An evaluation of screening for postnatal depression against NSC criteria

Dr Christine Hill
February 2010 (updated Sept10)

SECTION 1

Introduction

Postnatal depression

Postnatal depression (PND) is an important category of depression, with a prevalence within the first few postnatal months of 13% .[1] Most cases develop within the first three months,[2;3] with a peak incidence of about 4-6 weeks.[3;4] PND has a substantial impact on the mother, her partner[5], her family[6], mother-baby interactions[7] and the longer term emotional and cognitive development of the baby, especially when depression occurs in the first year of life. [8] [9;10] Although clinical and cost-effective treatments are available, less than half of women with PND are detected in routine clinical practice.[9-11]

Five strategies have been identified for case finding or screening for PND:[3]

- Postnatal identification using specially developed standardised postnatal questionnaires such as the Edinburgh Postnatal Depression Scale (EPDS);
- Postnatal identification using standardised generic questionnaires for depression e.g. the Beck Depression Inventory (BDI);
- Prenatal screening using standardised depression questionnaires to identify those with pre-existing depression and those at risk of developing significant depression postnatally.
- Prenatal screening using known risk factors for PND (such as previous history of depression and lack of social support) to identify those who are likely to develop depression postnatally.
- The use of training packages targeted at healthcare professionals designed to enhance awareness and recognition of clinical signs of PND, and to ensure that a thorough psychosocial assessment is provided on a routine basis.[3]

Screening for postnatal depression is not currently recommended by the UK National Screening Committee (NSC).

Previous reviews of screening for postnatal depression

The NSC last reviewed screening for PND in 2001.[10] The conclusion of the NSC was that there was insufficient clinical and economic evidence to support the implementation of screening strategies for PND.[9;10] The NSC also expressed concern at the widespread use of the EPDS as a screening tool.

However national guidance has been inconsistent, in 1999, the National Service Framework (NSF) for mental health, made it a requirement for all local areas to have protocols for the management of PND[12]. This resulted in case finding strategies focusing on the administration of a screening tool in the postnatal period; EPDS is the most widely used and is normally administered by a health visitor in the postnatal period.

The use of ad hoc screening in this way, on a national basis, has attracted much criticism, mainly concerning the screening tools, and screening in the absence of any evidence that it leads to either effective management of depression or improved outcomes.

Recent evidence

Since the 2001 review by the NSC, there have been a number of studies published including systematic reviews of screening/case finding strategies for PND, national guidance by NICE and a health technology assessment (HTA). The conclusions and recommendations of these studies are discussed below.

2010 July

Universal prevention of depression in women postnatally: cluster randomized trial evidence in primary care.[13]

This study is a randomised cluster trial in 101 GP practices. The practices were randomised into care as usual (CAU) and health visitor (HV) intervention. The study recruited women in the antenatal period and then used the EPDS to identify those women who were not depressed 6 weeks postnatally by excluding women who scored 12 or more at 6 weeks on a postal self-administered EPDS.(Hewitt[14] says 12 is an optimal cutoff for major depression).

The women had care from trained health visitors (HV) in the following 6-18 months postnatally and were sent postal EPDS by the research office at 6, 12 and 18 months postnatally. The HVs in the intervention group were trained in trained in identifying depressive symptoms using the EPDS and face to face clinical assessment and in providing 'psychologically orientated sessions based on cognitive behavioural principles'. The primary outcome measure was the proportion of women scoring ≥ 12 on the EPDS at 6 months postnatally. At 6 months, there was no difference between the two groups for those women who had an EPDS of 5 or less. At 6 months, there was a significant difference between groups for women who had scored between 5 and 12 (subthreshold group) in that 17.2% had become depressed (i.e. had an EPDS of 12 or more) while 12.5% of the intervention group had become depressed (EPDS of 12 or more). OR 0.68 (95%CI 0.48-0.97). The data was adjusted for living alone and history of PND and life events. The authors conclude that what seems to matter is the difference between being registered with a practice that has adopted the HV training and one that is not, as this affects the practitioners themselves through increased confidence, who then focus on the mothers well being rather than just the child's.

