Identifying perinatal depression with case-finding instruments: a mixed-methods study (BaBY PaNDA – Born and Bred in Yorkshire PeriNatal Depression Diagnostic Accuracy)

Elizabeth Littlewood, Shehzad Ali, Lisa Dyson, Ada Keding, Pat Ansell, Della Bailey, Debrah Bates, Catherine Baxter, Jules Beresford-Dent, Arabella Clarke, Samantha Gascoyne, Carol Gray, Lisa Hackney, Catherine Hewitt, Dorothy Hutchinson, Laura Jefferson, Rachel Mann, David Marshall, Dean McMillan, Alice North, Sarah Nutbrown, Emily Peckham, Jodi Pervin, Zoe Richardson, Kelly Swan, Holly Taylor, Bev Waterhouse, Louise Wills, Rebecca Woodhouse and Simon Gilbody

NHS
National Institute for Health Research

Identifying perinatal depression with case-finding instruments: a mixed-methods study (BaBY PaNDA – Born and Bred in Yorkshire PeriNatal Depression Diagnostic Accuracy)

Elizabeth Littlewood,[1] Shehzad Ali,[1] Lisa Dyson,[2] Ada Keding,[2] Pat Ansell,[3] Della Bailey,[1] Debrah Bates,[4] Catherine Baxter,[5] Jules Beresford-Dent,[5] Arabella Clarke,[2] Samantha Gascoyne,[1] Carol Gray,[4] Lisa Hackney,[5] Catherine Hewitt,[2] Dorothy Hutchinson,[4] Laura Jefferson,[2] Rachel Mann,[6] David Marshall,[5] Dean McMillan,[1,7] Alice North,[8] Sarah Nutbrown,[1] Emily Peckham,[1] Jodi Pervin,[1] Zoe Richardson,[2] Kelly Swan,[3] Holly Taylor,[5] Bev Waterhouse,[3] Louise Wills,[9] Rebecca Woodhouse[1] and Simon Gilbody[1,7]*

[1]Mental Health and Addiction Research Group, Department of Health Sciences, University of York, York, UK
[2]York Trials Unit, Department of Health Sciences, University of York, York, UK
[3]Epidemiology and Cancer Statistics Group, Department of Health Sciences, University of York, York, UK
[4]Northern Lincolnshire and Goole NHS Foundation Trust, Scunthorpe General Hospital, Scunthorpe, UK
[5]Leeds and York Partnership NHS Foundation Trust, Bootham Park Hospital, York, UK
[6]Department of Health Sciences, University of York, York, UK
[7]Hull York Medical School, York, UK
[8]Patient and public involvement representative, York, UK
[9]Harrogate and District NHS Foundation Trust, Harrogate District Hospital, Harrogate, UK

*Corresponding author

This report should be referenced as follows:

Littlewood E, Ali S, Dyson L, Keding A, Ansell P, Bailey D, *et al*. Identifying perinatal depression with case-finding instruments: a mixed-methods study (BaBY PaNDA – Born and Bred in Yorkshire PeriNatal Depression Diagnostic Accuracy). *Health Serv Deliv Res* 2018;**6**(6).

Health Services and Delivery Research

Criteria for inclusion in the *Health Services and Delivery Research* journal
Reports are published in *Health Services and Delivery Research* (HS&DR) if (1) they have resulted from work for the HS&DR programme or programmes which preceded the HS&DR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

HS&DR programme

The Health Services and Delivery Research (HS&DR) programme, part of the National Institute for Health Research (NIHR), was established to fund a broad range of research. It combines the strengths and contributions of two previous NIHR research programmes: the Health Services Research (HSR) programme and the Service Delivery and Organisation (SDO) programme, which were merged in January 2012.

The HS&DR programme aims to produce rigorous and relevant evidence on the quality, access and organisation of health services including costs and outcomes, as well as research on implementation. The programme will enhance the strategic focus on research that matters to the NHS and is keen to support ambitious evaluative research to improve health services.

For more information about the HS&DR programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hsdr

This report

Abstract

Identifying perinatal depression with case-finding instruments: a mixed-methods study (BaBY PaNDA – Born and Bred in Yorkshire PeriNatal Depression Diagnostic Accuracy)

Elizabeth Littlewood,[1] Shehzad Ali,[1] Lisa Dyson,[2] Ada Keding,[2]
Pat Ansell,[3] Della Bailey,[1] Debrah Bates,[4] Catherine Baxter,[5]
Jules Beresford-Dent,[5] Arabella Clarke,[2] Samantha Gascoyne,[1]
Carol Gray,[4] Lisa Hackney,[5] Catherine Hewitt,[2] Dorothy Hutchinson,[4]
Laura Jefferson,[2] Rachel Mann,[6] David Marshall,[5] Dean McMillan,[1,7]
Alice North,[8] Sarah Nutbrown,[1] Emily Peckham,[1] Jodi Pervin,[1]
Zoe Richardson,[2] Kelly Swan,[3] Holly Taylor,[5] Bev Waterhouse,[3]
Louise Wills,[9] Rebecca Woodhouse[1] and Simon Gilbody[1,7]*

[1]Mental Health and Addiction Research Group, Department of Health Sciences,
 University of York, York, UK
[2]York Trials Unit, Department of Health Sciences, University of York, York, UK
[3]Epidemiology and Cancer Statistics Group, Department of Health Sciences, University of York,
 York, UK
[4]Northern Lincolnshire and Goole NHS Foundation Trust, Scunthorpe General Hospital,
 Scunthorpe, UK
[5]Leeds and York Partnership NHS Foundation Trust, Bootham Park Hospital, York, UK
[6]Department of Health Sciences, University of York, York, UK
[7]Hull York Medical School, York, UK
[8]Patient and public involvement representative, York, UK
[9]Harrogate and District NHS Foundation Trust, Harrogate District Hospital, Harrogate, UK

*Corresponding author simon.gilbody@york.ac.uk

Background: Perinatal depression is well recognised as a mental health condition but < 50% of cases are identified in routine practice. A case-finding strategy using the Whooley questions is currently recommended by the National Institute for Health and Care Excellence.

Objectives: To determine the diagnostic accuracy, acceptability and cost-effectiveness of the Whooley questions and the Edinburgh Postnatal Depression Scale (EPDS) to identify perinatal depression.

Design: A prospective diagnostic accuracy cohort study, with concurrent qualitative and economic evaluations.

Setting: Maternity services in England.

Participants: A total of 391 pregnant women.

Main outcome measures: Women completed the Whooley questions, EPDS and a diagnostic reference standard (Clinical Interview Schedule – Revised) during pregnancy (20 weeks) and postnatally (3–4 months). Qualitative interviews were conducted with health professionals (HPs) and a subsample of women.

Results: *Diagnostic accuracy results*: depression prevalence rates were 10.3% during pregnancy and 10.5% postnatally. The Whooley questions and EPDS (cut-off point of ≥ 10) performed reasonably well, with comparable sensitivity [pregnancy: Whooley questions 85.0%, 95% confidence interval (CI) 70.2% to 94.3%; EPDS 82.5%, 95% CI 67.2% to 92.7%; postnatally: Whooley questions 85.7%, 95% CI 69.7% to 95.2%; EPDS 82.9%, 95% CI 66.4% to 93.4%] and specificity (pregnancy: Whooley questions 83.7%, 95% CI 79.4% to 87.4%; EPDS 86.6%, 95% CI 82.5% to 90.0%; postnatally: Whooley questions 80.6%, 95% CI 75.7% to 84.9%; EPDS 87.6%, 95% CI 83.3% to 91.1%). Diagnostic accuracy of the EPDS (cut-off point of ≥ 13) was poor at both time points (pregnancy: sensitivity 45%, 95% CI 29.3% to 61.5%, and specificity 95.7%, 95% CI 93.0% to 97.6%; postnatally: sensitivity 62.9%, 95% CI 44.9% to 78.5%, and specificity 95.7%, 95% CI 92.7% to 97.7%). *Qualitative evaluation*: women and HPs were supportive of screening/case-finding for perinatal depression. The EPDS was preferred to the Whooley questions by women and HPs, mainly because of its 'softer' wording. Whooley question 1 was thought to be less acceptable, largely because of the terms 'depressed' and 'hopeless', leading to women not revealing their depressive symptoms. HPs identified a 'patient-centred' environment that focused on the mother and baby to promote discussion about mental health. *Cost-effectiveness results*: screening/case-finding using the Whooley questions or the EPDS alone was not the most cost-effective strategy. A two-stage strategy, 'Whooley questions followed by the Patient Health Questionnaire' (a measure assessing depression symptomatology), was the most cost-effective strategy in the range between £20,000 and £30,000 per quality-adjusted life-year in both the prenatal and postnatal decision models.

Limitations: Perinatal depression diagnosis was not cross-referenced with women's medical records so the proportion of new cases identified is unknown. The clinical effectiveness and cost-effectiveness of screening/case-finding strategies was not assessed as part of a randomised controlled trial.

Conclusions: The Whooley questions and EPDS had acceptable sensitivity and specificity, but their use in practice might be limited by low predictive value and variation in their acceptability. A two-stage strategy was more cost-effective than single-stage strategies. Neither case-finding instrument met National Screening Committee criteria.

Future work: The yield of screening/case-finding should be established with reference to health-care records. The clinical effectiveness and cost-effectiveness of screening/case-finding for perinatal depression needs to be tested in a randomised controlled trial.

Funding: The National Institute for Health Research Health Services and Delivery Research programme.

Contents

List of tables

List of figures

List of boxes

List of abbreviations

AUC	area under the curve	MCS	mental component score
BaBY	Born and Bred in Yorkshire	MW	midwife
BaBY PaNDA	Born and Bred in Yorkshire PeriNatal Depression Diagnostic Accuracy	NCT	National Childbirth Trust
		NICE	National Institute for Health and Care Excellence
CBT	cognitive–behavioural therapy	NIHR	National Institute for Health Research
CI	confidence interval		
CIS-R	Clinical Interview Schedule – Revised	NLR	negative likelihood ratio
		NPV	negative predictive value
DSM	*Diagnostic and Statistical Manual of Mental Disorders*	NSC	National Screening Committee
		PARS	participant assessment record sheet
EoA	evaluation of acceptability		
EPDS	Edinburgh Postnatal Depression Scale	PCS	physical component score
		PHQ-15	Patient Health Questionnaire-15
EQ-5D	European Quality of Life-5 Dimensions	PHQ-9	Patient Health Questionnaire-9
		PLR	positive likelihood ratio
FN	false negative	PPI	patient and public involvement
FP	false positive	PPV	positive predictive value
FPR	false-positive rate	PSA	probabilistic sensitivity analysis
FSH	facilitated self-help	QALY	quality-adjusted life-year
GAD-2	Generalised Anxiety Disorder-2	R&D	research and development
GAD-7	Generalised Anxiety Disorder-7	ROC	receiver operating characteristic
GP	general practitioner	SCID	Structured Clinical Interview for *Diagnostic and Statistical Manual of Mental Disorders*
HP	health professional		
HRQoL	health-related quality of life		
HSDR	Health Services and Delivery Research	SD	standard deviation
		SF-12v2®	Short-Form questionnaire-12 items version 2
HV	health visitor		
IAPT	Improving Access to Psychological Therapies	STARD	standards for the reporting of diagnostic accuracy studies
		TN	true negative
ICD-10	*International Classification of Diseases,* Tenth Edition	TP	true positive
		WTP	willingness to pay
ICER	incremental cost-effectiveness ratio		
IVF	in vitro fertilisation		

Plain English summary

Depression during pregnancy and early motherhood is a common mental health problem. Less than half of women suffering from depression at this time are identified as depressed by health professionals (HPs) (e.g. midwives, health visitors).

Health professionals use a 10-item questionnaire, the Edinburgh Postnatal Depression Scale (EPDS), to detect symptoms of depression. The National Institute for Health and Care Excellence recommends a shorter questionnaire (the 'Whooley questions'). We compared these two questionnaires to see how well they identify women with depression. We looked at how cost-effective these questionnaires are and conducted in-depth interviews with HPs and women to determine how acceptable these questionnaires are to them.

A total of 391 women participated in the study. Women completed several questionnaires when they were 20 weeks pregnant and 3–4 months after birth. Some of the women took part in longer interviews with researchers.

The study found that women experienced similar rates of depression during pregnancy (10.3%) and the early postnatal period (10.5%). The EPDS and Whooley questions were comparable in ability to detect depression in women. Women and HPs agreed that there was a need to screen women for depression during pregnancy and the postnatal period. Women preferred the longer questionnaire (EPDS) to the shorter questionnaire (Whooley questions). Neither questionnaire represented value for money when used on its own and did not meet criteria required for a national screening programme for perinatal depression.

Scientific summary

Background

Perinatal depression is a substantial public health problem, affecting 7–20% of women during pregnancy and up to 22% of women during the first postnatal year. It is associated with a range of adverse outcomes for the mother, her baby and the family. Despite this, < 50% of cases are detected by health professionals (HPs) in clinical practice. Historically, screening or case-finding strategies for perinatal depression have been adopted in routine practice, the most common of which is the Edinburgh Postnatal Depression Scale (EPDS). Guidelines on antenatal and postnatal mental health issued in 2007 by the National Institute for Health and Care Excellence (NICE) advocate a case-finding approach, recommending that HPs ask women two brief 'case-finding' questions (the 'Whooley questions') to detect depression during the perinatal period. This recommendation was made in the absence of any validation studies of these questions in a perinatal population; instead, a research recommendation was made for a validation study of the effectiveness of the Whooley questions during the perinatal period.

Although the EPDS has been found to be acceptable to women and HPs, limited research has been conducted on the acceptability of the Whooley questions. Furthermore, evidence for the cost-effectiveness of a screening/case-finding strategy for perinatal depression in routine practice has been limited. An existing decision-analytic model suggests that formal screening/case identification strategies for postnatal depression, including the Whooley questions and the EPDS, do not represent value for money, mainly because of the cost of false positives (FPs). However, data on the diagnostic performance of such strategies were derived from studies of depressed non-perinatal populations.

Research on the diagnostic properties, acceptability and cost-effectiveness of the Whooley questions, compared with the EPDS, to identify perinatal depression will provide evidence to inform future iterations of NICE guidance and policy recommendations on perinatal depression.

Objectives

The objectives of the Born and Bred in Yorkshire PeriNatal Depression Diagnostic Accuracy (BaBY PaNDA) study were:

- to determine the diagnostic accuracy of the Whooley questions and the EPDS against a diagnostic reference standard to identify depression during pregnancy (around 20 weeks) and the early postnatal period (around 3–4 months after birth)
- to assess the temporal stability of positive and negative depression screens between pregnancy and the early postnatal period, and to ascertain whether or not there is an optimal time to screen for perinatal depression
- to investigate the coexistence of depressive symptoms alongside other common mental health problems during the perinatal period
- to determine the acceptability of the Whooley questions and the EPDS to expectant and new mothers, and to HPs, and to determine the potential implications for the care pathway during the perinatal period
- to assess the cost-effectiveness of the Whooley questions and the EPDS as a case-finding strategy for the identification of perinatal depression.

Method

Design

A prospective diagnostic accuracy cohort study in which two depression case-finding instruments – the Whooley questions and the EPDS – were validated against a diagnostic reference standard [the Clinical Interview Schedule – Revised (CIS-R)] during pregnancy (around 20 weeks) and the early postnatal period (around 3–4 months postnatally). Women were followed up for 12 months postnatally to provide a longitudinal assessment of psychological comorbidity during the perinatal period. The study included a concurrent qualitative evaluation of the acceptability of the Whooley questions and the EPDS to women and HPs, and a concurrent economic evaluation of the cost-effectiveness of these case-finding instruments to identify perinatal depression.

Setting

Maternity services in Yorkshire, the Humber and North Lincolnshire.

Participants

A total of 391 pregnant women participated.

Main outcome measures

Women completed the Whooley questions, EPDS and the diagnostic reference standard (CIS-R) during pregnancy and the early postnatal period. Further measures assessed psychological comorbidity, health-related quality of life, acceptability and resource utilisation. A subsample of women participated in in-depth qualitative interviews to discuss their views and experiences of the Whooley questions and the EPDS and, when applicable, their experience of the care pathway. Qualitative interviews were conducted with a sample of HPs [midwives (MWs) and health visitors (HVs)] to discuss their views and experiences of using the case-finding instruments in clinical practice.

The sensitivity, specificity and predictive values of the Whooley questions and the EPDS were calculated against the CIS-R during pregnancy and the early postnatal period. Receiver operating characteristic curves were constructed. The temporal stability of responses to the Whooley questions and the EPDS between pregnancy and the early postnatal period were explored and the longitudinal pattern of psychological comorbidity from pregnancy to 1 postnatal year is described. Qualitative interviews with women and HPs were analysed using phenomenological research methods. A decision-analytic model was developed based on the most recent NICE guidance (2014) to determine the impact of true and false identification of perinatal depression and subsequent treatment on costs and quality-adjusted life-years (QALYs). The cost-effectiveness analysis compared one-stage strategies (i.e. Whooley questions or EPDS) and two-stage strategies (i.e. Whooley questions followed by EPDS). Cost-effectiveness acceptability curves represented the probability of cost-effectiveness of each screening/case-finding strategy for a range of willingness-to-pay (WTP) thresholds.

Results

Diagnostic accuracy results

The Whooley questions, EPDS and CIS-R were completed by 390 women during pregnancy (20 weeks) and 334 (86%) women during the early postnatal period (3–4 months). Prevalence rates for depression were 10.3% during pregnancy and 10.5% postnatally. Diagnostic performance characteristics were reasonable and similar for the Whooley questions and the EPDS (using a cut-off point of ≥ 10) both during pregnancy [Whooley questions: sensitivity 85.0%, 95% confidence interval (CI) 70.2% to 94.3%, and specificity 83.7%, 95% CI 79.4% to 87.4%; EPDS: sensitivity 82.5%, 95% CI 67.2% to 92.7%, and specificity 86.6%, 95% CI 82.5% to 90.0%] and postnatally (Whooley questions: sensitivity 85.7%, 95% CI 69.7% to 95.2%, and specificity 80.6%, 95% CI 75.7% to 84.9%; EPDS: sensitivity 82.9%, 95% CI 66.4% to 93.4% and specificity 87.6%, 95% CI 83.3% to 91.1%). Diagnostic performance characteristics were poorer for the

EPDS using a cut-off point of ≥ 13 at both time points (pregnancy: sensitivity 45%, 95% CI 29.3% to 61.5%, and specificity 95.7%, 95% CI 93.0% to 97.6%; postnatally: sensitivity 62.9%, 95% CI 44.9% to 78.5%, and specificity 95.7%, 95% CI 92.7% to 97.7%). The overidentification of cases by both the Whooley questions and the EPDS (using a cut-off point of ≥ 10) was reflected in relatively low positive predictive values (PPVs) both during pregnancy (Whooley questions: PPV 37.4, 95% CI 27.4 to 48.1; EPDS: PPV 41.3, 95% CI 30.4 to 52.8) and postnatally (Whooley questions: PPV 31.4, 95% CI 24.3 to 45.0; EPDS: PPV 43.9, 95% CI 31.7 to 56.7).

In terms of temporal stability, depression caseness was stable between pregnancy and the postnatal period for most women, with no statistically significant differences between the two time points. Approximately half of depression cases during pregnancy became non-cases during the postnatal period (across both the Whooley questions and the EPDS), and around 7% of non-depressed women during pregnancy became depressed during the postnatal period. This pattern was similar across both case-finding instruments. Depression caseness, based on either of the case-finding instruments, was associated with higher scores on measures of anxiety and somatic symptoms.

Qualitative results

Women and HPs were supportive of screening and case-finding for perinatal depression. The results of the acceptability survey suggested that most women preferred the EPDS to the Whooley questions. EPDS questions were rated as more comfortable to answer, easier to understand and easier to remember, and women had more confidence in their answers, mainly because of the use of 'softer' wording of questions and answers. Conversely, Whooley question 1 (During the past month have you often been bothered by feeling down, depressed or hopeless?) was thought to be less acceptable by women and HPs, largely owing to the use of the terms 'depressed' and 'hopeless', resulting in women not revealing their feelings of depression. MWs and HVs share these concerns resulting in the adaptation of Whooley question 1 in clinical practice to instead ask a general question about a woman's feelings or mood. Women and HPs perceived that it is not socially acceptable for an expectant or new mother to be depressed or feel hopeless and they will wish to avoid the negative sociocultural stigma attached to these terms, including the potential consequences of losing their baby. Training of HPs and a patient-centred environment, in which HPs focus on the mother as well as the baby to promote disclosure about feelings, were identified as important facilitators to improve screening/case-finding to identify perinatal depression.

Cost-effectiveness results

A one-stage strategy using the Whooley questions alone or the EPDS alone was never the most cost-effective strategy, although EPDS alone had a higher probability of being cost-effective than the Whooley questions. Such one-stage strategies were either dominated (or extendedly dominated) or had incremental cost-effectiveness ratios that were above the conventional threshold of £20,000–30,000 per QALY. A two-stage strategy was more cost-effective than a one-stage strategy. In the prenatal period, 'Whooley questions followed by the Patient Health Questionnaire-9' (PHQ-9; a secondary outcome measure used to assess depression symptomatology) had the highest probability of being cost-effective at a WTP threshold of £20,000 (probability = 0.47) and £30,000 per QALY (probability = 0.48); this was closely followed by 'Whooley questions followed by EPDS (cut-off point of 13)'. Similarly, in the postnatal period, 'Whooley questions followed by PHQ-9' had the highest probability of being cost-effective at thresholds of £20,000 (probability = 0.43) and £30,000 per QALY (probability = 0.35). The difference in net monetary benefit (i.e. QALYs × WTP – cost) between the two most cost-effective strategies in both prenatal and postnatal periods was relatively small (< £15) at £20,000 and £30,000 thresholds.

The results show that specificity plays a significant role in the cost-effectiveness analysis because of the cost of treating FPs. For instance, a difference in specificity of 10 percentage points in the postnatal period would result in additional 90 FP cases per 1000 women screened, resulting in an unnecessary cost of £15,488. Our sensitivity analyses showed that results were generally robust to varying prevalence rates, assumed disutility of FP diagnosis and varying resource use by FPs. Our results are in agreement with the

most recent NICE model (2014) of postnatal depression, which also found 'Whooley questions followed by PHQ-9' to be the most cost-effective strategy.

Limitations

Perinatal depression diagnosis was not cross-referenced with women's general practitioner (GP) medical records so the proportion of new cases identified is unknown. The sample was predominantly white, English-speaking women so the results may not be generalisable to other social and ethnic groups. The clinical effectiveness and cost-effectiveness of screening/case-finding strategies was not assessed as part of a randomised controlled trial.

Conclusions

The Whooley questions and the EPDS had acceptable sensitivity and specificity, but their use in practice might be limited by low predictive value and variation in the acceptability of instruments. Economic analysis found that using single case-finding instruments (e.g. the Whooley questions alone or the EPDS alone) was not the most cost-effective strategy. Instead, a two-stage screening/case-finding strategy was the most cost-effective, although it was subject to important uncertainties. Screening and case-finding for perinatal depression using the Whooley questions or the EPDS did not meet the strict criteria laid down by the National Screening Committee (NSC).

Implications for health care

Our research findings suggest the following implications for health care.

- Depression was present in around 1 in 10 women in both the prenatal and postnatal periods. This research suggests that HPs need to be vigilant to the presence of depression in the prenatal period as well as the postnatal period.
- Around half of cases of depression in the prenatal period will resolve after the birth of a baby. A significant portion of depression in the postnatal period is 'new' and HPs need to be alert to this and the need for intervention.
- If case-finding questionnaires are used in clinical practice, then the Whooley questions and the EPDS have been shown to have acceptable psychometric properties.
- The Whooley questions have the advantage of being very brief with good psychometric properties, but there may be limits to the use of these questions in terms of acceptability to pregnant women and new mothers.
- The EPDS has the advantage of having been developed for use in the perinatal period. Although it is a longer instrument, it was found to be acceptable to women and HPs.
- This research suggests that HPs need to be vigilant to the presence of other common mental disorders and may consider using additional instruments or case-finding questions to detect these.
- Screening and case-finding may need to be undertaken within the context of a careful patient-centred consultation. HPs may benefit from training in the administration and interpretation of case-finding instruments such as the Whooley questions and EPDS. They need to be aware that women see depression as being stigmatised and may not answer questions truthfully.
- Screening and case-finding is a large public health undertaking, and the optimal strategy based on costs and benefits is not robustly supported by cost-effectiveness findings, which were subject to important uncertainties.
- Screening and case-finding for depression in the perinatal period did not meet the strict criteria laid down by the NSC.

Recommendations for future research

- The present research did not cross-reference perinatal depression diagnosis with GP medical records and so was unable to determine what proportion of new cases was identified by case-finding instruments. Research is needed that links the results of case-finding instruments with routine clinical records in order to establish the 'yield' of case-finding approaches. On the basis of these results, further determination of the cost-effectiveness of screening and case-finding needs to be undertaken.

- Future validation of case-finding approaches for perinatal depression may benefit from consideration of perinatal populations that include diversity in social, cultural and ethnic backgrounds. Such research would expand our understanding of the feasibility and impact of case-finding for perinatal depression given differences in women's cultural understanding and experiences of perinatal depression.

- Using criteria developed by the NSC, this research found important uncertainties in recommending the optimal strategy for routine use of screening and case-finding. Further research is needed within the context of a randomised controlled trial that assesses the clinical effectiveness and cost-effectiveness of treatment for perinatal depression following screening/case-finding procedures (such as those reported here).

- The present research highlights the importance of depression in the perinatal period and indicates that there is comorbidity with other common mental disorders, such as anxiety. Further research is needed to understand the inter-relationship between different common mental disorders and how these might best be managed for women in the perinatal period.

- Qualitative research has demonstrated variation among health-care practitioners in the mode of administration, level of confidence and delivery of screening/case-finding instruments. Research is needed to develop training programmes and optimise the delivery of screening/case-finding by HPs.

Funding

Chapter 1 Background

Perinatal depression: the health problem

Depression accounts for the greatest burden of disease among all mental health problems and is estimated to become the leading cause of disease burden in high-income countries by 2030.[1] Depressive symptoms are also a common occurrence in women during pregnancy and the postnatal period (collectively referred to as the perinatal period).[2] Perinatal depression is now well recognised and has become an important category of depression in its own right.[3] A number of high-income countries, including the UK, USA and Australia, have issued specific guidelines for the identification and management of the condition in clinical practice.[4–8]

Perinatal depression is often identified in clinical practice and in research using reliable and valid measures of depression that include clinical assessments of depression [using diagnostic criteria such as the *Diagnostic and Statistical Manual of Mental Disorders* (DSM)[9] or the *International Classification of Diseases,* Tenth Edition (ICD-10)[10]] and self-report symptom questionnaires. The core features of a depressive episode are a sustained low mood or loss of interest in pleasurable activities for most of the day or nearly every day, for at least 2 weeks, along with the presence of additional symptoms (such as fatigue or loss of energy, poor concentration, disturbed sleep/insomnia, feelings of worthlessness or guilt). The *DSM* Fifth Edition[9] provides a 'peripartum specifier' for major depressive episodes with onset during pregnancy or within 4 weeks after delivery. In contrast, the ICD-10 criteria do not include a prenatal specifier for depression, although there is a separate code for disorders occurring within 6 weeks after delivery.

However, a common broader definition of perinatal depression is that it encompasses major and minor depressive episodes that occur during pregnancy (prenatal depression) or the first postnatal year (postnatal depression)[11] and includes new and existing episodes of depression.[12]

Prevalence and time course of perinatal depression

Prevalence rates of perinatal depression vary but well-cited estimates indicate that approximately 7.4–20.0% of women meet the criteria for a diagnosis of depression at some point during pregnancy,[13,14] with up to 22% of women experiencing a depressive episode during the postnatal period.[11,15]

The time course of perinatal depression remains under investigation. Historically, the majority of research has been conducted on postnatal depression, although research on prenatal depression (also often referred to as antenatal depression) is gaining momentum.[16] Some studies[11,17,18] suggest that the prevalence of depression is comparable between pregnant and non-pregnant women, whereas others[2,17,19,20] report that the prevalence of depression in women during the postnatal period is higher than at other times in a woman's life.

Although perinatal depression may not be different from depression occurring in non-pregnant women,[8,21] some women will experience depression for the first time during pregnancy or during the postnatal period; some women with diagnosed prenatal depression will continue to be symptomatic in the postnatal period; and some women will have pre-existing chronic or relapsing depression. Research investigating the trajectory of depression experienced by women during the perinatal period has reported mixed results. Evidence from a large longitudinal cohort study conducted in the UK[22,23] suggests higher rates of depressive symptoms in women during pregnancy than during the postnatal period (up to 8 months postnatally), whereas a large US study[2] found that episodes of depression began more frequently in the postnatal period (40.1%) than during pregnancy (33.4%) (with 26.5% of depressive episodes beginning before pregnancy). A systematic review[15] of longitudinal studies similarly reported a higher prevalence of depression in the first year after birth, particularly the first 6 months after birth, than during pregnancy. Recent studies[2,22] suggest that at least a third of cases of postnatal depression begin in pregnancy or before pregnancy.

Although around 50% of women will experience low mood in the first few weeks following birth (often referred to as the 'baby blues'),[24] this is often mild and transient.[25] The majority of cases of postnatal depression are believed to develop within the first 3 postnatal months.[26] Although some cases last around 3 months from onset and resolve spontaneously without treatment,[26] > 50% of women remain depressed for > 6 months.[27] Moreover, up to 50% of women remain depressed for > 1 year following birth[26,28] and around 14.5% of women continue to show depressive symptoms at 4 years post partum.[29] Less research has been conducted on the time course of depression during pregnancy. Systematic reviews[11,13] report similar estimates of point prevalence of depression across the three trimesters of pregnancy, ranging from 7.4% to 12.8%.

Variations in reported estimates of the prevalence and time course of perinatal depression are typically dependent on various factors, such as the size, representativeness and country of the identified sample, the point at which depression was measured during pregnancy (e.g. first, second or third trimester) or the postnatal period (e.g. first 6 postnatal weeks, first 3 postnatal months), and the instruments or criteria adopted to determine the presence of depressive symptoms (e.g. self-report questionnaires using various cut-off scores) or a depressive episode (e.g. diagnosis via structured clinical assessment).[11]

Risk factors for perinatal depression

A substantial literature exists on the evidence of risk factors for perinatal depression.[25,30] The strongest predictor of depression both during pregnancy and the postnatal period is a previous history of depression.[15,22,31–34] Significant risk factors for postnatal depression include anxiety during pregnancy (antenatal anxiety),[22,35] poor marital or partner relationship,[15,34] poor social support,[15,31,32,35] stressful life events during pregnancy or the first postnatal months,[15,31,32] low socioeconomic status,[33] unintended pregnancy[15,33] and domestic violence.[36,37] Many of these risk factors are also known to increase the likelihood of depressive symptoms during pregnancy.[33,35,37,38]

Impact of perinatal depression

Perinatal depression is associated with a range of adverse outcomes for the mother, her baby and the family, in the short and longer term.[14,39,40] Perinatal depression can profoundly affect a woman's well-being, quality of life and relationships[16] and can lead to an increased risk of a range of psychological, behavioural and developmental problems in children.

Depression experienced during pregnancy has been shown to increase the presence of somatic symptoms (such as headaches and gastrointestinal problems)[41] and has been linked to an increased risk of poor obstetric and neonatal outcomes including premature births, low birthweight and a decrease in breastfeeding initiation.[42–44] Such adverse consequences of prenatal depression have been shown to be associated with poor self-reported health and functioning,[45] decreased perinatal care and poor use of antenatal/prenatal care services, and an increased risk of smoking, substance use and alcohol abuse.[14,46]

Perinatal depression can lead to difficulties with parenting, particularly early mother–infant interactions[47,48] and reduced maternal sensitivity and attachment with the infant,[49] which can lead to an increased risk of poor child emotional, behavioural and cognitive outcomes.[39,50,51] The offspring of mothers with perinatal depression are more likely to experience emotional problems,[52,53] such as difficulties with early emotional regulation and poor social skills in school years,[54,55] and are at an increased risk of developing clinical depression during adolescence and at age 18 years.[51,56] Studies have also shown a link between perinatal depression and behaviour problems in children, particularly attention deficit hyperactivity disorder[57] and conduct disorder.[30] Postnatal depression has shown consistent associations with poor cognitive functioning in children, including infant ability to learn and achieve developmental milestones, language and general cognitive development[58–60] (see Stein *et al.*[39]); this is particularly important when postnatal depression persists through the first year of life.[58,60,61] Women with perinatal depression are at an increased risk of suicide[62] and postnatal depression has been linked with infanticide.[62,63]

Maternal perinatal depression has been shown to be moderately correlated with depression in fathers, with an estimated 10% of men experiencing depression in the perinatal period.[64] Emerging evidence suggests an association between paternal perinatal depression and negative child emotional, behavioural and developmental outcomes.[65,66] Paternal postnatal depression also increases the risk of depression in offspring at 18 years.[56] Depression in mothers or fathers negatively impacts on the relationship, leading to decreased marital satisfaction and increased marital discord.[67]

Perinatal depression and comorbidity

Depression is not always experienced in isolation, and epidemiological research shows that depression commonly coexists with other common mental health disorders, such as general anxiety and somatoform complaints.[68] Guidance issued by the National Institute for Health and Care Excellence (NICE)[68] has highlighted the importance of recognising and assessing for coexisting psychological comorbidities to avoid the risk of delivering suboptimal treatment strategies. Limited research has been undertaken on the comorbidity of psychological problems with respect to perinatal depression.

A large study conducted in the USA[17] reported a 13% prevalence of any anxiety disorder during the perinatal period, comparable to the prevalence found among non-pregnant women. Studies suggest that perinatal anxiety (including panic disorder and phobias) is commonly comorbid with perinatal depression,[22,69] with around two-thirds of women with perinatal depression also having a comorbid anxiety disorder.[2,70]

Identification of perinatal depression

Although perinatal depression is well recognised as a mental health condition, it often goes undetected by health professionals (HPs) in routine clinical practice, with recognition rates of < 50% for both prenatal and postnatal depression.[71,72] Despite the finding that clinicians are generally supportive of the need for routine strategies to identify perinatal depression, there is a lack of consistency in the translation of such strategies into routine practice.[73]

Strategies for identifying perinatal depression

Screening or case-finding strategies have been advocated as a method of improving the identification, management and treatment of perinatal depression.[74] A distinction can be drawn between 'screening' (offering a test to a defined population) and 'case-finding' (offering a test to those at the highest risk of having the condition within a defined population), although the two terms are often used interchangeably in the wider literature. Screening is considered appropriate when the condition in question is an important and prevalent health problem, can be effectively treated and cannot be readily detected without screening.[75] The UK National Screening Committee (NSC) are responsible for making recommendations to ministers and the UK NHS regarding the adoption of a screening strategy and as such have a number of clear criteria by which a screening programme is assessed.[76] Based on these criteria, screening programmes should involve a screening test that is simple, safe and accurate and is acceptable to the target population, results in more effective treatment leading to better outcomes and has an acceptable ratio of costs to benefits.[77]

In the UK, national guidance on screening/case-finding for perinatal depression has been inconsistent.[3] In 1999, the National Service Framework for mental health made an explicit requirement for all local areas to have protocols in place for the management of postnatal depression.[78] This resulted in the use of screening and case-finding strategies that focused on the routine or ad hoc administration of screening/case-finding instruments, most notably the Edinburgh Postnatal Depression Scale (EPDS; a 10-item self-report questionnaire asking about symptoms of depression over the past week[79]), in the postnatal period, to the extent that the EPDS became the most widely used screening/case-finding instrument to detect symptoms of postnatal depression.[3,76] In 2001, the NSC concluded that there was insufficient clinical and economic evidence to support the implementation of screening strategies for postnatal depression, highlighting the lack of evidence for the validity of the EPDS as a screening tool.[3,80,81] As a result, resources for the treatment

of postnatal depression by health visitors (HVs) was withdrawn, prompting the NSC to modify its guidance.[82] Although the NSC continued to recommend that the EPDS should not be used as a screening tool for postnatal depression until further research had been conducted, it acknowledged that it could be used as part of the mood assessment in conjunction with professional clinical judgement and clinical interview. The initial NSC recommendation was reaffirmed in its 2010 review,[3] reiterating that there was no evidence that postnatal screening would improve maternal and infant health outcomes for women.

The use of a national screening/case-finding strategy based on ad hoc screening/case-finding using the EPDS has attracted much criticism.[3,74] Such criticisms are based on the ethics of mass screening, concerns regarding the psychometric properties of available screening/case-finding instruments, the acceptability of such screening/case-finding strategies to women or HPs, the paucity of evidence for the cost-effectiveness of screening/case-finding strategies (particularly the costs associated with the management of incorrectly identified cases of perinatal depression) and the absence of evidence that screening/case-finding leads to effective management of perinatal depression and improved mother and infant outcomes.[3,74,76,81]

Clinical guidelines for the identification of perinatal depression

Whether or not screening/case-finding for perinatal depression is recommended differs across various countries. Although clinical guidelines have been issued in several countries recommending the mental health assessment of women during pregnancy and after childbirth,[4,7,8,83–85] these do not necessarily advocate a screening/case-finding strategy per se. For example, the US Preventive Services Task Force[5,6] recommends screening for depression (including in pregnant and postpartum women), but only when this is done within the context of a package of integrated care to include accurate diagnosis, effective treatment and appropriate follow-up. Clinical recommendations issued in Australia[4] recommend screening/case-finding for perinatal depression, using the EPDS, as part of an integrated and holistic package of perinatal care. In contrast, universal screening for depression has not been recommended in the UK as part of NICE guidelines,[86,87] with these guidelines instead recommending that clinicians be alert to possible depression, particularly when there is a previous history of depression, and that they ask about symptoms of depression when there is a concern. Despite such differences, however, all guidelines agree that screening/case-finding instruments alone should not be used to diagnose women with perinatal depression, but should be used to identify those women at risk of perinatal depression who require a fuller assessment of their mental health.

Case-finding instruments for perinatal depression

In 2007, NICE issued guidelines on antenatal and postnatal mental health[7,88] that set out recommendations for the detection and treatment of mental health problems during the perinatal period. These guidelines strongly advocated a 'case-finding' approach (rather than a screening approach per se) and recommended the use of two brief 'case-finding' questions (often referred to as the 'Whooley questions'[89]) as a new strategy for HPs to identify perinatal depression. The NICE guidance[7] states:

> At a woman's first contact with primary care, at her booking visit and postnatally (usually at 4 to 6 weeks and 3 to 4 months), healthcare professionals (including midwives, obstetricians, health visitors and GPs) should ask two questions to identify possible depression:
>
> - 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?
>
> A third question should be considered if the woman answers 'yes' to either of the initial questions:
>
> - Is this something you feel you need or want help with?
>
> Reproduced with permission, pp. 117–18[7]

The guidance[7] also recommended that 'healthcare professionals may consider the use of self-report measures such as the EPDS, the HADS [Hospital Anxiety and Depression Scale[90]] or PHQ-9 [Patient Health

Questionnaire-9[91]] as part of subsequent assessment or for the routine monitoring of outcomes'
(reproduced with permission pp. 117–18).

The NICE recommendation to use the Whooley case-finding questions to detect depression in women
during the perinatal period was based on the current NICE guidelines for depression at that time,[92] along
with evidence from two studies[89,93] of the diagnostic accuracy of the Whooley questions. Neither of these
validation studies were conducted in a perinatal population. Although NICE identified eight studies that
assessed the validity of the EPDS, the diagnostic performance of the instrument varied considerably
between studies, as did the reported prevalence of perinatal depression.[7] As a result, NICE concluded that
although the sensitivity of the EPDS was reasonably good, the lower specificity would mean that almost
half of all women referred for further assessment (based on a positive screen) would be referred
unnecessarily.[7] NICE argued that the value of the Whooley questions was in their brevity, making them
appropriate for use across busy clinical practice settings, and that they lend themselves to use across the
perinatal period.[7]

The Whooley questions, derived from the Primary Care Evaluation of Mental Disorders,[94] were originally
validated in a primary care population. In a US cross-sectional study,[89] the two questions, along with a
diagnostic interview (the Quick Diagnostic Interview Schedule-III-R[95]), were administered to 536 male
veterans attending a medical centre. A positive response to either of the two questions had a sensitivity of
96% and a specificity of 57%, indicating acceptable evidence for the questions' ability to rule out depression,
but with a high number of false positives (FPs) (i.e. a high number of men were incorrectly identified as being
depressed). An additional study[93] examined the Whooley questions plus the third 'help' question in 936
patients attending primary care; the addition of the 'help' question increased specificity to 89%, without
reducing sensitivity (96%) (as validated against the Composite International Diagnostic Interview[96]).

Given the lack of any validation studies of the Whooley questions in a perinatal population, NICE made a
research recommendation for a validation study of the effectiveness of the Whooley questions (compared with
a psychiatric interview) in women during the perinatal period.[7] A subsequent Health Technology Assessment,[76]
which involved a comprehensive systematic review of methods to identify postnatal depression in primary care,
failed to identify any diagnostic accuracy studies that validated the use of the Whooley questions (against
various gold standard diagnostic interviews conducted according to internationally recognised criteria such as
ICD or DSM) for the detection of depression during pregnancy or the postnatal period.[76] Rather, this report
found that the EPDS was the most commonly used instrument to detect postnatal depression, and was
considered to perform reasonably well; sensitivity ranged from 60% (specificity 97%) to 96% (specificity 45%)
for detecting major depression, and from 31% (specificity 99%) to 91% (specificity 67%) for detecting major
or minor depression in the postnatal period. The authors concluded that research comparing the performance
of the Whooley questions (and the 'help' question), the EPDS and a generic depression measure was needed.[76]

A later systematic review[97] identified one diagnostic accuracy study of the two Whooley questions against
diagnostic gold standard criteria. The US study[98] administered the two Whooley questions, along with
depression component of the Structured Clinical Interview for *Diagnostic and Statistical Manual of Mental
Disorders*, Fourth Edition (SCID)[99] as the gold standard, to 506 women attending well-child visits during the
first 9 postnatal months; sensitivity and specificity across the duration of the study was 100% and 44%
respectively. The review[97] failed to identify any studies that validated the two Whooley questions and the
additional help question against a gold standard diagnostic measure.

The Born and Bred in Yorkshire PeriNatal Depression Diagnostic Accuracy (BaBY PaNDA) study was
therefore born out of the need to provide rigorous evidence of the diagnostic properties of the Whooley
questions (including the additional help question) compared with the EPDS to identify depression in
women during pregnancy and the postnatal period. Such evidence will inform future NICE guidelines and
NSC policy recommendations in this area.

Given that the BaBY PaNDA study was conducted in response to the NICE research recommendation for a validation study of the effectiveness of the Whooley questions to identify perinatal depression, and that NICE has advocated the use of the Whooley questions as 'case-finding' questions in two iterations of NICE guidance,[7,8] we will refer to the Whooley questions and the EPDS as 'case-finding' instruments throughout this report. However, we acknowledge that the BaBY PaNDA study could be viewed by some as evaluating the Whooley questions and the EPDS as 'screening' instruments. We therefore refer to both 'screening' and 'case-finding' as potential strategies to identify perinatal depression.

Research conducted since the design of the BaBY PaNDA study

The BaBY PaNDA study was in part informed by the results of a pilot diagnostic accuracy study,[100] conducted by authors of this report, in which the NICE case-finding approached was evaluated. The diagnostic accuracy of the Whooley questions – and the additional help question – was assessed against gold standard psychiatric diagnostic criteria (SCID[99]) in a small but diverse sample of women during pregnancy and the postnatal period. A total of 126 women completed the Whooley questions and a diagnostic gold standard interview during pregnancy, of which 94 went on to complete these during the first 3 postnatal months. A positive response to either of the two Whooley questions had a sensitivity of 100% [95% confidence interval (CI) 77% to 100%] and specificity of 68% (95% CI 58% to 76%) during pregnancy, and a sensitivity of 100% (95% CI 78% to 100%) and specificity of 65% (95% CI 53% to 75%) during the postnatal phase.

These results, along with the previous validation study,[98] provide initial support for a case-finding approach to rule out perinatal depression; however, rates of FPs were high in both studies. For those women with a positive screen, the additional question about the need for help improved specificity but, in contrast to the original validation study,[93] resulted in lower sensitivity. Although the additional help question improved specificity and the ability to rule in perinatal depression, the reduced sensitivity increased the risk of perinatal depression being missed.

In 2014, after commencement of the BaBY PaNDA study, NICE updated its clinical guidelines on antenatal and postnatal mental health.[8,101] As part of this update it conducted a comprehensive systematic review aimed at evaluating appropriate methods and instruments for identifying depression (and anxiety) in women during the perinatal period. Case identification instruments reviewed included the EPDS, the PHQ-9, the Whooley questions and the Kessler-10 (K10).[102] The review identified 57 studies that evaluated the EPDS and two studies that evaluated the Whooley questions (both described above)[98,100] against diagnostic gold standard criteria during pregnancy and/or the postnatal period. Based on the results of this review, along with preliminary evidence regarding the cost-effectiveness of a case-finding strategy (see *Cost-effectiveness of case-finding instruments for perinatal depression*), NICE continued to recommend the use of the two Whooley questions for the identification of depression during pregnancy and the early postnatal period, although the recommendation no longer included the additional question about the need for help.[8] In addition, the use of the EPDS or the PHQ-9, or referral to a general practitioner (GP) or mental health practitioner, was recommended as part of a full assessment if a woman screens positive on the Whooley questions, is at risk of developing a mental health problem or there is a clinical concern.

We are unaware of the publication of any subsequent validation studies that have evaluated the Whooley questions against gold standard diagnostic criteria since NICE issued its updated guidelines (although studies have since been published assessing the Whooley questions against other self-report measures of perinatal depression, e.g. Darwin *et al.*[103]).

Acceptability of case-finding instruments for perinatal depression

The acceptability of the EPDS among most women is widely documented.[76,104] Hewitt *et al.*[76] conducted a systematic review evaluating the acceptability of the EPDS in the perinatal period. In the majority of the studies reviewed, the EPDS was found to be acceptable to women and HPs when a number of factors

were considered, such as the need to forewarn women that they will be asked questions about their mental health; women's preference for the EPDS to be administered in their home; and the importance of the interpersonal relationship between the woman and the HP.

Since the introduction of the routine use of the Whooley questions for case-finding for perinatal depression in the UK,[88] there have been two evaluations[103,105] of the acceptability of the Whooley questions among pregnant women and mothers in the postnatal period. Mann et al.[105] conducted an intramixed-methods cohort study, set in an antenatal clinic in an inner-city hospital in the north of England. They asked pregnant women attending a routine appointment at approximately 26 weeks' gestation to self-complete the Whooley questions, along with a quantitative acceptability survey that included free-text commentaries on their views of the Whooley questions. Women were recontacted at 5–6 weeks postnatally and invited to complete these questionnaires again. A total of 152 and 97 self-report surveys were available for analysis in the prenatal and postnatal phases, respectively. The majority of pregnant (94.7%, $n = 144$) and postnatal (92.8%, $n = 90$) women answered 'yes' to the question about desirability of asking women about their mental well-being; eight (5.3%) pregnant women and seven (7.2%) postnatal women answered 'not sure'. The majority of pregnant (99.3%, $n = 151$) and postnatal (96.9%, $n = 94$) women indicated that they felt 'fairly' to 'very' comfortable answering the case-finding questions. Free-text comments revealed that depressed and non-depressed women found the questions 'easy, simple and straightforward to answer'.[105]

The study by Darwin et al.[103] was conducted in an inner-city hospital in Scotland, also in the prenatal period, at the earlier time of around the woman's booking appointment, typically around 8–10 weeks' gestation. A purposive subsample of 22 women who had previously completed the Whooley questions (including the help question) as part of their routine clinical care followed by a research questionnaire containing the EPDS ($n = 191$) were subsequently interviewed. Women's accounts identified that the context of assessment for depression and the perceived relevance of depression to maternity services influenced a woman's approach to answering the Whooley questions. The importance of enabling environments has previously been identified when routine assessment for domestic abuse was introduced[106] and in previous literature in relation to the acceptability of the EPDS. Practical considerations include the distractions and discomfort associated with completing the questions in a routine health-care clinic rather than at home, where there would be more privacy and time.[104] Furthermore, time constraints and work pressures have been identified as having a negative impact on patient-centredness, whereby a woman's relationship with her HP is central to her feelings of comfort or discomfort in answering the questions.[107]

In light of this limited but insightful evidence, the evaluation of acceptability (EoA) for the BaBY PaNDA study planned to utilise a cognitive framework to further investigate the acceptability of the Whooley questions and the EPDS as case-finding instruments and as individual questions. Cognitive interviewing is a research technique that identifies 'interpretative measurement error' as opposed to traditional components of measurement error, such as not reading the question as worded, or recording answers inaccurately. The cognitive interviewing technique provides a framework to measure performance in terms of a woman's comfort in answering the question, her ease of understanding of the question and her ability to remember, and have confidence in, her answer. One recent study[108] used cognitive interviewing techniques to explore patterns in answer mapping and comprehension of the PHQ-9 to ascertain whether or not the measure captures meaningful symptoms of low mood within a general health questionnaire. This study recruited 18 participants at the point of entry to a longitudinal primary care depression cohort study. Cognitive interviewing revealed that items on the PHQ-9 are interpreted in a range of ways, that patients often cannot 'fit' their experience into the response options, and therefore often feel that the questionnaire is misrepresenting their experience of meaningful symptoms of low mood.

The EoA for the BaBY PaNDA study utilised a mixed-methods approach with concurrent qualitative and quantitative phases and integrated data analysis.[109] Concurrent in-depth interviews with expectant and new mothers and HPs aim to achieve an in-depth understanding of cognitive acceptability of the Whooley questions and EPDS as case-finding instruments and their delivery in routine care, together with a large-scale acceptability survey to test generalisability of the cognitive performance of the Whooley questions

and the EPDS. The opportunity to nest the cognitive evaluation within the cohort of women whose views and experiences of completing the case-finding instruments are grounded within the research study and their routine health-care assessments in the prenatal and postnatal periods will further enhance the strength of the in-depth interview technique.[110–112]

Cost-effectiveness of case-finding instruments for perinatal depression

Perinatal depression is a growing public health concern; a substantial body of evidence suggests that it is associated with significant personal burden for the woman, her partner and child, both in the short and longer term (see above). Less is known about the economic costs of perinatal depression. A small UK-based study[113] estimated the economic burden of postnatal depression to UK health and social services to be around £54M annually (range from £52M to £65M), although it has been suggested this figure may represent an underestimation of the true economic cost of the condition.[8] Mean mother–infant costs over 18 months after birth have been found to cost an additional £591 (uplifted to 2013 prices[8]) for women experiencing postnatal depression.[113] The health and social care costs associated with fathers with depression or fathers at risk of depression during the postnatal period is an additional £159 and £130, respectively, compared with fathers with no depression (£945) (at 2008 prices). More recent UK estimates calculate the cost consequences of adverse child outcomes for perinatal depression at around £8190 per child [including public sector costs, costs of reduced earning and costs associated with reduced health-related quality of life (HRQoL)].[114] This evidence makes the case for identifying cost-effective approaches for screening/case-finding and treating perinatal depression.

Cost-effectiveness analysis is fundamental to decision-making. It brings together data on diagnostic performance of screening/case-finding strategies, prevalence of depression, relative treatment effects of therapies offered to positive cases, opportunity cost of incorrect diagnoses, and the costs and health benefits associated with diagnostic outcomes. The analysis is commonly conducted within the framework of a decision model that explicitly evaluates the clinical pathways and associated costs and outcomes;[115] it then considers the value for money of screening/case-finding strategies in relation to willingness to pay (WTP) for health benefits as determined by the decision-maker.[116]

In the UK context, we identified two recent cost-effectiveness models that evaluated case-finding strategies in the postnatal period: the decision model developed by Paulden et al.[81] and a second model reported in the NICE guidance on antenatal and perinatal mental health, issued in 2014.[8,81] Here we discuss both models in detail. We did not find any cost-effectiveness models for the prenatal period.

Paulden et al.[81] evaluated cost-effectiveness of using case-finding instruments to complement routine care case identification by a GP or HV. The case-finding instruments were given only to women who were identified as depressed in routine care (hence, they were not used as primary case-finding strategies). Paulden et al.[81] compared the EPDS, Beck Depression Inventory[117] and Whooley questions with no complementary screening/case-finding. The model had two linked components: a diagnostic model and a treatment model. In the diagnostic model, women who were identified by their GP/HV as possibly depressed were given one of the above-mentioned instruments, and were then classified as true positive (TP), FP, true negative (TN) or false negative (FN) by case-finding strategies. A treatment model followed the diagnostic model for those who were identified as TP or FP. In line with the NICE guidance available at the time, positive cases were offered structured psychological therapy with additional care.[7] Women who were FN were assumed either to improve spontaneously or to be identified by their GP through routine care and offered the same treatment as above. TN cases did not receive any further treatment. The model derived diagnostic accuracy data by conducting a diagnostic meta-analysis. Resource use data were based on published data sources, whereas unit costs were derived from national cost databases. The model used the NHS and social services perspective and a time horizon of 1 year after screening/case-finding.

Paulden et al.[81] found that complementing routine case-finding with a formal screening strategy (i.e. EPDS, Beck Depression Inventory or Whooley questions) was not cost-effective at conventional WTP thresholds used in the UK (£20,000–30,000) for a gain of 1 quality-adjusted life-year (QALY). All of the above case-finding strategies were associated with an incremental cost-effectiveness ratio (ICER) of > £40,000 per QALY. In a sensitivity analysis that assumed that FP cases would incur lower costs of treatment, the EPDS (cut-off point of ≥ 10) was found to have an ICER of £29,186 per QALY (i.e. just below the highest conventionally used threshold). Another conclusion of the study was that using Whooley questions was never a cost-effective option as a complementary case-finding instrument.

The Paulden et al.[81] model has a number of limitations. First, it evaluates case-finding instruments as complementary to routine case-finding, rather than as primary case-finding strategies. This is an important limitation for the UK decision-making context because the above instruments are commonly used as primary case-finding strategies by HVs. Moreover, it is not clear whether or not one or more of these instruments are already part of the initial case identification by the GP. Second, the study assumes that women who are identified as depressed by the HV/GP but incorrectly diagnosed as non-depressed by the case-finding instrument will still be considered positive and offered treatment. The NICE report also highlights that the Paulden et al.[81] model assumed a zero false-positive rate (FPR) for standard care, which is unrealistic.[8] Third, the Paulden et al.[81] model did not consider two-stage sequential screening/case-finding after implementing Whooley questions to improve specificity of case-finding – this is highly relevant to the context of perinatal depression owing to the high FPR of Whooley questions. Finally, some of the key parameters in the model (including, diagnostic performance of Whooley questions, and quality-of-life data) were based on the general population of depressed patients in primary care, rather than the postnatal population.

Based on the above limitations, the most recent NICE guidance,[8] issued in 2014, developed a new model to compare cost-effectiveness of case-finding instruments as primary case-finding strategies. Our cost-effectiveness analysis is based on this NICE model; hence, a detailed description of this model is presented in Chapter 7 of this report and a brief summary of the model is presented here. The NICE cost-effectiveness model also had diagnostic and treatment components. However, unlike the Paulden et al.[81] model, it evaluated both one- and two-stage case-finding strategies using Whooley questions only, EPDS only (cut-off point of ≥ 10), Whooley questions followed by EPDS (cut-off point of ≥ 10), Whooley questions followed by PHQ-9 (cut-off point of ≥ 10) and routine care case-finding. TP and FP cases received either facilitated self-help (FSH) (72% of cases with mild to moderate depression), intensive psychological therapy (20% of cases with moderate to severe depression) or pharmacotherapy with sertraline (Zoloft®, Pfizer) (8% of cases with moderate to severe depression). FP cases were assumed to receive 20% of the TP treatment (based on the clinical advice from the Guidance Development Group), and were also assumed to have utility reduction by 2% as a result of false diagnosis. As in the case of the Paulden et al.[81] model, the NICE model[8] assumed that FN cases would spontaneously recover (33%), be diagnosed through routine care (8%) or remain depressed. The model used the NHS and social services perspective and a time horizon of 1 year from birth.

