UK National Screening Committee

Screening for antenatal and postnatal mental health problems

External review against programme appraisal criteria for the UK National Screening Committee

Version: Final

Author: Solutions for Public Health

Date: February 2019

The UK National Screening Committee secretariat is hosted by Public Health England.

About the UK National Screening Committee (UK NSC)

The UK NSC advises ministers and the NHS in the 4 UK countries about all aspects of population screening and supports implementation of screening programmes. Conditions are reviewed against evidence review criteria according to the UK NSC's evidence review process.

Read a complete list of UK NSC recommendations.

UK NSC, Floor 5, North Wing, Wellington House, 133-155 Waterloo Road, London, SE1 8UG
www.gov.uk/uknsc
Twitter: @PHE_Screening Blog: phescreening.blog.gov.uk

For queries relating to this document, please contact: phe.screeninghelpdesk@nhs.net

Published: April 2019

Contents

Plain English summary

Between 10% and 20% of women develop mental health problems during pregnancy (the antenatal period) and up to a year after birth (the postnatal period). This means that over 80,000 women per year in the UK might experience mental health problems during this time.

Such problems include various types of depression and anxiety. They can affect women in different ways and need different types of treatment and support. If untreated, these problems can harm the woman, her baby and her family. They can also cause long term problems for children throughout their childhood. For example, severe depression can result in the baby being born very small or premature. This can harm their health later on in life. The aim of a screening programme in pregnancy would be to reduce the possibility of these problems.

The UK NSC looked at screening for antenatal mental health problems in 2006 and for postnatal depression in 2011. The UK NSC decided not to recommend universal screening for both these conditions. This was because at the time there was not enough evidence that such screening programmes would benefit women and children.

This review is looking at the evidence now available for screening for antenatal and postnatal mental health problems.

This review looked for evidence on:

- the negative effect on women and their children of mental health problems during pregnancy and after giving birth
- whether the tests available at the moment can predict which women are at risk of such problems
- whether the treatment for such disorders can help a woman and her baby
- whether the national guidance on how to help women with these problems is being followed.

The review found that:

- there is strong evidence that common mental health problems in this period cause harm to the mother and her child
- there is evidence that some tests can identify depression
- there is a lack of evidence on the effectiveness of screening tests for other common mental disorders, such as anxiety

- the evidence on such screening tests comes from studies with low numbers of women; so studies looking at a large number of women would help to understand better how good are such tests
- the evidence on the effectiveness of treatments is not good enough to make recommendations at the current time; it is not clear which of the treatments work well for different common mental health problems before and after giving birth
- current mental health services in the UK are not fully implementing the guidance on how to look after these women; most pregnant women are asked about their mental health, but the help offered varies across services.

In conclusion, there is not enough new evidence for the UK NSC to change its recommendation. This means that screening for antenatal and postnatal mental health problems is still not recommended.

Executive summary

Purpose of the review

This review on antenatal and postnatal mental health problems will update 2 UK NSC's recommendations; the 2011'Screening for postnatal depression' update and the 2006 'Psychiatric illness in pregnancy' review. Following on from the conclusions of these reviews, the update will assess the quality and volume of evidence published since 2006 for evidence on antenatal screening for mental health problems and 2011 screening for postnatal depression.

Background

Mental health problems in pregnancy and up to 1 year postpartum (the perinatal period) develop in 10 to 20% of women. Examples of these problems include antenatal and postnatal anxiety and depression, obsessive compulsive disorder, post-traumatic stress disorder and postpartum psychosis. These conditions range from mild to severe, requiring different kinds of care or treatment and support provided by universal and specialist services. Moreover, the management and treatment of mental health problems during pregnancy and the postnatal period is different from other times because such interventions need to take in to consideration the potential impact they have not only on the women but also on her baby.

There is some evidence that untreated mental health problems in pregnancy may be associated with poorer long-term outcomes for children beyond the immediate postnatal period. This includes adverse outcomes for the foetus such as low birth weight and premature birth, emotional and behavioural difficulties in children and depression in adolescents. Postnatal mental health problems in women can be associated with adverse cognitive and developmental outcomes for their children which may be mediated through impaired mother–infant interactions.

There is national guidance on how these antenatal and postnatal services should be organised and what they should provide for women presenting with mental illness in the perinatal period. This includes how women should be tested for mental health problems with different validated questionnaires followed with a structured clinical interview to confirm diagnosis. Treatment can be pharmacological such as antidepressants or non-pharmacological, such as cognitive behaviour therapy.

Focus of the review

This evidence summary includes studies published up to February 2018. It considers 6 key questions relating to adverse outcomes, the screening test and the intervention for mental health problems in women during pregnancy and in the postnatal period.

- What adverse outcomes have been reported from common mental health problems in pregnancy and in the postnatal period?
- What is the reported accuracy of screening tools to detect common mental health problems during pregnancy?
- What is the reported accuracy of screening tools to detect postnatal depression?
- What are the benefits of pharmacological and non-pharmacological intervention (alone or in combination) in women with screen-detected common mental health problems during pregnancy?
- What are the benefits of early pharmacological and non-pharmacological intervention (alone or in combination) in women with screen-detected postnatal depression
- Is clinical detection and management currently well implemented in the UK?

Recommendation under review

The current UK NSC's recommendation is that systematic population screening programmes for common mental health problems in pregnancy and for postnatal depression are not recommended in the UK.

The aim of such programmes would be to reduce negative outcomes associated with mental ill health by an early intervention. However, a previous 2011 UK NSC's review found inconclusive evidence on the effectiveness of postnatal screening programmes for depression and the UK NSC's recommendation in 2006 not to screen for antenatal mental health disorders was also based on a lack of evidence of effectiveness.

Findings and gaps in the evidence of this review

The main conclusions are:

1. There is a large volume of evidence about adverse outcomes associated with common mental health problems experienced by women during pregnancy and postpartum. Some outcomes are consistently reported such as preterm birth and low birth weight whilst others are less consistently observed.
2. There is a paucity of evidence for effective screening tests for common mental disorders such as generalised anxiety disorder, panic disorder, phobias, social anxiety disorder, obsessive-compulsive disorder and post-traumatic stress disorder during pregnancy. The total body of

evidence during pregnancy is based on studies with low numbers of participants and in general the accuracy of data reported across studies is variable. There is also a substantial heterogeneity between the studies in relation to study design, population sampled and the diagnostic criteria used.

3. Data on postpartum depression are obtained from studies with larger numbers of participants than in the antenatal period. However, the numbers are still low when considered as evidence for population-based screening. Moreover, these studies suffer from the same heterogeneity problems noted in the antenatal studies. Even though, the quantity, quality, consistency and applicability of the evidence indicate that there are screening tools (Edinburgh Postnatal Depression Scale (EPDS), Patient Health Questionnaire (PHQ) version 2) that could be used as part of an overall screening programme, the fact that the evidence is based on small cohorts makes it difficult for decision-makers to extrapolate a conclusion that is appropriate for national population-based screening policies.
4. The EPDS, screening test for major depression disorder has a high sensitivity and specificity, but low positive predictive value. Therefore, a high proportion of women with a positive screen referred for a full psychosocial assessment are likely not to have major depressive disorder.
5. Firm conclusions about the effectiveness of each of the pharmacological and non-pharmacological interventions for women with screen-detected common mental disorders including, generalised anxiety disorder, panic disorder, phobias, social anxiety disorder, obsessive-compulsive disorder and post-traumatic stress disorder in pregnancy, cannot be drawn from the number of small studies available as they have considerable heterogeneity in their methodology, level of bias and consistency of results.
6. The evidence base for the effectiveness of pharmacological and non-pharmacological interventions for women with screen-detected antenatal and postnatal depression remain very limited. Although some evidence indicates that Cognitive Behavioural Therapy (CBT) is likely to lead to a small reduction in the severity of the condition, these conclusions are based on a very limited volume of evidence with generally weak measures of effect. Only one study of pharmacological interventions in screen-detected women was identified.
7. Antenatal and postnatal mental health problems are important mental health and public health issues and this review found evidence of their influence on pregnancy and neonatal outcomes, it also found that current mental health services in the UK are not implementing the NICE guidance in its entirety. Although most women are likely to be asked about their mental health, actions to address those problems, by onward referral, support, advice and treatment is variable.

The development of antenatal and postnatal mental health problems is complex with serious life implications for the woman, her family and her baby. This evidence summary showed that there is still a lack of clarity on the population to be identified by screening and on the ability of the current screening tools to identify with sufficient accuracy women who would benefit more from treatment.

This has important consequences, as over-detection will mean some women will have unnecessary further psychosocial assessment and may experience stigmatisation associated with mental health problems. This is complicated by the fact that there is still

insufficient evidence that universal screening and subsequent intervention improve the health outcomes for the mother or the baby.

Nevertheless, it is still important to recognise that, as part of a comprehensive clinical assessment, health professionals should be alert to the possibility of antenatal and postnatal depression and manage it according to current guidance.

Recommendations on screening

The current recommendation not to introduce a systematic UK antenatal and postnatal population screening programme for mental health problems should be retained.

Limitations

This review was limited by the lack of evidence specific to the population of interest for population-based screening particularly relating to the performance of screening tests for detecting generalised anxiety disorder, panic disorder, phobias, social anxiety disorder, obsessive compulsive disorder and post traumatic stress disorder. There is also limited evidence of the most effective intervention for those women who are screen-detected with those conditions. The majority of the evidence base for effectiveness of interventions for women with pre and postnatal depression reported outcomes from populations that were not explicitly screen-detected.

Evidence uncertainties

There are a wide range of potential screening tests, symptoms of common mental health problems and interventions investigated that have led to a large volume of evidence for this topic. However, the studies are typically small and vary significantly in methodology, level of bias and consistency of results; therefore it is difficult to use such evidence to guide policy makers in their decision on national screening programmes. Larger studies on test accuracy and treatment effectiveness would provide a better estimate than current evidence. Such evidence would need to use an agreed definition of common mental health problems and studies on treatment would also need to use agreed clinically meaningful outcomes.

Introduction and approach

Background

Mental health problems in pregnancy and up to 1 year postpartum (the perinatal period) develop in 10 to 20% of women[1].In the UK this equates to between 73,647 and 136,396 of those women who gave birth in 2016[2,3]. Examples of these problems include antenatal and postnatal depression and anxiety, obsessive compulsive disorder (OCD), generalised anxiety disorder (GAD), social phobia, panic disorder, post-traumatic stress disorder (PTSD), and postpartum psychosis. These conditions range from mild to severe, requiring different kinds of care or treatment and support provided by universal and specialist services. Moreover, the management and treatment of mental health problems during pregnancy and the postnatal period is different from other times in the women's life, because such interventions need to take in to consideration the potential impact they have not only on the women but also on her baby.

There is some evidence that untreated mental health problems in pregnancy may be associated with poorer long-term outcomes for children beyond the immediate postnatal period[1]. This includes emotional and behavioural difficulties in children and depression in adolescent offspring. Postnatal mental health problems in women have been associated with adverse cognitive and developmental outcomes for their children which may be mediated through impaired mother-infant interactions[1].

Common mental health disorders have been defined by the National Collaborating Centre for Mental Health at the Royal College of Psychiatrists in the guidance they developed for NICE clinical guideline 123 (2011)[4]. This clinical guideline described the identification and pathways to care for adults with common mental disorders, a term that includes depression, generalised anxiety disorder (GAD), panic disorder, phobias, social anxiety disorder, OCD and PTSD.

In the UK there is guidance from the Scottish Intercollegiate Guideline Network (2012)[5] and the National Institute of Clinical and Care Excellence (NICE) (2014)[1] on how services should be organised and what they should provide for women presenting with mental illness in the perinatal period.

The clinical management and service guidance for antenatal and postnatal mental health (NICE 2014)[1] covers detection, evaluation and treatment of mental health problems in pregnant women or women who are planning to have a baby or who have been pregnant in the past year. It covers a range of mental illnesses with onset during pregnancy or

diagnosed prior to the woman becoming pregnant. It gives recommendations on early detection and management of mental health problems during the perinatal period with the aim of improving outcomes for women and their families. This includes how women should be tested for mental health problems with different validated questionnaires followed by a structured clinical interview to confirm diagnosis. Treatment can be pharmacological such as antidepressants or non-pharmacological, such as cognitive behaviour therapy.

The 2012 SIGN guideline 127[5] 'Management of perinatal mood disorders' provides recommendations on the management of antenatal and postnatal mood and anxiety disorders. It covers the prediction, detection and prevention, as well as management (including psychotropic medications in pregnancy and during breastfeeding) of such conditions.

The recommendations in these guidelines are based on varying levels of quality of evidence, which is nevertheless the best available and applicable to clinicians at the current time, when considering how to diagnose and manage conditions in individuals. For example the 2014 NICE guideline[1] recommends asking women two questions[*] (the Whooley questions) as part of their general discussion about a woman's mental health and wellbeing at the first contact with primary care when pregnant (or at the booking visit), and during the early postnatal period. If the woman responds with an affirmative response to either of these questions then the heath care provider should consider using a longer mental health instrument or scale to assess the need for onward referral. The systematic review on which this recommendation is based, identified and evaluated the diagnostic accuracy of 4 instruments which were most likely to be used in UK clinical practice to assess anxiety and depression, rather than all instruments with peer reviewed research reporting performance.

The systematic review for the NICE guideline 2014[1] found a lack of evidence on detection of women with some mental health conditions such as anxiety disorders. To overcome this limitation, guideline developers agreed that evidence on case identification in non-pregnant populations could be used to formulate the recommendations.

For depression, the evidence available was sufficient to provide data on effectiveness of the various case identification tools, and also to support development of the health economic model. However, several concerns were raised by the reviewers about the methodological quality of this evidence. The concerns include: the wide range of reported sensitivity and specificity measures across studies and a substantial heterogeneity between the studies

[*] The 2 Whooley questions -1.during the past month, have you often been bothered by feeling down, depressed or hopeless?
2. during the past month, have you often been bothered by having little interest or pleasure in doing things?

(such as study design, population sampled, timing of testing, different language version of the identification tool and diagnostic criteria used).

Though there was a paucity of evidence available on women in pregnancy and in the postnatal period (except for women who may have depression), the guideline developers judged that it was important to issue a recommendation to ensure that such problems do not go unrecognised, and therefore untreated, as this could have serious consequences for a woman and her baby.

Similarly, the 2012 SIGN guideline 127[5] noted that there is insufficient evidence to recommend the use of current mental health instruments with sufficient accuracy in either the antenatal or postnatal period. However, the experts' consensus was that their use is likely to have benefit in initiating a discussion with the woman and it would provide a tool for ongoing clinical monitoring and recommended that they should be used. The evidence on which these recommendations were based (using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) tool[6]) was D, the lowest level. This indicates that the evidence used to inform these recommendations included non-analytic studies (e.g. case reports, case series), experts' opinion or evidence extrapolated from well conducted case control or cohort studies with a low risk of confounding factors or bias and a moderate probability between cause and effect.

In contrast to the individual clinical approach that concerns the NICE[1,4] and Sign (2012)[5] guidance, the population perspective of a potential national screening programme for antenatal and postnatal mental health has to balance the benefit to women who are correctly identified by a test, and receive effective treatment, against the harm to women receiving an incorrect result and undergoing further unnecessary testing or ongoing suffering from an undiagnosed condition. Of those women who are correctly identified with mental health problems in the antenatal or postnatal period only a small proportion of them will receive an effective treatment that does no harm. The remainder will find that: the treatment is effective, but brings side effects; the treatment is not effective; the treatment is both not effective and brings side effects (Raffle and Gray 2007)[7]. Therefore, when setting out to invite people for screening, who do not consider themselves to have a condition, the identification tools and interventions used must be highly effective to ensure as least harm as possible is experienced by all those taking up the offer of screening. Therefore, screening programmes need to be based on high quality evidence.

Current policy context and previous reviews

The current UK NSC's recommendation is that systematic population screening programmes for common mental health problems in pregnancy and for postnatal depression are not recommended in the UK. In the past, the UK NSC made recommendations on screening for postnatal depression and screening for psychiatric illness in pregnancy separately. In 2006, the UK NSC made the decision to not recommend introducing a national screening programme for antenatal mental health disorders. In 2011, a similar decision about systematic screening for postnatal depression was made based on a UK NSC's evidence review published in 2011[8]. This review outlined a lack of clarity on the population to be identified by screening who would benefit from an intervention and the accuracy of the screening tools to identify those at risk.

This evidence summary on antenatal and postnatal mental health problems will provide the evidence to update the previous UK NSC's recommendations. The update will assess the quality and volume of evidence published since 2006 for evidence on common mental health problems in the antenatal period and from 2011 onwards for screening for postnatal depression.

Objectives

The current review will focus on describing the evidence for the detection and management of women identified with common antenatal mental health problems and postnatal depression. This includes:

- the association of common antenatal mental health disorders and postnatal depression with adverse outcomes
- test accuracy, timing and repetition of tests to detect common antenatal mental health problems and postnatal depression
- improved outcomes for screen-detected women from interventions
- current practice and the implementation of current national guidance.

The key questions addressed in the current review were developed by the UK NSC with input from Solutions for Public Health.

The key questions and the UK NSC criteria that they relate to are presented in Table 1 below.

Table 1. Key questions for the evidence summary and relationship to UK NSC screening criteria

	Criterion	Key questions	Studies Included
	THE CONDITION		
1	The condition should be an important health problem as judged by its frequency and/or severity. The epidemiology, incidence, prevalence and natural history of the condition should be understood, including development from latent to declared disease and/or there should be robust evidence about the association between the risk or disease marker and serious or treatable disease.	Question 1: What adverse outcomes are associated with common mental health problems in pregnancy and in the postnatal period?	11
	THE TEST		
4	There should be a simple, safe, precise and validated screening test.	Question 2: What is the reported accuracy of screening tools to detect common mental health problems during pregnancy?	5
5	The distribution of test values in the target population should be known and a suitable cut off level defined and agreed.	Question 3: What is the reported accuracy of screening tools to detect postnatal depression?	3
	THE INTERVENTION		
9	There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care. Evidence relating to wider benefits of screening, for example those relating to family members, should be taken into account where available. However, where there is no prospect of benefit for the individual screened then the screening programme shouldn't be further considered.	Question 4: What are the benefits of pharmacological and non-pharmacological intervention (alone or in combination) in women with screen-detected common mental health problems during pregnancy?	7
		Question 5: What are the benefits of early pharmacological and non-pharmacological intervention (alone or in combination) in women with screen-detected postnatal depression?	5
	IMPLEMENTATION		
15	All the cost-effective primary prevention interventions should have been implemented as far as practicable.	Question 6: Is clinical detection and management currently well implemented in the UK?	6

Methods

The current review was conducted by Solutions for Public Health (SPH), in keeping with the UK National Screening Committee evidence review process. Database searches were conducted on 14th February 2018 to identify studies relevant to the questions detailed in

Table 1.

Eligibility for inclusion in the review

The following review process was followed:

1. Each abstract was reviewed against the inclusion/ exclusion criteria by 1 reviewer. Where the applicability of the inclusion criteria was unclear from the abstract, the article was included at this stage in order to ensure that all potentially relevant studies were captured.
2. Full text articles required for the full text review stage were acquired.
3. Each full-text article was reviewed against the inclusion/ exclusion criteria by 1 reviewer, who determined whether the article was relevant to 1 or more of the review questions.
4. Any queries at the abstract or full text stage were resolved through discussion with a second reviewer.
5. The review was quality assured by a second senior reviewer, not involved with the writing of the review in accordance with SPH's quality assurance process.

Eligibility criteria for each key question are presented in Table 2 below. For questions 1, 2, 4 and 6 between September 2006 and February 2018 were eligible for consideration in the review. For questions 3 and 5, only peer reviewed studies published in English between 2011 and February 2018 were eligible for consideration in the review.

A total of 12,991 unique references were identified and sifted by an information scientist by title and abstract for potential relevance to the review. Of the unique references identified 746 titles and abstracts were reviewed by an SPH reviewer for further appraisal and possible inclusion in the final review.

Overall, 103 studies were identified as possibly relevant during title and abstract sifting and further assessed at full text.

Table 2. Inclusion and exclusion criteria for the key questions.

Key question	Inclusion criteria						Exclusion criteria
	Population	**Target condition**	**Intervention**	**Comparator**	**Outcome**	**Study type**	
Question 1 **What adverse outcomes are associated with common mental health problems in pregnancy and in the postnatal period?**	Pregnant women or up to 3 months postpartum, without previous clinical diagnosis of mental health problems. Women with experience of miscarriage stillbirth and neonatal death and women under 18 years of age are excluded.	Common mental health problems	N/A	N/A	*Maternal* a) Danger to self or infant b) Increased risk of subsequent depression c) Puerperal psychosis d) Increased risk of subsequent anxiety disorders e) Increased risk of subsequent personality disorders f) Substance misuse g) Eating disorders h) Mother-infant interaction problems i) Quality of life j) Adherence or persistence with treatment. *Baby* k) Preterm birth l) Small for gestational age m) Large for gestational age n) Low birth weight o) Neonatal Intensive Care Unit admission p) Growth and development up to 1 year of age. q) Mealtime conflict *Child* r) Emotional and behavioural adjustment s) Social adjustment t) Cognitive development *Adolescent* u) Depression	Observational descriptive studies; cross-sectional studies, cohort studies; case-control studies and systematic reviews of these.	Case reports, case series, reviews, non-peer reviewed literature

Question 2 What is the reported accuracy of screening tools to detect common mental health problems during pregnancy?	Pregnant women without previous clinical diagnosis of mental health problems	Common mental health disorders	Screening tools to detect common mental health disorders during the pregnancy period	Diagnosis confirmed with clinical interview	a) Sensitivity b) Specificity c) False positive rate d) False negative rate e) PPV/NPV	Studies of consecutively enrolled populations and systematic reviews of these	Case reports, case series, non-systematic reviews, case control studies, non-peer reviewed literature
Question 3 **What is the reported accuracy of screening tool to detect postnatal depression?**	Postpartum women	Postnatal depression	Questionnaires in common use	Diagnosis confirmed with clinical interview	a) Sensitivity b) Specificity c) False positive rate d) False negative rate e) PPV/NPV	Studies of consecutively enrolled populations and systematic reviews of these	Case reports, case series, non-systematic reviews, case-control studies, non-peer reviewed literature
Question 4 **What are the benefits of pharmacological and non-pharmacological intervention (alone or in combination) in women with screen-detected common mental health problems during pregnancy?**	Pregnant women with screen-detected common mental health problems during pregnancy	Common mental health disorders	Pharmacological and non-pharmacological interventions used in pregnancy	Compared to usual care or placebo (for pharmacological) or no active intervention	*Maternal* a) Danger to self or infant b) Increased risk of subsequent depression c) Puerperal psychosis d) Increased risk of subsequent anxiety disorders e) Increased risk of subsequent personality disorders f) Substance misuse g) Eating disorders h) Mother-infant interaction problems i) Quality of life j) Adherence or persistence with treatment. *Baby* k) Preterm birth l) Small for gestational age m) Large for gestational age n) Low birth weight o) Neonatal Intensive Care Unit admission p) Growth and development up to 1 year of age.	Randomised and quasi-randomised controlled trials and observational studies (cohort studies, and cross-sectional studies) or systematic reviews of these	Case reports, case series, non-systematic reviews, case-control studies, non-peer reviewed literature

					q)Mealtime conflict *Child* r) Emotional and behavioural adjustment s) Social adjustment t) Cognitive development *Adolescent* U)Depression		
Question 5 **What are the benefits of early pharmacological and non-pharmacological intervention (alone or in combination) in women with screen-detected postnatal depression?**	Postpartum women.	Postnatal depression	Pharmacological and non-pharmacological Interventions used in the postnatal period	Compared with placebo or no active intervention or usual care	*Maternal* a) Danger to self or infant b) Increased risk of subsequent depression c) Puerperal psychosis d) Increased risk of subsequent anxiety disorders e) Increased risk of subsequent personality disorders f) Substance misuse g) Eating disorders h) Mother-infant interaction problems i) Quality of life j) Adherence or persistence with treatment. *Baby* k) Preterm birth l) Small for gestational age m) Large for gestational age n) Low birth weight o) Neonatal Intensive Care Unit admission p) Growth and development up to 1 year of age. q) Mealtime conflict *Child* r) Emotional and behavioural adjustment s) Social adjustment t) Cognitive development *Adolescent* u) Depression	Randomised and quasi-randomised controlled trials and observational studies (cohort studies, and cross-sectional studies) or systematic reviews of these	Case reports, case series, non-systematic reviews, case-control studies, non-peer reviewed literature

Question 6 Is clinical detection and management currently well implemented in the UK?	All pregnant women	Common mental health problems	Current clinical management in the UK	For outcome a disease known prevalence For outcomes b, c and d : N/A	a) Proportion of mental health problems detected b) Proportion of women asked questions according to guidance c) Proportion of women with mental health problems referred for intervention d) Proportion of women attending/complying with intervention e) User experiences	Audit data, cross sectional study, cohort study (prospective and retrospective), systematic reviews	Non-UK studies, non-systematic reviews, case studies

Appraisal for quality/risk of bias tool

The following tools were used to assess the quality and risk of bias of each study included in the review:

- systematic reviews: Critical Appraisal Skills Programme (CASP) Systematic Review Checklist
- systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both: A Measurement Tool to Assess systematic Reviews 2 (AMSTAR 2)
- cohort studies: Critical Appraisal Skills Programme (CASP) Cohort Study Checklist
- diagnostic accuracy studies: Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool
- randomised controlled trials (RCTs): Cochrane Collaboration's "Risk of Bias" Tool
- qualitative studies: Critical Appraisal Skills Programme (CASP) qualitative research checklist.

Results of the quality assessments are presented in the summary and appraisal of individual studies in Appendix 3.

Databases/sources searched

A systematic search of 3 databases (Medline, Embase and Cochrane) was conducted on 14th February 2018 to identify studies relevant to the questions detailed in Table 1. The search strategy is presented in Appendix 1.

Question level synthesis

Criterion 1

The condition should be an important health problem as judged by its frequency and/or severity. The epidemiology, incidence, prevalence and natural history of the condition should be understood, including development from latent to declared disease and/or there should be robust evidence about the association between the risk or disease marker and serious or treatable disease

Question 1 – What adverse outcomes are associated with common mental health problems in pregnancy and in the postnatal period?

Eligibility for inclusion in the review

The inclusion criteria for this question are outlined below:

Population: Pregnant women or up to 3 months postpartum, without previous clinical diagnosis of mental health problems. Women with experience of miscarriage, stillbirth and neonatal death and women under 18 years of age are excluded.

Outcomes: Adverse outcomes cover those experienced by the women with mental health issues and their offspring as babies, children and adolescents. The specific outcomes are:
Maternal

- danger to self or infant
- increased risk of subsequent depression
- puerperal psychosis
- increased risk of subsequent anxiety disorders
- increased risk of subsequent personality disorders
- substance misuse
- eating disorders
- mother-infant interaction problems
- quality of life
- adherence or persistence with treatment.

Baby

- preterm birth
- small for gestational age
- large for gestational age

- low birth weight
- neonatal Intensive Care Unit admission
- growth and development up to 1 year of age.
- mealtime conflict

Child
- emotional and behavioural adjustment
- social adjustment
- cognitive development

Adolescent
- depression

Study design: Observational descriptive studies; cross-sectional studies, cohort studies; case-control studies and systematic reviews of these.

Date and language: English peer reviewed publications from 2006 onwards.

Description of the evidence

The searches carried out for question 1 identified a large volume of evidence describing the association of mental health problems in women for both antenatal and postnatal adverse outcomes either for themselves or their offspring. Searches yielded 746 results of which 196 were judged to be relevant to this question. This evidence included 76 systematic reviews and 120 individual primary studies. In order to highlight the best evidence available, systematic reviews and meta-analyses were prioritised.

After reviewing the abstracts, 19 studies met the criteria for full text review. Following review of the full texts, 11 studies were included of which 6 were systematic reviews or meta-analyses and 5 were primary studies. Eight articles did not meet the eligibility criteria for this review and the reasons for exclusion are outlined in Table 14 in Appendix 2.

In 2014 the Lancet published a series review on Perinatal Mental Health with the aim of summarising the evidence on the effects of perinatal mental disorders on mothers and children (Howard et al 2014, Jones et al 2014, Stein et al 2014)[9,10,11].These types of articles(Lancet Commissions) are commissioned for the Lancet group by the Lancet Global Health Journal on particular topics. The reports are developed by a panel of academics that use their expertise and the best evidence available to discuss and evaluate a topic of international interest. When possible they aim at proposing a series of policy recommendations, and suggestions for original research. The three papers in the Perinatal

Mental Health review were excluded from this review as they did not meet the inclusion criteria (ie they were not primary observational descriptive studies; cross-sectional studies, cohort studies; case-control studies or systematic reviews of these). However, as these publications are key summaries of the evidence relevant to this review and includes the consensus agreement of specialists in the field of antenatal and postnatal mental health there is a brief summary of key points at the end of the 'Discussion of findings' section.

The 11 included studies cover some, but not all, mental health problems detected in women in the pre or postnatal period and describe some, but not all of the possible adverse outcomes listed in the eligibility criteria. The included publications primarily focussed on women with depression, anxiety and stress.

Discussion of findings

A study-level summary of data extracted from each included publication is presented in the appraisal of individual studies in Appendix 3.

Of the 11 papers included for question 1, 6 were systematic reviews or meta-analyses and 5 were analyses from data of longitudinal studies. Table 3 summarises key details from each of the studies.

There is a large volume of literature concerned with the adverse outcomes of mental health problems in pregnancy and the postpartum period on women and their offspring. The studies included for question 1 of this review were, on the whole, of high quality, with study participant numbers ranging from 10,569 to 10,184 for individual studies and inclusion of between 11 and 54 studies for meta-analyses and systematic reviews. However, studies could not always control for confounding factors such as co morbidity with other mental health disorders and socioeconomic status which hampered the accurate attribution of adverse outcomes in children to maternal antenatal and postnatal mental health problems.

Mother and baby outcomes

One systematic review focussed on the impact of suicide ideation during pregnancy on the woman (Gelaye et al 2016[12]).Two further systematic reviews reported the relationship between antenatal PTSD (Cook et al 2018)[13] depression (Jarde et al 2016)[14] and birth outcomes. One longitudinal study reported the relationship between birth outcomes and antenatal stressful life events during pregnancy (Mackinnon et al 2018)[15].

Table 3. Summary of included studies for key question 1

Mental health condition of mother	Outcome	Number	Study outcome	Quality appraisal
Antenatal depression				
Antenatal suicide ideation	Maternal and baby outcomes; Danger to self and infant, preterm birth, low birth weight	57 studies in total. Results from 4 studies meeting the review inclusion /exclusion criteria reported here	1 study showed risk of death from suicide due to suicidal ideation during pregnancy increased by 92%compared to the general population (HR=1.92, 95% CI, 1.53–2.41) 1 study showed infants born to women who attempted suicide during pregnancy were more likely to be low birth weight (<2500 g) (OR=1.25, 95% CI, 1.08–1.44) and have respiratory distress syndrome (OR=1.41, 95% CI, 1.07–1.86) 1 study reported that risks for preterm birth (OR=1.34, 95% CI, 1.01–1.77), low birth weight (OR=1.49, 95% CI, 1.04–2.12), and circulatory system congenital anomalies (OR=2.17, 95% CI, 1.02–4.59) were elevated among women who attempted self-poisoning during pregnancy as compared with those who did not 1 study showed that infants of mothers reporting depressive symptoms with suicidal ideation weighed 239.5 g (95% CI, 3.9–475.1) less on average than infants born to mothers reporting depressive symptoms without suicidal ideation	**Gelaye et al (2016)**[12] **Systematic review** Of the 57 studies included 28 were retrospective, 19 were cross sectional, and one was a case control study so the majority of the studies were subject to recall bias. Moreover, the tools used to assess the suicidal ideation varied across studies and this might explain the heterogeneity in research findings
Antenatal depression	Baby outcomes (to 1 year); Preterm birth, small for gestational age, low birth weight	23 studies The total number of women was 25 440 across 23 research studies	Preterm birth <week 37 (14 studies, 21 048 women) crude OR= 1.56 (95%CI,1.25 to 1.94) p<0.001 Low birth weight<2500g (8 studies, 3262 women) crude OR=1.96 (95% CI,1.24 to 3.10) p<0.004 Small for gestational age <10th percentile (1 study, 4044 women) crude OR=1.37 (95% CI,1.10 to 1.70) p<0.005	**Jarde et al (2016)**[14] **Systematic review and meta-analysis** The CASP checklist for systematic reviews was used to assess the quality of this review. There were no areas of concern about the quality of the review

				The quality appraisal performed by the authors of the systematic review and meta-analysis assessing the quality of the included studies was done using Duval and Tweedie's(2000)[16] trim and fill method and a modified version of the NewcastleOttawa Scale Wells et al 2015)[17] for assessing the quality of observational studies. Several of the studies (more than half) did perform a rigorous diagnostic assessment of depression, but used only screening tools for the diagnosis. Therefore, it was not possible to confirm the trajectory of the symptoms measured nor whether such symptoms constituted a major depressive episode. The effects of several important confounding variables could not be taken into account due to a lack of reporting of adjusted data in most studies.
Antenatal stress and anxiety	Child outcomes; Cognitive development (0 to 4 years	11 studies (n=5903 with sample size range of 66 to 3139)	Low but significant relationship between maternal antenatal stress or anxiety and early child cognitive development (r‡=-0.05, p<0.001 95% CI,-.07 to-.02‡) Life events and cognitive development has a significantly greater effect (r‡=-0.31; p<0.0001; 95% CI, -0.40 to -0.22‡) than pregnancy related stress and anxiety(r=-0.08; p<0.001; 95% CI, -0.14 to-0.02) or stress and anxiety unrelated to pregnancy (r=-0.02; p<0.001; 95%CI, -0.05 to -0.01‡).	**Tarabulsy et al (2014)[21] Meta-analysis** The AMSTAR-2 checklist for systematic reviews and meta-analyses studies was used to assess the quality of this review. The authors did not report the age of the offspring when cognitive development was assessed although from the titles of the included papers this appears to be from infancy to early child pre-school age. The quality appraisal performed by the authors of the meta-analysis assessing the quality of the included studies and there were concerns that covariates such as prenatal and post natal depression and other mood disorders were not taken into account when attributing child cognitive development to antenatal stress and anxiety. There were differences in effect sizes between prospective (k=8, r=0.03,

				p<0.001,95% CI -0.06 to -0.01) and retrospective studies (k=3, r=-0.11, p<0.001,95% CI, -0.16 to -0.06) with retrospective studies showing a greater effect than prospective studies.
Antenatal stressful events	Child outcomes; Emotional and behavioural adjustment (7 to 16 years)	10,184 women and offspring pairs	Stressful event burden**Error! Bookmark not defined.** was positively associated with: • preterm birth (p=0.011) • low birth weight (p=0.013) • highest conduct disorder symptom trajectory in children associated with the highest quartile of prenatal stress (p<0.01) and continuous stress (p<0.05). • highest hyperactivity symptom trajectory in children associated with the highest quartile of prenatal stress (p<0.05)	**Mackinnon et al (2018)[15] Longitudinal study** The CASP checklist for cohort studies was used to assess the quality of this review. Attrition over the 16 year follow up period was observed and was associated with both prenatal stress and conduct disorder and hyperactivity at 7 years of age. This factor biases the results to be more conservative than they are likely to be, had no attrition occurred.
Antenatal stressful events	Adolescent outcomes; Depression (17 to 18 years)	10,569 women and offspring pairs	Multi-nomal logistic regression of the adjusted data showed an increase of 1 unit on the maternal Stressful Life Event**Error! Bookmark not defined.** score during gestation was associated with increased offspring depressive symptoms (β††**Error! Bookmark not defined.**=0.07, p<0.01) and major depression at 17 to 18 years (OR=1.03, 95% CI,1.01 to 1.05) This corresponds to a 12.6% increased risk of internalising disorders in offspring aged 17 to 18 if an event occurred that greatly affected the mother during pregnancy	**Kingsbury et al (2016)[19] Longitudinal study** The CASP checklist for cohort studies was used to assess the quality of this review. There were no areas of concern about the quality of the review Missing data on outcomes and confounders were estimated and 5 estimated datasets were created and estimates pooled. Results were similar to the complete case sample
Antenatal depression and anxiety	Child outcomes; Emotional and behavioural adjustment (10 to 11 years)	2,891 women and offspring	Mothers with antenatal depression were more likely to report their children with higher levels of: • hyperactivity (p<0.05) • emotional symptoms (p<0.01) • conduct problems (p<0.05) • total difficulties (p<0.001) Mothers with antenatal anxiety were more likely to report their children with higher levels of • emotional symptoms(p<0.001) • total difficulties(p<0.05)	**Leis et al (2014)[25] Longitudinal study** The CASP checklist for cohort studies was used to assess the quality of this review. There were no areas of concern about the quality of the review

Antenatal and postnatal depression				
Antenatal and postnatal PTSD	Baby outcomes (to 1 year); Mother-infant interaction, preterm birth, low birth weight, growth and development, (from 0 to 1 year).	26 publications reporting 21 studies The total number of women was 9942 across 21 research studies	Risk of preterm birth in women with likely PTSD (Adjusted OR = 1.22, 95% CI: 0.57 to 2.61) Studies were inconsistent in findings of mother infant interaction, infant interactional behaviour, cognitive development and maternal sensitivity and control (no figures given) No studies reported an association with gestational age at birth (no figures given)	**Cook et al (2018)[13] Systematic review** The CASP checklist for systematic reviews was used to assess the quality of this review. Validated measures were not used to collect postpartum birth outcome information in any of the 11 studies examining birth outcomes included in this review, which may limit the reliability and validity of the findings. Very few figures were extracted from the papers to support the findings described The quality appraisal performed by the authors of the systematic review assessing the quality of all the included studies was done using a checklist based on Downs and Black's (1998) feasibility checklist[18]. They concluded that the methodological quality of the studies included was poor to medium with scores ranging from 8 to 16 on a 30-point scale, and over 19% of papers scoring over 15. Methodological weaknesses of the studies included insufficient sample size, use of invalidated measures, and limited external validity.
Antenatal and Postnatal depression	Child outcomes Cognitive development (0to 5 years	14 studies (n=not reported)	Univariate analysis showed perinatal depression was related to lower cognitive scores in children ≤56 months (Cohen's *d** = −0.25, 95% CI, −0.39 to −0.12) Postpartum association when analysis restricted to 3 studies where maternal depressive symptoms were measured during 6–8 weeks, (β††=−4.17, 95% CI, −8.01 to −0.32, p = 0.03) Average score of 4.2 points lower on the Mental Development Index for Infants and Toddlers whose mothers reported high depressive symptoms	**Lui et al (2017)[20] Meta-analysis** The AMSTAR-2 checklist for systematic reviews and meta analyses studies was used to assess the quality of this review. The method of assessing the quality of the 14 included studies was not described although the authors report the quality as 'variable'. The paper mentions confounding factors that may effect the results of the unadjusted univariate analysis but does not describe what they are. The 3 studies

			compared to those reporting low depressive symptoms	included in an adjusted multivariate analysis highlighted 2 (gender and parental education) of the potential confounding factors but did not name others.
Common mental disorders in antenatal and postnatal period	Child outcomes Emotional and behavioural problems (3 years of age)	10,078 women and offspring pairs	Women with treated and untreated common mental disorders were more likely to have children rated as having socio-emotional difficulties (p<0.05) at age 3 than women at low risk of common mental disorders.	**Prady et al (2016)[24] [24]Longitudinal study** A CASP checklist for longitudinal studies was undertaken Limitations include not identifying specific mental health disorders and treatment. Self-report of mood which suggested an 'untreated' CMD may be sub-clinical or transient distress that might not meet current criteria for treatment
Postnatal mental health problems				
Postnatal depression	Child outcomes Emotional and behavioural adjustment (8 years of age)	473 women and offspring pairs	Postpartum Edinburgh Postnatal Depression Scale (EPDS)score >10 at 2, 3 or 6 months associated with higher risk of mental health problems in children at age 8 (p=0.017) PND at 3 months postpartum was associated with : • anxiety/depression,(p=0.007), • social problems (p=0.001), • attention problems(p=0.010), • rule-breaking behaviour (p=0.037) • aggressive problems (p=0.005).	**Closa-Monsterolo et al (2017)[23] Longitudinal study** The CASP checklist for RCTs was used to assess the quality of this review. There were no areas of concern about the quality of the review.
Postnatal depression	Adolescent outcomes Emotional behavioural, psychosocial and cognitive problems (11 to18 years	16 studies A number of studies included in this review used the same cohorts The total number of families	Two studies showed a marked effect of PND on cognitive development in boys (F**=5.13, p<0.001 and (F=4.18, p<0.05). Socioemotional development, externalising problems, internalising problems and overall psychopathology showed less consistent findings (no figures given).	**Sanger et al (2015)[22], Systematic review.** The CASP checklist for systematic reviews was used to assess the quality of this review. This review included studies with a wide range of different types of scoring systems which were not possible to compare easily No formal quality appraisal was performed by the authors of the systematic review; however, the authors noted that a possible

	included in this review at baseline was 13,199 across eight cohorts	limitation of this review was that 16 studies included came from only eight cohorts. Furthermore, two cohorts reflecting a small baseline sample of n=100 and n=179, respectively, represented the majority of the articles included Therefore, the findings based on these sample sizes have the potential to skew the overall findings of the review

* Cohen's d is effect size as a measure of the difference in the two groups' means divided by the average of their standard deviations(d=0.2 is small effect, d=0.5 is medium effect and d=0.8 is a large effect)

†† The regression β coefficient is the degree of change in the outcome variable for every 1-unit of change in the predictor variable. If the β coefficient is not statistically significant, the variable does not significantly predict the outcome

**F-value is a ratio of two quantities that are expected to be roughly equal under the null hypothesis, which produces an F-statistic of approximately 1. In order to reject the null hypothesis that group means are equal, (ie that there is a difference between groups) the F-value must be high.