There are a number of problems with this paper, such as the primary outcomes in the protocol, *"changes in symptoms, health outcomes, hospital admissions, NHS service use, family well-being and infant progress to eighteen months"* are quite different from the primary outcome in the paper which was the *"proportion of women with a six week EPDS >= 12, who had a six month EPDS score >= 12."* There was a higher loss to follow-up in the treatment group 133/404 (33%) compared with controls 44/191 (23%) which is unlikely to have occurred by chance. If more severely affected participants were more likely to be lost to follow-up, this would bias the results in favour of the treatment. And the only outcomes reported, apart from self-reported anti-depressant prescription, were questionnaire scores. Thus suicides, suicide attempts, and hospital admissions etc. would be reported elsewhere.[15]

This study does not add to the evidence that the EPDS is a good test for depression.

2009 December.

A cost-effectiveness analysis of routine screening for PND in primary care undertaken by Paulden and colleagues at the Centre for Health Economics, the University of York. [9]

The objective of this study was to evaluate the cost-effectiveness of routine screening for PND in primary care.[9] Some of the authors of this paper were authors of the HTA discussed below. This publication is intended as a summary of the cost-effectiveness analysis from the HTA, and a summary of the policy implications.

The authors note that there is a lack of reliable evidence from RCTs of the costs of the identification and treatment of PND. This includes the cost of administering the test, the cost of any subsequent treatment, and the cost associated with an incorrect diagnosis. Thus the authors established estimates of costs from published data or from expert opinion.

Using a conservative approach in their base case analysis, and comparing with routine care, the most cost-effective formal identification method (EDPS with a cut-off 16) showed an incremental cost-effectiveness ratio (ICER) of £41,103 per QALY, which is well above the conventional threshold of willingness to pay of £20,000 to £30,000 per QALY.[9]

Even when the authors used a very optimistic view of the cost of a false positive i.e. a single visit to a GP who would then make the correct diagnosis, compared with routine care the most cost-effective formal identification method (EDPS with a cut-off 10), had an ICER just below the upper limit of £30,000.

The authors' conclusion was that the use of formal identification methods for PND did not seem to represent value for money. They note that these conclusions are primarily driven by the costs of managing false positives i.e. women who are diagnosed as being depressed but who do not subsequently turn out to have PND. When the cost of treating a false positive is high, the specificity of the screening tool is an important contributor to its cost-effectiveness.

The study also found that adopting a structured interview as a confirmatory test for those with a positive test, proved to be cost saving compared with the equivalent strategy without an interview, however it was not proved to be cost-effective compared with routine care only.

The authors also suggest that the NICE guidance published in 2007 recommending the use of the Whooley questions (Appendix A), and the practice of the routine or ad hoc administration of the EPDS, did not result in value for money. The authors conclude that there is no current method of screening for PND that satisfies the NSC's criteria for the adoption of a screening strategy because none represent a favourable cost-benefit ratio. [9]

2009 July

A health technology assessment was commissioned in 2006 on PND screening and was published in 2009, undertaken by Hewitt and colleagues of the Department of Health Sciences, University of York.[14].

The aim of the HTA was to provide an overview of all available methods to identify PND and to assess their validity; to assess the acceptability of methods to identify PND; to assess the clinical and cost-effectiveness of methods to identify PND in improving maternal and infant outcomes; to identify research priorities and the value of further research into

methods to identify PND, and to assess whether methods to identify PND meet the minimum criteria outlined by the NSC.[14]

A total of 14 identification strategies for depression were found to have been validated among women during pregnancy or the postnatal period. PND screening/case finding strategies were the EPDS, Postpartum Depression Screening Scale (PDSS), Pregnancy Risk Questionnaire, and Predictive Index. The most frequently used strategy across all reviews was the EPDS.