The NICE model[8] found that 'Whooley questions followed by PHQ-9' was the most cost-effective case-finding strategy in the postnatal period. This was also the cheapest strategy, at £75,354 per 1000 women screened, with total QALYs of 751.98. The ICER compared with the next cheapest strategy, 'Whooley questions followed by EPDS (cut-off point of ≥ 10)', was £45,593 per QALY, which is above the conventionally used WTP threshold. Moreover, the model found that 'EPDS only' and routine care case identification were dominated by other strategies. One-way sensitivity analysis found that 'Whooley questions followed by EPDS (cut-off point of ≥ 10)' would be cost-effective only if the prevalence of postnatal depression was approximately 20%, which is higher than the prevalence reported in most published studies. The report also highlighted that the results were highly sensitive to the diagnostic performance of the instruments, although 'Whooley questions alone' was never a cost-effective strategy. Finally, the results were also sensitive to the consultation time required to conduct screening/case-finding

such that, when consultation time of EPDS was reduced, 'Whooley questions followed by EPDS' became the most cost-effective strategy.

It should be noted that, although the differences presented per 1000 women appear to be large, these are small differences per individual woman. For instance, per woman screened, the difference in QALYs between 'Whooley questions followed by EDPS' and 'Whooley questions followed by PHQ-9' was only 0.0001 and the difference in cost was only £5.20. This gives a large ICER of £45,593 per QALY; however, the difference in terms of net monetary benefit [i.e. (QALYs × WTP threshold) – cost] per woman was only £2.96 at a WTP threshold of £20,000 per QALY. This shows that the difference between strategies in terms of net monetary benefit is very small.

The NICE model is highly relevant to the UK decision-making context and makes useful comparisons between one- and two-stage strategies that are relevant to clinical practice. However, the model has a few limitations. First, diagnostic performance data for PHQ-9 are not specific to the postnatal population. Second, owing to lack of data on second-stage diagnostic performance, the model assumes that sensitivity and specificity of Whooley questions and subsequent use of EPDS or PHQ-9 are independent of each other (i.e. sensitivity and specificity of first-stage screening/case-finding holds at the second stage). This is clearly an important limitation of the model. Third, the model is deterministic and does not take account of uncertainty in parameter estimates [i.e. it does not include a probabilistic sensitivity analysis (PSA) to account for distributions of input parameters]. This is important in estimating the probability of each strategy being cost-effective at different WTP threshold levels. Given small differences in the deterministic results, the probabilistic analysis may play an important role in understanding decision uncertainty.

We note that both models highlight the significance of diagnostic performance of screening/case-finding strategies in general, and specificity in particular. Both models found that the screening/case-finding strategies that had the highest specificity had the highest probability of being cost-effective. This is because, given the prevalence of postnatal depression, a small change in specificity results in a large number of FP cases that are unnecessarily treated and therefore incur high costs. Moreover, regarding the cases that are missed due to low sensitivity, a significant proportion of them either recover spontaneously during the follow-up period or are identified as depressed in routine care. Likewise, the relatively modest treatment effects of therapies imply that the health benefit of higher sensitivity only results in modest gains in health benefits that are partly offset by reduction in quality of life of FP cases. In conclusion, specificity has been found to be an important factor in the cost-effectiveness of screening/case-finding strategies in perinatal depression.

Chapter 2 Research objectives

The overall purpose of this research was to determine the diagnostic accuracy, acceptability and cost-effectiveness of case-finding questions to identify perinatal depression in women. The objectives of the study were as follows.

1. Instrument validation: to determine the diagnostic accuracy of the Whooley questions and the EPDS against a diagnostic reference standard for the identification of depression during pregnancy (around 20 weeks) and the early postnatal period (around 3–4 months after birth).
2. Longitudinal assessment: to assess the temporal stability of positive and negative depression screens between pregnancy and the early postnatal period, and to ascertain whether or not there is an optimal time to screen for perinatal depression.
3. Assessment of comorbidity: to investigate the coexistence of depressive symptoms alongside other common mental health problems during the perinatal period.
4. Evaluation of Acceptability: to determine the acceptability of the Whooley questions and the EPDS to expectant and new mothers and to HPs, and to determine the potential implications for the care pathway during the perinatal period.
5. Estimates of cost-effectiveness: to assess the cost-effectiveness of the Whooley questions and the EPDS as case-finding strategies for the identification of perinatal depression.

Chapter 3 Methods

Some of this information is reported in Littlewood et al.[118]

Study design

The BaBY PaNDA study was a prospective diagnostic accuracy study of two depression case-finding instruments in a UK perinatal population. The NICE-endorsed ('ultra-brief') Whooley questions and the EPDS (the index tests) were validated against a diagnostic gold standard clinical assessment of depression [the Clinical Interview Schedule – Revised (CIS-R)[119] as the diagnostic reference standard], during pregnancy (around 20 weeks) and the early postnatal period (around 3–4 months postnatally). Women were followed up 12 months postnatally to provide a longitudinal assessment of psychological comorbidity during pregnancy and the first postnatal year. The study included a concurrent qualitative evaluation of the acceptability of the Whooley questions and the EPDS to women and HPs, and a concurrent economic evaluation of the cost-effectiveness of routine screening/case-finding for perinatal depression. The study was embedded within the framework of the existing Born and Bred in Yorkshire (BaBY) cohort study (described in *BaBY cohort*).

Approvals

The BaBY PaNDA study was submitted as a substudy of the existing BaBY cohort study and received approval from North East – York Research Ethics Committee (REC) on 23 April 2013 (reference number: 11/NE/0022) and was subsequently approved by the relevant NHS trust's research and development (R&D) committees.

BaBY cohort

Born and Bred in Yorkshire is a population-based cohort of babies and their parents. It was established in 2011 and recruited women during pregnancy, their partners and babies. Data were collected on a range of clinical outcomes during pregnancy, labour and the neonatal period, including obstetric and maternal morbidity, mental and physical health, and infant health. Data were also collected on the psychological well-being of women and their partners during pregnancy (around 26 weeks of gestation) and for 12 months after their baby's birth (at around 8–12 weeks and 12 months postnatally). The cohort initially recruited for 1 year in York and later extended recruitment to Harrogate, Hull, and Scunthorpe and Goole – representing a target area with around 13,500 births each year. Recruitment to BaBY was estimated at around 60% of women booked for delivery within each of the four regions (further details about recruitment to BaBY are described in *Recruitment*).

The BaBY cohort was a collaboration between the Mental Health and Addiction Research Group and the Epidemiology and Cancer Statistics Group within the Department of Health Sciences at the University of York, Hull York Medical School and colleagues within local NHS trusts. The overarching aim of the BaBY study was to establish a local infrastructure, accessible to the NHS and other health researchers, that should provide a high-quality resource with which to explore questions of real clinical importance to the NHS and other health researchers. Data collected as part of the BaBY study gives potential for wide-ranging research in many areas of adult and child health, including investigation of short- and long-term outcomes.

Study sites

The BaBY PaNDA study was conducted across four regions within Yorkshire and the Humber, and North Lincolnshire: Harrogate, Hull, Scunthorpe and Goole, and York.

Participant eligibility

Eligible women were identified from the wider BaBY cohort. All women invited to take part in the BaBY PaNDA study had previously provided their consent to take part in the BaBY cohort study.

Inclusion criteria
Limited inclusion criteria were applied to ensure recruitment of a representative sample of pregnant women from the wider BaBY cohort sample. Pregnant women were eligible to take part in the BaBY PaNDA study if they met the following criteria.

- They had consented to take part in the wider BaBY cohort and had agreed to be contacted again as part of that consent.
- They were < 20 weeks pregnant.
- They were aged ≥ 16 years.
- They currently lived in an area covered by one of the four study sites.

Exclusion criteria
Women were excluded from taking part in the study if they met any of the following criteria:

- They were non-English speaking.
- They were > 24 weeks pregnant at the time of receipt of a completed consent form.

Women with literacy difficulties were not excluded from taking part; in these cases, all study information and questionnaires were read out to them by the study researcher.

Recruitment

Recruitment to the BaBY cohort
Recruitment to the BaBY cohort took place between August 2011 and March 2015. Women were invited to take part in the BaBY study at around 12–14 weeks of gestation by a member of the midwifery team responsible for their care, or by a trained member of the BaBY team, either at their first antenatal appointment (which may be with their community midwife or in a hospital antenatal clinic) or at the time of their first ultrasound scan. At this point, women were given a study information pack, containing a parent information leaflet, maternal and partner consent forms, a family details sheet and a pre-paid return envelope. Consent forms for both parents were given to the woman; it was for the woman to decide whether or not to give the consent form to her partner. Parents could either provide their written informed consent by completing a consent form at the time of discussing the study with a member of the midwifery or BaBY team or by posting a completed consent form to the BaBY research team. Parents could also choose to consent to the study online using a secure website.

Recruitment to the BaBY PaNDA study
Recruitment to the BaBY PaNDA study took place over a 14-month consecutive period from July 2013 to August 2014. All pregnant women who consented to participate in BaBY and who met all the BaBY PaNDA inclusion criteria (including having provided their consent to be contacted again) were invited to take part in the BaBY PaNDA study.

Eligible women were sent a study information pack at around 15–18 weeks of gestation containing an invitation letter, participant information sheet (providing full details of the study), a summary information sheet (describing the key aspects of the study), a consent form, a contact details sheet and a pre-paid return envelope (see *Appendix 1*). Both the participant and summary information sheets provided contact details for the project team should any women wish to discuss the study in further detail before making a decision on whether or not to participate. The summary information sheet was also made available to women who were interested in taking part in the BaBY cohort either during visits to hospital antenatal clinics and/or during contact with members of the BaBY study team. Women interested in taking part in the study were asked to complete and return the consent form and contact details sheet to the BaBY PaNDA research team. Women who did not return a completed consent form within 2 weeks of being sent the study information pack were contacted by a member of the BaBY team to provide an opportunity to discuss the BaBY PaNDA study and to provide them with an opportunity to ask any questions they may have about the study. Women were contacted by telephone, e-mail or text, depending on the method(s) of contact they had provided as part of their consent to the BaBY cohort study. Women who were interested in participating in the BaBY PaNDA study following this contact were still required to complete a consent form and return this to the BaBY PaNDA research team.

On receipt of a completed consent form, women were contacted by a member of the BaBY PaNDA research team to arrange their prenatal assessment (see *Assessments*). During this contact, the researcher provided an overview of the study, confirmed that the woman understood why the research was being conducted and what she would be asked to do during the study, and answered any questions that the woman may have about the study. One copy of the woman's completed consent form was sent to her GP (with the woman's consent) along with a letter informing them that their patient had been included in the study and a copy of the participant information sheet. Women who were > 24 weeks pregnant at the time of receiving their completed consent form were advised that they were no longer eligible to take part in the BaBY PaNDA study.

Information about the BaBY cohort and BaBY PaNDA study was sent to all GP practices in the recruiting regions and were displayed (when possible) in locations where pregnant women attend as part of their maternity care pathway [e.g. antenatal clinics, National Childbirth Trust (NCT) classes, GP surgeries, children's centres] and was provided online via the BaBY cohort website. Information about the studies was also promoted via regional press releases and local press activities.

Assessments

Women undertook three assessments during the course of the study:

1. prenatal stage – prenatal assessment (at around 20 weeks of gestation)
2. postnatal stage – postnatal assessment (at around 3–4 months postnatally)
3. follow-up stage – follow-up assessment (at around 12 months postnatally).

The diagnostic accuracy of the two depression case-finding instruments (the Whooley questions and the EPDS) was determined against a diagnostic gold standard clinical assessment of depression (the CIS-R[119]) at two time points: the prenatal stage and the postnatal stage.

The Whooley questions, EPDS and CIS-R were also completed at the follow-up stage to determine the prevalence of depressive symptomatology 12 months postnatally. Additional measures were also completed at each of the three stages (prenatal, postnatal and follow-up) to provide a longitudinal assessment of psychological comorbidity, HRQoL, acceptability and resource use (see *Additional outcome measures*).

Diagnostic accuracy measures

The study involved validating two separate index tests, the Whooley questions and the EPDS, against the same diagnostic reference standard, the CIS-R, at two time points: the prenatal stage and the postnatal stage. The index tests were completed before the diagnostic reference standard. At each of the two stages, the index tests and diagnostic reference standard were completed within the same assessment session by one researcher. If it was not possible for women to complete the index tests and diagnostic reference standard within the same assessment session, the diagnostic reference standard was completed within 2 weeks of women completing the index tests; if this time elapsed, the diagnostic reference standard was not completed.

Index tests

Whooley questions

Women were asked the two Whooley[89] questions verbatim by a study researcher at the prenatal, postnatal and follow-up stages:

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?

A 'yes' or 'no' response was required for each question. A 'yes' response to either question 1 or 2 was considered a 'positive screen' for perinatal depression; in these cases, women were then asked the 'help' question (question 3), which required a 'yes', 'yes, but not today' or 'no' response:

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

Women who responded 'yes' to this help question were advised to speak to their GP [and any other HP, such as a midwife (MW) or HV] about their feelings and symptoms and that information about perinatal depression could be found on the NHS website.

The Whooley questions have been previously validated in primary care populations[89,93] and a number of clinical populations.[120–122] As detailed in *Chapter 1*, the Whooley questions have since been validated against diagnostic criteria in two perinatal populations;[98,100] these reported sensitivities of 100% and specificity in the range 44–68%.

The Whooley questions were chosen as the primary index test of investigation, as these were the questions recommended by NICE for use by HPs to aid the identification of depression during the perinatal period[7] and validation studies conducted in a perinatal population were limited (at the time of commencement of the study).

Edinburgh Postnatal Depression Scale

Women self-completed the EPDS[79] at the prenatal, postnatal and follow-up stages. The EPDS is a 10-item self-report questionnaire measuring depression symptoms over the past 7 days (e.g. 'I have been anxious or worried for no good reason', 'I have been so unhappy that I have had difficulty sleeping', 'I have felt sad or miserable'). Each item is scored on a 4-point Likert scale (0–3) with a total score ranging from 0 to 30, with higher scores indicating a greater degree of depression symptomatology. Hewitt *et al.*[76] report an optimal cut-off point of 13 to detect major depression and a cut-off point of 10 to detect major and minor depression combined. A recent comprehensive review by NICE[8] reported pooled sensitivity in the range of 68–83% and specificity of 85–92% for detecting combined depression in the postnatal period when applying these cut-off points. For detecting combined depression in pregnancy, pooled sensitivity in the range of 61–74% and specificity in the range of 86–94% have been reported when applying these cut-off points.

The EPDS was chosen as one of the index tests of investigation as it has been shown to be the most commonly used measure to detect symptoms of postnatal depression in maternity and child services.[123,124] It has also been validated for use in pregnancy.[125] The EPDS is widely used in perinatal mental health research including both prenatal and postnatal populations.[126]

Time taken to administer index tests
The time taken to administer each of the index tests (Whooley questions and EPDS), to include time taken to introduce the questionnaire, time taken to ask the questions (Whooley questions) or provide women with the questionnaire for self-completion (EPDS), time taken for women to provide their responses, and time taken for researchers to record women's responses, were recorded at the prenatal and postnatal stages using a pre-specified template (see *Appendix 2*). These data were used to calculate administration costs as part of the economic evaluation (see *Chapter 7*).

Diagnostic reference standard
Women self-completed the computer-based version of the CIS-R.[119] The CIS-R is a fully structured assessment that assesses 14 symptom domains including depression/depressive ideas, anxiety, somatic symptoms, worry, sleep, fatigue, panic, phobias, compulsions, obsessions, irritability, concentration and worry over physical health. It generates diagnostic categories (including depression severity and diagnosis) according to the ICD-10 criteria[10,127] and provides basic demographic information (e.g. age, sex, ethnic group, employment status and type, and housing type).

The CIS-R has been validated in primary care samples with good reliability and has been used in national psychiatric morbidity surveys.[119,128] It has also been validated for use over the telephone.[129] It has been included as a diagnostic gold standard in systematic reviews of case-finding instruments for common mental health problems, including perinatal depression.[8,76,130]

In an attempt to minimise the potential for incompletion of the diagnostic reference standard, and thereby reducing the study's statistical power to estimate the diagnostic performance characteristics of the index tests, the study design involved one researcher administering both the index tests and the diagnostic reference standard in a single session. The CIS-R was therefore chosen as the diagnostic reference standard due to its self-complete format; the researcher was required only to set up and initiate the CIS-R program. Instructions on how to complete the CIS-R were provided as part of the program (although the researcher was present to ask any questions if necessary; see *Blinding of index test and diagnostic reference standard results*).

Blinding of index test and diagnostic reference standard results
The index tests were completed before the diagnostic reference standard and both the index tests and the diagnostic reference standard were administered in the same session by one researcher. Therefore, the results of the index tests were known to the researcher before the woman completed the diagnostic reference standard. However, despite this design, the level of potential bias within an assessment session was considered minimal; the EPDS (index test) and CIS-R (diagnostic reference standard) are both self-reported measures that women self-completed on paper (EPDS) or on a computer (CIS-R) with only minimal interaction with the researcher (as described above).

However, although the researcher did not specifically ask, women could provide information about their circumstances (current or past) at any point during the assessment session. To capture any potential sources of bias during the assessment session, researchers completed a 'participant assessment record sheet' (PARS) following all assessment sessions with women (see *Appendix 3*). This included recording details of any questions raised by women during completion of the index tests and diagnostic reference standard; any other additional outcome measures completed as part of the assessment session (see *Additional outcome measures*); and any information provided by the woman about her circumstances (current or past), such as whether she currently has or has previously had depression, if she is receiving any treatment for depression, etc.

In order to maintain blinding of index tests and diagnostic reference standard results across the different assessment stages (prenatal stage to postnatal stage, and postnatal stage to follow-up stage), the assessments of each woman were conducted by different researchers (unless it was deemed more sensitive for the same researcher to conduct all assessment sessions).

The level of potential bias recorded within the study is reported in *Appendix 4*.

Additional outcome measures

Women completed a number of additional outcome measures at each of the three stages (prenatal, postnatal and follow-up) to provide a longitudinal assessment of psychological comorbidity, HRQoL, acceptability and resource use during pregnancy and the first postnatal year.

Psychological comorbidity

Outcome measures assessed a range of psychological comorbidities using a number of validated self-report questionnaires administered at the prenatal, postnatal and follow-up stages. Women completed the following questionnaires.

Patient Health Questionnaire-9

The PHQ-9[91] is a nine-item self-report questionnaire that assesses depression severity and symptomatology over the previous 2 weeks based on DSM diagnostic criteria for major depressive disorder. Each item is rated on a scale of 0–3 based on the frequency of depressive symptoms (0 = 'not at all' to 3 = 'nearly every day'). A cut-off point of ≥ 10 is known to detect clinical depression in a UK primary care population.[131] It has been shown to have good sensitivity (74–85%) and specificity (73–84%) for detecting major depression in pregnancy, and excellent sensitivity (82–89%) and moderate specificity (65–84%) for detecting major depression in the postnatal period.[8] It is one of the instruments (along with the EPDS) recommended by NICE as part of further assessment if perinatal depression is suspected following initial case-finding with the Whooley questions.[7,8]

Generalised Anxiety Disorder-7

The Generalised Anxiety Disorder-7 (GAD-7)[132] is a seven-item self-report questionnaire rating anxiety symptoms over the previous 2 weeks. Each item is scored on a four-point scale (0 = 'not at all' to 3 = 'nearly every day'). It has been validated in primary care samples,[132–134] although current diagnostic accuracy estimates are based on non-pregnant populations.[8,130] The GAD-7 is currently recommended by NICE as part of further assessment if anxiety symptoms are suspected in women during the perinatal period following initial case-finding with the shortened version of the GAD-7 – the Generalised Anxiety Disorder-2 (GAD-2).[101,133]

Patient Health Questionnaire-15

The Patient Health Questionnaire-15 (PHQ-15)[135] is a 15-item self-report questionnaire that assesses somatic symptom severity (the experience of physical symptoms in response to psychological problems/stressors) over the previous 4 weeks. Each item is scored on a 3-point scale (0 = 'not bothered at all', 1 = 'bothered a little', 2 = 'bothered a lot'). It has been validated in primary care populations.[135,136]

Clinical Interview Schedule – Revised

The CIS-R[119] was also used to identify other common mental health problems, including mixed anxiety and depressive disorder, generalised anxiety disorder, panic disorder, obsessive–compulsive disorder and phobias.

Health-related quality of life

Women completed two generic measures about their HRQoL at the prenatal, postnatal and follow-up stages.

Short-Form Health Survey-12 items

The Short-Form Health Survey-12 items version 2 (SF-12v2®)[137] is a 12-item self-report questionnaire that assesses individuals' perception of their general health over the past 4 weeks (standard version). The SF-12v2® Health Survey is copyrighted by QualityMetric Incorporated [SF-12v2™ Health Survey © 1992–2002 by Health Assessment Lab, Medical Outcomes Trust and QualityMetric Incorporated. All rights reserved. SF-12® is a registered trademark of Medical Outcomes Trust. (IQOLA SF-12v2 Standard, English (United Kingdom) 8/02)]. It measures health on eight dimensions (physical functioning, role physical, bodily pain, general health, vitality, role emotional, social functioning and mental health) and yields two summary scores [physical component score (PCS) and mental component score (MCS)]. The SF-12v2 is a reliable and well-validated questionnaire[138] and has been used as an outcome measure in a wide variety of patient groups, including patients with depression[139] and women with postnatal depression.[140]

European Quality of Life-5 Dimensions

The European Quality of Life-5 Dimensions (EQ-5D)[141] is a generic preference-based measure of health state covering five dimensions (mobility, self-care, usual activity, pain/discomfort and anxiety/depression) according to three possible levels of severity (no problems, some problems and major problems). Individuals are asked to rate their personal health state on that day for each of the five dimensions and also on a visual analogue scale, for which 100 represents the best imaginable health state and 0 represents the worst imaginable health state. The EQ-5D has been validated in UK populations and has been used to measure HRQoL in patients with depression[139,142] and women with postnatal depression.[143]

Acceptability survey

Women completed a self-report acceptability survey at the prenatal and postnatal stages; this was completed immediately following completion of the Whooley questions and the EPDS (index tests). The acceptability survey was originally developed to assess acceptability of the EPDS in a postnatal sample of Australian women[144] and later adapted to include an assessment of the acceptability of the Whooley questions in women during pregnancy and the early postnatal period.[105] The survey was further adapted for use within the BaBY PaNDA study (see *Appendix 5*) in order to assess a range of concepts regarding the acceptability of the Whooley questions and the EPDS as perinatal depression case-finding instruments, along with the processes for completing the questions (a more detailed description of the acceptability survey is provided in *Chapter 6*).

Resource use questionnaire

Women completed a self-report questionnaire to assess resource use at the prenatal, postnatal and follow-up stages (see *Appendix 5*). The questionnaire asked women about their use of health and social care services (including hospital- and community-based services) and their use of medication for mental health problems (when applicable) in the previous 6 months (when completed at the prenatal stage) and in the time since their previous completion of the same questionnaire (when completed at the postnatal and follow-up stages). Women were also asked to rate how happy they were feeling at that moment on a visual analogue scale, in which 0 represented completely unhappy and 10 represented completely happy. Women were asked about their employment status (including type of occupation, gross pay, weekly hours worked, days absent from work, weeks unemployed) at the prenatal stage only.

Biographical information

Women were asked to complete a brief bespoke biographical questionnaire at the prenatal stage only (see *Appendix 5*). This asked about their highest educational qualification and whether or not they had children already (and if so, how many children). It also asked if they had ever suffered with anxiety or depression. If the woman reported that she was currently suffering with anxiety or depression or had previously suffered with anxiety or depression, she was asked to indicate if she had been prescribed antidepressants, and whether or not she had seen anyone other than her GP for help with her anxiety or depression, and if so, who she saw (e.g. psychiatrist, psychologist, counsellor, etc.).

Method of data collection

Assessment sessions were conducted face to face at the prenatal and postnatal stages. At the follow-up stage, and for those women unable to attend a face-to-face session at the postnatal stage, outcome measures were collected by telephone (initial option) or a combination of post (self-report questionnaires including the Whooley questions and EPDS) and telephone (CIS-R). Face-to-face sessions were arranged for those women who specifically requested this method of data collection at the follow-up stage. Face-to-face assessment sessions were conducted at a time and place of the woman's choice (e.g. hospital antenatal clinic, the woman's home). Women were advised via the participant information sheet and during discussion with the study researchers that each assessment session would last approximately 30–40 minutes.

All study researchers underwent training in all aspects of the study including study recruitment, study protocol, and administration and interpretation of all outcome measures, including the Whooley questions and EPDS (index tests) and the CIS-R (diagnostic reference standard). Robust protocols were in place to deal with any risk issues that may arise during the assessment sessions, including the identification of women at risk of depression (see *Assessment of risk*), and all study researchers underwent training on these risk protocols.

Assessment of risk

Robust protocols were in place to deal with instances when cases of depression or anxiety were identified. Women identified as currently experiencing depression or anxiety as the probable primary diagnosis on the CIS-R were advised that their responses suggested that they may be experiencing some symptoms of depression or anxiety and were advised to discuss their feelings and symptoms with their GP or any other HPs (i.e. MW, HV). In addition, for cases when the probable primary diagnosis was a depressive episode, consent was also sought to write to the woman's GP to advise them of the outcome of the assessment (to include information regarding the probable primary diagnosis and the PHQ-9 score). Researchers were advised to discuss any participant concerns regarding risk issues with a clinical member of the team. Researchers were required to document all cases of risk of depression or anxiety identified during an assessment session on a depression/anxiety risk form that was countersigned by a clinical member of the team. Protocols were also in place to deal with any instances of risk of self-harm or suicide identified during the assessment session (via the CIS-R or question 9 of the PHQ-9).

In addition, any women who screened positive on the Whooley questions and subsequently responded 'yes' to the help question 'Is this something you feel you need or want help with?' (see *Whooley questions*) were advised to speak to their GP (and any other HP, such as a MW or HV) about their feelings and symptoms and that information about perinatal depression could be found on the NHS website.

Sample size

The sample size calculation was based on a previously developed method for design accuracy in diagnostic accuracy studies.[145] For an expected sensitivity of 95% and a minimal acceptable lower 95% CI of 80% to be detectable with 95% probability, a total of 50 women with depression (according to the diagnostic reference standard) in the perinatal period would be required. With an estimated prevalence of perinatal (prenatal and postnatal) depression of 20%[11] and assuming 34% attrition between the prenatal (stage 1) and postnatal (stage 2) assessment stages (based on a previous validation study of the Whooley questions in a perinatal population[100]), the required recruitment target was 379 pregnant women.

Study completion

Women were deemed to have exited the study when they:

- had completed their assessment session at the follow-up stage (around 12 months postnatally)
- expressed a wish to withdraw from the study.

Data were retained for any women who wished to withdraw from the study up to the date of withdrawal, unless they specifically requested that this information be removed.

Following consent to the study and completion of the assessment session at the prenatal stage (around 20 weeks of gestation), any women who subsequently suffered a fetal loss (such as a miscarriage, termination of pregnancy, stillbirth or neonatal death) was not automatically withdrawn from the study. Women were contacted by letter to offer the study team's condolences, thank them for their interest and/or involvement in the study and to provide them with the opportunity to remain in the study should they wish to do so. Women who did not respond to this letter were not contacted again by the BaBY PaNDA team (the letter advised of this). Data were retained for these women up to the date of the event, unless they specifically requested that their information be removed. Robust protocols were in place via the wider BaBY cohort study (and in liaison with clinical teams using routine NHS systems at each of the study sites), which acted to alert the BaBY PaNDA study team of such events.

Statistical analysis of clinical data

Methods for conducting the statistical analysis of collected data, including diagnostic accuracy at each time point, will be reported alongside the statistical results in *Chapter 5*.

Evaluation of acceptability

A concurrent qualitative evaluation was conducted with women and HPs to examine the acceptability of the Whooley questions and the EPDS, the extent to which these instruments capture appropriate information for the effective case-finding of perinatal depression in routine perinatal care, and the implications for the care pathway of delivering such perinatal depression case-finding instruments in routine care.

Overview of acceptability study design
The qualitative evaluation utilised a mixed-methods approach that included the use of both a quantitative survey tool (the adapted acceptability survey) and in-depth interviews.

Interviews with women
All women in the study completed the adapted acceptability survey (see *Acceptability survey* and *Chapter 6*) at the prenatal and postnatal stages to provide an overview of acceptability and process-related issues for the Whooley questions and the EPDS. These issues were explored in more detail through in-depth interviews conducted with a purposive subsample of 25–30 women. The interview sampling framework aimed for a maximum variation on the basis of socioeconomic background, age, parity, positive/negative screen on the Whooley questions, and study site. Information about the nested qualitative study was provided in the BaBY PaNDA participant information sheet and women provided their consent to participate in these interviews at the time of providing their initial consent to the BaBY PaNDA study. Women who chose not to provide consent to take part in the in-depth interviews could still continue to participate in the BaBY PaNDA study. Women participated in a maximum of three in-depth interviews. Interviews were conducted on a subsequent and separate occasion to completion of the BaBY PaNDA assessment session at the prenatal, postnatal and follow-up stages of the study. Interviews were conducted at a time and location of the woman's choice

(e.g. hospital, the woman's home). The interviews were guided by the use of a semistructured topic guide based on cognitive interviewing methodology (see *Chapter 6*) and open-ended probes (see *Appendix 6*).

Interviews with health professionals

In-depth semistructured single interviews were conducted with a purposive sample of up to 12 HPs (six MWs, six HVs) to include diversity in age, job role, professional grade, experience and study site. Interviews explored HPs' views and experiences of delivering the Whooley questions and the EPDS in routine clinical practice within their NHS trust, and their associated training needs, against descriptions of recommended routine practice and policy for perinatal mental health care in their study site. HPs were provided with an information sheet that described the BaBY PaNDA study and the rationale for conducting interviews with HPs, and a consent form (see *Appendix 7*). Interested HPs were asked to return the consent form to the qualitative researcher who would arrange to conduct the interview at their preferred time and location.

More detailed methods relating to the qualitative study will be described alongside the qualitative results in *Chapter 6*.

Cost-effectiveness of case-finding strategies for perinatal depression

The method for conducting the economic analysis and decision modelling for the cost-effectiveness data will be reported alongside the cost-effectiveness results in *Chapter 7.*

Patient and public participation

The BaBY PaNDA study received patient and public involvement (PPI) throughout the course of the study, including involvement in the design, management, interpretation and dissemination stages. PPI members included women with lived experience of perinatal depression and users of maternity services.

Deborah Morgan, founder and Chief Executive Officer of Perinatal Illness UK, was a coapplicant on the funding application and was involved in the initial design and application stages of the study. She provided comments and feedback on drafts of the study protocol and participant information materials. At the time of Deborah's involvement, Perinatal Illness UK was a registered charity and stakeholder with NICE, providing support and advice to women and their families affected by perinatal illness. Deborah was unable to contribute further to the study once the study had started to recruit owing to other work and personal commitments.

Drafts of the funding applications and participant information materials were also reviewed by several new mothers, who at that time were participating in the BaBY cohort, and also by the founder and co-ordinator of a local online support group for women with perinatal depression. A number of these women went on to contribute to the management and conduct of the study; in particular, we sought their advice on the sensitive issue relating to the procedure for dealing with women who suffered a fetal loss after consenting to the study. We were able to utilise their views and feedback to help us to develop an appropriate and sensitive protocol for contacting women who experienced such events. This procedure was subsequently approved by our Research Ethics Committee and relevant trust R&D committees.

Alice North became involved with the study in October 2014, following her participation in the BaBY cohort, and has played an integral role in the study. Alice has been involved with the local branch of the NCT, a large charity for parents, for > 9 years and co-ordinates a weekly group providing support for pregnant women and new mothers. She also has links to the UK charity Bliss and is a trained peer supporter for the Treasure Chest Breastfeeding Group, providing breastfeeding support and information for new parents, and assisting with the delivery of monthly antenatal information workshops.

Alice has been a member of the Study Management Group and Study Steering Committee and, as such, has been involved in many discussions regarding the continued management of the study, and, more recently the interpretation of the results, with particular contribution to the findings of the qualitative interviews conducted with women involved in the study. She has been involved with marketing activities associated with the study (and the wider BaBY cohort) including press releases and co-drafting a BaBY PaNDA update that was sent to women ahead of their follow-up assessment at 12 months postnatally. Alice also drafted the plain English summary for this report and is an author on the report. She is also keen to continue working with the team on future planned dissemination activities relating to the study (e.g. co-drafting the study findings summary to be sent to those women who participated in the study, press releases, etc.), and has agreed to liaise with the NCT to investigate using its newsletter as a platform to disseminate the study findings to their audience.

The input we have received from the PPI members has been invaluable and has contributed to the successful delivery of the study.

The full study protocol can be accessed via the National Institute for Health Research (NIHR) Health Services and Delivery Research (HSDR) website [www.nets.nihr.ac.uk/projects/hsdr/11200423 (accessed July 2016)].

The study is reported according to the standards for the reporting of diagnostic accuracy studies (STARD) statement.[146,147]

Chapter 4 Protocol changes

Assessment of risk

Our original protocol for the identification of women at risk of depression or anxiety involved writing to a woman's GP (with the woman's consent) for all cases in which depression or anxiety was identified as the probable primary diagnosis on the CIS-R (the diagnostic reference standard) (in addition to advising women to discuss their feelings and symptoms with their GP and any other HP, such as a MW or HV). However, following consultations with GPs and primary care clinicians, and given that the primary focus of the BaBY PaNDA study was the identification of risk of depression (rather than risk of anxiety), it was felt that advising GPs of those women who met diagnostic criteria (on the CIS-R) for an anxiety disorder would overburden GPs. The protocol was therefore revised so that a woman's GP was contacted only if she was identified as meeting diagnostic criteria for a depressive episode on the CIS-R. Women who met diagnostic criteria for an anxiety disorder were advised to discuss their feelings and symptoms with their GP (or any other HP). This revised protocol was implemented from August 2013 (by which point five women had completed their prenatal assessment).

Additional recruitment strategies

Initial recruitment to the study was slower than anticipated and so a number of additional recruitment strategies were implemented. We developed a one-page summary information sheet that described the key aspects of the study (see *Appendix 1*). This was included in the study information pack sent to all eligible women (this also included the participant information sheet which provided full details of the study, as described in *Chapter 2*). The summary information sheet was also made available to those women who may be interested in taking part in the wider BaBY cohort study during visits to the antenatal clinic and/or during contact with members of the BaBY team. It was hoped that providing information about the study in a summary format may encourage more women to take part in the BaBY PaNDA study.

Women who were eligible to take part in the BaBY PaNDA study were sent a study information pack through the post. We received approval to contact those women who had not returned a completed consent form within 2 weeks of being sent the information pack, to provide them with an opportunity to ask any questions they may have about the study. Women were contacted by a member of the BaBY cohort team, by telephone, e-mail or text, depending on the method(s) of contact they had provided during consent to the BaBY cohort. Women who were interested in participating in the BaBY PaNDA study were required to complete a consent form before the prenatal assessment session was conducted.

The initial participant information sheet stated that the questionnaire sessions would last around 1 hour. In practice, however, these sessions took around 30–40 minutes to complete (on average). The participant information sheet was therefore amended to more accurately reflect the time taken to complete the questionnaire sessions.

We amended the time frame the study information packs were sent out to eligible women, from 16–18 weeks of gestation to 15–18 weeks of gestation, to allow researchers more time to contact women on receipt of a completed consent form and to conduct their prenatal assessment session (to be conducted at around 20 weeks of gestation).

We also extended our recruitment period from 12 months to 14 months.

These additional recruitment strategies were implemented from April 2014 (by which point approximately 180 women had consented to take part in the study).

Eligibility criteria

The participant information sheet advised women that their first questionnaire session would take place at around 20 weeks of gestation. However, during the initial recruitment period, we received a number of consent forms from women who were between 24 and 27 weeks pregnant, and with whom we subsequently conducted the prenatal assessment session. In order to promote consistency of timing of the prenatal assessments (which were to be completed around 20 weeks of gestation), and following advice from our study statistician, we took the decision to recruit only those women who returned their completed consent forms before they were 24 weeks pregnant. Following this, any women who returned completed consent forms after they were 24 weeks pregnant were contacted to advise them that they were no longer eligible to take part in the study given the time frame of the initial questionnaire session and to thank them for their interest in the study. This protocol change was implemented from March 2014 (by which point approximately 160 women had completed their prenatal assessment).

Contact with women following an adverse event

In our original protocol, women who suffered a fetal loss (such as miscarriage, termination of pregnancy, stillbirth or neonatal loss) following consent to the BaBY PaNDA study would not be contacted again by the BaBY PaNDA study team, unless they indicated that they would like to actively continue with the study. However, following the first case of such a fetal loss, we sought advice from members of our PPI group as to whether or not our protocol was appropriate. Our PPI group felt that women suffering a fetal loss would most likely appreciate being contacted by the study team, even though they may not wish to continue taking part in the study. They felt that it would be appropriate to contact women in this instance and advised that a letter from the study team would be more appropriate than a personal card or a telephone call, and that providing an 'opt-in' option would be the most appropriate and sensitive method of indicating their wish to continue with the study. The protocol was therefore revised to allow a letter to be sent to any woman (and her partner, if known via the BaBY cohort study) who suffered a fetal loss following consent to the BaBY PaNDA study. The letter was sent from the BaBY PaNDA study team offering our condolences, thanking them for their interest in the study and advising them to contact the study team if they wish to continue taking part in the study. Women who did not respond to this letter were not contacted again by the BaBY PaNDA study team and women were made aware of this in the letter. This protocol change was implemented from June 2014 (by which point approximately 250 women had consented to take part in the study).

Outcome measures

Although one of the objectives of the study was to determine the diagnostic accuracy of the Whooley questions and the EPDS during pregnancy and the early postnatal period, evidence suggests that many women report depressive symptoms in the second 6 months after birth.[36] Therefore, in order to obtain information as to the prevalence of depressive symptoms in women 12 months after birth, we revised the protocol to include the Whooley questions and EPDS in the range of outcome measures to be completed at the follow-up stage of the study. This protocol change was implemented from November 2014 (no women had completed their follow-up assessment at this stage).

Retention

We originally planned to conduct assessment sessions with women at the postnatal stage (around 3–4 months postnatally) during a face-to-face interview, with the option of conducting these sessions over the telephone for those women unable to attend a face-to-face session at this time point. However, we experienced difficulty in contacting some women at this time point, and some women when contacted felt they did not have the time to meet face to face with a researcher or to conduct the assessment over the telephone. In order to provide greater flexibility for women to participate and to allow us to collect study data at this important time point, we amended the protocol so that the self-report questionnaires (i.e. all study questionnaires with the exception of the CIS-R) were sent out in the post to those women who we had been unable to make contact with, or who advised that they were unable to complete the assessments sessions face to face or over the telephone. This would allow these women to complete the questionnaires at home in their own time and return these to the research team using a pre-paid envelope. We also included a contact details sheet for women to complete and return along with their completed questionnaires. On receipt of the completed questionnaires, women were contacted to complete the CIS-R (as the diagnostic reference standard) over the telephone (this was only conducted if the questionnaires had been completed within < 2 weeks of women having completed the Whooley questions and the EPDS, as indicated by women on these questionnaires). The risk protocol for the identification of self-harm and suicide was revised to describe the procedure to follow if a potential self-harm/suicide risk was identified (via the PHQ-9) on return of a postal questionnaire. This protocol change was implemented from November 2014 (by which point approximately 150 women had been approached to arrange their postnatal assessment).

Qualitative study

A number of minor changes were implemented in relation to the design and conduct of the qualitative study.

We initially aimed to conduct qualitative interviews with women around 1 week after they had completed their BaBY PaNDA assessment session (at the prenatal, postnatal and follow-up stages) to ensure that the interview sampling criteria for positive/negative screens on the Whooley questions for risk of depression were still valid at the time of the qualitative interview. In practice, we found this difficult to achieve based on the availability of women to attend the interview at such short notice following completion of their BaBY PaNDA assessment session. We therefore extended this time frame to allow the qualitative interview to take place initially within 6 weeks post BaBY PaNDA assessment. We felt this longer time frame would not affect the reliability of the interview sampling framework given the often persistent nature of symptoms of depression and/or anxiety. Moreover, the cognitive interview methodology applied asks women to listen to, answer and consider various aspects of the individual questions for the two case-finding instruments (Whooley questions and EPDS); thus, the woman is not required to recall her experience of answering these questions during her previous BaBY PaNDA assessment session. Given this, the extended time frame between the assessment session and the qualitative interview was not a concern with respect to memory recall; with this in mind, the time frame originally imposed was later removed to allow women more flexibility to participate in their second or third qualitative interviews.

We had initially planned to apply the cognitive interviewing methodology to all qualitative interviews conducted with women (women were invited to participate in up to three qualitative interviews). However, based on qualitative data from the first round of qualitative interviews, it was anticipated that the cognitive interview methodology based on the case-finding questions may not identify crucially different views for women between the different time periods (e.g. from the prenatal stage to the postnatal stage, and from the postnatal stage to the follow-up stage). In this case, it was expected that data saturation would be achieved relatively quickly and that the primary focus of the second and third interviews would be on the experiences of the care pathway for the subsample of women experiencing symptoms of depression or anxiety. The cognitive interviewing methodology was therefore not applied to all second interviews

conducted with women (these protocol changes were implemented in November 2014 before any women completed their second interviews).

We originally intended to conduct all qualitative interviews with participants (women participants and HPs) face to face; however, some participants expressed difficulty with committing to a face-to-face interview (usually because of work commitments). We therefore offered participants the option of undertaking the qualitative interviews over the telephone, if this was their preferred option. This would therefore allow those participants interested in participating in the interviews, but who may struggle to do this face to face, the opportunity to still participate. Consent was still sought to record the interviews irrespective of the method of interview delivery. This protocol change was implemented from May 2015 (by which point no women had completed their second or third interviews, and no interviews had been conducted with HPs).

Additional protocol changes

A number of additional minor protocol changes were implemented during the course of the study. These included (1) sending out a 'BaBY PaNDA update' to all women ahead of their follow-up questionnaire session at 12 months postnatally (this provided information on how recruitment had progressed in the BaBY PaNDA and BaBY cohort studies, and reminded women that their 12 month follow-up session was due soon), (2) extending the length of the study by 5 months following approval from the funder (the NIHR HSDR programme) and (3) general clarification of existing procedures.

Chapter 5 Statistical analysis of clinical data

Summary of analysis plan

Overview

The progression of women through the study in terms of numbers eligible, consented and followed up were collated and illustrated with the help of a flow diagram, including reasons for exclusion. Frequencies of dropouts were presented with reasons when available. Baseline characteristics collected as part of the BaBY cohort study and BaBY PaNDA, as well as demographic data collected as part of the CIS-R interview, were presented descriptively, including age, education, marital and employment status, number of previous children, and history of depression and anxiety. Baseline measures that were available for the whole BaBY cohort as well as the BaBY PaNDA sample population (age, PHQ-8, GAD-7 and PHQ-15) were compared in order to establish the representativeness of the BaBY PaNDA study sample. PHQ-8 data were derived from PHQ-9 responses provided in the BaBY PaNDA sample, as the question on death and self-harm was not collected for the whole BaBY cohort.

Further statistical analyses addressed, in particular, the BaBY PaNDA objectives of instrument validation (objective 1), temporal stability (objective 2) and assessment of comorbidity (objective 3). All analyses were conducted in Stata version 13.1 (StataCorp LP, College Station, TX, USA). Full analyses are specified in the study's statistical analysis plan (version 1.3).

Instrument validation (objective 1)

Descriptive statistics for the three tests under investigation (the Whooley questions, EPDS and the CIS-R) were calculated for each study site at each stage of the analysis (stage 1, prenatal assessment at around 20 weeks' gestation; and stage 2, postnatal assessment at around 3–4 months postnatally).

Diagnostic accuracy studies aim to measure the extent of the agreement between the results of the test under investigation, referred to as the 'index test' (in this case the Whooley questions and the EPDS), and the outcome of the reference standard (the CIS-R, as the diagnostic clinical assessment of depression). There are four possible test outcomes: an individual is depressed based on the CIS-R, and the index test is positive (TP) or the index test is negative (FN); or an individual is not depressed based on the CIS-R and the index test is positive (FP) or the index test is negative (TN). Test outcomes were reported as 2 × 2 contingency tables of the index tests (Whooley questions and EPDS) against the diagnostic reference standard (CIS-R).

In order to measure the performance of each index test (Whooley questions and EPDS), a number of diagnostic test statistics were calculated, including sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (PLR), negative likelihood ratio (NLR) and area under the curve (AUC). See *Appendix 8* for the derivation and interpretation of these statistics in the context of this study. Statistics will be presented with 95% CIs to express uncertainty around the estimates.

Values for each statistic depend on the cut-off point applied to the index test. For the Whooley questions, the cut-off point was a positive response to at least one of the two case-finding questions. For the EPDS, the original protocol considered the EPDS cut-off score of ≥ 13 points. This high cut-off point was considered appropriate for a population of women during pregnancy and the postnatal period.[125] However, following sight of initial BaBY PaNDA data, it became apparent that the lower EPDS cut-off score of ≥ 10 points should additionally be considered for identification of perinatal depression (based on the CIS-R), as this has been shown to detect both transient and enduring distress in pregnancy.[148]

Receiver operating characteristic (ROC) curves plot sensitivity against 1 – specificity for each cut-off point of the index test to evaluate the trade-off between sensitivity and specificity across the score range. ROC curves were therefore constructed for the Whooley questions and EPDS to determine each test's performance profile.

All diagnostic accuracy analyses were conducted separately for stages 1 and 2. Only women with complete outcomes on the CIS-R, Whooley questions and EPDS at each time point were included in the diagnostic accuracy analyses. The numbers of women with indeterminate or missing results at one or more time points were summarised descriptively, together with reasons for these when available. Demographic characteristics of women with missing results at any time point were compared with those who had complete data sets.

Longitudinal assessment and temporal stability
Depression outcomes from each index test instrument were cross-tabulated across study stages according to the number of depression cases and non-cases at stage 1 (20 weeks' gestation) that remained or changed caseness at stage 2 (3–4 months postnatally). The temporal stability of responses between stages 1 and 2 for each of the depression index tests was explored using McNemar's test, an appropriate test for 2 × 2 contingency tables for repeated measures data. The chi-square values at each time point and corresponding p-values were reported.

Assessment of comorbidity
The coexistence of depressive symptoms alongside other common mental health problems (based on secondary outcomes of PHQ-9, GAD-7, SF-12v2, EQ-5D and PHQ-15) was summarised descriptively at each time point. Summaries were presented by screen-positive/screen-negative results of perinatal depression based on each index test instrument.

Study progression

Following invitation via the BABY cohort, 426 pregnant women consented to take part in the BaBY PaNDA study, of whom 391 completed the assessment at stage 1. Of these, 390 were analysed at stage 1 (20 weeks' gestation) and 334 were analysed at stage 2 (3–4 months postnatally). Attrition between stages 1 and 2 (15%) was less than anticipated in the sample size (34%). Study progression is illustrated in the flow diagram in *Figure 1*.

Baseline characteristics

Characteristics of mothers in the BaBY PaNDA study
Age for consenting women was extracted from the BaBY cohort data and was available for 422 of 426 consenting women. Women in Harrogate [$n = 67$, mean age 33.1 years, standard deviation (SD) 4.94 years] and York ($n = 198$, mean age 31.2 years, SD 4.84 years) tended to be older than those in Hull ($n = 108$, mean age 29.8 years, SD 5.43 years), and Scunthorpe and Goole ($n = 49$, mean age 29.4 years, SD 5.85 years).

Further demographic data were collected as baseline characteristics at the first study assessment at stage 1 and are reported for women analysed at stages 1 and 2 in *Table 1*. The study population were predominantly white, in their early thirties, educated to undergraduate or postgraduate degree level and employed, with approximately one-quarter having experienced anxiety or depression at some point in the past. The demographic profile was similar between women at stage 1 and those who completed the study assessment at stage 2.
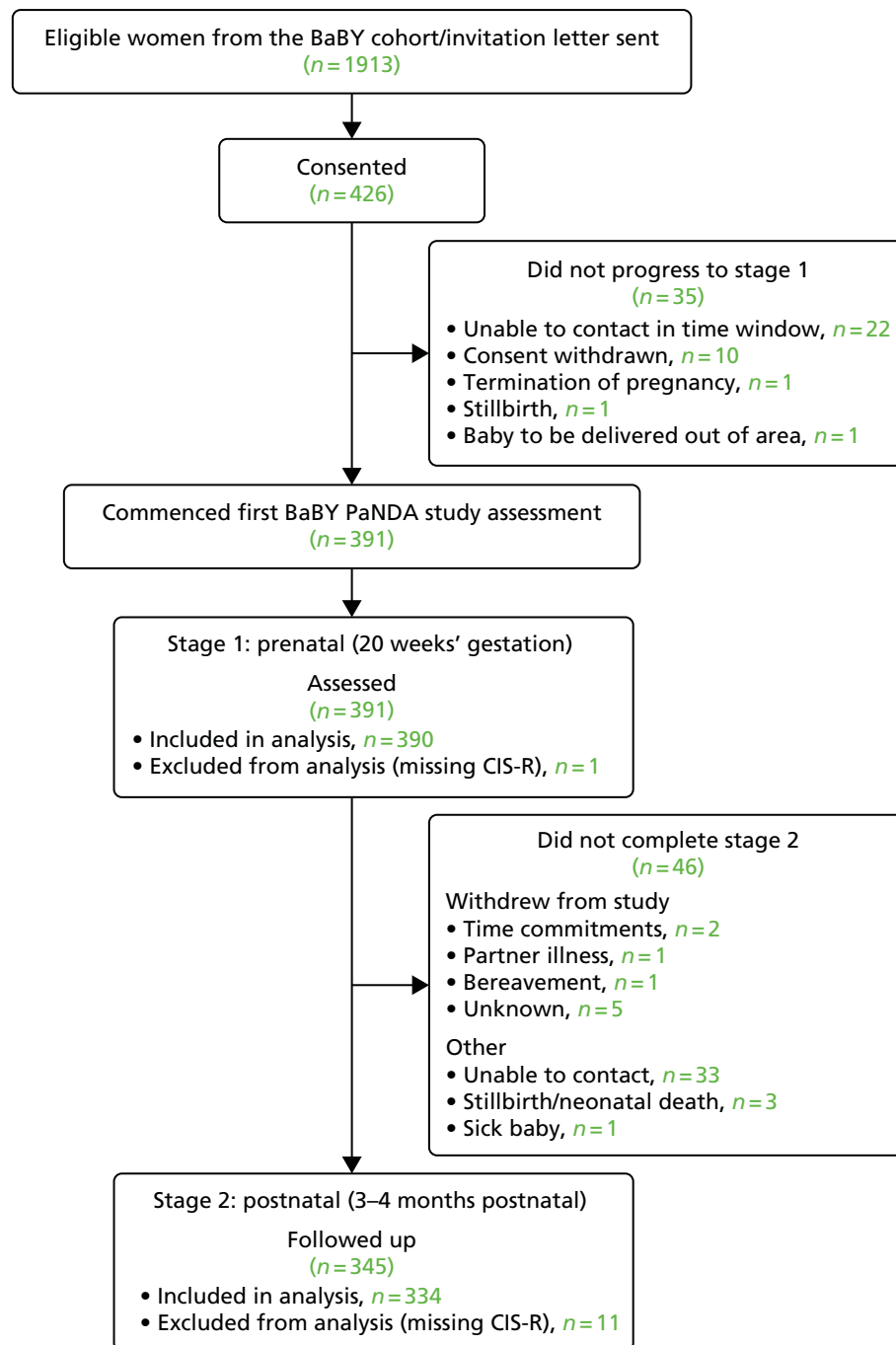
FIGURE 1 The BaBY PaNDA study flow diagram.

Representativeness of sample from within BaBY cohort

In order to establish the representativeness of the BaBY PaNDA study population, participant characteristics were compared with those of women in the full BaBY cohort (*Table 2*). BaBY cohort data (including GAD-7 anxiety, PHQ-8 depression and PHQ-15 somatic symptom severity) were collected at 26 weeks of gestation and at 8 weeks and at 1 year postnatally. Data at 26 weeks' gestation were treated as a proxy for comparable baseline characteristics. As the questionnaire completion rate within the BaBY cohort was low (e.g. only 270 of 390 mothers included in the BaBY PaNDA stage 1 analysis had valid GAD-7 data at 26 weeks of gestation), BaBY PaNDA data at stage 1 (20 weeks' gestation) were added to the comparisons for the outcomes of interest.

TABLE 1 Baseline characteristics of women included in stage 1 and 2 analyses

Characteristic at baseline	Women included in stage 1 (20 weeks' gestation) ($N = 390$)	Women included in stage 2 (3–4 months postnatally) ($N = 334$)
Site, n (%)		
York	187 (48)	162 (49)
Hull	95 (24)	75 (22)
Harrogate	66 (17)	66 (20)
Scunthorpe and Goole	42 (11)	31 (9)
Age at consent (years)		
n	389	334
Mean (SD)	31.2 (5.12)	31.8 (4.79)
Minimum, maximum	16, 46	16, 46
Missing	1	0
Ethnicity,[a] n (%)		
White	382 (98.0)	328 (98.2)
Other[b]	8 (2.1)	5 (1.5)
Missing	0 (0.0)	1 (0.3)
Education, n (%)		
No qualifications	7 (2)	5 (2)
GCSEs/O-levels	26 (7)	18 (5)
AS/A-levels	34 (9)	27 (8)
NVQ or vocational	84 (22)	62 (19)
Undergraduate degree	110 (28)	103 (31)
Postgraduate degree	88 (23)	80 (24)
Doctoral degree	11 (3)	11 (3)
Professional degree	24 (6)	23 (7)
Missing	6 (2)	5 (2)
Employment,[a] n (%)		
Working full time	215 (55.1)	193 (57.8)
Working part time	115 (29.5)	96 (28.7)
Student	10 (2.6)	9 (2.7)
Homemaker	38 (9.7)	31 (9.3)
Unemployed[c]	12 (3.1)	5 (1.5)
Marital status,[a] n (%)		
Married/living as married	345 (88.5)	304 (91.0)
Single	40 (10.3)	25 (7.5)
Separated/divorced[d]	5 (1.3)	4 (1.2)
Missing	0 (0.0)	1 (0.3)
Existing children, n (%)		
Yes	198 (51)	166 (50)
No	192 (49)	168 (50)

TABLE 1 Baseline characteristics of women included in stage 1 and 2 analyses (*continued*)

Characteristic at baseline	Women included in stage 1 (20 weeks' gestation) (*N* = 390)	Women included in stage 2 (3–4 months postnatally) (*N* = 334)
History of anxiety, *n* (%)		
Yes, currently	12 (3)	10 (3)
Yes, in the past	99 (25)	84 (25)
Yes, currently and in the past	26 (7)	21 (6)
No	231 (59)	199 (60)
Do not know	22 (6)	20 (6)
GAD-7 anxiety		
n	390	334
Mean (SD)	3.1 (3.56)	3.0 (3.38)
Minimum, maximum	0, 18	0, 18
History of depression, *n* (%)		
Yes, currently	7 (2)	4 (1)
Yes, in the past	113 (29)	92 (28)
Yes, currently and in the past	13 (3)	12 (4)
No	236 (61)	207 (62)
Do not know	20 (5)	18 (5)
Missing	1 (0.3)	1 (0.3)
PHQ-9 depression		
n	390	334
Mean (SD)	3.8 (3.97)	3.7 (3.84)
Minimum, maximum	0, 22	0, 22

A-level, Advanced level; GCSE, General Certificate of Secondary Education; NVQ, National Vocational Qualification; O-level, Ordinary level.
a Items extracted from CIS-R interview.
b Original data included separate categories for black, Asian, Chinese, mixed and other, but becaues of their low response frequencies have been combined here to safeguard participant anonymity.
c Original data included separate categories for jobseekers and inability to work due to ill health, but because of their low response frequencies have been combined here to safeguard participant anonymity.
d Original data included separate categories for separated and divorced, but because of their low response frequencies have been combined here to safeguard participant anonymity.

The women in the BaBY PaNDA study were on average 1.6 years older and had lower levels of anxiety, depression and somatic symptoms than women in the whole BaBY cohort. Differences between the populations at 26 weeks' gestation alone were more marked than comparing differences between the cohort data at 26 weeks' gestation and BaBY PANDA data at 20 weeks' gestation (stage 1). As BaBY completion rates were low, the estimates at 26 weeks' gestation are likely to be more biased towards the characteristics of those women who had good compliance (i.e. older women with better mental and physical outcomes).

Results at stage 1 (prenatal assessment: 20 weeks' gestation)

Outcome descriptives

Descriptive statistics for the study outcome measures are presented for all women who were included in the analysis at stage 1 [i.e. women with valid outcome data across all three assessment instruments (Whooley questions, EPDS and CIS-R); *n* = 390]. None of the reference standard (CIS-R) or index tests (Whooley questions or EPDS) had indeterminate results.