Suicide ideation

The 57 studies included in the systematic review about suicide ideation (Gelaye et al 2016)[12], of which 4 were relevant to the outcomes for question 1.The remaining studies either used cohorts of women with previous mental health problems, or were in countries that were not comparable for population and health care systems to the UK. A meta-analysis of all studies (4,835,313 participants of which 4,833,286 were part of one population based retrospective cohort study), that did and did not meet the inclusion criteria for this review, found that women with suicide ideation in pregnancy were at increased risk of attempting suicide, (Hazard Ratio [HR] =1.92, 95% Confidence interval [CI]: 1.53–2.41). The systematic review reported adverse outcomes for babies born to mothers with suicide ideation or that attempted suicide from 3 studies that met the inclusion criteria of this review with 4,834,010 participants of which 4,833,286 were part of one population based retrospective cohort study. These women were more likely to have babies:

- of low birth weight: (in 2 studies Odds Ratio (OR)=1.25, 95% CI, 1.08 to 1.44 and OR =1.49, 95% CI, 1.04 to 2.12)
- weighing less than women with no suicide ideation (239.5g lighter (95% CI, 3.9 to 475.1)
- born preterm (OR=1.34, 95% CI, 1.01 to 1.77)
- born with respiratory distress (OR=1.41, 95% CI, 1.07 to 1.86).

Post-Traumatic Stress Disorder

A systematic review by Cooke et al (2018)[13] that included 11 studies had 1 study that explicitly focussed on PTSD that reported an increase in the likelihood of preterm birth (Adjusted OR = 1.22, 95% CI, 0.57 to 2.61, 2487 participants 129 with PTSD symptomology). The findings were inconsistent for an effect of PTSD on mother-infant interaction, infant interactional behaviour and cognitive development. No studies in the systematic review reported an association with gestational age at birth and PTSD.

Birth outcomes of women with untreated antenatal depression

One systematic review that included 23 studies focused on birth outcomes of women with untreated antenatal depression studies (Jarde et al 2016)[14]. This review showed that untreated antenatal depression increased the likelihood of preterm birth, low birth weight and babies small for gestational age:

- preterm birth <week 37 (14 studies, 21 048 women, OR=1.56 (95% CI,1.25 to 1.94) p<0.001
- low birth weight<2500g (8 studies, 3262 women, OR=1.96, 95% CI,1.24 to 3.10) p<0.004
- small for gestational age <10th percentile (1study, 4044 women, OR=1.37, 95% CI, 1.10 to 1.70 p<0.005).

Stressful events during pregnancy
One longitudinal study (Kingsbury et al 2016)[19] in the UK reported that the number and impact of stressful events†experienced by women in pregnancy increased the likelihood of preterm birth (p=0.011) and low birth weight (p=0.013)[23].

Outcomes in children and adolescents

Two meta-analyses (Lui et al 2017 and Tarbulsy et al 2014)[20,21,] 1 systematic review (Sanger et al 2015)[22] and 5 papers reporting data from longitudinal studies (Mackinnon et al 2018, Kingsbury et al 2016,Closa-Monsterolo et al 2017, Prady et al 2016, Leis et al 2014,)[15,19 23,24,25] explored the relationship between antenatal and/or postnatal depression, anxiety and stressful events experienced by women and a range of adverse outcomes reported in children and adolescents.

Emotional and behavioural problems in the offspring
Three longitudinal cohort studies examined the relationship between mental health conditions in pregnancy and emotional and behavioural problems in their 3, 10 to 11, and 7 to 16 year old offspring. In the UK, Mackinnon et al (2018)[15], examined the impact of stressful events occurring during pregnancy and emotional and behavioural outcomes in children aged 7 to 16. This showed that the steepest conduct disorder trajectory and hyperactivity symptom trajectory was associated with the highest quartile of prenatal stress ($p<0.01$, $p<0.05$ respectively).

The relationship between antenatal depression and other common mental health disorders during pregnancy and emotional and behavioural problems was reported in 2 UK longitudinal studies. Prady et al. (2016)[24] concluded that women with either treated or untreated common mental disorders during pregnancy were more likely to rate their children as having socio-emotional difficulties at age 3 ($p<0.05$) than those at low risk of common mental disorders. Similarly, Leis et al. (2014)[25] found the children aged 10 to 11 born to mothers with antenatal depression were more likely to have higher levels of hyperactivity ($p<0.05$), emotional symptoms ($p<0.01$), conduct problems ($p<0.05$) and total overall difficulties ($p<0.001$).

Finally, 1 longitudinal study investigated the impact of PND on emotional and behavioural problems in 8 year old children (Closa-Monasterolo et al 2017)[23]. Authors reported that ,

† Stressful life events (SLEs) are defined as discrete experiences that disrupt an individual's usual activities, causing a substantial change and readjustment. SLEs have been linked to depression and, to a lesser extent, to anxiety (Phillips A, Carroll D and Der G Negative life events and symptoms of depression and anxiety: Stress causation and/or stress generation *Anxiety Stress Coping*. 2015; 28(4): 357–371)

postpartum PND at 3 months was associated with anxiety/depression (p=0.007), social problems (p=0.001), attention problems (p=0.010), rule-breaking behaviour (p=0.037), and aggressive problems (p=0.005).

Cognitive development

One meta-analysis of 11studies with 5903 participants studies (Tarabulsy et al 2014)[21] examined the association between antenatal depression and cognitive development prior to school entry. The authors found that life events have a greater adverse effect on cognitive development (r‡=-0.31; p<0.0001, 95% CI, -0.40 to -0.22,) than pregnancy related stress and anxiety (r=-0.08, p<0.001;95% CI,-0.14 to -0.02,) or stress and anxiety unrelated to pregnancy issues (r=-0.02; p<0.001, 95%CI, -0.05 to -0.01), though all of the effect sizes are small. Likewise, another meta-analysis (Lui et al 2017)[20] found a small effect size in terms of the impact of perinatal depression adversely affecting cognitive scores in children ≤56 months (Cohen's *d*§ = −0.25, 95% CI, −0.39 to −0.12).

The relationship between postnatal depression and children's later wellbeing was examined in a systematic review of 16 studies(Sanger et al 2015)[22] of which 2 studies examined the impact of PND on cognitive development and showed a marked effect of PND on cognitive development in adolescent boys (F=5.13, p<.001 and F=4.18, p<.05)**.

Depression

Mothers from a population-based cohort, (n=10569 of the Avon Longitudinal Study of Parents and Children (ALSPAC) reported the occurrence and impact of 42 prenatal stressful life events. Depressive symptoms in children were assessed using a computerized version of the Clinical Interview Schedule-Revised (CIS-R) at age 17 to 18, as well as 13 self-report statements from the Short Mood and Feelings Questionnaire (SMFQ) at 6 time points from ages 10 to 11 to 18 to 19.In adolescents aged 17 to 18 there was an association between increased offspring depressive symptoms (β††=0.07, p<0.01) and major depression (OR=1.03, 95% CI,1.01 to 1.05) and an increase in number and impact of antenatal stressful events **Error! Bookmark not defined.**(Kingsbury et al 2 016)[19].

‡ r, regression coefficient, a measure of strength of a relationship with no relation being 0 and a perfect relationship being 1

§ Cohen's d is effect size as a measure of the difference in the two groups' means divided by the average of their standard deviations(d=0.2 is small effect, d=0.5 is medium effect and d=0.8 is a large effect)

** F-value is a ratio of two quantities that are expected to be roughly equal under the null hypothesis, which produces an F-statistic of approximately 1. In order to reject the null hypothesis that group means are equal, (ie that there is a difference between groups) the F-value must be high.

†† The regression β coefficient is the degree of change in the outcome variable for every 1-unit of change in the predictor variable. If the β coefficient is not statistically significant, the variable does not significantly predict the outcome

The Lancet Commission of the 2014 Perinatal Mental Health series looked at perinatal mental disorders and the effects they might have on the fetus and child [9,10,11]. The conclusion of these articles were that there is evidence linking perinatal mental health disorders experienced by the mother with negative outcomes in her child that might persist to late adolescence. However, these are modifiable risks with small or moderate effect sizes, (in the absence of severe or chronic maternal mental disorder). There is also growing evidence suggesting that paternal mental health is also associated with the offspring's negative outcomes. The authors suggest that research should prioritise investigating the effectiveness of interventions in reducing risk to the child and reducing symptoms in the affected parent. Quality of parenting by both parents was considered to be one of the most important mediators of the effect of perinatal mental disorders on children's outcomes. Understanding how the quality of parenting and specific parenting behaviour leads to different child outcomes is important and should be targeted as a key modifiable risk factor. Identifying and evaluating interventions that can bring about a change in parenting behaviour in the context of the family's circumstances is an important future focus of research.

Summary of Findings Relevant to Criterion 1: Criterion met‡‡

There is a large volume of evidence including some high quality meta-analyses and systematic reviews about the adverse outcomes associated with mental health problems experienced by women in pregnancy and postpartum that is applicable to the UK population. Some adverse outcomes, such as preterm birth and low birth weight were consistently reported, whilst other outcomes, such as the effect of postnatal depression or depression in adolescence were less consistently found across studies. This may be due to the heterogeneity of the individual study designs used to measure these outcomes as described in the meta analyses and systematic reviews of the studies. For these adverse outcomes with little or inconsistent evidence, further research may change the understanding of the impact of perinatal mental health problems on them.

Criteria 4 and 5

There should be a simple, safe, precise and validated screening test.

‡‡ **Met** -for example, this should be applied in circumstances in which there is a sufficient volume of evidence of sufficient quality to judge an outcome or effect which is unlikely to be changed by further research or systematic review.
Not Met - for example, this should be applied in circumstances where there is insufficient evidence to clearly judge an outcome or effect or where there is sufficient evidence of poor performance.
Uncertain -for example, this should be applied in circumstances in which the constraints of an evidence summary prevent a reliable answer to the question. An example of this may be when the need for a systematic review and meta-analysis is identified by the rapid review.

The distribution of test values in the target population should be known and a suitable cut off level defined and agreed.

This section aims to evaluate the effectiveness of screening tests to identify mental health problems in women during the antenatal and postnatal periods with a view to find a suitable screening test that can be used in the context of a national population based screening programme.

Question 2 – What is the reported accuracy of screening tools to detect common mental health problems during pregnancy?

The question of the accuracy of screening tools to detect common mental health problems during pregnancy was not formally addressed by the UK NSC in previous reviews.

Eligibility for inclusion in the review

The inclusion criteria for this question are outlined below:

Population: Pregnant women without previous clinical diagnosis of mental health problems

Intervention: Screening tools used during pregnancy to detect common mental health disorders, including depression, GAD, panic disorder, phobias, social anxiety disorder, OCD and PTSD.

Comparator: Diagnosis confirmed with clinical interview

Outcomes:
Test parameters:
- sensitivity
- specificity
- false positive rate
- false negative rate
- positive Predictive Value (PPV)
- negative Predictive Value (NPV)

Study design:
- inclusions: Consecutively enrolled populations and systematic reviews of these.
- exclusions: Case reports, case series, non-systematic reviews, case control studies, non-peer reviewed literature.

Date and language: English peer reviewed publications from 2006 onwards.

Description of the evidence

Database searches yielded 746 results, of which 162 were judged to be relevant to this question. The search carried out for question 3 identified 12 systematic reviews and 150 publications describing individual studies. In order to identify the best available publications systematic reviews and meta-analyses were prioritised over individual studies.

After reviewing the abstracts 22 studies met the criteria for full text review and following review of the full texts, 5 studies were included. Seventeen articles did not meet the eligibility criteria for this review and the reasons for exclusion are outlined in Table 14 in Appendix 2.

Summary of findings

A study-level summary of data extracted from each included publication is presented in 'Summary and appraisal of individual studies' (Appendix 3) and are stratified by question. Table 4 summarises the data reported by the 5 publications and includes the scale used, the cut-off points, sensitivity, specificity, PPV and NPV. The reference standard used in all but one publication[28] was the structured clinical interview for the diagnostic statistical manual of mental disorders (4th edition).

Table 4. Summary of included studies for key question 2

Screening test	Antenatal mental health condition	Test cut off	Sensitivity	Specificity	PPV	NPV	Study systematic reviews (SR) are indicated
EPDS	Major depression	≥14	1.00	0.94	0.50	NR	O'Connor et al (2016)[26] SR
		≥13	1.00	0.87	0.33	NR	O'Connor et al (2016)[26] SR
		≥12	1.00	0.79	0.23	NR	O'Connor et al (2016)[26] SR
		≥10	0.80	0.83	NR	NR	Owora et al (2016)[27] SR
		≥10	0.94	0.87	NR	NR	Owora et al (2016)[27] SR
EPDS	Depression	≥14	0.57	0.95	0.66	NR	O'Connor et al (2016)[26] SR
		≥13	0.64	0.90	0.50	NR	O'Connor et al (2016)[26] SR
		≥12	0.64	0.80	0.35	NR	O'Connor et al (2016)[26] SR
		≥11	0.71	0.72	0.29	NR	O'Connor et al (2016)[26] SR
		13	0.59	0.94	0.52	0.95	Howard L, et al (2018)[30]
		12	0.68	0.92	0.48	0.96	Howard L, et al (2018)[30]
		11	0.73	0.88	0.40	0.97	Howard L, et al (2018)[30]
EPDS	Any MH condition	12/13	0.30	0.95	0.67	0.79	Howard L, et al (2018)[30]
PHQ8	Major depression	≥11	0.77	0.68	NR	NR	O'Connor et al (2016)[26] SR
		≥10	0.77	0.62	NR	NR	O'Connor et al (2016)[26] SR
PHQ9	Major depression	≥10	0.86	0.84	NR	NR	Owora et al (2016)[27] SR
PHQ-2 (L)	Major depression	≥4	0.62	0.79	NR	NR	O'Connor et al (2016)[26] SR
		≥3	0.77	0.59	NR	NR	O'Connor et al (2016)[26] SR
PHQ2 (N)	Depression	≥1	1.00	0.68	NR	NR	O'Connor et al (2016)[26] SR
		1	0.41	0.95	0.45	0.93	Howard , et al (2018)[30]
		2	0.25	0.98	0.61	0.84	Howard , et al (2018)[30]
PHQ2 (N)	Any MH condition	1	0.23	0.96	0.66	0.78	Howard , et al (2018)[30]
		2	0.06	0.99	0.78	0.66	Howard , et al (2018)[30]
PHQ2 (N)	vs HADS for depression	≥1	0.81(0.74-0.88)	0.79(0.77-0.81)	0.19(0.16-0.21)	0.99(0.98-0.99)	Vlenterie , et al (2017)[28]
PHQ2 (N)	vs EPDS for depression	≥1	0.84(0.77-0.91)	0.84(0.82-0.85)	0.21(0.17-0.25)	0.99(0.98-0.99)	Vlenterie , et al (2017)[28]
PHQ2 (N)	vs HADS for depression	≥1	0.69(0.62-0.76)	0.79(0.77-0.81)	0.26(0.23-0.30)	0.96(0.94-0.97)	Vlenterie , et al (2017)[28]
GHQ30	Anxiety	5/6	0.77	0.78	0.52	0.96	Meades et al (2011)[29] SR

EPDS – Edinburgh Postnatal Depression Score, PHQ 8 Patient Health Questionnaire 8 question version PHQ 9 – Patient Health Questionnaire 9 questions version PHQ2(L) Patient Health Questionnaire 2 question version Likert scale, PHQ2(N) – Patient Health Questionnaire 2 question version nominal yes/no scale, GHQ 30 – General Health Questionnaire 30 question version, MH – mental health

The 5 papers included to examine this question were 3 systematic reviews and 2 studies published after the systematic reviews covering 6 tests. The included publications focus on scales to detect depression, anxiety and stress. Two systematic reviews and one study(O'Connor et al (2016)[26], Owora et al (2016)[27], Vlenterie et al 2016)[28] looked at depression; one systematic review looked at anxiety (Meades et al 2011)[29], and one study at any mental health condition (Howard et al 2018)[30]. Abstracts of studies reported by each of the systematic reviews were checked for relevance to a UK population in respect of:

- use of an English language version of the questionnaire
- participants from the community or primary care settings
- participants were from the general population and not a specific sub group (eg; a particular ethnicity or deprivation level or those with known mental health problems).

All 3 were well conducted systematic reviews O'Connor et al (2016)[26], Owora et al (2016)[27] and Meades et al (2011)[29] undertook quality assessments of the studies included in their publications. On the whole the studies reported varied widely in the instruments used, cut offs chosen, method of administration and reporting of results of performance characteristics thus precluding meta-analyses. The quality assessment performed by O'Connor et al (2016)[26] in relation to the performance characteristics of the EPDS identification tool indicated that the studies were in general small, with limited and often old data on the English-language version (8 English-language versions), and with 2 of the 8 English-language version studies not reporting the interval between EPDS and reference standard. The number of observations, training and fidelity associated with the reference standard were rarely reported. Other possible causes of bias was the inconsistent use of optimal cut-offs. However, this was not the case for the majority of English-version EPDS studies which reported a cut-off. None of the studies examined by Owora et al (2016)[27] looked at the potential for confounding or effect modification by study participant characteristics on case-finding instrument diagnostic performance. Only a third of the included studies were judged to be of good overall quality with the rest having a fair quality rating. The studies were in general small with a number of participants ranging from 32 to 1011. The author of the 2011 systematic review described the majority of the included studies (73%) as of good quality. In total, O'Connor et al (2016)[26] looked at 26 studies with 23 of these reporting results from EPDS (a total of 5,398 participants, of which 100 were in the antenatal period using English language EPDS and 843 used non-English language EPDS). The remaining 3 studies reported PHQ accuracy (a total of 777 participants, of which 339 were in the antenatal period using English language PHQ). Owora et al (2016)[27] included 14 articles with data on 21 case finding instruments for a total of 6,716 participants. Of these 4 studies using 6 case finding instruments reporting a total of 3421 test results were in the antenatal period (total number of women uncertain as some carried out more than one test on more than one occasion). Meades et al (2011)[29] looked at 30 studies with 21 case finding instruments for a total of 7,872 participants. Of these 1 study

from the UK has been included here that tested the GHQ 30 in 179 women from 2 general practices. The remaining studies were testing scales in postnatal populations, did not report sensitivity, specificity PPV and NPV or were using non-English language versions of the scales.

The two recent studies included in this review, Howard et al (2018)[30]**Error! Bookmark not defined.** (545 pregnant women), and Vlenterie et al (2017)[28]**Error! Bookmark not defined.** (3033 pregnant women) were checked against the QUADAS-2 checklist which showed low bias for all areas apart from in Vlenterie et al (2017)[28] where authors used another self-report instrument as a reference standard rather than a structured diagnostic interview.

In general, the number of participants included in the studies was low. These studies are difficult to compare because of the wide range of statistical data reported across them and the substantial heterogeneity between them, such as study design, population sampled and diagnostic criteria used (see Appendix 3).

Edinburgh Postnatal Depression Scale (EPDS)

Overall there was agreement amongst the study authors that the EPDS was a useful and reasonable instrument to use as a way to determine likely major depression disorder during pregnancy.

Two systematic reviews from 2016[26,27] and 1 study from 2018[28] reported the performance characteristics of the EPDS. The systematic review by O'Connor et al (2016)[26] focussed on the use of the EPDS for antenatal and postnatal depression which formed the basis of the US Preventative Task Force (USPTF) policy about screening for depression in pregnancy and postpartum. Owora et al (2016)[27] reported data from studies focussed only on major depression disorder (MDD) during pregnancy and postpartum. The last paper by Howard et al (2018)[30] is a UK study looking at 2 screening tools and the recommended strategy by NICE (2014)[1] that suggests asking the 2 Whooley questions and then if there is an affirmative response to either of them, using the EPDS. The study looks at how the instruments work together in practice in the UK to detect depression, and any other mental health problems. All studies reported used a form of structured diagnostic interview as a reference standard.

O'Connor et al (2016)[26] reported 1 study using the English version of the EPDS during pregnancy that tested a range of cut-off points from 11 to 14 and Owora et al (2016)[27] reported a further study using a cut-off of 10. Howard et al (2018)[30] reported cut-offs from 11 to 15. For MDD cut-offs from 10 to 14 all had good sensitivities and specificities (sensitivity ranged from 0.80 to 1.00 and specificity 0.79 to 0.94) with the highest PPV

(0.50) at a cut-off of ≥14. For any form of depression, a pattern of higher sensitivities at lower cut-offs (0.73 and 0.71 at cut-off ≥11 versus 0.64 and 0.59 at ≥13) and lower PPVs at lower cut-offs (0.29 and 0.40 at ≥11 versus 0.50 and 0.52 at ≥13) was observed.

However, for detecting any mental health problem, the sensitivity of EPDS at a cut-off of 12/13 was lower at 0.30, specificity was 0.95, PPV was 0.67 and NPV was 0.79[30].

Owora et al (2016)[27] reported a pattern of higher diagnostic performance of the EPDS in detecting MDD during the second and third trimesters than during the first trimester across different studies.

Patient Health Questionnaire

Two systematic reviews from 2016[26,27] and 1 study from 2017[28] looked at the performance of the Patient Health Questionnaire (PHQ). A number of versions of the validated PHQ have been developed. Publications included here report on 4 different versions of the scale, the PHQ 9, PHQ 8, PHQ 2 (Likert scale) and PHQ 2(nominal 'yes/no' scale). In the UK the PHQ 2 'yes/no' scale is also known as the Whooley questions. Overall, PHQ-9 had good sensitivity and specificity for MDD and both versions of PHQ-2 had good test performance for depression but not for other common mental health conditions.

For PHQ 8 and 9 studies focussed only on MDD and cut-offs of ≥10 or ≥11. Sensitivities and specificities were good (0.77 to 0.86 and 0.62 to 0.84 respectively), but no PPVs or NPVs were reported so the proportion of false positives and false negatives is unknown. (Owora et al 2016 and O'Connor et al 2016)[26,27].

For PHQ 2 (Likert) cut-offs of ≥3 and ≥4 had a sensitivity of 0.77 and 0.62 and specificity of 0.59 and 0.79 respectively (O'Connor et al 2016)[26]

For PHQ (nominal 'yes/no') scale, two studies and 1 systematic review reported performance characteristics. Howard et al (2018)[30] and O'Connor et al (2016)[26] found major or minor depression sensitivity ranged widely from 0.25 to 1.00, specificity from 0.68 to 0.98, PPV from 0.45 to 0.61 and NPV from 0.84 to 0.93. Vlenterie et al (2017)[28] used either EPDS or the Hospital Anxiety and Depression Scale (HADS) as a reference standard and reported sensitivities of 0.69 to 0.81, specificities of 0.79 to 0.84 and NPV of 0.96 to 0.99, but low PPVs (0.19 to 0.26).

General Health Questionnaire

One systematic review (Meades et al 2011)[29] reported performance of the General Health Questionnaire (GHQ 30), administered to detect anxiety during pregnancy, with data obtained from 1 small study in the UK in 2 general practices with 179 women of which 62 were high scorers and 18 were confirmed cases by diagnostic interview and which produced a sensitivity of 0.77, specificity of 0.78, PPV of 0.52 and NPV of 0.96.

Question 3 - What is the reported accuracy of screening tools to detect postnatal depression?

Sub question - When is the optimum timing to perform the screening test?

In 2011 the UK NSC's evidence update focussed on a Health Technology Assessment (HTA) by Hewitt et al (2009)[31] that addressed the question of screening instruments for postnatal depression. In the HTA, the EPDS was found to have a sensitivity that ranged from 0.60 to 0.96 and specificity 0.45 to 0.97 for major depression; and a sensitivity of 0.31 to 0.91 and a specificity of 0.67 to 0.99 for major and minor depression respectively. For any mental health problem the sensitivity was 0.38 to 0.86 and specificity of 0.87 to 0.99. The conclusion of the HTA[31] was that the criterion had been met for the EPDS as an instrument to screen for postnatal depression, but noted a lack of evidence for a full screening, referral and management strategy. It also found no evidence to support the use of the Whooley questions as recommended by NICE guidance[31]. The UK NSC's evidence update in 2011[8] did not address the question of timing of screening for PND. The 2011 UK NSC's evidence update; however, looked at whether there was evidence from high quality RCTs of screening programmes which included evidence that the tests accurately measures risk. Evidence from the Hewitt et al (2009)[31] HTA showed that there was insufficient evidence to conclude that identification strategies are effective in improving maternal and infant outcomes. They noted that there was some suggestive evidence indicating that the EPDS, possibly with some enhancement of care, may lead to reductions in the number of women with EPDS scores above a certain threshold, or reductions in EPDS scores. The 2011 evidence update concluded that to identify the optimal screening tools for PND more research was required in order to compare the performance of the Whooley questions with the EPDS, and a generic depression measure.

Eligibility for inclusion in the review

The inclusion criteria for this question are outlined below:

Population: Postpartum women

Intervention: Screening tools used with women postpartum to detect postpartum depression

Comparator : Placebo or no active intervention or usual care

Outcomes:
Test parameters:

- sensitivity
- specificity
- false positive rate
- false negative rate
- positive Predictive Value (PPV)/
- negative Predictive Value (NPV)

Study design:
- Inclusions: Consecutively enrolled populations and systematic reviews of these.
- Exclusions: Case reports, case series, non-systematic reviews, case control studies, non-peer reviewed literature.

Date and language: English language peer reviewed publications from 2011 onwards.

Description of the evidence

Database searches yielded 746 results, of which 82 were judged to be relevant to this question. The search carried out for question 4 identified 7 systematic reviews and 75 publications describing individual studies. The results of systematic reviews and meta-analyses were prioritised over individual studies. The included publications focus on scales to detect depression, anxiety, panic disorder and stress.

After reviewing the abstracts 17 studies met the criteria for full text review and following review of the full texts, 3 studies were included looking at depression disorder. Table 14 in Appendix 2 outlines the reasons for exclusion of 14 articles.

Summary of findings

A study-level summary of data extracted from each included publication is presented in ‘Summary and appraisal of individual studies’ (Appendix 3) and are stratified by question.

The 3 papers included to examine this question include 2 systematic reviews O’Connor et al (2016)[26], Owora et al (2016)[27] and 1 more recent primary study Gollan et al (2017)[32] (n = 15,172) covering a total of 9 scales. As mentioned above, the 2 systematic reviews undertook quality assessments of the studies included in their publications reporting an inconsistency between the studies that did not permit to conduct meta-analyses. The primary study is a secondary analysis of a dataset from the identification and therapy of a postpartum depression study. The limitation of this study is its two-phase design in which the clinical interview was designated as the gold standard in identifying depression, and the

use of a phone-based approach as identification instrument which might have influenced the reliability of the responses.

In total O'Connor et al (2016)[26] looked at 26 studies with 23 studies reporting results from EPDS (for a total of 5,398 participants, of which 1,805 were in the postpartum period using English language EPDS and 2,650 used non-English language EPDS). The remaining 3 studies reported PHQ accuracy (for a total of 777 participants: of these, 438 were in the postpartum period using English language PHQ). Owora et al (2016)[27] included 14 articles with data on 21 case finding instruments for a total of 6,716 participants. Of these 10 studies using 18 case finding instruments reporting a total of 24,615 test results were in the postnatal period (total number of women uncertain as some carried out more than one test on more than one occasion).

Table 5 summarises the data reported by the 3 publications and includes the scale used, the cut-off points, sensitivity, specificity, PPV and NPV where available. Abstracts of studies reported by each of the systematic reviews were checked for relevance to a UK population in respect of:

- use of an English language version of the questionnaire
- participants from the community or primary care settings
- participants were from the general population and not a specific sub group (eg; a particular ethnicity or deprivation level or those with known mental health problems)

Overall, the quality of the evidence was good although some of the systematic patterns or variation in diagnostic performance observed among instruments may be explained by either the complex pathophysiology of depressive symptoms among mothers, or social desirability bias[§§]. Methodological heterogeneity in how studies administered a wide variety of instruments, varying cut offs, patient characteristics and reference standard used, means it was not possible for the authors to combine results using meta-analyses. In one systematic review a higher prevalence of MDD was associated with a lower sensitivity of any instrument which may indicate higher surveillance for MDD (spectrum bias).

[§§] Social desirability bias is the tendency of some respondents to report an answer in a way they feel to be more socially acceptable than what they consider is the 'true' answer.

Table 5. Summary of included studies for key question 3

Postnatal screening test	Condition	Screening test cut off	Sensitivity range	Specificity range	PPV range	NPV range	Study (the systematic reviews (SR) are indicated)
BDII-I	Major depression	11	0.80	0.63	NR	NR	Owora et al (2016)[27] SR
	Major depression	12	0.78	0.69	NR	NR	Owora et al (2016)[27] SR
	Major depression	13	0.78	0.73	NR	NR	Owora et al (2016)[27]**Error! B ookmark not defined.** SR
	Major depression	14	0.60-0.80	0.80-0.84	NR	NR	Owora et al (2016)[27] SR
	Major depression	15	0.69	0.81	NR	NR	Owora et al (2016)[27]SR
	Major depression	20	0.45-0.55	0.91-1.00	NR	NR	Owora et al (2016)[27] SR
EPDS 10	Depression	≥14	0.83	0.66	NR	NR	Gollan et al (2017)[32] SR
	Major depression	≥13	0.75-0.95	0.70-0.99	0.24 – 0.93	0.96	O'Connor et al (2016)[26] SR
	Major depression	≥13	0.55-0.67	0.81-0.91	NR	NR	Owora et al (2016)[27] SR
	Major depression	≥12	0.88-0.92	0.62-0.76	0.2	NR	O'Connor et al (2016)[26] SR
	Major depression	≥12	0.74-0.76	0.73-0.99	NR	NR	Owora et al (2016)[27] SR
	Major depression	≥11	0.88	0.73	0.18	NR	O'Connor et al (2016)[26] SR
	Major depression	≥10	0.88-1.00	0.71-82	0.17	NR	O'Connor et al (2016)[26] SR
	Major depression	≥10	0.61-0.80	0.66-0.89	NR	NR	Owora et al (2016)[27] SR
	Major depression	≥9	0.71-0.85	0.73-0.85	NR	NR	Owora et al (2016)[27] SR
	Depression	≥13	0.79	0.76	NR	NR	O'Connor et al (2016)[26] SR
	Depression	≥12	0.76-0.86	0.67-0.81	0.33-0.44	NR	O'Connor et al (2016)[26] SR
	Depression	≥11	0.76	0.79	0.41	NR	O'Connor et al (2016)[26] SR
	Depression	≥10	0.81	0.77	0.41	NR	O'Connor et al (2016)[26] SR
	Depression	≥9	0.59	0.86	0.64	0.82	O'Connor et al (2016)[26] SR
EPDS 7	Depression	≥8	0.76	0.68	NR	NR	Gollan et al (2017)[32] SR
IDAS-GD	Major depression& anxiety	51	0.63	0.81	NR	NR	Owora et al (2016)[27] SR
	Major depression & anxiety	44	0.81	0.65	NR	NR	Owora et al (2016)[27] SR
	Major depression& anxiety	39	0.93	0.55	NR	NR	Owora et al (2016)[27] SR
PDSS	Major depression	80	0.67-0.92	0.68-0.97	NR	NR	Owora et al (2016)[27] SR
PHQ 9	Major depression	≥10	0.75	0.91	0.28	0.99	O'Connor et al (2016)[26] SR
	Major depression	10	0.32-0.80	0.09-0.91	NR	NR	Owora et al (2016)[27] SR
	Major depression	12	0.99	1.00	NR	NR	Owora et al (2016)[27] SR
PHQ-2 (N)	Major depression	≥1	1.00	0.62	0.11	1.00	O'Connor et al (2016)[26] SR
PHQ2 (L)	Major depression	≥2	0.75	0.88	0.24	0.99	O'Connor et al (2016)[26] SR
	Major depression	3	0.75-0.84	0.88-0.79	NR	NR	Owora et al (2016)[27] SR
	Major depression	1	0.98-0.99	0.44-0.62	NR	NR	Owora et al (2016)[27] SR
PRAMS6	Major depression	17	1.00	0.84	0.81	NR	Owora et al (2016)[27] SR
	Major depression	15	0.80	0.47	0.63	NR	Owora et al (2016)[27] SR

Beck Depression Inventory-II (BDI-II)

BDI-II is a commonly used scale to measure severity of depression. Owora et al (2016)[27] included 4 studies examining BDI-II performance with postpartum women to detect major depression disorder cut offs between 11 and 20. At a cut-off of 14 studies reported the best performance for the test with a sensitivity ranging from 0.60 to 0.80 and specificity 0.80 to 0.84. No PPV or NPV was reported.

Edinburgh Postnatal Depression Scale (EPDS)

Two versions of the EPDS, the full version with 10 questions (EPDS 10) and the cut down version, EPDS 7 (with 7 questions), were evaluated. Owora et al (2016)[27] included studies just screening for MDD whilst O'Connor et al (2016)[26] reported a range of studies for MDD and minor depression using EPDS 10. Gollan et al (2017)[32] examined performance of both EPDS 10 and EPDS 7 in postpartum women for postnatal depression and reported a sensitivity of 0 83 vs 0.76 and a specificity of 0.68 vs 0.66 respectively (PPV and NPV were not reported).

Inventory of Depression and Anxiety Symptoms-General Depression (IDAS-GD)

The IDAS-GD is a scale designed to assess specific symptom dimensions of major depression and related anxiety disorders. Owora et al (2016)[27] included 1 study reporting 3 cut offs with sensitivities ranging from 0.63 to 0.93 and specificities from 0.55 to 0.80. No PPV or NPV was reported.

Panic Disorder Severity Scale (PDSS)

The PDSS is a 7-item instrument developed to rate overall severity of panic disorder. Owora et al (2016)[27] reported that the PDSS (diagnostic threshold/cut-point scores: 50 to 80) showed least variation and highest diagnostic performance across different postpartum periods and diagnostic thresholds (sensitivity range: 0.67 to 0.95 and specificity range: 0.68 to 0.97) compared to other tests.

Patient Health Questionnaire

A number of versions of the validated Patient Health Questionnaire have been developed. Publications included here report on 3 different versions of the scale, the PHQ 9, PHQ 2 (Likert scale) and PHQ 2 (nominal 'yes/no' scale). Overall sensitivities and specificities were good but where PPVs were reported these were low. For example, at a cut off of ≥10 for PHQ 9, sensitivity was 0.75, specificity 0.91, PPV 0.28 and NPV 0.99. Similarly, the

PHQ 2 nominal version of the scale recommended by NICE (2014)[1] as routine questions posed to women postpartum had a sensitivity of 1.00, specificity of 0.62, PPV of 0.11 and NPV of 1.00.

Pregnancy and Risk Assessment Monitoring System (PRAMS6)

The PRAMS6 is a questionnaire comprising 6 items covering both depression and anxiety symptoms common to postpartum women. One study reported by Owora et al (2016)[27] with 1011 participants reported a sensitivity of 1.00 and specificity of 0.84 and a PPV of 0.81 at a cut off of 17.

Timing of screening

Only one study, (Owora et al 2016)[27], considered timing of screening in the postnatal period and reported a pattern of lower and more variable diagnostic performance during the first 2 months postpartum than later for all instruments where performance at different times could be compared (EPDS 10, PRAM6, PHQ9, PDSS).

Summary of Findings Relevant to Criteria 4 and 5: Criteria not met***

The most extensively studied instruments used to detect mental health problems in pregnant women in the 3 good quality systematic reviews and 2 primary studies were the EPDS and PHQ scales.

Antenatal period
The majority of the evidence found looked at the accuracy of screening tools to detect antenatal depression rather than, more broadly, common mental health problems during pregnancy.

