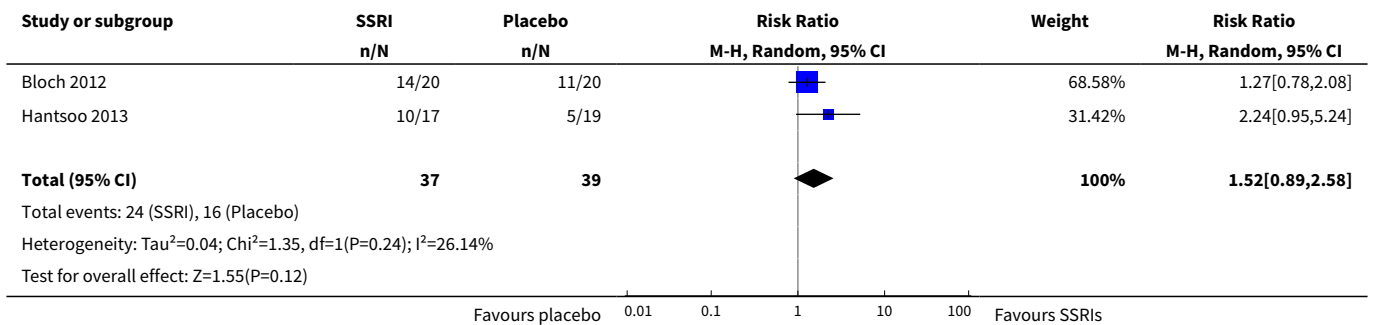
**Analysis 5.2. Comparison 5 Sensitivity analysis: excluding trials with combined treatment, Outcome 2 SSRIs versus placebo: outcome 1.2 remission rate at post-treatment.**



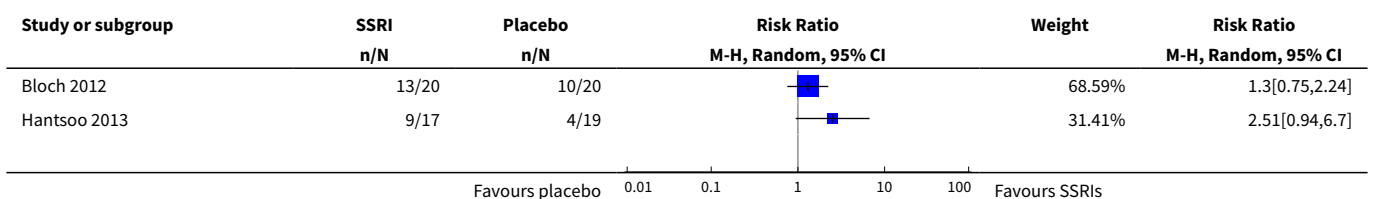
**Comparison 6. Sensitivity analysis: removing studies with high dropout or high risk of bias in any domain**

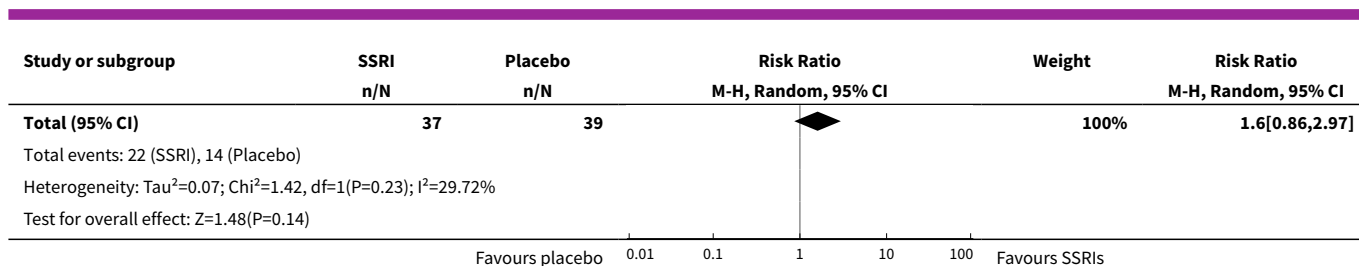
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 SSRIs versus placebo: outcome 1.1 response rate at post-treatment	2	76	Risk Ratio (M-H, Random, 95% CI)	1.52 [0.89, 2.58]
2 SSRIs versus placebo: outcome 1.2 remission rate at post-treatment	2	76	Risk Ratio (M-H, Random, 95% CI)	1.60 [0.86, 2.97]