TABLE 2 Baseline characteristics from BaBY cohort data

Characteristic	BaBY cohort (n = 4160)	Women included in stage 1 (20 weeks' gestation) (n = 390)	Women included in stage 2 (3–4 months postnatally) (n = 334)		
Age at projected delivery date (years)					
n	4130	389	334		
Mean (SD)	30.0 (5.68)	31.6 (5.09)	32.3 (4.76)		
Minimum, maximum	15, 52	16, 46	16, 46		
Missing	30	1	0		
	Week 26 BaBY assessment	*Week 20 BaBY PaNDA assessment*	*Week 26 BaBY assessment*	*Week 20 BaBY PaNDA assessment*	*Week 26 BaBY assessment*
GAD-7 anxiety					
n	2024	390	270	334	239
Mean (SD)	3.4 (3.82)	3.1 (3.56)	2.8 (3.53)	3.0 (3.38)	2.8 (3.56)
Minimum, maximum	0, 21	0, 18	0, 19	0, 18	0, 19
PHQ-8 depression					
n	2059	390	277	334	246
Mean (SD)	4.3 (3.67)	3.8 (3.95)	3.5 (3.29)	3.7 (3.81)	3.4 (3.27)
Minimum, maximum	0, 24	0, 22	0, 22	0, 22	0, 22
PHQ-15 somatic symptoms					
n	2006	389	266	333	235
Mean (SD)	7.6 (4.16)	6.4 (3.84)	7.1 (4.04)	6.3 (3.76)	6.8 (3.92)
Minimum, maximum	0, 23	0, 20	0, 22	0, 20	0, 22

Diagnostic reference standard: Clinical Interview Schedule – Revised

As part of the diagnostic clinical assessment using the CIS-R, a primary diagnosis, if applicable, was recorded (*Table 3*). The overall prevalence of depression was 10.3% (40/390), consisting of 15 mild, 22 moderate and three severe cases. The most commonly recorded other primary diagnoses were mixed anxiety and depressive disorder (4.4%) and generalised anxiety disorder (2.8%). Profiles between study sites were similar. The CIS-R was not available for one woman (as the CIS-R could not be completed within 2 weeks of completion of the index tests).

TABLE 3 Clinical Interview Schedule – Revised primary diagnosis at stage 1 (20 weeks' gestation)

Diagnosis	Study site, n (%)				Total (N = 390), n (%)
	York (N = 187)	Hull (N = 95)	Harrogate (N = 66)	Scunthorpe (N = 42)	
No diagnosis identified	155 (82.9)	72 (75.8)	56 (84.8)	34 (81.0)	317 (81.3)
Mild, moderate or severe depressive episode	16 (8.6)	14 (14.7)	6 (9.1)	4 (9.5)	40 (10.3)
Generalised anxiety disorder	5 (2.7)	5 (5.3)	0 (0.0)	1 (2.4)	11 (2.8)
Panic disorder/social phobia/specific isolated phobia	2 (1.1)	2 (2.1)	0 (0.0)	1 (2.4)	5 (1.3)
Mixed anxiety and depressive disorder	9 (4.8)	2 (2.1)	4 (6.1)	2 (4.8)	17 (4.4)

Whooley questions

Using the case-finding threshold of at least one of the two Whooley questions being answered as true, 23.3% (91/390) of women were identified as depressed cases (*Table 4*). A greater number of women tended to respond positively to the first Whooley question (During the past month, have you often been bothered by feeling down, depressed or hopeless?) than the second Whooley question (During the past month, have you often been bothered by having little interest or pleasure in doing things?), which was particularly the case in Hull. Based on the additional help question (Is this something you feel you need or want help with?), of those 91 women who did reach the test threshold, eight (8.8%) indicated they would like help ('yes'), 65 (71.4%) indicated that they would like help but not today ('yes, but not today') and 18 (19.8%) said that they did not need or want help ('no').

Edinburgh Postnatal Depression Scale

The average EPDS scores by study site and their depression case-finding cut-off points are presented in *Table 5*. Using the ≥ 10-point case-finding threshold, 20.5% (80/390) of women were identified as being depressed, compared with only 8.5% (33/390) when using the ≥ 13-point threshold. Profiles between study sites were in line with the local differences identified for the diagnostic reference standard.

Diagnostic accuracy

Using the depression case-finding cut-off points of each instrument, 2 × 2 contingency tables of the index tests against the diagnostic reference standard are presented in *Table 6*. There were 40 true depression cases based on the CIS-R, of which 34 were identified by the Whooley questions, 33 by the EPDS (≥ 10-point cut-off point) and 18 by the EPDS (≥ 13-point cut-off point). The study was designed for 95% power on the basis of 50 true depression cases, and the available 40 cases at stage 1 resulted in a drop of statistical power to 90%.

The diagnostic accuracy statistics (*Table 7*) showed reasonable predictive power of the Whooley questions and the EPDS (≥ 10-point cut-off point). The Whooley questions displayed marginally better sensitivity and the EPDS better specificity. AUC values were approximately equal (*Figures 2* and *3*). The overidentification

TABLE 4 Whooley questions: descriptives at stage 1

Whooley question	Study site, *n* (%)				Total (*N* = 390), *n* (%)
	York (*N* = 187)	Hull (*N* = 95)	Harrogate (*N* = 66)	Scunthorpe (*N* = 42)	
Question 1: feeling down, depressed or hopeless	26 (13.9)	28 (29.5)	11 (16.7)	8 (19.0)	73 (18.7)
Question 2: having little interest or pleasure in doing things	24 (12.8)	17 (17.9)	8 (12.1)	5 (11.9)	54 (13.8)
At least one question true	35 (18.7)	31 (32.6)	14 (21.2)	11 (26.2)	91 (23.3)
Both questions true	15 (8.0)	14 (14.7)	5 (7.6)	2 (4.8)	36 (9.2)

TABLE 5 Edinburgh Postnatal Depression Scale: descriptives at stage 1

EPDS score	Study site				Total (*N* = 390)
	York (*N* = 187)	Hull (*N* = 95)	Harrogate (*N* = 66)	Scunthorpe (*N* = 42)	
Mean (SD)	5.8 (4.10)	6.3 (5.59)	4.6 (4.19)	5.7 (4.58)	5.7 (4.58)
Median	5	5	3	5	5
Minimum, maximum	0, 19	0, 22	0, 20	0, 18	0, 22
EPDS score of ≥ 10, *n* (%)	36 (19.3)	26 (27.4)	10 (15.2)	8 (19.0)	80 (20.5)
EPDS score of ≥ 13, *n* (%)	13 (7.0)	13 (13.7)	3 (5.5)	4 (9.5)	33 (8.5)

TABLE 6 Contingency tables for index test tools and reference standard at stage 1

Case-finding instrument	CIS-R depressed (N = 40), n (%)	CIS-R non-depressed (N = 350), n (%)	Total (N = 390), n (%)
Whooley: positive screen	34 (85)	57 (16)	91 (23)
Whooley: negative screen	6 (15)	293 (84)	299 (77)
EPDS score of ≥ 10	33 (83)	47 (13)	80 (21)
EPDS score of < 10	7 (18)	303 (87)	310 (79)
EPDS score of ≥ 13	18 (45)	15 (4)	33 (8)
EPDS score of < 13	22 (55)	335 (96)	357 (92)

TABLE 7 Diagnostic accuracy statistics at stage 1

Accuracy statistic	Whooley questions (95% CI)	EPDS (≥ 10-point cut-off point) (95% CI)	EPDS (≥ 13-point cut-off point) (95% CI)
Sensitivity	85.0 (70.2 to 94.3)	82.5 (67.2 to 92.7)	45.0 (29.3 to 61.5)
Specificity	83.7 (79.4 to 87.4)	86.6 (82.5 to 90.0)	95.7 (93.0 to 97.6)
PPV	37.4 (27.4 to 48.1)	41.3 (30.4 to 52.8)	54.5 (36.4 to 71.9)
NPV	98.0 (95.7 to 99.3)	97.7 (95.4 to 99.1)	93.8 (90.8 to 96.1)
PLR	5.22 (3.98 to 6.84)	6.14 (4.54 to 8.31)	10.50 (5.75 to 19.17)
NLR	0.18 (0.09 to 0.38)	0.20 (0.10 to 0.40)	0.57 (0.43 to 0.76)
Diagnostic odds ratio	29.13 (11.95 to 70.72)	30.39 (12.94 to 71.12)	18.27 (8.21 to 40.73)
AUC	0.84 (0.78 to 0.90)	0.85 (0.78 to 0.91)	0.70 (0.62 to 0.78)
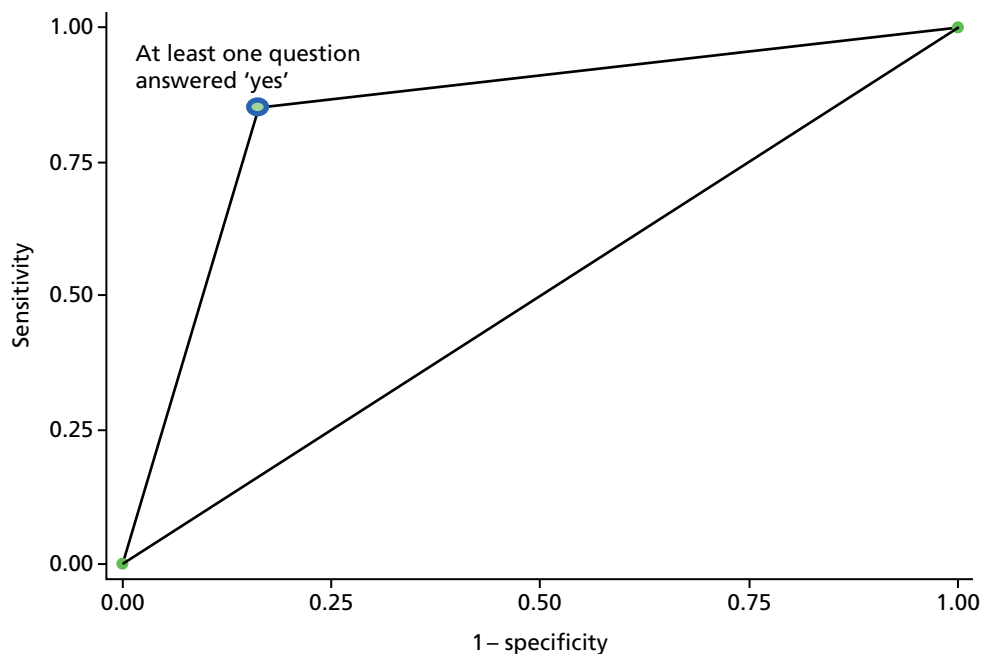
FIGURE 2 Receiver operating characteristics curve for Whooley questions at stage 1.
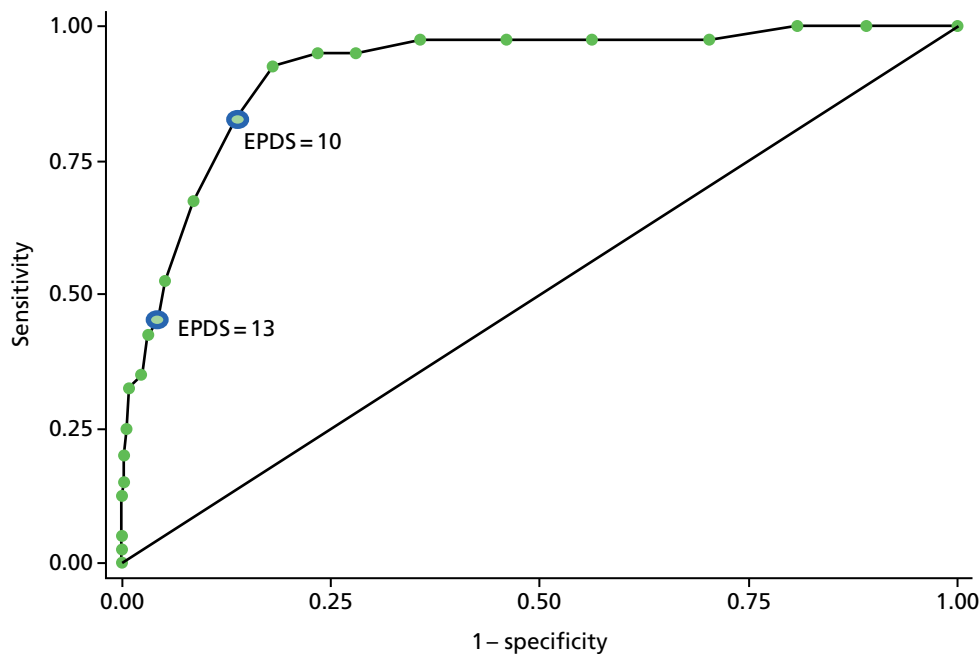
FIGURE 3 Receiver operating characteristics curve for EPDS at stage 1.

of cases by both case-finding instruments was reflected in relatively low PPVs. The EPDS (≥ 13-point cut-off point) was found to be a very conservative case-finding threshold with very low sensitivity (even when considering only moderate and severe depression cases as the reference threshold in an exploratory analysis, sensitivity only improved to 52.0%). The ≥ 10-point cut-off point was confirmed as the optimum EPDS case-finding threshold based on the ROC curve (see *Figure 3*).

Comorbidities

Descriptive statistics for concurrently collected mental and physical outcomes are presented for test-positive (depressed) cases in *Table 8* and for test-negative (non-depressed) cases in *Table 9* based on the case thresholds of each instrument.

Depression caseness was associated with higher PHQ-9 depression severity scores, higher levels of GAD-7 anxiety and poorer SF-12v2 MCS. In addition, worse PHQ-15 somatic symptom scores were observed for depression cases, whereas the SF-12v2 PCS were similar to those of non-cases.

Results at stage 2 (postnatal assessment: 3–4 months postnatal)

Outcome descriptives

Descriptive statistics for the study outcome measures are presented for all women who were included in the analysis at stage 2 [i.e. women with valid outcome data across all three instruments (Whooley questions, EPDS and CIS-R); *n* = 334]. None of the reference standard or index test assessments had indeterminate results.

Diagnostic reference standard: Clinical Interview Schedule – Revised

The CIS-R primary diagnosis at stage 2 is recorded in *Table 10*. The prevalence of depression was 10.5% (35/334), consisting of 13 mild, 20 moderate and two severe depression cases. The most commonly recorded other primary diagnosis was mixed anxiety and depressive disorder (5.4%). The CIS-R was not available for 11 women (as the CIS-R could not be completed within 2 weeks of completion of the index tests).

TABLE 8 Comorbidities for positive depression cases at stage 1

| Outcome | Positive depression status | | | |
	CIS-R	Whooley questions	EPDS (≥ 10-point cut-off point)	EPDS (≥ 13-point cut-off point)
PHQ-9				
n	40	91	80	33
Mean (SD)	11.7 (4.69)	8.0 (4.89)	8.9 (4.74)	11.9 (4.69)
Median (minimum, maximum)	12.5 (3, 22)	7 (1, 22)	8 (2, 22)	12 (5, 22)
PHQ-9 depression group, n (%)				
No	3 (8)	24 (26)	12 (15)	0 (0)
Mild	11 (28)	37 (41)	40 (50)	13 (39)
Moderate	14 (35)	18 (20)	16 (20)	9 (27)
Moderately severe	10 (25)	10 (11)	10 (13)	9 (27)
Severe	2 (5)	2 (2)	2 (3)	2 (6)
GAD-7				
n	40	91	80	33
Mean (SD)	8.7 (3.79)	6.3 (4.01)	7.3 (3.91)	8.7 (4.28)
Median (minimum, maximum)	8 (1, 17)	6 (0, 17)	6.5 (0, 18)	8 (0, 17)
GAD-7 anxiety group, n (%)				
No	2 (5)	30 (33)	16 (20)	5 (15)
Mild	22 (55)	41 (45)	42 (53)	14 (42)
Moderate	13 (33)	17 (19)	18 (23)	11 (33)
Severe	3 (8)	3 (3)	4 (5)	3 (9)
SF-12v2: PCS				
n	40	91	80	33
Mean (SD)	44.5 (9.91)	46.1 (9.37)	46.0 (9.44)	47 (9.12)
Median (minimum, maximum)	46 (16, 63)	46 (16, 63)	47 (16, 63)	50 (30, 63)
SF-12v2: MCS				
n	40	91	80	33
Mean (SD)	33.5 (8.15)	41.1 (10.80)	38.7 (9.68)	33.0 (9.92)
Median (minimum, maximum)	34 (18, 53)	41 (18, 65)	39 (18, 59)	32 (18, 59)
PHQ-15				
n	40	91	80	33
Mean (SD)	11.9 (3.95)	9.3 (4.37)	9.8 (3.85)	10.8 (4.12)
Median (minimum, maximum)	12 (3, 19)	9 (0, 20)	10 (0, 19)	11 (4, 19)
Symptomatic severity, n (%)				
No somatic symptoms	1 (3)	11 (12)	5 (6)	2 (6)
Low severity	9 (23)	38 (42)	33 (41)	9 (27)
Medium severity	20 (50)	32 (35)	33 (41)	16 (48)
High severity	10 (25)	10 (11)	9 (11)	6 (18)

TABLE 9 Comorbidities for negative depression cases at stage 1

Outcome	Negative depression status			
	CIS-R	Whooley questions	EPDS (≥ 10-point cut-off point)	EPDS (≥ 13-point cut-off point)
PHQ-9				
n	350	299	310	357
Mean (SD)	2.9 (2.68)	2.5 (2.51)	2.5 (2.43)	3.1 (2.94)
Median (minimum, maximum)	2 (0, 15)	2 (0, 15)	2 (0, 15)	2 (0, 15)
PHQ-9 depression group, *n* (%)				
No	272 (78)	251 (84)	263 (85)	275 (77)
Mild	68 (19)	42 (14)	39 (13)	66 (18)
Moderate	9 (3)	5 (2)	7 (2)	14 (4)
Moderately severe	1 (0.3)	1 (0.3)	1 (0.3)	2 (0.6)
Severe	0 (0)	0 (0)	0 (0)	0 (0)
GAD-7				
n	346	295	306	353
Mean (SD)	2.4 (2.68)	2.0 (2.43)	1.9 (2.17)	2.5 (2.78)
Median (minimum, maximum)	2 (0, 18)	1 (0, 18)	1 (0, 12)	2 (0, 18)
GAD-7 anxiety group, *n* (%)				
No	284 (82)	256 (87)	270 (88)	281 (80)
Mild	53 (15)	34 (12)	33 (11)	61 (17)
Moderate	8 (2)	4 (1)	3 (1)	10 (3)
Severe	1 (0.3)	1 (0.3)	0 (0)	1 (0.3)
SF-12v2: PCS				
n	344	293	304	351
Mean (SD)	49.7 (7.64)	50.1 (7.35)	50.0 (7.43)	49.3 (7.94)
Median (minimum, maximum)	52 (13, 62)	52 (13, 62)	52 (13, 62)	51 (13, 62)
SF-12v2: MCS				
n	344	293	304	351
Mean (SD)	52.5 (7.01)	53.5 (6.24)	53.7 (5.99)	52.2 (7.22)
Median (minimum, maximum)	54 (26, 67)	55 (26, 67)	55 (26, 67)	54 (23, 67)
PHQ-15				
n	345	294	305	352
Mean (SD)	5.8 (3.28)	5.6 (3.18)	5.6 (3.30)	6.0 (3.54)
Median (minimum, maximum)	6 (0, 20)	5 (0, 17)	5 (0, 20)	6 (0, 20)
Symptomatic severity, *n* (%)				
No somatic symptoms	126 (37)	116 (39)	122 (40)	125 (36)
Low severity	174 (50)	145 (49)	150 (49)	174 (49)
Medium severity	42 (12)	30 (10)	29 (10)	46 (13)
High severity	3 (0.9)	3 (1)	4 (1)	7 (2)

TABLE 10 Clinical Interview Schedule – Revised primary diagnosis at stage 2 (3–4 months postnatal)

Diagnosis	Study site, *n* (%)				Total (*N* = 334), *n* (%)
	York (*N* = 162)	Hull (*N* = 75)	Harrogate (*N* = 66)	Scunthorpe (*N* = 31)	
No diagnosis identified	133 (82.1)	54 (72.0)	60 (90.9)	27 (87.1)	274 (82.0)
Mild, moderate or severe depressive episode	18 (11.1)	11 (14.7)	4 (6.1)	2 (6.5)	35 (10.5)
Generalised anxiety disorder	1 (0.6)	0 (0.0)	1 (1.5)	0 (0.0)	2 (0.6)
Agoraphobia/specific isolated phobia	4 (2.5)	1 (1.3)	0 (0.0)	0 (0.0)	5 (1.5)
Mixed anxiety and depressive disorder	6 (3.7)	9 (12.0)	1 (1.5)	2 (6.5)	18 (5.4)

Whooley questions

Using the case-finding threshold of at least one of the two Whooley questions being answered as true, 26.3% (88/334) of women were identified as depressed cases (*Table 11*). The majority of depression cases were identified by the first Whooley question alone, which was consistent across study sites. Based on the additional help question, of those 88 women who did reach the test threshold, 16 (18.2%, more than at stage 1) indicated they would like help, 57 (64.8%) indicated that they would like help but not today, and 15 (17.1%) said that they did not need or want help.

Edinburgh Postnatal Depression Scale

The average EPDS scores by study site and their depression case-finding cut-off points are presented in *Table 12*. Using the ≥ 10-point case-finding threshold, 19.8% (66/334) of women were identified as being depressed, compared with only 10.5% (35/334) when using the ≥ 13-point threshold. Profiles between study sites were in line with the local differences identified for the diagnostic reference standard.

TABLE 11 Whooley questions: descriptives at stage 2

Whooley question	Study site, *n* (%)				Total (*N* = 334), *n* (%)
	York (*N* = 162)	Hull (*N* = 75)	Harrogate (*N* = 66)	Scunthorpe (*N* = 31)	
Question 1: feeling down, depressed or hopeless	40 (24.7)	22 (29.3)	12 (18.2)	11 (35.5)	85 (25.4)
Question 2: having little interest or pleasure in doing things	12 (7.4)	8 (10.7)	6 (9.1)	4 (12.9)	30 (9.0)
At least one question true	40 (24.7)	25 (33.3)	12 (18.2)	11 (35.5)	88 (26.3)
Both questions true	12 (7.4)	5 (6.7)	6 (9.1)	4 (12.9)	27 (8.1)

TABLE 12 Edinburgh Postnatal Depression Scale: descriptives at stage 2

EPDS score	Study site				Total (*N* = 334)
	York (*N* = 162)	Hull (*N* = 75)	Harrogate (*N* = 66)	Scunthorpe (*N* = 31)	
Mean (SD)	5.9 (4.55)	6.1 (5.45)	4.8 (4.67)	5.9 (5.16)	5.7 (4.84)
Median	5	4	3.5	5	5
Minimum, maximum	0, 22	0, 19	0, 21	0, 21	0, 22
EPDS score of ≥ 10 (%)	30 (18.5)	20 (26.7)	10 (15.2)	6 (19.4)	66 (19.8)
EPDS score of ≥ 13 (%)	16 (9.9)	11 (14.7)	5 (7.6)	3 (9.7)	35 (10.5)

Diagnostic accuracy

Using the depression case-finding cut-off points of each case-finding instrument, 2 × 2 contingency tables of the index test tools against the reference standard are presented in *Table 13*. Of the 334 mothers with valid outcome data across all three instruments, there were 35 true depression cases based on the CIS-R, of which 30 were identified by the Whooley questions, 29 by the EPDS (≥ 10-point cut-off point) and 22 by the EPDS (≥ 13-point cut-off point). The study was designed for 95% power on the basis of 50 true depression cases, and the available 35 cases at stage 2 resulted in a drop of statistical power to 85%.

Similar to stage 1, the diagnostic accuracy statistics (*Table 14*) showed reasonable predictive power of the Whooley questions and the EPDS (≥ 10-point cut-off point). The Whooley questions displayed marginally better sensitivity and the EPDS better specificity. AUC values were approximately equal (*Figures 4* and *5*). The overidentification of cases by both instruments was reflected in relatively low PPVs. The EPDS (≥ 13-point cut-off point) was again found to be a very conservative case-finding threshold with very low sensitivity (in contrast to stage 1, sensitivity could be improved to 81.8% when only considering moderate and severe depression cases as the reference threshold in an exploratory analysis; however, the optimum cut-off point for this outcome was still lower at 12 score points). For the existing reference standard of mild, moderate or severe depression, the ≥ 10-point cut-off point was confirmed as the optimum EPDS case-finding threshold based on the ROC curve (see *Figure 5*).

TABLE 13 Contingency tables for index test tools and reference standard at stage 2

Case-finding instrument	CIS-R depressed (*N* = 35), *n* (%)	CIS-R non-depressed (*N* = 299), *n* (%)	Total (*N* = 334), *n* (%)
Whooley: positive screen	30 (86)	58 (19)	88 (26)
Whooley: negative screen	5 (14)	241 (81)	246 (74)
EPDS score of ≥ 10	29 (83)	37 (12)	66 (20)
EPDS score of < 10	6 (17)	262 (88)	268 (80)
EPDS score of ≥ 13	22 (63)	13 (4)	35 (10)
EPDS score of < 13	13 (37)	286 (96)	299 (90)

TABLE 14 Diagnostic accuracy statistics at stage 2

Accuracy statistic	Whooley questions (95% CI)	EPDS (≥ 10-point cut-off point) (95% CI)	EPDS (≥ 13-point cut-off point) (95% CI)
Sensitivity	85.7 (69.7 to 95.2)	82.9 (66.4 to 93.4)	62.9 (44.9 to 78.5)
Specificity	80.6 (75.7 to 84.9)	87.6 (83.3 to 91.1)	95.7 (92.7 to 97.7)
PPV	34.1 (24.3 to 45.0)	43.9 (31.7 to 56.7)	62.9 (44.9 to 78.5)
NPV	98.0 (95.3 to 99.3)	97.8 (95.2 to 99.2)	95.7 (92.7 to 97.7)
PLR	4.42 (3.38 to 5.78)	6.70 (4.78 to 9.38)	14.46 (8.02 to 26.07)
NLR	0.18 (0.08 to 0.40)	0.20 (0.09 to 0.41)	0.39 (0.25 to 0.60)
Diagnostic odds ratio	24.93 (9.54 to 6.84)	34.23 (13.60 to 85.67)	37.23 (15.55 to 89.26)
AUC	0.83 (0.77 to 0.89)	0.85 (0.79 to 0.92)	0.79 (0.71 to 0.87)
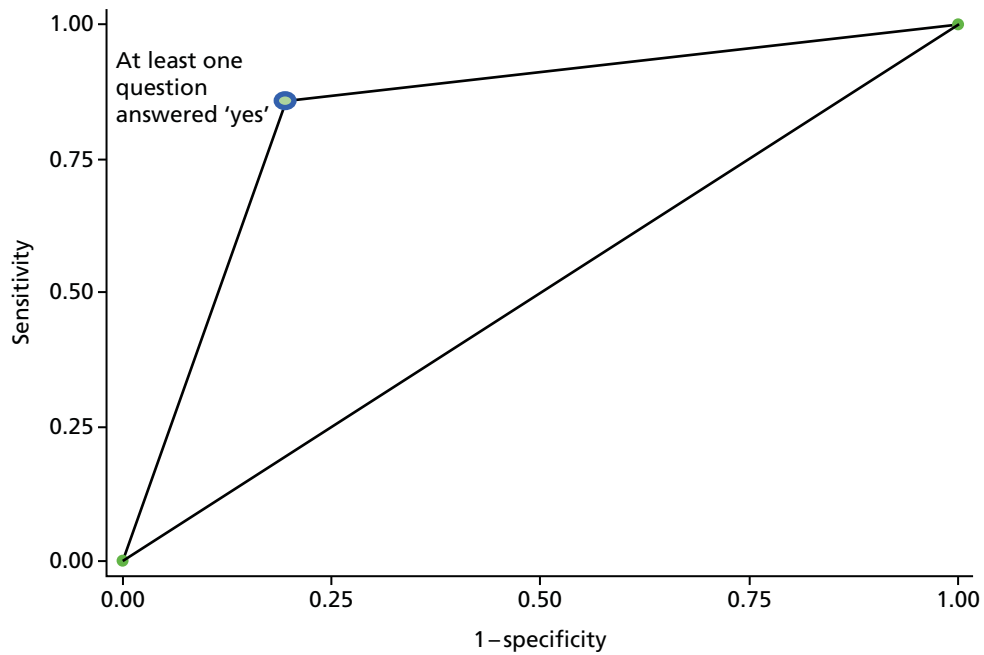
FIGURE 4 Receiver operating characteristics curve for Whooley questions at stage 2.
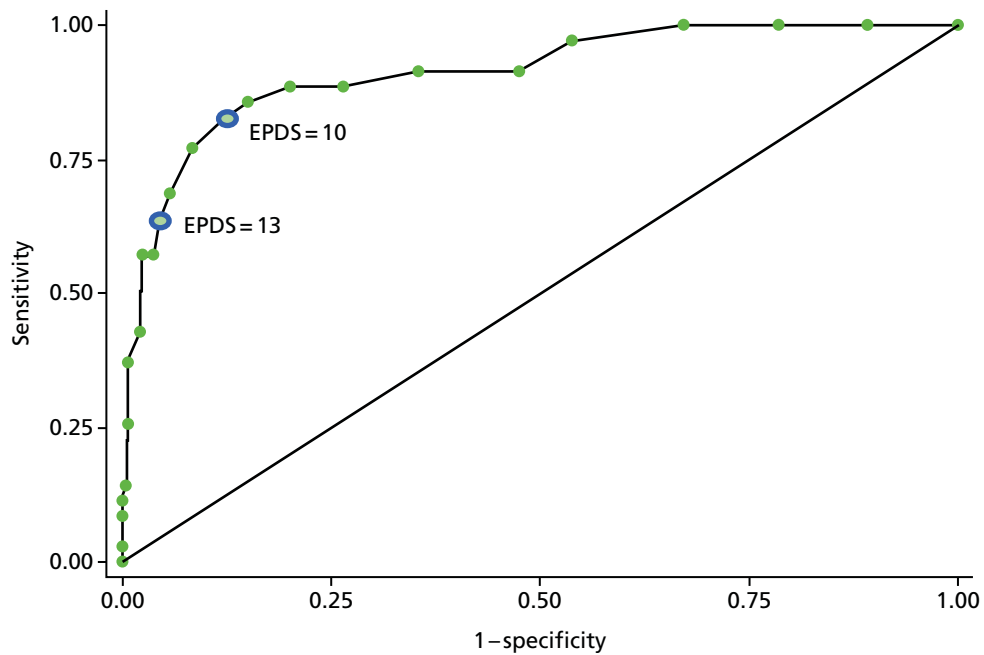
FIGURE 5 Receiver operating characteristics curve for EPDS at stage 2.

Comorbidities

Descriptive statistics for concurrently collected mental and physical outcomes are presented for test-positive (depressed) cases in *Table 15* and for test-negative (non-depressed) cases in *Table 16* based on the test thresholds of each instrument.

Depression caseness was associated with higher PHQ-9 depression severity scores, higher levels of GAD-7 anxiety and poorer SF-12v2 MCS. In addition, worse PHQ-15 somatic symptom scores were observed for depression cases, whereas the SF-12v2 PCS were similar to those of non-cases.

TABLE 15 Comorbidities for positive depression cases at stage 2

| Outcome | Positive depression status | | | |
	CIS-R	Whooley	EPDS ($\geq$ 10-point cut-off point)	EPDS ($\geq$ 13-point cut-off point)
PHQ-9				
n	34	75	66	27
Mean (SD)	8.1 (5.52)	5.9 (5.30)	6.9 (5.29)	7.2 (5.98)
Median (minimum, maximum)	6 (0, 18)	4 (0, 18)	6 (0,18)	6 (0, 18)
PHQ-9 depression group, *n* (%)				
No	12 (35)	41 (55)	28 (42)	11 (41)
Mild	7 (21)	16 (21)	19 (29)	6 (22)
Moderate	10 (29)	11 (15)	12 (18)	6 (22)
Moderately severe	5 (15)	7 (9)	7 (11)	4 (15)
Severe	0 (0)	0 (0)	0 (0)	0 (0)
GAD-7				
n	32	67	60	27
Mean (SD)	7.6 (5.73)	6.0 (5.25)	6.7 (5.15)	7.4 (6.36)
Median (minimum, maximum)	6 (0, 20)	5 (0, 20)	5 (0, 20)	6 (0, 20)
GAD-7 anxiety group, *n* (%)				
No	10 (31)	33 (49)	24 (40)	12 (44)
Mild	11 (34)	18 (27)	21 (35)	6 (22)
Moderate	7 (22)	12 (18)	10 (17)	5 (19)
Severe	4 (13)	4 (6)	5 (8)	4 (15)
SF-12v2: PCS				
n	31	66	59	26
Mean (SD)	50.4 (9.33)	52.5 (9.89)	51.7 (10.11)	51.2 (11.82)
Median (minimum, maximum)	53 (28, 65)	56 (16, 66)	54 (16, 66)	54 (16, 66)
SF-12v2: MCS				
n	31	66	59	26
Mean (SD)	39.6 (11.25)	42.9 (12.33)	41.4 (12.08)	39.6 (12.95)
Median (minimum, maximum)	41 (18, 57)	46 (18, 60)	42 (18, 61)	38 (18, 60)
PHQ-15				
n	32	67	60	27
Mean (SD)	7.4 (4.47)	6.4 (4.38)	6.9 (4.28)	7.0 (4.73)
Median (minimum, maximum)	6.5 (0, 16)	5 (0, 17)	6 (0, 17)	6 (0, 16)
Symptomatic severity, *n* (%)				
No somatic symptoms	7 (22)	24 (36)	14 (23)	7 (26)
Low severity	14 (44)	25 (37)	30 (50)	12 (44)
Medium severity	9 (28)	14 (21)	12 (20)	5 (19)
High severity	2 (6)	4 (6)	4 (7)	3 (11)

TABLE 16 Comorbidities for negative depression cases at stage 2

| Outcome | Negative depression status | | | |
	CIS-R	Whooley	EPDS (≥ 10-point cut-off point)	EPDS (≥ 13-point cut-off point)
PHQ-9				
n	310	269	278	317
Mean (SD)	3.1 (3.53)	2.9 (3.34)	2.8 (3.24)	3.3 (3.69)
Median (minimum, maximum)	2 (0, 19)	2 (0, 19)	2 (0, 19)	2 (0, 19)
PHQ-9 depression group, n (%)				
No	242 (78)	213 (79)	226 (81)	243 (77)
Mild	52 (17)	43 (16)	40 (14)	53 (17)
Moderate	10 (3)	9 (3)	8 (3)	14 (4)
Moderately severe	6 (2)	4 (1)	4 (1)	7 (2)
Severe	0 (0)	0 (0)	0 (0)	0 (0)
GAD-7				
n	278	243	250	283
Mean (SD)	2.7 (3.37)	2.4 (3.12)	2.4 (3.08)	2.8 (3.41)
Median (minimum, maximum)	2 (0, 18)	2 (0, 18)	2 (0, 18)	2 (0, 18)
GAD-7 anxiety group, n (%)				
No	229 (82)	206 (85)	215 (86)	227 (80)
Mild	35 (13)	28 (12)	25 (10)	40 (14)
Moderate	9 (3)	4 (2)	6 (2)	11 (4)
Severe	5 (2)	5 (2)	4 (2)	5 (2)
SF-12v2: PCS				
n	277	242	249	282
Mean (SD)	53.8 (7.79)	53.7 (7.41)	53.8 (7.39)	53.6 (7.55)
Median (minimum, maximum)	56 (11, 66)	56 (11, 66)	56 (11, 64)	56 (11, 66)
SF-12v2: MCS				
n	277	242	249	282
Mean (SD)	49.8 (9.20)	50.4 (8.46)	50.5 (8.43)	49.6 (9.13)
Median (minimum, maximum)	53 (18, 65)	53 (19, 65)	53 (18, 65)	52 (18, 65)
PHQ-15				
n	278	243	250	283
Mean (SD)	4.1 (3.19)	3.9 (2.98)	3.8 (2.97)	4.2 (3.24)
Median (minimum, maximum)	3.5 (0, 17)	3 (0, 17)	3 (0, 17)	4 (0, 17)
Symptomatic severity, n (%)				
No somatic symptoms	171 (62)	154 (63)	164 (66)	171 (60)
Low severity	86 (31)	75 (31)	70 (28)	88 (31)
Medium severity	18 (6)	13 (5)	15 (6)	22 (8)
High severity	3 (1)	1 (0.4)	1 (0.4)	2 (0.7)

Missing data

Of all women assessed at stage 1, there were 58 (14.8% of 391) with missing data for the diagnostic reference standard (CIS-R) or either of the index test assessments (Whooley questions or EPDS) either at stage 1 or at stage 2. These were either 'lost to follow-up' between time points, mainly because they could not be contacted to arrange the assessment or did not respond to postal questionnaires (see *Figure 1*), or did not complete the CIS-R as part of their assessment (as this could not be completed within 2 weeks of completing the index tests).

Baseline characteristics for women with complete and incomplete data (*Table 17*) revealed that women with missing data were more likely to be younger, less well educated, not working full time and single. They also tended to have marginally higher anxiety and depression levels. Notably, all women from the Harrogate study site had complete assessment data across stages 1 and 2.

TABLE 17 Baseline characteristics of women with/without missing data

Characteristic	Women with complete data at stages 1 and 2 (*N* = 333), *n* (%)	Women with incomplete data at stages 1 or 2 (*N* = 58), *n* (%)
Site, *n* (%)		
York	161 (48.4)	27 (46.6)
Hull	75 (22.5)	20 (34.5)
Harrogate	66 (19.8)	0 (0.0)
Scunthorpe and Goole	31 (9.3)	11 (19.0)
Age at consent (years)		
n	333	57
Mean (SD)	31.8 (4.80)	27.4 (5.23)
Minimum, maximum	16, 46	18, 39
Missing	0	1
Ethnicity,[a] *n* (%)		
White	328 (98.5)	54 (93.1)
Other[b]	5 (1.5)	3 (5.2)
Missing	0 (0.0)	1 (1.7)
Education, *n* (%)		
No qualifications	5 (1.5)	2 (3.5)
GCSEs/O-levels	18 (5.4)	8 (13.8)
AS/A-levels	27 (8.1)	7 (12.1)
NVQ or vocational	61 (18.3)	24 (41.4)
Undergraduate degree	103 (30.9)	7 (12.1)
Postgraduate degree	80 (24.0)	8 (13.8)
Doctoral degree	11 (3.3)	0 (0.0)
Professional degree	23 (6.9)	1 (1.7)
Missing	5 (1.5)	1 (1.7)
Employment,[a] *n* (%)		
Working full time	193 (58.0)	22 (37.9)
Working part time	96 (28.3)	19 (32.8)

TABLE 17 Baseline characteristics of women with/without missing data (*continued*)

Characteristic	Women with complete data at stages 1 and 2 ($N = 333$), n (%)	Women with incomplete data at stages 1 or 2 ($N = 58$), n (%)
Student	9 (2.7)	1 (1.7)
Homemaker	31 (9.3)	7 (12.1)
Unemployed[c]	4 (1.2)	8 (13.8)
Missing	0 (0.0)	1 (1.7)
Marital status,[a] n (%)		
Married/living as married	304 (91.3)	41 (70.7)
Single	25 (7.5)	15 (25.9)
Separated/divorced[d]	4 (1.2)	1 (1.7)
Missing	0 (0.0)	1 (1.7)
Existing children, n (%)		
Yes	165 (49.6)	34 (58.6)
No	168 (50.5)	24 (41.4)
History of anxiety, n (%)		
Yes, currently	10 (3.0)	2 (3.5)
Yes, in the past	83 (24.9)	17 (29.3)
Yes, currently and in the past	21 (6.3)	5 (8.6)
No	199 (59.8)	32 (55.2)
Do not know	20 (6.0)	2 (3.5)
GAD-7 anxiety		
n	333	58
Mean (SD)	3.0 (3.38)	4.0 (4.39)
Minimum, maximum	0, 18	0, 18
History of depression, n (%)		
Yes, currently	4 (1.2)	3 (5.2)
Yes, in the past	92 (27.6)	21 (36.2)
Yes, currently and in the past	12 (3.6)	1 (1.7)
No	206 (61.9)	31 (53.5)
Do not know	18 (5.4)	2 (3.5)
Missing	1 (0.3)	0 (0.0)
PHQ-9 depression		
n	333	58
Mean (SD)	3.7 (3.85)	4.6 (4.57)
Minimum, maximum	0, 22	0, 17

A-level, Advanced level; GCSE, General Certificate of Secondary Education; NVQ, National Vocational Qualification; O-level, Ordinary level.
a Items extracted from CIS-R interview.
b Original data included separate categories for black, Asian, Chinese, mixed and other, but because of their low response frequencies have been combined here to safeguard participant anonymity.
c Original data included separate categories for jobseekers and inability to work owing to ill health, but because of their low response frequencies have been combined here to safeguard participant anonymity.
d Original data included separate categories for separated and divorced, but because of their low response frequencies have been combined here to safeguard participant anonymity.

Temporal stability

Depression assessment according to each index test tool was compared at stages 1 and 2 for the same women. The resulting cross-tabulation is presented in *Table 18*, including the test results for statistical differences between time points (McNemar's test).

Approximately half of depression cases at stage 1 became non-cases at stage 2, which was evident across both case-finding instruments. Of the non-cases at stage 1, between 7% and 19% were identified as having become depressed at stage 2, depending on the case-finding instrument. When using the Whooley questions, a larger proportion of women who had not been depressed at stage 1 were identified as depressed at stage 2, compared with other case-finding instruments. However, there was no evidence for statistically significant differences between the two assessment points for any of the assessments.

Chapter summary

Across the four study sites, 390 pregnant women completed the reference standard and index tests at 20 weeks' gestation (stage 1), and 334 completed these at 3–4 months postnatally (stage 2). Attrition between stages 1 and 2 (15%) was less than anticipated for the sample size (34%). Women with missing outcomes at either assessment were more likely to be younger and less well educated and to have marginally higher anxiety and depression levels.

The prevalence of depression based on the CIS-R reference standard was approximately 10% at either stage, lower than anticipated. Depression prevalence was higher in Hull than at other study sites.

TABLE 18 Temporal stability of depression caseness between stages 1 and 2

| Tool | Follow-up at 20 weeks' gestation and 3–4 months postnatally, n (% of 391 total) | Assessment at 20 weeks' gestation, n | | Assessment at 3–4 months postnatally, n (% of those at 20 weeks' gestation) | McNemar's test | |
					χ^2	*p*-value
CIS-R	333 (85.3)	Cases	32	Cases	15 (46.9)	0.24	0.622
				Non-cases	17 (53.1)		
		Non-cases	301	Cases	20 (6.6)		
				Non-cases	281 (93.4)		
Whooley	345 (88.2)	Cases	76	Cases	41 (54.0)	2.98	0.085
				Non-cases	35 (46.0)		
		Non-cases	269	Cases	51 (19.0)		
				Non-cases	218 (81.0)		
EPDS score of ≥ 10	345 (88.2)	Cases	67	Cases	30 (44.8)	0.05	0.819
				Non-cases	37 (55.2)		
		Non-cases	278	Cases	39 (14.0)		
				Non-cases	239 (86.0)		
EPDS score of ≥ 13	345 (88.2)	Cases	27	Cases	12 (44.4)	2.08	0.150
				Non-cases	15 (55.6)		
		Non-cases	318	Cases	24 (7.6)		
				Non-cases	294 (92.4)		

Diagnostic accuracy tests were conducted for the Whooley questions and the EPDS (using two cut-off scores suggested by the literature) against the CIS-R reference standard. Although prevalence was lower than expected, the analyses retained 90% statistical power at stage 1 and 85% power at stage 2.

At both assessment time points, the Whooley questions and EPDS ($\geq$ 10-point cut-off point) were found to be reasonable case-finding instruments with comparable sensitivity (stage 1 Whooley questions 85.0%, EPDS 82.5%; stage 2 Whooley questions 85.7%, EPDS 82.9%) and specificity (stage 1 Whooley questions 83.7%, EPDS 86.6%; stage 2 Whooley questions 80.6%, EPDS: 87.6%). The diagnostic accuracy of the EPDS ($\geq$ 13-point cut-off point) was poor, even when considering only women with moderate to severe depression. Risk of bias related to interpretation of the results of the index tests and diagnostic reference standard was considered minimal (see *Appendix 4*).

Depression caseness was found to be stable between stages 1 and 2 for most women with no statistically significant difference between stages. Approximately half of depressed women at stage 1 became non-cases at stage 2, and around 7% of non-depressed women at stage 1 became depressed cases at stage 2. This pattern was similar across case-finding instruments, apart from the Whooley questions and EPDS ($\geq$ 10-point cut-off point), which picked up more new cases at stage 2.

Depression caseness based on any of the case-finding instruments was associated with higher PHQ-9 depression severity scores, higher levels of GAD-7 anxiety and poorer SF-12v2 MCS. In addition, worse PHQ-15 somatic symptom scores were observed for depression cases, whereas the SF-12v2 PCS were similar to those of non-cases.

Chapter 6 Evaluation of acceptability

The aim of the EoA was to investigate the acceptability and impact of the Whooley questions and EPDS depression case-finding questions and the extent to which they each capture appropriate information for effective screening/case-finding of depression in perinatal care.

The objectives of the EoA were:

1. to assess the acceptability of the Whooley questions and EPDS depression case-finding questions for women and HPs
2. to understand women's processes of answering the Whooley questions and EPDS depression case-finding questions in terms of their understanding, confidence and recall methods
3. to explore women's perceived effectiveness of the Whooley questions and EPDS depression case-finding questions in relation to their current and historical symptoms
4. to explore the impact of the Whooley questions and EPDS depression case-finding questions in relation to women's subsequent experience on the care pathway.

Method

Complementary methods, including a quantitative acceptability survey (see *Chapter 3* and *Appendix 5*) and in-depth interviews among expectant and new mothers and HPs, were conducted to collect acceptability and related data to address the aim and objectives of the EoA study.

Both sets of acceptability data were collected independently by different researchers and prior to analysis of either data set. This approach aimed to provide two complementary and independent sets of cognitive evaluation data: (1) a quantitative description from a large cohort of women and (2) a qualitative in-depth analysis from a subsample of those same women.

Sampling

Recruitment and consent of women participants
Recruitment and consent of women participants are described in *Chapter 3*. All women who consented to take part in the BaBY PaNDA study completed the quantitative acceptability survey tool as part of the data collection procedure for the study. Women were assured anonymity and confidentiality at the beginning of all interviews and all women were reminded of the study's aims and were given the opportunity to ask any questions they may have relating to the study.

Recruitment and consent of health professional participants
The heads of midwifery and health visiting at each study site (Harrogate, Hull, Scunthorpe and Goole, and York) were contacted via e-mail and/or telephone by a qualitative researcher (AC) to obtain their permission to contact trust staff. The heads of midwifery and health visiting were also asked to provide contact details for staff meeting the sampling criteria; these staff were then e-mailed about the study and provided with a participant information sheet and a consent form (see *Appendix 7*). Interested HPs were asked to contact the qualitative researcher directly to arrange a time and location for the interview to take place, should they wish to take part.

Data collection
Both the acceptability survey and in-depth interviews employed a cognitive framework[149] to collect data from individual women on their processing and understanding of each question and their given answer on that day for both case-finding questionnaires: the Whooley questions and EPDS. After a general question on whether or not women thought screening/case-finding in pregnancy and early postnatal periods was a

good idea for each case-finding instrument, questions within the cognitive framework included women's views on the degree of comfort, understanding, ease of completion and confidence in the accuracy of the answers for the individual questions within each case-finding instrument (EoA objectives 1 and 2).

The acceptability survey asked women to rate their views on a five-point scale of 'very comfortable'/'very easy'/'very sure' through to 'uncomfortable'/'difficult'/'very unsure' for each question (see *Appendix 5*).

The semistructured interviews explored and examined women's views and their reported experiences on the cognitive acceptability of the Whooley questions and EPDS questions in more detail. The in-depth interviews with women also collected detailed data on women's self-reported historical and current experience of depression and/or anxiety, including self-harm and/or suicide (EoA objective 3) and their historical and current experiences of being pregnant and/or a mother of a new baby. These contextual data provide important insights into the interpretation of cognitive findings of acceptability (EoA objectives 1, 2 and 3). All women who had experienced or were currently experiencing depression, with or without anxiety, based on the CIS-R diagnostic reference standard provided data on their experience of the care pathway (EoA objective 4). This included a subgroup of longitudinal interviews for women specifically to collect data on their experience of the care pathway over a series of two or three interviews from pregnancy through to 12 months postnatally.

Health professionals were interviewed to explore their experience of delivering the depression case-finding instruments in routine clinical practice and were asked to describe their training needs and suggestions to improve routine practice in their trust sites.

Interviews with women and HPs were conducted using semistructured topic guides (see *Appendix 6*). For interviews conducted with women, separate topic guides were used at 20 weeks' gestation, 3–4 months post birth and 1 year post birth (see *Appendix 6*).

Additional open-ended probes examined the case history of, and experience on the care pathway for, individual women. Interviews with women lasted between 25 minutes and 119 minutes and were conducted at the woman's home or workplace, or an alternative location if preferred. Interviews with HPs lasted between 36 and 83 minutes and were conducted in a private room at the trust site or, if participants preferred, at the University of York. All interviews were audio-recorded and were fully transcribed and anonymised before data analysis.

Interviews with women were conducted face to face when possible or over the telephone between February 2014 and February 2016. HP interviews were conducted between November 2015 and March 2016.

Data analysis
Women's categorical ratings for the cognitive questions within the acceptability survey were counted and summarised descriptively (*n* = number of women per category, % of total women) by the lead statistician.

In-depth interview data were coded and mapped for analysis within a framework approach, following the stages as outlined by Ritchie and Spencer:[150] familiarisation, identifying a thematic framework, indexing, charting and mapping, and interpretation. *Familiarisation* involved the researcher immersing themselves in the raw data by reading interview transcripts and listening to audio-recordings while making preliminary notes of recurrent themes and initial ideas. A *thematic framework* was then developed based on the researcher's initial perceptions, issues raised by interviewees, the study's research questions and in discussion with the senior qualitative researcher (LD). The framework was revised a number of times, starting as descriptive and based on a priori issues with it then being applied to a number of transcripts, and refined as it became more responsive to emergent and analytical themes. *Indexing* of data involved the systematic application of the numerical codes in the thematic framework to the transcripts. Codes were then checked, refined and reviewed further. Data were then organised by *mapping* onto the relevant

part of the thematic framework to form charts. One chart per theme was created in NVivo 10 for Windows (QSR International, Warrington, UK), including the theme, subthemes and direct links to verbatim text in the transcript for each participant. The coding and mapping work was conducted by three researchers, two for the women's interview data (LJ and ZR) and one researcher for the HP interview data (AC). Quality assurance of data coding was completed by LD, the lead qualitative researcher, to independently check a 10% sample of transcripts for comparability of coding and data entry.

Analysis and interpretation of the interview data from women was undertaken by the lead researcher (LD), who had conducted the first interviews among women in the prenatal and postnatal assessment periods ($n = 25$). Analysis of the HP data was conducted by the researcher who had conducted the HP interviews and coding and mapping of data within the framework (AC). Analysis involved exploring associations between themes, creating new themes and providing explanations for the findings.

The lead qualitative researcher (LD) integrated the findings from the acceptability survey and data from the interviews conducted with women and HPs including recommendations for research and/or implications for practice.

Results

Sampling frame

Women participants

Quantitative acceptability survey
A total of 391 expectant mothers completed the quantitative acceptability survey as part of their BaBY PaNDA prenatal assessment (at 20 weeks); 345 of these women completed the survey as part of their BaBY PaNDA postnatal assessment (at 3–4 months).

In-depth interviews
A total of 25 women completed at least one interview during the perinatal period. This included first interviews for 17 pregnant women following their BaBY PaNDA prenatal assessment (at 20 weeks) and eight new mothers following their BaBY PaNDA postnatal assessment (at 3–4 months).

Sampling targets were achieved for the number of women who participated in at least one interview (25 women, target of 25). Slightly fewer women than targeted participated in at least three interviews (three women, target of five). The majority of interviews (31/39) were completed during pregnancy and the early postnatal period, exceeding the target of 20–25 interviews in this period.

Although the number of women who completed a third interview was slightly below target, a total of eight women completed an interview (first or second) at 12 months postnatally, within the target of 5–10. *Figure 6* shows the flow of in-depth interviews conducted throughout the study.

Data saturation was achieved within the sample of 25 participants on the complex issues underpinning acceptability of the two depression case-finding instruments for women of different ages, parity, positive/ negative screens for depression (based on the primary diagnosis from the CIS-R diagnostic reference standard) and study sites. Oversampling of women with a current episode of depression (56% of women interviewees compared with 10% of the BaBY PaNDA sample) was achieved to increase the scope for data saturation on cognitive views of the case-finding instruments among the primary target population group and the care pathway experience.
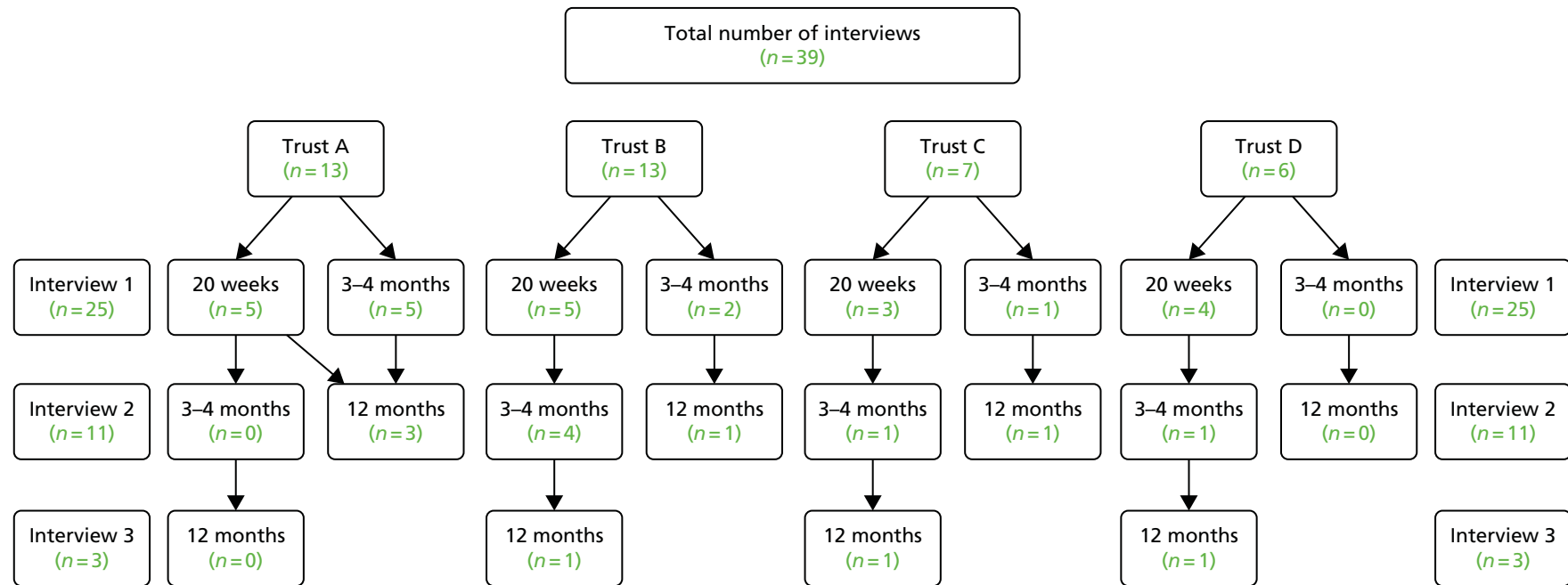
FIGURE 6 Participant interviews by site, time point (20 weeks' gestation; 3–4 months postnatally; 12 months postnatally) and interview number.

In summary, recruitment targets in terms of numbers and diversity were achieved, aiding the achievement of data saturation for all EoA objectives and confidence in the transferability of findings to similar population groups beyond this study sample (see *Discussion*).

Health professional participants

Characteristics of health professional participants

Qualitative semistructured interviews were conducted with 11 HPs across the four study sites. To preserve anonymity, study sites will be referred to as trust A, trust B, trust C and trust D in relation to interviews conducted with HPs. The sample included seven HVs and four MWs from across the four study sites: trust A ($n = 3$), trust B ($n = 3$), trust C ($n = 2$) and trust D ($n = 3$). HPs in one study site were mental health champions and included the perinatal mental health lead for the area. Alongside clinical practice, these individuals were involved in the delivery and development of mental health training and pathways. Characteristics of HP participants are presented in *Table 19*.