EPDS was considered by study authors to be a reasonable scale to screen for antenatal depression and the psychometric properties varied least across studies. One study reported that the EPDS did not perform well in detecting other mental health problems apart from antenatal depression. The performance of the PHQ scales also varied

*** **Met** -for example, this should be applied in circumstances in which there is a sufficient volume of evidence of sufficient quality to judge an outcome or effect which is unlikely to be changed by further research or systematic review.
Not Met - for example, this should be applied in circumstances where there is insufficient evidence to clearly judge an outcome or effect or where there is sufficient evidence of poor performance.
Uncertain -for example, this should be applied in circumstances in which the constraints of an evidence summary prevent a reliable answer to the question. An example of this may be when the need for a systematic review and meta-analysis is identified by the rapid review.

substantially across studies depending on number of questions and cut-off threshold. However, the total body of evidence about detecting mental health problems during pregnancy is based on studies with low number of participants, and in general the accuracy of data reported across studies is variable. Substantial heterogeneity between the studies was also noted by the systematic reviews' authors in relation to study design, population sampled and the diagnostic criteria used. This precluded the use of meta-analyses to combine results.

Postpartum period

While the data included in this evidence summary in relation to postpartum depression comes from studies with larger numbers of participants compared to studies in the antenatal period, the number remains low when considering screening in whole populations. Moreover, these studies suffer from the same heterogeneity problems noted in those in the antenatal period.

There are a wide range of instruments, cut-off thresholds and methods used to detect postpartum depression reported in the 2 good quality systematic reviews and 1 study included in this review. The performance of the EPDS, PHQ 9 and 2 were the most widely reported scales. Both the EPDS and PHQ scales had good sensitivities and specificities. A positive screen would necessitate a comprehensive psychosocial assessment to identify the type and severity of the problem. The reported PPVs indicate a high proportion of women who would turn out not to have postnatal depression after undergoing such in-depth psychosocial assessment. The timing of administering the EPDS following birth indicated less variable performance after 2 months postpartum.

Overall the quantity, quality, consistency and applicability of the evidence indicates that there are screening tools (EPDS, PHQ 2) that could be used as part of an overall screening strategy that would identify women with antenatal and postpartum depression. However, this evidence is based on small cohorts and therefore larger studies looking at the accuracy of such identification tools are needed. Larger studies would also provide a better estimate on the proportion of true positive and true negative results, which would help to ascertain if such tools are susceptible to over-detection. There is also limited evidence that these or other tools are effective in screen detecting women, with GAD, panic disorder, phobias, social anxiety disorder, OCD and PTSD besides depression, during pregnancy.

Criterion 9

There should be an effective intervention for patients identified through screening, with evidence that intervention at a pre-symptomatic phase leads to better outcomes for the screened individual compared with usual care. Evidence relating to wider benefits of screening, for example those relating to family members, should be taken into account where available. However, where there is no prospect of benefit for the individual screened then the screening programme should not be further considered.

Question 4 – What are the benefits of pharmacological and non-pharmacological interventions (alone or in combination) used by women with screen-detected mental health problems during pregnancy?

Sub question – How do pharmacological and non-pharmacological interventions (alone or in combination) affect maternal pregnancy outcomes?

Sub question – How do pharmacological and non-pharmacological interventions (alone or in combination) used by the mother affect outcomes in babies?

The issue of screening to detect common mental health problems during pregnancy was not formally addressed by the UK NSC in previous reviews.

Eligibility for inclusion in the review

Population: Pregnant women with screen-detected common mental health problems during pregnancy.

Intervention: Pharmacological and non-pharmacological interventions used in pregnancy.

Comparator: Placebo (for pharmacological) or no active intervention.

Outcomes: The specific outcomes as an effect of an intervention:
Maternal

- danger to self or infant
- increased risk of subsequent depression
- puerperal psychosis
- increased risk of subsequent anxiety disorders
- increased risk of subsequent personality disorders
- substance misuse
- eating disorders
- mother-infant interaction problems

- quality of life
- adherence or persistence with treatment.

Baby

- preterm birth
- small for gestational age
- large for gestational age
- low birth weight
- neonatal Intensive Care Unit admission
- growth and development up to 1 year of age.
- mealtime conflict

Child

- emotional and behavioural adjustment
- social adjustment
- cognitive development

Adolescent

- depression

Study design:

Inclusions:

- randomised and quasi-randomised controlled trials and observational studies (cohort studies, and cross-sectional studies) or systematic reviews of these

Exclusions:

- case reports, case series, non-systematic reviews, case-control studies, non-peer reviewed literature

Date and language: English peer reviewed publications from 2006 onwards.

Description of the evidence

Database searches yielded 746 results, of which 116 were judged to be relevant to this question. The search carried out for question 5 identified a large volume of evidence describing interventions for common mental health problems in women who are pregnant reporting outcomes either for women themselves or their subsequent offspring. This evidence included 19 systematic reviews and 97 publications describing individual studies. The results of systematic reviews and meta-analyses were prioritised over individual studies.

After reviewing the abstracts, 19 studies met the criteria for full text review. Of those 19 publications, 1RCT (Grote et al 2016)[33], 1 quasi-experimental trial (Weinreb et al 2017)[34] and 1 systematic review (O'Connor et al 2016)[26] were included that targeted screen-

detected cohorts of women with postpartum depression or PTSD. To provide a broader description of the evidence available about treatments for pregnant women with common mental disorders (in whom the mode of detection was uncertain) an overview of 4 recent systematic reviews has been included. These describe:

- non-pharmacological interventions and maternal outcomes(Van Ravesteyn et al 2017)[35]
- non-pharmacological interventions and other outcomes(Letourneau et al 2017)[36]
- pharmacological interventions and maternal outcomes(McDonagh et al 2014)[37]
- pharmacological interventions and other outcomes(McDonagh et al 2014[37], O'Connor et al 2016[26], Prady et al 2017[38])

Twelve articles did not meet the eligibility criteria for this review and the reasons for exclusion are outlined in Table 14 in Appendix 2.

Summary of findings

A study-level summary of data extracted from each included publication is presented in 'Summary and appraisal of individual studies' (Appendix 3) and are stratified by question.

The evidence base is comprised mostly of studies reporting results of non-pharmacological interventions offered during pregnancy. There are many fewer studies reporting results about the benefits of the use of antidepressants during pregnancy. Even when including some broader systematic reviews of studies of non-screen detected women there was considerable heterogeneity across trials and inconsistent results making it difficult to make judgements about which therapies are most effective for women with common mental health disorders during pregnancy.

Trials that were reported were on the whole small and limited to depressive conditions with less emphasis on anxiety, and other common mental health disorders.

Non-pharmacological interventions

Results on the effectiveness of the non-pharmacological intervention in the antenatal period are detailed in Table 6.

Table 6. Summary of included studies for key question 4 in relation to non-pharmacological interventions

Paper	Condition	Treatment	Outcomes	Number	Study outcome	Quality appraisal
Screen-detected population						
Weinreb et al (2017)[34] **Quasi – experimental**	Screen-detectedPTSD	Psychosocial educational sessions	Severity of PTSD, maternal depression, positive and negative coping skills	149 women	CBT based sessions for PTSD – change in severity of symptoms, effect size Cohen's *d**; 0.29, p=0.032	The JBI tool for assessing bias in quasi experimental trials was used to test for quality of this trial and there were the following concerns: • Group equivalence and confounding between groups: At baseline compared to the comparison group the intervention group had a significantly higher proportion of women who: o Used tobacco(p<0.002) o Used drugs in the past year (p<0.02) o Used negative coping strategies(p<0.02) o Ever had an abortion(p=0.041) o Spoke English (p<0.02) o Were born in the USA or Puerto Rico (p<0.001) • Attrition bias: o Of the 149 participants at baseline, 128 (86%) remained for the postpartum interview, including 72 (81%) of intervention and 56 (93%) of control participants ($\chi 2 = 8.71$, df = 1, p = .005). o There was a trend for those who were lost to postpartum follow-up to have a higher baseline PTSD symptom score (t147 = 1.84, p = .07), and a trend for them to be part of the intervention condition ($\chi 2 = 4.58$, df = 1, p = .053).
O'Connor et al (2016)[26] Systematic review	Screen-detected depression	CBT	Antenatal depression	3 studies n=25 to 181	3 trials were of fair quality Effect sizes (relative risk) favoured the intervention of: 1.55 (95%CI,0.98 to 2.44) 1.25 (95%CI,0.69 to 2.27) the 1.36 (95% CI,1.13 to 1.65)	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns. The authors of the systematic review used criteria developed by the US Preventive Services Task Force for randomized clinical trials,[39] the Newcastle-Ottowa Scale[40] for observational studies, and the AMSTAR-2[41] for systematic reviews. The studies included in the review in relation to the treatment of women for depression in pregnancy were judged by the reviewers to be mainly fair. All studies were small with possible bias due to the different definitions of remission.
Grote et al (2016)[33] **RCT**	Screen-detecteddepression with PTSD		Maternal depression	164 women with MDD or dysthymi	Women with MDD and PTSD responded more to MOMCare collaborative intervention compared with standard care at 18 months with group differences for:	Cochrane collaboration risk of bias tool for RCTs was used to assed the quality of this study. Low bias was noted for all questions apart from r 'Blinding of participants and researchers(performance bias)' which was considered to be high

				a of which 106 also had PTSD	• change in depression severity p<0.004 • PTSD severity p<0.02 • Wald effect size 0.39 favouring MOMCare No difference between intervention groups for women with just depression.	because the participants were not blinded to the intervention and 'Other bias' for which it was unclear.
Not explicitly screen-detected populations						
Letourneau et al (2017)[36] **Systematic review**	Depression and anxiety	CBT, IPT, massage	Parenting and offspring outcomes	7 Studies, 11 publications, (1280) women	Improvement of foetal attachment found with: IPT (p=0.03) Parenting education programme(p=0.007) Social and clinical supports improvement in Depression and Anxiety Stress Scales in treatment group (anxiety, p<0.01, stress, p<0.01). Depression measured with the Beck Depression Inventory (II) lowered (p<0.01), Parental stress measured with the Parental Stress Index was lower in the intervention group than the control (p<0.05).	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns The authors used the Cochrane Collaborations risk of bias tool for RCTs[42] (. A summary assessment of risk of bias was made for each paper with a score of 0-5 being low risk, 6 being moderate risk and 7 upwards high risk. Scores were 6 or below for 9 of 11 publications. Areas of high risk of bias were allocation concealment of the intervention, blinding of participants and personnel and blinding of outcome of assessment.
Van Ravesteyn et al (2017)[35] **Systematic review**	Depression and anxiety	CBT, IPT, Bright light therapy, body oriented intervention	Maternal outcome	29 studies involving 2779 women	**Robust benefit** reducing depressive symptoms was found for: CBT – 8 trials (g = -0.61; 95%CI:-0.73 to -0.49, I^2 = 0%[§§§]; k = 7) (Z***-value = 10.04; p<0.001). IPT- 4 trials (g = -0.67; 95%CI:-1.27 to -0.07; I^2 = 79%;Z-value = 2.20; p = 0.03). **Medium benefit in** reducing depressive symptoms was found for: Body oriented interventions 7 trials (g = -0.43; 95%CI:-0.61 to -0.25; I^2 = 17%;Z-value = 4.62; p<0.001). Acupuncture 2 trials (g = -0.43; 95%CI:-0.80 to -0.06; I^2 = 0%; Z-value = 2.30; p = 0.02). **No benefit** was found in reducing depressive symptoms for:	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns The quality appraisal was performed by the authors of the systematic review for each trial according to the Cochrane Risk of Bias. For 12 trials there were concerns on the procedure of concealment of allocation and randomization. The majority of the trials did not publish a protocol and/or was registered in a trial register. There was good inter-rater agreement with regards to the quality assessment. The quality assessment did not indicate publication bias.

	Bright light therapy 3 trials (g = -0.59; 95%CI:-1.25 to 0.06; I^2 = 0%, Z-value = 0.77; p = 0.08). Food supplements trials 3 trials (g = -0.51; 95%CI:-1.02 to 0.01; I^2 = 20%; Z-value = 1.92; p = 0.06)

* Cohen's d is effect size as a measure of the difference in the two groups' means divided by the average of their standard deviations(d=0.2 is small effect, d=0.5 is medium effect and d=0.8 is a large effect)

Screen-detected populations

The systematic review by O'Connor et al (2016)[26] and trials by Grote et al (2016)[33] and Weinreb et al (2017)[34] were explicitly focussed on screen-detected populations and maternal outcomes. The systematic review looked only at depressive disorders while the two trials looked at women with depression and PTDS.

The effect of CBTon antenatal depression was reported by 3 trials included in O'Connor et al (2016)[26]. All 3 trials were small (n=25 to 181) of fair quality with 1 or more methodological limitations and reported small to moderate effect sizes (relative risk) favouring the intervention (1.55, 95%CI: 0.98 to 2.44, 1.25, 95%CI; 0.69 to 2.27 and 1.36 95%CI; 1.13 to 1.65).

The effect of a psychoeducational intervention on the severity of PTSD in 149 women was examined in a quasi-experimental study reported in Weinreb et al (2017)[34]. A statistically significant difference (p<0.032) and small effect on PTSD severity (Cohen's *d§*; 0.29, where 0.2 is a small effect and 0.5 a medium effect) was reported over time between the control and intervention groups. The control group with usual care showed improvement in severity of PTSD symptoms but the intervention group showed significantly greater improvements. There was no effect on other outcomes including preterm birth, low birth weight, increase in social support, positive coping skills or negative coping skills. At baseline more than half the women also had co-morbid depression but sub analysis of this group showed no improvement in depression symptoms following the intervention. There was a trend for those who were lost to postpartum follow-up to have a higher baseline PTSD symptom score (p =0.07), and a trend for them to be part of the intervention condition (p =0.05).

A small, low bias trial reported by Grote et al (2017)[33] looked at combined non-pharmacological and pharmacological interventions in 164 pregnant women who had a PHQ-9 score of ≥10 during routine antenatal screening. At 18 months women with both MDD and PTSD improved more with the MOMCare collaborative intervention of combined pharmacological and non-pharmacological interventions, compared with standard care, at 18 months with significant group differences for change in depression severity (p<0.004) and PTSD severity (p<0.02). No difference was reported between intervention groups for women with just depression. Using the Wald test††† the effect size favouring MOMCare was larger in women with MDD plus PTSD (0.39) compared to those with just MDD.

††† *The Wald test is a parametric statistical test used whenever a relationship within or between data items can be expressed as a statistical model with parameters to be estimated from a sample, the Wald test can be used to test the true value of the parameter based on the sample estimate(small effect is 0.02-0.149, medium effect is of 0.15 to 0.349 and large effect 0.35 upwards)

Overall there is limited evidence of a small effect in the population of choice in relation to interventions for screen-detected antenatal depression.

Not explicitly screen-detected populations

Four systematic reviews included studies with populations of pregnant women where it was not clear how women with mental health problems were identified (ie there is no description of whether or not they were screen-detected in the systematic review) reported the effect on antenatal depression of a range of interventions. These included, CBT, interpersonal therapy (IPT), body oriented activities, acupuncture, food supplements and bright light therapy in terms of both maternal and other outcomes.

One systematic review (van Ravesteyn et al 2017)[35] found a large effect of CBT (Hedges score, g)‡‡‡ = -0.61; 95%CI,-0.73 to -0.49, I^2 = 0%§§§; Z-value = 10.04; p<0.001) and IPT (g = -0.67; 95%CI,-1.27 to -0.07; I^2 = 79%; Z-value**** = 2.20; p = 0.03) on reducing depressive symptoms in pregnant women and a medium benefit from body oriented interventions such as yoga, tai chi and massage (g = -0.43; 95%CI,-0.61 to -0.25; I^2 = 17%; Z-value = 4.62; p<0.001) and acupuncture (g = -0.43; 95%CI,-0.80 to -0.06; I^2 = 0%, Z-value = 2.30; p = 0.02). No effect was found for bright light therapy or food supplements.

Only one systematic review (Letourneau et al 2017)[36] examined outcomes than following treatment for mental health conditions in pregnancy. Data from 7 trials showed that CBT, IPT and social and clinical support started during pregnancy produced large effects on parenting and maternal mood. Higher levels of maternal-foetal attachment were reported when IPT (p=0.03) and parenting education (p=0.007) was administered, while an increase in attention towards infant distress (p=0.034) was noted when CBT was used. Social and clinical support comprising telephone based counsellors, play groups, mothers groups and self-help work book interventions was associated with a statistically significant improvement in Depression and Anxiety Stress Scales in the intervention group (anxiety, p<0.01, stress, p<0.01). In addition depression measured with the Beck Depression Inventory (II) was lowered (p<0.01) and parental stress measured with the Parental Stress Index was lower in the intervention group than the control (p<0.05).

---

‡‡‡ g, Hedge's g score is very similar to Cohen's d statistic in that it also measures the difference between the means using almost the same method(g=0.2 is small effect, g=0.5 is medium effect and g=0.8 is a large effect)

§§§ I^2, ranging from 0-100%, measures the degree of inconsistency across studies in a meta-analysis with 0% = no heterogeneity.

**** Z score enables a comparison of two scores that are from different normal distributions. A Z score of 0 indicates that the means of the two groups are similar, whereas a Z score of 2 indicates the intervention group is 2 standard deviations different to the mean.

Pharmacological interventions

Results on the effectiveness of the pharmacological intervention in the antenatal period are detailed in Table 7.

The use of antidepressants during pregnancy and negative offspring outcomes were reported in all three included systematic reviews[26,38,37] but no study reported the benefits to women or their offspring from pharmacological interventions. None of studies used screening as the method to detect depression in women prior to treatment with antidepressants.

Table 7. Summary of included studies for key question 4 in relation to pharmacological interventions

Paper	Mental health condition of mother	Treatment	Outcomes	Number	Study outcome	Quality appraisal
Not explicitly screen-detected populations						
O'Connor et al (2016)[26] **Systematic review**	Antenatal depression	CBT Anti-depressants	Maternal depression	11 studies Involving 4,759,707 women of which 5 (n=2,068,049) related to outcomes of interest for inclusion in this review.	**Vaginal bleeding during pregnancy or postpartum haemorrhage** Increased risk in depressed women (1 study) SSRI+venlafaxine, current (n=8,917): RR, 1.46 (95% CI, 1.29 to 1.65) SSRI+venlafaxine, recent (n=4,344): RR, 1.28 (95% CI, 1.08 to 1.52) Atypical antidepressant, current (n=1,012): RR, 1.52 (95% CI, 1.12 to 2.06) Controlling for depressed women(1 study) Increased risk: Citalopram, current (n=891): RR, 1.48 (95% CI, 1.07 to 2.04) Escitalopram, current (n=1,022): RR, 1.56 (95% CI, 1.16 to 2.09) Fluoxetine, current (n=3,322): RR, 1.51 (95% CI, 1.27 to 1.79) Paroxetine, current (n=2,055): RR, 1.36 (95% CI, 1.09 to 1.71); recent (n=962): adjusted RR, 1.52 (95% CI, 1.12 to 2.07) Sertraline, current (n=4,526): RR, 1.31 (95% CI, 1.12 to 1.54); recent (n=2,266): RR, 1.27 (95% CI, 1.01 to 1.59) Venlafaxine, current (n=763): RR, 2.24 (95% CI, 1.69 to 2.97) Bupropion, past (n=1,666): RR, 1.33 (95% CI, 1.03 to 1.71) **Miscarriage and/or spontaneous abortion** Increased risk (1 study) Venlafaxine (n=NR): RR, 1.80 (95% CI, 1.19 to 2.72) Duloxetine (n=NR): RR, 3.12 (95% CI, 1.55 to 6.31) Mirtazapine (n=NR): RR, 2.23 (95% CI, 1.34 to 3.70) **Pre-term birth or early gestational age** Increased risk: (1 study) Any antidepressant (mostly SSRIs), % born gestational weeks 32-36:1-2 prescriptions (n=10,700): OR 1.91(95% CI, 1.77 to 2.07)	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns. The authors of the systematic review used criteria developed by the US Preventive Services Task Force[39] for randomized clinical trials, the NOS for observational studies[40], and the AMSTAR for systematic reviews[41]. All studies were observational studies so causality could not be clearly determined. Because in some studies it was not possible to confirm the depression status of some women, this could have exaggerated the potential for confounding by indication. No study reported data examining the harms in relation to the dose; however, some studies looked at harms caused by length of exposure. No reporting bias was noted by the systematic review authors. The overall quality of these studies was assessed to be good

					3+ prescriptions (n=6,196): OR1.12 95% CI, 1.03 to 1.23) Unknown Depression Status in Control Group N=228,876) Increased risk: SSRIs in 2nd trimester (mean difference in days, n=NR for all, nulliparous women): 1 prescription: -2.6 (95% CI, -1.3 to -3.9) 2 prescriptions: -5.8 (95% CI, -3.8 to -7.8) 3+ prescriptions: -6.6 (95% CI,-4.6 to -8.6) Decreased risk: SSRIs in 3rd trimester (mean difference in days, n=NR for all, nulliparous women): 1 prescription: 0.9 (95% CI, 0.3 to 1.6) 2 prescriptions: 1.8 (95% CI, 0.9 to 2.7) 3+ prescriptions: 6.4 (95% CI, 5.5 to 7.3) **Low birth weight or small for gestational age (SGA)** All Women, Controlling for Depression (1 study) Increased risk SSRIs during pregnancy (n=NR): HR, 1.22 (95% CI, 1.13 to 1.32)	
Prady et al (2017)[38] **Systematic review**	Depression and anxiety	Anti-depressants	Offspring outcomes	11	2 studies did not identify any effect of anti-depressants on low birth weight 2 studies reported effect ratios (n=3966, effect ratio1.42, 95% CI 1.16 to 1.7 and n=1622, effect ratio 1.69, 95% CI 1.14 to 2.52); however, neither of these latter studies controlled for severity of depression 1 study with low risk of bias showed no neonate behaviour effect by exposure group 1 study showed a significant difference in average time (additional 13.6 days) to walking unsupported (retrospectively reported). A second study reported a significant increase of 28.9 days to walking unsupported for children of women exposed to anti-depressants in the 2nd and 3rd trimesters of pregnancy. ADHD and comorbid disorders – 1 included study with a higher bias found a statistically	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns. The authors performed a quality appraisal of the studies included in the systematic review using a a modified Newcastle Ottawa Scale[40]The authors were not able to perform a meta-analyses due to a high risk of bias and variation in study design. All the studies scored between 5 and 6 out of 8 on this scale, but none met the authors criteria for low risk of bias.

					significant effect size (11.4 95% CI 1.42 to 91.8) in children aged 3 to 7	
McDonagh et al (2014)[37] **Systematic review**	Depression and anxiety	Anti-depressants	Maternal depression	3 studies	Use of SSRIs and respiratory distress (3 small trials) pooled OR= 1.91; 95% CI, 1.63-2.24 I^2=0%.	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns. The authors assessed the internal validity of the studies based on criteria by McDonagh et al (2012)[43] and graded the evidence using the AHRQ grading methodology[44]. All 3 studies were assessed as of medium quality with concerns about controlling for confounding variables. All 7 studies were assessed as of medium quality with concerns about controlling for confounding variables, small sample sizes and poor reporting of health outcomes and adverse events.

Screen-detected populations

No studies were identified that examined the use of antidepressants in pregnant women who were screen-detected with a mental health condition.

Not explicitly screen-detected populations

The systematic review by O'Connor et al (2016)[26] reported negative outcomes regarding the use of antidepressants during pregnancy from 12 large observational studies which were not limited to women with screen-detected depression (n=4,759,707) of which 8 were good and 1 fair quality cohort studies, and 1 good and 2 fair case control studies.

These studies used second-generation antidepressant†††† agents: selective serotonin reuptake inhibitors (SSRIs), selective norepinephrine reuptake inhibitors, bupropion, nefazodone, trazodone, and mirtazapine. The studies included provided some evidence that use of some antidepressants during pregnancy increased adverse birth and infant outcomes. In particular, SSRIs and venlafaxine were associated with increased risk of postpartum haemorrhage, and miscarriage as well as, preterm birth, small for gestational age, neonatal seizures and neonatal respiratory distress. The absolute increase in risk for most infant outcomes was very small (see table 7), and sometimes occurred only with higher levels of exposure. For example, a large retrospective cohort study reported a more than doubling of seizure occurrence in infants of depressed women who had been provided 3 or more prescription fills for antidepressants of any kind (but primarily SSRIs). However, the absolute risk remained quite small (0.66% among exposed infants vs 0.28% in unexposed infants; moderate unadjusted OR=2.39 (95% CI, 1.57 to 3.64). These studies were all observational and it is not possible to control for all confounders related to depression, especially as women with more severe depression are more likely to take antidepressants during pregnancy.

One systematic review (McDonagh et al 2014)[37] reported neonatal respiratory distress among 7.8% of infants not exposed to SSRIs in utero, compared with 13.9% of exposed infants, and a pooled estimate combining results of 3 medium quality studies (1,547 participants) showed an increased odds of respiratory distress with SSRI exposure (pooled OR=1.91; 95% CI, 1.63 to 2.24 I^2 = 0%)[37] As with the other observational studies reported here it is not possible to control for all confounders related to depression.

A systematic review by Prady et al (2017)[38] reported some short and long term outcomes of offspring whose mothers took (SSRIs) antidepressants during pregnancy for antenatal depression including low birth weight, neurodevelopmental and neuro-behavioural

†††† There are a class of antidepressant identify by the period in which they were introduced (approximately 1970s to 1980s) rather then they pharmacological effects or chemical structure.

outcomes. With respect to low birth, two of the included studies in the review by Prady et al (2017)[38]**Error! Bookmark not defined.** did not identify any effect of antidepressants on low birth w eight whereas another two found a small/moderate/large effect (see table 11) However, neither of these latter studies controlled for severity of depression.

In terms of neurodevelopment and neuro-behaviour, 1 study in Prady et al (2017)[38] showed no effect on neonatal behaviour by exposure group. Another 3 studies reported by Prady et al (2017)[38] with a higher risk of bias investigated infant and toddler development and found conflicting results. One found no differences between groups of infants exposed and not exposed to SSRIs, whilst 2 others reported a significant difference in average time to walking unsupported.

In terms of child behaviour, of 15 outcomes reported across the 3 studies in Prady et al (2017)[38] there was no statistical difference between groups exposed to SSRIs and non-exposure groups. Outcomes of autistic behaviour in children aged 6 was not statistically different between groups of pregnant women exposed to SSRIs and those not exposed. In 1 study of higher bias, use of SSRIs was statistically associated with attention deficit and hyperactivity disorder (ADHD) (11.4 95% CI, 1.42 to 91.8) in children aged 3 to 7.

Question 5 - What are the benefits of early pharmacological and non-pharmacological intervention (alone or in combination) in women with screen-detected postnatal depression?

The 2011 UK NSC's evidence update did not address systematically the question of the benefits of pharmacological or non-pharmacological treatments for women with screen-detected postpartum depression. However, it noted that in their review, Gaynes et al (2005)[45] suggested that targeting major depression alone in screening studies would be a preferable strategy than looking at cohorts with mixed populations with minor and major depression. This is because the evidence on screening tools and on treatment in women with only major depression showed that the tests are more accurate and treatment more effective than in heterogeneous groups. The authors note that, because of the paucity of evidence in this area, and considering the significance and complexity of PND as a mental health and public health problem, it would be necessary to collect evidence from large scale studies to guide national policy given that the available small size studies are not appropriate for this aim[45].

The 2011 UK NSC's review update included also evidence from a cost-effectiveness analysis of routine screening for PND in primary care (Paulden et al 2009)[46] in which the authors noted that there was a lack of reliable evidence from RCTs on the costs of the identification and treatment of PND.

Eligibility for inclusion in the review

Population Women with screen-detected postpartum depression

Intervention: Pharmacological and non-pharmacological interventions

Comparator: Usual care or placebo (for pharmacological) or no active intervention

Outcomes: The specific outcomes as an effect of an intervention:
Maternal

- anger to self or infant
- increased risk of subsequent depression
- puerperal psychosis
- increased risk of subsequent anxiety disorders
- increased risk of subsequent personality disorders
- substance misuse
- eating disorders
- mother-infant interaction problems

- quality of life
- adherence or persistence with treatment.

Baby
- preterm birth
- small for gestational age
- large for gestational age
- low birth weight
- neonatal Intensive Care Unit admission
- growth and development up to 1 year of age.
- mealtime conflict

Child
- emotional and behavioural adjustment
- social adjustment
- cognitive development

Adolescent
- depression

Study design:
Inclusions:
- randomised and quasi-randomised controlled trials and observational studies (cohort studies, and cross-sectional studies) or systematic reviews of these

Exclusions:
- case reports, case series, non-systematic reviews, case-control studies, non-peer reviewed literature

Date and language: English peer reviewed publications from 2011 onwards.

Description of the evidence

Database searches yielded 746 results, of which 165 were judged to be relevant to this question. The search carried out for question 6 identified a large volume of evidence describing interventions for postnatal depression in postpartum women reporting outcomes either for women themselves or their offspring. This evidence included 50 systematic reviews and 115 publications describing individual studies. The results of systematic reviews and meta-analyses were prioritised over individual studies and those included were checked to ensure PND had been screen detected.

After reviewing the abstracts, 21 studies met the criteria for full text review of which one systematic review was eligible to be included (O'Connor et al 2016)[26]. To provide a broader description of the evidence available about treatments for women diagnosed with postnatal depression an overview of an additional 4 recent systematic reviews has been included. However, it is not clear in these reviews whether the participants in the included studies had mental health conditions that were screen-detected and so the findings may not be applicable to a screen-detected population.
The 4 systematic reviews describe:

- non-pharmacological interventions and maternal outcomes (Morrell et al 2016, O'Connor et al 2016, Pritchett et al 2017)[47],[26,48]
- non-pharmacological interventions and other outcomes(Letourneau et al 2017)[36]
- pharmacological interventions and maternal outcomes(McDonagh et al 2014)[37]
- pharmacological interventions and other outcomes(McDonagh et al 2014)[37]**Error! Bookmark not defined.**

Table 14 in Appendix 2 outlines the reasons for exclusion of 15 studies which are generally due to the publication of a more recent systematic review.

Summary of findings

A study-level summary of data extracted from each included publication is presented in 'Summary and appraisal of individual studies' (Appendix 3) and are stratified by question.

Table 8 summarises the outcomes of the included papers.

Non-pharmacological interventions

One systematic review focussed on studies using screen- detected cohorts of women with postpartum depression (O'Connor et al 2016)[26]. In addition, one meta-analysis (Pritchett et al 2017)[48] and 2 systematic reviews (Letourneau et al 2017, Morrell et al 2016)[36,47,] evaluated the effectiveness of a range of non-pharmacological interventions in studies that targeted populations of women with postpartum depression who may or may not be screen detected.

Table 8. Summary of included studies for key question 5

Paper	Mental health condition of mother	Treatment	Outcomes	Number studies	Study outcome	**Quality appraisal**
Screen-detected population						
O'Connor et al (2016)[26] Systematic review	Postpartum depression	CBT, IPT	Maternal depression	15 With 1074 participants in total ranging from 37 to 192	For reduction in depression symptom severity, likelihood of remission over the short term (<8mo) • 7 CBT postpartum trials had relative risk ratios ranging from 1.09 (95% CI, 0.84 -1.42) to 1.97 (95%CI, 0.88 - 4.44). • 3 trials of nondirective therapeutic approaches reported relative risk of 0.96 (95%CI,0.72 – 1.27) to 3.20 (1.32 - 7.76)	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns. The authors of the systematic review used criteria developed by the US Preventive Services Task Force[39] for randomized clinical trials, the NOS for observational studies[40], and the AMSTAR for systematic reviews[41]. The studies included in the review in relation to the treatment of women for depression in pregnancy were judged by the reviewers to be mainly fair. All studies were small with possible bias due to the different definitions of remission.
Not explicitly screen-detected populations						
Letourneau et al (2017)[36]	Depression and anxiety	CBT, IPT, peer support, maternal-child interaction guidance	Parenting and offspring outcomes	28 of which 3 (n= 112) reported outcomes relevant to this review	Maternal-child interaction guidance (3 trials). Effect sizes in 3 of the trials for some outcomes were reported: effect of verbal feedback to improve maternal parenting behaviour (effect size (r) 0.51) quality of interactive behaviour (effect size (r) 0.68 and 0.66)	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns The authors used the Cochrane Collaborations risk of bias tool for RCTs[42]. A summary assessment of risk of bias was made for each paper with a score of 0-5 being low risk, 6 being moderate risk and 7 upwards high risk. Scores ranged from 0 to 6 for the 28 publications. Areas of high risk of bias were allocation concealment of the intervention, blinding of participants and personnel and blinding of outcome of assessment..
Pritchett et al (2017)[48] Meta analysis	Postpartum depression	Exercise	Maternal outcomes	13 The total population of the combined	Overall exercise interventions significantly reduced depressive symptoms (p<0.006) but effect size (SMD) had wide confidence intervals reducing the precision of the point estimate (SMD –	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns. The authors of the meta-analysis used the Cochrane Collaboration's risk of bias tool to assess the quality of

				studies was 1734 ranging from 20 to 679	0.44‡‡‡‡, 95% CI, –0.75 to –0.12, n = 1307, I^2 85%, 13 trials). This was a similar pattern for all exercise interventions. -Depressed postpartum populations, p<0.05, SMD –0.32, 95% CI, –0.63 to 0.00, I^2 55% -General postpartum populations, p<0.04, SMD-0.57, 95% CI, –1.12 to –0.02, I^2 92% -Exercise with co-interventions, p<0.03 SMD –0.35, 95% CI, –0.66 to –0.04, I^2 72%. -Exercise-only interventions p<0.05 SMD –0.56, 95% CI, –1.13 to 0.01, I^2 89%. -Group exercise interventions, p<0.02SMD difference –1.10, 95% CI, -1.99 to –0.21, I^2 93% -Participant choice interventions such as exercise counselling with personal choice of exercise (often exercise alone) p<0.003 SMD –0.20, 95% CI, –0.33 to –0.06, I^2 0%.	the included papers. They noted that a substantial level of heterogeneity was present in the trials design. The methodological quality of several of the included trials was low. The main causes of risk of bias were: the failing of performing an intention-to- treat analysis; a lack of clarity on how the selected outcomes were reported, a lack of robust sequence generation and concealment of randomisation procedures, and unclear blinding.
Morrell et al (2016)[47] HTA	Postpartum depression	CBT, IPT	Maternal outcomes	86 A total of 66,418 participants were randomised across the 86 trials, with the individual trial sample sizes ranging from 25 to	The most beneficial interventions at 12 months, shown by difference in the mean EPDS score:midwifery redesigned postnatal care (–1.43, 95% credible interval (CrI) –4.00 to 1.36), person-centred approach (PCA)-based intervention (–0.97, 95% CrI, –3.54 to 1.71) cognitive–behavioural therapy (CBT)-based intervention (–0.78, 95% CrI –3.41 to 1.91	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns The authors of the systematic review evaluated the overall risk of bias of randomised controlled trials only. No quality assessment was undertaken for the systematic reviews. The majority of the studies were considered to have low selection bias. Higher level of risk was associated with allocation concealment. Due to the nature of the interventions, blinding of participants or caregivers was not always possible, but it was considered that it was unlikely that the lack of blinding had an effect on the results. Therefore, the majority of the studies were rated as being at low risk

‡‡‡‡ SMD(standardised mean difference - effect size 0.2= small, 0.5 = medium, 0.8 = strong, 1.3 = very strong

				18,555 participants		of performance bias. Around 60% of the RCTs were considered at low risk of attrition bias, over 30% were considered unclear and around 10% were assessed as being at high risk for selective outcome and/or analysis bias
McDonagh et al (2014)[37] Systematic review	Postpartum depression	Anti-depressants	Maternal and offspring outcomes	7	There was low strength evidence that symptom response was not improved when sertraline was added to psychotherapy or when CBT was added to paroxetine (figures not reported).	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns. The authors assessed the internal validity of the studies based on criteria by McDonagh et al (2012)[43] and graded the evidence using the AHRQ grading methodology[44]). All 7 studies were assessed as of medium quality with concerns about controlling for confounding variables, small sample sizes and poor reporting of health outcomes and adverse events.

Screen-detected populations

O'Connor et al (2016)[26] reported that overall CBT and related therapeutic approaches were associated with an increased likelihood of remission in the short term (<8 months) with relative risk from 7 trials (n=695) with relative risk ranging from no effect (1.09, 95%CI, 0.84 to 1.42) to weak effect (1.97,95% CI, 0.88 to 4.44) with wide confidence intervals indicating non significance. Relative risk was generally associated with greater contact hours; however, this was confounded with other important sources of heterogeneity. Use of non-directive therapeutic approaches in 3 trials reported relative risk of between 0.96 to 3.20 (95%CI, 0.72 to 1.27 and 95%CI, 1.32 to 7.76). Data were insufficient to evaluate other treatment approaches, including stepped care. Trials tended to be small (the largest was n=193) with 1 or more methodological limitations and of fair quality.

Not explicitly screen-detected populations

The HTA systematic review by Morrell et al (2016)47concluded that the most beneficial interventions at 12 months postpartum shown by difference in the mean EPDS score was midwifery led postnatal care [–1.43, 95% credible interval[§§§§] (CrI) –4.00 to 1.36], a person-centred approach (PCA)-based intervention (–0.97, 95% CrI –3.54 to 1.71) and a cognitive–behavioural therapy (CBT)-based intervention (–0.78, 95% CrI –3.41 to 1.91). However, the effect size confidence intervals were wide limiting the precision of the point estimate. The quality of the clinical and cost effectiveness evidence available was rated as poor, indicating that replication of some studies is needed within good-quality RCTs.

A meta-analysis of 13 trials with 1734 participants by Pritchett et al (2017)[48] of aerobic exercise interventions found that overall they statistically reduced depressive symptoms (p<0.006) but effect size (Standardised mean difference, SMD) had wide confidence intervals reducing the precision of the point estimate and the strength of any effect (SMD –0.44, 95% CI, –0.75 to –0.12, n = 1307, I^2 85%). Each of the exercise intervention types showed the same pattern of significant difference between exercise and usual care but weak or no effect due to point estimates with wide confidence intervals.

A meta-analysis carried out by Letourneau et al (2017)[36] on CBT interventions showed no overall effect on the parenting stress index (SMD=0.154, 95%CI, -0.005 to -0.313, p=0.057). Peer support investigated by two trials reported by Letourneau et al (2017)[36] found no difference between the intervention and usual care groups. Maternal interaction guidance was investigated by 6 trials but in the 3 that reported on parenting behaviour no statistically significant effect was found. One trial reported a large effect size of verbal feedback to parents to improve maternal parenting (effect size 0.51) and 2 trials showed an

[§§§§] CrI -credible interval is the interval in which an (unobserved) parameter has a given probability

improvement in the interactive behaviour between parents and offspring. (effect size 0.68 and 0.66).

Pharmacological interventions

Screen-detected populations

O'Connor et al (2016)[26] found 1 trial that used a screen-detected cohort of women to examine the benefit of fluoxetine on depression symptoms that reported a 10 point reduction in EPDS with fluoxetine after 12 weeks compared with a 7 point reduction with a placebo (p=0.05). No other evidence was available about the benefits of the use of anti-depressants postpartum either to women or offspring.

Not explicitly screen-detected populations

The systematic review by McDonagh et al (2014)[37] reported low quality evidence that postpartum depression symptom response was not improved when sertraline was added to psychotherapy or when CBT was added to paroxetine. There was insufficient evidence to determine other outcomes such as functional capacity, breastfeeding and infant child development.