The choice of cut-off point on the test is the key factor in indicating whether a woman has a positive test or a negative test. Variations on the choice of cut-off will lead to variations in measures of the accuracy of the test i.e. sensitivity and specificity. Using data from studies of the EPDS using multiple cut-off points (from 7-16):[14]

- Sensitivity ranges from 0.6 – 0.96 for major depression only; and specificity from 0.97 – 0.45 for major depression only.
- Sensitivity ranges from 0.31 – 0.91 for major or minor depression; and specificity from 0.99 – 0.67 for major and minor depression.
- Some women will be wrongly identified as being depressed and some depressed women will not be identified.

The authors showed that optimal cut-off points were 12 for major depression only, 10 for major and minor depression and 9 for any psychiatric disorder. To maximise sensitivity from a clinical perspective then they found that a cut-off of 7 for major depression, 8 for major and minor depression and 9 for any psychiatric disorder should be used. To maximise cost-effectiveness, an EPDS of greater than 10 should be used.

The EPDS was found to be a simple, safe, precise and validated screening test, and a suitable cut-off level could be defined. But while it was found to be acceptable to the population, the evidence surrounding the clinical and cost-effectiveness of screening with the EPDS was lacking.

None of the other case finding strategies could be assessed against NSC criteria as there was insufficient evidence.

Further, no evidence could be found across four systematic reviews of the validity, acceptability, clinical effectiveness and cost-effectiveness, for the Whooley questions (Appendix A) to be used in a postnatal population, as recommended in the NICE guidance[16].

The HTA specifically addressed the question of whether the identification of PND should be implemented as a national screening policy according to NSC criteria. The HTA conclusion was that the criteria for a PND screening programme using the EPDS were not currently met. (Their comments have been tabulated against NSC criteria in Section 2.)

2009 May

A systematic review of the clinical and cost-effectiveness of antenatal and postnatal identification of depressive symptoms by Hewitt and Gilbody, the Department of Health Sciences , University of York .[3]

This paper by two of the authors of the HTA, aimed to explore whether the routine use of screening/case finding of antenatal and postnatal depressive symptoms, or the integration of these methods with enhanced care, is clinically, and cost-effective. [3]

Five studies were identified that compared formal use of a method to identify PND, with or without enhancement of care, with not using a formal method, or usual care. All of the studies used the EPDS to identify women with PND.

One of the issues the authors encountered was the cut-off point on the EPDS scale used to distinguish those women who were deemed to have a positive test and thus were at higher risk of PND, and those who were not. Variations in the choice of cut-off will lead to variations in measures of the accuracy of the test i.e. sensitivity and specificity. The authors noted that this created difficulties within the review. [3]

The NSC criteria require a screening test to be precise, validated, and to have a suitable cut-off level that has been defined and agreed; and it must accurately measure risk. The authors had concerns regarding the validity of this case finding method and felt that there was insufficient evidence available to conclude that the EPDS was effective in improving maternal and infant outcomes. However they found some evidence that indicated that the EPDS, with some enhancement of care, may lead to reductions the number of women with EPDS scores above a certain threshold or reduction in EPDS scores, but the criterion for reducing morbidity or mortality would not be met. [3]

The authors concluded that the evidence surrounding the clinical and cost-effectiveness of PND screening is lacking and further research is required to address this gap.[3]

2007

NICE published service guidance, and guidance on the clinical management, of Antenatal and Postnatal Mental Health[16]

Clinical guidance on the management of antenatal and postnatal mental health care was issued which included a review of methods to identify depression during the postnatal period. The review considered two methods of identification, the EPDS and case finding questions (the Whooley questions plus help question).

The guidance recommended the use of the Whooley questions to identify possible PND, with the use of self report measures such as the EDPS, the hospital anxiety and depression scale (HADS) or the patient health questionnaire (PHQ-9) as part of subsequent assessment, or for routine monitoring [9;16]

The authors point out that the Whooley questions have been validated in older men but not in postnatal women [9;16].