Our original recruitment target was to purposively sample 12 HPs (six MWs and six HVs) across four study sites according to profession, professional grade, age, location and experience. However, delays in obtaining letters of access at study sites and difficulties with identifying and contacting heads of midwifery and health visiting at study sites, limited recruitment. Although the sample of HPs is considered diverse as different study sites, professions and levels of experience are represented, the relatively small numbers of MWs ($n = 4$) in comparison with HVs ($n = 7$) and the fact that HPs in one study site were are all HVs with a specific interest in perinatal mental health are considered limitations on the achievement of data saturation and the generalisability of findings beyond this study sample (see *Discussion*).

Evaluation of acceptability

This section presents the combined findings of both the quantitative acceptability survey and the in-depth interviews conducted among expectant and new mothers and HPs regarding the EoA and effectiveness of the questions and EPDS case-finding instruments (study aim and EoA objectives 1, 2 and 3). A breakdown of the findings from the quantitative acceptability survey can be found in *Appendix 9*.

The findings of the acceptability evaluation are presented in three main sections. The sections and their associated emergent themes are reported as follows.

TABLE 19 Characteristics of HP participants

Participant ID	Profession	Trust location	Years experience
01	HV	D	9
02	HV	D	26
03	HV	D	13 (10 years MW, 3 years HV)
04	HV	A	20
05[a]	MW	A	–
06	MW	A	11
07	MW	B	26
08	HV	B	20
09	HV	B	7
10	MW	C	–
11	HV	C	–

a Missing data due to a recording error.

- Section 1: acceptability of screening/case-finding in the perinatal period

  - Theme 1: acceptability of screening/case-finding in the perinatal period.
  - Theme 2: ease of understanding Whooley and EPDS case-finding questions.
  - Theme 3: ease of remembering Whooley and EPDS case-finding questions.
  - Theme 4: confidence in answers to the Whooley and EPDS case-finding questions.
  - Theme 5: comfort to answer the Whooley and EPDS case-finding questions.
  - Theme 6: overall comparisons of the Whooley and EPDS case-finding questions.

- Section 2: experiences of depression in the perinatal period

  - Theme 7: when is feeling tired and hopeless a problem in the perinatal period?
  - Theme 8: social expectations of pregnancy and motherhood.

- Section 3: care pathway issues

  - Theme 9: facilitators of, and barriers to, screening/case-finding.

Section 1: acceptability of screening/case-finding in the perinatal period

Theme 1: acceptability of screening/case-finding in the perinatal period

Quantitative acceptability survey Of the 391 pregnant women who completed the acceptability survey, 95% felt that it was a good idea to ask women about how they were feeling during pregnancy using the Whooley questions or the EPDS. A similar proportion of the 345 mothers who completed the acceptability survey at around 3–4 months postnatally felt that it was a good idea to ask women about how they were feeling during early motherhood using either the Whooley questions (95%) or the EPDS (97%).

With the exception of two women who participated in the in-depth interviews, pregnant women and new mothers widely supported the idea of screening/case-finding for depression in the prenatal and postnatal periods as an acceptable, welcome and important thing to do. This included views of expectant and new mothers with diversity in age, parity, location across the Yorkshire and the Humber, and Northern Lincolnshire, study regions and women with some or no experience of historical or current depression and/or anxiety.

Some of the main reasons why screening/case-finding was reported as important included recognition of depression as an important issue that can affect anyone when they are pregnant or have had a baby; the need to identify women who are experiencing depression who 'soldier on' without the help they need; and to identify those women as early as possible. Illustrative quotes for these reasons are provided below (see also *Appendix 10*). The following text also indicates some underlying concerns regarding a lack of priority given to depression as an issue for expectant and/or new mothers and the potential for cases of perinatal depression to go undetected and unsupported.

Women's views on why screening/case-finding in the perinatal period is a good idea (1) Taking depression seriously as it is an important issue that can affect anyone (see *Appendix 10* for examples of more quotations):

> *I think people treat it [depression] so differently [to physical health problems], so yeah, I would see it as a good thing, definitely. 'Cos if you don't have the opportunity, if you're not specifically asked, you could probably just, you just keep trying to soldier on.*
>
> *Trust C 2273 (20 weeks prenatal)*

(2) Good to identify women with depression and as early as possible (see *Appendix 10*):

> *I think that it's [screening/case-finding] good actually 'cos before, you didn't get that [depression or anxiety] before, when you was pregnant and that, you didn't get that. But it's like when me sister had the first one [baby], she got depression . . . and no-one even noticed 'till later on when she had the baby, a bit later on after when the little un was growing up.*
>
> *Trust D 2639 (20 weeks prenatal)*

(3) Important to screen throughout the perinatal period as things change or do not get picked up initially (see *Appendix 10*):

> *I think with you asking me, had anyone asked about me after the early stages, had it sort of, had I been getting depressed later on, I don't think anybody would have noticed if I hadn't said anything. So continuing it for longer with, yeah, the sort or run of the mill questions.*
>
> *Trust A 3913 (12 months postnatal)*

Health professionals who participated in the in-depth interviews considered that, overall, screening/case-finding was beneficial, particularly given the potential implications of not identifying mental health issues during the perinatal period. The case-finding instruments were described as a useful 'way in' to discussing mental health, particularly for women who were reluctant to discuss the subject, or who showed no obvious signs of mental health issues in how they were presenting. A minority of HPs also felt that the standardised approach to screening/case-finding reduces the potential for women to feel singled out or stigmatised and provides a mechanism for measuring whether or not women's mental health is improving throughout the perinatal period:

> *Mums have said yes, that they do feel like harming themselves and yet I wouldn't have known that if it hadn't have been for asking that question.*
>
> *HV, trust A, 04*

There was a consensus among HPs that, although the case-finding instruments are important, they are not essential, with HPs often using a range of other methods to identify depression and anxiety; particularly in situations for which there is a perceived discrepancy between how women are presenting and their screening/case-finding scores.

Particular emphasis was placed on discussions with women about their mental and emotional well-being, highlighting the need to establish a good rapport with women. Discussions were mainly used for women who score in the mild–moderate depression range on the case-finding instruments to ascertain whether or not these scores are due to underlying mental health issues or other factors such as sleep deprivation or financial concerns. Discussions were also used to distinguish between normal emotional fluctuations during the perinatal period; demonstrating the need for HPs to use their experience and clinical judgement. Finally, when case-finding instruments cannot be used, discussion forms the main method for identifying mental health issues. Additional methods commonly used to identify depression and anxiety included women's body language; non-verbal cues; attachment and bonding; and previous, current or family history of mental health problems. The importance of women disclosing any previous history of mental health problems and their readiness to admit issues, particularly in severe cases, was also frequently mentioned:

> *I've had occasions where people have answered no to the screening questions but it's really obvious that there's something going on, and that's maybe not even that they're not not telling you but perhaps the way the questions are doesn't represent their mental illness or their episode of what's going on with them, so then you'd be looking for signs; with them it could be communication, body language, lack of, you know, lack of eye contact, for example, talking to the ground, monotone voice, that kinda thing. It could be that you walk into the house and you see a deterioration in the house.*
>
> *HV, trust D, 03*

Theme 2: ease of understanding Whooley questions and Edinburgh Postnatal Depression Scale case-finding questions

Whooley questions 1 and 2 were considered 'easy to very easy' to *understand* by 97% and 98% of women, respectively, who completed the prenatal acceptability survey. Similar results were reported for new mothers completing the postnatal acceptability survey.

Data from the in-depth interviews suggest that there are directly conflicting views on how women view the questions and their answer within the context of being pregnant or a new mother. One woman who was interviewed in both the prenatal and postnatal periods thought that Whooley question 2 may be confusing because some people may be less interested in doing things because they do not have the energy to do so since having a new baby, rather than because they are depressed. Another pregnant woman related a lack of interest or pleasure in doing things to the tiredness of pregnancy but was not bothered by this:

> *I just think it's part of the pregnancy and you've got to get on with it, but I know I have less interest.*
>
> *Trust D 2919 (20 weeks prenatal)*

These issues highlight the complex differences in views and experiences of different women in how to answer and interpret the case-finding questions given the overlap between symptoms of tiredness and change in mood in pregnancy and motherhood and true depression (see *Theme 7: when is feeling tired and hopeless a problem in the perinatal period?* for further detail).

Only one-fifth of the 25 women completing in-depth interviews, comprising two expectant mothers and three new mothers, explicitly reported that they found the terms in Whooley question 1 clear and understandable. This compared with the majority of women who found Whooley question 2 easy to understand and having a clear meaning.

The majority of women reported in the interviews that they had a clear understanding about the meaning of the term 'depressed', defining it as 'being sad all the time, struggling to function, to do day-to-day things, feeling down over a long time or crying more than usual'. A small number of women understood 'feeling down' and 'depressed' as quite similar in meaning, whereas most women defined 'feeling down' as something that everyone experiences at some point and which will usually pass in time. Five of the six women who provided a definition of their understanding of the term 'hopeless' described this as struggling to cope at all and unable to see a way out, lacking in self-confidence as a person and/or a mother and being linked to thoughts of suicide.

Most expectant and new mothers defined the overall meaning of Whooley question 2 as losing interest in things that they used to do and found this to be a more personal approach than that in Whooley question 1 to identify if you are depressed.

The term 'bothered', included in Whooley questions 1 and 2, generated a lot of discussion among many of the expectant and new mothers, although most had a clear understanding of the term. The most common interpretations were described as whether your feelings have been affecting or worrying you, 'got to you in some way' or interrupted your usual activities. Three women had strong views that the term 'bothered' was unclear and confusing. One expectant mother expressed her concern that not being bothered about anything is part of feeling depressed and therefore inappropriate:

> *. . . 'cos a lot of people [with depression] don't realise that they have been down or, you know, they can't be bothered to do things and such as that.*
>
> *Trust A 2699*

The term 'often', used in both Whooley questions 1 and 2, was defined differently by women, ranging from every day (the most common interpretation) to as infrequently as 3 or 4 days in a month. This finding may suggest that individual women's understanding of Whooley question 1 or 2 is not consistent within

the population group of expectant or new mothers. This approach places the emphasis on women to decide when feelings of being down, depressed, hopeless or losing interest or pleasure in things is occurring sufficiently frequently to be a problem. This assumes that women are aware of their experience of these feelings and willing to admit them to others (see *Discussion*) and that women are clear when physical and emotional changes commonly associated with being pregnant or a new mother are occurring and/or pose a threat to their mental health status (see *Theme 7: when is feeling tired and hopeless a problem in the perinatal period?*).

The 10 questions in the EPDS (see *Appendix 5*) were rated as 'easy to very easy to understand' overall by the majority of women who completed the acceptability survey during pregnancy (99%) and 3–4 months postnatally (99%). Individual questions within the EPDS that were rated the most difficult to understand during pregnancy were questions 1 (3%) and 5 (2%) and at 3–4 months postnatally were questions 1 (2%) and 3 (2%) (see *Appendix 9*).

Twenty-one of the 25 women who participated in the in-depth interviews and answered the question on the clarity of meaning of EPDS question 1 found it to be clear. Some women did express minor concerns regarding selecting the correct answer from the options available. With regard to difficulties answering question 3, two women expressed concerns about the lack of clarity for the meaning of the term 'blamed myself unnecessarily'. Five expectant and new mothers had some concerns regarding the lack of objectivity in the meaning of the term 'for no good reason' in question 5.

Theme 3: ease of remembering Whooley and Edinburgh Postnatal Depression Scale case-finding questions

Around 95% of all expectant women found it 'easy to very easy' to remember their feelings when answering Whooley questions 1 (95%) and 2 (96%). The level of ease of remembering did vary between questions 1 and 2, however, with 10% more pregnant women finding it 'very easy' to remember their feelings to answer question 2 (64%) than to question 1 (54%) (see *Appendix 9*).

Similar findings were reported for new mothers at 3–4 months postnatally, who found it 'easy to very easy' to remember their feelings when answering Whooley questions 1 (96%) and 2 (97%). Again, relatively more mothers of new babies found it 'very easy' to remember their feelings to question 2 (63%) than to question 1 (57%) (see *Appendix 9*).

Expectant and new mothers who participated in the in-depth interviews were more divided about the ability to remember their feelings within the last month to answer Whooley question 1. More women ($n = 10$) expressed difficulties regarding the period of 1 month for remembering their feelings when answering question 1. This included nine women who preferred a shorter time period as they were finding, or would expect to find it difficult to remember what they did last week now life is so busy having had (or expecting to have) their new baby. Eight women considered a month to be relatively easy and straightforward. The majority of the women who preferred a shorter time period had already had their new baby and felt very busy, whereas most women who were happy with a month were around 20 weeks pregnant.

In contrast, there were twice as many women in the interviews who found they could, compared with those who could not, remember their feelings within the last month to answer Whooley question 2. Both expectant and new mothers thought question 2 was more practical and could think of examples straight away, making it easier to remember over a 1-month time period than question 1.

With regards to the EPDS questions overall, the majority of pregnant women (98%) and new mothers (98%) found it 'easy to very easy' to recall their feelings in the past 7 days to answer the questions in the EPDS. This is consistent with findings for data from the in-depth interviews, in which pregnant women and mothers of new babies universally reported that it was easier to remember their feelings to answer the questions over a 7-day period than a 1-month period.

When asked about each of the 10 questions within the EPDS case-finding tool, pregnant women and new mothers found responses to questions 1 (2%/2%), 2 (2%/2%), 4 (2%/2%) and 6 (2%/1%) the most difficult to recall both during pregnancy and postnatally. No specific concerns were raised regarding difficulties of remembering feelings to inform the answers to any individual questions within the EPDS by women participating in the in-depth interviews.

Theme 4: confidence in answers to Whooley and Edinburgh Postnatal Depression Scale case-finding questions

Nearly all expectant women were 'sure to very sure' about their answers to Whooley questions 1 (98%) and 2 (99%) (see *Appendix 9*). In contrast, approximately two-thirds of women who participated in the in-depth interviews were confident about their answer to question 1 and one-third were not confident. Nearly all the women who reported not feeling confident about their answers to question 1 explained that this was a lack of confidence in being able to provide an honest and accurate answer due to the extreme nature of what question 1 is asking. Feeling down is considered a more acceptable term for discussion of and potential identification of depression than depressed or hopeless (see illustrative quotes below and *Appendix 10*).

Discomfort for women to answer wording of Whooley question 1 honestly See *Appendix 10* for further examples of quotations:

> *I think some people might be scared of the word 'depressed'. Box ticking makes it too easy to play down the symptoms.*
>
> *Trust C 2288 (20 weeks prenatal and 12 months postnatal)*

> *I can relate to feeling down, hopeless is such an extreme term . . . If people just answer honestly, it's hard if, you have to answer, the questions that you, are a bit risky.*
>
> *Trust B 3171 (20 weeks prenatal)*

Importantly, women's lack of confidence in their ability to provide an honest answer to Whooley question 1 results in approximately half of the women interviewed in this qualitative study sample stating that this question would not generate a truthful response. These women conclude that the nature of the question is likely to result in women understating their feelings and experiences of depression and not getting the help and support they genuinely need. The strength of feeling regarding these concerns for Whooley question 1 was deep felt as well as widespread.

Similar concerns were expressed by HPs and may have influenced the way the questions were used and how the questions were asked. Although one HV from trust D felt that the questions were 'loaded with the symptoms of depression', others criticised the questions' wording, describing them as 'stilted' and 'wordy' and saying that they interrupted the flow of their relationship with women. The terms 'bothered' and 'hopeless' were particularly criticised and described as 'flippant', 'vague' and 'scary':

> *Hopeless is such a strange word in our society to use, 'cos hopeless kind of makes you feel like you should be just like on the floor not doing anything, doesn't it?*
>
> *MW, trust A, 06*

In contrast, all nine pregnant women and new mothers who answered the question regarding their confidence in their answer to Whooley question 2 were confident in their answer and the truthfulness of that answer. Again, this was explained by the practical nature of the question, which women could relate to without difficulty or concern.

These findings were also consistent for mothers of new babies, who were 'sure to very sure' about their answers to Whooley questions 1 and 2 (99% and 99%, respectively). The proportion of mothers in the early postnatal period who were 'very sure' about their answer to Whooley question 1 (71%) was similar to that for pregnant women (70%). However, a slightly higher proportion of mothers were 'very sure'

about their answer for question 2 about feeling little interest when asked in the postnatal period (75%) compared with being asked during pregnancy (71%) or compared with question 1 about feeling down in the postnatal period (71%) (see *Appendix 9*).

Women who answered 'yes' to either Whooley question 1 or 2 were asked a third 'help' question: 'Is this something you feel you need or want help with?'.

The majority of women were 'sure to very sure' about their answer to the third help question if asked during pregnancy (93%) or the postnatal period (92%); around 5% of responses to this question were missing. Only 55–60% of women were 'very sure' about their answer for question 3 regarding help, compared with 70–75% for questions 1 and 2 regarding feelings of being depressed or having little interest.

The relatively lower scores regarding confidence in answering the third 'help' question compared with scores for questions 1 and 2 are supported in the findings of women's views on and experiences of this question as discussed in the in-depth interviews. There was widespread support for the offer of help to women who had answered yes to either question 1 or 2 and particularly the use of both terms 'need' or 'want' in relation to help. In fact, one pregnant woman was of the view that help should be offered without asking question 3 if a woman answered 'yes' to questions 1 or 2 (trust A, 2493).

> *I think that's a really good question to ask actually, I think that's a really good question because, yeah, do you, do you want help; who from I suppose. Is this something you feel you need or want help with from a, maybe you could tick who, who you would, which health professional you would like help from.*
> *Trust B 1296 (3–4 months postnatal)*

> *It's, it's like somebody giving out like a helping hand. It's, it's somebody asking the question that you're kind wanting to ask them.*
> *Trust D 3486 (20 weeks prenatal)*

Two main concerns influencing women's ability to answer the help question, however, were expressed by several women: (1) a lack of knowledge regarding what type of help is available and a genuine fear of what the intervention might be and (2) it assumes women have admitted they are depressed to themselves and are willing to admit they need help to a HP (see illustrative quotations below and *Discussion* for underpinning theoretical concept of 'surrendering').

Reasons why women are less confident about their answer to Whooley question 3 regarding help for depression: Reason 1: lack of knowledge of what help is available (see *Appendix 10*):

> *'Cos I'd, I'd say no but then they [the health professional] might say something, and you think, oh yeah, that'd be a good idea. But I don't feel like there is anything but I suppose I don't, yeah, I don't know what's out there.*
> *Trust D 2919 (20 weeks prenatal)*

Reason 2: need to admit depression and need for help to self and others (see *Appendix 10*):

> *You'd have to acknowledge that something was wrong and be confident enough to ask for the help, not be embarrassed to.*
> *Trust B 3171 (20 weeks prenatal and 3–4 months postnatal)*

With regards to the EPDS questions overall, the majority of pregnant women (99%) and new mothers (99%) were 'sure to very sure' about their answers to questions in the EPDS when completing the acceptability survey (see *Appendix 9*). This is consistent with findings from the in-depth interviews, in which nearly all of the women who answered the question regarding confidence in their answers answered 'yes' to all the questions.

When asked about individual questions within the EPDS as part of the acceptability survey, pregnant women were 'particularly unsure' about questions 3 (4%), 4 (3%) and 6 (2%), whereas postnatal women were 'particularly unsure' about questions 1 (3%), 2 (2%) and 9 (2%) (see *Appendix 9*). As detailed above, concerns regarding the meaning of the term 'blaming myself unnecessarily' may be a factor affecting women's confidence in answering EPDS question 3. One woman was uncertain whether her unhappiness and frequent crying were as a result of a change in hormones due to being pregnant or because she was feeling down. This may partly explain why some women were less sure about answering EPDS question 9.

Interview women expressed very positive views regarding EPDS questions 6 and 8, which were compared favourably against Whooley question 1. Interestingly, question 10 regarding self-harm was acknowledged as important and well worded despite being difficult to answer.

Theme 5: comfort to answer Whooley and Edinburgh Postnatal Depression Scale case-finding questions

Nearly all women felt 'comfortable' or 'very comfortable' completing the Whooley questions (98%) and the EPDS (99%) during pregnancy. However, although 80% of expectant mothers were 'very comfortable' completing the EPDS questions, only 75% reported being 'very comfortable' answering the Whooley questions.

Similar findings were reported for women's views in the postnatal period, with 99% of mothers reporting feeling 'comfortable to very comfortable' completing either the Whooley questions or the EPDS. Again, the proportion of mothers who in the early postnatal period felt 'very comfortable' completing the EPDS questions (80%) was higher than the proprotion who felt 'very comfortable' answering the Whooley questions (76%). Perhaps surprisingly, no women reported feeling uncomfortable completing either case-finding instrument in either the prenatal or postnatal periods (see *Appendix 9*).

The high levels of comfort in answering the case-finding questions were not supported by the findings from the in-depth interviews with respect to Whooley question 1. The majority of women reported in the interviews that, although it is easy to answer the questions, it is not comfortable to give an honest answer to Whooley question 1. For most women, the reason for such discomfort in giving an honest answer was due to the severity of the wording of question 1, which includes the terms 'depressed' and 'hopeless' (as detailed in *Theme 4: confidence in answers to Whooley and Edinburgh Postnatal Depression Scale case-finding questions*). The underlying cause of such discomfort appears to be grounded in a widespread perception among our interview sample that having depression or feeling hopeless is associated with highly negative social and cultural stigma. Such stigma is exacerbated by the pressure of sociocultural norms around pregnancy and motherhood as a time of jubilation and fulfilment (see *Theme 8: social expectations of pregnancy and motherhood* for more detail).

In addition, deeply held fears of the sociopsychological implications of being diagnosed with depression, namely, losing your baby or being monitored in your own home and judged by HPs as a bad mother, add to the discomfort of answering Whooley question 1 regarding feeling depressed or hopeless openly and honestly (see illustrative quotations below).

Social and cultural reasons causing discomfort for women to answer Whooley questions honestly (1) Social stigma associated with depression, especially for expectant and new mothers (see *Appendix 10*):

> It think it's [depression] got a lot of stigma . . . I think I just didn't want to be labelled . . . that word and, and, and it, it felt like I didn't want a, I didn't want to accept it because I should be a really, a good mum. You're just expected to be perfect and oh, everything's going lovely and look at the fantastic little baby and you wanna go, oh but I'm really struggling and I'm finding it hard.
> *Trust B 1653 (3–4 months postnatal and 12 months postnatal)*

(2) Fear of consequences of having depression (see *Appendix 10*):

> *. . . it just felt like I couldn't do it, I was going through the motions . . . I just wanted to cry . . . I used to go the doctors and not say anything obviously, but wanted them to ask, and then didn't want to say, 'cos then I thought they might take my children away from me.*
>
> *Trust A 1839 (3–4 months postnatal)*

The same feelings of discomfort about answering honestly and openly were not expressed for Whooley question 2. In general, this question was not perceived to have the high level of negative stigma associated with the terms used in question 1. One notable exception is the potential for expectant and new mothers to be concerned about being judged negatively if feeling little interest or pleasure in their new baby against the sociocultural, and sometimes personal, expectation to be the 'perfect mother', full of joy and fulfilment (see *Theme 8: social expectations of pregnancy and motherhood* for more detail).

Findings from the in-depth interviews regarding views on the ease and comfort of answering questions honestly for the EPDS were consistent with those reported in the acceptability survey. Overall, women who participated in the in-depth interviews expressed a very positive view of the EPDS. The main reasons for this were particular characteristics, including softly worded individual questions that are not threatening to the woman; softly worded answers to choose from that reflect the complexity and range of emotions experienced by women in the perinatal period; a gentle build-up to the more difficult questions strategically placed at the end of the questionnaire; and the nature and number of the questions creating an environment of perceived genuine interest in the woman's emotional well-being rather than a tick-box exercise within an otherwise medical assessment (see illustrative quotations below):

> *. . . to me it [EPDS Q6 compared Whooley Q1] means that that things that you probably could deal with before are now overwhelming yer somewhat and you're not coping, so it's a softer version of those words like hopelessness.*
>
> *Trust B 3776 (3–4 months postnatal)*

> *. . . yeah the more direct ones [EPDS Q9 & 10] are at the end, which is good, because you've sorta built up that relationship first.*
>
> *Trust A 2415 (20 weeks prenatal)*

> *. . . probably it's [EPDS Q10] one of the most important, you know, trigger questions to get someone help . . . and you know, it's, it's nicely worded, it's clear and it's not yeah, it's easy to answer. So I think probably, yeah, needs to stay."*
>
> *Trust A 2493 (20 weeks prenatal)*

Comfort for HPs to routinely ask Whooley questions In accordance with NICE guidance (2014),[101] all HPs reported using the Whooley questions in routine practice. For the majority, the Whooley questions were used as part of more general discussions about women's mental health and emotional well-being. Although this may reflect a lack of training and awareness of why questions should be asked verbatim, how questions were asked and whether or not they were asked verbatim also varied according to time constraints, how women present (e.g. questions may not be asked if a woman is threatening harm) and women's intelligence quotient and English-language skills:

> *I don't ask them verbatim at all, if I'm honest, it's sort of you know, I'll ask about their mood, I'll ask about how they're feeling, you know, are you feeling.*
>
> *HV, trust B, 08*

Some issues with asking some case-finding questions verbatim were reported, with one HP also stating that they 'wouldn't dare ask if a woman had harmed their baby directly'. Despite this, the case-finding instruments were considered an important part of routine care, with the majority of HPs reporting feeling

comfortable asking women about their mental health and using the case-finding instruments. This was largely attributed to mental health not being considered as sensitive as other topics discussed with women during the perinatal period (e.g. domestic violence, paternity).

> *I don't feel particularly like it's a hard thing to ask, I'm quite happy to ask, 'cos it's important, you know, at the end of the day. It's like, I mean you're asking them other things in here, like have they had any incontinence.*
>
> *MW, trust A, 06*

Situations in which more general discussions about women's mental health were used instead of the Whooley case-finding instrument were also discussed. For instance, despite being a core contact, a minority of HVs reported not feeling comfortable about conducting the routine screening/case-finding at birth visits with the Whooley case-finding instrument because of their concerns that, given women's potential emotional state, there would be a large number of FPs.

In addition, HPs did not feel comfortable using case-finding questions as routine practice if women had suffered significant trauma or bereavement or were experiencing a crisis, or had a history of depression or anxiety or had recently been screened by another HP. A midwife in trust C also reported how students were being told not to use the case-finding tools or ask women about their emotional well-being to avoid causing emotional distress:

> *. . . if I know that somebody has a history of depression and anxiety, for me personally, and they are seeing their GP, I don't see it as appropriate to ask these questions, because they're actually, they've already indicated that there are concerns and so might be able, an assessment tool or something you would do further down the line once they've started treatment, but I will discuss it with them.*
>
> *HV, trust C, 11*

Theme 6: overall comparisons of Whooley and Edinburgh Postnatal Depression Scale case-finding tools

The EPDS was rated more favourably than the Whooley questions for all cognitive criteria during pregnancy (at around 20 weeks):

1. The EPDS was slightly more comfortable to answer for women than the Whooley questions (mean 1.26/1.37).
2. The EPDS questions were slightly easier to understand for women than the Whooley questions (mean 1.28/1.32).
3. It was considerably easier for women to remember answers to the EPDS questions (mean 1.35) than to the Whooley questions 1 (mean 1.65) and 2 (mean 1.57).
4. Women were generally more sure about their answers to the EPDS questions (mean 1.35) than to Whooley questions 1 (mean 1.41), 2 (mean 1.41) and, particularly, 3 (mean 1.60).

Comparison of Whooley questions by key cognitive criteria between prenatal and postnatal periods Understanding of questions was generally better than remembrering of responses. Although ratings at 20 weeks of gestation and 3–4 months postnatally were similar for all acceptability questions, they tended to be marginally more favourable at the postnatal time point. Ratings were generally lower for question 3, the help question, than for questions 1 and 2.

Comparison of EPDS case-finding tool between prenatal and postnatal periods Although ratings at 20 weeks of gestation and 3–4 months postnatally were similar for all acceptability questions for the EPDS, they tended to be marginally more favourable at the postnatal time point.

Women in the interview sample (*n* = 25) were generally extremely positive about the EPDS as a case-finding instrument. The individual questions were considered to be appropriate and important, using single, 'soft' terms that, together, generate an overall assessment of depression and/or anxiety. Some women thought,

for example, that the phrase 'things have been getting on top of me' was much more appropriate than its equivalent term, 'hopeless', used in Whooley question 1. The choice of answers was also described as a strength of the EPDS, reflecting the complexity of women's feelings and experiences in the perinatal period, which cannot be readily broken down into a yes/no answer. Despite the length of the questionnaire, only two women expressed concerns regarding the acceptability of an individual question within the EPDS, namely the use of the expression 'blamed myself unnecessarily' in question 3:

> *I don't know why these [EPDS] are acceptable and the first one [Whooley] aren't. I think it's just, yeah, 'cos you, because there's more scope I think, just people, people are more complicated than yes or no.*
> *Trust B 3776 (3–4 months postnatal)*

Section 2: experiences of depression perinatally

Theme 7: when is feeling tired and hopeless a problem in the perinatal period?

A recurring theme cutting across many of the findings of cognitive acceptability of the case-finding instruments is the lack of clarity on when common feelings such as tiredness and change in mood are considered a normal and acceptable part of being pregnant or a new mother and when they should be considered symptoms of depression necessitating identification, diagnosis and treatment.

Many women in our interview sample expressed concern and resulting unnecessary worry that their symptoms may be interpreted by the HP as them being depressed even though themselves did not believe this to be the case. Given the reluctance of many women to admit to being depressed in the perinatal period because of fears of the social stigma and its potential consequences (for example, losing their baby), it may be that some of these women are in fact experiencing depression but prefer to attribute the symptoms to the more socially acceptable explanation of pregnancy and/or new motherhood. This complicating factor specific to the identification of new cases of depression among women in the perinatal period creates an added level of uncertainty that requires understanding, discussion and guidance on how it should be managed consistently and effectively.

Four case studies are provided in *Appendix 12*. Case studies A–C illustrate the complex and diverse scenarios in which each woman expresses the view that her symptoms of depression or anxiety are as a result of being pregnant or a new mother, some of which are correct and some incorrect in this assessment, some are supported with the presence of clinically assessed symptoms of depression and/or anxiety and some are not, and some receive the help and support they need as others go undetected. The final case (D) illustrates the woman's concern that her worries about being a good mother to her premature baby may be assessed as having a mental health problem.

Theme 8: social expectations of pregnancy and motherhood

Another recurring theme cutting across many of the findings of cognitive acceptability of the screening/ case-finding tools is the social and personal pressure women face to feel joyful and jubilant as an expectant and new mother of a beautiful new and precious baby. Sadly, the daily reality of being pregnant, giving birth and parenting a newborn baby may not always feel as joyful and jubilant, despite the overwhelming love for the precious new life in the woman's care. This personal disappointment can be a contributing factor towards feeling low mood in itself:

> *I think that's what it was, I had an expectation in my head [about being a wonderful mother] and it wasn't . . . working how I wanted it to work and I got myself in a bit of a mess about it.*
> *Trust B 1653 (3–4 months and 12 months postnatal)*

Social expectations can also add to a woman's existing concerns about admitting to herself or others that she feels down or worried, making it less likely that women will give an honest answer to the case-finding questions:

> *. . . but also I think anyway, just, I mean even just the, the label 'depression' you know, people are very keen not to have that associated with them, very few people want to say they think they're*

depressed, you know, particularly at a time in your life when you've just got this amazing new baby and you're meant to be happy, you're meant to be coping, and if I think, if you're not, to tell someone that actually you're not doing these things that all the bloody books and everything will tell you you're meant to be doing. I think there's a real problem there, to get people to answer truthfully (laughs) about how they're feeling.

Trust A 1839 (3–4 months postnatal)

. . . you're just expected to be perfect and, oh everything's going lovely and look at the fantastic little baby and you wanna go, oh but I'm really struggling and I'm finding it hard.

Trust B 1653 (3–4 months and 12 months postnatal)

Everyone just expects you to be happy when you've had a baby and just get on with it.

Trust B 2321 (20 weeks prenatal)

As in the case of the complex overlap between symptoms of depression and physical and hormonal changes in the perinatal period discussed above, an understanding of the additional social pressures of motherhood by both the women and HPs and a commitment to discuss them openly may be considered a method to improve the effectiveness of the screening/case-finding process for depression among expectant and new mothers.

Section 3: care pathway issues

Theme 9: facilitators of, and barriers to, screening/case-finding

A number of facilitators of, and barriers to, appropriate and effective screening/case-finding were identified by expectant and new mothers and HPs during the in-depth interviews. Despite the open, unprompted nature of the question aimed at eliciting this information, both groups identified the same facilitators and barriers based on their views and experiences of screening/case-finding for depression in the perinatal period.

Location of screening/case-finding Of the women who expressed a preference for a specific location, several identified their home as their preferred option on the grounds that this was more relaxed and there were fewer distractions. One woman, however, reported that her home visits from her HV felt rushed as the HV usually needed to get to her next appointment.

The over-riding requirement identified by most women was the need for the location to be private and quiet and without the presence of the woman's partner, family members, a trainee HP or other patients. The midwifery clinic was explicitly identified as unsuitable owing to the lack of privacy and quiet space:

. . . he'd [partner] be listening upstairs, and then after they've gone, he'd go, why did you say that, do you know what I mean.

Trust D 3486 (20 weeks prenatal and 12 months postnatal)

Nearly half of the HPs identified the barrier of another person being present at home while conducting screening/case-finding and difficulties in asking that person to leave:

. . . grannies can be real barriers because, you know, they might just be a neighbour from down the road come to see the baby you see and you're, you're quite, you don't want to kick 'em out 'cos you're in somebody's house.

HV, trust D, 01

Health professional to conduct screening/case-finding In most cases, women who expressed a preference for a particular HP to conduct the screening/case-finding identified their HV or MW, as the continuity of care from the same team throughout the perinatal period had allowed them to develop a

relationship with some of those HPs. In some cases, however, a poor relationship between the woman and her MW or HV specifically precluded this routine carer as the preferred option. Most women expressed a clear preference for their GP not to be the HP of choice owing to a lack of continuity of care and relationship with a single GP nowadays and/or their GP's medical approach to mental health issues.

Some women would prefer to be screened by a mental health practitioner as they had experienced a rushed 'box ticking' exercise with their routine carer. A mental health practitioner is also considered by most women to have undergone the relevant training and have the relevant skills and experience to discuss and consider the questions and answers appropriately for the specific needs of the individual woman. For example, both HPs and women identified the need for the HP to be skilled in reading women's body language as part of screening/case-finding.

The majority of HPs who participated in an in-depth interview also identified the importance of a good rapport and relationship with women on their caseload as one of the over-riding criteria for effective screening/case-finding for depression:

> If you've built up that relationship before then they trust you as well, then you're going to get, you know, somebody who's going to open up more.
>
> *MW, trust A, 04*

In summary, therefore, the HP conducting the screening/case-finding should have regular contact with the same woman over time, have developed a good working relationship with the woman, and should be trained and have the relevant skills and experience to discuss the screening/case-finding questions and/or answers appropriately for each woman.

Discussion of case-finding questions Several expectant and new mothers expressed a preference to be asked the case-finding questions in person using discussion rather than a case-finding question-led approach. The willingness of a HP to spend the time doing that is an important part of the process as it creates an environment of patient centredness in which the woman's emotional well-being and not the baby's physical health, is the focus of the appointment. This approach assumes that the HP can plan the necessary time to conduct the screening/case-finding in this way, a barrier which is not easily overcome, and the HP has the appropriate personal communication skills to conduct the screening/case-finding appropriately. Training to improve communication skills for screening/case-finding of depression and anxiety should be standard practice to minimise the current wide variation in practice and communication styles:

> I think if you had the relevant training you would know how to handle that situation and to be able to probe to get the effective answer.
>
> *Trust C 1489, mother (3–4 months postnatal)*

> Well I mean she [HV] was just like, you know, well she, basically phrased it in a way where, yes I feel fine, you would have to be a very strong person not to give that answer, you know. So it was very much phrased in a term, 'Now, you know, so I, I need to just ask you about depression, you know, so I mean you've been feeling fine haven't you?
>
> *Trust A 2493, expectant mother (20 weeks prenatal)*

Advance warning and clear explanation about being asked screening/case-finding questions Both expectant and new mothers and HPs identified the importance of being told in advance about the appointment to discuss the screening/case-finding questions. This would also help with the practicalities of arranging a suitable quiet and private location for the screening/case-finding:

> A way to improve the verbal screening is if it is clearly explained beforehand, 'I want to ask you a few questions about how your are feeling'.
>
> *Trust C 2288 (20 weeks prenatal)*

. . . if they're asked properly and if it's explained, if it's explained fully why we're asking, what we're asking and then people are getting the information that these are the things we are concerned about and why . . . then yeah it works, but if you're just asking them again as a tick box, but that never works.

HV, trust C, 10

Summary

In summary, women and HPs have clearly identified a broad set of requirements that create an appropriate environment to facilitate improved and more effective screening/case-finding of expectant and new mothers for depression in the perinatal period. The environment should be patient-centred with a clear focus on discussing the woman's mental health; pre-warning of the intended focus on the woman's mental health; sufficient time to discuss the issues properly; a quiet and private space for the discussion; a good relationship between the mother and the HP conducting the screening/case-finding; a good understanding by the person conducting the screening/case-finding of the social stigma regarding being labelled as having depression and its potential impact on a woman's ability to be open; and good communication skills of the person conducting the screening/case-finding to overcome a woman's lack of openness resulting from her personal fears and/or social concerns.

This environment is more consistent with that experienced by pregnant women and new mothers when they completed their screening/case-finding questionnaires with an interviewer within the BaBY PaNDA study. Conversely, most of these environmental requirements are not typically in place when screening/case-finding for depression by a HP is routinely implemented in practice with a pregnant woman or new mother:

In terms of the scoring and knowing I was borderline but thinking no, I want help, . . . I was trying to give the right answer. I can't remember exactly, and it was like well, is it, is it not, and I couldn't decide which answer to go in and, the best way I can describe it is I felt like I was in an exam [with the HP] whereas with BaBY PaNDA I didn't.

Trust A 1148 (20 weeks prenatal)

Evaluation of case-finding and care pathway

This aspect of the EoA aimed to explore the impact of the Whooley and EPDS case-finding questions in relation to women's subsequent experience on the care pathway (EoA objective 4).

This objective was informed by data from the in-depth interviews among women participants and HPs. Their views on and experiences of screening/case-finding, referral, diagnosis, treatment and management of care for depression with or without anxiety are presented here.

Screening/case-finding instruments used for depression in the perinatal period

All HPs reported using the Whooley questions in routine practice (*Table 20*). HVs in two study sites were the only HPs who reported using the EPDS, with the tool used to determine severity and/or a need for referral in the case of women who responded positively to the Whooley questions. Among those HPs not using the EPDS, a couple described how their trusts had previously used the EPDS, with the perinatal lead for mental health in one study site being the only HP to provide an explanation for this change in screening/case-finding practice:

. . . we used to use the EPDS some years ago and then when NICE recommended PHQ-9 that's when we moved and, and everyone moved, and now NICE recommend either, and so do the Institute of Health Visiting recommend either, but because everybody moved to PHQ-9 and we all use it, GPs and everybody now, 'cos at the time they were using summat different, we were using, and it was, you couldn't possibly assess somebody's mental health if you're all using different tools, so we all agreed we'd go with the PHQ-9.

HV, trust D, 01

TABLE 20 Screening/case-finding instruments used in routine care by study site and profession

| NHS trust/HP | Whooley questions | | EPDS | | GAD-7 | | GAD-2 | | PHQ-9 | |
	Used	Time point	Used	Time point	Used	Time point	Used	Time point	Used	Time point
Trust A										
HV	✓	Five contacts, antenatal contact, birth visit, 6–8 weeks, 1 year, 2 years	✓	Used only if women respond positively to Whooley questions						
MW	✓[a]	Booking appointment, 28 weeks (antenatal)								
Trust B										
HV	✓	Antenatal (25–37 weeks), postnatal (10–14 days), 6–8 weeks, 3–4 months, 7–12 months								
MW	✓	Booking appointment								
Trust C										
HV	✓	Birth visit, 6–8 weeks, 3–4 month contact	✓	Used only if women respond positively to Whooley questions					✓	Used only if there are anxiety-related concerns
MW	✓[a]	Booking appointment, antenatal (28 weeks) if there is a concern, when discharged to HV (10 days after birth)								
Trust D										
HV 01	✓	Six core contacts, antenatal visit (28 weeks); birth visit at 10–14 days; 6–12 weeks; 3–4 months; 6–12 months; 2–2.5 years and in between contacts if necessary			✓	Six core contacts and in between contacts if necessary. Used only if women respond positively to Whooley questions	✓		✓	Used only if women respond positively to Whooley questions
MW	✓	Booking appointment								

a Includes additional anxiety question.

Impact of screening/case-finding on health services

A large proportion of HPs felt that the screening/case-finding instruments bring added time pressures, particularly when women respond positively and require help. This was considered particularly problematic during busy clinics, with HPs discussing the dilemma of having a woman who needs help, while being aware of other appointments and short appointment times. For some, time pressures led to HPs worrying whether or not they had sufficient time to invest in exploring and identifying issues, and for two HPs, as a result, to save time, did not ask screening/case-finding questions at every appointment. More general time pressures and HP's large caseloads also led to some HPs being concerned that they would not have time to make additional visits to women and provide support when needed. In contrast, one HP felt that there were no added time pressures as 'women had always been asked about their mental health':

. . . only you and you've got the, all these ladies waiting, you've got like an hour to see them all and you know you've got another visit to do and this person starts crying, how much time can you honestly invest in finding out what it is?

MW, trust C, 10

Although a minority of HPs felt that, because women had always been asked about their mental health, screening/case-finding had not increased demand on services, others felt that screening/case-finding may have increased pressure on other services, particularly GPs. However, HPs considered it the responsibility of these services to respond appropriately:

I'm aware that by the process of screening we will increase the pressure of you know, into primary care, it's inevitable I think with screening that that would happen, yeah and I'm comfortable with that. I think, you know we need to, there's no point in having a screening tool if you're not going to back it up with primary care.

HV, trust D, 03

In one study site, the organisation was in a 'transition period' at the time of the study, with HVs soon to be commissioned by the local authority, resulting in uncertainty as to how their practice would be affected and staffing issues already emerging. For instance, staffing shortages, which in addition to high caseloads and the organisation's perceived reluctance to cover the staffing deficit either through hiring bank staff or new staff or increasing current staff's hours, led to women sometimes not being seen during pregnancy because of time constraints. As a result, HPs raised concerns relating to their own health and the potential for mental health issues to be missed:

At the moment people are leaving, they're not replacing them, we've people off sick; so we have to prioritise the work that we've got, and obviously child protection and things like that will take priority. So I can't speak for everybody else but I am managing to see all my ladies but I think some people are actually struggling to see them.

HV, trust B, 08

Health professionals had varied experiences of co-ordinated, or collaborative, care between health-care services and sectors to diagnose, refer and treat women with depression in the perinatal period. This includes variations between and across study sites. For instance, although MWs in three study sites discussed how HVs are alerted to women with mental health issues by MWs, examples of communication problems between MWs and between MWs and other hospital staff were reported in one study site (e.g. labour MWs not reporting concerns to postnatal MWs).

Referral and treatment pathways for patients with depression in perinatal periods

The complex referral and mental health-care pathways for women in the prenatal and postnatal periods as self-reported by HPs in the in-depth interviews are illustrated by trust, profession and severity of depression in *Figures 7* and *8*. The complexity of the care pathways between sites is largely a result of the large
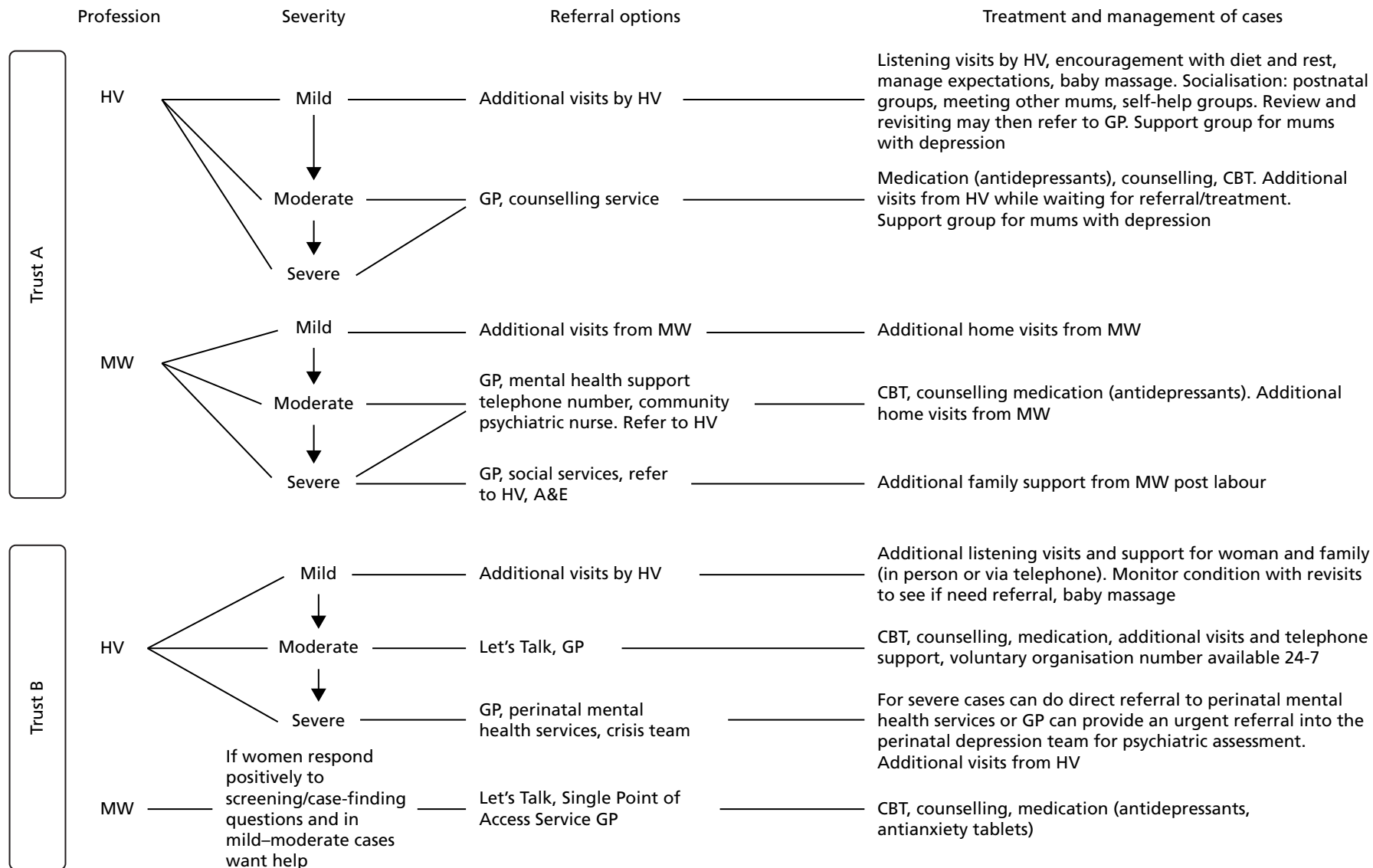
FIGURE 7 Referral and mental health-care pathways in trust A (top) and trust B (bottom). A&E, accident and emergency; CBT, cognitive–behavioural therapy; IAPT, Improving Access to Psychological Therapies.
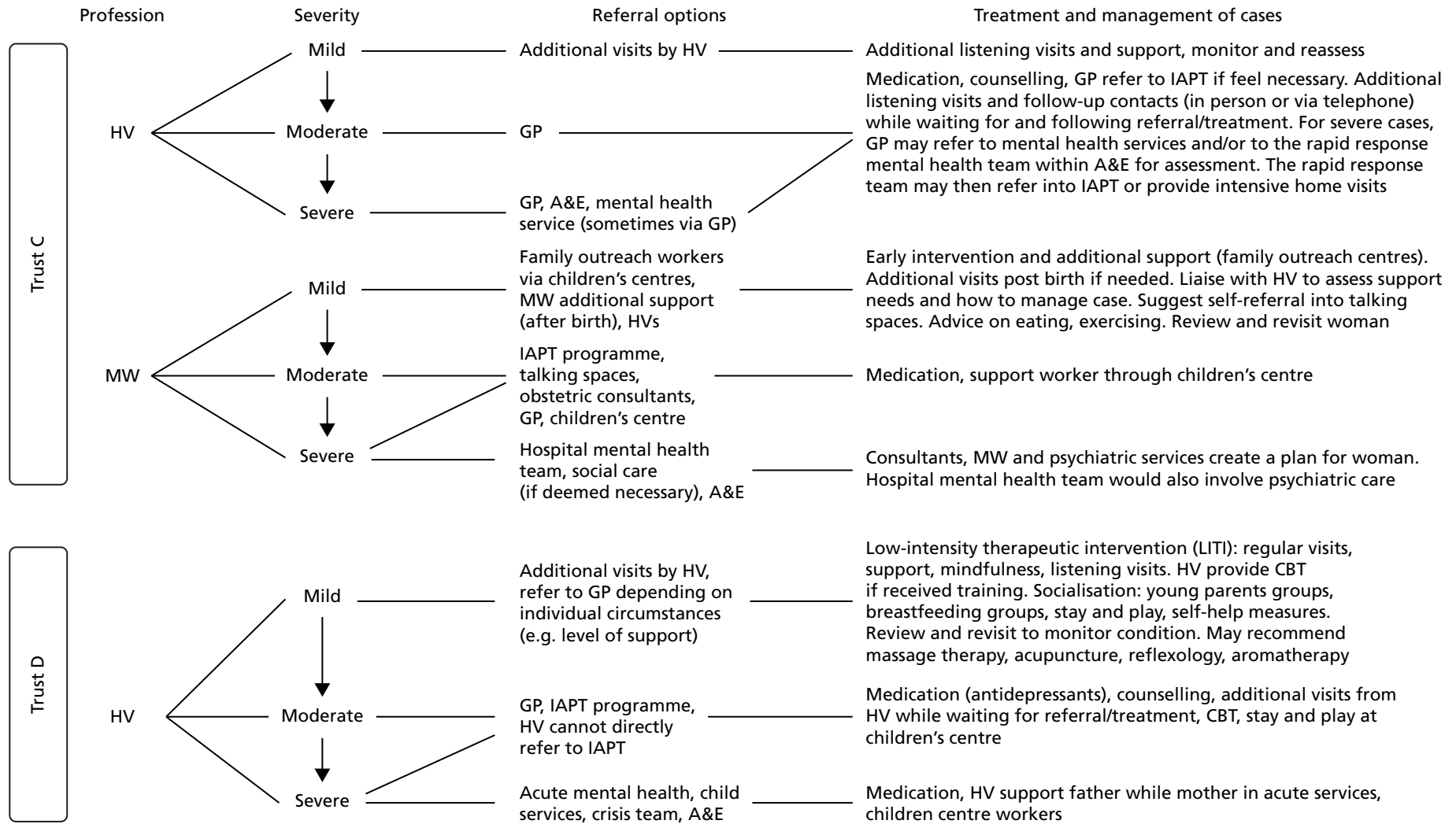
FIGURE 8 Referral and mental health-care pathways in trust C (top) and trust D (bottom). A&E, accident and emergency; CBT, cognitive–behavioural therapy; IAPT, Improving Access to Psychological Therapies.

numbers of providers for mental health services and local decisions regarding the organisation of care to deliver those, often multifaceted and complementary, services.

Despite the inevitable local variation in organisation of care, universal standards for a core provision of care are evident across sites. Women who are considered to have mild depression and anxiety will receive additional visits and monitoring of symptoms as standard care. Women who are identified as having moderate depression, and sometimes women with moderate to severe depression, are usually referred to their GP when there is a clear need for medication and/or when a GP referral may be needed to access core counselling and support from primary mental health-care services such as cognitive–behavioural therapy (CBT) and an Improving Access to Psychological Therapies (IAPT) programme.

Barriers to referral and treatment

General practitioners' awareness of mental health and their lack of knowledge regarding different treatment options and their responsibilities were frequently cited barriers to referral and treatment, which has in some cases led to women who have been referred to GPs being turned away. In regards to treatment, HPs gave conflicting accounts of antidepressant prescribing, with some perceiving GPs to be reluctant to prescribe, particularly during pregnancy, because of 'the toxic effect on the fetus', whereas others criticised GPs for rarely referring women for alternative treatment (e.g. counselling) and overprescribing medication. One HP also criticised GPs for not eliminating other potential causes of low mood (e.g. thyroid) before prescribing. A further issue with GPs was their failure to alert HPs when women's medication had been stopped:

> You get one GP who says, I like this antidepressant and I'll give them that, you'll get another one who'll do something completely different; one will prescribe, one who won't prescribe. So we do send people into a bit of a lottery, that's my frustration.
>
> *HV, trust D, 03*

Another commonly discussed issue was women's reluctance or refusal to accept treatment. HPs perceived women to be particularly reluctant to be referred to GPs. This was often linked to perceptions that women are concerned about taking medication during pregnancy. However, women's reluctance to receive treatment was also attributed to a desire to deal with issues on their own, previous poor experiences of treatment and a fear of having their baby 'taken from them':

> Some women just don't wanna go to the GP, they don't wanna think about medication.
>
> *HV, trust B, 08*

On a practical level, for HPs in two study sites, the extensive waiting times for women to receive counselling and CBT were considered barriers to both treatment and referral. For HPs in one study site, the closure of a local mental health hospital and subsequent lack of a mother-and-baby unit were of particular concern. A number of HPs also discussed how they receive no official feedback on the impact of screening/case-finding from GPs or other services and are reliant on women to inform them on current waiting times for treatment and referral:

> Well I, I think all the cuts that are happening at the moment with the mental health services; I mean [name of local mental health hospital] just closed down, we've no close mother & baby unit and, as I say, the counselling, the waiting list for counsellors and CBT is so long it's just ridiculous.
>
> *HV, trust A, 04*

Although the mental health specialists in one study site showed an awareness of referral pathways and reported following NICE guidance, a small number of HPs from other study sites reported having limited knowledge of referral pathways. This was attributed by the HPs from two study sites to insufficient

training. Associated with this, a minority of HPs acknowledged the importance of having the confidence to refer, which was perceived to be a particular issue for student HVs:

> *. . . it's not clear to be honest, but I think we either have a form we can fill in and send off, but generally we ring up and there's a, a number to ring, and this number changes a lot; so it's a little bit confusing.*
>
> *MW, trust B, 07*

Additional barriers to providing treatment that were identified by a minority of HPs included insufficient time for HPs to provide additional visits; low-intensity support; and a lack of psychologists to determine the causes of depression and anxiety.

Women's views and experiences of the mental health-care pathway

Nearly a quarter of women interviewed (6/25) were not aware they had been screened for depression during their recent pregnancy, birth or postnatal period. Of the women who were aware of having been screened, half (9/18) recalled being asked the Whooley questions and half recalled being asked how things were going or how were they feeling in general conversation. One woman was screened opportunistically for depression by her GP following an appointment regarding a cough, during which the issue of insomnia emerged.

Of the 25 women who participated in the in-depth interviews, 17 were classified as having a history, or current episode, of depression with or without anxiety. The classification was based on data collected for the CIS-R diagnostic reference standard completed during the BaBY PaNDA assessments. The views and experiences of being referred into, and treated and managed within, the local mental health-care pathway, are presented for several of these 17 women in *Appendix 11* as individual case studies. Women's experiences ranged from very positive to mixed and negative. The 17 women included nine women with a current episode of depression identified by the HP, three women with a history but no current episode of depression and five women who were identified as having a current episode of depression during the BaBY PaNDA assessment but not by the HP.