Summary of Findings Relevant to Criterion 9: Criterion not met*****

Antenatal mental health

This evidence summary showed that although there is a large volume of evidence looking at the benefits of pharmacological and non-pharmacological interventions for common mental health disorders during pregnancy, there are very few studies reporting results from screen-detected cohorts.

Evidence on the benefits from non-pharmacological interventions in screen-detected women came from 3 studies involving very small cohorts (25 to 149 participants). In terms of outcomes the risk ratios indicate a weak effect and the confidence intervals undermine certainty in these estimates.

***** **Met** -for example, this should be applied in circumstances in which there is a sufficient volume of evidence of sufficient quality to judge an outcome or effect which is unlikely to be changed by further research or systematic review.
Not Met - for example, this should be applied in circumstances where there is insufficient evidence to clearly judge an outcome or effect or where there is sufficient evidence of poor performance.
Uncertain -for example, this should be applied in circumstances in which the constraints of an evidence summary prevent a reliable answer to the question. An example of this may be when the need for a systematic review and meta-analysis is identified by the rapid review.

The only evidence reported in this evidence summary on the benefits of pharmacological interventions in screen-detected women came from one systematic review. The studies considered in the systematic review were large, but as they were they were all observational, causality could not be confirmed and it was not possible to control for all cofounders related to depression. This systematic review looked only at harms from to second-generation antidepressants agents which might not be the current preferred treatment.

In general for this question there was a lack of evidence in relation to the majority of the predefined outcomes, depression and PTSD being the main outcomes of interests for these studies.

There is a paucity of good quality evidence about the effective pharmacological and non-pharmacological interventions for women with other screen-detected common mental illnesses during pregnancy including generalised anxiety disorder (GAD), panic disorder, phobias, social anxiety disorder, OCD and PTSD so firm conclusions about treatment for these disorders cannot be drawn. The low strength of evidence for effective interventions for women with screen-detected common mental illnesses indicates future studies are needed to inform decision-makers on this issue and therefore criterion 9 is not met.

Postnatal depression
Evidence from the 2011 UK NSC's review showed that the evidence base on intervention for postnatal depression was relatively small and based only on small size studies. It noted that this paucity of evidence limits the capability to make recommendations in this area. The current review suggests that this is still the case in both screen-detected and non-screen-detected populations.

For example, in screen-detected women a systematic review of non-pharmacological interventions reported a very limited volume of evidence with generally weak measures of effect. The same review found only one study of pharmacological interventions in screen-detected women.

The problem of small study size is replicated in studies of non-screen-detected populations. Though the overall number of studies is larger than in screen-detected women, measures of effect tend to be weak with confidence intervals undermining certainty in these estimates.

In the absence of the large studies as recommended in the previous review and with very little direct evidence relating to screen-detected women, this criterion remains not met for interventions for postnatal depression.

Criterion 15

Clinical management of the condition and patient outcomes should be optimised in all health care providers prior to participation in a screening programme.

Question 6 – Is clinical detection and management currently well implemented in the UK?

Sub question –What is the proportion of such conditions that remain undiagnosed?
Key question 2 aims to describe the evidence available that indicates how well clinical detection, referral and treatment of perinatal mental health problems are currently managed in the UK.

This question was not formally addressed by the UK NSC in previous reviews.

Clinical guidelines developed by NICE (Clinical Guideline 192)[1] updated in 2014 and SIGN guideline 127(2012)[5] set out how clinicians should identify common antenatal mental health problems and postpartum depression. These are summarised in the table below.

It is important to know how well the guidance is being used by midwives and GPs and women's views about how perinatal mental health is broached by primary care professionals. The purpose of key question 2 is to identify the proportion of women currently identified with common antenatal mental health disorders and postpartum depression, the proportion of missed cases and how well guidance recommendations are reflected in the day to day recognition and management of these disorders.

Table 9. NICE CG 192 (2014) and SIGN guideline 127 (2012) on recognising antenatal and postnatal mental health problems in pregnancy and the postnatal period.

Condition	Timing	Mental health problem identification
Depression and anxiety (NICE CG192)[1]	At all contacts after the first contact with primary care or the booking visit, the health visitor, and other healthcare professionals who have regular contact with a woman in pregnancy and the postnatal period (first year after birth)	Consider asking the following questions (also known as the Whooley questions): during the past month, have you often been bothered by feeling down, depressed or hopeless? during the past month, have you often been bothered by having little interest or pleasure in doing things? If either question is affirmatively answered then consider using the EPDS, PHQ-9 and GAD-2 scale prior to onward referral.
Depression and anxiety (SIGN 127)[5]	Enquiry about depressive symptoms should be made, at minimum, on booking in and postnatally at four to six weeks and three to four months.	The EPDS or the Whooley questions may be used in the antenatal and postnatal period as an aid to clinical monitoring and discussion of emotional issues. If there are concerns about the presence of depression, women should be re-evaluated after 2 weeks. If symptoms persist, or if at initial evaluation there is evidence of severe illness or suicidality, women should be referred to their GPor mental health service for further evaluation.

EPDS – Edinburgh Postnatal Depression Scale, PHQ 9 – Patient Health Questionnaire (9 questions), GAD – General Anxiety Disorder Scale (version 2)

Eligibility for inclusion in the review

The inclusion criteria for this question are outlined below:

Population: All UK pregnant women
Outcomes

- proportion of mental health problems detected (prevalence)
- proportion of women asked questions according to guidance
- proportion of women with mental health problems referred for intervention
- proportion of women attending / complying with intervention
- user experiences

Study design: Audit data, cross sectional study, cohort study (prospective and retrospective), systematic reviews.

Date and language: English peer reviewed publications from 2006 onwards.

Description of the evidence

Database searches yielded 746 results, of which 21 were judged to be relevant to this question.

After reviewing the abstracts 15 studies met the criteria for full text review and following review of the full texts, 6 studies were included. Table 14 in Appendix 2 outlines the reasons for exclusion of 9 publications. Appendix 1 includes the full PRISMA flow diagram, along with a table of the included publications and details of which questions these publications were identified as being relevant to (Table 13). The publications included were focussed on common mental health problems such as depression and anxiety.

Summary of findings

A study-level summary of data extracted from each included publication is presented in ‘Summary and appraisal of individual studies’ (Appendix 3) and are stratified by question.

Of the 6 papers included for question 6, 1 was a meta-synthesis (Ford et al 2017)**Error! Bookmark not defined.** of qualitative research about GP views of regarding the identification of mental health problems in pregnant women, 1 paper was qualitative study (McGlone et al 2016)**Error! Bookmark not defined.** of midwives experiences and 1 gathered views from both midwives and pregnant women (Williams et al 2016)**Error! Bookmark not defined.**. One paper reported national survey data (Redshaw et al 2016)[49]**Error! Bookmark not defined.** and one was a l

ongitudinal study of women in Bradford reporting rates of common mental disorders in prior to conception, during pregnancy and postpartum (Prady et al 2016)[50]. The final paper was a cohort study that tracked the documenting of mental health conversations, referrals and actions in women's hand held notes and health records (Darwin et al 2015)[51]. Table 10 summarises key details from each of the studies.

Table 10: Summary of included papers for key question 6

Type of study	Paper	Number	Study results	Quality appraisal
Proportion of mental health problems detected in pregnancy and postpartum	Prady et al (2016)[50] Longitudinal study	8991 women	% Period Prevalence (identified through GP records) for common mental health disorders (95% CI) • 6 months prior to conception and pregnancy 9.5%(8.9 to 10.1) • 1st year postpartum13.1%(12.4 to 13.7) • 2nd year postpartum12.8%(11.9 to 13.6) • 3rd year postpartum 14.0%(13.2 to 14.9) Incidence (identified through GP records) as CMD Rate per 1000 person years at risk (95% CI) • pregnancy 37.5 (33.1 to 42.5) • 1st year postpartum 102.4(95.5 to 109.8) • 2nd year postpartum 64.9 (59.1 to 72.4) • 3rd year postpartum 68.2 (61.1 to76.2) % Potentially missed cases of CMD (95% CI) • pre-birth (6 months prior to conception and pregnancy) GHQ-28 threshold ≥15 31.3% (28.7 to 33.8) • pre-birth (6 months prior to conception and pregnancy) GHQ-28 threshold ≥9 46.8% (44.7 to 49.0) Fewer than 13% of women had read codes in their records indicating screening and case finding of CMD in the first postnatal year.	A CASP checklist for longitudinal studies was undertaken. Limitations of the study include lack of knowledge of the quantity of missing data from primary care records, only administering the GHQ-28 once in mid to late pregnancy and applying conservative weighting of results so estimates are likely to be under rather than overestimated. The population had low socioeconomic diversity and it was disadvantaged and served by one maternity unit. This may make the results less generalisable across the rest of the UK

Proportion of women asked questions according to guidance: Secondary analysis of study about who is offered depression screening	Redshaw et al (2016)[Error! Bookmark not defined.]
	
	National survey (2014)
		4572 women	Antenatal cohort (4569)
			• asked about mental health – 3703 (81.9%)
			• women disclosing mental health problems of those asked 946 (21%)
			• offered help (of those disclosing MH problems - 337(35.7%)
			  ○ support was given to - 215 (65.4%)
			  ○ advice was given to – 229 (71.3%)
			  ○ treatment was received by – 119 (45.2%)
			Postnatal cohort (4500)
			• asked about mental health - 4043 (89.8%)
			• women offered help (of total postnatal cohort) - 522(11.6%)
			• support was given to - 263 (63.4%)
				The CASP cohort study checklist was used to assess the quality of this study. Limitations of the work include an under representation of survey respondents who were younger, single, born outside of the UK and living in disadvantaged areas making the results less generalisable to the whole of the UK

			• advice was given to – 165 (64.4%) • treatment was given to – 262 (50.2%)	
Proportion of women with mental health problems referred for intervention	Darwin et al (2015)[51] Mixed methods cohort study	191	Midwife appointments and mental health questions: • 167 (87.4%) responses to mental health questions at booking were present in both hand held notes and health records • 5 (3%)of these did not have completed questions • 0% had completed questions in trimesters 2 and 3 30 (18.5%) women had ticked one of the questions as applying to them • 21 (70%) of these had comments documented in the hand held notes • 8 (27%) were consistent between handheld notes and health records Referrals 23/191 (12%) women were referred to a mental health midwife service • 3 of these referrals had a note 'will contact patient' • 12 (52%) referrals either 'did not meet criteria' or had 'no plans to contact' • 8 (35%) no evidence of response from the mental health midwife	A CASP checklist for longitudinal studies was undertaken and no concerns were identified. It is not known whether the findings of this local study are generalisable across the UK.
User experiences: GPs views of diagnosing depression and or anxiety in pregnant women and postpartum	Ford et al (2017)[Error! Bookmark not defined.] Meta synthesis of qualitative information	4 publications of 3 UK studies Reporting on views in total from 323 GPs only 70 of these came from UK GPs	Themes from the 4 UK publications included • reluctance to identify a condition with a diagnostic label (2/4 publications) • clinical judgement vs use of guidelines with a preference for GPs to rely on their own intuition to pick up the signs if something was wrong (3/4 publications) • care and management of women identified – approach was determined by GP own experience and resources available (4/4 publications) • use of medication – approach depends on what non-pharmacological resources are available and balance of risk of treatment vs no treatment (1/4 publications) • isolation and role of other professionals – links between GPs and Health visitors is much reduced so less communication about women who might have problems (2/4 publications)	Ford et al (2017): There are currently no checklists available to critically appraise a meta synthesis of qualitative studies so the CASP checklist for systematic reviews was used. The authors used a checklist derived from Atkins et al (2008)[52] to indicate the range of quality of the qualitative research. Out of 11 points all studies scored 9 or 10. The research papers included reported a total of 70 UK GP views (there are approximately 34,000 UK GPs) and so may not be representative.

User experiences: Midwives experiences of women completing screening using Whooley questions	McGlone et al (2016)[Error! Bookmark not defined.] Survey of user experiences	8 midwives	Themes and sub themes from the interviews with the 8 midwives include:
			• no clear understanding of the purpose of the Whooley questions
			  ○ midwife discomfort when women disclose signs of depression and anxiety
			• feeling pressurised from lack of time and frustration at unable to fulfil their role
			• resultant dissatisfaction to fulfil their role
			  ○ lack of knowledge and how to refer
			  ○ lack of training around perinatal mental illness
			  ○ rely on experience and intuition rather than training
				The CASP checklist for qualitative research was used to assess the quality of this study. Overall there were no areas of concern about the quality of the study although the number of midwives interviewed was low and they were all from one maternity unit which may not be representative of other units across the country

User experiences: Prospective cohort of women and midwives experiences	Williams et al (2016)[Error! Bookmark not defined.] Qualitative study	15 midwives and 20 pregnant women	Key themes midwives: • asking the questions at the booking appointment was not always appropriate and some did not ask the questions if that was the case • some midwives felt the questions were a bit blunt • when women could not speak English or where they had learning disabilities there was a greater challenge in communicating a sensitive topic and more reluctance on the part of the midwife • some midwives were uncomfortable asking the questions • the 2 Whooley questions were seen as on the whole a good way to broach the subject and drew out issues that might not otherwise arise • midwives felt the Arroll 'help' question helped them understand how the women were feeling but there were not a lot of options for support Key points pregnant women: • most women did not expect to be asked about mental health issues at the booking appointment • some felt the questions were a bit blunt • many women found it a good opportunity to talk about pregnancy anxiety and whether what they were experiencing was normal • women felt both that it was an advantage to be able to talk to the midwife about how they were feeling but also felt admitting they were feeling low during pregnancy carried social stigma	The CASP checklist for qualitative research was used to assess the quality of this study. Overall there were no concerns about the quality of the publication. The sample size was small (35 participants in total) and recruitment was via a validation study already being undertaken in Bristol which may limit generalisability to the rest of the UK. All views are reported by women and midwives rather than from direct observation of the booking appointment

		• how the midwife was perceived (brisk or distracted versus empathic and non-judgemental) had an impact on the women's response to the questions and their likelihood of disclosing issues • all the women were happy to be asked the Whooley questions as it was an opportunity to talk about how they felt. • women understood the Arroll 'help' question variably – women thought it could mean seeing a GP for medication, a counsellor, specialist support group or psychiatrist, or some form of self-help. Some did not know what help would consist of and the midwives did not clarify this.

Overview of results

The evidence available for question 6 is on the whole of good quality and each of the 6 studies adds to the overall picture of how the identification and onward referral of women with common mental disorders during pregnancy and postpartum is undertaken in the UK. However, most studies are carried out in a single maternity service catchment which may not represent wider day-to-day practice across other UK Trusts. In a national survey from 2014 a high proportion of women were asked the 2 questions recommended by NICE guidance[1]. For a variety of reasons there is ambivalence by GPs and midwives about using these questions and how to respond to women if they disclose a mental health problem. The type and range of support, advice and treatment available to women with mental health issues is reportedly variable which is 1 reason given for ambivalence in prompting women to disclose problems. Referral to a service did not always mean women would be offered help.

Pregnant women on the whole welcomed being asked about their mental health although they varied in their understanding of what help would be offered to them if they disclosed any issues.

One study of 8891 women from one maternity service reported prevalence and incident rates of common mental disorders identified by primary care and calculated the proportion of missed cases. This geographical area is characterised as an area of high deprivation and might be expected to have higher rates of common mental disorders than other areas of the UK.

Proportion of mental health problems detected

Prady et al (2016)[50] reported the period prevalence††††† of common mental disorders through the maternal period. GP documented common mental health disorders were recorded in the notes of 9.5% of a cohort of 8991 women during pregnancy and up to 6 months prior to conception. In order to determine missed cases, pregnant women were asked to complete the General Health Questionnaire (GHQ-28) at 26 to 28 weeks gestation and the results linked to their mental health records up to 6 months prior to conception and 3 years postpartum. When a threshold score of ≥9 was used (the standard threshold range is between 5 and 8) 46.8% (95% CI, 44.7 to 49.0) of women above the threshold had no indication in their medical records that a common mental health disorder had been detected earlier in pregnancy (and up to 6 months prior to conception). Of those pregnant women with a GHQ-28 score of ≥15, the authors estimated that, 31.3% (95%CI,

††††† Period prevalence is proportion of people in a population who have a particular disease or attribute at a specified point in time or over a specified period of time

28.7 to 33.8) women with common mental disorders were missed compared to GP documented cases. The study reported that the incidence of common mental disorders reported in the antenatal period tended to increase in the postnatal period, and then, it remains higher in the following two years (Table10).

Proportion of women asked questions according to guidance

Redshaw et al (2016)[49] analysed data from the National Maternity Survey carried out in 2014. They reported that the majority of women during pregnancy (81.9%) were asked about their mental health and less than a quarter (21%) disclosed having mental health problems. Of those, only around a third (35.7%) were offered help, which comprised support, such as extra visits from midwives and counselling(65.4%), advice (71.3%) and treatment (45.2%).

During the postnatal period, a slightly higher proportion of women (89.8%) were asked about their mental health but the proportion disclosing mental health problems was not reported. Of the 4043 asked about their mental health 522 (11.6%) were offered help which included support (63.4%), advice (64.4%) and treatment (50.2%).

Prady et al (2016)[24] reported that fewer than 13% of women had information in their records indicating that screening or case finding for common mental disorders had taken place in the first postnatal year.

Proportion of women with mental health problems referred for intervention

Darwin et al (2015)**Error! Bookmark not defined.** undertook a mixed methods cohort study of 191 p regnant women of whom 162 had complete documented responses to the 2 Whooley questions at the antenatal booking appointment. Of the 162 responses 30 (18.5%) women had ticked at least 1 of the 2 questions (which indicate further assessment). Of these, 8 (26.6%) had the information documented consistently in both the hand held notes and health records. Of the 30 women assessed, 23 (76.6%) were referred to a specialist mental health midwife. Of referred women, 12 (52%) were recorded by the service as ‘did not meet the criteria’ or ‘no plans to contact’, 8 (35%) had no documented response from the service and 3 (13%) had a note which said ‘will contact patient’.

Views of professionals and women about recognition of mental health problems during pregnancy and postpartum

Three papers, Ford et al (2017)[Error! Bookmark not defined.], McGlone et al (2016)[Error! Bookmark not defined.] and Williams et al (2016)[Error! Bookmark not defined.] describe the views of GPs, midwives and pregnant women and their experiences of recognition, referral and management of people with mental health issues in the perinatal period.

Common themes for professionals were:

- tension between using clinical guidance and what felt appropriate to talk about with individual women
- the requirement to ask the mental health questions and complete a field on their computer system before being able to continue to the next field (and the next question at the booking in appointment) was seen by some midwives as unhelpful as it might not be the right time to ask the questions
- a reluctance/discomfort in talking about mental health problems
- lack of resources to support women when they were identified
- the decision to use pharmacological or non-pharmacological interventions was largely based on availability of options balanced with the risk of not treating someone and GP experience
- lack of training about perinatal mental health issues
- the requirement to ask about mental health problems was a good way of broaching a subject that would be difficult to otherwise introduce to find out what women were experiencing
- lack of time to explore the issues properly

Key themes[51] voiced by pregnant women[Error! Bookmark not defined.] were:

- most women did not expect to be asked about mental health issues at the booking appointment
- many women found it a good opportunity to talk about pregnancy anxiety and whether what they were experiencing was normal
- women felt both that it was an advantage to be able to talk to the midwife about how they were feeling but also felt that admitting they were feeling low during pregnancy carried social stigma
- women varied in their perception of what support might entail
- how the midwife was perceived (brisk or distracted versus empathic and non-judgemental) had an impact on the women's response to the questions and their likelihood of disclosing issues.

Summary of Findings Relevant to Criterion 15: Criterion not met‡‡‡‡‡

There is a small volume of good quality evidence about the current clinical detection and management of antenatal and postpartum mental health problems in the UK. The 3 qualitative studies varied in size from 8 midwives to 8991 women and consistently reported the same issues with identifying and referring women with mental health problems during pregnancy and postpartum. These included reluctance to broach the subject of mental health problems with women at the booking appointment, a lack of resources to support women when they were identified and a lack of understanding about perinatal mental health problems. All but 1 study (the National Maternity Survey 2014) was based on data collected in relation to a single maternity service so it is difficult to know whether the results of the studies are applicable across the country.

A large study, albeit from one maternity service, reported 46.8% of missed cases of common mental health disorders amongst pregnant women and it would be helpful to know the proportion of likely missed cases in other parts of the country.

Overall the studies suggest that the clinical detection and management of antenatal and postpartum mental health problems is not effectively implemented in the areas of the UK were the studies were carried out. Most women were likely to be asked about their mental health; however, addressing those problems by onward referral, support, advice and treatment is variable. However, is not clear if this finding is likely to be the case across the whole of the UK. This criterion is therefore not met.

‡‡‡‡‡ **Met** -for example, this should be applied in circumstances in which there is a sufficient volume of evidence of sufficient quality to judge an outcome or effect which is unlikely to be changed by further research or systematic review.
Not Met - for example, this should be applied in circumstances where there is insufficient evidence to clearly judge an outcome or effect or where there is sufficient evidence of poor performance.
Uncertain -for example, this should be applied in circumstances in which the constraints of an evidence summary prevent a reliable answer to the question. An example of this may be when the need for a systematic review and meta-analysis is identified by the rapid review.

Review summary

Conclusions and implications for policy

This report is an update review on systematic antenatal and postnatal screening for mental health problems against select UK NSC criteria for appraising the viability, effectiveness and appropriateness of a screening programme. This review assessed key questions to determine if new evidence published since 2006 for antenatal screening for common mental health problems and 2011 for screening for postnatal depression suggests that reconsideration of the current recommendation for screening in the UK is required.

The volume, quality and direction of new evidence published since 2006 and 2011 does not indicate that the changes in the evidence base since the previous recommendations were made are sufficient to reconsider the current UK NSC's recommendation for systematic screening for mental health problems during pregnancy and postpartum period.

The main conclusions are:

1. There is a large volume of evidence about adverse outcomes associated with common mental health problems experienced by women during pregnancy and postpartum. Some outcomes are consistently reported such as preterm birth and low birth weight whilst others are less consistently observed.
2. There is a paucity of evidence for effective screening tests for common mental disorders such as generalised anxiety disorder, panic disorder, phobias, social anxiety disorder, obsessive-compulsive disorder and post-traumatic stress disorder during pregnancy. The total body of evidence during pregnancy is based on studies with low numbers of participants and in general the accuracy of data reported across studies is variable. There is also a substantial heterogeneity between the studies in relation to study design, population sampled and the diagnostic criteria used.
3. Data on postpartum depression come from studies with larger numbers of participants than in the antenatal period. However, the numbers are still low when considered as evidence for population-based screening. Moreover, these studies suffer from the same heterogeneity problems noted in the antenatal studies. Even though, the quantity, quality, consistency and applicability of the evidence indicate that there are screening tools (EPDS, PHQ 2) that could be used as part of an overall screening programme, the fact that the evidence is based on small cohorts makes it difficult for decision-makers to extrapolate a conclusion that is appropriate for national population-based screening policies.
4. The EPDS, screening test for major depression disorder has a high sensitivity and specificity, but low positive predictive value. Therefore, a high proportion of women with a positive screen referred for a full psychosocial assessment are likely not to have major depressive disorder.
5. Firm conclusions about the effectiveness of each of the pharmacological and non-pharmacological interventions for women with screen-detected common mental disorders

including, generalised anxiety disorder, panic disorder, phobias, social anxiety disorder, obsessive-compulsive disorder and post-traumatic stress disorder in pregnancy, cannot be drawn from the number of small studies available as they have considerable heterogeneity in their methodology, level of bias and consistency of results.

6. The evidence base for the effectiveness of pharmacological and non-pharmacological interventions for women with screen-detected antenatal and postnatal depression remain very limited. Although some evidence indicates that CBT is likely to lead to a small reduction in the severity of the condition, these conclusions are based on a very limited volume of evidence with generally weak measures of effect. Only one study of pharmacological interventions in screen-detected women was identified.
7. Antenatal and postnatal mental health problems are important mental health and public health issues and this review found evidence of their influence on pregnancy and neonatal outcomes, it also found that current mental health services in the UK are not implementing the NICE guidance in its entirety. Although most women are likely to be asked about their mental health, actions to address those problems, by onward referral, support, advice and treatment is variable.

The development of antenatal and postnatal mental health problems is complex with serious life implications for the woman, her family and her baby. This evidence summary showed that there is still a lack of clarity on the population to be identified by screening and on the ability of the current screening tools to identify with sufficient accuracy women who would benefit more from treatment.

This has important consequences, as over-detection will mean some women will have unnecessary further psychosocial assessment and may experience stigmatisation associated with mental health problems. This is complicated by the fact that there is still insufficient evidence that universal screening and subsequent intervention improve the health outcomes for the mother or the baby.

Nevertheless, it is still important to recognise that, as part of a comprehensive clinical assessment, health professionals should be alert to the possibility of antenatal and postnatal depression and manage it according to current guidance.

The current recommendation not to introduce a UK systematic antenatal and postnatal population screening programme for mental health problems should be retained.
It is recognised that the development of psychiatric illness in pregnancy and depression in the postnatal period are complex issues and the benefit and harms of a screening programme need to be carefully balanced. There is a lack of clarity on the population to be identified by screening and evidence that the use of current screening tools cannot identify risk with sufficient accuracy. There is also insufficient evidence that universal screening and subsequent intervention improve the health outcomes for the mother or the baby.

Limitations

This review was limited by the lack of evidence specific to the population of interest for population-based screening particularly relating to the performance of screening tests for detecting generalised anxiety disorder, panic disorder, phobias, social anxiety disorder, obsessive compulsive disorder and post traumatic stress disorder. Also the evidence of the effectiveness of interventions for those women who are screen-detected comes from small size studies, and therefore it is not suitable for making national policy recommendations. The majority of the evidence base for effectiveness of interventions for women with pre and postnatal depression reported outcomes from populations which were not explicitly screen detected.

This rapid review process was conducted over a condensed period of time (approximately 12 weeks). Searching was limited to 3 bibliographic databases and did not include grey literature sources. The review was guided by a protocol developed a priori. The literature search and first appraisal of search results were undertaken by 1 information scientist, and further appraisal and study selection by 1 reviewer. Any queries at both stages were resolved through discussion with a second reviewer. Studies not available in the English language, abstracts and poster presentations, were not included. Studies that were not published in peer-reviewed journals were not reviewed.

Appendix 1 —Search strategy

Electronic databases

The search strategy included searches of Medline, PsycINFO, Cocharane and Embase databases shown in Table11.

Table 11. Summary of electronic database searches and dates

Database	Platform	Date of search	Date range of search
Medline	Ovid SP	14th February 2018	2006 to Present(Q1,2,3,5) 2011 to present (Q4 & 6)
Embase	Ovid SP	14th February 2018	2006 to present (Q1,2,3,5) 2011 to present (Q4 & 6)
PsycINFO	Ovid SP	14th February 2018	2006 to present (Q1,2,3,5) 2011 to present (Q4 & 6)
The Cochrane Library, including: - Cochrane Database of Systematic Reviews (CDSR) - Cochrane Central Register of Controlled Trials (CENTRAL) - Database of Abstracts of Reviews of Effects (DARE)	Wiley Online	14th February 2018	2006 to present (Q1,2,3,5) 2011 to present (Q4 & 6)

Search Terms

Search terms for Medline, Embase, PsycINFO are shown in Table 14. Search terms included combinations of free text and subject headings

Table 12. Search strategy for MEDLINE, Embase and PsycINFO

Question 1

#	Searches
▲	
1	Adjustment Disorders/
2	Anxiety Disorders/
3	Dissociative Disorders/
4	Eating Disorders/
5	Mood Disorders/
6	Neurotic Disorders/
7	Personality Disorders/
8	exp Schizophrenia/
9	Somatoform Disorders/
10	Substance-Related Disorders/

11 Depression/
12 (Seasonal affective disorder$2 or depress$4 or Dysthym$4).tw.
13 melanchol$3.tw.
14 ((bipolar or bi polar) and (disorder$2 or depress$4)).tw.
15 ((((cyclothymi$3 or rapid) adj cycl$3) or ultradian) adj5 cycl$3).tw.
16 (mania or manic or hypomania).tw.
17 Anorexia/
18 (eating and disorder$2).tw.
19 anorexia nervosa.tw.
20 Bulimia/
21 (bulimia or hyperphagia).tw.
22 (bing$4 or overeat$3 or (compulsive and (eat$3 or vomit$3)) or (self induce$2 and vomit$3) or (restrict$3 and eat$3)).tw.
23 Anxiety/
24 Anxiety, Separation/
25 Panic Disorder/
26 (anxious or anxiety or panic or phobia or phobic).tw.
27 Stress, Psychological/
28 ((((((post adj traumatic$2) or posttraumatic$2 or stress) adj disorder$2) or acute) adj stress) or PTSD or ASD or DESNOS).tw.
29 (((((((combat or concentration) adj camp adj syndrome) or extreme) adj stress) or flash) adj back$2) or flashback$2).tw.
30 (((((hypervigilan$2 or hypervigilen$2 or psycholog$4) adj stress) or psycho) adj (trauma or traumatic)) or psychotrauma or psychtraumatic).tw.
31 (((((((((railway adj spine) or rape) and trauma$5) or torture) adj syndrome) or traumatic) adj neuros$2) or traumatic) adj stress).tw.
32 (trauma$5 and (avoidance or grief or horror or death$1 or (night adj mare$1) or nightmare$1 or emotion$2)).tw.
33 (recurr$5 adj thought$2).tw.
34 Obsessive Behavior/
35 obsessi$4.tw.
36 OCD.tw.
37 osteochondr$3.tw.
38 36 not 37
39 Compulsive Behavior/
40 compuls$5.tw.
41 (personalit$3 adj (disorder$1 or difficult$3)).tw.
42 (antisocial or anti-social or psychopath$3 or borderline or hysteri$3).tw.
43 (dissociat$3 adj (disorder$ or difficult$3 or personalit$3 or disturb$5 or trauma$4)).tw.
44 schizo$9.tw.
45 (psychos$2 or psychotic$2 or parano$3).tw.
46 (somatoform or somatization or somatic or hyperchondria$3 or neurasthenia$3).tw.
47 (conversion adj disorder$1).tw.

48 ((body adj dysmorphic) or dysmorphobi$2 or (briquet adj syndrome$1) or (syndrome adj briquet)).tw.
49 ((attach$ or bond or bonding) adj5 (ambivalent or anxious$4 or avoid$4 or difficult$3 or disinhibit$3 or disorder$2 or disorganis$4 or disruptiv$4 or dissociat$4 or dysregula$4 or disorientat$4 or disturbance$2 or impair$5 or inadequate or inhibit$3 or injur$4 or insecur$5 or poor or style$2)).tw.
50 ((adjust$4 or reactive) adj5 (disorder$1 or disturb$5 or difficult$3)).tw.
51 (transient situational adj (disturb$5 or disorder$1)).tw.
52 kleptomani$3.tw.
53 ((drug$2 or substance$2) adj (abuse or misuse or depend$6 or addict$4)).tw.
54 (neurotic$2 or neuros?s).tw.
55 (affective adj5 disorder$1).tw.
56 Mental Disorders/
57 (mental adj (illness$2 or disease$1 or disorder$1)).tw.
58 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57
59 (pregnan$ or antenatal$ or prenatal$ or perinatal$ or postnatal$ or postpart$).ti.
60 exp Pregnancy/
61 59 or 60
62 58 and 61
63 Depression, Postpartum/
64 62 or 63
65 Time/
66 ((longterm or long-term) adj3 (outcome? or effect? or consequence? or impact?)).ti,ab.
67 ((longitudinal or cohort or longterm or long-term) and (adverse adj3 (outcome? or effect? or consequence? or impact?))).ti,ab.
68 (adverse adj3 (outcome? or effect? or consequence? or impact?)).ti.
69 Mother-Child Relations/
70 ((mother* or maternal) adj3 (child* or infant*) adj5 (interact* or attach* or relation*)).ti,ab.
71 "treatment adherence and compliance"/ or patient compliance/ or medication adherence/ or treatment refusal/
72 ((treat* or therap* or medication? or medicine?) adj5 (adhere* or nonadhere* or complian* or noncomplian* or refus*)).ti,ab.
73 Family Relations/
74 ((family or parental or ((mother* or maternal) adj2 (father* or paternal))) adj5 (relation* or interact*)).ti,ab.
75 obstetric labor complications/ or exp obstetric labor, premature/
76 exp Infant, Low Birth Weight/
77 Fetal Growth Retardation/
78 Intensive Care Units, Neonatal/
79 Emotional Adjustment/
80 child development/
81 ((preterm or pre-term or premature) adj3 (birth* or childbirth* or infant*)).ti,ab.

82	(((small or large) adj2 gestational age) or (low adj2 (birthweight or birth weight)) or (growth adj2 restrict*)).ti,ab.
83	((infant* or child) adj2 (development or adjustment)).ti,ab.
84	((infant* or child*) and ((social or emotional or behavio?ral) adj (adjustment or development))).ti,ab.
85	((child* or adolescen* or teen* or paediatric? or pediatric?) adj2 (depression or mental health or mental illness)).ti,ab.
86	65 or 66 or 67 or 68 or 69 or 70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85
87	64 and 86
88	Meta-Analysis/
89	meta-analys?s.tw.
90	systematic review.tw.
91	case-control studies/ or cohort studies/ or follow-up studies/ or longitudinal studies/ or cross-sectional studies/
92	((cohort or longitudinal or case-control* or cross-sectional) adj stud*).ti,ab.
93	88 or 89 or 90 or 91 or 92
94	87 and 93
95	limit 94 to (english language and yr="2006 -Current")
96	(case reports or comment or editorial or letter or news or "review").pt.
97	95 not 96
98	exp animals/ not humans.sh.
99	97 not 98

Question 2.

# ▲	Searches
1	Adjustment Disorders/
2	Anxiety Disorders/
3	Dissociative Disorders/
4	Eating Disorders/
5	Mood Disorders/
6	Neurotic Disorders/
7	Personality Disorders/
8	exp Schizophrenia/
9	Somatoform Disorders/
10	Substance-Related Disorders/
11	Depression/
12	(Seasonal affective disorder$2 or depress$4 or Dysthym$4).tw.
13	melanchol$3.tw.
14	((bipolar or bi polar) and (disorder$2 or depress$4)).tw.
15	((((cyclothymi$3 or rapid) adj cycl$3) or ultradian) adj5 cycl$3).tw.
16	(mania or manic or hypomania).tw.
17	Anorexia/
18	(eating and disorder$2).tw.

19 anorexia nervosa.tw.
20 Bulimia/
21 (bulimia or hyperphagia).tw.
22 (bing$4 or overeat$3 or (compulsive and (eat$3 or vomit$3)) or (self induce$2 and vomit$3) or (restrict$3 and eat$3)).tw.
23 Anxiety/
24 Anxiety, Separation/
25 Panic Disorder/
26 (anxious or anxiety or panic or phobia or phobic).tw.
27 Stress, Psychological/
28 ((((((post adj traumatic$2) or posttraumatic$2 or stress) adj disorder$2) or acute) adj stress) or PTSD or ASD or DESNOS).tw.
29 (((((((combat or concentration) adj camp adj syndrome) or extreme) adj stress) or flash) adj back$2) or flashback$2).tw.
30 (((((hypervigilan$2 or hypervigilen$2 or psycholog$4) adj stress) or psycho) adj (trauma or traumatic)) or psychotrauma or psychtraumatic).tw.
31 (((((((((railway adj spine) or rape) and trauma$5) or torture) adj syndrome) or traumatic) adj neuros$2) or traumatic) adj stress).tw.
32 (trauma$5 and (avoidance or grief or horror or death$1 or (night adj mare$1) or nightmare$1 or emotion$2)).tw.
33 (recurr$5 adj thought$2).tw.
34 Obsessive Behavior/
35 obsessi$4.tw.
36 OCD.tw.
37 osteochondr$3.tw.
38 36 not 37
39 Compulsive Behavior/
40 compuls$5.tw.
41 (personalit$3 adj (disorder$1 or difficult$3)).tw.
42 (antisocial or anti-social or psychopath$3 or borderline or hysteri$3).tw.
43 (dissociat$3 adj (disorder$ or difficult$3 or personalit$3 or disturb$5 or trauma$4)).tw.
44 schizo$9.tw.
45 (psychos$2 or psychotic$2 or parano$3).tw.
46 (somatoform or somatization or somatic or hyperchondria$3 or neurasthenia$3).tw.
47 (conversion adj disorder$1).tw.
48 ((body adj dysmorphic) or dysmorphobi$2 or (briquet adj syndrome$1) or (syndrome adj briquet)).tw.
49 ((attach$ or bond or bonding) adj5 (ambivalent or anxious$4 or avoid$4 or difficult$3 or disinhibit$3 or disorder$2 or disorganis$4 or disruptiv$4 or dissociat$4 or dysregula$4 or disorientat$4 or disturbance$2 or impair$5 or inadequate or inhibit$3 or injur$4 or insecur$5 or poor or style$2)).tw.
50 ((adjust$4 or reactive) adj5 (disorder$1 or disturb$5 or difficult$3)).tw.
51 (transient situational adj (disturb$5 or disorder$1)).tw.
52 kleptomani$3.tw.

53	((drug$2 or substance$2) adj (abuse or misuse or depend$6 or addict$4)).tw.
54	(neurotic$2 or neuros?s).tw.
55	(affective adj5 disorder$1).tw.
56	Mental Disorders/
57	(mental adj (illness$2 or disease$1 or disorder$1)).tw.
58	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57
59	Prenatal Diagnosis/
60	58 and 59
61	(screen$3 or detect$3 or test or question$5 or surveillance).tw.
62	Mass screening/
63	61 or 62
64	(pregnan$ or antenatal$ or prenatal$).tw.
65	exp Pregnancy/
66	64 or 65
67	58 and 63 and 66
68	60 or 67
69	exp United Kingdom/
70	(united kingdom or uk or britain or gb or wales or scotland or england or ireland).ti,ab,in.
71	68 and 70
72	limit 71 to (english language and yr="2006 -Current")
73	(case reports or comment or editorial or letter or news or "review").pt.
74	72 not 73

Question 3.

# ▲	Searches
1	Adjustment Disorders/
2	Anxiety Disorders/
3	Dissociative Disorders/
4	Eating Disorders/
5	Mood Disorders/
6	Neurotic Disorders/
7	Personality Disorders/
8	exp Schizophrenia/
9	Somatoform Disorders/
10	Substance-Related Disorders/
11	Depression/
12	(Seasonal affective disorder$2 or depress$4 or Dysthym$4).tw.
13	melanchol$3.tw.
14	((bipolar or bi polar) and (disorder$2 or depress$4)).tw.
15	((((cyclothymi$3 or rapid) adj cycl$3) or ultradian) adj5 cycl$3).tw.
16	(mania or manic or hypomania).tw.