The NICE guidelines did not consider the cost-effectiveness of using these screening/case finding strategies. [9]

2005

An evidence report/technology assessment by the Agency for Healthcare Research and Quality on Perinatal Depression: Prevalence, Screening Accuracy, and Screening Outcomes. RTI-University of North Carolina Evidence-based Practice Center.[17]

This was a systematic review of the evidence for the prevalence and incidence of PND, the accuracy of different screening instruments, and the effectiveness of interventions for women screened as high risk for PND.

The authors were surprised there was a such as paucity of evidence in this area considering the significance of PND as a mental health and public health problem, and commented that large scale studies are needed. They note that the small number and small size of relevant studies are not adequate to guide national policy. [17] They found that there was no good evidence on whether PND rates differ among various ethnic groups. This information is necessary in order to know where to target screening programs and for researchers to know where to target studies. The authors also note that it needs to be clarified if the incidence of PND is greater than the incidence of depression in non-childbearing women of similar ages.[17]

It was also noted that future research should continue to assess and directly compare multiple screening instruments, and to evaluate which screening instrument is more accurate in the postnatal setting. Studies are also needed to evaluate the cost-effectiveness of screening, specifically the costs of false negative and false positives, the degree of provider burden, and patient acceptability of the screening strategy. These will provide insights on how to consider target sensitivity and specificity when attempting to maximise cost-effectiveness.

The authors also make an important point that studies should carefully consider whether to target major depression alone, for which beneficial treatments exist, or a combined category of major and minor depression, a heterogeneous group for which treatment benefit is unclear. They suggest that available screening tools identify major depression alone more accurately, and as there appears to be greater benefit from interventions for major depression alone, this they feel suggests that targeting major depression alone in screening/case finding would be a preferable strategy.

Most studies identified for this review were conducted in the first 3 months postpartum but the authors suggest that the depression may remain high for several months so more studies are needed to include 6 weeks, 3 months, 6 months, and 12 months. They also recommend that researchers should consider developing and using standard screening measures, using similar cut-off points, so that studies can be compared. They note that the evidence base was quite limited but from their review of screening instruments it would appear that for major depression alone, an EPDS cut-off of >13 or a PDSS cut-off of >81 were reasonably supported by the evidence as thresholds to use.

SECTION 2

NSC criteria for screening programmes

	YES	NO	INCONCLUSIVE EVIDENCE
THE CONDITION			
1. The condition should be an important health problem	PND is an important category of depression with over 11% of women experiencing major or minor PND six weeks postnatally.[9;10] There is now considerable evidence to show that PND has a substantial impact on the mother, her partner[5], her family[6], mother-baby interactions[7] and the longer term emotional and cognitive development of the baby, especially when depression occurs in the first year of life.[8-10] Though clinical and cost-effective treatments are available, less than half of cases of PND are detected in routine clinical practice.[9-11]		
2. The epidemiology and natural history of the condition, including development from latent to declared disease, should be adequately understood and there should be a detectable risk factor, disease marker or early symptomatic stage	The precise level of the prevalence and incidence of PND is uncertain. Published estimates of the rate of major and minor depression in the postpartum period range widely from 5 percent to more than 25 percent of new mothers, depending on the assessment method, the timing of the assessment, and population characteristics.[17] A meta analysis of 59 studies (12810 women mainly from first world countries) found the prevalence of depression within the first few postnatal months to be 13% (95%CI 12.3-13.4%).[1] Most cases develop within the first three postnatal months[2;3] with a peak incidence of about 4-6 weeks[3;4]. One study showed that most cases last around 3 months and resolve spontaneously without treatment[2;3], another study demonstrated that 50% have depression lasting over 6 months and some still being present after 4 years[3;18]. Postpartum psychosis can occur in the postnatal period but unlike PND, postpartum psychosis is a relatively rare event with a range of estimated incidence of 1.1 to 4.0 cases per	PND encompasses major and minor depressive episodes that occur either during pregnancy or within the first 12 months following delivery. When referring to depression in this population, researchers and clinicians frequently have not been clear about whether they are referring to major depression alone or to both major and minor depression. Major depression is a distinct clinical syndrome for which treatment is clearly indicated, whereas the definition and management of minor depression is less clear.[17]	The HTA stated that it would be informative to identify the natural history of PND over time and to identify the clinical effectiveness of the most valid and acceptable method to identify PND.[14]