One woman who had been identified as having depression by her HP had a mixed experience on the care pathway. She had received good support from her HPs in the prenatal and postnatal periods, but was unable to access CBT without significant delay because there was a waiting list (trust C 2941). This resulted in her accessing CBT privately, although financial considerations resulted in this not being affordable once she was on maternity leave. This woman's case of mild depression, historically and currently, was directly related to her devastating experiences of being an expectant mother. At age 42 years, she had experienced two multiple pregnancies with her partner through the in vitro fertilisation (IVF) programme; her first set of twin daughters died at 23 weeks' gestation and one of her babies in her second twin pregnancy also died. The second twin, a baby boy, survived pregnancy and birth despite some initial health problems:

> *. . . so yeah, this pregnancy is very precious, . . . But yeah, it did make, there was, we had, we had to be very brave, let's say, to start again.*
>
> *Trust C 2941 (20 weeks prenatal)*

The main findings emerging from the women's diverse experiences suggest that the main factor influencing the care for women who are already on the mental health-care pathway appears to be the individual approach of the HP and the woman's relationship with that HP. This is particularly important for women who are experiencing depression and are more likely to be feeling vulnerable and in need of familiarity and support. A woman's state of readiness to ask for, or accept an offer of, help to treat depression has also been shown to be an important factor in the woman's experience of the care pathway. The need for HPs to show a more balanced interest in both the physical health and development of the baby and the emotional well-being of the mother appears to be an important factor influencing effective screening/case-finding from the women's perspective.

Overall, it is clear that some women are not being identified as a new case of depression and are not being appropriately referred for monitoring and/or treatment within the existing screening/case-finding, diagnostic, referral and treatment pathway. In some cases, this may be due to the lack of expertise and/or confidence of the individual HP to conduct a holistic and accurate assessment of the woman's circumstances and readiness for help. In one case, this was the experience of a woman who was self-harming and asking for help. Conversely, there are clearly examples of exemplary practice when primary care professionals provide continuity of highly skilled and supportive care to ensure the safety of women and their babies is maintained until each woman is ready to receive, and/or additional services are available to provide, the full care she needs. These experiences highlight the underlying variation in care depending on the skill and commitment of individual HPs and the importance of core and refresher training to improve the standard of care across all HP groups. These experiences also highlight the need for women and HPs to be aware of the personal journey of admittance and readiness for help as a key factor influencing the management of individual women into and on the care pathway. An integral factor to a woman's progression on this journey appears to be a lack of focus on the emotional well-being of the woman throughout the prenatal and postnatal periods. Women's experiences of contacts with HPs focusing, virtually exclusively, on the health and development of the baby do not create a patient-centred environment conducive to them opening up about their feelings or being ready to admit that they need help.

Strategies to improve screening/case-finding and mental health-care pathway

A number of strategies to improve access to and provision of mental health-care services for all women identified as having depression in the pre- or postnatal periods were proposed by the HP interview participants, including two specialist mental health HPs. Strategies include training, organisation of care, dedicated appointments for mental health assessment, women support groups, joined-up screening/case-finding with children's centres and screening/case-finding for fathers and partners (see *Appendix 13* for details).

Discussion

Consideration of study findings within the literature

Screening and case-finding for depression in the prenatal and postnatal periods is widely supported by pregnant women, mothers of new babies up to 12 months of age and HPs caring for those women and babies who participated in the EoA study (as described above). These women and HPs consider depression to be an important and often neglected health issue, which would benefit from being detected more effectively and earlier. This is consistent with findings from two systematic reviews, one regarding the acceptability of postnatal depression screening/case-finding[76] and another on the acceptability of perinatal depression screening/case-finding.[151]

Use of the cognitive evaluation method[149] in the acceptability survey ($n = 391$ prenatal women, $n = 345$ postnatal women) identified that women viewed the EPDS questions as easier to understand, easier to remember and easier to answer more confidently than the Whooley questions. There are widespread concerns regarding the lack of appropriateness and effectiveness of Whooley question 1 among expectant and new mothers and HPs who were interviewed in this study. An in-depth examination of these cognitive differences among the women interviewed identified that some pregnant women and new mothers were 'not comfortable' with Whooley question 1, resulting in approximately half of the women interviewed admitting that their answer to this question was not honest and was an understatement of their feelings of depression. MWs and HVs share these concerns, stating that women's body language and visible symptoms do not appear to match their answers to the Whooley questions resulting in most HPs adapting Whooley question 1 to instead ask a more general question about a woman's feelings or mood.

The reason for the reported discomfort and inability to answer the Whooley question openly and honestly was clearly stated by women and HP interviewees as the 'harshness' and 'severity' of terms in the question, namely 'depressed' and 'hopeless'. These women and HPs perceive that it is not socially acceptable for an

expectant or new mother to be depressed or to feel hopeless, and believe that they will naturally wish to avoid the negative sociocultural stigma attached to these terms, particularly as a new mother. Stigma associated with depression in expectant mothers and new mothers who are traditionally expected to be experiencing joy and fulfilment in their new role is widely documented and theorised in the literature.[103,152–155] It is the mother's own recognition of the discrepancy between the emotions she is feeling and the emotions society perceives as normal for motherhood that initiates the mother's symptoms of disappointment and depression. Robertson *et al.*[155] and Nicholson[156] have written extensively about the 'dangerous myths' operating among both professionals and lay people that equate becoming a mother with total fulfilment and happiness. Eight of the 18 studies in a meta-synthesis[155] centred on the role of conflicting expectations and experiences of motherhood in the development of postnatal depression.

Furthermore, women and HPs identified that many women are genuinely fearful of the potential consequences of being identified as depressed or feeling hopeless, which could result in the mother being 'monitored' in her own home and ultimately losing her baby.[103,152,157,158]

Another reason women and HP interviewees gave for women finding it difficult to answer the case-finding questions correctly is uncertainty of whether the common symptoms of tiredness and change in mood are associated with depression or with being pregnant or a new mother. The theoretical normalisation of symptoms of tiredness, changes in mood, for example, as a 'normal' part of motherhood and not depression has been widely documented.[103,152,159] The lack of differentiation between 'transient distress' and 'enduring distress' as a measure to discern when these symptoms are no longer appropriate has also been debated.[160]

Another factor that we identified as an important influence on a woman's willingness to answer a screening/ case-finding question honestly and/or accept help is whether or not the woman recognises that she has symptoms of depression and, if so, whether or not she has come to terms sufficiently with the social stigma to be prepared to admit these symptoms to a HP who is obliged to prioritise the welfare of the child over the mother. This finding is recognised in the literature as the concepts of 'making gains' and 'surrendering'.[155,161] 'Surrendering' is a big part of a mother's recovery from postnatal depression. The concept of surrendering means realising that something is very wrong and that help is needed. Unfortunately, women's initial interactions of 'surrendering' to their HPs can cause more distress; women reported that their concerns were ignored or minimised and 'feelings of disappointment, frustration, humiliation and anger were commonplace' (p. 55).[155,161] This resulted in women failing to 'make gains' and halting their progression on the journey of self-recognition through to surrendering.

In our study, women and HP interviewees identified a 'patient-centred environment' as a prerequisite to asking questions about mental health.[107,162] Many of the practical characteristics of such an environment are widely evidenced, including a private and quiet location, usually the home if privacy from the partner and family members can be achieved; a trained and skilled HP with good communication skills and ideally a good pre-existing relationship with the mother; a discursive approach around general well-being and the individual questions; and forewarning women of the process.[76,103,151,154,157–159,163] Importantly, the women interviewed also believed that a patient-centred environment would give equal priority to the mother's emotional well-being and the baby's physical development and would create a non-pressured and non-judgemental environment for women to feel that the HP is genuinely interested in supporting the woman if she is open about her feelings.[107,164]

Darwin *et al.*[103] described this patient-centred environment as the 'context of disclosure', and suggested that it will be more readily achieved in a research environment than in a routine care environment. Darwin *et al.*[103] explain this as the reason why different studies show different results for validation of the same screening/case-finding instruments depending on whether the evaluation of validity was conducted by a researcher in a research setting[100,165] or by a HP in a routine care setting.[162] This may partly explain why the BaBY PaNDA study found the Whooley and EPDS questions to have similar levels of sensitivity and specificity when compared to the CIS-R reference standard. In the case of the BaBY PaNDA study,

completion of the screening/case-finding instruments to assess their validity was conducted in a research setting by a researcher. However, within this research setting, the Whooley questions were asked and answers written down by an interviewer, the ideal scenario to generate the most open response from women, and the EPDS questions were self-completed by the expectant or new mother, a less than favourable scenario to generate the most open response from women. In the case of the in-depth qualitative interviews, both the Whooley and EPDS questions were asked individually to each woman by the interviewer and were discussed at length using the cognitive framework to ascertain women's views on their understanding, ability to remember and confidence to answer each question. Based on the above-stated importance of an appropriate 'context of disclosure' to elicit genuine responses from women, the patient-centred environment created for the in-depth qualitative interviews would have been most likely to have generated open, honest responses from these women for both the Whooley and EPDS questions. This would suggest that the concerns expressed by approximately half of the 25 women about the appropriateness and effectiveness of Whooley question 1 are internally valid and can be interpreted with confidence.

Hewitt *et al.*[76] reported that women who were found to have postnatal depression (EPDS score of ≥ 13) experienced statistically significantly more discomfort in completing the EPDS than women who were not found to have postnatal depression ($p < 0.0001$) and did not experience discomfort in completing the EPDS. A preliminary analysis of the interview data with women by their mental health status (based on a diagnosis from the CIS-R reference standard and self-reported history of depression and/or anxiety) does suggest that, within this sample of expectant and new mothers, women who were experiencing a current episode of depression were more likely to experience discomfort in answering Whooley question 1, whereas the small group of women who had a history of depression or anxiety with no current episode and some women who had a history with a positive care pathway experience followed by a current episode were more likely to have experienced a high level of comfort in answering Whooley question 1. These findings may be a result of women with a history of depression or anxiety having progressed positively through their personal journey of self-recognition and 'making gains' to actively seeking help at an early stage of their next episode of depression or anxiety and depression, and/or increased confidence in the health-care support they may receive.

Overall, these findings suggest that, even in the ideal scenario of HPs conducting their screening/case-finding assessments in a patient-centred environment with a patient-centred approach to screening/case-finding, women who have a current episode of depression without a positive care experience for a historical episode are likely to underestimate their feelings and feel uncomfortable answering Whooley question 1 honestly and openly.

Strengths and weaknesses of the evaluation of acceptability
The EoA achieved its overall aim of ascertaining valuable large-scale descriptive information from women about the acceptability of the Whooley and EPDS screening/case-finding instruments and explanatory in-depth information from women and HPs about the extent to which the questions each capture appropriate information for ascertaining the presence of depression in both the prenatal and postnatal periods. The purposive oversampling of women with a history and/or current case of depression for the in-depth interviews achieved a comprehensive exploration of the impact of being referred onto the mental health-care pathway in both the pre- and postnatal periods. Furthermore, the subsample of longitudinal interviews for a maximum of three occasions from 20 weeks' gestation through to 12 months postnatally provided a novel insight into the whole care pathway experience throughout the perinatal period. In-depth interviews with HPs provided a complementary perspective on the barriers to and facilitators of screening/case-finding and the provision of mental health care within the local organisation of care. Importantly, the combination of in-depth cognitive and care pathway experience from each woman with a validated diagnostic reference standard classification of her mental health status has afforded a unique insight into the complex and often inter-related explanatory factors underpinning the lack of appropriateness or effectiveness of individual case-finding questions (as detailed above).

As detailed in *Chapter 5*, the large and diverse sample of women completing the acceptability survey gives confidence in the generalisability of findings to women from similar population groups across the UK. One notable limitation is the under-representation of women from diverse ethnic backgrounds living in the study locations. The validity of the data may be compromised for some questions for some women who find a question confusing or do not agree with any of the answers available (for an example see Ritchie and Spencer[150]).

The sample of women participating in the in-depth interviews achieved target sample sizes (see *Results*) and diversity of characteristics including approximately equal numbers of women having their first or second child, a spread of different ages across the typical childbearing age range, a spread across study sites and an over-representation of women with a history, or current case, of depression. A key limitation in terms of generalisability to other locations is the under-representation of women from different ethnic backgrounds, a reflection of the general population living in the study locations. The effective sampling and achievement of data saturation within the large number of interviews provide confidence in the generalisability of findings to other women with similar characteristics regarding women's views on acceptability, appropriateness and effectiveness of the screening/case-finding questions and their experiences on the care pathway.[166,167]

Difficulties in recruiting HPs resulted in slightly fewer HPs taking part in the in-depth interviews than targeted (11 and 12, respectively). A broad range of issues was identified by individual HPs on the key topics of interest. Given the adherence to data collection protocols, topic guides and the skills and experience of the interviewer, the findings of individual HPs as a generic group of professionals who conduct screening/ case-finding for depression in the perinatal period can be treated with confidence in terms of the internal validity of data.[150,168] However, the small sample sizes of subgroups of MWs, HVs and HVs with a speciality in mental health warrants caution in generalising findings beyond these subgroups.

Women were asked to answer each screening/case-finding question based on how they were feeling that day and to discuss their cognitive views of each question based on their views and experiences that day. Participant recall was not a factor that may have affected the reliability of data.

Triangulation of study findings across the three data sources identified agreements and inconsistencies between the findings. Further analysis of the data to investigate descriptive inconsistencies generated meaningful explanatory insight into the underlying reasons why women and HPs had concerns regarding Whooley case-finding question 1.[169]

In summary, the appropriateness of the EoA research methods nested within the diagnostic accuracy study to collect data from a subsample of women and HPs who have direct experience of the screening/ case-finding instruments gives confidence in the internal validity of the findings. The acceptability survey generated large-scale, reliable descriptive data, although it may have some limitations regarding its validity for some women due to the uncertainty or limited options available for a self-complete method of a closed questionnaire.[170] Interview data from the HPs are considered valid but of limited generalisability and interview data from the women are considered both valid and generalisable. Confidence is at its highest for findings that are in agreement across all three data sources.

Recommendation for research
Mixed-methods research is needed to investigate the scope and complexity of relationships between a history of depression and/or anxiety, a positive or negative historical experience of the mental health-care pathway and women's honesty in responding to the Whooley screening/case-finding questions for subsequent screening/ case-finding episodes.

Implications for practice

The combined findings of the acceptability survey and in-depth interviews among expectant and new mothers, MWs and HVs to evaluate the acceptability of the Whooley and EPDS screening/case-finding instruments and their impact on the care pathway suggest the following implications for practice:

- The Whooley questions may have limitations in their acceptability, whereas fewer concerns were expressed about the EPDS.
- The evidence suggests that a patient-centred approach to screening/case-finding within a patient-centred environment will maximise the acceptability and effectiveness of any screening/case-finding instrument.
- Training of HPs on the appropriate implementation of any screening/case-finding instrument and discussion about emotional well-being within a patient-centred approach and environment has been evidenced as important.
- The normality of depression among women in the perinatal period could be emphasised as a first step to reducing the negative sociocultural stigma associated with depression.

Chapter 7 Cost-effectiveness of screening/ case-finding strategies for perinatal depression

The overall aim of the economic evaluation was to compare a range of screening/case-finding strategies for prenatal and postnatal depression, in terms of costs and health benefits, to establish their relative cost-effectiveness in the context of UK NHS. The analysis was conducted from the NHS/Personal Social Services perspective, with costs expressed in 2015/16 prices and health outcomes expressed in terms of QALYs. The time horizon for the postnatal model was 1 year, whereas the time horizon for the prenatal model was up to the point of birth; as such, no discounting of costs and health outcomes was required.

Methods

Overall approach of cost-effectiveness analysis
A decision model was developed to evaluate the costs and outcomes of screening/case-finding strategies for perinatal depression. Decision modelling is a systematic approach to decision-making under conditions of uncertainty. It is based on logical and temporal sequence of events (such as positive diagnosis followed by treatment) that would flow from a set of alternative options being evaluated. The likelihood of each pathway is expressed in terms of probabilities, and consequences are expressed in terms of costs and outcomes. Subsequently, the expected costs and outcomes of each consequence are calculated as the sum of the costs and outcomes weighted by the probability of each pathway.

Decision models also take into account uncertainty in the evidence base, including uncertainty in diagnostic accuracy of screening/case-finding strategies, costs and outcomes. Hence, these input data (i.e. parameters) enter the model as probability distributions to reflect uncertainty around the mean estimates. To simultaneously assess the joint impact of uncertainty in model inputs, we conduct PSA using Monte Carlo simulations. This is the recommended approach of propagating uncertainty in model parameters.[171] Based on this probabilistic analysis, the results provide the probability of each strategy being the most cost-effective, conditional on decision-makers' WTP for gain in health benefit. Finally, we also conducted a series of sensitivity analyses to evaluate the impact of model assumptions on the overall decision.

The following sections provide a detailed description of the screening/case-finding strategies evaluated in the structure of the decision model and the sources and assumptions about parameter inputs in the model. The cost-effectiveness analysis is reported according to the Consolidated Health Economic Evaluation Reporting Standards statement.[172]

Screening/case-finding strategies evaluated in the cost-effectiveness analysis
A number of strategies were considered for the cost-effectiveness analysis, based on the most recent NICE guidance (2014) on clinical management of antenatal and postnatal depression,[8] and reflecting clinical practice in the NHS sites involved in the BaBY PaNDA study. Both one- and two-stage screening/case-finding strategies were compared; these included the following:

- Whooley questions only
- EPDS only
- Whooley questions followed by EPDS
- Whooley questions followed by PHQ-9
- standard care case identification.

For the EPDS instrument, cut-off points of ≥ 10 and ≥ 13 were used in the cost-effectiveness analysis, which is in line with the findings of the statistical analysis reported earlier (see *Chapter 5*). For two-stage

screening/case-finding strategies, Whooley questions were used as the first instrument because they have higher sensitivity than EPDS and PHQ-9, and are also less time-intensive than the other instruments. In line with NICE guidance (Clinical Guideline 192),[8] the standard care case identification refers to routine clinical assessment that HPs would undertake to arrive at a diagnosis of depression in the perinatal period without the formal use of a diagnostic instrument. This was assumed to be provided by a GP and/or HV.

Structure of the decision model

We developed a decision tree model to evaluate the relative cost-effectiveness of screening/case-finding strategies. The model was based on the most recent NICE guidance on clinical management of antenatal and postnatal mental health.[8] The model had two linked components: (1) a case-finding model, which uses diagnostic performance data for each screening/case-finding strategy to determine rates of TP, FP, TN and FN outcomes, and applies the associated administrative cost; and (2) a treatment model, which evaluates the costs and health outcomes (measured in terms of QALYs) of each of the four diagnostic outcomes of part (1), TP, FP, TN and FN.

The decision model then attaches costs and health consequences to each pathway, and estimates the expected costs and outcomes of each screening/case-finding strategy as a product of probability of the pathway and the outcome/cost. The model was developed using Microsoft Excel 2013 (Microsoft Corporation, Redmond, WA, USA).

Screening/case-finding model

The structure of the screening/case-finding model follows the model developed as part of NICE guidance[8] and is presented in *Figures 9* and *10*. Both prenatal and postnatal models follow the same structure. The model pathways are described in detail below.

At the start of the model, all women in the prenatal or postnatal period (whether depressed or non-depressed) undergo screening/case-finding for depression. The prenatal model starts at 20 weeks into pregnancy whereas the postnatal model starts at 12 weeks after childbirth. The model assumes that screening/case-finding strategies are implemented by MWs during pregnancy and by HVs postnatally. This is in line with the time at which screening/case-finding strategies were implemented in this study. Women undergo either one- or two-stage screening/case-finding, as described above. If undergoing a one-stage strategy (including 'Whooley questions only' or 'EPDS only'), the model assumes that no further screening/case-finding will be offered. Two-stage screening/case-finding involves answering the Whooley questions first, followed by either EPDS or PHQ-9 if a woman is found to be depressed (positive) based on the Whooley questions.
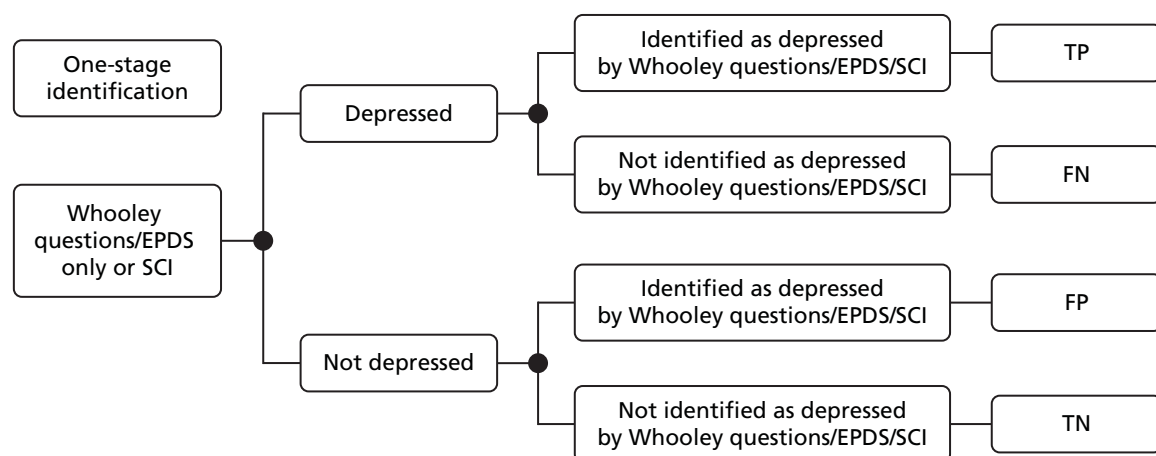
FIGURE 9 Diagnostic pathways: single screen. SCI, standard care case identification.
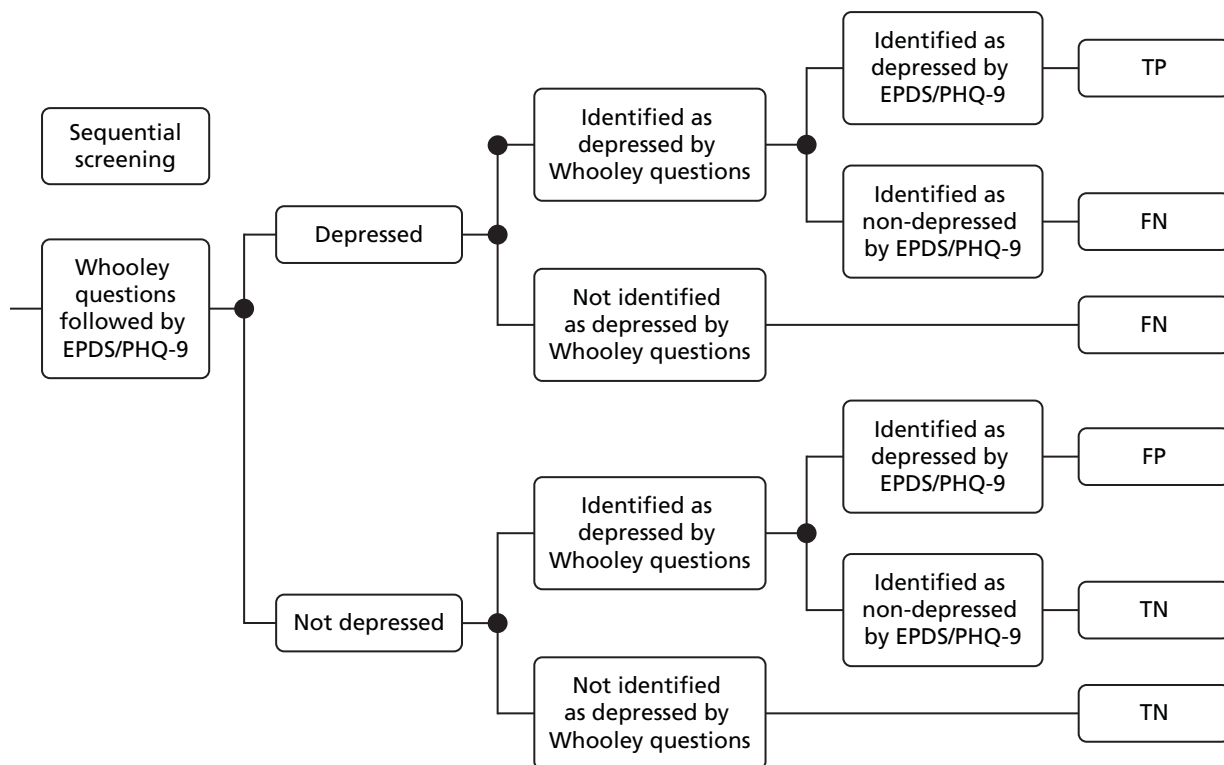
FIGURE 10 Diagnostic pathways: sequential screening/case-finding.

The sensitivity and specificity of the Whooley questions, EPDS and PHQ-9 were established using a diagnostic gold standard clinical assessment of depression (i.e. the CIS-R) (see *Chapter 5* for further details). Based on the outcome of the screening/case-finding strategy, women are classified in one of the four categories: (1) TP, (2) FP, (3) TN or (4) FN. Based on this classification, they are either treated or not treated. To simplify the model description, each of the pathways is described in detail below.

True positive Women who were depressed and correctly identified as depressed were classed as TP. The probability of TP is based on the sensitivity of a screening/case-finding strategy. For a single-stage strategy it is the sensitivity of the screening/case-finding instrument (such as Whooley questions or EPDS). For two-stage screening/case-finding, the overall sensitivity is equal to the product of the sensitivity of the first screening/case-finding strategy and the second strategy (as defined below).

Net sensitivity of two-stage screening = sensitivity of screen 1 × sensitivity of screen 2.   (1)

Alternatively, this can be represented as:

$$\text{Net sensitivity or net TP rate} = \left(\frac{TP_1}{TP_1 + FN_1}\right) \times \left(\frac{TP_2}{TP_2 + FN_2}\right). \tag{2}$$

Here TP and FN represent true positives and false negatives, respectively, 1 represents the first screening instrument (i.e. Whooley questions) and 2 represents the second instrument (i.e. EPDS or PHQ-9). The sensitivity of the second stage is estimated using the sample of positive cases (both true and false) from the first stage. Hence, the combined sensitivity is always less than the sensitivity of one-stage screening/case-finding (unless the sensitivity of the second stage is 1, in which case the combined sensitivity will be the same as the first-stage sensitivity).

Based on expert opinion in the NICE guidance,[8] all positive cases (including both TPs and FPs) were assumed to undergo further assessment by the HV or MW lasting approximately 1 hour; this was

considered only in terms of cost of the HV's/MW's time and not in terms of the impact on treatment pathway, as no studies were available that reported if further assessment had an impact on subsequent treatment pathway.

False positive Women who were not depressed but incorrectly identified as depressed were classed as FP. The probability of FP depends on the specificity of a screening/case-finding strategy. For single-stage screening/case-finding, the FPR is simply equal to 1 minus the specificity of the screening/case-finding strategy. For two-stage screening/case-finding, the overall specificity is equal to the sum of the two specificities minus the product of the two specificities, and the FPR is 1 minus the combined specificity (see below). Hence, the combined specificity of two-stage testing is higher than one-stage specificity. Therefore, FPR of a combined test is lower than the FPR of the first stage only.

True negative Women who were not depressed and correctly identified as non-depressed were classed as TN. The probability of TN depends on the specificity of a screening/case-finding strategy (i.e. specificity is the TN rate).

Net specificity of two-stage screening or net TN rate

$$= \left(\left(\frac{TN_1}{TN_1 + FP_1} \right) + \left(\frac{TN_2}{TN_2 + FP_2} \right) \right) - \left(\left(\frac{TN_1}{TN_1 + FP_1} \right) \times \left(\frac{TN_2}{TN_2 + FP_2} \right) \right). \tag{3}$$

Moreover, we note that:

Net FPR of two-stage screening $= 1-$ (net specificity of two-stage screening). $\tag{4}$

False negative Women who were depressed but falsely identified as non-depressed were classed as FNs. The probability of being FN depends on the sensitivity of a test (i.e. the FN rate is 1 minus sensitivity of a test). Hence, two-stage strategies are likely to have higher FN rates (unless sensitivity of the second stage is 1).

Treatment model
Based on the case-finding outcome, women were assigned to a care pathway and followed this pathway until the end of the follow-up period (i.e. at 1 year after screening/case-finding in the postnatal model or until birth in the prenatal model). The treatment pathway was also based on interventions recommended in the NICE guidance (2014) and used in the NICE model.[8] Below we describe the treatment pathways in detail and in *Figures 11–14* we present the schematics.

True positives and false positives Women who were TP or FP were assumed to receive one of the following treatment options, in proportions reflecting severity of depression as reported in NICE guidance: 72% of women were assumed to have subthreshold or mild to moderate depression and received FSH; the remaining women (28%) were assumed to receive intensive psychological therapy in the prenatal model, and either intensive psychological therapy (20%) or pharmacological treatment (8%). In the postnatal model intensive psychological therapy may be offered for up to 16 sessions; however, most studies used to derive the treatment effect reported in NICE guidance used between 8 and 12 sessions.[8] Therefore, we assumed 10 sessions of intensive psychological therapy with a CI of between 8 and 12 sessions. The pharmacological therapy in the postnatal model consisted of sertraline for 8 weeks plus 6 months of maintenance treatment.

In line with the NICE model, we assumed that FP women would receive the same treatments in the same proportions as TP women[8] but that they would stop the treatment earlier and would consume only 20% of the treatment-related health-care resources (based on consensus of the NICE Guidance Development Group).[8] However, we conducted sensitivity analysis of this assumption by varying the resource use incurred by FPs to be equal to 10% or 30% of the resource use by TPs.
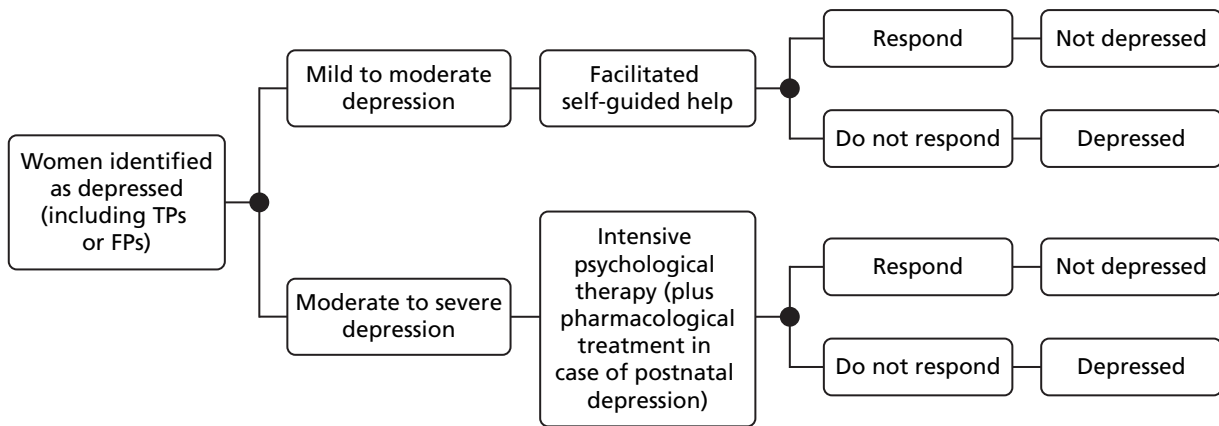
FIGURE 11 Treatment pathways: TP.
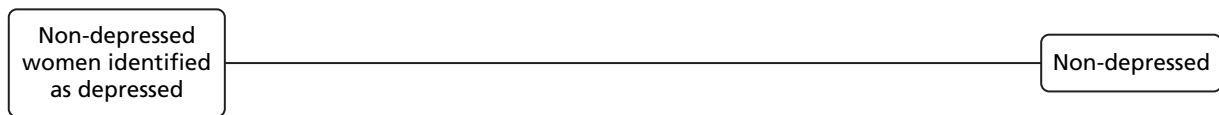
FIGURE 12 Treatment pathways: FP.
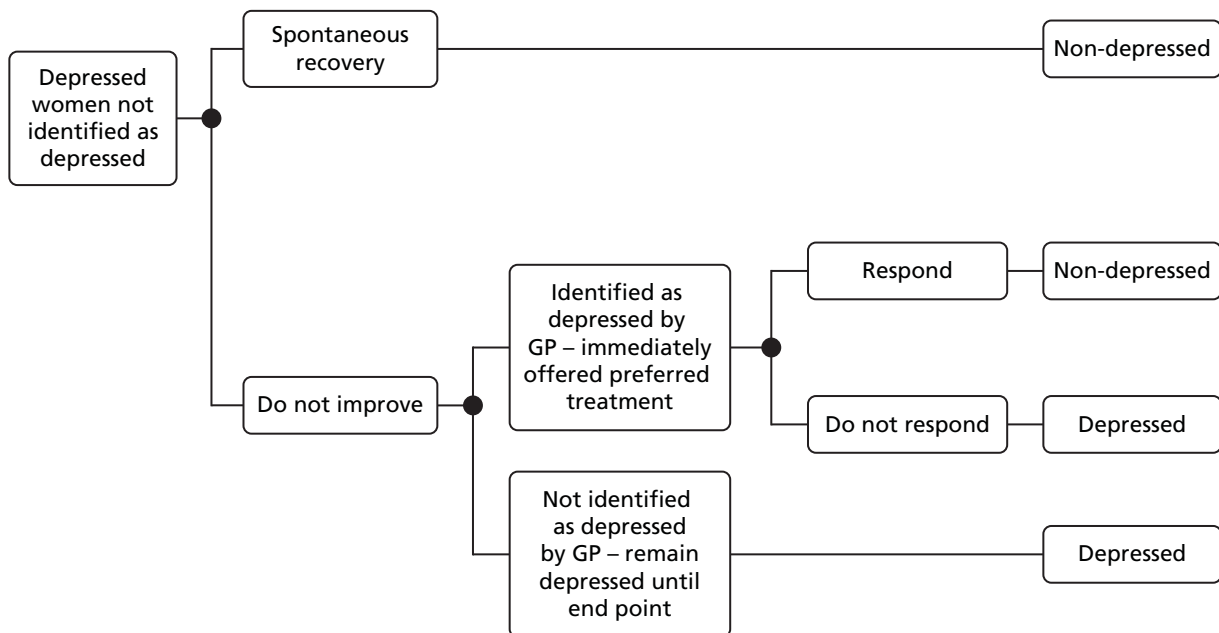
FIGURE 13 Treatment pathways: FN.
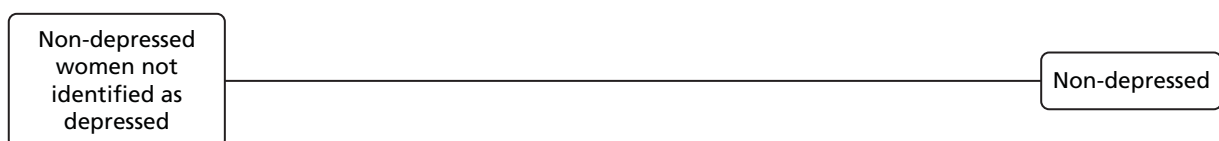
FIGURE 14 Treatment pathways: TN.

True negatives and false negatives Women who were TN were assumed to incur no further costs related to depression. It was assumed that a proportion of FN women will recover on their own. In line with NICE guidance, we assumed that women who do recover without treatment will incur additional health and social care costs while they are depressed – recovery was assumed to happen 6–7 weeks after implementing a screening/case-finding strategy (as in the NICE model).[8] However, if women did not recover spontaneously, they were assumed to have one GP visit half-way through the follow-up period during which their depression could be detected and in which case treatment would be offered in the same proportion as to TP women (described above). Finally, if women were not detected by their GP during the follow-up period, then they were assumed to stay depressed and also incur health and social care costs until the end point of the model. Owing to lack of relevant data and short time horizon, relapse was not modelled.

Parameters and assumptions of the decision model

Parameters and assumptions in the model are summarised in *Table 21* and discussed in detail below.

TABLE 21 Model parameters, assumptions and sources

Parameter	Value	Source
Prevalence of depression		
Prenatal	10.3% (95% CI 7.23% to 13.28)	BaBY PaNDA
Postnatal	10.5% (95% CI 7.18% to 13.78)	
Sensitivity of case-finding strategies[a] *Prenatal*	*Mean (log-odds, variance)*	
EPDS (≥ 10-point cut-off point)	82.5% (1.5506, 0.2033)	BaBY PaNDA
EPDS (≥ 13-point cut-off point)	45.0% (–0.2007, 0.1039)	
Whooley questions only	85.0% (1.7346, 0.2416)	
EPDS-10 (after Whooley questions)[b]	91.2% (2.3354, 0.3556)	
EPDS-13 (after Whooley questions)[b]	50.0% (0.000, 0.1197)	
PHQ-9 (after Whooley questions)	70.6% (0.8755, 0.1708)	
Postnatal	*Mean (log-odds, variance)*	
EPDS (≥ 10-point cut-off point)	82.9% (1.5755, 0.2490)	BaBY PaNDA
EPDS (≥ 13-point cut-off point)	62.9% (0.5261, 0.1360)	
Whooley questions only	85.7% (1.7918, 0.3024)	
EPDS-10 (after Whooley questions)[b]	93.3% (2.6391, 0.3339)	
EPDS-13 (after Whooley questions)[b]	73.3% (1.0116, 0.2178)	
PHQ-9 (after Whooley questions)	63.3% (0.5465, 0.1717)	
Specificity of case-finding strategies[a] *Prenatal*	*Mean (log-odds, variance)*	
EPDS (≥ 10-point cut-off point)	86.6% (1.8636, 0.0247)	BaBY PaNDA
EPDS (≥ 13-point cut-off point)	95.7% (3.1061, 0.0727)	
Whooley questions only	83.7% (1.6371, 0.0211)	
EPDS-10 (after Whooley questions)[b]	43.9% (–0.2469, 0.0752)	
EPDS-13 (after Whooley questions)[b]	77.2% (1.2192, 0.1079)	
PHQ-9 (after Whooley questions)	89.5% (2.1401, 0.2335)	

TABLE 21 Model parameters, assumptions and sources (*continued*)

Parameter	Value	Source
Postnatal	*Mean (log-odds, variance)*	
EPDS (≥ 10-point cut-off point)	87.6% (1.9574, 0.0313)	BaBY PaNDA
EPDS (≥ 13-point cut-off point)	95.7% (3.0910, 0.0860)	
Whooley questions only	80.6% (1.4244, 0.0213)	
EPDS-10 (after Whooley questions)[b]	53.4% (0.1382, 0.0795)	
EPDS-13 (after Whooley questions)[b]	79.3% (1.3437, 0.1188)	
PHQ-9 (after Whooley questions)	86.2% (1.8326, 0.1737)	
Cost of screening/case-finding		
Whooley questions	0.71 minutes (SE = 0.045) + 1 minute for scoring (HV/MW)	BaBY PaNDA
EPDS and PHQ-9	1.54 minutes (SE = 0.055) + 2 minutes for scoring (HV) for EPDS. PHQ-9 assumed to take the same time	BaBY PaNDA
Standard care case identification	11.7 minutes (one GP consultation) = £44 (SE = 4.4)	NICE (2014);[8] Curtis and Netten (2015).[173] SE assumed
Treatments		
TP	FSH (seven sessions): received by 72% (SE 0.072)	NICE (2014)[8]
	IPT (10 sessions; 95% CI 8 to 12 sessions), received by 28% in prenatal period and 20% in postnatal period	NICE (2014)[8]
	Pharmacological therapy (sertraline) for postnatal depression (received by 8%): 8 weeks of initial therapy followed by 6 months of maintenance therapy	NICE (2014)[8]
FP	Same as TP but assumed to be 20% of therapy received in base case	
FN and depressed women who do not respond to treatment	Additional health and social care	Petrou *et al.* (2002)[113]
Relative risk compared with usual care		
FSH	0.73 (95% CI 0.53 to 0.99)	Meta-analyses reported in NICE (2014)[8]
IPT or pharmacological therapy	0.48 (95% CI 0.39 to 0.60)	Meta-analyses reported in NICE (2014)[8]
Absolute risk of no improvement		
Usual care	0.667 (SE = 0.05)	NICE (2014)[8]
Cost of treatment		
FSH	Seven sessions (25 minutes per session) provided by psychological well-being practitioner (assumed rate of mental health nurse at £75 per hour of face-to-face contact; 95% CI £56 to £97); plus guided self-help manual costing £9.09.[174] In addition, three GP consultations at £44 [Curtis and Netten (2015)[173]]	Treatment assumptions based on NICE (2014)[8] and unit costs based on Curtis and Netten (2015)[173]
	Total cost of FSH = £360	

TABLE 21 Model parameters, assumptions and sources (*continued*)

Parameter	Value	Source
IPT	10 sessions (55 minutes each) at £94 per session (95% CI £83 to £115). In addition, three GP consultations lasting 11.7 minutes at £44 [Curtis and Netten (2015)[173]]	As above
	Total cost of IPT = £1072	
Pharmacological therapy	Sertraline for 8 weeks plus 6 months' maintenance. Unit cost of sertraline £1.57 per 28 tablets, 50-mg tablets[c]	As above. Drug cost from *British National Formulary* (2015)[175]
Weekly cost of health and social care incurred by women with depression	£9.00 per week	Petrou *et al.* (2002)[113]/NICE (2014),[8] adjusted to 2016
Utility of health states		
Depressed	Prenatal: 0.678 (SE = 0.04)	BaBY PaNDA
	Postnatal: 0.771 (SE = 0.03)	
Not depressed	Prenatal: 0.888 (SE = 0.01)	BaBY PaNDA
	Postnatal: 0.907 (SE = 0.01)	
Probability of being identified by GP (half-way through the follow-up period)	8.3%	Kessler *et al.* (2002)[176]
Spontaneous recovery rate	33% (SE = 6.6%; assumed)	Dennis *et al.* (2009)[177]

IPT, intensive psychological therapy; SE, standard error.
a Sensitivities and specificities are for each stage of screening/case-finding. Combined sensitivity of a two-stage screening/case-finding strategy = sensitivity of first stage × sensitivity of second stage. Combined specificity of a two-stage screening/case-finding strategy = specificity of stage 1 + specificity of stage 2 – (specificity of stage 1 × specificity of stage 2).
b EPDS-10 and EPDS-13 imply cut-off points of ≥10 and ≥13, respectively, on the EPDS scale.
c See *Methods* section for consultations for pharmacotherapy.

Clinical input parameters

The prevalence of depression was based on our primary study (as reported in *Chapter 5*). The prevalence of prenatal depression was 10.3% (95% CI 7.2% to 13.3%), whereas the prevalence of postnatal depression was 10.5% (95% CI 7.2% to 13.8%). Data on the diagnostic accuracy of screening/case-finding strategies were also based on our primary study (as reported in *Chapter 5*). Sensitivity and specificity data were specific to prenatal (20 weeks) and postnatal (3–4 months) periods. Diagnostic performance for second-stage screening was estimated conditional on the outcome of the first stage. We note that one of the limitations of the NICE model (due to lack of available evidence) was that the sensitivity and specificity of the second-stage screen was assumed to be independent of the first stage.

For standard care case identification, we used the sensitivity and specificity parameters reported in NICE guidance (2014), which were based on a meta-analysis of 118 studies that assessed accuracy of identifying depression by GPs.[8,178] The authors reported weighted sensitivity and specificity of 50.1% and 81.3%, respectively, which were used in our model to approximate sensitivity and specificity of routine care case identification (as in the NICE model). These parameters are based on cases identified without the use of formal screening/case-finding instruments but as part of routine assessment. It should be noted that standard care case identification does not include conducting a gold standard interview to confirm initial diagnosis. Hence, HPs may incorrectly identify positive and negative cases.

Treatment effects of FSH and high-intensity psychological therapy were based on the meta-analysis reported in NICE guidance (2014).[8] The rate of response to pharmacological therapy was assumed to be the same as the rate of response to high-intensity psychological therapy. See NICE guidance for further details about the meta-analysis,[8] and see *Table 21* for estimates of treatment effects. Note that treatment

effects are reported as relative risk of no improvement (compared with usual care); hence, a relative risk of < 1 is desirable. More specifically, the absolute risk of no improvement in the usual care group is 0.667 [i.e. 33.3% (1 – 0.667) of women with depression will recover under usual care]. However, women receiving FSH have a relative risk of no improvement of 0.73, which is multiplied by the absolute risk of no improvement in the usual care group to arrive at an absolute risk of no improvement of 0.487 (= 0.73 × 0.667) in the FSH group. In other words, women receiving FSH have a 51.3% (1 – 0.487) chance of recovering under FSH. Owing to lack of data on treatment effect in the prenatal period, our prenatal model assumed the same treatment effect as the postnatal model.

As mentioned earlier, a proportion of women will spontaneously recover from perinatal depression. A review conducted by Dennis *et al.*[177] reported that the rate of spontaneous recovery from postnatal depression was between 25% and 40% in the control arms of randomised trials. Consistent with the NICE model, we used the mid-point of this range to represent the proportion of women who would enter remission. This is also consistent with the meta-analysis of standard care arms reported in NICE guidance.[8] We assumed the same recovery rate for prenatal depression.

Women who do not spontaneously enter remission may be detected by their GP through routine care. The model assumed that these women will have one GP consultation halfway through the follow-up period, during which their depression could be detected and treatment offered. Previous models[8,81] have used an estimate from a study in general practice.[176] Based on this study, 8% of women who did not enter spontaneous remission would be detected halfway through the follow-up period.

Resource use and cost data

Resource use data were based on a number of sources, including our primary study, NICE guidance (2014)[8] and national cost databases. The times required to administer the Whooley questions (under 1 minute) and EPDS (under 2 minutes) were based on our primary study. We added 1 and 2 minutes, respectively, for scoring of Whooley questions and EPDS. PHQ-9 was assumed to require the same amount of time as EPDS. The cost of a HV's time was based on NHS reference cost database (2014–15).[179] Routine care case identification was assumed to require one GP consultation lasting 11.7 minutes.[173]

The cost of treatments was based on NICE guidance (2014).[8] Of the positive cases, 72% received a mean of seven sessions of FSH (each lasting 25 minutes) with a psychological well-being practitioner whose costs were assumed to be equal to those of a mental health nurse (£75 per hour of face-to-face contact[173]). In addition, a self-help manual was costed at £9.09 (www.amazon.co.uk).[174]

For intensive psychological therapy, the NICE guidance (2014) included six studies in the meta-analysis to derive the treatment effect.[8] The number of sessions in these studies varied between studies from three to more than 15; however, most studies used 8–12 sessions. Hence, for the economic analysis, we assumed 10 sessions for the base case and conducted sensitivity analyses using 8 and 12 sessions. Each session lasted 55 minutes and was costed at £94 [NHS reference cost database (2014–15)[179]]. Based on NICE expert group opinion (2014),[8] women receiving intensive psychological therapy or self-help would receive additional care that would comprise three GP consultations.

Women receiving pharmacological therapy in the postnatal model would receive 8 weeks of initial therapy with sertraline followed by 6 months of maintenance therapy (i.e. 50 mg per day). The cost of sertraline was based on the most recent version of the *British National Formulary*.[175] In addition, these women were actively monitored in primary or secondary care. Based on NICE guidance (2014),[8] it was assumed that 15% would receive two consultant psychiatrist visits (one lasting 30 minutes and one lasting 15 minutes) and two consultations with the GP; the remaining 85% of women receiving pharmacological therapy were assumed to be managed in primary care and to have four GP consultations. The unit cost of a GP consultation was £44.[173]

False-negative women (i.e. those not identified as depressed by screening/case-finding strategies) were assumed to incur additional health and social care costs while they are depressed. Following the NICE

model, the cost of health and social care was based on Petrou *et al.*,[113] who estimated the cost of health and social care in a cohort of women with postnatal depression (including cost of primary and secondary care). These costs were adjusted to 2016 and were £9.00 per week.

All costs were expressed in 2015–16 prices. Discounting of costs and outcomes was not necessary owing to the short time horizon of the model.

Utility data and estimation of quality-adjusted life-years
Utility data were based on primary data collected in our study using the EQ-5D,[141] which is the standard generic instrument used to estimate HRQoL. Responses to the EQ-5D were converted into QALYs using the UK tariff, and QALYs were estimated using an AUC approach.

Women were assigned utility values based on the diagnostic and treatment pathways. TP cases were assigned the utility of depression at the time of screening/case-finding. Women who recovered after treatment were assumed to have a linear improvement in quality of life until the end of therapy (estimated using linear interpolation). Women who did not respond to treatment were assumed to continue with the utility value of depression. FP cases were assumed to have no utility decrement due to false diagnosis in the base-case model; however, in a sensitivity analysis we assumed a 2% reduction in utility due to a FP diagnosis (in line with the NICE model). TN cases had the utility level of non-depressed women. FN women who entered spontaneous remission were assumed to recover linearly over 7 weeks after initial screening/ case-finding (as per NICE model), whereas those who did not recover continued with the utility of depression unless identified and treated in routine care (as per the model pathway).

Model parameters and assumptions are summarised in *Table 21*.

Analysis
Pathways in the decision tree were evaluated using probabilities based on the sensitivity and specificity of screening/case-finding strategies and subsequent probabilities of treatment and response. More specifically, the model started with a cohort of depressed and non-depressed women, with the proportion of depressed women determined by the prevalence of depression in prenatal and postnatal periods (see *Table 21* for parameters). For instance, based on a prevalence estimate of 10.5% in the postnatal period, 105 women in a cohort of 1000 women would be depressed. Both depressed and non-depressed groups were screened using either a one- or two-stage screening/case-finding strategy. Based on the sensitivity and specificity of the screening/case-finding strategy, the outcomes were TP, FP, TN or FN. For instance, using the Whooley questions as the first screen, 85.7% of depressed women (= the sensitivity of Whooley questions) in the postnatal period were identified as positive [i.e. TP = 90 women in a cohort of 1000 women (this is calculated as 105 depressed women times 85.7% sensitivity)]. As a result, 14.3% of depressed women were missed by this strategy (i.e. FN). Also, based on the specificity of Whooley questions, 80.6% of non-depressed women were correctly identified (i.e. TN), whereas 19.4% were incorrectly identified as depressed (i.e. FP). In the one-stage strategy based on Whooley questions alone, both women identified as TP and those identified as FP were offered treatment at this stage, as described earlier; however, those identified as TN or FN were not offered treatment. In the two-stage strategy, women identified as TP or FP were screened again, which reduced the number of FPs, but some more TPs were incorrectly identified as FNs. Subsequently, costs and outcomes were applied to each pathway, as mentioned earlier. Expected costs and QALYs were calculated for each pathway. The ICER was calculated as a ratio of the difference in costs and difference in QALYs compared with the next most effective screening/case-finding strategy. A screening/case-finding strategy is considered 'dominated' (and therefore ruled out) if it has higher costs and lower QALYs than another strategy. Moreover, 'extended dominance' rules out any strategy for which the ICER is higher than that of the next most effective strategy. Net monetary benefit was calculated as the total QALYs of each strategy times the WTP threshold minus the total cost of each strategy.

Probabilistic sensitivity analysis was conducted using 10,000 Monte Carlo simulations. For PSA, normal distribution was assumed for the following parameters: prevalence of depression; sensitivity and specificity

of screening/case-finding tests; and absolute and relative risk of improvement. A beta distribution was assumed for utility weights and for the probability that depression is identified by GP during the follow-up period. Unit costs were assumed to be fixed. Using simulation results, cost-effectiveness acceptability curves and frontiers were plotted to present the probability of each screening/case-finding strategy being most cost-effective for a WTP threshold between £0 and £100,000. Further sensitivity analyses were conducted by (1) varying the prevalence of perinatal depression, (2) assuming 2% reduction in utility in FP cases (as per the NICE model) and (3) varying the level of resource use by FPs (for further details of the sensitivity analyses see *Results*).

Another health economist who was independent of the research group and not part of the model development process checked the decision model. The model was also checked for logical consistency by setting input parameters to null and extreme values and examining the direction of change of results.

Results

The results are presented separately for the prenatal and postnatal models. For each model, the base-case results are presented using the data sources and assumptions discussed in the *Methods* section; this is followed by sensitivity analyses to assess the impact of using alternative parameters and assumptions. We present the postnatal results first, followed by the prenatal model results.

Base-case results: postnatal depression

Table 22 presents a summary of the results of base-case cost-effectiveness analysis of screening/case-finding strategies for postnatal depression. Strategies were ranked based on estimates of average cost per woman screened (from the least expensive to the most expensive). Deterministic analysis shows that the average cost per woman varies between £66.30 for 'Whooley questions followed by PHQ-9' and £103.50 for 'Whooley questions only'. The difference between the most and least expensive strategy in terms of QALYs was only 0.00078 per woman screened. The ICERs were compared with the next cheapest (non-dominated) strategy. *Table 22* also shows that routine care case identification strategy was dominated by other strategies because it was more expensive and less effective than the next low-cost strategy in the ordered list. One-stage EPDS (≥ 13-point cut-off point) was extendedly dominated, whereas one-stage EPDS (≥ 10-point cut-off point) had

TABLE 22 Total costs, QALYs and probability of cost-effectiveness of screening/case-finding strategies for postnatal depression

Strategy	Total cost (£)	Total QALYs	ICER (£)[a]	Probability of cost-effectiveness for maximum WTP		
				£13,000	£20,000	£30,000
Whooley questions followed by PHQ-9	66.30	0.8986	–	0.554	0.432	0.345
Whooley questions followed by EPDS (≥ 13-point cut-off point)	71.80	0.8989	25,929	0.116	0.100	0.080
EPDS (≥ 13-point cut-off point) only	73.40	0.8989	Extendedly dominated	0.102	0.099	0.087
Routine care case identification	79.30	0.8985	Dominated	0.007	0.010	0.016
Whooley questions followed by EPDS (≥ 10-point cut-off point)	86.40	0.8993	34,489	0.112	0.154	0.168
EPDS (≥ 10-point cut-off point) only	93.80	0.8994	105,352	0.091	0.153	0.207
Whooley questions only	103.50	0.8994	137,883	0.020	0.052	0.097

a Dominated implies that the strategy has higher costs hand lower QALYs than the comparator. Extended dominance implies that the ICER of this strategy is higher than the ICER of the next most effective strategy.
Strategies with the highest probability of being cost-effective at different WTP thresholds are shaded.

an ICER of £105,352 compared with the next cheapest strategy. One-stage Whooley questions had an ICER of £137,883 compared with EPDS ($\geq$ 10-point cut-off point) only. Hence, one-stage identification strategies were either dominated (or extendedly dominated) by two-stage screening or had an ICER that was significantly above the conventional WTP threshold range of £20,000–30,000 per QALY.

Costs and QALYs were jointly evaluated using net monetary benefit at conventionally used cost-effectiveness thresholds of £20,000 and £30,000 per QALY (see *Appendix 14*). The results show a very small difference between strategies in terms of net monetary benefit. The most cost-effective strategy in the range between £20,000 and £30,000 per QALY was 'Whooley questions followed by PHQ-9'. However, the difference between this and the next most likely cost-effective strategy, in terms of net monetary benefit, was < £15 per woman for the conventionally used WTP thresholds.

Given uncertainties in costs and QALYs, ICERs and net benefits should be interpreted in the light of the probabilistic cost-effectiveness analysis presented in *Table 22*, and also graphically presented as cost-effectiveness acceptability frontier in *Figure 15*; the figure shows that 'Whooley questions followed by PHQ-9' is most cost-effective in the range between £20,000 and £30,000 per QALY (see *Appendix 14* for cost-effectiveness acceptability curves for all strategies). It was also the most cost-effective strategy at a threshold of £13,000 per QALY (note: this is the lower threshold proposed in a recent report).[180] *Figure 15* shows that 'Whooley questions followed by PHQ-9' is cost-effective up to a WTP threshold of £51,500, beyond which EPDS ($\geq$ 10-point cut-off point) becomes the most cost-effective strategy. However, these results should be interpreted in the light of net monetary benefits, which show a very small monetary difference between strategies.

Our analysis also found that Whooley questions alone was never a cost-effective strategy even at threshold of £100,000 per QALY. Moreover, when comparing EPDS alone with Whooley questions in terms of probability of cost-effectiveness, the EPDS alone always had a higher probability of being cost-effective than the Whooley questions alone.