**Analysis 6.1. Comparison 6 Sensitivity analysis: removing studies with high dropout or high risk of bias in any domain, Outcome 1 SSRIs versus placebo: outcome 1.1 response rate at post-treatment.**



**Analysis 6.2. Comparison 6 Sensitivity analysis: removing studies with high dropout or high risk of bias in any domain, Outcome 2 SSRIs versus placebo: outcome 1.2 remission rate at post-treatment.**





## ADDITIONAL TABLES

**Table 1. Summary of results and GRADE assessments**

Antidepressants compared with treatment as usual for postnatal depression					
Outcomes	Raw data by group % (no of women)	Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)	Comments
Remission	Antidepressants: 37% (48/129)  Treatment as usual: 18% (22/125)	Sharp 2010:  RR 2.11 (1.36 to 3.28)	254 (1 study)	⊕⊕⊕⊕ very low <sup>1,2</sup>	Sharp 2010: re- mission: < 13 EPDS (4 weeks)

GRADE Working Group grades of evidence

**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low quality:** We are very uncertain about the estimate.

<sup>1</sup> downgraded twice due to high risk of bias in two domains (lack of blinding of outcome assessors and low adherence)

<sup>2</sup> downgraded due to imprecision (only one study available for this comparison)

EPDS: Edinburgh Postnatal Depression Scale; RR: risk ratio.

**Table 2. Summary of results and GRADE assessments**

Antidepressants compared with listening visits for postnatal depression					
Outcomes	Raw data by group % (no of women)	Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)	Comments
Remission	Antidepressants: 47% (60/129)  Listening visits: 45% (56/125)	Sharp 2010:  RR 1.04 (0.79 to 1.36)	254  (1 study)	⊕⊕⊕⊕ very low <sup>1,2</sup>	Sharp 2010: re- mission: < 13 EPDS (18 weeks)

GRADE Working Group grades of evidence

**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

Antidepressant treatment for postnatal depression (Review)

Copyright © 2014 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

**Table 2. Summary of results and GRADE assessments** (Continued)

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low quality:** We are very uncertain about the estimate.

<sup>1</sup> downgraded twice due to high risk of bias in two domains (lack of blinding of outcome assessors and low adherence)

<sup>2</sup> downgraded due to imprecision (only one study available for this comparison)

EPDS: Edinburgh Postnatal Depression Scale; RR: risk ratio.

**Table 3. Summary of results and GRADE assessments**
**Antidepressant (sertraline) compared with other antidepressant (nortriptyline) for postnatal depression**

Outcomes	Raw data by group % (no of women)	Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)	Comments
Response	Sertraline: 56% (31/55)  Nortriptyline: 69% (37/54)	Wisner 2006:  <b>RR</b> 0.82 (0.61 to 1.10)	109 (1 study)	⊕⊕⊕⊕  <b>low<sup>1,2</sup></b>	Wisner 2006: response: 50% reduction in HAM-D from baseline (at 8 weeks)
Remission	Sertraline: 46% (25/55)  Nortriptyline: 48% (26/54)	Wisner 2006:  <b>RR</b> 0.94 (0.63 to 1.41)	109 (1 study)	⊕⊕⊕⊕  <b>low<sup>1,2</sup></b>	Wisner 2006: remission: HAM-D < 7 (at 8 weeks)

GRADE Working Group grades of evidence

**High quality:** Further research is very unlikely to change our confidence in the estimate of effect.

**Moderate quality:** Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

**Low quality:** Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

**Very low quality:** We are very uncertain about the estimate.

<sup>1</sup> downgraded due to risk of bias (incomplete outcome data owing to loss to follow-up)

<sup>2</sup> downgraded due to imprecision (only 1 study available for this comparison)

HAM-D: Hamilton Rating Scale for Depression; RR: risk ratio.