	1,000 deliveries. The onset of postpartum psychosis is usually acute, within the first 2 weeks of delivery, and appears to be more common in women with a family history of bipolar or schizoaffective disorder. [17]		
3. All the cost-effective primary prevention interventions should have been implemented as far as practicable.	Not reviewed.		
4. If the carriers of a mutation are identified as a result of screening, the natural history of people with this status should be understood, including the psychological implications.	Not applicable.		

THE TEST			
5. There should be a simple, safe, precise and validated screening test.	Paulden et al [9] identified numerous generic and specific screening strategies for PND, 14 of which had been validated among women during pregnancy or the postnatal period. The most commonly used identification method was the EPDS.[9] The HTA identified the EPDS as the only instrument for which sufficient data were available to combine studies at multiple cut-off points. [14] In the HTA, the EPDS was found to have a sensitivity that ranged from 0.60 (specificity 0.97) to 0.96 (specificity 0.45) for major depression only; from 0.31 (specificity 0.99) to 0.91 (specificity 0.67) for major or minor depression; and from 0.38 (specificity 0.99) to 0.86 (specificity 0.87) for any psychiatric disorder. Although the EPDS was shown to have a reasonable sensitivity and specificity, some women with PND will be unidentified and some women without PND will be wrongly identified as having PND. [14] ***The HTA conclusion was that criterion to be met for the EPDS but note a lack of evidence for the other potential identification strategies identified.***[14]	Formal strategies for screening and case identification (such as standardised postnatal questionnaires, standardised generic questionnaires for depression, and prenatal screening for known risk factors for PND have been advocated but are controversial. [9] Screening and case finding instruments are imperfect diagnostic instruments and might identify some women who are depressed and not picked up on routine care, but might incorrectly identify some women who are not depressed. [9] The cut-off point chosen to distinguish between a positive and a negative test is critical. Screening for PND using these methods generates a score on a particular scale, when a patient scores at or above a cut-off then clinically significant depression is assumed to be present. [9] The Whooley questions. NICE guidance recommends that healthcare professionals ask 2 questions at a woman's first contact with primary care, again at her booking visit, and again postnatally (usually at 4-6 weeks and 3-4 months). A third question should be considered if the woman answers 'yes' to the either of the initial questions.[9;16] The HTA found no evidence to support their use in screening for PND.	

6. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed.	A wide variety of cut points for the EPDS have been reported. The original validation study of the EPDS recommended a cut-off point of 10 for 'possible depression' and 13 for 'probable depression'.[19] The lower the cut point used to distinguish between cases and non-cases the higher the sensitivity becomes. Increasing the sensitivity will lead to fewer women with PND being unidentified but a lower specificity leading to an increase in the number of women wrongly diagnosed with PND. The HTA reported that the optimal cut point, in terms of the trade-off between sensitivity and specificity, was 12 for major depression only, 10 for major or minor depression and 9 for any psychiatric disorder. If the cut point was chosen to maximise sensitivity then from this analysis the optimal cut point was 7 for major depression only, 8 for major or minor depression and 9 for any psychiatric disorder. The results suggested that in the scenarios considered the most cost-effective identification approach would be the EPDS at a cut point of 10 or higher. ***The HTA conclusion was that the criterion had been met in principle for the EPDS. There was a lack of evidence for other potential identification strategies identified.***[14]		
7. The test should be acceptable to the population.	The HTA reported that they had identified 16 studies that explored the acceptability of screening strategies for PND and had found that the most frequently explored .They found that overall, the majority of studies indicated that the EPDS was acceptable when undertaken in the home, with training for the health visitor and careful management of positive responses to the question on self harm. [14] ***The HTA conclusion was that this criterion had been met for the EPDS. There was a lack of evidence for other potential identification strategies identified.***[14]		