To better understand these results, *Table 23* compares screening/case-finding strategies in terms of sensitivity, specificity and health gains for a hypothetical cohort of 1000 postnatal women who undergo screening/case-finding. *Table 23* shows that 'Whooley questions only' has the lowest specificity, of 80.6%, which implies a FPR of 19.4% (i.e. 100% – 80.6%). Given the prevalence of postnatal depression in the model, this FPR implies that 174 per 1000 women will be incorrectly identified as depressed. If these FP women receive 20% of the full treatment (as assumed in the base-case model), this implies that
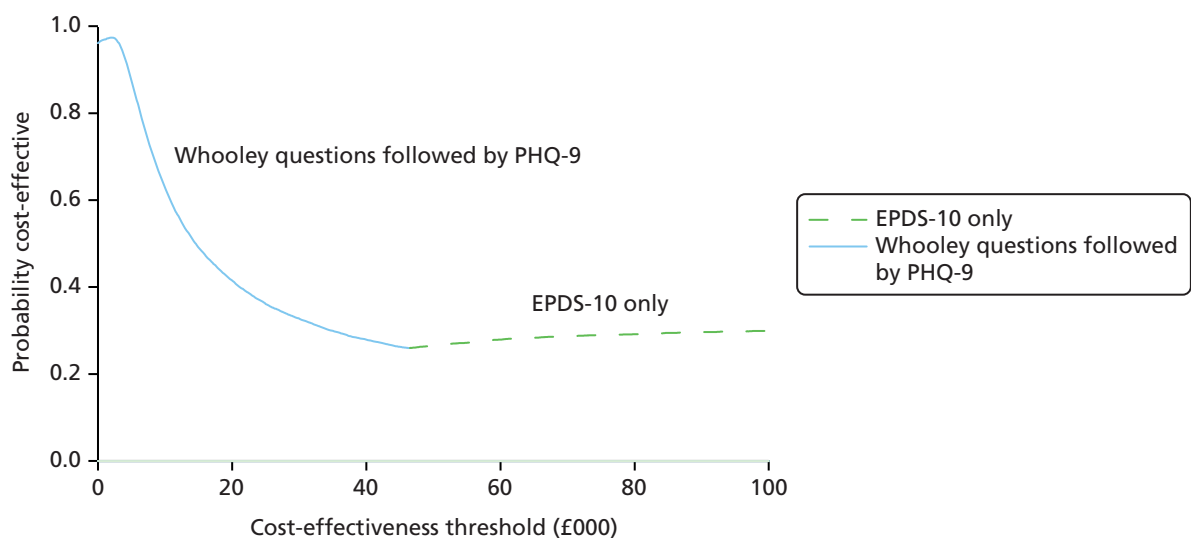
FIGURE 15 Cost-effectiveness acceptability frontier of diagnostic strategies for postnatal depression.

TABLE 23 Sensitivity, specificity and diagnostic outcomes of a hypothetical cohort of 1000 women screened for postnatal depression

Screening/ case-finding strategy	TPs per 1000 women screened		FPs per 1000 women screened		TNs per 1000 women screened		FNs per 1000 women screened		
	Number of TPs	QALYs of TPs	Number of FPs	QALYs of FPs	Number of TNs	QALYs of TNs	Number of FNs	QALYs of FNs	Total QALYs
Whooley questions followed by PHQ-9	57	48	24	22	871	790	48	39	898.65
Whooley questions followed by EPDS (≥ 13-point cut-off point)	66	55	36	33	859	779	39	32	898.86
EPDS (≥ 13-point cut-off point) only	66	55	39	35	856	777	39	32	898.86
Whooley questions followed by EPDS (≥ 10-point cut-off point)	84	70	81	73	814	739	21	17	898.55
Routine care case identification	52	44	167	152	728	660	52	43	899.28
EPDS (≥ 10-point cut-off point) only	87	73	111	100	784	711	18	15	899.36
Whooley questions only	90	75	174	158	722	654	15	12	899.43

£30,043 (= 174 FP women × £173 cost of FP treatment) will be unnecessarily spent as a result of incorrect identification of non-depressed women. In comparison, the 'Whooley questions followed by PHQ-9' strategy has a FPR of 2.7% (combined specificity = 97.3%; see *Methods*) and would therefore incorrectly identify only 24 women per 1000 as depressed, resulting in relatively low unnecessary treatment cost of £4144 (i.e. £25,899 less than Whooley questions alone). This shows why specificity is a key driver of the cost-effectiveness results. It should also be noted that the base-case model makes the conservative assumption that, while a FP identification is associated with additional costs, it does not cause loss of quality of life as a result of false diagnosis. Relaxing this assumption (see *Sensitivity analysis*) makes 'Whooley questions followed by PHQ-9' even more likely to be cost-effective.

Table 23 shows that the 'Whooley questions followed by PHQ-9' strategy has the second lowest sensitivity (after routine care case identification) whereas 'Whooley questions only' has the highest sensitivity, of 85.7%. High sensitivity of 'Whooley questions only' implies that more cases of depression will be identified and given treatment – this is reflected in the highest number of cases identified and total QALYs in the 'Whooley questions only' strategy. However, the incremental difference in QALYs between 'Whooley questions only' and 'Whooley questions followed by PHQ-9' is relatively low. This is because of a number of factors, including (1) the prevalence of perinatal depression in the model – this implies that a gain in sensitivity by 10% would result in only 10.5 additional TP diagnoses per 1000 women; (2) the magnitude of the relative treatment effect for therapies offered to women identified as TP – the most common treatment in the model (i.e. FSH) reduces the relative risk of no improvement by 27%; (3) the spontaneous recovery rate of FN cases – based on the NICE guidance document, the model assumes that 33% of FN women will enter spontaneous remission within 6–7 weeks of screening/case-finding; and (4) the routine care identification of missed cases of depression (i.e. FNs) during the follow-up period. Taking account of these factors (see *Methods* for the parameter values), the difference in QALYs between 'Whooley questions only' and 'Whooley questions followed by PHQ-9' strategies is 0.78 per 1000 women screened. This shows why a difference in sensitivity has a relatively smaller impact on cost-effectiveness compared with a similar improvement in specificity.

Base-case results: prenatal depression

Table 24 presents a summary of the results of the base-case cost-effectiveness analysis of screening/case-finding strategies for prenatal depression. Strategies were ranked based on estimates of average cost per woman screened (from the least expensive to the most expensive). Deterministic analysis shows that the average cost per woman varies between £49.20 for 'Whooley questions followed by EPDS ($\geq$ 13-point cut-off point)' and £91.60 for 'Whooley questions only'. Difference between the most and least expensive strategy in terms of QALYs was only 0.00049 per woman screened. One-stage EPDS at $\geq$ 10 and $\geq$ 13 cut-off points was extendedly dominated by less expensive strategies, whereas one-stage 'Whooley questions only' had an ICER of £171,990 compared with a two-stage strategy of 'Whooley questions followed by EPDS' ($\geq$ 10-point cut-off point). Hence, one-stage identification strategies were either extendedly dominated by two-stage screening or had an ICER that was significantly above the conventional WTP threshold range of £20,000–30,000 per QALY.

Costs and QALYs were jointly evaluated using net monetary benefit at conventionally used cost-effectiveness thresholds of £20,000 and £30,000 per QALY (see *Appendix 14*). The results show a very small difference between strategies in terms of net monetary benefit. Two strategies are cost-effective in the range between £20,000 and £30,000 per QALY [i.e. 'Whooley questions followed by PHQ-9' and 'Whooley questions followed by EPDS ($\geq$ 13-point cut-off point)']. The difference in net monetary benefit between these strategies at a WTP threshold of between £20,000 and £30,000 per QALY was < £4 per woman.

As in the case of postnatal depression, our analysis found that 'Whooley questions alone' were never a cost-effective strategy, even at threshold of £100,000 per QALY. Moreover, when comparing the EPDS alone with the Whooley questions in terms of probability of cost-effectiveness, the EPDS alone always had higher probability of being cost-effective than the Whooley questions alone.

The cost-effectiveness results were evaluated probabilistically and are presented in *Table 24* and also graphically as a cost-effectiveness acceptability frontier in *Figure 16*. *Table 24* and *Figure 16* show that 'Whooley questions followed by PHQ-9' is most cost-effective in the range between £20,000 and £30,000 per QALY, although at £20,000 per QALY 'Whooley questions followed by EPDS ($\geq$ 13-point cut-off point)'

TABLE 24 Total costs, QALYs and probability of cost-effectiveness of screening/case-finding strategies for prenatal depression

Strategy	Total cost (£)	Total QALYs	ICER (£)[a]	Probability of cost-effectiveness for maximum WTP		
				£13,000	£20,000	£30,000
Whooley questions followed by EPDS ($\geq$ 13-point cut-off point)	49.20	0.33607	–	0.650	0.461	0.336
EPDS ($\geq$ 13-point cut-off point) only	52.80	0.33610	Extendedly dominated	0.033	0.039	0.047
Whooley questions followed by PHQ-9	55.10	0.33627	28,958	0.316	0.466	0.480
Routine care case identification	68.40	0.33615	Dominated	0.001	0.001	0.001
Whooley questions followed by EPDS ($\geq$ 10-point cut-off point)	76.80	0.33647	107,222	0.001	0.019	0.065
EPDS ($\geq$ 10-point cut-off point) only	87.80	0.33653	Extendedly dominated	0.000	0.011	0.052
Whooley questions only	91.60	0.33656	171,990	0.000	0.002	0.020

a Dominated implies that the strategy has higher costs and lower QALYs than the comparator. Extended dominance implies that the ICER of this strategy is higher than the ICER of the next most effective strategy.
Strategies with the highest probability of being cost-effective at different WTP thresholds are shaded.
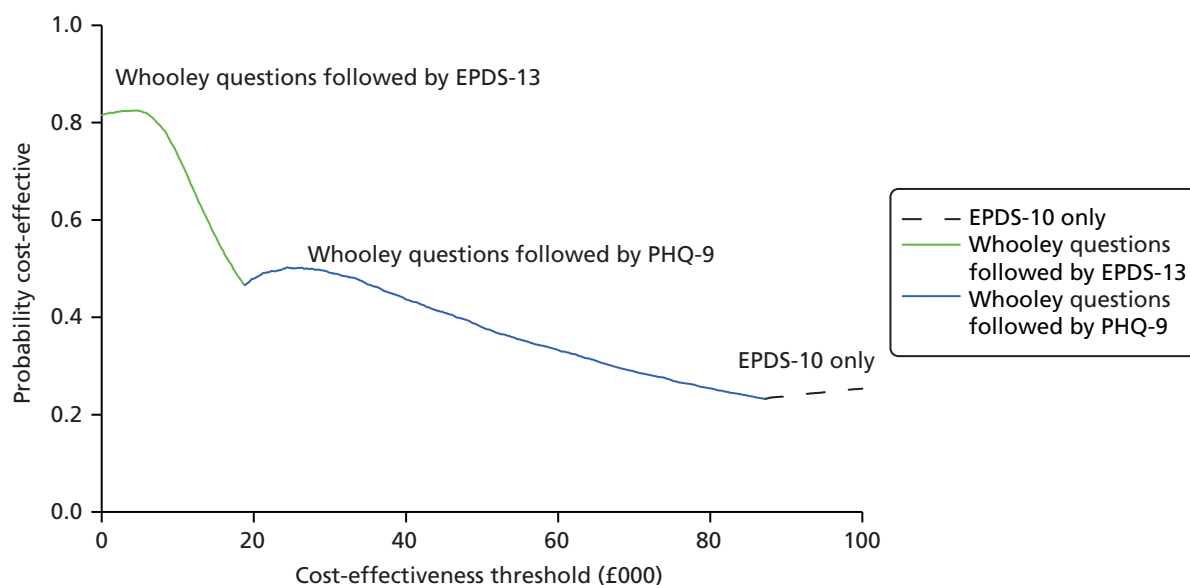
FIGURE 16 Cost-effectiveness acceptability frontier of diagnostic strategies for prenatal depression.

is almost equally cost-effective. Using a lower threshold of £13,000 per QALY, 'Whooley questions followed by EPDS (≥ 13-point cut-off point)' would be the most cost-effective strategy. However, these results should be interpreted in the light of net monetary benefits, which show a very small monetary difference between strategies.

Table 25 compares screening/case-finding strategies in terms of sensitivity, specificity and health gains for a hypothetical cohort of 1000 prenatal women who undergo screening/case-finding. The table shows that routine care case identification and 'Whooley questions only' strategies have the lowest specificities of

TABLE 25 Sensitivity, specificity and diagnostic outcomes of a hypothetical cohort of 1000 women screened for prenatal depression

Screening/ case-finding strategy	TPs per 1000 women screened		FPs per 1000 women screened		TNs per 1000 women screened		FNs per 1000 women screened		
	Number of TPs	QALYs of TPs	Number of FPs	QALYs of FPs	Number of TNs	QALYs of TNs	Number of FNs	QALYs of FNs	Total QALYs
Whooley questions followed by EPDS (≥ 13-point cut-off point)	44	13	33	11	864	295	59	17	336.07
EPDS (≥ 13-point cut-off point) only	46	14	38	13	859	293	56	16	336.10
Whooley questions followed by PHQ-9	62	18	15	5	882	301	41	12	336.27
Routine care case identification	51	15	168	57	730	249	51	14	336.15
Whooley questions followed by EPDS (≥ 10-point cut-off point)	79	23	82	28	815	278	23	7	336.47
EPDS (≥ 10-point cut-off point) only	85	25	121	41	777	265	18	5	336.53
Whooley questions only	87	26	146	50	751	257	15	4	336.56

81.3% and 83.7%, respectively. For 'Whooley questions only' this implies a FPR of 16.3%. For the prevalence of prenatal depression in the model, this FPR implies that 146 per 1000 women will be incorrectly identified as depressed using 'Whooley questions only'. If these FP women receive 20% of the full treatment (as assumed in the base-case model), this implies that £27,309 will be unnecessarily spent due to incorrect identification of non-depressed women. In comparison, the 'Whooley questions followed by PHQ-9' strategy has a FPR of 1.7% and would therefore incorrectly identify only 15 women per 1000 as depressed, resulting in relatively low unnecessary treatment cost of £2875 (i.e. £24,434 less than 'Whooley questions alone'). Similar analysis can be done for 'Whooley questions followed by EPDS (≥ 13-point cut-off point)'. These results show why specificity is a key driver of the cost-effectiveness results. Also, as mentioned earlier, the base-case model makes the conservative assumption that although FP is associated with additional costs, it does not cause loss of quality of life due to false diagnosis. Relaxing this assumption (see *Sensitivity analysis*) makes 'Whooley questions followed by PHQ-9' and 'Whooley questions followed by EPDS (≥ 13-point cut-off point)' even more likely to be cost-effective.

In the prenatal model, sensitivity is found to be even less influential than the postnatal model. This is partly because of the short follow-up period (20 weeks) during which the QALY benefits of recovery from depression in the TP women are realised. Moreover, FN women (those not identified as depressed) who do not recover from depression do not have a long period with depression until birth. Therefore, in the prenatal model, specificity is even more influential in the cost-effectiveness analysis.

Sensitivity analyses: prenatal and postnatal depression
The base-case analysis made a number of assumptions about model parameters. We tested the robustness of our results by changing assumptions about model parameters. Three sets of sensitivity analyses were conducted, which are described below, along with the results.

Sensitivity analysis 1: prevalence of depression
The base-case model used prevalence estimates of prenatal and postnatal depression from the BaBY PaNDA study. However, lower and higher estimates have also been reported in the literature. For instance, the NICE postnatal model used a lower prevalence estimate of 8.7% based on a pragmatic randomised controlled trial of postnatal depression conducted in the UK.[8,143] Higher estimates have also been reported in the literature. Hence, we conducted sensitivity analyses by varying the prevalence parameter between 8.7% and 12% for both prenatal and postnatal depression.

Tables 26 and *27* show the results of sensitivity analyses of prenatal and postnatal models, respectively. When lower prevalence of 8.7% is assumed in the model, the probability of 'Whooley questions followed by PHQ-9' being the most cost-effective strategy increases in both prenatal and postnatal models, whereas the opposite is observed when prevalence increases to 12%. This is not surprising, since lower prevalence implies that specificity is even more important because of the higher number of women without depression who could be misdiagnosed as FP by a strategy with lower specificity, and therefore incur additional costs. The results of prenatal and postnatal models are generally similar to the base-case analysis, except that at higher prevalence of 12% 'Whooley questions followed by EPDS (≥ 13-point cut-off point)' is cost-effective at £20,000 per QALY in the prenatal model; however, in all other comparisons, 'Whooley questions followed by PHQ-9' remains the most effective strategy.

Sensitivity analysis 2: reduction in quality of life in false positives
Our base-case model assumed that FP diagnosis is only associated with excess health-care cost, but does not result in reduction in quality of life. This assumption is common in most diagnostic studies. However, we checked the robustness of our assumption by assuming a 2% reduction in quality of life due to FP diagnosis of perinatal depression (as assumed in the NICE model).[8]

The results for prenatal and postnatal models are presented in *Table 28*. Both prenatal and postnatal results show that 'Whooley questions followed by PHQ-9' is the most cost-effective strategy in the range between £20,000 and £30,000 per QALY. Moreover, the probability of this strategy being the most

TABLE 26 Sensitivity analysis of varying the prevalence parameter in the prenatal model

Strategy	Prevalence = 8.7%				Prevalence = 12%			
	Total cost (£)	Total QALYs	Probability of being cost-effective at WTP of		Total cost (£)	Total QALYs	Probability of being cost-effective at WTP of	
			£20,000	£30,000			£20,000	£30,000
Whooley questions followed by EPDS (≥ 13-point cut-off point)	43.30	0.33690	0.420	0.315	55.90	0.33514	0.508	0.359
EPDS (≥ 13-point cut-off point) only	46.60	0.33692	0.030	0.037	59.60	0.33517	0.053	0.053
Whooley questions followed by PHQ-9	47.70	0.33707	0.524	0.533	63.40	0.33537	0.399	0.428
Routine care case identification	62.40	0.33697	0.000	0.000	75.10	0.33524	0.001	0.001
Whooley questions followed by EPDS (≥ 10-point cut-off point)	68.10	0.33724	0.015	0.059	86.40	0.33561	0.023	0.071
EPDS (≥ 10-point cut-off point) only	79.00	0.33729	0.009	0.042	97.70	0.33568	0.014	0.063
Whooley questions only	82.70	0.33731	0.002	0.015	101.70	0.33571	0.004	0.026

Strategies with the highest probability of being cost-effective at different WTP thresholds are shaded.

TABLE 27 Sensitivity analysis of varying the prevalence parameter in the postnatal model

Strategy	Prevalence = 8.7%				Prevalence = 12%			
	Total cost (£)	Total QALYs	Probability of cost-effectiveness at WTP of		Total cost (£)	Total QALYs	Probability of cost-effectiveness at WTP of	
			£20,000	£30,000			£20,000	£30,000
Whooley questions followed by PHQ-9	56.30	0.9001	0.444	0.358	74.80	0.8974	0.409	0.327
Whooley questions followed by EPDS (≥ 13-point cut-off point)	61.30	0.9002	0.106	0.087	80.70	0.8977	0.090	0.072
EPDS (≥ 13-point cut-off point) only	63.00	0.9002	0.113	0.103	82.30	0.8977	0.095	0.083
Routine care case identification	70.30	0.9000	0.004	0.009	86.90	0.8973	0.015	0.021
Whooley questions followed by EPDS (≥ 10-point cut-off point)	74.90	0.9006	0.151	0.169	96.20	0.8982	0.154	0.161
EPDS (≥ 10-point cut-off point) only	82.30	0.9007	0.141	0.195	103.70	0.8982	0.170	0.219
Whooley questions only	92.00	0.9007	0.041	0.079	113.40	0.8983	0.066	0.117

Strategies with the highest probability of being cost-effective at different WTP thresholds are shaded.

TABLE 28 Sensitivity analysis of assuming a 2% reduction in quality of life in FP cases

Strategy	Prenatal model				Postnatal model			
	Total cost (£)	Total QALYs	Probability of cost-effectiveness at WTP of		Total cost (£)	Total QALYs	Probability of cost-effectiveness at WTP of	
			£20,000	£30,000			£20,000	£30,000
Whooley questions followed by EPDS (≥ 13-point cut-off point)	49.20	0.33584	0.307	0.198	71.80	0.89821	0.150	0.153
EPDS (≥ 13-point cut-off point) only	52.80	0.33583	0.031	0.042	73.40	0.89815	0.165	0.174
Whooley questions followed by PHQ-9	55.10	0.33616	0.648	0.704	66.30	0.89821	0.508	0.450
Routine care case identification	68.40	0.33501	0.000	0.000	79.30	0.89551	0.000	0.000
Whooley questions followed by EPDS (≥ 10-point cut-off point)	76.80	0.33591	0.009	0.032	86.40	0.89782	0.099	0.121
EPDS (≥ 10-point cut-off point) only	87.80	0.33570	0.005	0.020	93.80	0.89735	0.071	0.093
Whooley questions only	91.60	0.33556	0.001	0.004	103.50	0.89628	0.006	0.009

Strategies with the highest probability of being cost-effective at different WTP thresholds are shaded.

cost-effective increased when assuming a QALY reduction. This is not surprising as the 'Whooley questions followed by PHQ-9' strategy has high specificity (and therefore a low FPR), which would result in limited loss of QALYs due to FP diagnosis. Moreover, the benefit of a low FPR partly offsets the lost opportunity of QALY gain from test sensitivity.

Sensitivity analysis 3: resource use by false-positive cases

Based on the NICE guidance, the base-case cost-effectiveness analysis assumed that FP cases incur 20% of the treatment cost incurred by TP cases.[8] However, the impact of this assumption needs to be evaluated. Hence, we conducted a sensitivity analysis assuming 10% and 30% resource use in the FP cases.

Tables 29 and *30* present the results of changing this assumption about resource use. When resource use by FP cases is assumed to be 10% of resource use by TP cases, the cost of FP cases reduces, which in turn reduces the influence of specificity of the test on the overall cost-effectiveness result. Hence, the probability of 'Whooley questions followed by PHQ-9' being the most cost-effective strategy decreases with this assumption. However, except at a threshold of £20,000 per QALY in the prenatal model when 'Whooley questions followed by EPDS (≥ 13-point cut-off point)' is the most cost-effective strategy, for all other analyses, 'Whooley questions followed by PHQ-9' remains the most cost-effective strategy.

In contrast, when assuming that a FP case incurs 30% of the cost of a TP case, the FPR becomes even more important than the base-case analysis because the unnecessary use of resources by FP cases is significant. As a result, the probability of 'Whooley questions followed by PHQ-9' being the most cost-effective strategy increases and it becomes the most cost-effective strategy in both prenatal and postnatal periods.

It should also be noted that two-stage screening/case-finding was found to be more cost-effective than one-stage screening/case-finding in all sensitivity analyses. Moreover, when comparing one-stage screening/case-finding strategies, 'Whooley questions alone' was always one of the least cost-effective strategies and therefore should not be preferred over a two-stage screening/case-finding or 'EPDS alone' strategy.

TABLE 29 Sensitivity analysis of varying the amount of resource use among the FP cases in the prenatal model

Strategy	FP treatment cost = 10% of TP treatment cost				FP treatment cost = 30% of TP treatment cost			
	Total cost (£)	Total QALYs	Probability of cost-effectiveness for a maximum WTP of		Total cost (£)	Total QALYs	Probability of cost-effectiveness for a maximum WTP of	
			£20,000	£30,000			£20,000	£30,000
Whooley questions followed by EPDS (≥ 13-point cut-off point)	47.40	0.33607	0.528	0.377	51.10	0.33607	0.403	0.298
EPDS (≥ 13-point cut-off point) only	50.60	0.33610	0.054	0.050	54.90	0.33610	0.033	0.040
Whooley questions followed by PHQ-9	54.20	0.33627	0.363	0.384	56.00	0.33627	0.544	0.563
Routine care case identification	59.00	0.33615	0.006	0.006	77.80	0.33615	0.000	0.000
Whooley questions followed by EPDS (≥ 10-point cut-off point)	72.20	0.33647	0.024	0.069	81.30	0.33647	0.013	0.050
EPDS (≥ 10-point cut-off point) only	81.10	0.33653	0.017	0.073	94.60	0.33653	0.007	0.039
Whooley questions only	83.50	0.33656	0.008	0.041	99.80	0.33656	0.002	0.011

Strategies with the highest probability of being cost-effective at different WTP thresholds are shaded.

TABLE 30 Sensitivity analysis of varying the amount of resource use among the FP cases in the postnatal model

Strategy	FP treatment cost = 10% of TP treatment cost				FP treatment cost = 30% of TP treatment cost			
	Total cost (£)	Total QALYs	Probability of cost-effectiveness at WTP of		Total cost (£)	Total QALYs	Probability of cost-effectiveness at WTP of	
			£20,000	£30,000			£20,000	£30,000
Whooley questions followed by PHQ-9	65.10	0.8986	0.382	0.298	67.40	0.8986	0.446	0.365
Whooley questions followed by EPDS (≥ 13-point cut-off point)	70.00	0.8989	0.081	0.066	73.50	0.8989	0.117	0.091
Routine care case identification	71.10	0.8985	0.048	0.053	87.50	0.8985	0.004	0.008
EPDS (≥ 13-point cut-off point) only	71.50	0.8989	0.078	0.071	75.30	0.8989	0.114	0.101
Whooley questions followed by EPDS (≥ 10-point cut-off point)	82.40	0.8993	0.153	0.153	90.30	0.8993	0.144	0.170
EPDS (≥ 10-point cut-off point) only	88.40	0.8994	0.177	0.223	99.20	0.8994	0.137	0.191
Whooley questions only	95.00	0.8994	0.081	0.137	112.00	0.8994	0.039	0.075

Strategies with the highest probability of being cost-effective at different WTP thresholds are shaded.

Discussion

In the postnatal period, 'Whooley questions followed by PHQ-9' had the highest probability of being cost-effective at conventionally used WTP thresholds in the UK. At values of £20,000 and £30,000 per QALY, 'Whooley questions followed by PHQ-9' had a probability of 0.43 and 0.35, respectively, of being the most cost-effective screening/case-finding strategy. The next most cost-effective strategy at these thresholds was 'EPDS only (≥ 10-point cut-off point)'. 'Whooley questions alone' was never a cost-effective strategy even at a threshold of £100,000 per QALY. Moreover, when comparing 'EPDS alone' with 'Whooley questions alone' in terms of probability of cost-effectiveness, the 'EPDS alone' strategy always had higher probability of being cost-effective compared with 'Whooley questions alone' strategy.

In the prenatal model, 'Whooley questions followed by PHQ-9' and 'Whooley questions followed by EPDS (≥ 13-point cut-off point)' are the most cost-effective strategies, with small difference between the strategies in terms of net monetary benefit between £20,000 and £30,000 per QALY.

Our decision model found that specificity of case-finding instruments is a key driver of the cost-effectiveness analysis of screening/case-finding strategies. This is because, given the prevalence of perinatal depression, small changes in specificity result in significant increase in the number of FP cases detected. For instance, for a hypothetical cohort of 1000 women, a reduction in specificity of a postnatal diagnostic strategy by 5% would result in unnecessary treatment of 45 women (given the prevalence of postnatal depression). This would result in unnecessary resource use of £7785 in terms of treatment cost (assuming that FP cases receive 20% of the total treatment). Both the NICE guidance[8] and Paulden et al.[81] model have highlighted the significant role of the specificity parameter in the cost-effectiveness results for this particular condition.

With regards to the impact of FP diagnosis on quality of life, our base-case analysis took a conservative approach and assumed that FP diagnosis is not associated with loss of quality of life; however, our sensitivity analysis found that assuming a 2% reduction is quality of life for FPs (as assumed in the NICE guidance model[8]) further improves the value for money of the 'Whooley questions followed by PHQ-9' strategy. Hence, specificity is a key driver of the cost-effectiveness results.

Another important finding of the cost-effectiveness analysis is the comparison of two-stage sequential testing with one-stage testing. The specificity of a two-stage screening/case-finding strategy is higher than that of a one-stage strategy (and, therefore, has lower net FPR). This is important given the prevalence rate of perinatal depression and the known specificity of the Whooley questions alone and EPDS alone. *Table 25* shows that use of a single instrument would result in 146 and 121 FP cases per 1000 women screened for Whooley questions alone and EPDS alone, respectively, in the prenatal period. Given that there were 871,038 conceptions in England and Wales in 2014,[181] one-stage screening/case-finding would result in 127,172 and 105,396 FP cases with Whooley questions alone and EPDS alone, respectively, in the prenatal period. However, this improved specificity comes at the cost of reduced sensitivity of sequential tests. However, the opportunity cost of FN diagnoses in terms of QALYs lost was found to be relatively small. This is because of spontaneous recovery of perinatal depression in the weeks following the initial screen, relatively modest treatment effect of therapies offered for perinatal depression, subsequent identification of cases missed by initial screen, and the known prevalence of postnatal depression. Hence, taking into account both sensitivity and specificity, the most cost-effective strategy in the perinatal period is 'Whooley questions followed by PHQ-9'.

Our findings are in line with the model reported in NICE guidance, which also found 'Whooley questions followed by PHQ-9' to be the most cost-effective strategy.[8] Our model used the same model structure as in the NICE guidance; however, we used primary data on diagnostic performance of case-finding instruments and prevalence of depression in the perinatal period. Moreover, by virtue of having primary data on diagnostic performance, we were able to overcome some of the assumptions made in the NICE model due to data limitations, such as second-stage estimates of sensitivity and specificity conditional on the outcome of the first stage. Finally, it should be noted that the cost-effectiveness results should be interpreted in the

light of net monetary benefit. In the case of both prenatal and postnatal depression, our model found that the difference in net monetary benefits between screening/case-finding strategies is small. Again, this is in line with the NICE model, which found the difference between the most cost-effective strategy (i.e. Whooley questions followed by PHQ-9) and the next most cost-effective strategy (i.e. Whooley questions followed by EPDS) to be < £3 (this is not explicitly reported in the report, but can be derived from the results). Given the small difference in expected net benefit in our study (and the previously published NICE model[8]), decision-makers should also consider other aspects of screening/case-finding, such as acceptability of screening/case-finding strategies.

Chapter 8 Assessment of study findings against National Screening Committee criteria

The UK NSC is responsible for making recommendations to ministers and the UK NHS regarding the implementation of any national population screening policy. A set of internationally recognised criteria provide the framework for reviewing the evidence for the effectiveness of a screening programme. Evidence reviewed includes existing research evidence, such as systematic reviews, randomised controlled trials, pilot programmes and economic evaluations.[76]

The NSC first reviewed the existing evidence base for a screening programme for postnatal depression in 2001.[80] The EPDS was identified by the NSC review[80] as the most currently adopted identification strategy at that time and as such was used as the main screening/case-finding tool against which the NSC criteria were assessed. The NSC concluded that there was insufficient clinical and economic evidence to support the implementation of a postnatal depression screening programme and recommended that the EPDS should not be used as a screening/case-finding tool for perinatal depression until further research had been conducted.[3,76,80,81]

A second NSC review was conducted in 2010,[3] in which evidence reviewed included systematic reviews of case-finding or screening strategies for postnatal depression;[123,182] a cluster randomised controlled trial that used the EPDS to identify women with and without postnatal depression;[183] the NICE clinical guidelines on antenatal and postnatal mental health issued in 2007;[88] and the Health Technology Assessment report conducted by Hewitt et al.,[76,81] which included a cost-effectiveness analysis of routine screening/case-finding for postnatal depression in primary care. Case identification methods reviewed by this research included the EPDS and the Whooley questions. Based on the evidence reviewed, the NSC reaffirmed its initial decision not to recommend the adoption of a screening programme for postnatal depression, concluding that there was no evidence that postnatal screening would improve maternal and infant health outcomes for women.[3] Neither the EPDS nor the Whooley questions were recommended for use as a screening/case-finding tool. Instead, the NSC made a recommendation that further research be carried out to identify the optimal strategy for screening or case-finding of postnatal depression (in terms of key diagnostic properties for postnatal populations) as well as research comparing the diagnostic performance of the Whooley questions with the EPDS and a generic depression measure.[3]

According to the NSC criteria, a screening programme should involve a screening test that is simple, safe, precise and validated (item 4), the distribution of test values in the target population should be known and a suitable cut-off level defined and agreed (item 5), and the test should be acceptable to the target population (item 7). In addition, the screening programme should represent value for money (item 16) (see Appendix 15).[76]

The BaBY PaNDA study evaluated and compared the diagnostic properties of two case identification strategies – the Whooley questions (including the help question, based on NICE 2007 guidelines)[88] and the EPDS – during pregnancy (around 20 weeks) and the early postnatal period (around 3–4 months). The acceptability of these case identification tools and the cost-effectiveness of such a screening/case-finding strategy for perinatal depression (to include prenatal and postnatal depression) was also evaluated. This provides the opportunity to assess our diagnostic accuracy, acceptability and cost-effectiveness findings against the NSC criteria (our study findings are assessed against the previous criteria as these were the criteria available at the time the study was designed and commissioned). Items 5, 6, 7 and 16 were revisited in the light of our results from the diagnostic accuracy, acceptability and cost-effectiveness analyses presented in previous chapters.

National Screening Committee criteria

Item 4: there should be a simple, safe, precise and validated screening test. On the basis of the findings of our research we judged that these criteria were met for both the Whooley questions and EPDS instruments. However, there were limitations in that we were unable to establish what proportion of cases of depression revealed by screening/case-finding were in fact 'new cases' that were not already known to HPs, or what proportion would be picked up as a matter of routine without the use of screening/case-finding procedures.

Item 5: the distribution of test values in the target population should be known and a suitable cut-off level defined and agreed. On the basis of the findings of our research we judged that these criteria were met for both the Whooley questions and EPDS instruments. The optimum cut-off point for the EPDS was 10, and for the Whooley questions the convention of positive responses was appropriate. There were acceptable estimates of sensitivity and specificity. However, in view of the prevalence of depression in the perinatal period (which we found to be present in around 1 in 10 women) there was poor positive predictive ability.

Item 7: the test should be acceptable to the population. This criterion was partly met with a number of caveats for the Whooley questions and for the EPDS. The EPDS was preferred over the Whooley questions in terms of its acceptability to women, but there were significant concerns raised about the application of such instruments in routine care. HPs were cautious about the routine use of such questions in view of the sensitivity of the items.

Item 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). Screening and case-finding using single instruments (i.e. the Whooley questions or the EPDS) did not fall within acceptable thresholds of WTP, and the most cost-effective approaches (high cut-off points with high specificity) did not represent an approach that was likely to be acceptable in clinical practice. Cost-effectiveness results were sensitive to plausible changes in key assumptions. The value for money of case-finding and screening approaches using single instruments only cannot be assumed.

Reflection on current policy and practice within the UK

The evidence generated from our research suggests that both the Whooley questions and the EPDS are simple, safe, precise and validated case-finding tests, with acceptable diagnostic properties to detect depression during the perinatal period. Suitable cut-off levels could be identified; the optimum cut-off point for the EPDS was 10, and for the Whooley questions the convention of positive responses was appropriate. Although the Whooley questions and the EPDS were found to be generally acceptable to the population, concerns were raised by both women (the target population) and HPs working in perinatal care regarding the application of such screening/case-finding instruments in routine clinical practice. A screening/case-finding strategy using single screening/case-finding instrument only was not shown to be cost-effective, and the most cost-effective approaches (high cut-off points with high specificity) did not represent an approach that was likely to be acceptable in clinical practice. We would therefore conclude that NSC criteria for a screening programme for perinatal depression would not currently be met.

Chapter 9 Discussion and conclusions

This programme of research is the first to address the specific gaps in the evidence highlighted in recent NICE guidance on the recognition and management of depression among pregnant women and new mothers.[8] We have used a mixed-methods approach and a number of findings emerge that will be helpful in informing the care of women and in addressing unmet mental health needs for this population. These findings are discussed in turn with reference to their implications for practice, policy and future research.

Summary

Clinical results

In this programme of research, we assembled a longitudinal cohort of women in the perinatal period and examined the properties of commonly advocated mental health instruments, including those in current practice and those recommended in recent NICE guidelines (including the Whooley questions). To our knowledge, this is the first large-scale validation of these instruments and exploration of their acceptability and their cost-effectiveness in use. Across the four study sites, 390 pregnant women completed the diagnostic reference standard (CIS-R) and index tests (Whooley questions and EPDS). The prevalence of depression according to ICD-10 criteria was approximately 10% both during and after pregnancy, which is comparable to the rate in non-pregnant populations. Our first finding is therefore that depression is no more common in pregnancy than at other points in a woman's life.

Diagnostic accuracy tests were conducted for the Whooley questions and the EPDS (using two cut-off scores suggested by the literature[76,125,148]) against the diagnostic reference standard. At both assessment time points, the Whooley questions and EPDS (cut-off point of ≥ 10 points) were found to be reasonable case-finding instruments with comparable sensitivity (prenatal stage: Whooley 85.0%, EPDS 82.5%; postnatal stage: Whooley questions 85.7%, EPDS 82.9%) and specificity (prenatal stage: Whooley questions 83.7%, EPDS 86.6%; postnatal stage: Whooley questions 80.6%, EPDS: 87.6%). The diagnostic accuracy of the EPDS cut-off of ≥ 13 points was poor, even when considering only moderate to severe depression cases. The key message based on the psychometric evaluation is that there is little to choose between the Whooley questions and EPDS in their ability to detect and rule out depression. Based on psychometric considerations alone, then the briefer Whooley questionnaire might be preferred as it is much shorter and can be more readily administered within a shorter period of time. Use of the cognitive evaluation method[149] in the acceptability survey identified that women viewed the EPDS questions as easier to understand, easier to remember and easier to answer more confidently than the Whooley questions. There is therefore a tension between brevity in the Whooley questions and the more acceptable style of the longer EPDS questionnaire. Either instrument is fit for purpose based on psychometric considerations, and the research presented in this report will be helpful in making informed policy decisions regarding the choice of instrument in practice.

We also found that women who were found to be depressed in the prenatal period were commonly not found to be depressed at postnatal follow-up. Approximately half of depressed women during pregnancy became non-cases postnatally, and around 7% of non-depressed women during pregnancy became depressed cases postnatally.

There was also clear evidence that depression in the prenatal and postnatal periods is associated with other common mental disorders. Depression caseness based on any of the case-finding instruments was associated with higher levels of anxiety and poorer quality of life. In addition, higher somatic symptom scores were observed in women with depression. This highlights the importance of depression at any point in terms of its impact on quality of life, but also the importance of coexisting anxiety. This finding is in line with 2014 recommendations from NICE that highlight the importance of the presence of anxiety and the shortcomings of assessments of mental health among women in the perinatal period that focus solely on depression.[101]

Evaluation of acceptability

Prior to the conduct of this programme of work, there had been very little research into the acceptability of screening/case-finding approaches to depression among women in the perinatal period, particularly with respect to the use of the Whooley questions. The qualitative EoA revealed that women and HPs are generally supportive of the use of routine instruments to enquire about mental health, and that these could form the basis of routine practice. However there were several caveats.

We noted that there were some concerns regarding the wording of Whooley question 1 among expectant and new mothers and HPs who were interviewed in this study. Some pregnant women and new mothers were 'not comfortable' with Whooley question 1, and indicated that they might not reveal the true extent of their depression. MWs and HVs also shared these concerns, stating that women's body language and visible symptoms do not appear to match their answers to the Whooley question. Some HPs revealed that they adapted Whooley question 1 when they used this instrument in practice. Instead they asked a more general question about a woman's feelings or mood. The psychometric properties of the Whooley questionnaire should therefore be interpreted with this in mind.

Qualitative interviews also showed that women and HPs were concerned about the potential consequences of being identified as depressed. Women spoke of being 'monitored' in their own home and ultimately having their baby taken into care.[103,152,157,158] These findings are of importance for those who might use such instruments in practice. Women may not answer questions honestly depending on whether or not they recognise that they have symptoms of depression and whether or not they will admit these symptoms to a HP who is obliged to prioritise the welfare of the child over the mother.

Our qualitative research also found that HPs preferred to introduce case-finding questions in the context of a 'patient-centred environment' before asking questions about mental health.[107,164] We also found that prior experience of depression was important in determining whether or not someone gave honest answers to screening/case-finding questions. In qualitative interviews we found evidence that women with prior experience of depression were more willing to disclose feelings of current depression in the context of screening/case-finding questionnaires. This was a novel finding that could be important for clinical practice and the mode by which screening/case-finding instruments might be used.

Cost-effectiveness

Our economic decision model addressed the cost-effectiveness of screening and case-finding for depression in the prenatal and in the postnatal periods. We extended earlier work by Paulden et al.[81] and used the results of the empirical estimates of diagnostic test accuracy described above to populate our model. The decision model also builds on a decision analysis reported in the NICE guidance on antenatal and postnatal mental health, issued in 2014.[8,81] The most recent NICE guidance developed a new model to compare cost-effectiveness of screening/case-finding instruments as primary screening/case-finding strategies.[8] This NICE model included a 'two-stage procedure' as well as a single-stage screening/case-finding scenario in line with the consensus-developed care pathway presented in NICE guidelines.

When we examined the cost-effectiveness of screening/case-finding in the postnatal period using the NICE-described care pathways, 'Whooley questions followed by PHQ-9' had the highest probability of being cost-effective at conventionally used WTP thresholds in the UK. At values of £20,000 and £30,000 per QALY, 'Whooley questions followed by PHQ-9' had a probability of 0.43 and 0.35, respectively, of being the most cost-effective screening/case-finding strategy. The next most cost-effective strategy at these thresholds was 'EPDS only (≥ 10-point cut-off point)'. However, 'Whooley questions only' was never cost-effective even at WTP thresholds of up to £100,000 per QALY. The main finding is therefore that screening/case-finding using single instruments in the postnatal period was not cost-effective within conventional thresholds for cost-effectiveness, whereas using two-stage screening/case-finding, albeit relatively cost-effective, was subject to important uncertainties.

In the prenatal model, 'Whooley questions followed by PHQ-9' and 'Whooley questions followed by EPDS (≥ 13-point cut-off point)' were the most cost-effective strategies with very small difference between the strategies in terms of net monetary benefit between £20,000 and £30,000 per QALY. This represented screening/case-finding at a cut-off point that was not sufficiently optimal in terms of psychometric properties.

An important finding of our decision model was that specificity of the screening/case-finding instruments is a key driver of the cost-effectiveness analysis of screening/case-finding strategies. This is because, given the prevalence of perinatal depression found in our cohort, small changes in specificity result in significant increase in the number of FP cases detected. This finding replicates earlier work reported in the NICE guidance[8] and the Paulden et al.[81] model, both of which have highlighted the significant role of the specificity parameter in the cost-effectiveness results for this particular condition.

In summary, the findings of our cost-effectiveness study are in line with the model reported in NICE guidance, which also found 'Whooley questions followed by PHQ-9' to be the most cost-effective strategy.[8] Our model used the same model structure as in the NICE guidance; however, we used UK-derived primary data on diagnostic performance of screening/case-finding instruments and prevalence of depression in the perinatal period. Moreover, by virtue of having primary data on diagnostic performance, we were able to overcome some of the assumptions that were necessary in the NICE model owing to data limitations, such as second-stage estimates of sensitivity and specificity conditional on the outcome of the first stage.

Strengths and limitations

This study is based on a large UK cohort of women characterised in the prenatal period and then followed up in the postnatal period. Previous research has been largely of smaller sample size and based on cross-sectional data. We have used this cohort to generate empirically derived data, which has in turn populated a decision-analytic model. The results are therefore more robust and generalisable. This study therefore overcomes several of the limitations that were acknowledged in the NICE model,[8] including the fact that we are able to draw on the results of the first large-scale validation of the Whooley questions.

Recruitment to the cohort went to plan and we achieved our predetermined sample size with a good level of follow-up (85%), and our validation study is considerably larger than other published validation studies of the NICE-endorsed screening/case-finding instruments.[100]

We have also employed a mixed-methods approach and utilised qualitative methodology to understand the barriers to and facilitators of the NICE-endorsed case-finding instruments. This provides a rich data source by which to understand the strengths and limitations of screening and case-finding. We also note that the decision-analytic model is the first to explore screening/case-finding in the prenatal period.

The main limitation is that we were not able to establish whether or not perinatal depression (and other common mental disorders) had already been recognised by a woman's GP or perinatal services (although consent was obtained for this we were unable to secure the necessary funds to allow us to extract these data from GP records). This is a common limitation of psychometric diagnostic test accuracy studies (of which our research is an example). Thombs et al.[184] have empirically demonstrated that validation studies of depression screening/case-finding instruments rarely cross-reference GP treatment records to check what proportion of screen-positive patients are already known of or recognised. Prior research[185] has shown that a significant proportion of people who screen positive for depression are already known to be depressed or that this becomes apparent at follow-up. The 'yield' of screening/case-finding validation studies is therefore overestimated and this is a limitation of the present study. Additional research would be needed to cross-reference the findings in this cohort with women's GP records.

A further limitation is that the women participating in this study were disproportionately white, and all were English speakers, and caution is needed in extrapolating the results to other social and ethnic groups (although we note that our previous pilot research in this area included a small but more ethnically diverse sample of women[100]). The final limitation is that our study does not form a randomised evaluation of the clinical effectiveness and cost-effectiveness of screening and case-finding in the perinatal period. A core criterion when considering screening programmes is that the benefits and harms of screening should be evaluated within the context of a randomised controlled trial.[186] Trial-based evaluations of screening programmes of depression have generally shown little or no benefit to patients,[187] and to date there are no published robust trial-based evaluations of screening for depression in the perinatal period.[188]

Research recommendations

- The present research did not cross-reference perinatal depression diagnosis with GP medical records and so was unable to determine what proportion of new cases were identified by screening/case-finding instruments or what proportion of cases would go on to be recognised by health-care practitioners. Future research should link the results of case-finding instruments with routine clinical records to establish what proportion of women who screen positive for perinatal depression using screening/case-finding instruments are already known to be depressed and/or are receiving treatment for depression. This will provide an indication of the true 'yield' of case-finding instruments and the degree to which validation studies overestimate the ability of case-finding instruments to identify perinatal depression.
- Future validation of screening/case-finding approaches for perinatal depression may benefit from consideration of perinatal populations that include diversity in social, cultural and ethic backgrounds. Such research would expand our understanding of the feasibility and impact of screening/case-finding for perinatal depression given differences in women's cultural understanding and experiences of perinatal depression. Future research may also consider how to effectively deliver screening/case-finding approaches to women with moderate to severe learning difficulties.
- Using criteria developed by the NSC, this research found important uncertainties in recommending the optimal strategy for routine use of screening and case-finding. Further research is needed within the context of a randomised controlled trial that assesses the clinical effectiveness and cost-effectiveness of treatment for perinatal depression following screening/case-finding procedures (such as those reported here) and the extent to which such screening/case-finding strategies improve maternal and infant outcomes for women. Such research would also need to consider the effectiveness of various screening/diagnostic pathways, which may involve perinatal health-care practitioners, primary care HPs and specialist mental health services.
- The present research highlights the importance of depression in the perinatal period and indicates that there is comorbidity with other common mental disorders, such as anxiety. Further research is needed to understand the inter-relationship between different common mental disorders and how these might best be managed for women in the perinatal period. Such research may consider establishing whether or not comorbid common mental disorders predict future patterns of mental health problems in the perinatal period.
- Qualitative research has demonstrated variation among health-care practitioners in the mode of administration, level of confidence and delivery of screening/case-finding instruments. Research is needed to develop training programmes and optimise the delivery of screening/case-finding by HPs.

Implications for policy and practice

Our research findings suggest the following implications for policy and practice.

- Depression was found to be present in around 1 in 10 women in both the prenatal and postnatal periods. This research suggests that HPs need to be vigilant to the presence of depression in the prenatal period as well as the postnatal period.

- Around half of cases of depression in the prenatal period will resolve after the birth of a baby. A significant portion of depression in the postnatal period is 'new' and HPs need to be alert to this and the need for intervention.
- If case-finding instruments are used in clinical practice, then the Whooley questions and the EPDS have been shown to have acceptable psychometric properties.
- The Whooley questions have the advantage of being very brief with good psychometric properties, but there may be limits to the use of these questions in terms of their acceptability to pregnant women and new mothers.
- The EPDS has the advantage of having been developed for use in the perinatal period. Although it is a longer instrument, it was found to be acceptable to women and to HPs.
- This research suggests that HPs need to be vigilant to the presence of other common mental disorders and may consider using additional instruments or screening/case-finding questions to detect these.
- Screening and case-finding may need to be undertaken within the context of a careful patient-centred consultation. HPs may benefit from training in the administration and interpretation of screening/case-finding instruments such as the Whooley questions and EPDS. They need to be aware that women see depression as being stigmatised and may not answer questions truthfully.
- Screening and case-finding is a large public health undertaking, and the optimal strategy based on costs and benefits is not robustly supported by cost-effectiveness findings, which were subject to important uncertainties.
- Screening and case-finding for depression in the perinatal period did not meet the strict criteria laid down by the NSC.

Conclusions

Depression in the perinatal period is an important psychological disorder and there are instruments that can help in its diagnosis and management. The psychometric properties have been determined in the context of a large UK sample, and shorter screening/case-finding instruments such as the Whooley questions have reasonable psychometric properties, comparable with longer instruments such as the EPDS. There are advantages and limitations to the use of these screening/case-finding instruments in clinical practice, which are revealed in qualitative research. Economic analysis found that using single screening/case-finding instruments is not cost-effective, whereas two-stage screening/case-finding, albeit relatively cost-effective, was subject to significant uncertainty. Further research is needed in the form of a randomised controlled trial to establish the clinical effectiveness and cost-effectiveness of screening and case-finding approaches for psychological problems (including depression).

Acknowledgements

We would especially like to thank all the women who agreed to take part in this study; their time and contribution has been invaluable.

We would also like to thank all the researchers who assisted with recruitment, data collection and data coding: Anne Thomson and Phil Andersen for providing administrative support; Will Curson, who designed and maintained the study database; site principal investigators (Debrah Bates, Nicola Dean, Jan Cairns and Katherine Johnson); colleagues working with the BaBY team, particularly Helen Bexhell, Kathryn McNamara and Kim Mitchelson; Cath Jackson for providing advice and support with the qualitative interview methodology; Miqdad Asaria, Jan Boehnke, Sarah Byford and Mike Paulden for their advice on the diagnostic analysis and decision modelling; Claire Clarke for providing advice on perinatal mental health-care pathways to inform the decision model; Deborah Morgan from Perinatal Illness UK; Eve Roman for providing epidemiological expertise; Donna White for being a supportive HSDR programme manager; the NIHR Clinical Research Network: Yorkshire and the Humber, particularly Rebecca Hargate; staff at NHS R&D departments; and independent members of the Study Steering Committee. Special thanks go to all the individuals who formed the PPI group, including Amanda Hizzett, Alice North, Natalie Wheatley and Holly Williams. We are grateful to the reviewers of this report for their helpful and insightful comments.

This study is dedicated to the memory of Claire Brunskill (1975–2016), Research Officer for the BaBY PaNDA study and BaBY studies at Harrogate and District Hospital. Claire was a valued, committed and enthusiastic member of the BaBY PaNDA and BaBY research teams and she will be missed by all those who worked with her.

Collaborators

Deborah Morgan from Perinatal Illness UK was a coapplicant on the HSDR application and contributed to the study design and study protocol. Debrah Bates, Nicola Dean, Jan Cairns and Katherine Johnson were site principal investigators. Eve Roman was a member of the Study Management Group and Study Steering Committee.

Contributions of authors

Elizabeth Littlewood (Research Fellow, Applied Health Services Research) was a coapplicant on the HSDR application, contributed to the study design, drafted the study protocol, monitored the study as project manager and co-ordinated and drafted the final report.

Shehzad Ali (Research Fellow, Health Economics), **Lisa Dyson** (Research Fellow, Qualitative Research and Health Services Research), **Pat Ansell** (Senior Research Fellow, Reproductive Epidemiology and Child Health), **Catherine Hewitt** (Professor, Statistics), **Rachel Mann** (Research Fellow, Applied Health Services Research), **Dean McMillan** (Senior Lecturer in Mental Health, Applied Health Services Research), **Simon Gilbody** (Professor, Applied Health Services Research) and **Elizabeth Littlewood**, were coapplicants on the HSDR application, contributed to the study design and protocol, and were members of the Study Management Group.

Shehzad Ali designed and conducted the economic evaluation.

Ada Keding (Research Fellow, Statistics) drafted the Statistical Analysis Plan and conducted the clinical analysis.

Publication

Littlewood E, Ali S, Ansell P, Dyson L, Gascoyne S, Hewitt C, *et al.* Identification of depression in women during pregnancy and the early postnatal period using the Whooley questions and the Edinburgh Postnatal Depression Scale: protocol for the Born and Bred in Yorkshire: PeriNatal Depression Diagnostic Accuracy (BaBY PaNDA) study. *BMJ Open* 2016;**6**:e011223.

Data sharing statement

Requests for access to data should be addressed to the corresponding author.

References

1. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLOS Med* 2006;**3**:e442. https://doi.org/10.1371/journal.pmed.0030442

2. Wisner KL, Sit DK, McShea MC, Rizzo DM, Zoretich RA, Hughes CL, *et al.* Onset timing, thoughts of self-harm, and diagnoses in postpartum women with screen-positive depression findings. *JAMA Psychiatry* 2013;**70**:490–8. http://dx.doi.org/10.1001/jamapsychiatry.2013.87

3. Hill C. *An Evaluation of Screening for Postnatal Depression Against NSC Criteria*. London: UK National Screening Committee; 2010.

4. Austin M-P, Highet N, Guidelines Expert Advisory Committee. *Clinical Practice Guidelines for Depression and Related Disorders – Anxiety, Bipoloar Disorder and Puerperal Psychosis – In the Perinatal Period. A Guideline for Primary Care Health Professionals*. Melbourne, VIC: Beyondblue; 2011.

5. Siu AL, Bibbins-Domingo K, Grossman DC, Baumann LC, Davidson KW, Ebell M, *et al.* Screening for depression in adults: US Preventive Services Task Force recommendation statement. *JAMA* 2016;**315**:380–7. https://doi.org/10.1001/jama.2015.18392

6. US Preventive Services Task Force. Screening for depression in adults: US Preventive Services Task Force recommendation statement. *Ann Intern Med* 2009;**151**:784. https://doi.org/10.7326/0003-4819-151-11-200912010-00006

7. National Collaborating Centre for Mental Health. *Antenatal and Postnatal Mental Health: Clinical Management and Service Guidance*. Leicester and London: The British Psychological Society and the Royal College of Psychiatrists; 2007.

8. National Collaborating Centre for Mental Health. *Antenatal and Postnatal Mental Health: Clinical Management and Service Guidance (Updated Edition). National Clinical Guideline No. 192*. Leicester and London: The British Psychological Society and the Royal College of Psychiatrists; 2014.

9. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th edn. Arlington, VA: American Psychiatric Association; 2013.

10. World Health Organization (WHO). *International Statistical Classification of Diseases and Related Health Problems*. Tenth edn. Geneva: WHO; 2010.

11. Gavin NI, Gaynes BN, Lohr KN, Meltzer-Brody S, Gartlehner G, Swinson T. Perinatal depression: a systematic review of prevalence and incidence. *Obstet Gynecol* 2005;**106**:1071–83. https://doi.org/10.1097/01.AOG.0000183597.31630.db

12. Matthey S. Detection and treatment of postnatal depression (perinatal depression or anxiety). *Curr Opin Psychiatr* 2004;**17**:21–9.

13. Bennett HA, Einarson A, Taddio A, Koren G, Einarson TR. Prevalence of depression during pregnancy: systematic review. *Obstet Gynecol* 2004;**103**:698–709. http://dx.doi.org/10.1097/01.AOG.0000116689.75396.5f

14. Marcus SM. Depression during pregnancy: rates, risks and consequences – Motherisk update 2008. *Can J Clin Pharmacol* 2009;**16**:e15.

15. Schmied V, Johnson M, Naidoo N, Austin MP, Matthey S, Kemp L, *et al.* Maternal mental health in Australia and New Zealand: a review of longitudinal studies. *Women Birth* 2013;**26**:167–78. http://dx.doi.org/10.1016/j.wombi.2013.02.006

16. Vigod SN, Wilson CA, Howard LM. Depression in pregnancy. *BMJ* 2016;**352**:i1547. http://dx.doi.org/10.1136/bmj.i1547

17. Vesga-López O, Blanco C, Keyes K, Olfson M, Grant BF, Hasin DS. Psychiatric disorders in pregnant and postpartum women in the United States. *Arch Gen Psychiatry* 2008;**65**:805–15. http://dx.doi.org/10.1001/archpsyc.65.7.805

18. Ko JY, Farr SL, Dietz PM, Robbins CL. Depression and treatment among U.S. pregnant and nonpregnant women of reproductive age, 2005–2009. *J Womens Health* 2012;**21**:830–6. http://dx.doi.org/10.1089/jwh.2011.3466

19. Ban L, Gibson JE, West J, Fiaschi L, Oates MR, Tata LJ. Impact of socioeconomic deprivation on maternal perinatal mental illnesses presenting to UK general practice. *Br J Gen Pract* 2012;**62**:e671–8. http://dx.doi.org/10.3399/bjgp12X656801

20. Munk-Olsen T, Laursen TM, Pedersen CB, Mors O, Mortensen PB. New parents and mental disorders: a population-based register study. *JAMA* 2006;**296**:2582–9. https://doi.org/10.1001/jama.296.21.2582

21. 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 Technol Assess* 2016;**20**(37). http://dx.doi.org/10.3310/hta20370

22. Heron J, O'Connor TG, Evans J, Golding J, Glover V, ALSPAC Study Team. The course of anxiety and depression through pregnancy and the postpartum in a community sample. *J Affect Disord* 2004;**80**:65–73. http://dx.doi.org/10.1016/j.jad.2003.08.004

23. Evans J, Heron J, Francomb H, Oke S, Golding J. Cohort study of depressed mood during pregnancy and after childbirth. *BMJ* 2001;**323**:257–60. https://doi.org/10.1136/bmj.323.7307.257

24. Henshaw C. Mood disturbance in the early puerperium: a review. *Arch Womens Ment Health* 2003;**6**(Suppl. 2):33–42. http://dx.doi.org/10.1007/s00737-003-0004-x

25. Howard LM, Molyneaux E, Dennis CL, Rochat T, Stein A, Milgrom J. Non-psychotic mental disorders in the perinatal period. *Lancet* 2014;**384**:1775–88. http://dx.doi.org/10.1016/S0140-6736(14)61276-9

26. Cooper PJ, Murray L. Postnatal depression. *BMJ* 1998;**316**:1884. http://dx.doi.org/10.1136/bmj.316.7148.1884

27. Kumar R, Robson KM. A prospective study of emotional disorders in childbearing women. *Br J Psychiatry* 1984;**144**:35–47. https://doi.org/10.1192/bjp.144.1.35

28. Goodman JH. Postpartum depression beyond the early postpartum period. *J Obstet Gynecol Neonatal Nurs* 2004;**33**:410–20. https://doi.org/10.1177/0884217504266915

29. Woolhouse H, Gartland D, Mensah F, Brown S. Maternal depression from early pregnancy to 4 years postpartum in a prospective pregnancy cohort study: implications for primary health care. *Br J Obstet Gynaecol* 2015;**122**:312–21.