17 Anorexia/
18 (eating and disorder$2).tw.
19 anorexia nervosa.tw.
20 Bulimia/
21 (bulimia or hyperphagia).tw.
22 (bing$4 or overeat$3 or (compulsive and (eat$3 or vomit$3)) or (self induce$2 and vomit$3) or (restrict$3 and eat$3)).tw.
23 Anxiety/
24 Anxiety, Separation/
25 Panic Disorder/
26 (anxious or anxiety or panic or phobia or phobic).tw.
27 Stress, Psychological/
28 ((((((post adj traumatic$2) or posttraumatic$2 or stress) adj disorder$2) or acute) adj stress) or PTSD or ASD or DESNOS).tw.
29 (((((((combat or concentration) adj camp adj syndrome) or extreme) adj stress) or flash) adj back$2) or flashback$2).tw.
30 (((((hypervigilan$2 or hypervigilen$2 or psycholog$4) adj stress) or psycho) adj (trauma or traumatic)) or psychotrauma or psychtraumatic).tw.
31 (((((((((railway adj spine) or rape) and trauma$5) or torture) adj syndrome) or traumatic) adj neuros$2) or traumatic) adj stress).tw.
32 (trauma$5 and (avoidance or grief or horror or death$1 or (night adj mare$1) or nightmare$1 or emotion$2)).tw.
33 (recurr$5 adj thought$2).tw.
34 Obsessive Behavior/
35 obsessi$4.tw.
36 OCD.tw.
37 osteochondr$3.tw.
38 36 not 37
39 Compulsive Behavior/
40 compuls$5.tw.
41 (personalit$3 adj (disorder$1 or difficult$3)).tw.
42 (antisocial or anti-social or psychopath$3 or borderline or hysteri$3).tw.
43 (dissociat$3 adj (disorder$ or difficult$3 or personalit$3 or disturb$5 or trauma$4)).tw.
44 schizo$9.tw.
45 (psychos$2 or psychotic$2 or parano$3).tw.
46 (somatoform or somatization or somatic or hyperchondria$3 or neurasthenia$3).tw.
47 (conversion adj disorder$1).tw.
48 ((body adj dysmorphic) or dysmorphobi$2 or (briquet adj syndrome$1) or (syndrome adj briquet)).tw.
49 ((attach$ or bond or bonding) adj5 (ambivalent or anxious$4 or avoid$4 or difficult$3 or disinhibit$3 or disorder$2 or disorganis$4 or disruptiv$4 or dissociat$4 or dysregula$4 or disorientat$4 or disturbance$2 or impair$5 or inadequate or inhibit$3 or injur$4 or insecur$5 or poor or style$2)).tw.
50 ((adjust$4 or reactive) adj5 (disorder$1 or disturb$5 or difficult$3)).tw.

51	(transient situational adj (disturb$5 or disorder$1)).tw.
52	kleptomani$3.tw.
53	((drug$2 or substance$2) adj (abuse or misuse or depend$6 or addict$4)).tw.
54	(neurotic$2 or neuros?s).tw.
55	(affective adj5 disorder$1).tw.
56	Mental Disorders/
57	(mental adj (illness$2 or disease$1 or disorder$1)).tw.
58	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57
59	Prenatal Diagnosis/
60	58 and 59
61	(screen$3 or detect$3 or test or question$5 or surveillance).tw.
62	Mass screening/
63	61 or 62
64	(pregnan$ or antenatal$ or prenatal$).tw.
65	exp Pregnancy/
66	64 or 65
67	58 and 63 and 66
68	60 or 67
69	"Sensitivity and Specificity"/
70	(sensitiv$ or specific$).tw.
71	((false or true) adj (negative$ or positive$)).tw.
72	"Predictive Value of Tests"/
73	((positive or negative) adj predictive value$).tw.
74	69 or 70 or 71 or 72 or 73
75	68 and 74
76	limit 75 to (english language and yr="2006 -Current")
77	(case reports or comment or editorial or letter or news or "review").pt.
78	76 not 77
79	exp animals/ not humans.sh.
80	78 not 79
81	limit 76 to "reviews (maximizes specificity)"
82	80 or 81

Question 4.

# ▲	Searches
1	Adjustment Disorders/
2	Anxiety Disorders/
3	Dissociative Disorders/
4	Eating Disorders/
5	Mood Disorders/

6 Neurotic Disorders/
7 Personality Disorders/
8 exp Schizophrenia/
9 Somatoform Disorders/
10 Substance-Related Disorders/
11 Depression/
12 (Seasonal affective disorder$2 or depress$4 or Dysthym$4).tw.
13 melanchol$3.tw.
14 ((bipolar or bi polar) and (disorder$2 or depress$4)).tw.
15 ((((cyclothymi$3 or rapid) adj cycl$3) or ultradian) adj5 cycl$3).tw.
16 (mania or manic or hypomania).tw.
17 Anorexia/
18 (eating and disorder$2).tw.
19 anorexia nervosa.tw.
20 Bulimia/
21 (bulimia or hyperphagia).tw.
22 (bing$4 or overeat$3 or (compulsive and (eat$3 or vomit$3)) or (self induce$2 and vomit$3) or (restrict$3 and eat$3)).tw.
23 Anxiety/
24 Anxiety, Separation/
25 Panic Disorder/
26 (anxious or anxiety or panic or phobia or phobic).tw.
27 Stress, Psychological/
28 ((((((post adj traumatic$2) or posttraumatic$2 or stress) adj disorder$2) or acute) adj stress) or PTSD or ASD or DESNOS).tw.
29 (((((((combat or concentration) adj camp adj syndrome) or extreme) adj stress) or flash) adj back$2) or flashback$2).tw.
30 (((((hypervigilan$2 or hypervigilen$2 or psycholog$4) adj stress) or psycho) adj (trauma or traumatic)) or psychotrauma or psychtraumatic).tw.
31 (((((((((railway adj spine) or rape) and trauma$5) or torture) adj syndrome) or traumatic) adj neuros$2) or traumatic) adj stress).tw.
32 (trauma$5 and (avoidance or grief or horror or death$1 or (night adj mare$1) or nightmare$1 or emotion$2)).tw.
33 (recurr$5 adj thought$2).tw.
34 Obsessive Behavior/
35 obsessi$4.tw.
36 OCD.tw.
37 osteochondr$3.tw.
38 36 not 37
39 Compulsive Behavior/
40 compuls$5.tw.
41 (personalit$3 adj (disorder$1 or difficult$3)).tw.
42 (antisocial or anti-social or psychopath$3 or borderline or hysteri$3).tw.

43 (dissociat$3 adj (disorder$ or difficult$3 or personalit$3 or disturb$5 or trauma$4)).tw.
44 schizo$9.tw.
45 (psychos$2 or psychotic$2 or parano$3).tw.
46 (somatoform or somatization or somatic or hyperchondria$3 or neurasthenia$3).tw.
47 (conversion adj disorder$1).tw.
48 ((body adj dysmorphic) or dysmorphobi$2 or (briquet adj syndrome$1) or (syndrome adj briquet)).tw.
49 ((attach$ or bond or bonding) adj5 (ambivalent or anxious$4 or avoid$4 or difficult$3 or disinhibit$3 or disorder$2 or disorganis$4 or disruptiv$4 or dissociat$4 or dysregula$4 or disorientat$4 or disturbance$2 or impair$5 or inadequate or inhibit$3 or injur$4 or insecur$5 or poor or style$2)).tw.
50 ((adjust$4 or reactive) adj5 (disorder$1 or disturb$5 or difficult$3)).tw.
51 (transient situational adj (disturb$5 or disorder$1)).tw.
52 kleptomani$3.tw.
53 ((drug$2 or substance$2) adj (abuse or misuse or depend$6 or addict$4)).tw.
54 (neurotic$2 or neuros?s).tw.
55 (affective adj5 disorder$1).tw.
56 Mental Disorders/
57 (mental adj (illness$2 or disease$1 or disorder$1)).tw.
58 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57
59 Prenatal Diagnosis/
60 58 and 59
61 (screen$3 or detect$3 or test or question$5 or surveillance).tw.
62 Mass screening/
63 61 or 62
64 (pregnan$ or antenatal$ or prenatal$).tw.
65 exp Pregnancy/
66 64 or 65
67 58 and 63 and 66
68 60 or 67
69 Meta-Analysis/
70 meta-analys?s.tw.
71 systematic review.tw.
72 Clinical Trial/
73 Controlled Clinical Trial/
74 Cross-Over Studies/
75 Random Allocation/
76 RandomisedRandomised Controlled Trial/
77 (random* adj5 (alloca* or assign* or control*)).tw.
78 (clinical adj5 trial*).tw.
79 (compar* adj5 (report* or stud* or trial*)).tw.
80 Longitudinal Studies/

81	69 or 70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80
82	68 and 81
83	limit 82 to (english language and yr="2006 -Current")
84	(case reports or comment or editorial or letter or news or "review").pt.
85	83 not 84
86	exp animals/ not humans.sh.
87	85 not 86

Question 5.

# ▲	Searches
1	Depression, Postpartum/
2	((postpart* or post-part* or postnatal or post-natal or puerperal) adj5 (depress* or mood disorder? or mood disturbance? or melanchol*)).ti,ab.
3	1 or 2
4	(screen$3 or detect$3 or test or question$5 or surveillance).tw.
5	diagnos*.ti.
6	Mass Screening/
7	4 or 5 or 6
8	3 and 7
9	*Depression, Postpartum/di [Diagnosis]
10	8 or 9
11	"Sensitivity and Specificity"/
12	(sensitiv$ or specific$).tw.
13	((false or true) adj (negative$ or positive$)).tw.
14	"Predictive Value of Tests"/
15	((positive or negative) adj predictive value$).tw.
16	11 or 12 or 13 or 14 or 15
17	10 and 16
18	(case reports or comment or editorial or letter or news or "review").pt.
19	17 not 18
20	exp animals/ not humans.sh.
21	19 not 20
22	limit 21 to (english language and yr="2011 -Current")
23	limit 17 to "reviews (maximizes specificity)"
24	limit 23 to (english language and yr="2011 -Current")
25	22 or 24

Question 6.

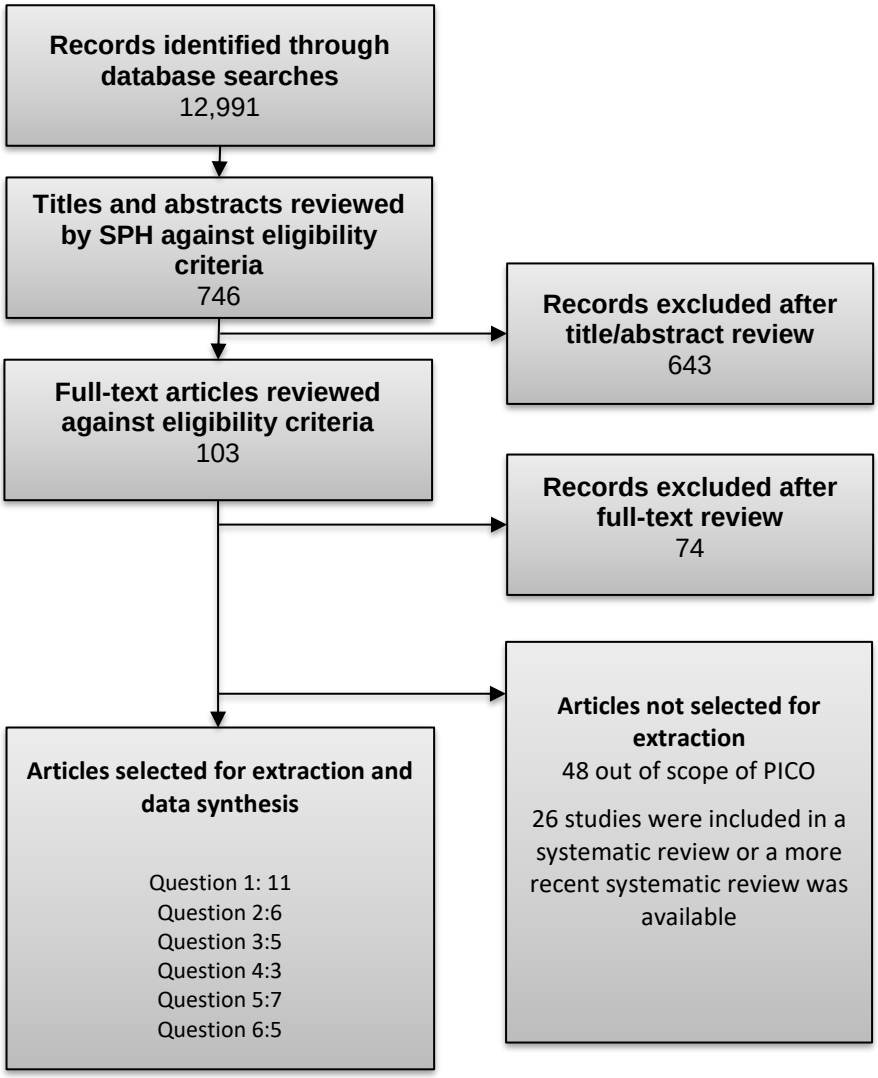
# ▲	Searches
1	Depression, Postpartum/
2	((postpart* or post-part* or postnatal or post-natal or puerperal) adj5 (depress* or mood disorder? or mood disturbance? or melanchol*)).ti,ab.
3	1 or 2

4 limit 3 to "reviews (maximizes specificity)"
5 limit 4 to (english language and yr="2011 -Current")
6 Clinical Trial/
7 Controlled Clinical Trial/
8 Cross-Over Studies/
9 Random Allocation/
10 RandomisedRandomised Controlled Trial/
11 (random* adj5 (alloca* or assign* or control*)).tw.
12 (clinical adj5 trial*).tw.
13 (compar* adj5 (report* or stud* or trial*)).tw.
14 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13
15 3 and 14
16 limit 15 to (english language and yr="2011 -Current")
17 (case reports or comment or editorial or letter or news or "review").pt.
18 16 not 17
19 5 or 18

Results were imported into EndNote and de-duplicated.

Appendix 2 — Included and excluded studies

Figure 1: PRISMA flowchart



1 summarises the volume of publications included and excluded at each stage of the review. 103 publications were ultimately judged to be relevant to one or more review questions and were considered for extraction. Publications that were included or excluded after the review of full-text articles are detailed below.

Publications included after review of full-text articles

The 29 publications included after review of full-texts are summarised in Table 15 below. It was planned *a priori* that the following approach would be taken to prioritise studies for extraction:

Systematic reviews and meta-analyses would be considered the highest quality of evidence if any were found. Studies relating to epidemiology would be prioritised if they considered a UK population

Publications reviewed at full text but not selected for extraction and data synthesis are clearly detailed in Table 16 below.

Table 13. Summary of publications included after review of full-text articles, and the question(s) each publication was identified as being relevant to

Study	The condition	Current adherence to guidance	The test	The intervention
.Cook et al (2018)	✓			
.Howard et al (2018)			✓	
.Mackinnon et al (2018)				
.Closa-Monsterolo et al (2017)	✓			
.Ford et al (2017)		✓		
.Gollan et al (2017)			✓	
.Letourneau et al (2017)				✓
.Lui et al (2017)	✓			
.Prady et al (2017)				✓
.Pritchett et al (2017)				✓
.van Ravesteyn et al (2017)				✓
.Vlenterie et al (2017)			✓	
.Weinreb et al (2017)				✓
.Gelaye et al 2016	✓			
Grote et al (2016)				✓
.Jarde et al 2016	✓			
McGlone et al (2016)		✓		
.Morrell et al (2016)				✓
.O'Connor et al (2016)			✓	✓
.Owora et al (2016)			✓	
.Prady et al (2016)	✓	✓		
.Redshaw et al (2016)		✓		
.Williams et al (2016)		✓		
.Darwin et al (2015)		✓		
.Sanger et al (2015)	✓			

.Leis et al (2014)	✓		
.McDonagh et al (2014)			✓
.Tarabulsy et al (2014)	✓		
.Meades et al (2011)		✓	

Publications excluded after review of full-text articles

Of the 103 publications included after the review of titles and abstracts, 74 were ultimately judged not to be relevant to this review. These publications, along with reasons for exclusion, are listed in Table 16.

Table 14. Publications excluded after review of full-text articles

	Reference	Reason for exclusion
Question 1 What adverse outcomes have been reported from common mental health problems in pregnancy and in the postnatal period?		
1	Farias-Antunez S, Xavier MO, Santos IS. Effect of maternal postpartum depression on offspring's growth. Journal of Affective Disorders. 2018 01 Mar;228:143-52	Examining PND and low income on offspring outcomes
2	Gentile S. Untreated depression during pregnancy: Short- and long-term effects in offspring. A systematic review. Neuroscience. 2017 07 Feb;342:154-66.	A more recent systematic review available.
3	Eyden J, Winsper C, Wolke D, Broome MR, MacCallum F. A systematic review of the parenting and outcomes experienced by offspring of mothers with borderline personality pathology: Potential mechanisms and clinical implications. Clinical Psychology Review. [Review]. 2016 01 Jul;47:85-105.	The study selection included those whose participants had been diagnosed with a mental health problem prior to pregnancy - not in PICO scope.
4	Rose MS, Pana G, Premji S. Prenatal Maternal Anxiety as a Risk Factor for Preterm Birth and the Effects of Heterogeneity on This Relationship: A Systematic Review and Meta-Analysis. BioMed Research International. 2016;2016 (no pagination)(8312158).	In many of the included studies prenatal maternal anxiety was not determined in a general population but rather participants were recruited as part of a sub group such as those at high risk of intrauterine growth restriction, low income, high medical risk. Out of PICO scope
5	Accortt EE, Cheadle AC, Dunkel Schetter C. Prenatal depression and adverse birth outcomes: an updated systematic review. Maternal and child health journal. 2015 01 Jun;19(6):1306-37.	More recent systematic reviews and meta-analyses are available
6	Kingston D, McDonald S, Austin MP, Tough S. Association between Prenatal and Postnatal Psychological Distress and Toddler Cognitive Development: A Systematic Review. PloS one. 2015;10(5):e0126929.	More recent systematic reviews and meta analyses are available
7	Kingston D, Tough S. Prenatal and postnatal maternal mental health and school-age child development: a systematic review. Maternal and child health journal. 2014 01 Sep;18(7):1728-41.	More recent systematic reviews and meta-analyses are available
8	Kingston D, Tough S, Whitfield H. Prenatal and postpartum maternal psychological	More recent systematic reviews and meta-analyses available

	distress and infant development: a systematic review. Child psychiatry and human development. 2012 Oct;43(5):683-714.	
Question 2 Is clinical detection and management currently well implemented in the UK		
9	Ford E, Shakespeare J, Elias F, Ayers S. Recognition and management of perinatal depression and anxiety by general practitioners: A systematic review. Family Practice. [Review]. 2017;34(1):11-9.	The 4 UK papers in Ford et al 2017 have been separately assessed for inclusion in this review
10	Marcano-Belisario JS, Gupta AK, O'Donoghue J, Ramchandani P, Morrison C, Car J. Implementation of depression screening in antenatal clinics through tablet computers: results of a feasibility study. BMC Medical Informatics & Decision Making. 2017;17(1):59	About screening technology not current practice
11	Matthey S, Souter K, Mortimer K, Stephens C, Sheridan-Magro A. Routine antenatal maternal screening for current mental health: evaluation of a change in the use of the Edinburgh Depression Scale in clinical practice. Archives of Women's Mental Health. [Evaluation Studies]. 2016;19(2):367-72.	Australian cohort not UK
12	Prady SL, Pickett KE, Petherick ES, Gilbody et al Variation and ethnic inequalities in treatment of common mental disorders before, during and after pregnancy: combined analysis of routine and research data in the Born in Bradford cohort BMC Psychiatry 2016;16:99	Women selected were diagnosed with mental health disorder prior to becoming pregnant
13	Amiel Castro RT, Schroeder K, Pinard C, Blochlinger P, Kunzli H, Riecher-Rossler A, et al. Perinatal mental health service provision in Switzerland and in the UK. Swiss Medical Weekly. [Comparative StudyResearch Support, Non-U.S. Gov't]. 2015;145:w14011.	Descriptive comparison of UK and Swiss systems
14	Edge D. Falling through the net - black and minority ethnic women and perinatal mental healthcare: health professionals' views. General Hospital Psychiatry. 2010;32(1):17-25	Before NICE guidance 2014 setting out guidance for what should be standard 'current' practice
15	Alder EM, Reid M, Sharp LJ, Cantwell R, Robertson K, Kearney E. Policy and practice in the management of postnatal depression in Scotland. Archives of Women's Mental Health. [Research Support, Non-U.S. Gov't]. 2008;11(3):213-9.	More recent studies are available. This is published 6 years prior to NICE guidance published in 2014
16	Stanley N, Borthwick R, Macleod A. Antenatal depression: Mothers' awareness and professional responses. Primary Health Care Research and Development. 2006 Jul;7(3):257-68.	More recent studies available. This is published 8 years prior to NICE guidance published in 2014
Question 3 What is the reported accuracy of screening tools to detect common mental health problems during pregnancy		

17	Coates R, Ayers S, de Visser R. Factor structure of the Edinburgh Postnatal Depression Scale in a population-based sample. Psychological Assessment. 2017 Aug;29(8):1016-27	Concerned with the factor structure of EPDS not accuracy. No Sensitivity, specificity, PPV, or NPV
18	Friesen K, Peterson WE, Squires J, Fortier C. Validation of the Edinburgh Postnatal Depression Scale for Use With Young Childbearing Women. Journal of Nursing Measurement. [Validation Studies]. 2017;25(1):1-16.	Population significant proportion under 18
19	Gelaye B, Zheng Y, Medina-Mora ME, Rondon MB, Sanchez SE, Williams MA. Validity of the posttraumatic stress disorders (PTSD) checklist in pregnant women. BMC Psychiatry. [Validation Studies]. 2017;17(1):179.	Study aimed to validate Spanish version of the checklist in Lima, Peru
20	Hirsch NM, Fingerhut R, Allison KC. The Prenatal Distress Measure: Adaptation of the postpartum distress measure for a prenatal sample. Journal of Women's Health. 2017 Nov;26(11):1193-200	No way of determining sensitivity specificity PPV or NPV
21	Venkatesh KK, Kaimal AJ, Castro VM, Perlis RH. Improving discrimination in antepartum depression screening using the Edinburgh Postnatal Depression Scale. Journal of Affective Disorders. [Observational Study]. 2017;214:1-7	Reported in later systematic review
22	Brodey BB, Goodman SH, Baldasaro RE, Brooks-DeWeese A, Wilson ME, Brodey IS, et al. Development of the Perinatal Depression Inventory (PDI)-14 using item response theory: A comparison of the BDI-II, EPDS, PDI, and PHQ-9. Archives of Women's Mental Health. 2016 Apr;19(2):307-16	No way of determining sensitivity specificity PPV or NPV
23	Darwin Z, McGowan L, Edozien LC. Identification of women at risk of depression in pregnancy: using women's accounts to understand the poor specificity of the Whooley and Arroll case finding questions in clinical practice. Archives of Women's Mental Health. [Research Support, Non-U.S. Gov't]. 2016;19(1):41-9.	More recent study available
24	Brunton RJ, Dryer R, Saliba A, Kohlhoff J. Pregnancy anxiety: A systematic review of current scales. Journal of Affective Disorders. 2015 01 May;176:24-34.	No way of determining sensitivity specificity PPV or NPV
25	Cunningham NK, Brown PM, Page AC. Does the Edinburgh Postnatal Depression Scale measure the same constructs across time? Archives of Women's Mental Health. [Research Support, Non-U.S. Gov't]. 2015;18(6):793-804.	No way of determining sensitivity specificity PPV or NPV

26	Evans K, Spiby H, Morrell C. A psychometric systematic review of self-report instruments to identify anxiety in pregnancy. Journal of Advanced Nursing. 2015 Sep;71(9):1986-2001	No way of determining sensitivity specificity PPV or NPV
27	Kozinszky Z, Dudas RB. Validation studies of the Edinburgh Postnatal Depression Scale for the antenatal period. Journal of Affective Disorders. 2015 May;176:95-105	More recent systematic review incorporating these studies is available
28	Husain N, Rahman A, Husain M, Khan SM, Vyas A, Tomenson B, et al. Detecting depression in pregnancy: validation of EPDS in British Pakistani mothers. Journal of Immigrant & Minority Health. [Research Support, Non-U.S. Gov't]. 2014;16(6):1085-92	More recent study available
29	Simpson W, Glazer M, Michalski N, Steiner M, Frey BN. Comparative efficacy of the generalized anxiety disorder 7-item scale and the Edinburgh Postnatal Depression Scale as screening tools for generalized anxiety disorder in pregnancy and the postpartum period. Canadian Journal of Psychiatry - Revue Canadienne de Psychiatrie. [Comparative Study Validation Studies]. 2014;59(8):434-40.	Not a screening population but already referred to psychiatric service
30	Prady SL, Miles JNV, Pickett KE, Fairley L, Bloor K, Gilbody S, et al. The psychometric properties of the subscales of the GHQ-28 in a multi-ethnic maternal sample: results from the Born in Bradford cohort. BMC Psychiatry. [Research Support, Non-U.S. Gov't]. 2013;13:55.	No way of determining sensitivity specificity PPV or NPV
31	Matthey S, Ross-Hamid C. Repeat testing on the Edinburgh Depression Scale and the HADS-A in pregnancy: Differentiating between transient and enduring distress. Journal of Affective Disorders. 2012 10 Dec;141(2-3):213-21.	Study included in more recent systematic review.
32	Ji S, Long Q, Newport D, Na H, Knight B, Zach EB, et al. Validity of depression rating scales during pregnancy and the postpartum period: Impact of trimester and parity. Journal of Psychiatric Research. 2011 Feb;45(2):213-9.	No way of determining sensitivity specificity PPV or NPV
33	Austin M-P, Priest SR, Sullivan EA. Antenatal psychosocial assessment for reducing perinatal mental health morbidity. Cochrane Database of Systematic Reviews [serial on the Internet]. 2008; (4):	Study testing the scale for validity
Question 4 What is the reported accuracy of screening tools to detect postnatal depression		

34	McKean M, Caughey AB, Yuracko McKean MA, Cabana MD, Flaherman VJ. Postpartum Depression: When Should Health Care Providers Identify Those at Risk? Clinical Pediatrics. 2017:9.	Not a consecutively enrolled population and no test statistics
35	Knights JE, Salvatore ML, Simpkins G, Hunter K, Khandelwal M. In search of best practice for postpartum depression screening: is once enough? European Journal of Obstetrics Gynecology and Reproductive Biology. 2016 01 Nov;206:99-104.	Not a consecutively enrolled population and no test statistics
36	Santos IS, Tavares BF, Munhoz TN, Manzolli P, de Avila GB, Jannke E, et al. Patient health questionnaire-9 versus Edinburgh postnatal depression scale in screening for major depressive episodes: a cross-sectional population-based study. BMC Research Notes. 2016;9(1):453	Population was not postpartum women
37	Yawn BP, Bertram S, Kurland M, Wollan PC. Repeated depression screening during the first postpartum year. Annals of Family Medicine. 2015 May-Jun;13(3):228-34	Not a consecutively enrolled population or assessing screening tool for accuracy
38	Walker LO, Gao J, Xie B. Postpartum psychosocial and behavioural health: A systematic review of self-administered scales validated for postpartum women in the United States. Women's Health Issues. 2015 Sep-Oct;25(5):586-600.	All studies in systematic review are reported in a more recent included review
39	Meijer JL, Beijers C, van Pampus MG, Verbeek T, Stolk RP, Milgrom J, et al. Predictive accuracy of Edinburgh postnatal depression scale assessment during pregnancy for the risk of developing postpartum depressive symptoms: a prospective cohort study. BJOG: An International Journal of Obstetrics & Gynaecology. [Research Support, Non-U.S. Gov't]. 2014;121(13):1604-10.	EPDS used in pregnancy to predict PND
40	Davis K, Pearlstein T, Stuart S, O'Hara M, Zlotnick C. Analysis of brief screening tools for the detection of postpartum depression: comparisons of the PRAMS 6-item instrument, PHQ-9, and structured interviews. Archives of Women's Mental Health. [Comparative StudyResearch Support, N.I.H., Extramura lResearch Support, Non-U.S. Gov't]. 2013;16(4):271-7.	Study is in a systematic review included in this review
41	Johnson M, Schmeid V, Lupton SJ, Austin MP, Matthey SM, Kemp L, et al. Measuring perinatal mental health risk. Archives of Women's Mental Health. [Review]. 2012;15(5):375-86.	More recent systematic review available

42	McCabe K, Blucker R, Gillaspy J, Cherry A, Mignogna M, Roddenberry A, et al. Reliability of the Postpartum Depression Screening Scale in the neonatal intensive care unit. Nursing Research. 2012 Nov-Dec;61(6):441-5	Not a screening population but based on parents of children in NICU
43	O'Hara MW, Stuart S, Watson D, Dietz PM, Farr SL, D'Angelo D. Brief scales to detect postpartum depression and anxiety symptoms. Journal of Women's Health. [Research Support, U.S. Gov't, P.H.S.]. 2012;21(12):1237-43	Study is reported in more recent systematic review
44	Yawn BP, Dietrich AJ, Wollan P, Bertram S, Graham D, Huff J, et al. TRIPPD: a practice-based network effectiveness study of postpartum depression screening and management. Annals of Family Medicine. [RandomisedRandomised Controlled TrialResearch Support, N.I.H., Extramural]. 2012;10(4):320-9.	Not a consecutively enrolled population or assessing screening tool for accuracy
45	Burton A, Patel S, Kaminsky L, Rosario GD, Young R, Fitzsimmons A, et al. Depression in pregnancy: Time of screening and access to psychiatric care. Journal of Maternal-Fetal and Neonatal Medicine. 2011 November;24(11):1321-4.	US study using EPDS focused on access to psychiatric care. More recent studies of EPDS in UK
46	Magnusson M, Lagerberg D, Sundelin C. How can we identify vulnerable mothers who do not reach the cut off 12 points in EPDS? Journal of Child Health Care. [Comparative StudyResearch Support, Non-U.S. Gov't]. 2011;15(1):39-49.	Study is not assessing screening tool for accuracy
47	Meades R, Ayers S. Anxiety measures validated in perinatal populations: a systematic review. Journal of Affective Disorders. 2011;133(1-2):1-15	Focus on anxiety instruments rather than postnatal depression
Question 5: What are the benefits of pharmacological and non-pharmacological intervention (alone or in combination) in women with screen-detected mental health problems during pregnancy?		
48	Andalib S, Emamhadi MR, Yousefzadeh-Chabok S, Shakouri SK, Hoilund-Carlsen PF, Vafaee MS, et al. Maternal SSRI exposure increases the risk of autistic offspring: A meta-analysis and systematic review. European Psychiatry: the Journal of the Association of European Psychiatrists. 2017;45:161-6.	Studies included are not based on screen detected
49	Dimidjian S, Goodman SH, Sherwood NE, Simon GE, Ludman E, Gallop R, et al. A pragmatic randomisedrandomised clinical trial of behavioural activation for depressed pregnant women. Journal of Consulting & Clinical Psychology. [RandomisedRandomised Controlled Trial]. 2017;85(1):26-36.	Cohort not screen-detected - receiving care at a mental health facility

50	Forsell E, Bendix M, Hollandare F, Szymanska von Schultz B, Nasiell J, Blomdahl-Wetterholm M, et al. Internet delivered cognitive behaviour therapy for antenatal depression: A randomised controlled trial. Journal of Affective Disorders. 2017;221:56-64.	Cohort not screen-detected but recruited through the media
51	Zhang T-N, Gao S-Y, Shen Z-Q, Li D, Liu C-X, Lv H-C, et al. Use of selective serotonin-reuptake inhibitors in the first trimester and risk of cardiovascular-related malformations: a meta-analysis of cohort studies. Scientific Reports. 2017;7:43085.	Studies included are not based on screen-detected populations
52	Eke AC, Saccone G, Berghella V. Selective serotonin reuptake inhibitor (SSRI) use during pregnancy and risk of preterm birth: a systematic review and meta-analysis. BJOG: An International Journal of Obstetrics and Gynaecology. 2016 01 Nov;123(12):1900-7	Most included studies reported effects in first trimester so are unlikely to be screen-detected cohorts
53	Sheffield KM, Woods-Giscombe CL. Efficacy, Feasibility, and Acceptability of Perinatal Yoga on Women's Mental Health and Well-Being: A Systematic Literature Review. Journal of holistic nursing : official journal of the American Holistic Nurses' Association. 2016 01 Mar;34(1):64-79	Studies varied in study participants with populations comprising women with no mental health problems, current psychiatric disorder, diagnosed clinical depression and/or anxiety. Not screen-detected populations
54	Milgrom J, Holt C, Holt CJ, Ross J, Ericksen J, Gemmill AW. Feasibility study and pilot randomised trial of an antenatal depression treatment with infant follow-up. Archives of Women's Mental Health. [RandomisedRandomised Controlled TrialResearch Support, Non-U.S. Gov't]. 2015;18(5):717-30	Included in recent systematic review
55	Bittner A, Peukert J, Zimmermann C, Junge-Hoffmeister J, Parker LS, Stobel-Richter Y, et al. Early intervention in pregnant women with elevated anxiety and depressive symptoms: Efficacy of a cognitive-behavioural group program. The Journal of Perinatal & Neonatal Nursing. 2014 Jul-Sep;28(3):185-95.	More recent systematic review available
56	Daley A, Foster L, Long G, Palmer C, Robinson O, Walmsley H, et al. The effectiveness of exercise for the prevention and treatment of antenatal depression: systematic review with meta-analysis (Provisional abstract). Database of Abstracts of Reviews of Effects [serial on the Internet]. 2014; (2	Studies varied in study participants with populations comprising women with no mental health problems, current psychiatric disorder, diagnosed clinical depression and/or anxiety. Not screen-detected populations
57	Fontein-Kuipers YJ, Nieuwenhuijze MJ, Ausems M, Bude L, de Vries R. Antenatal interventions to reduce maternal distress: a systematic review and meta-analysis of	More recent systematic review available

	randomised trials. BJOG: An International Journal of Obstetrics & Gynaecology. [Meta-AnalysisResearch Support, Non-U.S. Gov'tReview]. 2014;121(4):389-97.	
58	Marc I, Toureche N, Ernst E, Hodnett ED, Blanchet C, Dodin S, et al. Mind-body interventions during pregnancy for preventing or treating women's anxiety. Cochrane Database of Systematic Reviews [serial on the Internet]. 2011; (7):	More recent systematic review available
Question 6 - What are the benefits of early pharmacological and non-pharmacological intervention (alone or in combination) in women with screen detected screen-detected postnatal depression?		
59	Stein A, Netsi E, Lawrence PJ, Granger C, Kempton C, Craske MG, et al. Mitigating the effect of persistent postnatal depression on child outcomes through an intervention to treat depression and improve parenting: a randomised controlled trial. The Lancet Psychiatry. 2018;5(2):134-44.	Study did not use a screen-detected population
60	Felice E, Agius A, Sultana R, Felice EM, Calleja-Agius J. The effectiveness of psychosocial assessment in the detection and management of postpartum depression: a systematic review. Minerva Ginecologica. 2017;09:09.	Detection of mental health problems rather than treatment
61	Lau Y, Htun TP, Wong SN, Tam WSW, Klainin-Yobas P. Therapist-Supported Internet-Based Cognitive Behavior Therapy for Stress, Anxiety, and Depressive Symptoms Among Postpartum Women: A Systematic Review and Meta-Analysis. Journal of medical Internet research. 2017 28 Apr;19(4):e138.	Studies in this systematic review that meet the evidence update PICO criteria are included in other systematic reviews included in this review.
62	McCurdy AP, Boule NG, Sivak A, Davenport MH. Effects of Exercise on Mild-to-Moderate Depressive Symptoms in the Postpartum Period: A Meta-analysis. Obstetrics & Gynecology. [Review]. 2017;129(6):1087-97.	More recent systematic review available
63	Molyneaux E, Howard LM, McGeown HR, Karia AM, Trevillion K. Antidepressant treatment for postnatal depression. Issues in Mental Health Nursing. 2014 Feb;38(2):188-90.	More recent systematic review available
64	Wisner KL, Sit DKY, McShea M, Luther JF, Eng HF, Dills JL, et al. Telephone-Based Depression Care Management for Postpartum Women: A RandomisedRandomised Controlled Trial. Journal of Clinical Psychiatry. 2017;78(9):1369-75.	Comparator is not ‘usual care’
65	Wozney L, Olthuis J, Lingley-Pottie P, McGrath PJ, Chaplin W, Elgar F, et al.	Study did not use a screen-detected population

	Strongest Families TM Managing Our Mood (MOM): a randomisedrandomised controlled trial of a distance intervention for women with postpartum depression. Archives of Women's Mental Health. [RandomisedRandomised Controlled Trial]. 2017;20(4):525-37	
66	Ashford MT, Olander EK, Ayers S. Computer- or web-based interventions for perinatal mental health: A systematic review. Journal of Affective Disorders. [Review]. 2016;197:134-46	Not a well recognised intervention used in the perinatal period
67	Kempler L, Sharpe L, Miller CB, Bartlett DJ. Do psychosocial sleep interventions improve infant sleep or maternal mood in the postnatal period? A systematic review and meta-analysis of randomised controlled trials. Sleep Medicine Reviews. 2016 Oct;29:15-22.	Study not focussed on postnatal depression
68	Milgrom J, Danaher BG, Gemmill AW, Holt C, Holt CJ, Seeley JR, et al. Internet cognitive behavioural therapy for women with postnatal depression: A randomisedrandomised controlled trial of MumMoodBooster. Journal of Medical Internet Research. 2016 Mar;18(3):No Pagination Specified	Study did not use a screen-detected population
69	Pugh NE, Hadjistavropoulos HD, Dirkse D. A Randomised Controlled Trial of Therapist-Assisted, Internet-Delivered Cognitive Behaviour Therapy for Women with Maternal Depression. PLoS ONE [Electronic Resource]. [RandomisedRandomised Controlled Trial Research Support, Non-U.S. Gov't]. 2016;11(3):e0149186	Study did not use a screen-detected population
70	Stephens S, Ford E, Paudyal P, Smith H. Effectiveness of psychological interventions for postnatal depression in primary care: A meta-analysis. Annals of Family Medicine. 2016 Sep-Oct;14(5):463-72	UK HTA covering same period available
71	Barlow J, Bennett C, Midgley N, Larkin SK, Wei Y. Parent-infant psychotherapy for improving parental and infant mental health. Cochrane Database of Systematic Reviews. 2015 08 Jan;2017 (6)	More recent systematic review available
72	Milgrom J, Gemmill AW, Ericksen J, Burrows G, Buist A, Reece J. Treatment of postnatal depression with cognitive behavioural therapy, sertraline and combination therapy: A randomised controlled trial. Australian and New Zealand Journal of Psychiatry. 2015 Mar;49(3):236-45.	Study is reported in more recent systematic review
73	Tsivos ZL, Calam R, Sanders MR, Wittkowski A. Interventions for postnatal depression assessing the mother-infant relationship and child developmental outcomes: A systematic	More recent systematic review available

	review. International Journal of Women's Health. 2015 23 Apr;7:429-47.	
74	Claridge AM. Efficacy of systemically oriented psychotherapies in the treatment of perinatal depression: a meta-analysis. Archi`ves of Women's Mental Health. [Meta-Analysis]. 2014;17(1):3-15.	More recent systematic review available

Appendix 3 — Summary and appraisal of individual studies

Data Extraction

Studies relevant to question 1: what adverse outcomes have been reported from mental health problems in pregnancy and in the postnatal period?