8. There should be an agreed policy on the further diagnostic investigation of individuals with a positive test result and the choices available to those individuals.		Not explored in these systematic reviews.	Paulden and colleagues[9] found that adopting a structured interview as a confirmatory test for those with a positive test proved to be cost saving compared with the equivalent strategy without an interview, however no such strategy subsequently proved to be cost-effective compared with routine care only. This they suggest needs further research to find alternative approaches to management of those women who have positive tests.[9]
9. If the test is for mutations the criteria used to select the subset of mutations to be covered by screening, if all possible mutations are not being tested, should be clearly set out.	Not applicable		

THE TREATMENT			
10. There should be an effective treatment or intervention for patients identified through early detection, with evidence of early treatment leading to better outcomes than late treatment.	Not explored in these systematic reviews.		Gaynes and colleagues[17] make the point that studies should carefully consider whether to target major depression alone, for which beneficial treatments exist, or a combined category of major and minor depression, a heterogeneous group for which treatment benefit is unclear. They suggest that available screening tools identify major depression alone more accurately, and as there appears to be greater benefit from interventions for major depression alone, this recommends targeting major depression alone in screening/case finding tests.
11. There should be agreed evidence based policies covering which individuals should be offered treatment and the appropriate treatment to be offered.	Not explored in these systematic reviews.		
12. Clinical management of the condition and patient outcomes should be optimised in all health care providers prior to participation in a screening programme.	Not explored in these systematic reviews.		

THE SCREENING PROGRAMME			
13. There should be evidence from high quality Randomised Controlled Trials that the screening programme is effective in reducing mortality and morbidity. Where screening is aimed solely at providing information to allow the person being screened to make an 'informed choice' (e.g. Down's syndrome or cystic fibrosis screening) there must be evidence from high quality trials that the test accurately measures risk. The information that is provided about the test and its outcome must be of value and readily understood by the individual being screened.		Gaynes and colleagues stated that there was a paucity of evidence in this area considering the significance of PND as a mental health and public health problem, and commented that large scale studies are needed. They note that the small number and small size of relevant studies are not adequate to guide national policy.[17] The HTA demonstrated in their systematic review that insufficient evidence is available to conclude that identification strategies are effective in improving maternal and infant outcomes. Some suggestive evidence indicated 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 HTA conclusion was that this criterion had not been met[14]***	
14. There should be evidence that the complete screening programme (test, diagnostic procedures, treatment/intervention) is clinically, socially and ethically acceptable to health professionals and the public.	The majority of studies found that EPDS was acceptable when undertaken in the home, providing that the health visitor had undergone training and if the woman was forewarned of the test.[14]		
15. The benefit from the screening programme should outweigh the physical and psychological harm (caused by the test, diagnostic procedures and treatment)	Not examined in these reviews		
16. The opportunity cost of the screening programme (including testing, diagnosis and treatment, administration, training and quality assurance) should be economically balanced in relation to expenditure on medical care as a whole (i.e. value for money)		The HTA could not find evidence of any full economic evaluations of PND identification strategies in their systematic review. In the absence of cost-effectiveness studies, the authors developed a decision-analytic model. The results of the base-case analysis suggested that the use of formal identification strategies do not appear to represent value for money based on conventional thresholds of cost-effectiveness used in the NHS. The scenarios considered in the decision-analytic model demonstrated that this	