30. Milgrom J, Gemmill AW. *Identifying Perinatal Depression and Anxiety. Evidence-Based Practice in Screening, Psychosocial Assessment, and Management.* Chichester: John Wiley & Sons Ltd; 2015.

31. Dennis CL, Heaman M, Vigod S. Epidemiology of postpartum depressive symptoms among Canadian women: regional and national results from a cross-sectional survey. *Can J Psychiatry* 2012;**57**:537–46. https://doi.org/10.1177/070674371205700904

32. Robertson E, Grace S, Wallington T, Stewart DE. Antenatal risk factors for postpartum depression: a synthesis of recent literature. *Gen Hosp Psychiatry* 2004;**26**:289–95. http://dx.doi.org/10.1016/j.genhosppsych.2004.02.006

33. Rich-Edwards JW, Kleinman K, Abrams A, Harlow BL, McLaughlin TJ, Joffe H, *et al.* Sociodemographic predictors of antenatal and postpartum depressive symptoms among women in a medical group practice. *J Epidemiol Community Health* 2006;**60**:221–7.

34. Milgrom J, Gemmill AW, Bilszta JL, Hayes B, Barnett B, Brooks J, *et al.* Antenatal risk factors for postnatal depression: a large prospective study. *J Affect Disord* 2008;**108**:147–57. https://doi.org/10.1016/j.jad.2007.10.014

35. Leigh B, Milgrom J. Risk factors for antenatal depression, postnatal depression and parenting stress. *BMC Psychiatry* 2008;**8**:24. http://dx.doi.org/10.1186/1471-244X-8-24

36. Woolhouse H, Gartland D, Hegarty K, Donath S, Brown S. Depressive symptoms and intimate partner violence in the 12 months after childbirth: a prospective pregnancy cohort study. *Br J Obstet Gynaecol* 2012;**119**:315–23.

37. Howard LM, Oram S, Galley H, Trevillion K, Feder G. Domestic violence and perinatal mental disorders: a systematic review and meta-analysis. *PLOS Med* 2013;**10**:e1001452. http://dx.doi.org/10.1371/journal.pmed.1001452

38. Lancaster CA, Gold KJ, Flynn HA, Yoo H, Marcus SM, Davis MM. Risk factors for depressive symptoms during pregnancy: a systematic review. *Am J Obstet Gynecol* 2010;**202**:5–14. http://dx.doi.org/10.1016/j.ajog.2009.09.007

39. Stein A, Pearson RM, Goodman SH, Rapa E, Rahman A, McCallum M, *et al.* Effects of perinatal mental disorders on the fetus and child. *Lancet* 2014;**384**:1800–19. http://dx.doi.org/10.1016/S0140-6736(14)61277-0

40. O'Hara MW, Swain AM. Rates and risk of postpartum depression – a meta-analysis. *Int Rev Psychiatry* 1996;**8**:37–54.

41. Kelly RH, Russo J, Katon W. Somatic complaints among pregnant women cared for in obstetrics: normal pregnancy or depressive and anxiety symptom amplification revisited? *Gen Hosp Psychiatry* 2001;**23**:107–13. https://doi.org/10.1016/S0163-8343(01)00129-3

42. Grigoriadis S, VonderPorten EH, Mamisashvili L, Tomlinson G, Dennis CL, Koren G, *et al.* The impact of maternal depression during pregnancy on perinatal outcomes: a systematic review and meta-analysis. *J Clin Psychiatry* 2013;**74**:e321–41. http://dx.doi.org/10.4088/JCP.12r07968

43. Grote NK, Bridge JA, Gavin AR, Melville JL, Iyengar S, Katon WJ. A meta-analysis of depression during pregnancy and the risk of preterm birth, low birth weight, and intrauterine growth restriction. *Arch Gen Psychiatry* 2010;**67**:1012–24. http://dx.doi.org/10.1001/archgenpsychiatry.2010.111

44. Ibanez G, Charles MA, Forhan A, Magnin G, Thiebaugeorges O, Kaminski M, *et al.* Depression and anxiety in women during pregnancy and neonatal outcome: data from the EDEN mother-child cohort. *Early Hum Dev* 2012;**88**:643–9. http://dx.doi.org/10.1016/j.earlhumdev.2012.01.014

45. Orr ST, Blazer DG, James SA, Reiter JP. Depressive symptoms and indicators of maternal health status during pregnancy. *J Womens Health* 2007;**16**:535–42. http://dx.doi.org/10.1089/jwh.2006.0116

46. Flynn HA, Davis M, Marcus SM, Cunningham R, Blow FC. Rates of maternal depression in pediatric emergency department and relationship to child service utilization. *Gen Hosp Psychiatry* 2004;**26**:316–22. http://dx.doi.org/10.1016/j.genhosppsych.2004.03.009

47. 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**:2512–26. https://doi.org/10.2307/1131637

48. Milgrom J, Westley DT, Gemmill AW. The mediating role of maternal responsiveness in some longer term effects of postnatal depression on infant development. *Infant Behav Dev* 2004;**27**:443–54. https://doi.org/10.1016/j.infbeh.2004.03.003

49. Campbell SB, Brownell CA, Hungerford A, Spieker SI, Mohan R, Blessing JS. The course of maternal depressive symptoms and maternal sensitivity as predictors of attachment security at 36 months. *Dev Psychopathol* 2004;**16**:231–52. https://doi.org/10.1017/S0954579404044499

50. 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**:1259–71. https://doi.org/10.1111/1469-7610.00542

51. Murray L, Arteche A, Fearon P, Halligan S, Goodyer I, Cooper P. Maternal postnatal depression and the development of depression in offspring up to 16 years of age. *J Am Acad Child Adolesc Psychiatry* 2011;**50**:460–70. http://dx.doi.org/10.1016/j.jaac.2011.02.001

52. Velders FP, Dieleman G, Henrichs J, Jaddoe VW, Hofman A, Verhulst FC, *et al.* Prenatal and postnatal psychological symptoms of parents and family functioning: the impact on child emotional and behavioural problems. *Eur Child Adolesc Psychiatry* 2011;**20**:341–50. http://dx.doi.org/10.1007/s00787-011-0178-0

53. Goodman SH, Rouse MH, Connell AM, Broth MR, Hall CM, Heyward D. Maternal depression and child psychopathology: a meta-analytic review. *Clin Child Fam Psychol Rev* 2011;**14**:1–27. http://dx.doi.org/10.1007/s10567-010-0080-1

54. Field T. Postpartum depression effects on early interactions, parenting, and safety practices: a review. *Infant Behav Dev* 2010;**33**:1–6. http://dx.doi.org/10.1016/j.infbeh.2009.10.005

55. Kersten-Alvarez LE, Hosman CM, Riksen-Walraven JM, van Doesum KT, Smeekens S, Hoefnagels C. Early school outcomes for children of postpartum depressed mothers: comparison with a community sample. *Child Psychiatry Hum Dev* 2012;**43**:201–18. http://dx.doi.org/10.1007/s10578-011-0257-y

56. Pearson RM, Evans J, Kounali D, Lewis G, Heron J, Ramchandani PG, *et al.* Maternal depression during pregnancy and the postnatal period: risks and possible mechanisms for offspring depression at age 18 years. *JAMA Psychiatry* 2013;**70**:1312–19. http://dx.doi.org/10.1001/jamapsychiatry.2013.2163

57. Galéra C, Côté SM, Bouvard MP, Pingault JB, Melchior M, Michel G, *et al.* Early risk factors for hyperactivity-impulsivity and inattention trajectories from age 17 months to 8 years. *Arch Gen Psychiatry* 2011;**68**:1267–75. http://dx.doi.org/10.1001/archgenpsychiatry.2011.138

58. Kaplan PS, Danko CM, Diaz A, Kalinka CJ. An associative learning deficit in 1-year-old infants of depressed mothers: role of depression duration. *Infant Behav Dev* 2011;**34**:35–44. http://dx.doi.org/10.1016/j.infbeh.2010.07.014

59. Quevedo LA, Silva RA, Godoy R, Jansen K, Matos MB, Tavares Pinheiro KA, *et al.* The impact of maternal post-partum depression on the language development of children at 12 months. *Child Care Health Dev* 2012;**38**:420–4. http://dx.doi.org/10.1111/j.1365-2214.2011.01251.x

60. Cogill SR, Caplan HL, Alexandra H, Robson KM, Kumar R. Impact of maternal postnatal depression on cognitive development of young children. *Br Med J* 1986;**292**:1165–7. https://doi.org/10.1136/bmj.292.6529.1165

61. Sutter-Dallay A-L, Murray L, Dequae-Merchadou L, Glatigny-Dallay E, Bourgeois M-L, Verdoux H. A prospective longitudinal study of the impact of early postnatal vs. chronic maternal depressive symptoms on child development. *Eur Psychiatry* 2011;**26**:484–9. https://doi.org/10.1016/j.eurpsy.2010.05.004

62. Oates MR. Perinatal psychiatric syndromes: clinical features. *Psychiatry* 2009;**8**:1–6. https://doi.org/10.1016/j.mppsy.2008.10.014

63. Flynn SM, Shaw JJ, Abel KM. Homicide of infants: a cross-sectional study. *J Clin Psychiatry* 2007;**68**:1501–9. https://doi.org/10.4088/JCP.v68n1005

64. Paulson JF, Bazemore SD. Prenatal and postpartum depression in fathers and its association with maternal depression: a meta-analysis. *JAMA* 2010;**303**:1961–9. http://dx.doi.org/10.1001/jama.2010.605

65. Paulson JF, Keefe HA, Leiferman JA. Early parental depression and child language development. *J Child Psychol Psychiatry* 2009;**50**:254–62. http://dx.doi.org/10.1111/j.1469-7610.2008.01973.x

66. Ramchandani PG, Stein A, O'Connor TG, Heron J, Murray L, Evans J. Depression in men in the postnatal period and later child psychopathology: a population cohort study. *J Am Acad Child Adolesc Psychiatry* 2008;**47**:390–8. http://dx.doi.org/10.1097/CHI.0b013e31816429c2

67. Beck CT. A meta-analysis of predictors of postpartum depression. *Nurs Res* 1996;**45**:297–303. https://doi.org/10.1097/00006199-199609000-00008

68. NICE. *Common Mental Health Disorders: Identification and Pathways to Care CG123*. London: The British Psychological Society and The Royal College of Psychiatrists; 2011.

69. Matthey S, Barnett B, Howie P, Kavanagh DJ. Diagnosing postpartum depression in mothers and fathers: whatever happened to anxiety? *J Affect Disord* 2003;**74**:139–47. https://doi.org/10.1016/S0165-0327(02)00012-5

70. Lydsdottir LB, Howard LM, Olafsdottir H, Thome M, Tyrfingsson P, Sigurdsson JF. The mental health characteristics of pregnant women with depressive symptoms identified by the Edinburgh Postnatal Depression Scale. *J Clin Psychiatry* 2014;**75**:393–8. http://dx.doi.org/10.4088/JCP.13m08646

71. Goodman JH, Tyer-Viola L. Detection, treatment, and referral of perinatal depression and anxiety by obstetrical providers. *J Womens Health* 2010;**19**:477–90. http://dx.doi.org/10.1089/jwh.2008.1352

72. 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**:1064–6.

73. LaRocco-Cockburn A, Melville J, Bell M, Katon W. Depression screening attitudes and practices among obstetrician-gynecologists. *Obstet Gynecol* 2003;**101**:892–8. https://doi.org/10.1016/S0029-7844(03)00171-6

74. Henshaw C, Elliott S. *Screening for Perinatal Depression*. London: Jessica Kingsley Publishers; 2005.

75. Thombs BD, Coyne JC, Cuijpers P, de Jonge P, Gilbody S, Ioannidis JP, *et al.* Rethinking recommendations for screening for depression in primary care. *CMAJ* 2012;**184**:413–18. http://dx.doi.org/10.1503/cmaj.111035

76. 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 Technol Assess* 2009;**13**(36). http://dx.doi.org/10.3310/hta13360

77. Mant D, Fowler G. Mass screening: theory and ethics. *BMJ* 1990;**300**:916–18. https://doi.org/10.1136/bmj.300.6729.916

78. Secretary of State for Health. *National Service Framework – Mental Health*. London: HMSO; 1999.

79. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry* 1987;**150**:782–6. https://doi.org/10.1192/bjp.150.6.782

80. Shakespeare J. *Evaluation of Screening for Postnatal Depression Against the NSC Handbook Criteria. Prepared for a Working Party, June 2001*. London: UK National Screening Committee; 2002.

81. Paulden M, Palmer S, Hewitt C, Gilbody S. Screening for postnatal depression in primary care: cost effectiveness analysis. *BMJ* 2009;**339**:b5203. http://dx.doi.org/10.1136/bmj.b5203

82. Shakespeare J. *Screening: The Role and Recommendations of the UK National Screening Committee*. London: Jessica Kingsley Publishers; 2005.

83. American College of Obstetricians and Gynecologists. Committee on Obstetric Practice. Committee Opinion No. 453: Screening for depression during and after pregnancy. *Obstet Gynecol* 2010;**115**:394–5. http://dx.doi.org/10.1097/AOG.0b013e3181d035aa

84. American College of Obstetricians and Gynecologists. Psychosocial risk factors: perinatal screening and intervention. Committee Opinion No. 343. *Obstet Gynecol* 2006;**108**:469–77.

85. Scottish Intercollegiate Guidelines Network (SIGN). *Management of Perinatal Mood Disorders. SIGN Publication No. 127*. Edinburgh: SIGN; 2012.

86. Gilbody S, Sheldon T, House A. Screening and case-finding instruments for depression: a meta-analysis. *CMAJ* 2008;**178**:997–1003. http://dx.doi.org/10.1503/cmaj.070281

87. National Collaborating Centre for Mental Health. *Depression: Treatment and Management of Depression in Adults (Update)*. Leicester and London: The British Psychological Society and the Royal College of Psychiatrists; 2010.

88. NICE. *Antenatal and Postnatal Mental Health: Clinical Management and Service Guidance. NICE Clinical Guideline 45*. London: The British Psychological Society and The Royal College of Psychiatrists; 2007.

89. Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression. Two questions are as good as many. *J Gen Intern Med* 1997;**12**:439–45. https://doi.org/10.1046/j.1525-1497.1997.00076.x

90. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983;**67**:361–70. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x

91. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient Health Questionnaire. *JAMA* 1999;**282**:1737–44. https://doi.org/10.1001/jama.282.18.1737

92. NICE. *Depression in Adults: The Treatment and Management of Depression in Adults. NICE Clinical Guideline 90*. London: The British Psychological Society and The Royal College of Psychiatrists; 2009.

93. Arroll B, Goodyear-Smith F, Kerse N, Fishman T, Gunn J. Effect of the addition of a 'help' question to two screening questions on specificity for diagnosis of depression in general practice: diagnostic validity study. *BMJ* 2005;**331**:884. https://doi.org/10.1136/bmj.38607.464537.7C

94. Spitzer RL, Williams JB, Kroenkc K, Linzer M, de Gruy EV III, Hahn SR, *et al.* Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. *JAMA* 1994;**272**:1749–59. https://doi.org/10.1001/jama.1994.03520220043029

95. Robins LN, Helzer JE, Croughan J, Ratcliff KS. National Institute of Mental Health Diagnostic Interview Schedule. Its history, characteristics, and validity. *Arch Gen Psychiatry* 1981;**38**:381–9. https://doi.org/10.1001/archpsyc.1981.01780290015001

96. Jordanova V, Wickramesinghe C, Gerada C, Prince M. Validation of two survey diagnostic interviews among primary care attendees: a comparison of CIS-R and CIDI with SCAN ICD-10 diagnostic categories. *Psychol Med* 2004;**34**:1013–24. https://doi.org/10.1017/S0033291703001727

97. Mann R, Gilbody S. Validity of two case finding questions to detect postnatal depression: a review of diagnostic test accuracy. *J Affect Disord* 2011;**133**:388–97. http://dx.doi.org/10.1016/j.jad.2010.11.015

98. Gjerdingen D, Crow S, McGovern P, Miner M, Center B. Postpartum depression screening at well-child visits: validity of a 2-question screen and the PHQ-9. *Ann Fam Med* 2009;**7**:63–70. http://dx.doi.org/10.1370/afm.933

99. First M, Spitzer R, Gibbon M, Williams JB. *Structured Clinical Interview for DSM-IV Clinical Version (SCID-I/CV)*. Washington, DC: American Psychiatric Press; 1997.

100. Mann R, Adamson J, Gilbody SM. Diagnostic accuracy of case-finding questions to identify perinatal depression. *CMAJ* 2012;**184**:e424–30. http://dx.doi.org/10.1503/cmaj.111213

101. NICE. *Antenatal and Postnatal Mental Health: Clinical Management and Service Guidance. NICE Clinical Guideline 192*. London: The British Psychological Society and The Royal College of Psychiatrists; 2014.

102. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SL, *et al.* Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol Med* 2002;**32**:959–76. https://doi.org/10.1017/S0033291702006074

103. 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. *Arch Womens Ment Health* 2016;**19**:41–9. https://doi.org/10.1007/s00737-015-0508-1

104. Brealey SD, Hewitt C, Green JM, Morrell J, Gilbody S. Screening for postnatal depression – is it acceptable to women and healthcare professionals? A systematic review and meta-synthesis. *J Reprod Infant Psychol* 2010;**28**:328–44.

105. Mann R, Adamson J, Gilbody S. The acceptability of case-finding questions to identify perinatal depression. *Br J Midwifery* 2015;**23**:630–8. https://doi.org/10.12968/bjom.2015.23.9.630

106. Taket A, Nurse J, Smith K, Watson J, Shakespeare J, Lavis V, *et al.* Routinely asking women about domestic violence in health settings. *BMJ* 2003;**327**:673–6. http://dx.doi.org/10.1136/bmj.327.7416.673

107. Mead N, Bower P. Patient-centredness: a conceptual framework and review of the empirical literature. *Soc Sci Med* 2000;**51**:1087–110. https://doi.org/10.1016/S0277-9536(00)00098-8

108. Malpass A, Dowrick C, Gilbody S, Robinson J, Wiles N, Duffy L, *et al.* Usefulness of PHQ-9 in primary care to determine meaningful symptoms of low mood: a qualitative study. *Br J Gen Pract* 2016;**66**:e78–84. http://dx.doi.org/10.3399/bjgp16X683473

109. Johnson RB, Onwuegbuzie AJ. Mixed methods research: a research paradigm whose time has come. *Educ Res* 2004;**33**:14–26.

110. Bower P, Wallace P, Ward E, Graffy J, Miller J, Delaney B, *et al.* Improving recruitment to health research in primary care. *Fam Pract* 2009;**26**:391–7. http://dx.doi.org/10.1093/fampra/cmp037

111. O'Cathain A, Thomas KJ, Drabble SJ, Rudolph A, Goode J, Hewison J. Maximising the value of combining qualitative research and randomised controlled trials in health research: the QUAlitative Research in Trials (QUART) study – a mixed methods study. *Health Technol Assess* 2014;**18**(38). http://dx.doi.org/10.3310/hta18380

112. Salmon P, Peters S, Rogers A, Gask L, Clifford R, Iredale W, *et al.* Peering through the barriers in GPs' explanations for declining to participate in research: the role of professional autonomy and the economy of time. *Fam Pract* 2007;**24**:269–75. https://doi.org/10.1093/fampra/cmm015

113. Petrou S, Cooper P, Murray L, Davidson LL. Economic costs of post-natal depression in a high-risk British cohort. *Br J Psychiatry* 2002;**181**:505–12. https://doi.org/10.1192/bjp.181.6.505

114. Bauer A, Pawlby S, Plant DT, King D, Pariante CM, Knapp M. Perinatal depression and child development: exploring the economic consequences from a South London cohort. *Psychol Med* 2015;**45**:51–61. http://dx.doi.org/10.1017/S0033291714001044

115. Mushlin AI, Ruchlin HS, Callahan MA. Cost-effectiveness of diagnostic tests. *Lancet* 2001;**358**:1353–5. https://doi.org/10.1016/S0140-6736(01)06417-0

116. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press; 2015.

117. Beck AT, Ward CH, Mendelson M. An inventory for measuring depression. *Arch Gen Psychiatry* 1961;**4**:561–71. https://doi.org/10.1001/archpsyc.1961.01710120031004

118. Littlewood E, Ali S, Ansell P, Dyson L, Gascoyne S, Hewitt C, *et al.* Identification of depression in women during pregnancy and the early postnatal period using the Whooley questions and the Edinburgh Postnatal Depression Scale: protocol for the Born and Bred in Yorkshire: PeriNatal Depression Diagnostic Accuracy (BaBY PaNDA) study. *BMJ Open* 2016;**6**:e011223. http://dx.doi.org/10.1136/bmjopen-2016-011223

119. Lewis G, Pelosi AJ, Araya R, Dunn G. Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. *Psychol Med* 1992;**22**:465–86. https://doi.org/10.1017/S0033291700030415

120. Mohr DC, Hart SL, Julian L, Tasch ES. Screening for depression among patients with multiple sclerosis: two questions may be enough. *Mult Scler* 2007;**13**:215–19. https://doi.org/10.1177/1352458506070926

121. McManus D, Pipkin SS, Whooley MA. Screening for depression in patients with coronary heart disease (data from the Heart and Soul Study). *Am J Cardiol* 2005;**96**:1076–81. https://doi.org/10.1016/j.amjcard.2005.06.037

122. Mallen CD, Peat G. Screening older people with musculoskeletal pain for depressive symptoms in primary care. *Br J Gen Pract* 2008;**58**:688–93. http://dx.doi.org/10.3399/bjgp08X342228

123. 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* 2009;**116**:1019–27. http://dx.doi.org/10.1111/j.1471-0528.2009.02148.x

124. Hewitt CE, Gilbody SM, Mann R, Brealey S. Instruments to identify post-natal depression: which methods have been the most extensively validated, in what setting and in which language? *Int J Psychiatry Clin Pract* 2010;**14**:72–6.

125. Murray D, Cox JL. Screening for depression during pregnancy with the Edinburgh Depression Scale (EDDS). *J Reprod Infant Psychol* 1990;**8**:99–107.

126. Cox J, Holden J. *Perinatal Mental Health: A Guide to the Edinburgh Postnatal Depression Scale (EPDS)*. London: Royal College of Psychiatrists; 2003.

127. Gruenberg AM, Goldstein RD, Pincus HA. *Classification of Depression: Research and Diagnostic Criteria: DSM-IV and ICD-10*. Weinheim: Wiley; 2005. https://doi.org/10.1002/9783527619672.ch1

128. Singleton N, Bumpstead R, O'Brien M, Lee A, Meltzer H. *Psychiatric Morbidity Among Adults Living in Private Households, 2000*. London: The Stationery Office; 2001.

129. Evans M, Kessler D, Lewis G, Peters TJ, Sharp D. Assessing mental health in primary care research using standardized scales: can it be carried out over the telephone? *Psychol Med* 2004;**34**:157–62. https://doi.org/10.1017/S0033291703008055

130. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. *Gen Hosp Psychiatry* 2016;**39**:24–31. http://dx.doi.org/10.1016/j.genhosppsych.2015.11.005

131. Gilbody S, Richards D, Barkham M. Diagnosing depression in primary care using self-completed instruments: UK validation of PHQ-9 and CORE-OM. *Br J Gen Pract* 2007;**57**:650–2.

132. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006;**166**:1092–7. https://doi.org/10.1001/archinte.166.10.1092

133. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Ann Intern Med* 2007;**146**:317–25. https://doi.org/10.7326/0003-4819-146-5-200703060-00004

134. García-Campayo J, Zamorano E, Ruiz MA, Pardo A, Pérez-Páramo M, López-Gómez V, *et al.* Cultural adaptation into Spanish of the generalized anxiety disorder-7 (GAD-7) scale as a screening tool. *Health Qual Life Outcomes* 2010;**8**:8. http://dx.doi.org/10.1186/1477-7525-8-8

135. Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosom Med* 2002;**64**:258–66. https://doi.org/10.1097/00006842-200203000-00008

136. van Ravesteijn H, Wittkampf K, Lucassen P, van de Lisdonk E, van den Hoogen H, van Weert H, *et al.* Detecting somatoform disorders in primary care with the PHQ-15. *Ann Fam Med* 2009;**7**:232–8. http://dx.doi.org/10.1370/afm.985

137. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;**34**:220–33. https://doi.org/10.1097/00005650-199603000-00003

138. Jenkinson C, Layte R, Lawrence K. Development and testing of the Medical Outcomes Study 36-item short form health survey summary scale scores in the United Kingdom: results from a large-scale survey and a clinical trial. *Med Care* 1997;**35**:410–16.

139. Kessler D, Lewis G, Kaur S, Wiles N, King M, Weich S, *et al.* Therapist-delivered internet psychotherapy for depression in primary care: a randomised controlled trial. *Lancet* 2009;**374**:628–34. http://dx.doi.org/10.1016/S0140-6736(09)61257-5

140. Morrell C, Warner R, Slade P, Dixon S, Walters S, Paley G, *et al.* Psychological interventions for postnatal depression: cluster randomised trial and economic evaluation. The PoNDER trial. *Health Technol Assess* 2009;**13**(30). https://doi.org/10.3310/hta13300

141. EuroQol Group. EuroQol – a new facility for the measurement of health-related quality of life. *Health Policy* 1990;**16**:199–208.

142. Sobocki P, Ekman M, Agren H, Krakau I, Runeson B, Mårtensson B, *et al.* Health-related quality of life measured with EQ-5D in patients treated for depression in primary care. *Value Health* 2007;**10**:153–60. https://doi.org/10.1111/j.1524-4733.2006.00162.x

143. Sharp DJ, Chew-Graham C, Tylee A, Lewis G, Howard L, Anderson I, *et al.* A pragmatic randomised controlled trial to compare antidepressants with a community-based psychosocial intervention for the treatment of women with postnatal depression: the RESPOND trial. *Health Technol Assess* 2010;**14**(43). http://dx.doi.org/10.3310/hta14430

144. Gemmill AW, Leigh B, Ericksen J, Milgrom J. A survey of the clinical acceptability of screening for postnatal depression in depressed and non-depressed women. *BMC Public Health* 2006;**6**:211. https://doi.org/10.1186/1471-2458-6-211

145. Flahault A, Cadilhac M, Thomas G. Sample size calculation should be performed for design accuracy in diagnostic test studies. *J Clin Epidemiol* 2005;**58**:859–62. https://doi.org/10.1016/j.jclinepi.2004.12.009

146. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, *et al.* STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015;**351**:h5527. http://dx.doi.org/10.1136/bmj.h5527

147. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, *et al.* Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ* 2003;**326**:41. http://dx.doi.org/10.1136/bmj.326.7379.41

148. Matthey S, Ross-Hamid C. Repeat testing on the Edinburgh Depression Scale and the HADS-A in pregnancy: differentiating between transient and enduring distress. *J Affect Disord* 2012;**141**:213–21. http://dx.doi.org/10.1016/j.jad.2012.02.037

149. Tourangeau R. Cognitive Science and Survey Methods. In Jabine TB, Straf ML, Tanur JM, Tourangeau R, editors. *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines*. Washington, DC: National Academy Press; 1984. pp. 73–100.

150. Ritchie J, Spencer L. Qualitative Data Analysis for Applied Policy Research. In Huberman AM, Miles MB, editors. *The Qualitative Researcher's Companion*. Thousand Oaks, CA: Sage Publications Inc.; 2002. pp. 305–29.

151. El-Den S, O'Reilly CL, Chen TF. A systematic review on the acceptability of perinatal depression screening. *J Affect Disord* 2015;**188**:284–303. http://dx.doi.org/10.1016/j.jad.2015.06.015

152. Dennis CL, Chung-Lee L. Postpartum depression help-seeking barriers and maternal treatment preferences: a qualitative systematic review. *Birth* 2006;**33**:323–31. https://doi.org/10.1111/j.1523-536X.2006.00130.x

153. Beck CT. Theoretical perspectives of postpartum depression and their treatment implications. *MCN Am J Matern Child Nurs* 2002;**27**:282–7. https://doi.org/10.1097/00005721-200209000-00008

154. Goodman JH. Women's attitudes, preferences, and perceived barriers to treatment for perinatal depression. *Birth* 2009;**36**:60–9. http://dx.doi.org/10.1111/j.1523-536X.2008.00296.x

155. Robertson E, Celasum N, Stewart DE. Risk Factors for Postpartum Depression. In Stewart DE, Robertson E, Dennis C-L, Grace SL, Wallington T, editors. *Postpartum Depression: Literature Review of Risk Factors and Interventions. Maternal Mental Health & Child Health and Development*. Geneva: Department of Mental Health and Substance Abuse, World Health Organization; 2008.

156. Nicolson P. Understanding postnatal depression: a mother-centred approach. *J Adv Nurs* 1990;**15**:689–95. https://doi.org/10.1111/j.1365-2648.1990.tb01892.x

157. Armstrong SJ, Small RE. The paradox of screening: rural women's views on screening for postnatal depression. *BMC Public Health* 2010;**10**:744. http://dx.doi.org/10.1186/1471-2458-10-744

158. Slade P, Morrell CJ, Rigby A, Ricci K, Spittlehouse J, Brugha TS. Postnatal women's experiences of management of depressive symptoms: a qualitative study. *Br J Gen Pract* 2010;**60**:e440–8. http://dx.doi.org/10.3399/bjgp10X532611

159. Kim JJ, La Porte LM, Corcoran M, Magasi S, Batza J, Silver RK. Barriers to mental health treatment among obstetric patients at risk for depression. *Am J Obstet Gynecol* 2010;**202**:312.e1–5. http://dx.doi.org/10.1016/j.ajog.2010.01.004

160. Matthey S. Are we overpathologising motherhood? *J Affect Disord* 2010;**120**:263–6. http://dx.doi.org/10.1016/j.jad.2009.05.004

161. Berggren-Clive K. Out of the darkness and into the light: women's experiences with depression after childbirth. *Can J Commun Ment Health* 1998;**17**:103–20. https://doi.org/10.7870/cjcmh-1998-0006

162. Bennett IM, Coco A, Coyne JC, Mitchell AJ, Nicholson J, Johnson E, *et al.* Efficiency of a two-item pre-screen to reduce the burden of depression screening in pregnancy and postpartum: an IMPLICIT network study. *J Am Board Fam Med* 2008;**21**:317–25. http://dx.doi.org/10.3122/jabfm.2008.04.080048

163. Beauchamp H. What factors influence the use of the Whooley questions by health visitors? *J Health Visit* 2014;**2**:378–87.

164. Bennett IM, Palmer S, Marcus S, Nicholson JM, Hantsoo L, Bellamy S, *et al.* 'One end has nothing to do with the other:' Patient attitudes regarding help seeking intention for depression in gynecologic and obstetric settings. *Arch Womens Ment Health* 2009;**12**:301–8.

165. Smith MV, Gotman N, Lin H, Yonkers KA. Do the PHQ-8 and the PHQ-2 accurately screen for depressive disorders in a sample of pregnant women? *Gen Hosp Psychiatry* 2010;**32**:544–8. http://dx.doi.org/10.1016/j.genhosppsych.2010.04.011

166. Donovan J, Sanders C. Key Issues in the Analysis of Qualitative Data in Health Services Research. In Bowling A, Ebrahim S, editors. *Handbook of Health Research Methods*. Berkshire: Open University Press; 2005. pp. 515–32.

167. Green J, Thorogood N. *Qualitative Methods for Health Research*. 3rd edn. Los Angeles, CA: Sage; 2009.

168. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. *BMJ* 2000;**320**:114–16. https://doi.org/10.1136/bmj.320.7227.114

169. Mertens DM, Hesse-Biber S. Triangulation and mixed methods research provocative positions. *J Mixed Meth Res* 2012;**6**:75–9.

170. Miles MB, Huberman AM. *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks, CA: Sage; 1994.

171. NICE. *Guide to the Methods of Technology Appraisal*. London: NICE; 2013.

172. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, *et al.* Consolidated health economic evaluation reporting standards (CHEERS) statement. *Cost Eff Resour Alloc* 2013;**11**:1. https://doi.org/10.1186/1478-7547-11-6

173. Curtis L, Netten A. *Unit Costs of Health and Social Care 2015*. Canterbury: Personal Social Services Research Unit, University of Kent; 2015.

174. Gilbert P. *Overcoming Depression: A Self-Help Guide Using Cognitive Behavioural Techniques*. 2nd edn. London: Robinson; 2009.

175. Joint Formulary Committee 2015. *British National Formulary (Online)*. London: BMJ Group and Pharmaceutical Press; 2015. URL: www.bnf.org (accessed 16 July 2016).

176. Kessler D, Bennewith O, Lewis G, Sharp D. Detection of depression and anxiety in primary care: follow up study. *BMJ* 2002;**325**:1016–17. https://doi.org/10.1136/bmj.325.7371.1016

177. Dennis CL, Hodnett E, Kenton L, Weston J, Zupancic J, Stewart DE, *et al.* Effect of peer support on prevention of postnatal depression among high risk women: multisite randomised controlled trial. *BMJ* 2009;**338**:a3064. http://dx.doi.org/10.1136/bmj.a3064

178. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. *Lancet* 2009;**374**:609–19. http://dx.doi.org/10.1016/S0140-6736(09)60879-5

179. Department of Health (DH). *NHS Reference Costs*. London: DH; 2014. URL: www.gov.uk/government/collections/nhs-reference-costs (accessed 16 July 2016).

180. Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S, *et al.* Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health Technol Assess* 2015;**19**(14). http://dx.doi.org/10.3310/hta19140

181. Office for National Statistics (ONS). *Conceptions in England and Wales, 2014*. London: ONS; 2014. URL: www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/conceptionandfertilityrates/bulletins/conceptionstatistics/2014 (accessed June 2016).

182. 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* 2005;**119**:1–8.

183. Brugha TS, Morrell CJ, Slade P, Walters SJ. Universal prevention of depression in women postnatally: cluster randomized trial evidence in primary care. *Psychol Med* 2011;**41**:739–48. http://dx.doi.org/10.1017/S0033291710001467

184. Thombs BD, Arthurs E, El-Baalbaki G, Meijer A, Ziegelstein RC, Steele RJ. Risk of bias from inclusion of patients who already have diagnosis of or are undergoing treatment for depression in diagnostic accuracy studies of screening tools for depression: systematic review. *BMJ* 2011;**343**:d4825. http://dx.doi.org/10.1136/bmj.d4825

185. Kessler D, Lloyd K, Lewis G, Gray DP. Cross sectional study of symptom attribution and recognition of depression and anxiety in primary care. *BMJ* 1999;**318**:436–9. https://doi.org/10.1136/bmj.318.7181.436

186. Raffle AE, Gray JAM. *Screening: Evidence and Practice*. Oxford: Oxford University Press; 2007. https://doi.org/10.1093/acprof:oso/9780199214495.001.0001

187. Gilbody S, Sheldon T, Wessely S. Should we screen for depression? *BMJ* 2006;**332**:1027–30. https://doi.org/10.1136/bmj.332.7548.1027

188. Thombs BD, Arthurs E, Coronado-Montoya S, Roseman M, Delisle VC, Leavens A, *et al.* Depression screening and patient outcomes in pregnancy or postpartum: a systematic review. *J Psychosom Res* 2014;**76**:433–46. http://dx.doi.org/10.1016/j.jpsychores.2014.01.006

Appendix 1 Study information for women

Our ref: *<BABY Study ID>*

<Participant Name>
<Address 1>
<Address 2>
<Address 3>
<Address 4>
<Postcode>

Dear <Participant Name>,

**Identifying Depression in Pregnancy & Early Motherhood:
BABY PaNDA Study**

Thank you for taking part in the BABY study. We are now writing to invite you to take part in BABY PaNDA, a special aspect of the BABY study.

The BABY PaNDA Study is being carried out by researchers in the Department of Health Sciences at the University of York in partnership with doctors, midwives and health visitors involved in the care of women during pregnancy and early motherhood in local NHS Trusts. The study has been funded by the government through the Health Services & Delivery Research Programme. The study has been reviewed and approved by the North East Northern and Yorkshire Research Ethics Committee.

The BABY PaNDA Study involves investigating screening for depression in pregnancy and early motherhood. We want to know whether a brief screening questionnaire is a good way of identifying women with depression during pregnancy and after birth. We hope to involve 350-400 women in the study. The enclosed information sheet describes the study in more detail. Please take time to read it carefully and discuss it with others if you wish. If you would like more information about the study, our contact details are at the top of this letter.

If you would like to take part in the study, please fill in the enclosed consent form and send it to the research team in the enclosed pre-paid stamped addressed envelope. You will also need to keep this letter and the information sheet. You will have the chance to ask more questions about the study when the researcher makes contact with you.

You are under no obligation to take part if you do not feel that the study is right for you. If you decide you are not interested, your normal care will continue.

Yours sincerely,

Simon Gilbody
Professor of Psychological Research & Health Services Research

Identifying Depression in Pregnancy & Early Motherhood: BABY PaNDA Study

Participant Information Sheet

Thank you for taking part in the BABY study. You are now being invited to take part in BABY PaNDA, a special aspect of BABY investigating screening for depression in pregnancy and early motherhood. Before you decide, please take time to read the following information carefully and discuss it with family, friends or your GP if you wish. Our contact details are on the back page if you would like any more information.

You can continue to take part in other aspects of the BABY study even if you decide to take part in BABY PaNDA.

What is the purpose of BABY PaNDA?

We would like to find out more about depression during pregnancy and early motherhood, and the best ways of identifying it. Approximately 1 in every 10 women is thought to experience depression during this time. Unfortunately, it is often not identified with around half the women affected not getting the help and support they need.

The National Institute for Health and Clinical Excellence (NICE) have recommended that a brief two-question screening questionnaire should be used to identify women with depression during pregnancy and after birth. However, it is not yet known how well this brief questionnaire works compared to other longer screening questionnaires. BABY PaNDA aims to find out.

Why have I been invited?

We are inviting women who have joined the BABY study early in their pregnancy to also take part in BABY PaNDA. We hope approximately 350-400 women will agree to take part in BABY PaNDA.

What does BABY PaNDA involve?

The study involves taking part in three questionnaire sessions with a specially trained researcher to find out how you feel at three time points during your pregnancy and

after your baby is born. Each session will last around 30 to 40 minutes and take place at approximately:-

1. 20 weeks of pregnancy
2. 3-4 months after the birth of your baby
3. 1 year after the birth of your baby

You will be asked some brief questions and you will also be asked to fill in some short questionnaires (we can help you with these if necessary).

A researcher will telephone you before each session to arrange them at a time and place of your choice. Your first two sessions will take place face-to-face. The final session, at around the time of your baby's first birthday, will take place by telephone; if you prefer, you can complete the questionnaires online through a secure website, or we can send some of the questionnaires to you by post for you to complete and return to us.

Additional Interviews

Approximately 25-30 women taking part in BABY PaNDA will be asked to take part in up to three additional face-to-face interviews. These are to help us collect more information, including women's views on the depression screening questionnaires and how they feel about them. This information is very important as we need to find out if women are happy with the questions used to identify depression during pregnancy and early motherhood. If you have given us consent to look at your medical records as part of the BABY study, we may also collect information from your medical records about the NHS services you have used whilst taking part in BABY PaNDA.

The additional interviews will take place one to two weeks after each BABY PaNDA questionnaire session has been carried out, and take up to one hour to complete. You will be interviewed by an experienced female researcher from the study team at a time and place of your choice. If you agree, the interviews will be recorded and any information that could identify you will be removed when the interview is written up afterwards.

If you decide not to take part in the additional interviews, you can still take part in BABY PaNDA by completing the questionnaire sessions at 20 weeks of pregnancy, and 3-4 months and 1 year following the birth of your baby.

Do I have to take part in BABY PaNDA?

It is up to you to decide whether to not to take part. If you do decide to take part, you are free to withdraw from the study at any time without giving a reason. You can remain in the BABY study even if you decide to withdraw from BABY PaNDA. Whatever your decision, it will not affect the standard of care you receive or your relationship with your doctor, midwife or health visitor.

If you take part, you may find that during your maternity care your midwife and/or health visitor may ask you to complete some of the same screening questions to identify depression that we are using in BABY PaNDA. Please do not worry if this happens. The questionnaires that you complete as part of BABY PaNDA will only be used for research purposes.

Why should I help?

Your contribution to BABY PaNDA will be very valuable in helping us understand whether the brief screening questionnaire is effective at identifying depression during pregnancy and early motherhood. There will be no benefits to your usual maternity care or the care you receive from your health visitor. Taking part in this study will involve some of your time to complete the questionnaires. We do not anticipate any other disadvantages to you taking part in the study.

Will my information be kept confidential?

Yes. All information collected about you during the study will be treated confidentially and in accordance with the Data Protection Act. Any personal information which could identify you will be kept separately from your study information and will only be accessed by members of the research team.

We will destroy all personal information immediately after the end of the study. We will store your study information securely for 20 years and then destroy it.

If you decide to withdraw from the study, your personal information will be securely destroyed. We would still, however, like to use other information collected from you for the BABY PaNDA study. Please let us know if you do not want us to keep this information if you withdraw from BABY PaNDA.

Will my GP be involved?

All GP practices in the study area have received information about BABY and BABY PaNDA. We will let your GP know if you decide to take part in BABY PaNDA. We will also tell your GP if you later decide to withdraw from the study.

If we are worried about your welfare or safety, or the welfare or safety of your baby or other people, the researcher may need to contact your GP so it is important your GP is aware that you are taking part in BABY PaNDA. Please ask the researcher if you have any concerns or questions about this.

You will continue to receive your usual care from your GP, midwife and health visitor throughout the duration of the study.

What will happen to the results at the end of the study?

The study findings will be published in a research report and in articles for health professionals. We will also send you a summary of the study findings at the end of BABY PaNDA. You will not be personally identified in any publications from this study.

Who has reviewed the research?

The research has been reviewed and approved by the North East - York Research Ethics Committee.

Who is organising and funding the research?

This research is being done by researchers in the Department of Health Sciences at the University of York and doctors, midwives and health visitors involved in the care of women during pregnancy and after birth in the local NHS Trust. The research is being funded by the National Institute for Health Research Health Services and Delivery Research (NIHR HS&DR) Programme.

What should I do now?

If you would like to take part, please read and sign the consent form and send it back to us in the pre-paid envelope enclosed. A researcher will then contact you to arrange your first questionnaire session. If you would like more information before deciding, please contact us – our contact details are on the back page.

If you decide that you do not wish to take part, no further action is required.

What do I do if I want to withdraw or complain?

Every care will be taken during this study. If you wish to withdraw from BABY PaNDA or have a concern about any aspect of this study, please contact the study coordinator using the details below. If you do not want to contact the study coordinator you can contact the Chief Investigator, Professor Simon Gilbody (telephone number: [*insert contact number* / email: [*insert email address]*]).

Standard NHS indemnity arrangements apply to this research. While we anticipate no harm or distress to anyone as a result of this study, if you are harmed during this study due to someone's negligence, then you may have grounds for legal action, but you may have to pay for it. If you wish to complain formally, the normal NHS complaint mechanisms are available to you.

Further information and contact details

For further general information about research, please see the INVOLVE website at http://www.invo.org.uk

For specific information about BABY PaNDA, please contact:

BaBY PaNDA Coordinator: [*insert name]*

Address: [*insert current address]*
Freephone: [*insert contact number)*
Email: [*insert email address]*
Website: [*insert website address]*

Thank you for reading this information.

The BABY PaNDA Study:
Identifying Depression in Pregnancy & Early Motherhood
Summary Information Sheet

BABY PaNDA is a special part of the BABY study set up to investigate screening for depression in pregnancy and early motherhood. Women who join the BaBY study early in pregnancy may be invited to take part in BaBY PaNDA too.

What is the purpose of BABY PaNDA?

To find out more about depression during pregnancy and early motherhood and the best ways to identify it. We know that around half the women affected are not getting the help and support they need because their depression is often not identified. BABY PaNDA hopes to find out whether brief screening questionnaires are a good way to identify women with depression during this time.

Who is being invited to take part in BaBY PaNDA?

Women who have recently joined the BaBY study early in their pregnancy. Invitation letters are being sent to eligible women between 15 and 18 weeks of pregnancy. We hope that 350-400 women will agree to take part in BaBY PaNDA too.

What does BABY PaNDA involve?

Three meetings with a specially trained researcher to find out how you feel at three time points; once during pregnancy and twice during your baby's first year. Women will be asked some brief questions and to fill in some short questionnaires. Each session lasts 30 – 40 minutes and takes place at approximately 20 weeks of pregnancy, and 3-4 months and 1 year after your baby is born.

A few women will be asked to take part in up to three additional face-to-face interviews. These are to collect information about women's views on the depression screening questionnaires and how they feel about them. Women may still take part in BABY PaNDA even if they prefer not to take part in the additional interviews.

What should I do if I would like to take part in BaBY PaNDA?

Please read the Participant Information Sheet enclosed with your PaNDA invitation letter. It contains more detailed information about what the study involves. Then read and complete the consent form and post it back to us in the pre-paid envelope

provided. A researcher will contact you before each meeting to arrange them at a time and place of your choice.

If you would like more information about taking part in BaBY PaNDA:
Please contact the study coordinator, [*insert name*]: Freephone: [*insert contact number*] or Email: [*insert email address*]

baby panda

☐☐ / ☐☐ ☐☐☐

If you would like to take part in the BABY PaNDA Study, it would be really helpful if you could complete the information below to allow us to keep our records up-to-date. Please send this form, along with your completed consent form, to the research team in the pre-paid envelope enclosed.

Name:	
Address: Postcode:	
Mobile No:	
Landline No:	
Email Address:	
GP Name:	
GP Address:	
Date & location of 20 week scan, if known/booked	

baby panda   THE UNIVERSITY *of York*   ☐☐ / ☐☐ ☐☐☐

Identifying Depression in Pregnancy & Early Motherhood:BABY PaNDA Study

Consent Form

Thank you for reading the information about this study. If you would like to take part, please read and sign this form. Please put your **initials** in the boxes of the statements you agree with:

1. I have read the patient information sheet and have been given a copy to keep. I have been able to ask questions about the study and I understand why the research is being done.   ☐

2. I understand that my participation in this research is entirely voluntary and that I will not receive any payment. I understand that I am free to withdraw my consent at any time without giving a reason and without my medical care being affected.   ☐

3. I am willing to complete the BABY PaNDA questionnaires at approximately 20 weeks of pregnancy and at approximately 3-4 months and 1 year after my baby's birth.   ☐

4. I am willing to be interviewed by an experienced researcher approximately one week after completing the BABY PaNDA questionnaires if I am approached to do so.   ☐

    a. I am happy for these interviews to be recorded and kept safe in an anonymised format.   ☐

    b. I understand that direct quotes from these interviews may be used in the publication of the study results but that these will be anonymised.   ☐

5. I understand that relevant sections of my medical notes and data collected during the study may be looked at by individuals from the University of York, from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give my permission for these individuals to have access to my records.   ☐

6. I agree to my GP being informed of my participation in this study and updated with information from this study relevant to my medical care.   ☐

………………………………………………   ………………..   …………………………………..
Name of patient (BLOCK CAPITALS)   **Date**   **Signature**

**PLEASE RETURN THIS FORM IN THE STAMPED ADDRESSED ENVELOPE PROVIDED /
PLEASE RETAIN THIS FORM FOR YOUR RECORDS**

Appendix 2 Screening questionnaires timings record

Screening Questionnaires: Timings Record

The following record should be completed by the researcher for all participant assessments completed at 20 weeks pregnancy (prenatal), 3-4 months post-birth (postnatal) and 12 months post-birth (follow-up).

Date Completed   [dd/mm/yyyy]

Staff Initials

Participant ID

Questionnaire Session   Prenatal / Postnatal   / Follow-Up (please circle)

Please indicate below the **approximate TIME it took to administer the depression screening questionnaires**.

NICE ultra-brief depression screening questionnaire (administered verbally):
(This will include the time taken to ask the participant the screening questions, the time taken for participants to provide their responses, and the time taken for the researcher to record the participant's responses)

NICE ultra-brief depression screening questionnaire   mins   secs

Edinburgh Postnatal Depression Scale (EPDS, self-completed by participants)
(This will begin when the researcher provides the participant with the EPDS questionnaire / begins to ask the participant the questions, and will end when the participant indicates that they have completed the questionnaire / when the researcher finishes recording the participant's responses)

Edinburgh Postnatal Depression Scale   mins   secs

Appendix 3 Participant assessment record sheet

baby panda　　　　**Participant Assessment Report Sheet**

Q1: Apart from the NICE Ultra-Brief Screening Questionnaire, was the participant assessment read out by the researcher or read by the participant?

Read by researcher　　　　　　　　　　Read by participant
☐　　　　　　　　　　　　　　　　　　☐

(Please record any additional details).

Q2: Did the participant ask any questions during the assessment (particularly with respect to the CIS-R, the NICE ultra-brief questionnaire, the EPDS and the PHQ9)?

Yes　　　　　　　　　　　　　　No
☐　　　　　　　　　　　　　　☐

(If 'Yes' please provide further details, indicating which questionnaire(s) the questions applied to). (It is important to report this information as it is possible that a researcher could unintentionally influence the participant's response in some way on the basis of knowledge the researcher has about the participant's response on another questionnaire – this could artificially inflate the diagnostic performance of the test).

Q3: Were there any un-blinding issues in relation to the assessment?

Yes　　　　　　　　　　　　　　No
☐　　　　　　　　　　　　　　☐

(If 'Yes' please provide further details). For example, the participant revealed their responses/outcome from a previous PaNDA assessment; the participant revealed that they have depression and/or that they are on medication for depression).

Q4: If yes to Q3, has this been reported to the Project Manager?　　Yes ☐　No ☐　N/A ☐

Participant ID:	☐☐ / ☐☐☐☐
Session:	20 weeks pregnancy / 3-4 months after birth / 12 months after birth (please circle)
Date of Session:	☐☐ / ☐☐ / ☐☐☐☐
Researcher Signature	

Appendix 4 Identifying potential sources of bias

Identifying potential sources of bias

Method

As discussed in *Chapter 2*, a PARS (see *Appendix 3*) was used to capture any instances of potential bias within the assessment sessions. The PARS invited researchers to note down any questions asked during the assessment, particularly those asked during completion of the index tests (Whooley questions and EPDS) or the diagnostic reference standard (CIS-R), and any information that the woman may have disclosed during the assessment session. In addition, researchers were also asked to document any unblinding issues in relation to the assessment. To reiterate from *Chapter 2*, women self-completed the EPDS (index test) and CIS-R (diagnostic reference standard), whereas researchers asked women the Whooley questions (index test) verbally.

All recorded information was reviewed and assessed for any potential bias with reference to STARD criteria, namely points 13a, 'whether clinical information and reference standard results were available to the performers/readers of the index test', and 13b, 'whether clinical information and index test results were available to the assessors of the reference standard'.[146] We adopted the terminology used in the STARD criteria: recorded information was categorised into (a) 'clinical information' and (b) 'questions asked'. 'Clinical information' was further classified into subcategories. The point during the assessment when clinical information was disclosed or questions were asked (e.g. before the assessment began, during completion of the index tests or the diagnostic reference standard, or any of the additional outcome measures, or after the assessment) was recorded for each of these categories (if this information was documented).

Clinical information

Clinical information was classified into three subcategories: (1) 'psychopathology', (2) 'mood issues/symptoms' and (3) 'external/life events'. Psychopathology included information about current and/or historic diagnoses of depression or anxiety or disclosure of current treatment (e.g. medication/talking therapies/service involvement); mood issues/symptoms included information disclosed about how the woman was feeling (e.g. reported feelings of 'feeling low', 'emotional', 'low mood', 'feeling overwhelmed'), but no reference was made to a formal diagnosis or treatment; external/life events included information that eluded to mood issues associated with an occurrence of an external or life event (e.g. family issues, bereavements). Examples of clinical information disclosed for each subcategory can be found in *Box 1*.

Questions

Questions asked by women were recorded. Information regarding which outcome measure (index tests or reference standard) the question related to and at what point during the assessment the question was posed to the researcher was documented.

Inter-rater reliability

Categorisation of information from the PARS was undertaken by JP in the first instance and then independently rated by EL. The concordance rate between the two reviewers was 93%. Discrepancies were resolved through discussion and a mutual decision was made on the final categories for the information.

Results

The results are presented separately for clinical information and questions women asked.

BOX 1 Examples of clinical information disclosed for each subcategory of psychopathology, mood issues/symptoms and external/life events

Psychopathology

. . . informed me that she was on anti-depressants – 7 month before baby born as she was noticing signs of post-natal depression (experienced with 1st child) this was when answering the Whooley Q's.

Participant revealed before the assessment began that she was currently taking anti-depressants and is receiving support from 'lifeline' for previous binge drinking) . . .

Revealed under care of CPN [community psychiatric nurse] and on medication (both before + during the assessment).

. . . disclosed at the beginning of the session that she had severe postnatal depression with a previous pregnancy and was admitted to hospital.

Mood issues/symptoms

During the correspondence we had to arrange our meeting, Pt [participant] wrote 'might help more with the way my moods etc at the moment'.

With reference to NICE – felt she had felt down, but not often. She therefore felt it more appropriate to answer 'No'.

. . . in response to 1st NICE Q she said she had felt low but she wouldn't say 'often'.

. . . relating to NICE Q1, had one or two episodes of feeling overwhelmed with work/uni module/moving house . . .

External/life events

Participant explained that her mother has been seriously ill which has caused considerable worry – re EPDS – she has been crying but felt that this was not caused by unhappiness just worry.

Participant said that she had been feeling down during the last month during the NICE ultra-brief questionnaire but said that this had nothing to do with the pregnancy and was because her mum had recently been diagnosed with breast cancer.

Participant disclosed that her previous child died, and she has been having some concerns going into this pregnancy (in relation to birth & acceptance of new baby).

. . . informed me she was answering positively on Q1 on NICE due to illness in family affects how she feels.

Clinical information

STARD criterion 13a: 'whether clinical information and reference standard results were available to the performers/readers of the index test'

Stage 1: prenatal assessment (20 weeks' gestation)   The index tests were completed by women before the diagnostic reference standard, therefore the results of the reference standard were not available to the readers/performers (i.e. the researchers) of the index tests. Clinical information recorded before the assessment began and during completion of the index tests are presented in *Table 31*.

The total number of participants who disclosed clinical information at the stage 1 assessment was 41 (10.5%). Seven (1.8%) women disclosed clinical information before the assessment and 11 (2.8%) women disclosed clinical information during each of the index tests. The point at which clinical information was disclosed was not known for 12 (3.1%) women.

Of the clinical information disclosed, 19 (4.9%) cases were deemed to fit in the psychopathology category (e.g. 'Told me at the beginning of the session that she is on medication for depression'), 15 (3.8%) were deemed to be related to mood issues/symptoms (e.g. 'Participant deliberated over Q1 on ultra-brief questions "have occasionally felt down as not feeling well" advised to answer as honestly as possible') and 7 (1.8%) were deemed to be external/life events (e.g. 'Participant said she had been feeling down during the last month . . . this has nothing to do with the pregnancy and was because her mum had recently been diagnosed with breast cancer').