Table 15. Cook et al (2018)

Publication	**Cook N, Ayers S, Horsch A. Maternal posttraumatic stress disorder during the perinatal period and child outcomes: A systematic review. Journal of Affective Disorders. [Review]. 2018;225:18-31.**
Study details	Systematic review
Study objectives	The study aimed to systematically review and summarize research investigating the association between maternal PTSD during the perinatal period (beginning of pregnancy until first year postpartum) and child outcomes.
Inclusions	Report of PTSD symptoms and/or diagnosis during the perinatal period, which was measured using a questionnaire, interview, or clinical code.
Exclusions	None stated
Population	26 publications reporting 21 studies
Intervention	N/A
Comparator	N/A

Outcomes	Child outcomes are grouped into postpartum birth outcomes, child development, and mother-infant relationship.
	• 11 papers reported birth outcomes of women with antenatal PTSD. The largest of the studies (n=2487 of which 129 had likely PTSD) showed risk of preterm birth in women with likely PTSD (Adjusted OR = 1.22, 95% CI,0.57 to 2.61). A further large study showed a 3x risk of preterm delivery with PTSD; however, 3 other studies showed no significant increased risk of preterm delivery. No association of PTSD and gestational age was reported.
	• 4 papers investigated the association between cognitive development of infants and mothers with postpartum PTSD and 2 reported poorer cognitive outcomes at 17 months. The 2 other studies did not find an association (no figures given).
	• 11 papers investigated postpartum PTSD and the mother-infant interaction.
	  ○ 2 studies found no association between mothers with higher scores for postpartum PTSD and maternal sensitivity or control whilst 2 other studies did report an association with these factors (figures not given).
	  ○ 3 studies investigating postpartum PTSD and infant interactional behaviour were inconsistent (no figures given)
	  ○ 7 studies investigating the mother infant bond were inconsistent with 3 reporting no relationship with postpartum PTSD and the quality of the mother infant bond whilst 4 reported that mothers perceived that

	the relationship with their baby was not optimal and were more likely to have a negative view, show more infant directed hostility or less desire for proximity (no figures given)
Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review Validated measures were not used to collect postpartum birth outcome information in any of the 11 studies examining birth outcomes included in this review, which may limit the reliability and validity of the findings. Very few figures were extracted from the papers to support the findings described.

Table 16. Farias-Antunez et al 2018

Publication	**Farias-Antunez S, Xavier MO, Santos IS. Effect of maternal postpartum depression on offspring's growth. Journal of Affective Disorders. 2018 01 Mar; 228:143-52.**
Study details	Systematic review
Study objectives	To evaluate the evidence about maternal postpartum depression on offspring growth outcomes to age 18
Inclusions	Cohort and case control studies that analysed exposure to maternal postnatal depression and effect on offspring growth up to 18 years published to February 2017.
Exclusions	Non peer reviewed articles
Population	20 articles of which 2 were case control and 18 were cohort studies with outcomes of mothers and offspring (from 0 to 11 years) published between 2003 and 2016.
Intervention	N/A
Comparator	N/A
Outcomes	All but 1 of the 6 studies reporting height and/or weight in children under 12 months were from low income counties and so results are less likely to be generalizable to the UK. Weight outcomes to age 12 months(n=5 studies) • All studies (Nigeria, India, Pakistan and Bangladesh populations) reported depressed mothers having a higher chance of an underweight child at age 3 to 12 months (range of effect size OR=2.8; 95% CI, 1.1 to 7.3 to OR=7.4 95% CI, 1.6 to 38.5) and in one study β= -0.38;(95% CI, -0.72 to -0.65) Height outcomes to age 12 months(n=5 studies) • All studies (US, India, Nigeria, Bangladesh, Pakistan) reported depressed mothers as having a higher chance of a child with lower height at age 3 to 12 months(range of effect size OR=1.61; 95%CI, 1.11 to 2.24 to OR=3.28;95% CI, 1.03 to 10.47) and in one study β=-0.26; 95% CI, 0.54 to 0.01). • The US study(n=4745 mothers) reported this effect only among those with a lower income (OR1.65, 95% CI 1.10-2.48) vs higher income(OR =1.38 95% CI, 0.54-3.54) Weight outcomes to age 11 years was reported in 13 studies from Denmark, US, South Africa, , Netherlands, Brazil, and a multicentre European study. • No effect of maternal postnatal depression (≤3months) on weight status of children (aged 12 months to 11 years) was reported by any study. Height outcomes to 11 years was reported in 7 studies from the US, Brazil and South Africa. • No effect of maternal postnatal depression (≤3months) on restricted height status of children (aged 12 months to 11 years) was reported by any study.

Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review. The studies chosen for inclusion in this systematic review were assessed for quality by the authors and those 2 studies with lower quality scores were as a result of poor adjustment for confounders. All but 1 of the 6 studies reporting height and/or weight in children under 12 months were from lower income counties and so results are less likely to be generalizable to the UK.

Table 17. MacKinnon et al 2018

Publication	**MacKinnon N, Kingsbury M, Mahedy L, Evans J, Colman I. The Association Between Prenatal Stress and Externalizing Symptoms in Childhood: Evidence From the Avon Longitudinal Study of Parents and Children. Biological Psychiatry. 2018;83(2):100-8.**
Study details	Longitudinal study
Study objectives	To examine whether children of mothers reporting higher levels of antenatal stressful events would be at higher risk of exhibiting symptoms of externalising disorders such as ADHD and conduct disorder between the ages of 7 to 16.
Inclusions	Prospective birth cohort data from the Avon Longitudinal Study of Parents and Children (ALSPAC). All individuals with outcome data were included in the analysis
Exclusions	None stated
Population	10,184 mother-offspring pairs
Intervention	N/A
Comparator	N/A
Outcomes	Mothers where asked to record number, type and impact of prenatal stressful events**Error! Bookmark not defined.** at 1 8 weeks gestation and a total stressful event score calculated and coded into quartiles. Care givers were asked to rate children's symptoms of externalising and internalising problems with the Strengths and Difficulties questionnaire at age 7, 9, 11, 13 and 16. **Baseline measures** Stressful event burden was positively associated with: • Preterm birth (p=0.011) • Low birth weight (p=0.013) Multi normal logistic regression used data adjusted for; gender, maternal education, low birth weight, preterm birth, ethnicity, socioeconomic status, prenatal smoking, prenatal alcohol use, maternal and parental history of depression, maternal stressful life events at 8 months, offspring stressful life events at ages 1 to 6 and offspring symptoms of depression/anxiety at 81 months. • Highest conduct disorder symptom trajectory in children aged 7 to 16 was significantly associated with the highest quartile of prenatal stress (p<0.01) and continuous stress (p<0.05). • Highest hyperactivity symptom trajectory in children aged 7 to 16 was significantly associated with the highest quartile of prenatal stress (p<0.05)

Using linear regression at each age time point the adjusted data showed a relationship with prenatal stress and conduct disorder and hyperactivity(table below)

Adjusted data	Age 7	Age 9	Age 11	Age13	Age 16
Conduct disorder	*β*, p value	β, p value	β, p value	B, p value	β, p value
Second quartile	0.10, <0.05	0.08, 0.075	0.06, 0.176	0.10, <0.05	0.05, 0.333
Third quartile	0.15,<0.01	0.09, 0.069	0.1, <0.05	0.10,<0.05	0.13,<0.05
Fourth quartile	0.18,<0.001	0.13, <0.05	0.13, <0.05	0.18, <0.01	0.18, <0.01
Stress	0.01, <0.05	0.01,<0.05	0.01, <0.001	0.01, <0.001	0.01, <0.001
Hyperactivity					
Second quartile	0.20, <0.01	0.12<0.018	0.00, 0.999	0.04, 0.537	0.043, 0.567
Third quartile	0.22, <0.01	0.12, 0.11	0.03, 0.717	0.11, 0.153	0.15, 0.068
Fourth quartile	0.31, <0.001	0.23, <0.001	0.12, 0.153	0.2, <0.05	0.19, <0.05
Stress	0.01, <0.05	0.01, 0.064	0.004, 0.282	0.01, <0.01	0.01, <0.05

Quality appraisal	The CASP checklist for cohort studies was used to assess the quality of this review. Attrition over the 16 year follow up period was observed and was associated with both prenatal stress and conduct disorder and hyperactivity at 7 years of age. This factor biases the results to be more conservative than they are likely to be had no attrition occurred.

Table 18. Closa-Monasterolo et al 2017

Publication	**Closa-Monasterolo R, Gispert-Llaurado M, Canals J, Luque V, Zaragoza-Jordana M, Koletzko B, et al. The effect of postpartum depression and current mental health problems of the mother on child behaviour at eight years. Maternal and Child Health Journal. 2017 Jul;21(7):1563-72**
Study details	Secondary analysis of data from the EU Childhood Obesity Project which was a randomisedrandomised controlled multi center study assessing the effect of higher- vs lower protein formula on overweight and obesity later in childhood.
Study objectives	To examine the effect of maternal postpartum depression and maternal current mental health problems on child behaviour problems at 8 years.
Inclusions	All women who completed the EPDS at 2, 3, and 6 months after delivery and the children at age 8 who were part of the study and whose parents completed a Childs Behaviour Checklist
Exclusions	None stated
Population	Participants were healthy, singleton, and term infants born between October 2002 and July 2004.
Intervention	N/A
Comparator	N/A
Outcomes	• Children whose mothers had postpartum EPDS at a cut off score of >10 at any time point (2,3, or 6 months postpartum) was associated with a significant risk of mental health problems ($p < 0.017$) compared with those children whose mothers did not have postpartum depression.

	• Children whose mothers had both postpartum depression and current mental problems exhibited the highest levels of psychological problems, followed by those whose mothers who had only current mental problems. Postnatal depression and women with current mental problems had a significant effect on child's total psychological problems (p = 0.033, p < 0.001, respectively). • Maternal postpartum depression at 3 months was linked to children at 8 years having a greater likelihood of problems compared to those whose mothers did not have postpartum depression including Anxiety/Depression,(p=0.007), social problems(p=0.001), attention problems(p=0.010), rule-breaking behaviour (p=0.037) and aggressive problems (p=0.005). • Maternal postpartum depression and current mental health problems, separately and synergistically, increase children's psychological problems at 8 years.
Quality appraisal	The CASP checklist for RCTs was used to assess the quality of this review. There were no areas of concern about the quality of the review.

Table 19. Liu et al 2017

Publication	**Liu Y, Kaaya S, Chai J, McCoy DC, Surkan PJ, Black MM, et al. Maternal depressive symptoms and early childhood cognitive development: a meta-analysis. Psychological medicine. 2017 01 Mar;47(4):680-9.**
Study details	Systematic review and meta-analysis
Study objectives	To describe the relationship between maternal depressive symptoms (antenatal and postnatal) and child cognitive development by systematically reviewing relevant literature and performing a meta-analysis.
Inclusions	Studies that quantitatively assessed relationships between maternal depression (antenatal and postnatal) or depressive symptoms and child cognitive development; (ii) was published in a peer-reviewed journal; (iii) was not a case study; and (iv) assessed the outcome in young children aged between 0 and 7 years
Exclusions	Studies of sub- groups of either women or children eg; only obese children
Population	14 studies of which 2 reported data of women with antenatal depressive symptoms and 12 women with postnatal depressive symptoms. Children were aged 0 to 56 months (approx. 4.5 years)
Intervention	N/A
Comparator	N/A
Outcomes	Infant toddler cognitive development: • Unadjusted meta-analysis of crude estimates from the 14 studies showed statistically significant relationships between maternal depressive symptoms and child cognitive development (Cohen's *dʂ*= −0.25, 95% CI −0.39 to −0.12), indicating a −0.25 S.D. difference in the mean cognitive scores for children whose mothers had high v. low scores on measures of depressive symptoms. • The heterogeneity of the findings was marginally significant (p = 0.05, I^2 = 41.8%), indicating possible variability in effect sizes across studies. Sub group analysis: • When analysis restricted to 3 studies where maternal depressive symptoms were measured during 6–8 weeks postpartum, the association was strengthened (Cohen's *dʂ*=−0.40, 95% CI −0.58 to −0.22, p <0.0005) and (p = 0.97, I^2 = 0.0%). All 3 studies adjusted for some confounders such as child gender and parental education, but differed in others.

	• There was a statistically significant negative association between postpartum depressive symptoms and child cognitive development (β=−4.17, 95% CI, −8.01 to −0.32, p = 0.03) with decreasing heterogeneity (, p = 0.072, I^2 = 62%). • The magnitude of the association was an average score of 4.2 points lower on the Mental Development Index for Infants and Toddlers whose mothers reported high depressive symptoms compared to those reporting low depressive symptoms.
Quality appraisal	The AMSTAR-2 checklist for systematic reviews and meta analyses studies was used to assess the quality of this review. The method of assessing the quality of the 14 included studies was not described. The paper mentions confounding factors that may effect the results of the unadjusted univariate analysis but does not describe what they are. The 4 studies included in an adjusted multivariate analysis highlighted 2 (gender and parental education)of the potential confounding factors but did not name others.

Table 20. Gelaye et al 2016

Publication	**Gelaye B, Kajeepeta S, Williams MA. Suicidal ideation in pregnancy: An epidemiologic review. Archives of Women's Mental Health. 2016 Oct;19(5):741-51**.
Study details	Systematic epidemiologic review
Study objectives	To synthesize available evidence about antepartum suicidal ideation
Inclusions	Observational studies with greater than 10 subjects, reporting prevalence estimates, risk factors, consequences or screening for suicidal ideation during pregnancy. Articles published up to June 3rd 2015 were included. 57 articles were included in the systematic review of which 4 were relevant to this review question.
Exclusions	Full text unavailable, study did not include primary data collection or analysis.
Population	57studies reporting on outcomes of suicide ideation in pregnant women
Intervention	N/A
Comparator	N/A
Outcomes	Prevalence: Suicide ideation in pregnancy was reported by 18 studies and varied from 3 to 33% with the highest figures from urban US studies (23-33%) Adverse maternal outcomes • 1 study reported the risk of death from suicide as increasing by 92%(Hazard Ratio=1.92, 95% CI: 1.53–2.41) within a few days or months of the onset of suicidal ideation • 1 study reported that 33% of those with suicidal ideation in pregnancy will plan suicide, 29% will attempt suicide and 60% are likely to transition from suicide ideation to attempted suicide within the first year of the onset of ideation. Baby outcomes • 1 study showed infants born to women who attempted suicide during pregnancy were more likely to be low birth weight (<2500 g) (OR=1.25, 95% CI, 1.08–1.44) and have respiratory distress syndrome (OR=1.41, 95% CI, 1.07–1.86).

	• 1 study reported that risks for preterm birth (OR=1.34, 95% CI, 1.01–1.77), low birth weight (OR=1.49, 95% CI, 1.04–2.12), and circulatory system congenital anomalies (OR=2.17, 95% CI, 1.02–4.59) were elevated among women who attempted self-poisoning during pregnancy as compared with those who did not. • 1 study showed that infants of mothers reporting depressive symptoms with suicidal ideation weighed 239.5g (95% CI, 3.9–475.1) less on average than infants born to mothers reporting depressive symptoms without suicidal ideation.
Quality appraisal	Of the 57 studies included 28 were retrospective, 19 were cross sectional, and one was a case control study so the majority of the studies were subject to recall bias.

Table 21. Jarde et al 2016

Publication	**Jarde A, Morais M, Kingston D, Giallo R, MacQueen GM, Giglia L, et al. Neonatal outcomes in women with untreated antenatal depression compared with women without depression: A systematic review and meta-analysis. JAMA Psychiatry. 2016 August;73(8):826-37.**
Study details	Systematic review and meta analysis
Study objectives	A systematic review and meta-analysis to understand the effect of untreated antenatal depression on neonatal outcomes.
Inclusions	Studies that assessed depression using either a clinical interview/diagnosis or a screening tool or scale at any time during pregnancy and where anti-depressants and non-pharmacological interventions had not been used or where an adjustment had been made for their effect.
Exclusions	Studies not stratifying outcomes by multiple pregnancies women experiencing domestic violence, or with other reported comorbid psychiatric diseases (eg, anxiety or bipolar disorders).
Population	23 studies of women assessed for antenatal depression were included in the systematic review and 18 studies were used in the meta-analysis
Intervention	N/A
Comparator	N/A
Outcomes	

Outcome	Number studies	Number women	Odds ratio	P value	I^2
Preterm birth <week 37	14	21 048	1.56 (1.25 to 1.94)	<0.001	39%
Low birth weight<2500g	8	3262	1.96(1.24 to 3.10)	<0.004	48%
Small for gestational age <10th percentile	1	4044	1.37(1.10 to 1.70)	<0.005	NA
NICU admission	2	200	1.12 (0.40 to 3.15)	0.83	0

There was a non-significant trend towards those with more severe depression having a higher likelihood of preterm birth.

Sub group analysis showed the odds of preterm birth more than doubled in studies reporting conflicts of interest (OR= 2.50; 95%CI, 1.70-3.67; 5 studies; I^2, 0%), studies not reporting such conflicts showed more moderate results (OR=1.34; 95%CI, 1.08-1.66; 9 studies; I^2, 30%)

	The study authors carried out a sub group analysis based on quality of studies included in their review and results suggested that there are significant differences (*P* = .06) between the results of high- or acceptable-quality studies (OR= 2.39; 95%CI, 1.72- 3.30; 5 studies; I^2, 0%) and low-quality studies (OR=0.89; 95% CI, 0.33-2.35; 3 studies; I^2, 36%) which was attributed to the rigour of selection of participants in each study and their use of pharmacological or non-pharma logical treatments for symptoms of depression.
Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review. There were no areas of concern about the quality of the review.

Table 22. Kingsbury et al 2016

Publication	**Kingsbury M, Weeks M, MacKinnon N, Evans J, Mahedy L, Dykxhoorn J, et al. Stressful Life Events During Pregnancy and Offspring Depression: Evidence From a Prospective Cohort Study. Journal of the American Academy of Child & Adolescent Psychiatry. 2016;55(8):709-16.e2**
Study details	Longitudinal study
Study objectives	Examine long term association between prenatal stressful events and offspring depression in adolescents
Inclusions	Prospective birth cohort data from the Avon Longitudinal Study of Parents and Children (ALSPAC). All individuals with outcome data were included in the analysis
Exclusions	None stated
Population	10,569 mother –offspring pairs
Intervention	N/A
Comparator	N/A
Outcomes	Mothers where asked to record number, type and impact of prenatal stressful events at 18 weeks gestation and a total stressful event score calculated and coded into quartiles. At 17 to18 years offspring completed a computerised version of the Clinical Interview Schedule- Revised (CIS-R). A continuous symptom score and a yes/no variable indicating primary diagnosis of major diagnosis were the primary outcome measures. Authors adjusted the crude data taking into account the following covariates; gender, maternal education, low birth weight, preterm birth, ethnicity, socioeconomic status, prenatal smoking, prenatal alcohol use, maternal history of depression, maternal stressful life events at 8 weeks and 8 months postpartum and maternal depression at 8 years postpartum. Multi normal logistic regression of the adjusted data showed: • An increase of 1 unit on the maternal Stressful Life Event score during gestation was associated with increased offspring depressive symptoms (β=0.07, p<0.01) and major depression at 17 to 18 years (OR=1.03, 95% CI,1.01 to 1.05) • This corresponds to a 12.6% increased risk of internalising disorders in offspring aged 17 to 18 if an event occurred that greatly affected the mother during pregnancy.
Quality appraisal	The CASP checklist for cohort studies was used to assess the quality of this review. There were no areas of concern about the quality of the review. Missing data on outcomes and confounders were estimated and 5 estimated datasets were created and estimates pooled. Results were similar to the complete case sample.

Table 23. Prady et al 2016

Publication	**Prady SL, Pickett KE, Croudace T, Mason D, Petherick ES, McEachan RRC, et al. Maternal psychological distress in primary care and association with child behavioural outcomes at age three. European Child & Adolescent Psychiatry. 2016;25(6):601-13**.
Study details	Longitudinal study
Study objectives	To describe variation in socio-emotional behavioural problems in offspring aged 3 of untreated women with common mental health disorders.
Inclusions	Women recruited consecutively during 26-28 weeks gestation born in Bradford
Exclusions	Women whose research data could not be linked to GP record, women with severe mental illness, second born twins, women where child SDQ not completed at 3 years.
Population	1078 women and offspring
Intervention	N/A
Comparator	N/A
Outcomes	Authors modelled chronicity of common mental disorders from late pregnancy to 3 years postpartum rather than discrete risks in the pre and postnatal periods. Latent class analysis was applied to the data and identified 3 groups: • Women unlikely (low risk) to have common mental disorders (70.1%) • Women treated for common mental disordersby GP (6.2%) • Women likely to have common mnetal disorders from self-report scores but 'untreated' (23.7%) Using the 3 classes, authors used linear regression to test for association with children's SDQ at 3 years. Data was adjusted for mothers ethnicity, child's gender and socioeconomic status, • Women with treated and untreated common mnetal disorders were more likely to have children rated as having socio-emotional difficulties (p<0.05).

	Cohen's d	95% CI
Low risk vs treated	**-0.27***	**-0.52 to -0.02**
Low risk vs untreated	**-0.32**	**-0.46 to -0.17**
Treated vs untreated	0.04	-0.23 to -0.31

*Figures in bold are significant (p<0.05)

Quality appraisal	A CASP checklist for longitudinal studies was undertaken. Limitations include not identifying specific mental health disorders and treatment. Self-report of mood which suggested an 'untreated' CMD may be sub clinical or transient distress that might not meet current criteria for treatment.

Table 24. Sanger et al 2015

Publication	**Sanger C, Iles JE, Andrew CS, Ramchandani PG. Associations between postnatal maternal depression and psychological outcomes in adolescent offspring: a systematic review. Archives of Women's Mental Health. 2015;18(2):147-62**.
Study details	Systematic review
Study objectives	To describe the impact of PND on offspring adolescent outcomes through a systematic review of the literature
Inclusions	Studies examining associations between maternal PND (to 1 year postpartum) and internalising problems, externalising problems, psychopathology, psychosocial, and cognitive outcomes of adolescent offspring were included (age 11 to 18).
Exclusions	None stated
Population	16 studies of women assessed for PND during the first year postpartum and their adolescent offspring assessed for internalising and externalising problems aged 11 to 18 were included.
Intervention	N/A
Comparator	N/A
Outcomes	**Cognitive development** (eg; outcome measures relating to academic, intellectual, IQ, language, other cognitive development measures) • 4 publications of 3 prospective longitudinal cohort studies of which 2 publications reported cognitive development measured by school examination results and 2 papers from the same study used a range of cognitive development tests such as SDQ (Strengths and Difficulties Questionnaire), WISC-III (Wechsler Intelligence Scale for Children-UK),and WASI (Wechsler Abbreviated Scale of Intelligence) • Impaired offspring cognitive outcomes reflected some of the most consistent findings. • In 2 studies at age 16 PND had a main effect (when accounting for confounding variables) and a more marked effect was seen in boys (F=5.13, p<.001 and (F=4.18, p<.05). **Socioemotional development (no figures given)** (eg; measures relating to social and emotional developmental outcomes) • Psychosocial outcomes in offspring adolescents indicated a specific adverse effect, although based on only 2 longitudinal prospective studies. • 1 study reported PND was associated with lower scores on the CBCL and YSR Social Competence subscales. Effects were more marked in male offspring • 1 study reported PND was associated with heightened emotional sensitivity (and depressive symptoms) in female offspring at 13 years **Internalising problems (no figures given)** (eg depression and/or anxiety disorders) • 10 publications reporting results from 8 studies examined adolescent depression and anxiety disorders of which 7 were longitudinal prospective cohort studies and one was a cross sectional study. • Conflicting evidence was found for an effect of PND on adolescent offspring internalising problems • 5 studies showed no relationship between PND and adolescent depression • 2 studies found that higher maternal CESDI scores postpartum resulted in higher at risk trajectories for internalising problems in adolescent offspring • 1 study showed a small but significant correlation between offspring internalising problems at 11 years and PND using the Child behaviour checklist but not with the Children's depression inventory.

	Externalising problems (no figures given) (eg; aggression, violence, conduct disorder (CD),oppositional defiant disorder (ODD), and attention deficit hyperactivity disorder (ADHD). • 7 publications reported results from 4 studies • Conflicting evidence was found for an effect of PND on adolescent offspring externalising problems • 5 studies reported no relationship between PND and externalising problems • 1 study reported maternal PND predicts increased severity and diversity of offspring violent behaviours associated with ADHD and anger management difficulties • 1 study reported PND was associated with externalising problems on the Youth Self Report questionnaire but not the Child Behaviour Checklist at 16-years **Overall psychopathology(no figures given)** (eg; outcomes that did not fit into either the internalising or externalising constructs, eg eating disorders, emergent bipolar, or psychosis). • Conflicting evidence was found for an effect of PND on overall psychopathology problems with 2 studies included, one was a longitudinal prospective cohort and the other a cross sectional study with follow up. • 1 study reported when PND was diagnosed using the CIS, index offspring were 4 times as likely to be diagnosed with a psychiatric disorder • No associations were found between PND and offspring DSM-IV diagnoses
Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review. This review included studies with a wide range of different types of scoring systems which were not possible to compare easily.

Table 25. Leis et al 2014

Publication	**Leis JA, Heron J, Stuart EA, Mendelson T. Associations between maternal mental health and child emotional and behavioral problems: does prenatal mental health matter? Journal of Abnormal Child Psychology. [Research Support, Non-U.S. Gov't]. 2014;42(1):161-71.**
Study details	Longitudinal study
Study objectives	To describe the associations between elevated symptoms of antenatal depression or anxiety and offspring emotional and behavioural problems during mid to late childhood taking into account the impact of later maternal mental health symptoms.
Inclusions	All pregnant women living in the Avon area of England with an expected date of delivery between April 1, 1991 and December 31, 1992, who gave birth to singleton babies still alive at 1 year of age who had complete data on antenatal mental health, and whose children had complete data from mothers and teachers on emotional and behavioural functioning in mid-late childhood.
Exclusions	None stated
Population	2,891 pregnant women and their offspring
Intervention	N/A
Comparator	N/A
Outcomes	Symptoms of depression were assessed multiple times during pregnancy (18 and 32 weeks gestation), the first year postpartum (8 weeks and 8 months postpartum), and when offspring were childhood 21, 33, 61, and 73 months and

	11 years with the Edinburgh Postnatal Depression Scale Children were assessed by mothers and teachers with the Strengths and Difficulties Questionnaire at age 10/11. Multivariable regression models were fit to address study hypotheses that antenatal depression or anxiety impacted on offspring emotional and behavioural problems in mid to late childhood. Multivariable associations were adjusted for psychological variables(eg: other periods of elevated symptoms of anxiety, ever having elevated symptoms of depression), maternal age at birth of child, child birth weight, child gender, maternal tobacco and alcohol use during pregnancy, maternal educational attainment, and maternal marital status at child age 10. **Association of antenatal depression or anxiety and offspring emotional/behavioural problems from linear regression models showing maternal and teacher reports (adjusted data).**

	Mother report (antenatal depression)	Teacher report	Mother report(antenatal anxiety)	Teacher report
Prosocial behaviour	−0.00 (0.10)	0.02 (0.14)	0.07 (0.11)	0.09 (0.14)
Hyperactivity	0.23 (0.10)*	0.41 (0.14)*	0.03 (0.13)	−0.15 (0.15)
Emotional symptoms	0.22 (0.07)**	0.06 (0.10)	0.43 (0.11)***	0.16 (0.11)
Conduct problems	0.17 (0.08)*	0.10 (0.07)	0.13 (0.08)	0.01 (0.08)
Peer problems	0.17 (0.09)	−0.01 (0.10)	0.02 (0.10)	−0.04 (0.11)
Total difficulties	0.91 (0.28)***	0.56 (0.30)	0.62 (0.30)*	−0.02 (0.32)
R2 (total difficulties)	0.10	0.08	0.10	0.08

*p<0.05, **p<0.01, ***p<0.001

Quality appraisal	The CASP checklist for cohort studies was used to assess the quality of this review. There were no areas of concern about the quality of the review.

Table 26. Tarabulsy et al 2014

Publication	**Tarabulsy GM, Pearson J, Vaillancourt-Morel M-P, Bussieres E-L, Madigan S, Lemelin J-P, et al. Meta-analytic findings of the relation between maternal prenatal stress and anxiety and child cognitive outcome. Journal of Developmental and Behavioral Pediatrics. 2014 Jan;35(1):38-43**.
Study details	Systematic review and meta-analysis
Study objectives	To examine the relationship between antenatal stress and anxiety and early child cognitive outcome by meta-analysis taking into account methodological moderators of the studies.
Inclusions	Articles between 1970 and 2011 where antenatal stress and child cognitive development from 0 to 60 months reported.
Exclusions	Only one article per study was included – the one with the most exhaustive reporting of the sample.
Population	11 studies
Intervention	N/A
Comparator	N/A

Outcomes	• Low but significant relationship between maternal antenatal stress or anxiety and early child cognitive development (r=-0.05, p<0.001[95% CI -.07 to-.02]) • Significant heterogeneity of the results across studies was reported (p<0.001). • No effect sizes differed by trimester when maternal assessments were made • Life events and cognitive development has a significantly greater effect (r=-0.31; p<0.0001; 95% CI, -0.40 to -0.22) than pregnancy related stress and anxiety(r=-0.08;p<0.001;95% CI-0.14 to-0.02) or stress and anxiety unrelated to pregnancy (r=-0.02;p<0.001; 95%CI -0.05 to -0/01) • 8 studies used prospective assessments and 3 used retrospective assessments of antenatal depression, stress, anxiety or life events. Retrospective postnatal assessments by mothers of antenatal periods reported significantly greater effect sizes (r=-0.11;p<0.001;95% CI -0.16 to -0.06) than prospective antenatal assessments (r=-0.03;p<0.001;95%CI -0.06 to -0.01). • Authors reported no effect of age of offspring when cognitive assessment was carried out although they do not report timing of offspring assessment from the included papers.
Quality appraisal	The AMSTAR-2 checklist for systematic reviews and meta analyses studies was used to assess the quality of this review. The authors did not report the age of the offspring when cognitive development was assessed although from the titles of the included papers this appears to be from infancy to early child preschool age.

Studies relevant to question 2 criterion 4 and 5; what is the reported accuracy of screening tools to detect mental health problems during pregnancy?

Table 27. Howard et al 2018

Publication	**Howard L, Ryan E, Trevillion K, Anderson F, Bick D, Bye A, Byford S, O'connor S, Sands P, Demiliw J, Milgrom J, Pickles Accuracy of the Whooley questions and the Edinburgh Postnatal Depression Scale in identifying depression and other mental disorders in early pregnancy. British Journal of Psychiatry 2018 212,50-56**
Study details	Cross sectional survey of women responding to PHQ 2 (also known as the Whooley questions)
Study objectives	Investigate prevalence of mental disorders in early pregnancy and accuracy of PHQ2 compared with EPDS and Structured Clinical Interview for DSM-IV-TR(SCID)
Inclusions	• All women saying 'yes' to either of the 2 PHQ 2 questions at booking appointment 10th November 2014 to 30th June 2016 • Random selection (either 1:6 or 1:4) of women responding 'no' to both PHQ 2 questions at booking appointment between 10th November 2014 and 30th June 2016.
Exclusions	Women ages <16 years
Population	545 pregnant women from South East London
Intervention	PH2 questions
Comparator	EPDS questionnaire – self report using iPad or pen and paper and clinical interview using SCID-I research version asking questions from modules about mood disorder, anxiety disorder and eating disorder and SCIDII questions about borderline personality disorder.

Outcomes

287 women responding 'yes' to PHQ 2, 258 women responding 'no' to PHQ 2.

Using weighted estimations the table below gives population prevalence of a SCID disorder.

Condition	Prevalence (%)	95% Confidence interval
Major depression	11	8-14
• Mild depression	6	4-9
• Moderate depression	4	3-8
• Severe depression	0.1	0-0.3
• Mixed anxiety and depression	0.4	0.2-0.6
Any anxiety disorder	15	11-19
Obsessive compulsive disorder	2	1-4
Eating disorder	2	0.4 -3
PTSD	0.8	0–1
Bipolar disorder	0.03	0-0.2
Borderline personality disorder	0.7	0-1
Any SCID	27	22-32

	Sensitivity	Specificity	PPV	NPV	+LR	-LR	AUC(95% CI)
Cut off score 12/13 for SCID depression	0.59	0.94	0.52	0.95	9.8	0.44	0.89(0.99-0.90)
Cut off score 12/13 for any SCID condition	0.30	0.95	0.67	0.79	6	0.74	0.74(0.73 0.75)

PHQ 2 questions performance compared with SCID for depression

Weighted SCID depression	Sensitivity	Specificity	PPV	NPV	+LR	-LR	AUC(95% CI)
PHQ2 either question answered 'yes'	0.41	0.95	0.45	0.93	8.2	0.62	0.37(9.34-0.40)
PHQ2 either question answered 'yes' and Arroll question 'yes'	0.08	0.99	0.66	0.83	8	0.93	0.21(0.19 to 0.23)
PHQ2 answering yes to both PHQ2 questions	0.25	0.98	0.61	0.84	12.5	0.77	0.24(0.21 to 0.26)
Weighted for any SCID condition							
PHQ2 either question answered 'yes'	0.23	0.96	0.66	0.78	5.8	0.80	0.21(0.2-0.23)
PHQ2 either question answered 'yes' and	0.05	1.00	0.86	0.65			0.11(0.10-0.12)

Arroll question 'yes'							
PHQ2 answering yes to both PHQ2 questions	0.06	0.99	0.78	0.66	6	0.95	0.12(0.11-0.13)

PPV positive predictive value, NPV negative predictive value, LR – likelihood ratio, AUC Area under the curve, SCID – Structured Clinical Interview DSM-IV, EPDS Edinburgh Postnatal Depression Scale, PHQ Patient health questionnaire.

Quality appraisal

QUADAS 2: Quality appraisal Howard et al 2018

Question	Assessment (Y, N, unclear)	Risk of Bias (low, high, unclear)	Supporting info
Domain I: Patient selection			
Consecutive or random sample of population enrolled?	yes	low	Sample was consecutive for those who tested positive and a stratified random sample was recruited for those testing negative
Case-control design avoided?	yes	low	Not a case control design
Inappropriate exclusions avoided?	yes	low	Under age 16, no longer pregnant when approached for interview, booking appointment was elsewhere
Domain II: Index Test			
Index test results interpreted without knowledge of reference standard results?	yes	low	Index test undertaken before reference standard undertaken.
Threshold pre-specified?	yes	low	For EPDS it was 12/13, for PHQ2 it was answering 'yes' to at least 1 of the 2 questions.
Domain II: Reference standard			
Reference standard likely to correctly classify condition?	yes	low	Diagnosis using clinical interview and DSMIV then consensus with other clinicians
Reference standard results interpreted without knowledge of index test results?	Unclear	Unclear	Researchers undertook gold standard clinical interview and met weekly with lead researcher to achieve consensus on diagnosis where clinical notes and responses to PHQ2 questions were available.
Domain IV: Test strategy flow and timing			

	Appropriate interval between index test and reference standard?	yes	low	Maximum of 3 weeks between booking appointment and reference test.	
	Did all participants receive same reference standard?	yes	low	All completed EPDS and SCID.	
	All patients included in analysis?	yes	low		
	Applicability				
	Applicable to UK screening population of interest?	yes	low	Carried out on population in UK	
	Applicable to UK screening test of interest?	yes	low	Tests are currently already used in UK for case finding.	
	Target condition measured by reference test applicable to UK screening condition of interest?	yes	low	Reference test (SCID) is used in UK	

Table 28. Vlenterie et al 2017

Publication	**Vlenterie R, van Ras HWP, Roeleveld N, Pop-Purceleanu M, van Gelder MMHJ. Epidemiological evaluation of the Patient Health Questionnaire-2 in a pregnant population. Journal of Psychosomatic Research. 2017;101:96-103**
Study details	Prospective cohort study in the Netherlands
Study objectives	Evaluate PHQ-2 compared with HADS-D and EPDS for detecting depression in pregnant women
Inclusions	Women aged ≥18 years with estimated date of delivery between February 2012 and May 2016.
Exclusions	Women who experienced miscarriage, stillbirth, termination of pregnancy or gave birth prior to week 34 of pregnancy. Women were baseline PHQ-2 scores were missing.
Population	3033 pregnant women
Intervention	PHQ-2 administered
Comparator	EPDS and HADS-D administered
Outcomes	Questionnaires were administered at 3 time points: Q1- 8-12 weeks gestation Q2- 17 weeks gestation Q3- 34 weeks gestation Cut off scores: PHQ-2 cut off ≥1 answer with 'yes' response, EPDS score cut off ≥13, HADS-D score cut off ≥8

At baseline(Q1) for PHQ-2, 24.2%(n=735)scored positive with a cut off of ≥1

Reference standard and timing	N	True positives	True negatives	False positives	False negatives
Q1 PHQ-2-HADS-D	2931	132	2189	579	31
Q2 PHQ-2-EPDS	2486	103	1976	388	19
Q3PHQ-2-HADS-D	2510	173	1778	480	79

Reference standard and timing	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV
Q1 PHQ-2-HADS-D	0.81(0.74-0.88)	0.79(0.77-0.81)	0.19(0.16-0.21)	0.99(0.98-0.99)
Q2 PHQ-2-EPDS	0.84(0.77-0.91)	0.84(0.82-0.85)	0.21(0.17-0.25)	0.99(0.98-0.99)
Q3PHQ-2-HADS-D	0.69(0.62-0.76)	0.79(0.77-0.81)	0.26(0.23-0.30)	0.96(0.94-0.97)

PHQ-2 Patient Health Questionnaire with 2 questions, HADS-D – Hospital Anxiety and Depression Scale –using subscale for depression, EPDS – Edinburgh Postnatal Depression Scale

Quality appraisal

Quality appraisal Vlenturie et al 2017

Question	Assessment (Y, N, unclear)	Risk of Bias (low, high, unclear)	Supporting info
Domain I: Patient selection			
Consecutive or random sample of population enrolled?	Yes	Low	Part of a prospective cohort of pregnant women attending first appointments
Case-control design avoided?	Yes	Low	Not a case control design
Inappropriate exclusions avoided?	Yes	Low	Exclusions were appropriate and included women who experienced miscarriage, stillbirth, and termination of pregnancy or gave birth prior to week 34 of pregnancy. Also excluded were women where baseline PHQ-2 scores were missing.
Domain II: Index Test			
Index test results interpreted without knowledge of reference standard results?	Yes	Low	Questionnaires were all self reports, carried out by all women using a web based format.