		conclusion was primarily driven by the costs of false positives. Alternative assumptions resulted in more favourable estimates of cost-effectiveness. When the cost of a false positive diagnosis was assumed to be the cost of a single GP attendance who would then make the correct diagnosis, the EPDS using a cut point of 10 was found to be the most cost-effective strategy. A definitive answer to the question as to whether formal identification strategies are cost-effective, and, if they are, which individual strategy is optimal in cost-effectiveness terms, requires further more reliable evidence. [14] ***The HTA conclusion was that the criterion had not been not met.***[14]	
17. There should be a plan for managing and monitoring the screening programme and an agreed set of quality assurance standards.	Not examined in these reviews.		
18. Adequate staffing and facilities for testing, diagnosis, treatment and programme management should be available prior to the commencement of the screening programme	Not examined in these reviews.		
19. All other options for managing the condition should have been considered (e.g. improving treatment, providing other services), to ensure that no more cost-effective intervention could be introduced or current interventions increased within the resources available.	Not examined in these reviews.		
20. Evidence based information, explaining the consequences of testing, investigation and treatment, should be made available to potential participants to assist them in making an informed decision.	Not examined in these reviews.		
21. Public pressure for widening the eligibility criteria, for reducing the screening interval, and for increasing the sensitivity of the testing process, should be anticipated. Decisions about these parameters should be scientifically justifiable to the public.	Not examined in these reviews.		

22. If screening is for a mutation the programme should be acceptable to people identified as carriers and to other family members.	Not applicable.		

Summary

Screening for PND fails to fulfill all of the criteria set by the NSC for the implementation of a population screening programme.

1. The EPDS is the most frequently used instrument to identify PND, however the choice of cut-off point to indicate whether a person has a positive test, is key as to whether a test is deemed positive or negative. Variations on the choice of cut-off will lead to variations in measures of the accuracy of the test i.e. sensitivity and specificity.

   Using data from studies using multiple cut-off points (from 7-16):[14]

   - Sensitivity ranges from 0.6 – 0.96 for major depression only; and specificity from 0.97 – 0.45 for major depression only.
   - Sensitivity ranges from 0.31 – 0.91 for major or minor depression; and specificity from 0.99 – 0.67 for major and minor depression.
   - Some women will be wrongly identified as being depressed and some depressed women will not be identified.

   Hewitt and colleagues[14] in their systematic review showed an optimal cut-off point was 12 for major depression only, 10 for major and minor depression and 9 for any psychiatric disorder. To maximise sensitivity from a clinical perspective, a cut-off of 7 for major depression, 8 for major and minor depression and 9 for any psychiatric disorder is recommended. To maximise cost-effectiveness then an EPDS of greater than 10 should be used.

   Gaynes and colleagues [17] suggest that for major depression alone, an EPDS cut-off of >13 is reasonably supported by the evidence as the optimal threshold to use.

   The original validation study of the EPDS recommended a cut-off point of 10 for 'possible depression' and 13 for 'probable depression'. [19]

2. The majority of studies found that the EPDS was acceptable when undertaken in the home, provided that the health visitor had received training and if the woman was forewarned of the test.[14]

3. Hewitt and colleagues[14] in the HTA found that there was insufficient evidence to conclude that any case finding or screening strategy was effective in improving maternal and infant outcomes, although there was some evidence to indicate that EPDS might, with enhanced care, lead to reductions in the number of women with scores above a certain EPDS threshold. [14]

4. Hewitt and colleagues[14] in the HTA developed a decision-analytic model and the results of the analysis suggested that the use of formal identification strategies did not appear to represent a value for money based on conventional thresholds of cost-effectiveness in the NHS, and that costs are driven by the cost of treating

false positives. They recommend more research into the costs of managing false positives. This finding was confirmed in a subsequent paper by the same investigators[9] using decision modeling.

5. One of the findings of the HTA was that there was no evidence to support the use of the Whooley questions as recommended by NICE guidance. [14]

6. Paulden and colleagues[9] found that adopting a structured interview as a confirmatory test for those with a positively test proved to be cost saving compared with the equivalent strategy without an interview, however no such strategy subsequently proved to be cost-effective compared with routine care only. This they suggest needs further research to find alternative approaches to management of those women who have positive tests.