**WHAT'S NEW**

Date	Event	Description
5 September 2014	New citation required and conclusions have changed	Review updated
5 September 2014	New search has been performed	Review updated, new searches conducted and new studies included

## HISTORY

Protocol first published: Issue 2, 2000

Review first published: Issue 2, 2001

Date	Event	Description
1 November 2008	Amended	Converted to new review format.
12 January 2001	New citation required and conclusions have changed	Substantive amendment

## CONTRIBUTIONS OF AUTHORS

KT (Kylee Trevillion), LH (Louise M Howard) and EM (Emma Molyneux) developed the protocol.

KT, HM (Helen McGeown) and AK (Amar Karia) carried out searches and screening.

HM and AK conducted the data extraction.

EM, HM and AK conducted risk of bias assessment.

HM, EM, KT, AK and LH wrote the results and conclusion.

## DECLARATIONS OF INTEREST

Louise M Howard is Chair of the National Institute for Health and Care Excellence (NICE) (update) guideline on antenatal and postnatal mental health. She is Chief Investigator of an NIHR Programme Grant for Applied Research on the effectiveness of perinatal mental health services (RP- RP-DG-1108-10012) and has funding from an NIHR Research Professorship on maternal mental health, and a grant from Tommy's baby charity (with the support of a corporate social responsibility grant from Johnson & Johnson) on antipsychotics in pregnancy. Her work is also supported by the NIHR Mental Health Biomedical Research Centre at the South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author and not necessarily those of the NHS, the NIHR or the Department of Health.

Kylee Trevillion is project manager on an NIHR Programme Grant for Applied Research on the effectiveness of perinatal mental health services (RP- RP-DG-1108-10012).

Emma Molyneux is supported by a Medical Research Council (MRC) PhD Studentship and Tommy's baby charity.

There are no other declarations of interest.

## DIFFERENCES BETWEEN PROTOCOL AND REVIEW

This update of the review includes an updated background, additional information on the included studies and participants, and uses the Cochrane 'Risk of bias' assessment tool. We have updated the primary outcome; this was originally "clinically significant improvement in depression", which was not specifically defined by many papers. We also added severity of depression and quality of life as additional secondary outcomes. Based on peer reviewer comments, we altered the inclusion criteria so that non-validated scales can be included as outcomes (but will be excluded in sensitivity analyses if meta-analyses are conducted in future updates of this review to examine their impact on findings). This does not affect the data included in this review. The 'other' category of antidepressants has been separated into the separate types included to reflect the diversity of antidepressants previously included in this single category. Again, this does not alter the analyses in this review, as we identified no studies including these antidepressants. The original protocol planned sub-group analyses for women with a history of bipolar disorder, which was not included in the current study owing to the lack of relevant studies. The original protocol stated that studies with greater than 50% drop-out would not be eligible for the review. However, owing to the small evidence base, we decided to include studies with greater than 50% drop-out ([Yonkers 2008](#) reported a 56% drop-out rate).

---

**INDEX TERMS****Medical Subject Headings (MeSH)**

Antidepressive Agents, Second-Generation [\*therapeutic use]; Clinical Trials as Topic; Depression, Postpartum [\*drug therapy]; Fluoxetine [therapeutic use]; Nortriptyline [therapeutic use]; Paroxetine [therapeutic use]; Randomized Controlled Trials as Topic; Selective Serotonin Reuptake Inhibitors [\*therapeutic use]; Sertraline [therapeutic use]

**MeSH check words**

Adult; Female; Humans