	Threshold pre-specified?	Yes	Low	PHQ-2 cut off ≥1 answer with 'yes' response, EPDS score cut off ≥13, HADS-D score cut off ≥8	
	Domain II: Reference standard				
	Reference standard likely to correctly classify condition?	No	High	Both HADS-D and EPDS are not gold standard reference stands although they are commonly used validated tools to detect depression in pregnant women.	
	Reference standard results interpreted without knowledge of index test results?	Yes	Low		
	Domain IV: Test strategy flow and timing				
	Appropriate interval between index test and reference standard?	Yes	Low	Index test and reference standard collected at the same time	
	Did all participants receive same reference standard?	Yes	Low	All patients carried out the same tests	
	All patients included in analysis?	Yes	Low		
	Applicability				
	Applicable to UK screening population of interest?	Yes	Low	Northern European population	
	Applicable to UK screening test of interest?	Yes	Low	All tests used are also used in the UK	
	Target condition measured by reference test applicable to UK screening condition of interest?	Yes	Low	Depression is the target condition of reference standard	

Table 29. O'Connor et al 2016

Publication	**O'Connor E, Rossom RC, Henninger M, Groom HC, Burda BU. Primary Care Screening for and Treatment of Depression in Pregnant and Postpartum Women: Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA. [Meta-Analysis=Research Support, U.S. Gov't, P.H.S.Review]. 2016;315(4):388-406.**
Study details	Systematic review

Study objectives	To systematically review the benefits and harms of depression screening treatment and accuracy of selected screening instruments for pregnant and postpartum women.
Inclusions	Studies of the PHQ and EPDS compared to a valid reference standard defined as a structured or semi structured diagnostic interview with a trained interviewer or non-brief interview with mental health clinician.
Exclusions	None stated
Population	26 studies were included in total.
Intervention	N/A
Comparator	N/A
Outcomes	23 studies (n=5398) reporting results from EPDS and 3 studies (n=777) reporting PHQ accuracy.

EPDS	Cut off	True Positives	False Positives	False Negatives	True Negatives	Sensitivity (%)*	Specificity (%)*	PPV (%)*	NPV %
MDD	≥ 14	6	6	0	88	100	94	50	NR
	≥ 13	6	12	0	82	100	87	33	NR
	≥ 12	6	20	0	74	100	79	23	NR
Major or minor depression	≥ 14	8	4	6	82	57	95	66	NR
	≥ 13	9	9	5	77	64	90	50	NR
	≥ 12	9	17	5	69	64	80	35	NR
	≥ 11	10	24	4	62	71	72	29	NR
PHQ-8									
MDD	≥ 11	10	64	3	136	77	68	NR	NR
	≥ 10	10	76	3	124	77	62	NR	NR
PHQ-2, Likert scale)									
MDD	≥ 4	8	42	5	158	62	79	NR	NR
	≥ 3	10	82	3	118	77	59	NR	NR
PHQ2 yes/no									
Major or minor depression	≥ 1	17	35	0	74	100	68	NR	NR

Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns.

Table 30. Owora et al 2016

Publication	**Owora AH, Helene C, Reese J, Garwe T. Diagnostic performance of major depression disorder case-finding instruments used among mothers of young children in the United States: A systematic review. Journal of Affective Disorders. [Review]. 2016;201:185-93.**
Study details	Systematic review
Study objectives	Identify the most valid case finding instrument for major depression in pregnant women and postpartum used in the US.
Inclusions	Studies using gold standard reference tests including SCID, CIDI, SADS and DIS
Exclusions	None stated
Population	14 articles with data on 21 case finding instruments
Intervention	N/A
Comparator	N/A
Outcomes	• The study reported the 10 item EPDS had the least variation and highest performance during the second and third trimesters and thresholds (sensitivity range 0.63-0.94 and specificity range 0.83 to 0.98). • Where there were different versions of the same scale the version with fewer questions typically had the highest sensitivity and lower specificity. • There was a pattern of increased accuracy of the instruments with lower prevalence rates of major depression. • Accuracy of instruments tended to be higher antepartum compared to postpartum Authors report sensitivity, specificity, reference standard and cut off point for a range of scales and studies. Data extracted here are for studies from the US with an adult population tested in a community or primary care setting.

	Reference standard	Cut off point	Trimester period	Sensitivity	Specificity
EPDS 10	CIDI	10	1	0.80(0.61 to 0.91)	0.83(0.80 to 0.85)
	CIDI	10	2	0.94(0.73 to 0.99)	0.87 (0.84 to 0.89)
PHQ 9	SCID	10	4	0.86(0.49 to 0.97)	0.84(0.81 to 0.87)

Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns.

Table 31. Meades et al 2011

Publication	**Meades R, Ayers S. Anxiety measures validated in perinatal populations: a systematic review. Journal of Affective Disorders. 2011;133(1-2):1-15**
Study details	Systematic review
Study objectives	Identify and review measures of anxiety validated in perinatal populations.
Inclusions	Studies were written in English, examined reliability or validity of a self-report anxiety measure the sample was of antenatal or postnatal women(to one year)

Exclusions	None stated
Population	30 studies with 21 scales
Intervention	N/A
Comparator	N/A
Outcomes	One out of 30 studies reported sensitivity, specificity, PPV and NPV results of the use of 1 scale the GHQ30 during the antenatal period in a population generalizable to the UK. At a cut off of 5/6 the sensitivity was 0.77, specificity 0.78, PPV 0.52 and NPV was 0.96. Other studies were testing scales in postnatal populations, did not report sensitivity, specificity PPV and NPV or were using non-English language versions of the scales.
Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns

Studies relevant to question 3 criteria 4 and 5: What is the reported accuracy of screening tool to detect postnatal depression?

Table 32. Gollan et al 2017

Publication	**Gollan JK, Wisniewski SR, Luther JF, Eng HF, Dills JL, Sit D, et al. Generating an efficient version of the Edinburgh Postnatal Depression Scale in an urban obstetrical population. Journal of Affective Disorders. 2017 Jan;208:615-20.**
Study details	Secondary analysis of a dataset from the identification and therapy of postpartum depression study a prospective cohort study.
Study objectives	Assess screening quality of 7 and 10-item EPDS to screen for postpartum depression
Inclusions	Women ≥ 18 years of age, who had delivered a live infant, consented to be contacted could be contacted by phone at 4 to 6 weeks postpartum.
Exclusions	None stated
Population	15,172 newly delivered women between 2006 and 2011
Intervention	10 item EPDS administered 4 to 6 weeks postpartum
Comparator	SIGH-ADS clinical interview (29 item structured interview comprising the HRSD scale plus a further 8 items)
Outcomes	EPDS scores were analysed by weighting the questions by severity of symptom they measured. Responses to all questions were used for the 10 item EPDS scale and responses to items 1,2,6,7,8,9,10 were used for the 7 item EPDS as comparison with the SIGH-ADS gold standard.

	Cut off	Sensitivity	Specificity
EPDS 10	≥14	0.83	0.66
EPDS 7	≥8	0.76	0.68

Quality appraisal	Two phase design where participants phoned and EPDS administered clinical interview only offered to those who scored above the EPDS cut off of >10.

Table 33. O'Connor et al 2016

Publication	**O'Connor E, Rossom RC, Henninger M, Groom HC, Burda BU. Primary Care Screening for and Treatment of Depression in Pregnant and Postpartum Women: Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA. [Meta-Analysis=Research Support, U.S. Gov't, P.H.S.Review]. 2016;315(4):388-406.**
Study details	Systematic review
Study objectives	To systematically review the benefits and harms of depression screening treatment and accuracy of selected screening instruments for pregnant and postpartum women.
Inclusions	Studies of the PHQ and EPDS compared to a valid reference standard defined as a structured or semi structured diagnostic interview with a trained interviewer or non-brief interview with mental health clinician.
Exclusions	None stated
Population	26 studies were included in total.
Intervention	N/A
Comparator	N/A
Outcomes	23 studies (n=5398) reporting results from EPDS and 3 studies (n=777) reporting PHQ accuracy. The data below is extracted from the systematic review for populations relevant to the UK for the accuracy of the EPDS and PHQ scales during the postnatal period

EPDS	Cut off	True Positives	False Positives	False Negatives	True Negatives	Sensitivity (%)*	Specificity (%)*	PPV (%)*	NPV %
MDD	≥ 13	14	1	4	131	78	99	93	96
Major or minor depression	≥ 9	27	15	19	89	59	86	64	82
MDD	≥ 13	6	19	2	101	75	84	24	NR
	≥ 12	7	29	1	91	88	76	20	NR
	≥ 11	7	32	1	88	88	73	18	NR
	≥ 10	7	25	1	85	88	71	17	NR
	≥ 9	8	48	0	72	NR	NR	NR	NR
Major or minor depression	≥ 13	13	12	8	95	62	89	52	NR
	≥ 12	16	20	5	87	76	81	44	NR
	≥ 11	16	23	5	84	76	79	41	NR
	≥ 10	17	25	4	82	81	77	41	NR
MDD	≥ 13	21	7	1	97	95.0	93.0	NR	NR
	≥ 10	22	19	0	85	100	82	NR	NR
	≥ 13	106	178	28	548	79.1	75.5	NR	NR

	Mild, moderate or severe depression	≥ 12	116	239	18	487	86.6	67.1	32.7	NR
	Moderate or severe depression	≥ 13	46	238	8	568	85.2	70.5	37.3	NR
		≥ 12	50	305	4	501	92.6	62.2	NR	NR
	PHQ 9									
	MDD	≥ 10	15	38	5	380	75	91	28	99
	PHQ2 Yes/no									
	MDD	≥ 1	20	159	0	259	100	62	11	100
	PHQ2 Likert									
	MDD	≥ 2	15	48	5	368	75	88	24	99

Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review and ….

Table 34. Owora et al 2016

Publication	**Owora AH, Helene C, Reese J, Garwe T. Diagnostic performance of major depression disorder case-finding instruments used among mothers of young children in the United States: A systematic review. Journal of Affective Disorders. [Review]. 2016;201:185-93.**
Study details	Systematic review
Study objectives	Identify the most valid case finding instrument for major depression in pregnant women and postpartum used in the US.
Inclusions	Studies using gold standard reference tests including SCID, CIDI, SADS and DIS
Exclusions	None stated
Population	14 articles with data on 21 case finding instruments
Intervention	N/A
Comparator	N/A
Outcomes	• Authors report sensitivity, specificity, reference standard and cut off point for a range of scales and studies. • Data extracted below are for studies from the US with an adult pregnant or postpartum population tested in a community or primary care setting. • The PDSS showed the least variation and highest performance across different postpartum periods and cut off thresholds (sensitivity range 0.67 to 0.95 and specificity range 0.68 to 0.97). • Where there were different versions of the same scale the version with fewer questions typically had the highest sensitivity and lower specificity. • There is a pattern of lower and more variable performance during the first 2 months than later.

- There was a pattern of increased accuracy of the instruments with lower prevalence rates of major depression.
- Accuracy of instruments tended to be higher antepartum compared to postpartum

Postpartum major depression disorder instrument accuracy data for populations relevant to the UK from the US

	Reference standard	Cut off	Month postpartum	Sensitivity	Specificity
BDI-II	DIS	20	0-2	0.55(0.34 to0.75)	1.00(0.97 to 1.00)
	SCID	14	0-14	0.74(0.63 to 0.82)	0.80(0.72 to 0.86)
	SCID	20	0-14	0.45(0.34 to0.57)	0.91(0.85 to 0.95)
	SCID	14	0-4	0.80(0.60 to 0.92)	0.84(0.71 to 0.92)
	SCID	14	4-8	0.72(0.53 to0.86)	0.80(0.65 to 0.89)
	SCID	14	9-14	0.60(0.41 to 0.76)	0.88(0.75 to 0.95)
	SCID	15	0-13	0.69(0.59 to0.78)	0.81(0.76 to 0.85)
	SCID	13	0-13	0.78(0.69 to 0.86)	0.73(0.67 to 0,78)
	SCID	12	0-13	0.78(0.69 to 0.86)	0.69(0.63 to 0.74)
	SCID	11	0-13	0.80(0.71 to 0.87)	0.63(0.57 to 0.69)
EPDS10	DIS	12	0-2	0.76(0.54 to0.90)	0.99(0.95 to 1.00)
	SCID	9	0-14	0.78(0.67 to 0.86)	0.76(0.68 to0.82)
	SCID	13	0-14	0.55(0.43 to 0.66)	0.91(0.85 to 0.95)
	SCID	9	0-4	0.85(0.65 to 0.94)	0.76(0.62 to 0.86)
	SCID	9	4-8	0.72(0.53 to0.86)	0.85(0.71 to 0.92)
	SCID	9	9-14	0.71(0.52 to 0.85)	0.73(0.58 to 0.84)
	DIS	10	1-2	0.61(0.36 to 0.81)	0.85(0.62 to 0.95)
	SCID	13	0-13	0.67(0.57 to 0.76)	0.81(0.76 to 0.85)
	SCID	12	0-13	0.74(0.64 to 0.82	0.73(0.67 to 0.78)
	SCID	10	0-13	0.80(0.71 to 0.87)	0.66(0.6.0 to 0.71)
	CIDI	10	2	0.75(0.44 to 0.92)	0.89(0.87 to 0.91)
IDAS-GD	SCID	51	0-13	0.63(0.53 to 0.73)	0.81(0.72 to 0.88)
	SCID	44	0-13	0.81(0.72 to 0.88)	0.65(0.59 to 0.70)
	SCID	39	0-13	0.93(0.85 to 0.97)	0.55(0.49 to 0.61)
PDSS	DIS	80	0-2	0.92(0.72 to 0.98)	0.97(0.93 to 0.99)
	SCID	80	0-14	0.71(0.60 to 0.80)	0.78(0.70 to 0.84)
	SCID	80	0-4	0.67(0.47 to 0.83)	0.78(0.64 to 0.87)
	SCID	80	4 - 8	0.80(0.61 to 0.91)	0.77(0.63 to 0.87)
	SCID	80	9-14	0.71(0.52 to 0.86)	0.68(0.52 to 0.80)
	SCID	80	0-14	0.71(0.60 to 0.80)	0.78(0.70 to 0.84)
PHQ 2	SCID	3	0-1	0.75(0.54 to0.88)	0.88(0.84 to 0.91)
	SCID	1	0-1	0.98(0.82 to 1.00)	0.62(0.57 to 0.67)
	SCID	3	0-9	0.84(0.71 to 0.92)	0.79(0.75 to 0.82)

		SCID	1	0-9	0.99(0.91 to 1.00)	0.44(0.40 to 0.49)
	PHQ 9	SCID	12	0-12	0.99(0.97 to 1.00)	1.00(0.99 to1.00)
		SCID	10	0-12	0.80(0.76 to 0.83)	0.64(0.60 to 0.68)
		SCID	10	0-1	0.75(0.54 to 0.88)	0.91(0.88 to 0.93)
		SCID	10	0-9	0.67(0.53 to 0.790	0.92(0.89 to 0.94)
		SCID	10	1-2	0.32(0.14 to 0.58)	0.09(0.02 to 0.31)
	PRAMS6	SCID	17	0-12	1.00(0.99 to1.00)	0.84(0.81 to 0.87)
		SCID	15	0-12	0.80(0.76 to 0.83)	0.47(0.43 to 0.51)
Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns.					

Question 4: What are the benefits of pharmacological and non-pharmacological intervention (alone or in combination) in women with screen-detected mental health problems during pregnancy?

Table 35. Letourneau et al 2017

Publication	**Letourneau NL, Dennis CL, Cosic N, Linder J. The effect of perinatal depression treatment for mothers on parenting and child development: A systematic review. Depression and Anxiety. 2017 October;34(10):928-66.**
Study details	Systematic review
Study objectives	To examine which treatment interventions for antenatal or postpartum depression are most effective in improving both parenting and child development.
Inclusions	Women defined as depressed in the antenatal or postpartum period using a valid assessment tool or been diagnosed by a physician.
Exclusions	None stated
Population	7 trials where women identified and treatment started (and sometimes completed) in the antenatal period.
Intervention	Any intervention
Comparator	
Outcomes	IPT, CBT and massage produced large effects on parenting and subsequent child development. • Higher levels of maternal foetal attachment with IPT (p=0.03) and parenting education programme (p=0.007) • Increase in attention bias indices towards infant distress (p=0.034) with CBT • Between group differences in 23/25 of the revised infant behaviour questionnaire (IBQ-R) subscales favouring the intervention group following CBT. • Social and clinical support – telephone based counsellors, play groups, mothers groups and self-help work book – significant improvement in Depression and Anxiety Stress Scales in treatment group (anxiety, p<0.01, stress, p<0.01). Depression measured with the Beck Depression Inventory (II) lowered (p<0.01), Parental stress measured with the Parental Stress Index was lower in the intervention group than the control (p<0.05). • Massage of mothers resulted in lower likelihood of infants being born prematurely, having low birth weight, and high cortisol levels (figures not reported).

Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns.

Table 36. Prady et al 2017

Publication	**Prady SL, Hanlon I, Fraser LK, Mikocka-Walus A. A systematic review of maternal antidepressant use in pregnancy and short- and long-term offspring's outcomes. Archives of Women's Mental Health. 2017 12 Oct:1-14**
Study details	Systematic review
Study objectives	Evaluate the literature comparing outcomes for children of women who took antidepressants compared to those whose mothers had common mental disorders, or symptoms, during pregnancy.
Inclusions	Children whose mothers who took antidepressants while pregnant
Exclusions	None stated
Population	11 studies; 4 cohort studies reporting a low birth weight outcome and 7 cohort studies reporting a neurodevelopmental outcome
Intervention	Antidepressants
Comparator	Children whose mothers were depressed or anxious and non-exposed to antidepressants (not treated or undergoing psychological, or alternative treatments such as light therapy, massage therapy, exercise or omega-3 fatty acid supplementation).
Outcomes	Maternal antidepressant use in pregnancy and offspring's Low Birth Weight • LBW of infant/neonate is birth weight < 2.500 kg or small for gestational age (SGA), defined as weight for gestation < 10th (or 5th) percentile or birth weight is lower than 2 standard deviations below the mean value for the gestational age. • All 4 cohort studies investigating LBW examined selective serotonin reuptake inhibitors (SSRIs) for antenatal depression • 2 studies did not identify any effect of anti-depressants on LBW • 2 did report statistically significant effect ratios (n=3966, effect ratio1.42, 95% CI, 1.16, 1.7 and n=1622, effect ratio 1.69, 95% CI, 1.14, 2.52); however, neither of these latter studies controlled for severity of depression. Maternal antidepressant use in pregnancy and offspring neurodevelopment and neurobehaviour • Neonate behaviour:1 study with low risk of bias showed no effect by exposure group • Infant and toddler development: 3 studies were included with higher risk of bias. One study showed a significant difference in average time (additional 13.6 days) to walking unsupported (retrospectively reported). A second study reported a significant increase of 28.9 days to walking unsupported for children of women exposed to anti-depressants in the 2nd and 3rd trimesters of pregnancy. No other between group differences in either study was significant. The third study did not find any significant differences between groups. • Child behavioural outcomes - none of the 3 included studies reporting 15 outcomes reported a significant difference between exposure and non-exposure groups. • Child autistic symptoms – the 1 included study did not find any differences between groups for autistic symptoms in children aged 6 with mothers who were using anti-depressants in pregnancy.

	• ADHD and comorbid disorders – 1 included study with a higher bias found a statistically significant effect size (11.4 95% CI, 1.42 to 91.8) in children aged 3 to 7.
Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns.

Table 37. van Ravesteyn et al 2017

Publication	**van Ravesteyn LM, Lambregtse-van den Berg MP, Hoogendijk WJG, Kamperman AM. Interventions to treat mental disorders during pregnancy: A systematic review and multiple treatment meta-analysis. PLoS ONE [Electronic Resource]. [Meta-Analysis]. 2017;12(3):e0173397**
Study details	Systematic review of 29 studies including 2779 participants
Study objectives	To provide an estimation of the overall effect size of a decrease of psychiatric symptoms at the end of treatment or postpartum, for each categorized intervention per mental disorder.
Inclusions	Pregnant women with antenatal mental disorder
Exclusions	Women with addiction and substance misuse
Population	Pregnant women with antepartum mental disorders diagnosed through psychiatric interview
Intervention	Various
Comparator	
Outcomes	• 28 trials focussed on treatment of depressive disorder, 1 trial focussed on anxiety disorders. • All trials used non-pharmacological interventions. **Depressive disorders – robust benefit reducing depressive symptoms was found for:** CBT – 8 trials (*g* = -0.61; 95%CI,-0.73 to -0.49, I_2 = 0%;Z-value = 10.04; p<0.001). IPT- 4 trials (*g* = -0.67; 95%CI,-1.27 to -0.07; I_2 = 79%;Z-value = 2.20; p = 0.03). **Medium benefit in reducing depressive symptoms was found for:** Body oriented interventions 7 trials (*g* = -0.43; 95%CI,-0.61 to -0.25; I_2 = 17,Z-value = 4.62; p<0.001). Acupuncture 2 trials (*g* = -0.43; 95%CI,-0.80 to -0.06; I_2 = 0%; Z-value = 2.30; p = 0.02). **No benefit was found in reducing depressive symptoms for:** Bright light therapy 3 trials (*g* = -0.59; 95%CI,-1.25 to 0.06; I_2 = 0%; Z-value = 0.77; p = 0.08). Food supplements trials 3 trials (*g* = -0.51; 95%CI,-1.02 to 0.01; I^2 = 20%; Z-value = 1.92; p = 0.06).
Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns

Table 38 Weinreb et al 2017

Publication	Weinreb L, Wenz-Gross M, Upshur C. Postpartum outcomes of a pilot prenatal care-based psychosocial intervention for PTSD during pregnancy. Archives of Women's Mental Health. 2017 07 Nov:1-14
Study details	Quasi-experimental pilot trial

<table>
<tr><td>Study objectives</td><td>To examine the subsequent effects on postpartum mental health and birth outcomes for study participants who screened positive for PTSD during pregnancy and received the Seeking Safety intervention compared to women with PTSD in the usual care setting</td></tr>
<tr><td>Inclusions</td><td>Women aged ≥18, below 27 weeks gestation who screened positive using the four-item Primary Care-PTSD Screen at first prenatal appointment from 2 community care centres in the NE US and who spoke 1 of the following languages; English, Spanish, Portuguese, Vietnamese, or Arabic.</td></tr>
<tr><td>Exclusions</td><td>None stated</td></tr>
<tr><td>Population</td><td>149 pregnant women with screen-detected antenatal depression</td></tr>
<tr><td>Intervention</td><td>At 1 of the 2 community care centres (randomly allocated) 89 women received usual prenatal services with the addition of a seeking safety course delivered between first prenatal appointment and delivery. It includes 25 topics, and is based on a cognitive behavioural theoretical framework that emphasizes safe coping, managing PTSD symptoms, and developing healthy behaviours.</td></tr>
<tr><td>Comparator</td><td>At the second (randomly allocated) community care centre usual prenatal services was offered to 60 women (controls) who met the eligibility criteria for the study.</td></tr>
<tr><td>Outcomes</td><td>Severity of PTSD was the only outcome which changed significantly both over time and by group with an effect score of 0.29.
<table>
<tr><th>Outcome</th><th>Baseline mean(SD)</th><th>Postpartum mean (SD)</th><th>P group by time</th><th>Change score mean (SD)</th><th>Change score Cohen’s d</th></tr>
<tr><td>PTSD severity</td><td></td><td></td><td>0.032</td><td>− 8.10 (10.0)</td><td>0.29</td></tr>
<tr><td>Intervention (n = 119)</td><td>21.45 (11.6)</td><td>13.16 (9.1)</td><td></td><td>− 5.11 (10.4)</td><td></td></tr>
<tr><td>Control(n = 135)</td><td>21.64 (10.5)</td><td>16.18 (12.8)</td><td></td><td></td><td></td></tr>
<tr><td>Total(n = 254)</td><td>21.56 (11.0)</td><td>14.76 (11.3)</td><td></td><td></td><td></td></tr>
</table>
No significant difference between groups in terms of:

preterm birth
low birth weight
increase in social support
positive coping skills
negative coping skills
improvements in depression symptoms
</td></tr>
<tr><td>Quality appraisal</td><td>The JBI tool for assessing bias in quasi experimental trials was used to test for quality of this trial and there were the following concerns:</td></tr>
</table>

- Group equivalence and confounding between groups: At baseline, compared to the comparison group the intervention group had a significantly higher proportion of women who:
    - Used tobacco(p<0.002)
    - Used drugs in the past year (p<0.02)
    - Used negative coping strategies(p<0.02)
    - Ever had an abortion(p=0.041)
    - Spoke English (p<0.02)
    - Were born in the USA or Puerto Rico (p<0.001)
- Attrition bias:
    - Of the 149 participants at baseline, 128 (86%) remained for the postpartum interview, including 72 (81%) of intervention and 56 (93%) of control participants ($\chi 2 = 8.71$, df = 1, p = .005).
    - There was a trend for those who were lost to postpartum follow-up to have a higher baseline PTSD symptom score (t147 = 1.84, p = .07), and a trend for them to be part of the intervention group($\chi 2 = 4.58$, df = 1, p = .053).

Table 39. Grote et al 2016

Publication	**Grote NK, Katon WJ, Russo JE, Lohr MJ, Curran M, Galvin E, et al. A Randomised Trial of Collaborative Care for Perinatal Depression in Socioeconomically Disadvantaged Women: The Impact of Comorbid Posttraumatic Stress Disorder. Journal of Clinical Psychiatry. [Comparative Study RandomisedRandomised Controlled Trial]. 2016;77(11):1527-37.**
Study details	Randomised controlled trial
Study objectives	Compare outcomes of pregnant women with comorbid depression and PTSD treated using the MOMCare collaborative care model of brief inter personal psychotherapy and /or anti-depressants versus intensive public health maternity support services (MSS-Plus).
Inclusions	Women aged ≥18 between 12-32 weeks gestation attending one of 10 health centres between January 2010 and July 2012 who scored ≥10 in the PHQ-9 and /or probable dysthymia based on the MINI-International Neuropsychiatric interview(MINI5.0.0)
Exclusions	Women with suicidal behaviour, schizophrenia, bipolar disorder, recent substance misuse, severe partner violence or currently seeing a psychiatrist or psychotherapist.
Population	164 with probable MDD and/or dysthymia
Intervention	MOMCare brief inter-personal psychotherapy (IPT) and /or anti-depressants
Comparator	Usual care – intensive maternity support services
Outcomes	Data about depression severity between groups of women with MDD with and without comorbid PTSD were compared separately by intervention. Overall women with MDD and PTSD showed significantly better improvement with MOMCare compared to MSS-Plus. There was no difference in outcomes between women with MDD who received enhanced maternity services -plus vs MOMCare, they showed similar improvement in both treatment groups. A large effect size (using the Wald test*) was reported at 18 month follow up of MOMCare for severity of depression for the women with MDD and PTSD but a much smaller effect was reported for those with just MDD.

	Women with MDD and PTSD	Women with MDD

	MOMCare vs MSS-Plus between group differences (95% CI)		MOMCare vs MSS-Plus between group differences(95% cl)	
Depression severity (SCL-20)		P value **<0.004**		P value 0.90
3 mo	-0.30(-0.55 to -0.06)	**p<0.02**	0.02 (-0.25 to 0.30)	p=0.88
6mo	-0.39(-0.67 to -0.12)	**p<0.006**	0.03(-0.23 to0.28)	p=0.84
12mo	-0.18(-0.45 to 0.08)	p=0.18	-0.08(-0.30 to 0.46)	p=0.68
18mo	-0.31(-0.58 to -0.95)	**p<0.02**	-0.14(-0.14 to0.42)	p=0.32
Effect size 18 mo	0.39 (favouring MOMCare)		0.11 (favouring MOMCare)	
PTSD severity		**p<0.02**		P=0.62
3mo	NR	NR	NR	NR
6mo	-4.85(-9.70 to0.01)	**P<0.05**	-0.95(-6.14 to 4.24)	P=0.72
12 mo	-3.38(-8.91 ro1.26)	P=0.14	-0.80(-7.30 to 5.70)	P=0.81
18 mo	-6.93(-12.00 to -1.85)	**p<0.008**	-1.96(-7.27 to 3.36)	P=0.47
Remission of depression symptoms(SCL score<0.05)		**p<0.02**		P=0.99
3mo	1.37(-0.43 to 4.30)	P=0.77	1.28(-0.36 to 4.52)	P=0.76
6mo	2.13(-0,89 to 5.07)	P=0.13	1.29(-0.41 to 4.00)	P=0.78
12mo	2.73(1.04 to 7.22)	**P<0.05**	0.65(-0.22 to 1.88)	P=0.59
18mo	2.14(-90 to 5.11)	P=0.13	1.47(-0.50 to 4.33)	P=0.59
*The Wald test is a parametric statistical test used whenever a relationship within or between data items can be expressed as a statistical model with parameters to be estimated from a sample, the Wald test can be used to test the true value of the parameter based on the sample estimate(small effect is 0.02-0.149, medium effect is of 0.15 to 0.349 and large effect 0.35 upwards)				

Quality appraisal

Cochrane collaboration risk of bias tool for RCTs

Bias	High or low bias	Evidence
Random sequence generation(selection bias)	Low	Computerised adaptive block randomisation stratified by initial depression severity and gestational age.
Allocation concealment (selection bias)	Low	Use of computerised randomisation system
Blinding of participants and researchers(performance bias)	High	Participants and researchers were not blinded to intervention.
Blinding of outcome assessment (detection bias)	Low	Telephone interviewer assessing outcomes at 3,6, 12 and 18 months was blinded.
Incomplete outcome data (attrition bias)	Low	5% equivalent across treatment groups

	Selective reporting (reporting bias)	Low	Reported the same primary and secondary outcomes across all treatments.	
	Other bias	Unclear	Unclear	

Table 40. O'Connor et al 2016

Publication	**O'Connor E, Rossom RC, Henninger M, Groom HC, Burda BU. Primary Care Screening for and Treatment of Depression in Pregnant and Postpartum Women: Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA. [Meta-Analysis=Research Support, U.S. Gov't, P.H.S.Review]. 2016;315(4):388-406.**
Study details	Systematic review
Study objectives	To systematically review the benefits and harms of depression screening treatment and accuracy of selected screening instruments for pregnant an postpartum women.
Inclusions	Trials of pharmacological or non-pharmacological interventions with a minimum of 6 weeks follow up after randomisation using population-based screening
Exclusions	Recruitment based on referral or those with known depression (from medical records) or volunteers recruited through media or other advertising.
Population	3 antenatal studies
Intervention	Cognitive behavioural therapy
Comparator	Usual care, no intervention, waitlist, minimal intervention
Outcomes	**Non-pharmacological benefits** Overall CBT and related therapeutic approaches were associated with an increased likelihood of remission; however, point estimates for effect were weak with larger effects generally associated with greater contact hours. Number of contact hours was confounded with other important sources of heterogeneity. The 3 antenatal trials were small (participant numbers 25 to 181) with effect sizes (relative risk) favouring the intervention of 1.55(0.98-2.44), 1.25(0.69-2.27) and 1.36(1.13-1.65). Most trials were small with1 or more methodological limitations and of fair quality **Pharmacological interventions** No studies examined the benefits or harms of the use of antidepressants in screen-detected pregnant women. 5 studies of women using antidepressants during pregnancy with outcomes of interest to this review were identified in this systematic review (n= There was evidence that use of some antidepressants during pregnancy, particularly SSRIs and venlafaxine, are associated with increased risk of preeclampsia, postpartum haemorrhage, and miscarriage as well as a number of adverse infant outcomes, including neonatal or post neonatal death, preterm birth, small for gestational age, neonatal seizures, serotonin withdrawal syndrome, neonatal respiratory distress, pulmonary hypertension, or major congenital malformations. The absolute increase in risk for most infant outcomes was very small, given the rarity of the events, and

sometimes occurred only with higher levels of exposure. For example, a large retrospective cohort study reported a more than doubling of seizure occurrence in infants of depressed women who had been provided 3 or more prescription fills for antidepressants of any kind (but primarily SSRIs). However, the absolute risk remained quite small (0.66% among exposed infants vs 0.28% in unexposed infants; unadjusted OR=2.39 (95% CI, 1.57-3.64)

Results of 5 studies

Vaginal bleeding during pregnancy or postpartum haemorrhage

Increased risk in depressed women (1 study)
SSRI+venlafaxine, current (n=8,917): RR, 1.46 (95% CI, 1.29 to 1.65)
SSRI+venlafaxine, recent (n=4,344): RR, 1.28 (95% CI, 1.08 to 1.52)
Atypical antidepressant, current (n=1,012): RR, 1.52 (95% CI, 1.12 to 2.06)
Controlling for depressed women(1 study)
Increased risk:
Citalopram, current (n=891): RR, 1.48 (95% CI, 1.07 to 2.04) Escitalopram, current (n=1,022): RR, 1.56 (95% CI, 1.16 to 2.09)
Fluoxetine, current (n=3,322): RR, 1.51 (95% CI, 1.27 to 1.79) Paroxetine, current (n=2,055): RR, 1.36 (95% CI, 1.09 to 1.71); recent (n=962): adjusted RR, 1.52 (95% CI, 1.12 to 2.07) Sertraline, current (n=4,526): RR, 1.31 (95% CI, 1.12 to 1.54); recent (n=2,266): RR, 1.27 (95% CI, 1.01 to 1.59)
Venlafaxine, current (n=763): RR, 2.24 (95% CI, 1.69 to 2.97) Bupropion, past (n=1,666): RR, 1.33 (95% CI, 1.03 to 1.71)

Miscarriage and/or spontaneous abortion

Increased risk (1 study)
Venlafaxine (n=NR): RR, 1.80 (95% CI, 1.19 to 2.72) Duloxetine (n=NR): RR, 3.12 (95% CI, 1.55 to 6.31) Mirtazapine (n=NR): RR, 2.23 (95% CI, 1.34 to 3.70)

Pre-term birth or early gestational age

Increased risk: (1 study)
Any antidepressant (mostly SSRIs), % born gestational weeks
32-36:1-2 prescriptions (n=10,700): OR 1.91(95% CI, 1.77 to 2.07)
3+ prescriptions (n=6,196): OR1.12 95% CI, 1.03 to 1.23)
Unknown Depression Status in Control Group N=228,876)
Increased risk:
SSRIs in 2nd trimester (mean difference in days, n=NR for all,
nulliparous women):
1 prescription: -2.6 (95% CI, -1.3 to -3.9)
2 prescriptions: -5.8 (95% CI, -3.8 to -7.8)
3+ prescriptions: -6.6 (95% CI,-4.6 to -8.6)
Decreased risk:
SSRIs in 3rd trimester (mean difference in days, n=NR for all, nulliparous women): 1 prescription: 0.9 (95% CI, 0.3 to 1.6)
2 prescriptions: 1.8 (95% CI, 0.9 to 2.7)
3+ prescriptions: 6.4 (95% CI, 5.5 to 7.3)

Low birth weight or small for gestational age (SGA)

All Women, Controlling for Depression (1 study)
Increased risk
SSRIs during pregnancy (n=NR): HR, 1.22 (95% CI, 1.13 to 1.32)

The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns.

Table 41. McDonagh et al 2014

Publication	**McDonagh MS, Matthews A, Phillipi C, Romm J, Peterson K, Thakurta S, et al. Depression drug treatment outcomes in pregnancy and the postpartum period: A systematic review and meta-analysis. Obstetrics and Gynaecology. 2014;124(3):526-34**
Study details	Systematic review and meta-analysis
Study objectives	To evaluate the comparative benefits and harms in both mother and child of antidepressant treatment for depression in pregnant or postpartum women.
Inclusions	Pregnant and postpartum women using antidepressants
Exclusions	None stated
Population	6 randomisedrandomised controlled trials and 15 observational studies.
Intervention	N/A
Comparator	N/A
Outcomes	Five small (n=44-225) observational studies from the US assessed outcomes associated with antidepressant treatment during pregnancy The outcomes with sufficient direct evidence (studies comparing interventions of interest in the population of interest and measuring outcomes of interest) were preterm birth (no increased risk), neonatal convulsions (no increased risk), and respiratory distress (significant increased risk, OR=1.91, 95% CI 1.63–2.24).
Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns.

Question 5: What are the benefits of early pharmacological and non-pharmacological intervention (alone or in combination) in women with screen-detected postnatal depression?

Table 42. Letourneau et al 2017

Publication	**Letourneau NL, Dennis CL, Cosic N, Linder J. The effect of perinatal depression treatment for mothers on parenting and child development: A systematic review. Depression and Anxiety. 2017 October;34(10):928-66.**
Study details	Systematic review
Study objectives	To examine which treatment interventions for antenatal or postpartum depression are most effective in improving both parenting and child development.
Inclusions	Women defined as depressed in the antenatal or postpartum period using a valid assessment tool or been diagnosed by a physician.
Exclusions	None stated
Population	28 studies of postpartum treatment
Intervention	Any intervention
Comparator	
Outcomes	• Impact of CBT on parenting stress index-short form showed no overall all effect following meta-analysis (SMD=0.154, 95%CI, -0.005 -0.313,p=0.057). Meta-analysis was not possible for other study outcomes and a range of small individual studies showed inconsistent results on the impact of CBT on PND.

	• 2 trials investigated the effect of peer support on a range of parent and child outcomes but none showed differences between the intervention and usual care groups.
	• Maternal interaction guidance was examined in 6 trials that reported varying findings. Effect sizes in 3 of the trials for some outcomes were reported:
	  ○ Effect of verbal feedback to improve maternal parenting behaviour (effect size 0.51)
	  ○ Quality of interactive behaviour (effect size 0.68 and 0.66)

Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns.

Table 43. Pritchett et al 2017

Publication	**Pritchett RV, Daley AJ, Jolly K. Does aerobic exercise reduce postpartum depressive symptoms? a systematic review and meta-analysis. British Journal of General Practice. 2017;67(663):e684-e91**
Study details	Meta-analysis
Study objectives	To undertake a meta-analysis of randomised controlled trials (RCTs) investigating the effect of aerobic exercise interventions, compared with usual care, on depressive symptoms in women up to 1 year postpartum.
Inclusions	• Populations of mothers <1 year postpartum • Interventions designed to increase aerobic exercise (activity causing increased heart rate, respiratory rate, and sweating), including those with co-interventions such as peer support, educational sessions about postpartum issues and dietary support • Depressive symptoms measured by questionnaire or diagnostic interview;
Exclusions	Trials comparing 2 types of exercise.
Population	13 RCTs reporting the effect of exercise on PND
Intervention	Exercise or exercise plus co intervention
Comparator	No care or any form of usual care
Outcomes	• Overall exercise interventions significantly reduced depressive symptoms (SMD –0.44, 95% CI: –0.75 to –0.12, n = 1307, I^2 85%, 13 trials) • Exercise interventions had a significant effect in reducing depressive symptoms in 'depressed' postpartum populations (SMD –0.32, 95% CI, –0.63 to –0.00), I 2 55% and in general postpartum populations (-0.57, 95% CI, –1.12 to –0.02, I^2 92%) • Exercise with co-interventions had a significant effect on reducing depressive symptoms (SMD–0.35, 95% CI, –0.66 to –0.04, I^2 72%). • Exercise-only interventions had a non-significant effect in reducing depressive symptoms (SMD –0.56, 95% CI,–1.13 to 0.01, I^2 89%). • Group exercise interventions had a significant effect in reducing depressive symptoms (SMD difference –1.10, 95% CI, -1.99 to –0.21, I^2 93%) • Participant choice interventions such as exercise counselling with personal choice of exercise (often exercise alone) had a significant effect in reducing depressive symptoms (–0.20, 95% CI,–0.33 to –0.06, I^2 0%).
Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns.