Implications for further research

The above systematic reviews have identified the following areas requiring further research:

1. To identify the optimal strategy for screening/case finding of PND in terms of key psychometric properties for postnatal populations. Further research is also required to compare the performance of the Whooley questions with the EPDS, and a generic depression measure.
2. To establish the acceptability of the different PND screening/case finding strategies.
3. To understand the natural history of PND over time in (different ethnic) populations where formal methods to identify PND have been used, and where they have not.
4. To identify the costs associated with the management of false positives.
5. To understand the impact of PND on health related quality of life.
6. To add to the current epidemiological data on prevalence of PND.
7. To gather evidence from further research within an RCT on clinical effectiveness of screening/case finding for PND.

Reference List

(1) Swain AM, O'Hara MW, Starr KR, Gorman LL. A prospective study of sleep, mood, and cognitive function in postpartum and nonpostpartum women. *Obstet Gynecol* 1997; 90(3):381-386.

(2) Cooper PJ, Campbell EA, Day A, Kennerley H, Bond A. Non-psychotic psychiatric disorder after childbirth. A prospective study of prevalence, incidence, course and nature. *Br J Psychiatry* 1988; 152:799-806.

(3) Hewitt CE, Gilbody SM. Is it clinically and cost effective to screen for postnatal depression: a systematic review of controlled clinical trials and economic evidence. *BJOG : an international journal of obstetrics and gynaecology* 2009; 116(8):1019-1027.

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

(5) Boath EH, Pryce AJ, Cox JL. Postnatal depression: The impact on the family. *Journal of Reproductive and Infant Psychology* 1998; 16(2):199-203.

(6) Lovestone S, Kumar R. Postnatal psychiatric illness: the impact on partners. *Br J Psychiatry* 1993; 163:210-216.

(7) Murray L, Fiori-Cowley A, Hooper R, Cooper P. The impact of postnatal depression and associated adversity on early mother-infant interactions and later infant outcome. *Child Dev* 1996; 67(5):2512-2526.

(8) Murray L, Sinclair D, Cooper P, Ducournau P, Turner P, Stein A. The socioemotional development of 5-year-old children of postnatally depressed mothers. *J Child Psychol Psychiatry* 1999; 40(8):1259-1271.

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

(10) Shakespeare J, National Screening Committee. Evaluation of screening for postnatal depression against NSC handbook criteria. 1-8-2001.
Ref Type: Report

(11) Hearn G, Iliff A, Jones I, Kirby A, Ormiston P, Parr P et al. Postnatal depression in the community. *Br J Gen Pract* 1998; 48(428):1064-1066.

(12) Department of Health. National Service Framework for Mental Health. London HMSO, editor. 1999.
Ref Type: Report

(13) Brugha TS, Morrell CJ, Slade P, Walters SJ. Universal prevention of depression in women postnatally: cluster randomized trial evidence in primary care. *Psychological Medicine* 2010; FirstView:1-10.

(14) 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.
Ref Type: Report

(15) Thornton J. Trial reporting problems - rapid response to The PoNDER Trial. *BMJ* 2009.

(16) NICE. Antenatal and postnatal mental health: clinical management and service guidance. 2007.
Ref Type: Report

(17) Gaynes BN, Gavin N, Meltzer-Brody S, Lohr KN, Swinson T, Gartlehner G et al. Perinatal depression: prevalence, screening accuracy, and screening outcomes. *Evid Rep Technol Assess (Summ)* 2005;(119):1-8.

(18) Kumar R, Robson KM. A prospective study of emotional disorders in childbearing women. *Br J Psychiatry* 1984; 144:35-47.

(19) Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry* 1987; 150:782-786.

Appendix A

The Whooley questions[16]

Clinical guidance issued by NICE in 2007 recommends that healthcare professionals ask two questions at a woman's first contact with primary care, again at her booking visit, and again postnatally (usually at 4-6 weeks and 3-4 months):

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 little interest or pleasure in doing things?

A third question should be considered if the woman answers "yes" to either of the initial questions:

3. Is this something you feel you need or want help with?