Table 44. Morrell et al 2016

Publication	**Morrell CJ, Sutcliffe P, Booth A, Stevens J, Scope A, Stevenson M, et al. A systematic review, evidence synthesis and meta-analysis of quantitative and qualitative studies evaluating the clinical effectiveness, the cost-effectiveness, safety and acceptability of interventions to prevent postnatal depression. Health Technology Assessment (Winchester, England). 2016;20(37):1-414**
Study details	Health Technology Assessment
Study objectives	To evaluate the clinical effectiveness, cost-effectiveness, acceptability and safety of antenatal and postnatal interventions for pregnant and postnatal women to prevent PND
Inclusions	Pregnant women (universal). • Postnatal women with a live baby born within the previous 6 weeks • Vulnerable pregnant or postnatal women who were aged less than 18 years; at risk of violence; an ethnic minority; HIV positive; living in deprivation, financial hardship or poverty; or single, socially disadvantaged or unsupported • Pregnant or postnatal women with a raised score on the antenatal risk questionnaire, Beck Depression Inventory, Center for Epidemiologic Studies Depression scale, the Cooper predictive index, depression symptom checklist, EPDS, HADS, Hamilton Depression Rating Scale, Health during pregnancy questionnaire; a past history of PND or major depression (indicated). • Pregnant women with a diagnosis of depression using Research Diagnostic Criteria or DSM-IV criteria (indicated).
Exclusions	• Postnatal women with a diagnosis of PND. • Pregnant women with comorbid psychiatric disorders. • Postnatal women with major medical problems.
Population	122 papers based on 86 RCTs
Intervention	Any intervention to prevent the development of PND
Comparator	
Outcomes	The most beneficial interventions at 12 months, shown by difference in the mean EPDS score: • midwifery redesigned postnatal care [–1.43, 95% credible interval (CrI) –4.00 to 1.36], • person-centred approach (PCA)-based intervention (–0.97, 95% CrI –3.54 to 1.71) • cognitive behavioural therapy (CBT)-based intervention (–0.78, 95% CrI –3.41 to 1.91). The HTA rated the quality of the clinical and cost effectiveness evidence available as poor, indicating that replication of some studies is needed within good-quality RCTs including: • PCA-based intervention • CBT-based intervention • Education on preparing for parenting, peer support and • IPT-based intervention • promoting parent–infant interaction • peer support

Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns.

Table 45. O'Connor et al 2016

Publication	**O'Connor E, Rossom RC, Henninger M, Groom HC, Burda BU. Primary Care Screening for and Treatment of Depression in Pregnant and Postpartum Women: Evidence Report and Systematic Review for the US Preventive Services Task Force. JAMA. [Meta-Analysis=Research Support, U.S. Gov't, P.H.S.Review]. 2016;315(4):388-406.**
Study details	Systematic review
Study objectives	To systematically review the benefits and harms of depression screening treatment and accuracy of selected screening instruments for pregnant and postpartum women.
Inclusions	Trials of pharmacological or non-pharmacological interventions with a minimum of 6 weeks follow up after randomisation using population-based screening
Exclusions	Recruitment based on referral or those with known depression (from medical records) or volunteers recruited through media or other advertising.
Population	15 trials reporting on postpartum interventions for PND
Intervention	Any intervention
Comparator	Usual care, no intervention, waitlist, minimal intervention
Outcomes	**Pharmacological benefits** Overall CBT and related therapeutic approaches were associated with an increased likelihood of remission. Larger effects were generally associated with greater contact hours; however, contact hours was confounded with other important sources of heterogeneity. For reduction in depression symptom severity, likelihood of remission over the short term (<8mo) • 7 CBT postpartum trials had small effect sizes ranging from 1.09 (95%CI, 0.84 -1.42) to 1.97 (95%CI, 0.88 - 4.44). • 3 trials of nondirective therapeutic approaches reported small effect sizes of 0.96 (95%CI, 0.72 –1.27) to 3.20 (95%CI,1.32 -7.76) **Pharmacological benefits** One trial of pregnant women examined the benefit of fluoxetine on depression symptoms and reported a 10 point reduction in EPDS with fluoxetine after 12 weeks compared with a 7 point reduction with a placebo (p=0.05) Most trials were small with 1 or more methodological limitations and of fair quality.
Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns.

Table 46. McDonagh et al 2014

Publication	**McDonagh MS, Matthews A, Phillipi C, Romm J, Peterson K, Thakurta S, et al. Depression drug treatment outcomes in pregnancy and the postpartum period: A systematic review and meta-analysis. Obstetrics and Gynaecology. 2014;124(3):526-34**
Study details	Systematic review and meta-analysis

Study objectives	To evaluate the comparative benefits and harms in both mother and child of antidepressant treatment for depression in pregnant or postpartum women.
Inclusions	Pregnant and postpartum women using antidepressants
Exclusions	None stated
Population	6 randomisedrandomised controlled trials and 1 observational studies
Intervention	Antidepressants
Comparator	Other antidepressants, no care or usual care
Outcomes	There was low strength evidence that symptom response was not improved when sertraline was added to psychotherapy or when CBT was added to paroxetine. There was insufficient evidence to determine other outcomes such as functional capacity, breastfeeding and infant child development. No figures were reported.
Quality appraisal	The CASP checklist for systematic reviews was used to assess the quality of this review and there were no concerns.

Studies relevant to question 6 criterion 15: Is clinical detection and management currently well implemented in the UK?

Table 47. Ford et al 2017

Publication	**Ford E, Lee S, Shakespeare J, Ayers S. Diagnosis and management of perinatal depression and anxiety in general practice: a meta-synthesis of qualitative studies. British Journal of General Practice. 2017;67(661):e538-e46**.
Study details	Meta synthesis of qualitative studies
Study objectives	Synthesising information about GP attitudes to recognition and management of anxiety and depression in pregnancy and postpartum
Inclusions	Papers and reports if they reported qualitatively on attitudes and decision making or routine clinical practice for diagnosis or treatment of anxiety and depression in pregnancy and postpartum.
Exclusions	Published before 1990 did not report original research, were not published in English, reported quantitative results, or if GPs were not the main participants of reported separately.
Population	5 papers including 4 from peer reviewed journals and one report from grey literature.
Intervention	N/A
Comparator	N/A
Outcomes	4 of the 5 publications were based on 3 UK studies. The remaining paper was based on a survey of Australian GPs. The 3 UK studies included a postal survey with 43 responses, in depth interviews of 19 GPS participating in an RCT about support for women with PND and a report of semi structured interviews with 8 GPs about their use and experience of guidelines for perinatal depression. Themes from the 4 UK publications included • Reluctance to identify a condition with a diagnostic label (2/4 publications) • Clinical judgement vs use of guidelines with a preference for GPs to rely on their own intuition to pick up the signs if something was wrong (3/4 publications).

	• Care and management of women identified – approach was determined by GP own experience and resources available (4/4 publications) • Use of medication – approach depends on what non-pharmacological resources are available and balance of risk of treatment vs no treatment.(1/4 publications) • Isolation and role of other professionals – links between GPs and Health visitors is much reduced so less communication about women who might have problems.(2/4 publications)
Quality appraisal	There are currently no checklists available to critically appraise a meta synthesis of qualitative studies so the CASP checklist for systematic reviews was used. There were no quality concerns identified with the methodology of the meta synthesis although the research papers included only reported a total of 70 UK GP views (there are approximately 34,000 UK GPs) and may not be representative. The authors used a checklist derived from Atkins et al 2008to indicate the range of quality of the qualitative research. Out of 11 points all studies scored 9 or 10.

Table 48. McGlone et al 2016

Publication	**McGlone C, Hollins Martin CJ, Furber C. Midwives' experiences of asking the Whooley questions to assess current mental health: a qualitative interpretive study. Journal of Reproductive and Infant Psychology. 2016 07 Aug;34(4):383-93**
Study details	Qualitative research - 1 to 1 semi structured interviews
Study objectives	To explore midwives experiences of asking the Whooley questions with pregnant women.
Inclusions	Midwives who agreed to 1 to1 interviews in 1 maternity unit in the UK and who had to regularly conduct antenatal booking visits
Exclusions	None stated
Population	8 midwives
Intervention	N/A
Comparator	N/A

Outcomes	Themes and sub themes from the interviews with the 8 midwives include:
	• No clear understanding of the purpose of the Whooley questions
	  ○ Midwife discomfort when women disclose signs of depression and anxiety
	• Feeling pressurised from lack of time and frustration at unable to fulfil their role
	• Resultant dissatisfaction to fulfil their role
	  ○ Lack of knowledge and how to refer
	  ○ Lack of training around perinatal mental illness
	  ○ Rely on experience and intuition rather than training

Quality appraisal	The CASP checklist for qualitative research was used to assess the quality of this study. Overall there were no areas of concern about the quality of the study although the number of midwives interviewed was low and they were all from one maternity unit which may not be representative of other units across the country.

Table 49. Prady et al 2016

Publication	**Prady SL, Pickett KE, Petherick ES, Gilbody S, Croudace T, Mason D, et al. Evaluation of ethnic disparities in detection of depression and anxiety in primary care during the maternal period: combined analysis of routine and cohort data. British Journal of Psychiatry. 2016;208(5):453-61**
Study details	Longitudinal study
Study objectives	To describe Common Mental Disorders(CMD) in primary care in before during and after pregnancy
Inclusions	Women consecutively recruited at 26-28 weeks pregnancy between 2007 and 2010 at routine hospital appointment
Exclusions	Women with no GP linked records because they moved practice either into or out of the study area during the study period. Women were excluded if they had severe or suspected severe mental illness or if they did not complete the study questionnaires.
Population	8991 women
Intervention	Data collected: • Women completed a 28 item General Health Questionnaire(GHQ-28) at 26-28 weeks gestation • GP linked records were downloaded from 6 months prior to conception to 3 years postpartum • Hospital maternity records of delivery date and gestational age at birth For women with no evidence of common mental disorders in the GP record a GHQ-28 summary threshold score of ≥15 as an indication of likely psychiatric morbidity was used. This was high (recommended threshold is 5 to 8) to minimise false positive GHQ-28 screens. Where common mental disorders were already in the GP record the GHQ-28 screen was ignored.
Comparator	
Outcomes	

Time period and parity	**Common mental health disorders (%)**	**Confidence intervals (95%)**
Period Prevalence (Identified through GP records)		
Pre-birth(6 months prior to conception and pregnancy)	9.5%	8.9 to 10.1
1st year postpartum	13.1%	12.4 to 13.7
2nd year postpartum	12.8%	11.9 to 13.6
3rd year postpartum	14.0%	13.2 to 14.9
Incidence (Identified through GP records)	**CMD Rate per 1000 person years at risk**	**CI (95%)**
Pregnancy	37.5	33.1 to 42.5
1st year postpartum	102.4	95.5 to 109.8
2nd year postpartum	64.9	59.1 to 72.4
3rd year postpartum	68.2	61.1 to76.2
Incident rate ratio difference in CMD between white British women and Pakistani women	**CMD Incidence rate ratio**	**CI (95%)**

Pregnancy : White British women vs Pakistani women	1.44	1.07 to 1.96
1st year postpartum: White British women vs Pakistani women	1.95	1.64 to 2.32
2nd year postpartum White British women vs Pakistani women	1.66	1.27 to 2.19
3rd year postpartum White British women vs Pakistani women	1.82	1.58 to 2.09
Potentially missed cases of CMD	**CMD %**	**CI (95%)**
Pre-birth (6 months prior to conception and pregnancy) GHQ-28 threshold ≥15	31.3%	28.7 to 33.8
Pre-birth (6 months prior to conception and pregnancy) GHQ-28 threshold ≥9	46.8%	44.7 to 49.0

Fewer than 13% of women had read codes in their records indicating screening and case finding of CMD in the first postnatal year; twice as many white British women had these codes compared to those of other ethnicities. Minority ethnic women had twice the rate of potentially missed cases and half the volume of screening or case finding records.

Quality appraisal	A CASP checklist for longitudinal studies was undertaken. Limitations of the study include lack of knowledge of the quantity of missing data from primary care records, only administering the GHQ-28 once in mid to late pregnancy and applying conservative weighting of results so estimates are likely to be under rather than over estimated. The population has low socioeconomic diversity and is disadvantaged and was served by one maternity unit. This may make the results less generalizable across the rest of the UK.

Table 50. Redshaw et al 2016

Publication	**Redshaw M, Henderson J. Who is actually asked about their mental health in pregnancy and the postnatal period? Findings from a national survey. BMC Psychiatry. 2016;16(1):**
Study details	Secondary data analysis of a national maternity survey carried out in 2014
Study objectives	To determine which women are asked about their mood and mental health during pregnancy and postnatally and about the offer and uptake of treatment.
Inclusions	Random selection of women who had given birth in first half of January 2014 identified by ONS birth statistics
Exclusions	Age <16, or baby had died
Population	4521 women
Intervention	N/A
Comparator	N/A
Outcomes	

	Detecting mental health problem		Mental health service provided		
	Asked about mental health (%)	Offered help	Received support	Received advice	Received treatment

	Antenatal total cohort	3703(81.9%)	Of those reporting an MH issue 337/946(35.7%)	215(65.4%)	229 (71.3%)	119(45.2%)
	Postnatal total cohort	4043 (89.8%)	Of total postnatal cohort 522(12.9%)	263(63.4%)	165(64.4%)	262(50.2%)
	Antenatal					
	White British women	3045(82.6)	277 (41.2)	177 (74.4)	188(76.4)	92(46.2)
	Women of mixed ethnicity	73(83.9)	9 (40.9)	4 (50.0)	4 (57.1)	3 (60.0)
	Asian women	331(75.7)	29 (19.6)	19 (52.8)	209(55.6)	14(46.7)
	Black women	130(83.9)	7 (17.9)	5 (50.0)	5 (71.4)	5(62.5)
	Women of other ethnicities	18(78.3)	3 (37.5)	2 (100)	2(100)	1(100)
	Total	**3705(81.9)^**	**325(36.6)^^**	**207(70.4)^^^**	**219 (73.5)**	**115(47.3)**
	Postnatal					
	White British women	3399 (91.9)	Not reported	217(67.4)	134(67.3)	216(53.8)
	Women of mixed ethnicity	74 (85.1)	Not reported	9(90.0)	3 (70.0)	7(50.0)
	Asian women	350 (80.1)	Not reported	24(42.9)	18(50.8)	31(33.3)
	Black women	124 (80.0)	Not reported	6(46.2)	5(58.3)	7(50.0)
	Women of other ethnicities	16 (69.6)	Not reported	2(50.0)	2(33.3)	1(50.0)
	Total	**3963 (90.1)****	**Not reported**	**258(62.7)***	**162(64.4)**	**262(50.2)**
	^p<0.05 During pregnancy Asian women were less likely to be asked about their mental health than other women ^^p<0.01 During pregnancy white British women more likely to be offered treatment than either Asian or Black women. ^^^p<0.05 During pregnancy white British women more likely to receive support than other women **p<0.01 Postpartum women who are white are more likely to be asked about their mental health *p<0.05 Postpartum women who are Asian or Black are less likely to receive support than other women.					
Quality appraisal	The CASP checklist cohort study checklist was used to assess the quality of this study. Limitations of the work include an under representation of survey respondents who were younger, single, born outside the UK and living in disadvantaged areas making the results less generalisable to the whole of the UK.					

Table 51. Williams et al 2016

Publication	**Williams CJ, Turner KM, Burns A, Evans J, Bennert K. Midwives and women's views on using UK recommended depression case finding questions in antenatal care. Midwifery. [Research Support, Non-U.S. Gov't]. 2016;35:39-46.**
Study details	Qualitative study - semi structured interviews with midwives and women

Study objectives	Exploring midwives and pregnant women's views and experiences of screening and case finding using the questions recommended by NICE guidance.
Inclusions	Midwives and pregnant women
Exclusions	None stated
Population	15 midwives and 20 pregnant women
Intervention	N/A
Comparator	N/A
Outcomes	Key themes midwives: • Asking the questions at the booking appointment was not always appropriate as women may have early transient pregnancy symptoms that coloured their response and it may be the first time the women and midwife had met so a relationship had not been established to facilitate disclosure • Some midwives felt the questions were a bit blunt • Sometimes midwives did not use the questions if it felt inappropriate eg if someone else was with the woman • When women could not speak English or where they had learning disabilities and there was a greater challenge in communicating a sensitive topic and more reluctance on the part of the midwife • Some midwives were more uncomfortable asking the questions than others • The 2Whooley questions were seen as on the whole a good way to broach the subject and drew out issues that might not otherwise arise • Midwives felt the Arroll 'help' question helped them to understand how the women were feeling but there were not a lot of options for support they could offer them. Key points pregnant women: • Most women did not expect to be asked about mental health issues at the booking appointment • Some felt the questions were a bit blunt • Many women found it a good opportunity to talk about pregnancy anxiety and whether what they were experiencing was normal • Women felt both that it was an advantage to be able to talk to the midwife about how they were feeling but also felt admitting they were feeling low during pregnancy carried social stigma. • How the midwife was perceived (brisk or distracted versus empathic and non-judgemental) had an impact on the women's response to the questions and their likelihood of disclosing issues. • All the women were happy to be asked the Whooley questions as it was an opportunity to talk about how they felt. • Women understood the Arroll 'help' question variably – women thought it could mean seeing a GP for medication, a counsellor, specialist support group or psychiatrist or in one case being offered a self-help book. Some did not know what help would consist of and the midwives didn't clarify this either.
Quality appraisal	The CASP checklist for qualitative research was used to assess the quality of this study. Overall there were no concerns about the quality of the publication. The sample size was small (35 participants in total) and recruitment was via a validation study already being undertaken in Bristol which may limit generalisability to the rest of the UK. All views are reported by women and midwives rather than from direct observation of the booking appointment.

Table 52. Darwin et al 2015

Publication	**Darwin Z, McGowan L, Edozien LC. Antenatal mental health referrals: review of local clinical practice and pregnant women's experiences in England. Midwifery. [Research Support, Non-U.S. Gov't]. 2015;31(3):e17-22.**
Study details	Mixed methods cohort study
Study objectives	To investigate the consistency and completeness of mental health assessment, subsequent management of pregnant women identified with mental health problems and women's experiences of the mental health referral process.
Inclusions	Women able to provide written consent and complete English language questionnaires unassisted.
Exclusions	None stated
Population	191 women early pregnancy
Intervention	N/A
Comparator	N/A
Outcomes	

Booking appointment	
191	Completed research questionnaire and mental health assessment in the hand held maternity notes (Whooley questions and Arroll 'help' question)
167/191	Responses to Whooley and Arroll questions present in both hand held records and health records
5/167	Whooley and Arroll questions not completed
30/162	Had ticked at least 1 of the 2 Whoolley items
21/30	Comments documented by midwives in hand held notes
8/21	Had consistency between handheld notes and main hospital notes
11/30	Midwives explored other factors in discussion indicating the symptoms were more pregnancy related physical symptoms such as nausea
6/162	Ticked the Arroll question wanting help or support
6/6	All those ticking the Arroll question for 'help' recorded a discussion in hand held notes
3/6	Who ticked the Arroll 'help' question was referred to Mental Health Specialist Midwife
3/6	Who ticked the Arroll 'help' question were referred elsewhere (GP, social services and treatment for hyperemesis gravidarum)
0	Completed the questionnaires in trimesters 2 and 3
Referrals	
23/191	Referred to Mental health Midwife service at booking all had current or past mental health histories documented
5/191	With mental health histories not referred as booking midwife documented there were no mental health concerns
20/23	Handheld notes were available to evaluate
10/20	Had ticked at least 1 Whooley question
3/10	Ticked the Arroll 'Help' item
11/23	Reference to current symptoms made in referral but not to the Whooley questions
Response of Mental Health Specialist Midwife	
3/23	'will contact patient'

	12/23	‘’Does not meet criteria’ or ‘no plans to contact’ with or without suggestion to contact GP for support (n=6) People who did not meet the criteria for support from the service included women currently taking medication, women with past and continuing PND, women with past and continuing anxiety, women with psychological stress combined with other difficulties such as unwanted pregnancy, eating problems a carers role or lack of support
	8/23	No evidence of response from Mental health Specialist Midwife
	Interviews were carried out with 22 of the 167 women responding to the research questionnaire whose hand held notes and health notes were available. Of those 22, 6 had been referred to the specialist midwifery service but none of them were contacted by the service or informed of the outcome of their referral. One person was re-referred by the research group due to disclosure of symptoms and history of severe mental illness and received specialist psychiatric support. This person was also the only one of the 6 who had any ongoing monitoring in the postnatal period. The fragmented communication with support focussed on severe mental illness is partly attributed to lack of resources in perinatal mental health care and lack of a jointly developed (GPs, midwives and the mental health service) care pathways for those with common mental disorders.	
Quality appraisal	A CASP checklist for longitudinal studies was undertaken and no concerns were identified. It is not known whether the findings of this local study are generalizable across the UK.	

Appendix 4 –Summary of NICE guidance (CG192): Antenatal and postnatal mental health: clinical management and service guidance

Recognising mental health problems in pregnancy and the postnatal period and referral

Recognise that women who have a mental health problem (or are worried that they might have) may be:

- unwilling to disclose or discuss their problem because of fear of stigma, negative perceptions of them as a mother or fear that their baby might be taken into care
- reluctant to engage, or have difficulty in engaging, in treatment because of avoidance associated with their mental health problem or dependence on alcohol or drugs.
- All healthcare professionals referring a woman to a maternity service should ensure that communications with that service (including those relating to initial referral) share information on any past and present mental health problem.

Depression and anxiety disorders

Recognise that the range and prevalence of anxiety disorders (including generalised anxiety disorder, obsessive-compulsive disorder, panic disorder, phobias, post-traumatic stress disorder and social anxiety disorder) and depression are under-recognised throughout pregnancy and the postnatal period.

At a woman's first contact with primary care or her booking visit, and during the early postnatal period, consider asking the following depression identification questions as part of a general discussion about a woman's mental health and wellbeing:

- during the past month, have you often been bothered by feeling down, depressed or hopeless?
- during the past month, have you often been bothered by having little interest or pleasure in doing things?

Also consider asking about anxiety using the 2-item Generalized Anxiety Disorder scale (GAD-2):

- over the last 2 weeks, how often have you been bothered by feeling nervous, anxious or on edge?
- over the last 2 weeks, how often have you been bothered by not being able to stop or control worrying?

If a woman responds positively to either of the depression identification questions, is at risk of developing a mental health problem, or there is clinical concern, consider: using the Edinburgh Postnatal Depression Scale (EPDS) or the Patient Health Questionnaire (PHQ-9) as part of a full assessment or referring the woman to her GP or, if a severe mental health problem is suspected, to a mental health professional.

If a woman scores 3 or more on the GAD-2 scale, consider:

- using the GAD-7 scale for further assessment or
- referring the woman to her GP or, if a severe mental health problem is suspected, to a mental health professional.

If a woman scores less than 3 on the GAD-2 scale, but you are still concerned she may have an anxiety disorder, ask the following question:

- do you find yourself avoiding places or activities and does this cause you problems?

If she responds positively, consider:

- using the GAD-7 scale for further assessment or
- referring the woman to her GP or, if a severe mental health problem is suspected, to a mental health professional.

At all contacts after the first contact with primary care or the booking visit, the health visitor, and other healthcare professionals who have regular contact with a woman in pregnancy and the postnatal period (first year after birth), should consider asking the 2 depression identification questions and the GAD-2 as part of a general discussion about her mental health and wellbeing and using the EPDS or the PHQ-9 as part of monitoring.

Appendix 5 – UK NSC reporting checklist for evidence summaries

All items on the UK NSC Reporting Checklist for Evidence Summaries have been addressed in this report. A summary of the checklist, along with the page or pages where each item can be found in this report, is presented in Table5.

Table 53. UK NSC reporting checklist for evidence summaries

	Section	Item	Page no.
1.	TITLE AND SUMMARIES		
1.1	Title sheet	Identify the review as a UK NSC evidence summary.	Title page
1.2	Plain English summary	Plain English description of the executive summary.	5
1.3	Executive summary	Structured overview of the whole report. To include: the purpose/aim of the review; background; previous recommendations; findings and gaps in the evidence; recommendations on the screening that can or cannot be made on the basis of the review.	7
2.	INTRODUCTION AND APPROACH		
2.1	Background and objectives	Background – Current policy context and rationale for the current review – for example, reference to details of previous reviews, basis for current recommendation, recommendations made, gaps identified, drivers for new reviews Objectives – What are the questions the current evidence summary intends to answer? – statement of the key questions for the current evidence summary, criteria they address, and number of studies included per question, description of the overall results of the literature search. Method – briefly outline the rapid review methods used.	11
2.2	Eligibility for inclusion in the review	State all criteria for inclusion and exclusion of studies to the review clearly (PICO, dates, language, study type, publication type, publication status etc.) To be decided *a priori*.	16, 22, 35, 42, 49, 63, 73

2.3	Appraisal for quality/risk of bias tool	Details of tool/checklist used to assess quality, e.g. QUADAS 2, CASP, SIGN, AMSTAR.	21
3.	SEARCH STRATEGY AND STUDY SELECTION (FOR EACH KEY QUESTION)		
3.1	Databases/ sources searched	Give details of all databases searched (including platform/interface and coverage dates) and date of final search.	86
3.2	Search strategy and results	Present the full search strategy for at least one database (usually a version of Medline), including limits and search filters if used. Provide details of the total number of (results from each database searched), number of duplicates removed, and the final number of unique records to consider for inclusion.	86
3.3	Study selection	State the process for selecting studies – inclusion and exclusion criteria, number of studies screened by title/abstract and full text, number of reviewers, any cross checking carried out.	98
4.	STUDY LEVEL REPORTING OF RESULTS (FOR EACH KEY QUESTION)		
4.1	Study level reporting, results and risk of bias assessment	For each study, produce a table that includes the full citation and a summary of the data relevant to the question (for example, study size, PICO, follow-up period, outcomes reported, statistical analyses etc.). Provide a simple summary of key measures, effect estimates and confidence intervals for each study where available. For each study, present the results of any assessment of quality/risk of bias.	Study level reporting: Quality assessment: 112
5.	QUESTION LEVEL SYNTHESIS		
5.1	Description of the evidence	For each question, give numbers of studies screened, assessed for eligibility, and included in the review, with summary reasons for exclusion.	23,36,43,50, 64,71
5.2	Combining and presenting the findings	Provide a balanced discussion of the body of evidence which avoids over reliance on one study or set of studies. Consideration of four components should inform the reviewer's judgement on whether the criterion is 'met', 'not met' or 'uncertain': quantity; quality; applicability and consistency.	23,37,48,55, 64,73

5.3	Summary of findings	Provide a description of the evidence reviewed and included for each question, with reference to their eligibility for inclusion. Summarise the main findings including the quality/risk of bias issues for each question. Have the criteria addressed been 'met', 'not met' or 'uncertain'?	34,43,51,72, 82
6.	REVIEW SUMMARY		
6.1	Conclusions and implications for policy	Do findings indicate whether screening should be recommended? Is further work warranted? Are there gaps in the evidence highlighted by the review?	83
6.2	Limitations	Discuss limitations of the available evidence and of the review methodology if relevant.	85

References

[1] National Institute for Clinical and Care Excellence (NICE) Antenatal and postnatal mental health: clinical management and service guidance Clinical guideline [CG192] Published 2014 updated April 2018

[2] Information Services Division Scotland; Births in Scottish hospitals 2016 http://www.isdscotland.org/Health-Topics/Maternity-and-Births/Publications/data-tables.asp accessed September 20th 2018

[3] Office for National Statistics. Statistical bulletin: Conceptions in England and Wales 2016 https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/conceptionandfertilityrates/bulletins/conceptionstatistics/2016 Accessed September 20th 2018

[4] NICE Common Mental Health Disorders: Identification and pathways to care (CG123), (2011) The British Psychological Society and The Royal College of Psychiatrists

[5] Scottish Intercollegiate Guidelines Network (SIGN). Management of perinatal mood disorders. Edinburgh: SIGN; 2012. (SIGN publication no. 127). [March 2012].Available from URL: http://www.sign.ac.uk

[6] GRADE working group. Grading quality of evidence and strength of recommendations BMJ 2004 Jun 19; 328(7454): 1490

[7] Raffle A and Gray M. Screening: Evidence and Practice 2007 Oxford University Press, Oxford

[8] Hill C (2011) An evaluation of screening for postnatal depression against NSC criteria

[9] Howard L Molyneaux E, Dennis C, Rochat T, Stein A, Milgrom J Non psychotic mental health disorders in the perinatal period 2014. Lancet; 384; 1775-85

[10] Jones I, Chandra P, Dazzan P, Howard L. Bipolar disorder, affective psychosis and schizophrenia in pregnancy and the postpartum period. 2014 Lancet; 384;1789-99

[11] Stein A, Pearson R, Goodman S, Rapa E, Rahman A, McCallum M, Howard L, Pariante C Effects of perinatal mental disorders on the fetus and child. 2014. Lancet; 384; 1800-1819

[12] Gelaye B, Kajeepeta S, Williams MA. Suicidal ideation in pregnancy: An epidemiologic review. Archives of Women's Mental Health. 2016 Oct;19(5):741-51

[13] Cook N, Ayers S, Horsch A. Maternal post-traumatic stress disorder during the perinatal period and child outcomes: A systematic review. Journal of Affective Disorders. [Review]. 2018;225:18-31

[14] Jarde A, Morais M, Kingston D, Giallo R, MacQueen GM, Giglia L, et al. Neonatal outcomes in women with untreated antenatal depression compared with women without depression: A systematic review and meta-analysis. JAMA Psychiatry. 2016 August;73(8):826-37

[15] MacKinnon N, Kingsbury M, Mahedy L, Evans J, Colman I. The Association Between Prenatal Stress and Externalizing Symptoms in Childhood: Evidence From the Avon Longitudinal Study of Parents and Children. Biological Psychiatry. 2018;83(2):100-8

[16] Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis.Biometrics. 2000;56(2):455-463

[17] Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed December 17, 2015.

[18] Down S and Black N 1998 The feasibility of creating a checklist for the assessment of the methodological quality of randomised and non randomised studies of health care intervention Journal of Epidemiology Community Health 52; 377-384

[19] Kingsbury M, Weeks M, MacKinnon N, Evans J, Mahedy L, Dykxhoorn J, et al. Stressful Life Events During Pregnancy and Offspring Depression: Evidence From a Prospective Cohort Study. Journal of the American Academy of Child & Adolescent Psychiatry. 2016;55(8):709-16.e2

20 Liu Y, Kaaya S, Chai J, McCoy DC, Surkan PJ, Black MM, et al. Maternal depressive symptoms and early childhood cognitive development: a meta-analysis. Psychological medicine. 2017 01 Mar;47(4):680-9.

21 Tarabulsy GM, Pearson J, Vaillancourt-Morel M-P, Bussieres E-L, Madigan S, Lemelin J-P, et al. Meta-analytic findings of the relation between maternal prenatal stress and anxiety and child cognitive outcome. Journal of Developmental and Behavioral Pediatrics. 2014 Jan;35(1):38-43

22 Sanger C, Iles JE, Andrew CS, Ramchandani PG. Associations between postnatal maternal depression and psychological outcomes in adolescent offspring: a systematic review. Archives of Women's Mental Health. 2015;18(2):147-62

[23]Closa-Monasterolo R, Gispert-Llaurado M, Canals J, Luque V, Zaragoza-Jordana M, Koletzko B, et al. The effect of postpartum depression and current mental health problems of the mother on child behaviour at eight years. Maternal and Child Health Journal. 2017 Jul;21(7):1563-72

24 Prady SL, Pickett KE, Croudace T, Mason D, Petherick ES, McEachan RRC, et al. Maternal psychological distress in primary care and association with child behavioural outcomes at age three. European Child & Adolescent Psychiatry. 2016;25(6):601-13

[25] Leis JA, Heron J, Stuart EA, Mendelson T. Associations between maternal mental health and child emotional and behavioral problems: does prenatal mental health matter? Journal of Abnormal Child Psychology. [Research Support, Non-U.S. Gov't]. 2014;42(1):161-71

[26] O'Connor E, Rossom R, Henniger M,Groom H,Burda B Primary care screening for treatment of depression in pregnant and postpartum women. Evidence report and systermatic review for the US Preventative Services Task Force 2016 JAMA315(4):388-406

[27] Owora AH, Helene C, Reese J, Garwe T. Diagnostic performance of major depression disorder case-finding instruments used among mothers of young children in the United States: A systematic review. Journal of Affective Disorders. [Review]. 2016;201:185-93.

[28] Vlenterie R, van Ras HWP, Roeleveld N, Pop-Purceleanu M, van Gelder MMHJ. Epidemiological evaluation of the Patient Health Questionnaire-2 in a pregnant population. Journal of Psychosomatic Research. 2017;101:96-103

[29] Meades R, Ayers S. Anxiety measures validated in perinatal populations: a systematic review. Journal of Affective Disorders. 2011;133(1-2):1-15.

[30] Howard L, Ryan E, Trevillion K, Anderson F, Bick D, Bye A, Byford S, O'connor S, Sands P, Demiliw J, Milgrom J, Pickles A Edinburgh Postnatal Depression Scale in identifying depression and other mental disorders in early pregnancy. British Journal of Psychiatry 2018 212,50-56

31 Hewitt C, Gilbody S, Brealey S, Paulden M, Palmer S, Mann R et al. Methods to identify postnatal depression in primary care: an integrated evidence synthesis and value of information analysis. Health Technology Assessment 13[36], 1-230. 2009. England., NLM.

32 Gollan JK, Wisniewski SR, Luther JF, Eng HF, Dills JL, Sit D, et al. Generating an efficient version of the Edinburgh Postnatal Depression Scale in an urban obstetrical population. Journal of Affective Disorders. 2017 Jan;208:615-20.

33 Grote NK, Katon WJ, Russo JE, Lohr MJ, Curran M, Galvin E, et al. A Randomized Trial of Collaborative Care for Perinatal Depression in Socioeconomically Disadvantaged Women: The Impact of Comorbid Posttraumatic Stress Disorder. Journal of Clinical Psychiatry. [Comparative StudyRandomized Controlled Trial]. 2016;77(11):1527-37.

34 Weinreb L, Wenz-Gross M, Upshur C. Postpartum outcomes of a pilot prenatal care-based psychosocial intervention for PTSD during pregnancy. Archives of Women's Mental Health. 2017 07 Nov:1-14

35 van Ravesteyn LM, Lambregtse-van den Berg MP, Hoogendijk WJG, Kamperman AM. Interventions to treat mental disorders during pregnancy: A systematic review and multiple treatment meta-analysis. PLoS ONE [Electronic Resource]. [Meta-Analysis]. 2017;12(3):e0173397

36 Letourneau NL, Dennis CL, Cosic N, Linder J. The effect of perinatal depression treatment for mothers on parenting and child development: A systematic review. Depression and Anxiety. 2017 October;34(10):928-66.

37 McDonagh MS, Matthews A, Phillipi C, Romm J, Peterson K, Thakurta S, et al. Depression drug treatment outcomes in pregnancy and the postpartum period: A systematic review and meta-analysis. Obstetrics and Gynecology. 2014;124(3):526-34.

38 Prady SL, Hanlon I, Fraser LK, Mikocka-Walus A. A systematic review of maternal antidepressant use in pregnancy and short- and long-term offspring's outcomes. Archives of Women's Mental Health. 2017 12 Oct:1-14

39 Harris RP, Helfand M, Woolf SH, et al (2001) Methods Work Group, Third US Preventative Task Force; a review of the process Am J Prev Med 2001;20(3)(suppl):21-35

40 Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P (2010). The Newcastle–Ottawa Scale(NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa, Ontario: Ottawa Health Research Institute.

41 Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. The BMJ. 2017;358:j4008. doi:10.1136/bmj.j4008.

42 Higgins Julian P T, Altman Douglas G, Gøtzsche Peter C, Jüni Peter, Moher David, Oxman Andrew D et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials BMJ 2011; 343 :d5928

43 McDonagh MS, Jonas DE, Gartlehner G, Little A, Petersen K, Carson S et al Methods for the drug effectiveness review project BMC Med Res Methodol 2012 ;12:140

44 Berkman N, Lob RK, Ansari M, McDonagh M, Balk E Whitlock E et al Grading the strength of a body of evidence when assessing health care interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: an update AHRQ Publication number 13(14)-EHC130-EF. Rockville(MD); Agency of Healthcare Research and Quality 2013.

45 Gaynes B, Gavin N, Meltzer-Brody S, Lohr K, Swinson T, Gartlehner G et al Perinatal depression prevalence, screening accuracy and screening outcomes. Evid Rep Technol Assess (sum) 2005;(119):1-8

[46] Paulden M, Palmer S, Hewitt C, Gilbody S. Screening for postnatal depression in primary care: cost effectiveness analysis. BMJ 2009; 339(dec22_1):b5203

[47] Morrell CJ, Sutcliffe P, Booth A, Stevens J, Scope A, Stevenson M, et al. A systematic review, evidence synthesis and meta-analysis of quantitative and qualitative studies evaluating the clinical effectiveness, the cost-effectiveness, safety and acceptability of interventions to prevent postnatal depression. Health Technology Assessment (Winchester, England). 2016;20(37):1-414

[48] Pritchett RV, Daley AJ, Jolly K. Does aerobic exercise reduce postpartum depressive symptoms? a systematic review and meta-analysis. British Journal of General Practice. 2017;67(663):e684-e91

[49] Redshaw M, Henderson J. Who is actually asked about their mental health in pregnancy and the postnatal period? Findings from a national survey. BMC Psychiatry. 2016;16(1):322

[50] Prady SL, Pickett KE, Petherick ES, Gilbody S, Croudace T, Mason D, et al. Evaluation of ethnic disparities in detection of depression and anxiety in primary care during the maternal period: combined analysis of routine and cohort data. British Journal of Psychiatry. 2016;208(5):453-61

[51] Darwin Z, McGowan L, Edozien LC. Antenatal mental health referrals: review of local clinical practice and pregnant women's experiences in England. Midwifery. [Research Support, Non-U.S. Gov't]. 2015;31(3):e17-22.

[52] Atkins S, Lewin S, Smith H, et al. Conducting a meta-ethnography of qualitative literature: lessons learnt. BMC Med Res Methodol 2008; 8: 21.