Stage 2: postnatal assessment (3–4 months postnatal)   As this was the second assessment women completed as part of the study, there was the potential for information relating to the results from stage 1 to be disclosed to the researcher (different researchers conducted the assessments at the two time points). Two women disclosed information to the researcher about their assessment at stage 1; one disclosed that a risk protocol was enacted and another stated that the previous assessment had made her reflect on a previous episode of depression. However, neither of these women went on to ask a question on any of the index tests or the reference standard, thereby minimising the risk of bias during these assessment sessions.

Clinical information recorded before the assessment began and during completion of the index tests is presented in *Table 32.* The total number of participants who disclosed clinical information at stage 2 was 30 (8.7%). Eight (2.3%) women disclosed clinical information before the assessment and 14 (4.1%) women disclosed clinical information during the Whooley questions. The point at which clinical information was disclosed was not known for eight (2.3%) women. Of the clinical information disclosed, 18 (5.2%) cases were deemed to fit in the psychopathology category, 10 (2.9%) were deemed to be related to mood issues/symptoms and two (0.6%) were deemed to be external/life events.

TABLE 31 Clinical information disclosed during stage 1 assessment

Point during assessment at which clinical information disclosed	Psychopathology, *n* (%)	Mood issues/ symptoms, *n* (%)	External/life events, *n* (%)	Total (*N* = 391), *n* (%)
Before assessment commenced	3 (0.8)	4 (1)	0 (0)	7 (1.8)
Whooley questions	5 (1.3)	5 (1.3)	1 (0.3)	11 (2.8)
EPDS	1 (0.3)	6 (1.5)	4 (1)	11 (2.8)
Not known	10 (2.6)	0 (0)	2 (0.5)	12 (3.1)

TABLE 32 Clinical information disclosed during stage 2 assessment

Point during assessment clinical information disclosed	Psychopathology, *n* (%)	Mood issues/ symptoms, *n* (%)	External/life events, *n* (%)	Total (*N* = 345), *n* (%)
Before assessment commenced	8 (2.3)	0 (0)	0 (0)	8 (2.3)
Whooley questions	6 (1.7)	7 (2)	1 (0.3)	14 (4.1)
EPDS	0 (0)	0 (0)	0 (0)	0 (0)
Not known	4 (1.2)	3 (0.9)	1 (0.3)	8 (2.3)

STARD criterion 13b: 'whether clinical information and index test results were available to the assessors of the reference standard'

Women's responses to the Whooley questions (index test) would have been known to the assessor (the researcher) of the reference standard at both stages 1 and 2 as these questions were asked verbally. The EPDS was completed as a self-report measure so the researcher would not have had access to the results of this index test before women completed the reference standard. Clinical information (classified as 'psychopathology') was disclosed by only one woman during completion of the reference standard as part of the postnatal assessment (stage 2), indicating a very low risk of bias. Figures relating to disclosure of clinical information on the index tests and reference standard have been discussed above (see *Tables 31* and *32*).

Questions

It is important to consider questions women asked during completion of the index tests and the reference standard as researcher response to questions could provide the opportunity for bias. The number of questions women asked during completion of the index tests and reference standard can be found in *Table 33*.

Few women asked questions during the assessment sessions. The total number of questions asked was 10 (2.6%) and 18 (5.2%) during the assessment sessions at stages 1 and 2, respectively. A total of five (0.7%) questions were asked on each of the index tests across the two stages. A total of 18 (2.4%) women asked questions on the reference standard, although these questions were predominantly (61%) related to either demographic questions (e.g. 'CIS-R queried marital status – has a partner – not single – answer – living as married') or the functionality of the program (e.g. 'Participant asked how to "page-back" on the CIS-R'). The number of questions recorded during completion of the additional outcome measures at both assessment stages is detailed in *Table 34*.

TABLE 33 Number of questions asked by women during assessments at stages 1 and 2

Point during assessment at which question asked	Stage 1: prenatal assessment (*N* = 391), *n* (%)	Stage 2: postnatal assessment (*N* = 345), *n* (%)	Total, *n* (%)
Before assessment commenced	0 (0)	0 (0)	0 (0)
Whooley questions	4 (1)	1 (0.3)	5 (0.7)
EPDS	3 (0.8)	2 (0.6)	5 (0.7)
CIS-R	3 (0.8)	15 (4.3)	18 (2.4)
Total	10 (2.6)	18 (5.2)	28 (3.8)

TABLE 34 Number of questions recorded during completion of the additional outcome measures at both assessment stages

Questionnaire/questions	Prenatal assessment (20 weeks) (*N* = 391), *n* (%)	Postnatal assessment (3–4 months) (*N* = 345), *n* (%)	Total, *n* (%)
PHQ-9	4 (1)	2 (0.6)	6 (0.8)
GAD-7	1 (0.3)	2 (0.6)	3 (0.4)
PHQ-15	3 (0.8)	0 (0)	3 (0.4)
SF-12v2	0 (0)	1 (0.3)	1 (0.1)
EQ-5D	0 (0)	1 (0.3)	1 (0.1)
Resource use questionnaire	16 (4.1)	18 (5.2)	34 (4.6)
Acceptability survey	2 (0.5)	1 (0.3)	3 (0.4)
Biographical questionnaire	2 (0.5)	0 (0)	2 (0.3)
General questions regarding study	3 (0.8)	0 (0)	3 (0.4)
Not known	0 (0)	0 (0)	0 (0)
Total	31 (7.9)	25 (7.2)	56 (7.6)

Combining the data from the clinical information disclosed by women and the number of questions women asked, only two women who revealed information about their current diagnosis and/or treatment for depression went on to ask a question during the same assessment session. Both of these women asked a question during completion of the reference standard (CIS-R); however, in both instances, the questions required minimal involvement from the researcher and were related to the demographic questions administered at the beginning of the reference standard. This indicates very low risk of bias.

Additional factors

As previously detailed in *Chapter 2*, assessment sessions conducted at the prenatal and postnatal stages were ideally completed face to face with women. At stage 1, all assessments were conducted face to face. For those women who were unable to complete the assessment face to face at stage 2, the assessment either took place over the telephone or the self-report study questionnaires were sent out in the post for women to complete at home and (when possible) the reference standard (CIS-R) was conducted over the telephone (within 14 days of completion of the postal questionnaires). Fifteen (4.3%) women completed their postnatal assessment over the telephone; nine (2.6%) women returned their questionnaire via post and then went on to complete the CIS-R over the telephone. For these nine women, blinding between the index tests and reference standard was maintained by having one researcher log and check the returned questionnaires and a different researcher conduct the CIS-R. The mean number of days between completion of the index tests by post and completion of the reference standard by telephone was 8.4 (range 3–13).

Additional information recorded on the PARS related to whether or not the assessment session had been conducted with other people present (i.e. the woman's partner, parents or other children) and if the same researcher had conducted the assessment at both stages. A total of 60 (15.3%) women had another person present during their assessment at stage 1; this reduced to eight (2.3%) women at the assessment at stage 2. The finding that 60 women had someone else (usually their partner) present during the prenatal assessment was to be expected, given that many of these assessments took place during visits to the hospital antenatal clinic. There were only two occasions when the same researcher conducted both the prenatal and postnatal assessments (for reasons of sensitivity to the women).

Summary

In summary, only 10.5% of women disclosed any sort of clinical information at stage 1, with 9% of women disclosing such information at stage 2. Few questions were asked during completion of the index tests or reference standard; 2.6% and 5.2% of women asked a question during the prenatal and postnatal assessment stages, respectively, and these questions related to demographic questions on the reference standard. For those women who asked questions, only two went onto to disclose any clinical information about themselves. In summary, the data collated from the PARS suggest the potential risk of bias (given that the index tests and reference standard were completed in the same assessment session by a single researcher) is considered minimal.

Appendix 5 Data collection forms

The EPDS has been reproduced with permission from: Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression: development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry* 1987;**150**:782–6,[79] and Wisner KL, Parry BL, Piontek CM. Postpartum depression. *N Engl J Med* 2002;**347**:194–9.[2]

Edinburgh Postnatal Depression Scale

As you are pregnant or have recently had a baby, we would like to know how you are feeling.
Please check the answer that comes closest to how you have felt **IN THE PAST 7 DAYS**,
not just how you feel today.

In the past 7 days:

1. I have been able to laugh and see the funny side of things
   - ☐ As much as I always could
   - ☐ Not quite so much now
   - ☐ Definitely not so much now
   - ☐ Not at all

2. I have looked forward with enjoyment to things
   - ☐ As much as I ever did
   - ☐ Rather less than I used to
   - ☐ Definitely less than I used to
   - ☐ Hardly at all

3. I have blamed myself unnecessarily when things went wrong
   - ☐ Yes, most of the time
   - ☐ Yes, some of the time
   - ☐ Not very often
   - ☐ No, never

4. I have been anxious or worried for no good reason
   - ☐ No, not at all
   - ☐ Hardly ever
   - ☐ Yes, sometimes
   - ☐ Yes, very often

5. I have felt scared or panicky for no very good reason
   - ☐ Yes, quite a lot
   - ☐ Yes, sometimes
   - ☐ No, not much
   - ☐ No, not at all

6. Things have been getting on top of me
   - ☐ Yes, most of the time I haven't been able to cope at all
   - ☐ Yes, sometimes I haven't been coping as well as usual
   - ☐ No, most of the time I have coped quite well
   - ☐ No, I have been coping as well as ever

7. I have been so unhappy that I have had difficulty sleeping
   - ☐ Yes, most of the time
   - ☐ Yes, sometimes
   - ☐ Not very often
   - ☐ No, not at all

8. I have felt sad or miserable
   - ☐ Yes, most of the time
   - ☐ Yes, quite often
   - ☐ Not very often
   - ☐ No, not at all

9. I have been so unhappy that I have been crying
   - ☐ Yes, most of the time
   - ☐ Yes , quite often
   - ☐ Only occasionally
   - ☐ No, never

10. The thought of harming myself has occurred to me
    - ☐ Yes, quite often
    - ☐ Sometimes
    - ☐ Hardly ever
    - ☐ Never

Source: Cox, J.L., Holden, J.M., and Sagovsky, R. 1987. Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry* 150:782-786.

Source: K. L. Wisner, B. L. Parry, C. M. Piontek, Postpartum Depression N Engl J Med vol. 347, No 3, July 18, 2002, 194-199

THE UNIVERSITY *of York*

— DEPARTMENT OF —
HEALTH SCIENCES

baby
panda

☐☐ / ☐☐☐☐☐

**Identifying Depression in Pregnancy &
Early Motherhood**

Acceptability Questionnaire (Prenatal)

We would like you to tell us your views about the types of questions used to find out about how you are feeling during pregnancy. Please take a few minutes to read the questions for the different screening questionnaires and respond to the following questions.

NICE Ultra Brief Screening Questionnaire

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?*

If yes to either of the above two questions, a third help question may be asked:

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

---

1. **Do you think it is a good idea to ask women about how you are feeling during pregnancy using the NICE Ultra Brief Screening Questionnaire?** *(please cross one box only for your answer)*

   Yes ☐   No ☐   Unsure ☐

2. **How comfortable did you feel when you were completing the questions for the NICE Ultra Brief Screening questionnaire?** *(please cross one box only for your answer)*

Very comfortable	**Fairly comfortable**	**Comfortable**	**Fairly uncomfortable**	**Uncomfortable**
☐	☐	☐	☐	☐

3. **Was it easy or difficult to understand the meaning of question 1?** *(please cross one box only for your answer)*

Very easy	**Fairly easy**	**Easy**	**Fairly difficult**	**Difficult**
☐	☐	☐	☐	☐

4. Was it easy or difficult to remember whether you have often been feeling down, depressed or hopeless in the last month?

Very easy	Fairly easy	Easy	Fairly difficult	Difficult
☐	☐	☐	☐	☐

5. How sure do you feel about your answer for question 1 about feeling down?

Very sure	Fairly sure	Sure	Fairly unsure	Very unsure
☐	☐	☐	☐	☐

6. Was it easy or difficult to understand the meaning of question 2? *(please cross one box only for your answer)*

Very easy	Fairly easy	Easy	Fairly difficult	Difficult
☐	☐	☐	☐	☐

7. Was it easy or difficult to remember whether you have often been bothered by little interest or pleasure in doing things in the past month?

Very easy	Fairly easy	Easy	Fairly difficult	Difficult
☐	☐	☐	☐	☐

8. How sure do you feel about your answer for question 2 about feeling little interest?

Very sure	Fairly sure	Sure	Fairly unsure	Very unsure
☐	☐	☐	☐	☐

9. How sure do you feel about your answer for question 3 about needing help (if asked)?
(please cross one box only for your answer)

Not asked

☐

Very sure	Fairly sure	Sure	Fairly unsure	Very unsure
☐	☐	☐	☐	☐

10. Please provide any additional information about completing the NICE ultra-brief screening questions you would like to share in the box below.

11. Do you think it is a good idea to ask women about how you are feeling during pregnancy using the EPDS *(see questions for your information below)***?** *(please cross one box only for your answer)*

Yes ☐   No ☐   Unsure ☐

Edinburgh Postnatal Depression Scale (EPDS)

In the past 7 days:

1. I have been able to laugh and see the funny side of things
   - As much as I always could
   - Not quite so much now
   - Definitely not so much now
   - Not at all

2. I have looked forward with enjoyment to things
   - As much as I ever did
   - Rather less than I used to
   - Definitely less than I used to
   - Hardly at all

3. I have blamed myself unnecessarily when things went wrong
   - Yes, most of the time
   - Yes, some of the time
   - Not very often
   - No, never

4. I have been anxious or worried for no good reason
   - No, not at all
   - Hardly ever
   - Yes, sometimes
   - Yes, very often

5. I have felt scared or panicky for no very good reason
   - Yes, quite a lot
   - Yes, sometimes
   - No, not much
   - No, not at all

6. Things have been getting on top of me
   - Yes, most of the time I haven't been able to cope at all
   - Yes, sometimes I haven't been coping as well as usual
   - No, most of the time I have coped quite well
   - No, I have been coping as well as ever

7. I have been so unhappy that I have had difficulty sleeping
   - Yes, most of the time
   - Yes, sometimes
   - Not very often
   - No, not at all

8. I have felt sad or miserable
   - Yes, most of the time
   - Yes, quite often
   - Not very often
   - No, not at all

9. I have been so unhappy that I have been crying
   - Yes, most of the time
   - Yes, quite often
   - Only occasionally
   - No, never

10. The thought of harming myself has occurred to me
    - Yes, quite often
    - Sometimes
    - Hardly ever
    - Never

Source: Cox, J.L., Holden, J.M., and Sagovsky, R. 1987. Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry* 150:782-786.

Source: K. L. Wisner, B. L. Parry, C. M. Piontek, Postpartum Depression N Engl J Med vol. 347, No 3, July 18, 2002, 194-199

12. How comfortable did you feel when you were completing the questions for the EPDS?
(please cross one box only for your answer)

Very comfortable	Fairly comfortable	Comfortable	Fairly uncomfortable	Uncomfortable
☐	☐	☐	☐	☐

13. Was it easy or difficult to understand the meaning of the 10 questions? *(please cross one box only for your answer)*

Very easy	Fairly easy	Easy	Fairly difficult	Difficult
☐	☐	☐	☐	☐

14. Were there any questions (from 1-10) which were particularly difficult to understand?
(please cross the box if you found that question difficult to understand)

1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

15. Was it easy or difficult to remember how you have been feeling in the past 7 days to answer the 10 questions for the EPDS?

Very easy	Fairly easy	Easy	Fairly difficult	Difficult
☐	☐	☐	☐	☐

16. Were there any questions (from 1-10) which were particularly difficult to remember how you were feeling? *(please cross the box if you found that question difficult to remember)*

1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

17. How sure do you feel about your answers for the 10 questions for the EPDS?

Very sure	Fairly sure	Sure	Fairly unsure	Very unsure
☐	☐	☐	☐	☐

18. **Were there any questions (from 1-10) which you were particularly unsure about your answer?** *(please cross the box if you were unsure about your answer to that question)*

1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

19. **Please provide any additional information about completing the EPDS screening questions you would like to share in the box below.**

Thank you for taking the time to complete this questionnaire

THE UNIVERSITY *of York*

— DEPARTMENT OF —
HEALTH SCIENCES

baby panda

☐☐ / ☐☐☐☐

**Identifying Depression in Pregnancy &
Early Motherhood**

Acceptability Questionnaire (Postnatal)

We would like you to tell us your views about the types of questions used to find out about how you are feeling during early motherhood. Please take a few minutes to read the questions for the different screening questionnaires and respond to the following questions.

NICE Ultra Brief Screening Questionnaire

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?*

If yes to either of the above two questions, a third help question may be asked:

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

---

1. **Do you think it is a good idea to ask women about how you are feeling during early motherhood using the NICE Ultra Brief Screening Questionnaire?** *(please cross one box only for your answer)*

Yes ☐　　　No ☐　　　Unsure ☐

2. **How comfortable did you feel when you were completing the questions for the NICE Ultra Brief Screening questionnaire?** *(please cross one box only for your answer)*

Very comfortable	Fairly comfortable	Comfortable	Fairly uncomfortable	Uncomfortable
☐	☐	☐	☐	☐

3. **Was it easy or difficult to understand the meaning of question 1?** *(please cross one box only for your answer)*

Very easy	Fairly easy	Easy	Fairly difficult	Difficult
☐	☐	☐	☐	☐

4. Was it easy or difficult to remember whether you have often been feeling down, depressed or hopeless in the last month?

Very easy	Fairly easy	Easy	Fairly difficult	Difficult
☐	☐	☐	☐	☐

5. How sure do you feel about your answer for question 1 about feeling down?

Very sure	Fairly sure	Sure	Fairly unsure	Very unsure
☐	☐	☐	☐	☐

6. Was it easy or difficult to understand the meaning of question 2? *(please cross one box only for your answer)*

Very easy	Fairly easy	Easy	Fairly difficult	Difficult
☐	☐	☐	☐	☐

7. Was it easy or difficult to remember whether you have often been bothered by little interest or pleasure in doing things in the past month?

Very easy	Fairly easy	Easy	Fairly difficult	Difficult
☐	☐	☐	☐	☐

8. How sure do you feel about your answer for question 2 about feeling little interest?

Very sure	Fairly sure	Sure	Fairly unsure	Very unsure
☐	☐	☐	☐	☐

9. How sure do you feel about your answer for question 3 about needing help (if asked)?
(please cross one box only for your answer)

Not asked

☐

Very sure	Fairly sure	Sure	Fairly unsure	Very unsure
☐	☐	☐	☐	☐

10. **Please provide any additional information about completing the NICE ultra-brief screening questions you would like to share in the box below.**

11. **Do you think it is a good idea to ask women about how you are feeling during early motherhood using the EPDS** *(see questions for your information below)*? *(please cross one box only for your answer)*

Yes ☐   No ☐   Unsure ☐

Edinburgh Postnatal Depression Scale (EPDS)

In the past 7 days:

1. I have been able to laugh and see the funny side of things
   - ☐ As much as I always could
   - ☐ Not quite so much now
   - ☐ Definitely not so much now
   - ☐ Not at all

2. I have looked forward with enjoyment to things
   - ☐ As much as I ever did
   - ☐ Rather less than I used to
   - ☐ Definitely less than I used to
   - ☐ Hardly at all

3. I have blamed myself unnecessarily when things went wrong
   - ☐ Yes, most of the time
   - ☐ Yes, some of the time
   - ☐ Not very often
   - ☐ No, never

4. I have been anxious or worried for no good reason
   - ☐ No, not at all
   - ☐ Hardly ever
   - ☐ Yes, sometimes
   - ☐ Yes, very often

5. I have felt scared or panicky for no very good reason
   - ☐ Yes, quite a lot
   - ☐ Yes, sometimes
   - ☐ No, not much
   - ☐ No, not at all

6. Things have been getting on top of me
   - ☐ Yes, most of the time I haven't been able to cope at all
   - ☐ Yes, sometimes I haven't been coping as well as usual
   - ☐ No, most of the time I have coped quite well
   - ☐ No, I have been coping as well as ever

7. I have been so unhappy that I have had difficulty sleeping
   - ☐ Yes, most of the time
   - ☐ Yes, sometimes
   - ☐ Not very often
   - ☐ No, not at all

8. I have felt sad or miserable
   - ☐ Yes, most of the time
   - ☐ Yes, quite often
   - ☐ Not very often
   - ☐ No, not at all

9. I have been so unhappy that I have been crying
   - ☐ Yes, most of the time
   - ☐ Yes, quite often
   - ☐ Only occasionally
   - ☐ No, never

10. The thought of harming myself has occurred to me
    - ☐ Yes, quite often
    - ☐ Sometimes
    - ☐ Hardly ever
    - ☐ Never

Source: Cox, J.L., Holden, J.M., and Sagovsky, R. 1987. Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry* 150:782-786.

Source: K. L. Wisner, B. L. Parry, C. M. Piontek, Postpartum Depression N Engl J Med vol. 347, No 3, July 18, 2002, 194-199

12. How comfortable did you feel when you were completing the questions for the EPDS?
(please cross one box only for your answer)

Very comfortable	Fairly comfortable	Comfortable	Fairly uncomfortable	Uncomfortable
☐	☐	☐	☐	☐

13. Was it easy or difficult to understand the meaning of the 10 questions? *(please cross one box only for your answer)*

Very easy	Fairly easy	Easy	Fairly difficult	Difficult
☐	☐	☐	☐	☐

14. Were there any questions (from 1-10) which were particularly difficult to understand?
(please cross the box if you found that question difficult to understand)

1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

15. Was it easy or difficult to remember how you have been feeling in the past 7 days to answer the 10 questions for the EPDS?

Very easy	Fairly easy	Easy	Fairly difficult	Difficult
☐	☐	☐	☐	☐

16. Were there any questions (from 1-10) which were particularly difficult to remember how you were feeling? *(please cross the box if you found that question difficult to remember)*

1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

17. How sure do you feel about your answers for the 10 questions for the EPDS?

Very sure	Fairly sure	Sure	Fairly unsure	Very unsure
☐	☐	☐	☐	☐

18. Were there any questions (from 1-10) which you were particularly unsure about your answer? *(please cross the box if you were unsure about your answer to that question)*

1	2	3	4	5	6	7	8	9	10
☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

19. Please provide any additional information about completing the EPDS screening questions you would like to share in the box below.

Thank you for taking the time to complete this questionnaire

ECONOMIC PATIENT QUESTIONNAIRE

(Prenatal)

Participant ID:	☐☐ / ☐☐☐☐☐
Date of Session:	☐☐ / ☐☐ / ☐☐☐☐

We would like to know how many (if any) contacts you have had with the services listed below **during your current pregnancy.**

SECTION 1: HOSPITAL-BASED SERVICES

Please tell us the **total number** of appointments with and visits to the following **HOSPITAL-BASED SERVICES during your current pregnancy.** If none, enter zero.

HOSPITAL-BASED SERVICES	Number of appointments or visits
Total number of **hospital outpatient appointments** (note: please do not count hospital admissions)	
Total number of **nights spent in hospital as an inpatient** (note: if you had more than one admission, please add all nights)	
Total number of **hospital day admissions** (without overnight stay)	
Total number of **Accidents and Emergency (A & E) visits**	

SECTION 2: NON-HOSPITAL-BASED SERVICES

Please tell us the **total number** of appointments with and visits to the following **COMMUNITY-BASED SERVICES during your current pregnancy.** If none, enter zero. Please do not count hospital appointments which were included in the previous question.

COMMUNITY-BASED SERVICES	Number of appointments or visits
General practitioner (GP) at the surgery	
General practitioner (GP) at your home	
General practitioner (GP) on the telephone	
Nurse at GP surgery (e.g. practice nurse, district nurse, midwife etc.)	
Nurse at your home (e.g. practice nurse, district nurse, midwife etc.)	
Nurse on the telephone	
Community psychiatric nurse	
Psychiatrist in the community	
Counsellor	
Clinical Psychologist	
Occupational therapist	
Art, drama or music therapy sessions in the community	
Health visitor (do not include midwife)	
Social worker	

COMMUNITY-BASED SERVICES	Number of appointments or visits
Marriage counselling service, e.g. Relate	
Advice service, e.g. Citizen's Advice Bureau	
Family therapist	
Family support worker	
Helpline (e.g. Samaritans, MIND)	
Drug/alcohol support worker	
Other services – please write details here	

SECTION 3: MEDICATION

Have you taken any medication for mental health problems <u>during your current pregnancy</u>?

Yes ☐   No ☐

If YES, write the details below. Please include e.g. medications for depression, anxiety, psychosis, sleep problems and other mental health problems.

Name of Medication	Date started	Dose (mg)*	Number per day	Continuing at present?	Date Stopped
e.g. Fluoxetine/Prozac	*01/04/2013*	*20*	*1*	*Yes/no*	*10/09/2013*

*For current medication use please tell us current dose; for medication no longer taken please tell us final dose.

SECTION 4: EMPLOYMENT

4.1   **What is your current occupational status? Please tick the appropriate box.**
   Full-time employment (30+ hours per week) ☐
   Part-time employment (<30 hours per week) ☐
   Voluntary work ☐
   Unemployed & looking for work ☐
   Unemployed & not looking for work (e.g. housewife) ☐
   Unemployed & not able to work for medical reasons ☐
   Student ☐
   Retired ☐
   Other (specify)   …………………………………

4.2   <u>**If employed:**</u>
Please state your occupation.

Manager
Professional (e.g. health, teaching, legal)
Associate professional (e.g. technical, nursing)
Clerical worker/secretary
Services/sales (e.g. retail)
Skilled agricultural/fishery worker
Skilled labourer/craftsman (e.g. building, electrical etc.)
Elementary occupation (e.g. domestic, caretaker, labourer)
Armed Forces
Other (specify)   ……………………………………

4.3   **If you are employed, what is your gross pay per year (before tax) for your current or most recent employment?**

£ [   ] **per year**

4.4   **In a typical week, how many hours do you work?**   [   ]   **hours per week**

4.5   **Have you missed any <u>full days</u> of work <u>during your current pregnancy</u> because of your health? Please include only days missed for your <u>own</u> health.**

[   ] Yes   [   ] No

If YES, how many days did you miss from work <u>during your current pregnancy</u>? [   ] **days**

4.6   **Did you miss <u>part of a day's</u> work during your current pregnancy because of your health? Please include only days missed for your <u>own</u> health.**

[   ] Yes   [   ] No

If YES, how many days did this occur <u>during your current pregnancy</u>? [   ] **days**

4.7   **How many <u>weeks</u> were you unemployed during your current pregnancy?**   [   ]   **weeks**

Section 5: SUBJECTIVE WELL-BEING

Please, indicate with a cross <u>how happy you are at this moment on the scale below</u>.

Completely unhappy   Completely happy

0   1   2   3   4   5   6   7   8   9   10
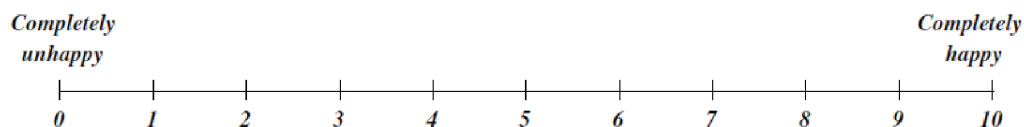
Thank you for taking the time to complete this questionnaire.

ECONOMIC PATIENT QUESTIONNAIRE

(Postnatal)

Participant ID:	☐☐ / ☐☐☐☐
Session:	3-4 months after birth / 12 months after birth (please circle)
Date of Session:	☐☐ / ☐☐ / ☐☐☐☐

We would like to know how many (if any) contacts you have had with the services listed below **since the last time you completed this questionnaire.**

SECTION 1: HOSPITAL-BASED SERVICES

Please tell us the **total number** of appointments with and visits to the following **HOSPITAL-BASED SERVICES since the last time you completed this questionnaire.** If none, enter zero. **Please report appointments, visits and admissions that were for you alone or mainly for you (please do NOT include appointments, visits or admissions that were mainly for your baby or your other children).**

HOSPITAL-BASED SERVICES	Number of appointments or visits
Total number of **hospital outpatient appointments** (note: please do not count hospital admissions)	
Total number of **nights spent in hospital as an inpatient** (note: if you had more than one admission, please add all nights)	
Total number of **hospital day admissions** (without overnight stay)	
Total number of **Accidents and Emergency (A & E) visits**	

SECTION 2: NON-HOSPITAL-BASED SERVICES

Please tell us the **total number** of appointments with and visits to the following **COMMUNITY-BASED SERVICES since the last time you completed this questionnaire.** If none, enter zero. Please do not count hospital appointments which were included in the previous question. **Please report appointments and visits that were for you alone or mainly for you (please do NOT include appointments and visits that were mainly for your baby or your other children).**

COMMUNITY-BASED SERVICES	Number of appointments or visits
General practitioner (GP) at the surgery	
General practitioner (GP) at your home	
General practitioner (GP) on the telephone	
Nurse at GP surgery (e.g. practice nurse, district nurse, midwife etc.)	
Nurse at your home (e.g. practice nurse, district nurse, midwife etc.)	
Nurse on the telephone	
Community psychiatric nurse	
Psychiatrist in the community	
Counsellor	

COMMUNITY-BASED SERVICES	Number of appointments or visits
Clinical Psychologist	
Occupational therapist	
Art, drama or music therapy sessions in the community	
Health visitor – include ALL visits to you and your baby. Do NOT include midwife visits.	
Social worker	
Marriage counselling service, e.g. Relate	
Advice service, e.g. Citizen's Advice Bureau	
Family therapist	
Family support worker	
Helpline (e.g. Samaritans, MIND)	
Drug/alcohol support worker	
Other services – please write here	

SECTION 3: MEDICATION

Have you been taking any medication for mental health problems <u>since the last time you completed this questionnaire</u>?

Yes ☐   No ☐

If YES, write the details below. Please include e.g. medications for depression, anxiety, psychosis, sleep problems and other mental health problems.

Name of Medication	Date started	Dose (mg)*	Number per day	Continuing at present?	Date Stopped
e.g. Fluoxetine/Prozac	*01/04/2013*	*20*	*1*	*Yes/no*	*10/09/2013*

*For current medication use please tell us current dose; for medication no longer taken please tell us final dose.

SECTION 4: SUBJECTIVE WELL-BEING

Please, indicate with a cross how happy you are at this moment on the scale below.

*Completely
unhappy*

*Completely
happy*

0　1　2　3　4　5　6　7　8　9　10
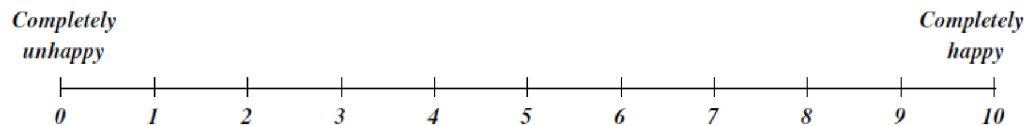
Thank you for taking the time to complete this questionnaire.

Identifying Depression in Pregnancy & Early Motherhood

Biographical Questionnaire

1. What is your highest educational qualification? (please tick one of the boxes)

No qualifications ☐

GCSEs/O-Levels ☐

AS/A-Levels ☐

NVQ or other vocational qualification ☐

Undergraduate degree ☐

Postgraduate degree ☐

Doctoral degree ☐

Professional degree (e.g. MD) ☐

2. Do you have children already? (please tick)

Yes ☐

No ☐

If yes, how many children do you have? (please write here) _____

We would like to ask you a few questions to find out whether you have ever suffered from anxiety or depression and whether you may have received help for

3. Have you ever suffered from anxiety? (please tick one)

a) Yes, I am currently suffering from anxiety ☐

b) Yes, I am currently suffering from anxiety and have suffered in the past ☐

c) Yes, I have suffered with anxiety in the past ☐

d) No ☐

e) Don't know ☐

f) No response ☐

3a. If yes to question 3, did you seek help for your anxiety?

a) Current anxiety   Yes ☐   No ☐   Don't know ☐

b) Previous anxiety   Yes ☐   No ☐   Don't know ☐

c) No response ☐

3b. If yes to question 3a, were you prescribed any medication to help with your anxiety?

a) Current anxiety   Yes ☐   No ☐   Don't know ☐

b) Previous anxiety   Yes ☐   No ☐   Don't know ☐

c) No response   ☐

3c. If yes to question 3a, have you ever seen anyone other than your GP for help with your anxiety?

a) Current anxiety   Yes ☐ No ☐   Don't know ☐

b) Previous anxiety   Yes ☐ No ☐   Don't know ☐

c) No response ☐

3d. If yes to question 3c, who did you see?

a) **Psychiatrist:** current anxiety ☐ previous anxiety ☐ don't know ☐

b) **Psychologist:** current anxiety ☐ previous anxiety ☐ don't know ☐

c) **Counsellor:** current anxiety ☐ previous anxiety ☐ don't know ☐

d) **Community psychiatric nurse:** current anxiety ☐ previous anxiety ☐

don't know ☐

e) **Social worker:** current anxiety ☐ previous anxiety ☐ don't know ☐

f) **Other statutory/voluntary agency:** current anxiety ☐ previous anxiety ☐

    don't know ☐

g) **Other:** current anxiety ☐ previous anxiety ☐ don't know ☐

h) **No response** ☐

4. Have you ever suffered with depression? (please tick one)

a) Yes, I am currently suffering with depression   ☐

b) Yes, I am currently suffering with depression and have suffered with depression
    in the past

            ☐

c) Yes, I have suffered with depression in the past   ☐

d) No   ☐

e) Don't know   ☐

4a. If yes to question 4, did you seek help for your depression?

a) Current depression Yes ☐   No ☐   Don't know ☐

b) Previous depression Yes ☐   No ☐   Don't know ☐

c) No response ☐

4b. If yes to question 4a, were you prescribed antidepressants?

a) Current depression   Yes ☐   No ☐   Don't know ☐

b) Previous depression   Yes ☐   No ☐   Don't know ☐

c) No response ☐

4c. If yes to question 4a, have you ever seen anyone other than your GP for help
with your depression?

a) Current depression   Yes ☐   No ☐   Don't know ☐

b) Previous depression Yes ☐   No ☐   Don't know ☐

c) No response ☐

4d. If yes to question 4c, who did you see?

a) **Psychiatrist:** current depression ☐ previous depression ☐   don't know ☐

b) **Psychologist:** current depression ☐ previous depression ☐   don't know ☐

c) **Counsellor:** current depression ☐ previous depression ☐   don't know ☐

d) **Community psychiatric nurse:** current depression ☐   previous depression ☐ don't know ☐

e) **Social worker:** current depression ☐ previous depression ☐ don't know ☐

f) **Other statutory/voluntary agency:** current depression ☐   previous depression ☐   don't know ☐

g) **Other:** current depression ☐ previous depression ☐   don't know ☐

h) **No response** ☐

Appendix 6 Topic guides for acceptability study

BaBY PaNDA

Topic Guide

Pregnant and Postnatal Women

1st interview

1. Background information
a) Number of other children
b) How she's been feeling during pregnancy/childbirth/new mum

2. Views of screening perinatal depression
a) Do you think it's a good idea for Hlth Profs to screen women for depression in pregnancy / postnatal care?
b) Who do you think should do the screening?
(1) Health professional (known GP/ midwife / health professional)
(2) Specialist in mental health independent of usual care

3. Experience of being routinely screened / referred by own health professional
a) Were you aware of being asked about whether you were feeling depressed or anxious during pregnancy by a health professional in any of your routine visits? (not BABY PaNDA study)

4. Experience of previous depression or anxiety
a) Have you suffered from depression or anxiety in the past?
i) Generally or perinatal?
ii) How long did it last?
iii) How severe was it?
iv) Was it diagnosed / treated?
v) Discuss in relation to findings of PaNDA assessment

5. Current experience of perinatal depression
a) Have you noticed a change of mood since you became pregnant?
b) Are you suffering from depression or anxiety at the moment?
c) If yes,
i) symptoms / severity
ii) whether been diagnosed
iii) care pathway

6. Acceptability of questions

a) How did you find the PaNDA assessment in terms of completing the various questionnaires?
b) Whooley questions
c) EPDS questions

State each question for b) and c) verbatim followed by probes below
i) Do you think this is an acceptable question to ask a pregnant woman?

ii) Can you tell me in your own words what that question means to you? / What goes through your mind as you hear and answer the question?

iii) What does the term x mean to you? / What did you understand by term x?

iv) How did you remember that? / What time period were you thinking about?

v) How sure of your answer are you? / How accurate would you say your answer is? Why?

vi) How easy or difficult was it to answer the question?

vii) How did you feel about answering the question?

7. Methods for screening questionnaires

a) Based on your experience today for answering the various questions, would you prefer to complete these questions on a written questionnaire or talk them through within a conversation?

b) Why?

8. Check if woman happy to have a similar interview after their next PaNDA assessment (3-4 mo or 12 mo) to see if any different views?

BaBY PaNDA

Topic Guide
Postnatal Women (3-4 months postnatal)
2nd interviews

1. Background/Ice breaker
a. How have you been feeling as a new Mum?

2. Current experience of post-natal depression/anxiety
a. Since your previous interview / assessment have you noticed a change of mood?
b. Are you suffering from depression or anxiety (or any other common mental health issue) at the moment?
c. If yes,
   i. symptoms / severity
   ii. whether been diagnosed
   iii. care pathway

3. Ongoing experience of being routinely screened/referred by health professional
a. Since your previous interview / assessment have you been aware of being asked about whether you were feeling depressed or anxious by a health professional in any of your routine visits (not BABY PaNDA study)
b. If yes,
   i. What happened?

4. Acceptability of questions

a. Whooley questions
b. EPDS questions

5. State each question verbatim followed by probes below

   i. Do you think this is an acceptable question to ask a pregnant woman?

   ii. Can you tell me in your own words what that question means to you? / What goes through your mind as you hear and answer the question?

   iii. What does the term x mean to you? / What did you understand by term x?

   iv. Why did you choose that answer?

    v.   How did you remember that? / What time period were you thinking about?

   vi.   How sure of your answer are you? / How accurate would you say your answer is? Why?

  vii.   How easy or difficult was it to answer the question?

 viii.   How did you feel about answering the question?

6. Acceptability of questions at different time points

a. Do you feel any different about being asked or answering these questions now you are a mum compared to before?

    i.   Why?(Familiar / Less stigma / thought about it from previous time)

Check if woman is happy to have a similar interview after their next (and final) PaNDA assessment.

BaBY PaNDA

Topic Guide
Postnatal Women (12 months postnatal)
3rd interviews

1. **Background/Ice breaker**
    a. How has your first year as a Mum been?

2. **Current experience of post-natal depression/anxiety**
    a. Since your previous interview / assessment (3-4 months) have you noticed a change of mood?
    b. Are you suffering from depression or anxiety (or any other common mental health issue) at the moment?
    c. If yes,
        i. symptoms / severity
        ii. whether been diagnosed
        iii. care pathway

3. **Ongoing experience of being routinely screened/referred by health professional**
    a. Since your previous interview / assessment (3-4 months) have you been aware of being asked about whether you were feeling depressed or anxious by a health professional in any of your routine visits (not BABY PaNDA study)
    b. If yes,
        i. What happened?

4. **Overall, how found answering the questions over time in the study (have these with you as a reminder)**
    a. Whooley questions
    b. EPDS questions

5. **Final reflections over time have been in the study**
    a. Post-natal depression/anxiety
    b. Routine screening/referral

6. **Anything else want to say?**

BaBY PaNDA

Topic Guide

Health Professionals

- 6 health visitors (HV)
- 6 midwives (MW)

(Recruit a mix by site and grade)

Purpose (from protocol): Explore their experience of delivering the depression screening tools in routine clinical practice and their training needs. This will be explored against descriptions of recommended and routine practice from health professionals in the respective site.

1. **Study information**
   a. ID number
   b. Study centre number
   c. Interviewer
   d. Date
   e. Go over study purpose and clarify anonymity/confidentiality and data reporting
   f. Background information of interviewee (years experience, job title)

2. **Background information on screening for perinatal depression**
   a. Can you provide some background information on how you screen for symptoms and causes of perinatal depression?
   b. How does screening for perinatal depression differ throughout pregnancy and motherhood?
   c. What are the impacts of screening and failing to screen/manage perinatal depression?
   d. How do you decide whether someone has perinatal depression?
   e. Can you explain the policy and practice for perinatal screening at this hospital?

3. **General views and experiences of screening, diagnosis and management of perinatal depression**
   a. When do you think is the most suitable time to diagnose for perinatal depression (pregnancy, post-natally at first health visit 6-8 weeks, or all of these) why is that?
   b. Does your management of perinatal depression differ depending on the severity of the case? why?
      i. How does the care pathway differ depending on severity of case? And why?

ii.   Do people get referred to GPs or specialists? Does this vary with case severity?

c.   Can you tell me how comfortable you are asking the screening questions? Why is that?

d.   What do you feel are the barriers to asking screening questions?

e.   (If participant does not raise barriers below discuss the following):

   i.   Do you find your ability or experience of asking screening questions differs or is affected when you first meet patients or by your relationship with patients?

   ii.   Do you need technical training to diagnose or refer or is screening enough to be referred?

      1.   If a patient says they need help is that triggering diagnosis and referrals?

   iii.   Do you feel that the personality / background of woman, particularly if have a history of mental health problems or there is a known lack of personal / family support affects your ability or experience of asking screening questions?

   iv.   Do you find that women need to be pre-warned about questions and if so why?

Concerns in referring to GPs and/or lack of resources for appropriate treatment and follow up

Experiences of delivering the study screening tools (take these tools with you ask the participant to look at them)

f.   Whooley questions

g.   EPDS questions

- How do you feel about using these tools? Why is that?
- What has your experience been of asking these questions?
- Are there any questions that you have found difficult/awkward or sensitive to ask, and if so which questions and why?
- Are there any questions you find difficult to understand? Which ones and why?
- Any question women appear to have found difficult/awkward/sensitive to answer? Which ones and why?
- Any questions the women appear to find difficult to understand? Which ones and why?

4. Potential impact of the screening tools on the care pathway

a.   How do you think the screening tools have impacted on the care pathway?

   i.   Have there been any issues surrounding treatment availability or waiting lists since using the screening tools?

   ii.   Do you have any ethical concerns regarding the screening tool?

   iii.   Do you think the tools have impacted on the continuity or appropriateness of care provided?

   iv.   What treatment is typically provided post-diagnosis?

    1. Are anti-depressant's routinely offered as first line of treatment?

   v. Has the screening tool impacted on the care pathway in any other ways?

  b. Do you feel the screening tools creates any increased demands on MW / HV (at initial visit, paperwork, at follow up) and why?

  c. Do you think the screening tools add any burden in terms of supporting women emotionally, for yourself and patients

   i. If yes why and how?

  d. No change – already implemented

5. Training needs

  a. Do you think there are any training needs to help staff screen for perinatal depression, or for staff's development of counselling and diagnostic skills?

  b. Are there any other training needs you can think of?

6. Change in organisational systems

  a. Are there any changes needed to current organisational systems for screening and management of perinatal depression? (If necessary prompt r.e. staff, referral processes, booking appointments)

  b. If so what types of changes are required? What do you feel is the most important change that is needed?

7. Is there anything else you would like to add?

8. Re-clarify anonymity and data reporting

Appendix 7 Qualitative study information for health professionals

**Identifying Depression in Pregnancy &
Early Motherhood: BABY PaNDA Study**

Health Professional Information Sheet

You are being invited to take part in the BABY PaNDA study, a sub-study within the BABY study investigating screening for depression in pregnancy and early motherhood. Before you decide, please take time to read the following information carefully. Contact details are on the back page if you would like any more information.

What is the purpose of BABY PaNDA?

We would like to find out more about depression during pregnancy and early motherhood, and the best ways of identifying it. Approximately 1 in every 10 women is thought to experience depression during this time. Unfortunately, it is often not identified with around half the women affected not getting the help and support they need.

The National Institute for Health and Clinical Excellence (NICE) have recommended that a brief two-question screening questionnaire should be used to identify women with depression during pregnancy and after birth. However, it is not yet known how well this brief questionnaire works compared to other longer screening questionnaires. BABY PaNDA aims to find out.

It is also important that the acceptability and feasibility of delivering screening tools in perinatal care from the perspective of health professionals is taken into account when assessing the effectiveness of the screening tools.

Why have I been invited?

We would like to interview 6 midwives and 6 health visitors about their views and experience of delivering screening tools for depression in routine perinatal care and implications for subsequent care.

What does BABY PaNDA involve for Health Professionals?

The study involves taking part in a single face-to-face semi-structured in-depth interview with a specially trained researcher. The interview is expected to last

approximately one hour and will take place at your work place or other convenient location during the next year. A researcher will telephone you before the interview to arrange a convenient time and location.

You will be asked some brief questions about your experience of screening for depression in routine perinatal care and any training needs you may have identified.

Do I have to take part in BABY PaNDA?

It is up to you to decide whether to not to take part. If you do decide to take part, you are free to withdraw from the study at any time without giving a reason.

Why should I help?

Your contribution to BABY PaNDA will be valuable in helping us understand whether the brief screening questionnaire and the EPDS are appropriate and feasible to implement from the perspective of health professionals. It will also help us explore the impact of screening tools on subsequent care pathways for women identified with depression.

Taking part in this study will involve your time for the interview. We do not anticipate any other disadvantages to you taking part in the study.

Will my information be kept confidential?

All information collected about you during the study will be treated confidentially and in accordance with the Data Protection Act. Any personal information which could identify you will be kept separately from your study information and will only be accessed by members of the research team.

We will destroy all personal information immediately after the end of the study. We will store your study information securely for 20 years and then destroy it.

If you decide to withdraw from the study, your personal information will be securely destroyed. We would still, however, like to use other information collected from you for the BABY PaNDA study.

What will happen to the results at the end of the study?

The study findings will be published in a research report and in articles for health professionals. We will also send you a summary of the study findings at the end of BABY PaNDA. You will not be personally identified in any publications from this study.

Who has reviewed the research?

The research has been reviewed and approved by the North East - York Research Ethics Committee.

Who is organising and funding the research?

This research is being done by researchers in the Department of Health Sciences at the University of York and doctors, midwives and health visitors involved in the care of women during pregnancy and after birth in the local NHS Trust. The research is being funded by the National Institute for Health Research Health Services and Delivery Research (NIHR HS&DR) Programme.

What should I do now?

If you would like to take part, please read and sign the consent form and send it back to us in the pre-paid envelope enclosed. A researcher will then contact you to arrange your interview. If you would like more information before deciding, please contact us using the details on the back page.

If you decide that you do not wish to take part, no further action is required.

What do I do if I want to withdraw or complain?

Every care will be taken during this study. If you wish to withdraw from BABY PaNDA or have a concern about any aspect of this study, please contact the study coordinator using the details below. If you do not want to contact the study coordinator you can contact the Chief Investigator, Professor Simon Gilbody (telephone number: [*insert contact number]* / email: [*insert email address]*).

Standard NHS indemnity arrangements apply to this research. While we anticipate no harm or distress to anyone as a result of this study, if you are harmed during this study due to someone's negligence, then you may have grounds for legal action, but you may have to pay for it. If you wish to complain formally, the normal NHS complaint mechanisms are available to you.

Further information and contact details

For further general information about research, please see the INVOLVE website at http://www.invo.org.uk

For specific information about BABY PaNDA, please contact:

Study Coordinator: [*insert name*]
                            [*insert current address*]
                            Telephone: [*insert contact number*]
                            Email: [*insert email address*]
                            Website: [*insert website address*]

Local Researcher(s)   [*insert researcher/s name*]
                            [*insert researcher contact details*]

Thank you for reading this information.

baby panda   THE UNIVERSITY *of York*   ☐☐ / ☐☐ ☐☐☐

Identifying Depression in Pregnancy & Early Motherhood: BABY PaNDA Study

Consent Form
Health Professionals

Thank you for reading the information about this study. If you would like to take part, please read and sign this form. Please put your **initials** in the boxes of the statements you agree with:

1. I have read the information sheet for Health Professionals and have been given a copy to keep. I have been able to ask questions about the study and I understand why the research is being done.   ☐

2. I understand that my participation in this research is entirely voluntary and that I will not receive any payment. I understand that I am free to withdraw my consent at any time without giving a reason.   ☐

3. I am willing to be interviewed by an experienced researcher about my experience of delivering screening tools for depression in routine perinatal care.   ☐

    a. I am happy for these interviews to be recorded and kept safe in an anonymised format.   ☐

    b. I understand that direct quotes from these interviews may be used in the publication of the study results but that these will be anonymised.   ☐

....
Name of health professional   **Date**   **Signature**
(BLOCK CAPITALS)

**PLEASE RETURN THE TOP COPY OF THIS FORM IN THE STAMPED-ADDRESSED ENVELOPE
PROVIDED. THE YELLOW COPY IS FOR YOU TO KEEP.**

Appendix 8 Interpretation of diagnostic test statistics

TABLE 35 Interpretation of diagnostic test statistics

Statistic	Description	Numerator	Denominator	Desirable
Sensitivity	Proportion of true depression cases correctly identified by the index tool	Correctly identified true depression cases by the index tool	All true depression cases	High values towards 100%
Specificity	Proportion of true non-depressed cases correctly identified by the index tool	Correctly identified non-depressed cases by the index tool	All true non-depressed cases	High values towards 100%
PPV	Proportion of positive (depressed) index tool outcomes that are correct (true depression cases)	Correctly identified true depression cases by index tool	All depression cases according to the index tool	High values towards 100%
NPV	Proportion of negative (non-depressed) index tool outcomes that are correct (true non-depressed cases)	Correctly identified non-depressed cases by the index tool	All non-depressed cases according to the index tool	High values towards 100%
PLR	Increase in likelihood of identifying a true depression case if the index tool outcome is positive (depressed)	Sensitivity	$1 - specificity$	Large values > 1 (moderate increase in likelihood > 5)
NLR	Increase in likelihood of identifying a true depression case if the index tool outcome is negative (non-depressed)	$1 - sensitivity$	Specificity	Small values towards 0 (moderate decrease in likelihood < 0.2)
AUC	Predictive accuracy of screening/case-finding tool, taking trade-off between sensitivity and specificity into account	n/a	n/a	High values towards 1

n/a, not applicable.

Appendix 9　Detailed breakdown of findings from quantitative acceptability survey

TABLE 36 Acceptability survey question: do you think it is a good idea to ask women about how you are feeling during early motherhood using the Whooley questions or the EPDS?

Participant response to question	Whooley questions, *n* (%)	EPDS, *n* (%)
Yes	327 (94.8)	334 (96.8)
No	5 (1.4)	2 (0.6)
Unsure	13 (3.8)	9 (2.6)
Not answered/missing	0 (0.0)	0 (0.0)

TABLE 37 Acceptability survey question: was it easy or difficult to understand the meaning of Whooley questions 1 and 2? (Prenatal assessment at 20 weeks)

Participant response to question	Whooley question 1,[a] *n* (%)	Whooley question 2,[b] *n* (%)
1: very easy	309 (79.0)	292 (74.7)
2: fairly easy	46 (11.8)	56 (14.3)
3: easy	25 (6.4)	34 (8.7)
4: fairly difficult	7 (1.8)	8 (2.0)
5: difficult	1 (0.3)	0 (0.0)
Not answered/missing	0 (0.0)	1 (0.3)

a Whooley question 1: during the past month, have you often been bothered by feeling down, depressed or hopeless? Yes/no.
b Whooley question 2: during the past month, have you often been bothered by having little interest or pleasure in doing things? Yes/no.

TABLE 38 Acceptability survey question: was it easy or difficult to understand the meaning of Whooley questions 1 and 2? (Postnatal assessment at 3–4 months)

Participant response to question	Whooley question 1, *n* (%)	Whooley question 2, *n* (%)
1: very easy	283 (82.0)	273 (79.1)
2: fairly easy	39 (11.3)	45 (13.0)
3: easy	15 (4.3)	23 (6.7)
4: fairly difficult	5 (1.4)	2 (0.6)
5: difficult	2 (0.6)	0 (0.0)
Not answered/missing	0 (0.0)	2 (0.6)

TABLE 39 Acceptability survey question: was it easy or difficult overall to understand the meaning of the 10 questions in the EPDS?

Participant response to question	Prenatal (20 weeks), n (%)	Postnatal (3–4 months), n (%)
1: very easy	311 (79.5)	282 (81.7)
2: fairly easy	55 (14.1)	49 (14.2)
3: easy	22 (5.6)	11 (3.2)
4: fairly difficult	2 (0.5)	1 (0.3)
5: difficult	1 (0.3)	0 (0.0)
Not answered/missing	0 (0.0)	2 (0.6)

TABLE 40 Acceptability survey question: were there any questions in the EPDS (from 1 to 10) which were particularly difficult to understand?

EPDS question number	Prenatal (20 weeks), n (%)	Postnatal (3–4 months), n (%)
1	10 (2.6)	7 (2.0)
2	3 (0.8)	3 (0.9)
3	6 (1.5)	7 (2.0)
4	6 (1.5)	1 (0.3)
5	7 (1.8)	1 (0.3)
6	5 (1.3)	1 (0.3)
7	4 (1.0)	5 (1.4)
8	2 (0.5)	0 (0.0)
9	2 (0.5)	2 (0.6)
10	0 (0.0)	0 (0.0)

TABLE 41 Acceptability survey question: how easy or difficult was it to remember how you were feeling to answer the Whooley questions 1 and 2? (Prenatal assessment at 20 weeks)

Participant response to question	Whooley question 1,[a] n (%)	Whooley question 2,[b] n (%)
1: very easy	212 (54.2)	241 (61.6)
2: fairly easy	124 (31.7)	94 (24.0)
3: easy	35 (9.0)	42 (10.7)
4: fairly difficult	18 (4.6)	12 (3.1)
5: difficult	1 (0.3)	2 (0.5)
Not answered/missing	1 (0.3)	0 (0.0)

a Prenatal acceptability survey: was it easy or difficult to remember whether you have often been feeling down, depressed or hopeless in the last month? (Question 4)
b Prenatal acceptability survey: was it easy or difficult to remember whether you have often been bothered by having little interest or pleasure in doing things in the past month? (Question 7)

TABLE 42 Acceptability survey question: ease of remembering how you were feeling to answer Whooley questions 1 and 2 (postnatal assessment at 3–4 months)

Participant response to question	Whooley question 1,[a] n (%)	Whooley question 2,[b] n (%)
1: very easy	197 (57.1)	216 (62.6)
2: fairly easy	106 (30.7)	91 (26.4)
3: easy	29 (8.4)	28 (8.1)
4: fairly difficult	12 (3.5)	9 (2.6%)
5: difficult	1 (0.3)	0 (0.0)
Not answered/missing	0 (0.0)	1 (0.3)

a Postnatal acceptability survey: was it easy or difficult to remember whether you have often been feeling down, depressed or hopeless in the last month? (Question 4)
b Postnatal acceptability survey: was it easy or difficult to remember whether you have often been bothered by having little interest or pleasure in doing things in the past month? (Question 7)

TABLE 43 Acceptability survey question: was it easy or difficult to remember how you have been feeling in the past 7 days to answer the 10 questions in the EPDS?

Participant response to question	Prenatal (20 week), n (%)	Postnatal (3–4 months), n (%)
1: very easy	281 (71.9)	248 (71.9)
2: fairly easy	82 (21.0)	71 (20.6)
3: easy	21 (5.4)	18 (5.2)
4: fairly difficult	4 (1.0)	3 (0.9)
5: difficult	0 (0.0)	0 (0.0)
Not answered/missing	3 (0.8)	3 (1.4)

TABLE 44 Acceptability survey question: were there any questions in the EPDS (from 1 to 10) that were particularly difficult to remember how you were feeling in the past 7 days?

EPDS question number	Prenatal (20 weeks), n (%)	Postnatal (3–4 months), n (%)
1	8 (2.0)	8 (2.3)
2	7 (1.8)	6 (1.7)
3	6 (1.5)	5 (1.4)
4	8 (2.0)	6 (1.7)
5	0 (0.0)	4 (1.2)
6	8 (2.0)	4 (1.2)
7	3 (0.8)	4 (1.2)
8	3 (0.8)	3 (0.9)
9	3 (0.8)	2 (0.6)
10	2 (0.5)	2 (0.6)

TABLE 45 Acceptability survey question: confidence in answers for Whooley screening/case-finding questions 1 and 2 (prenatal assessment at 20 weeks)

Participant response to question	Whooley question 1,[a] n (%)	Whooley question 2,[b] n (%)
1: very sure	272 (69.6)	277 (70.8)
2: fairly sure	83 (21.2)	73 (18.7)
3: sure	28 (7.2)	37 (9.5)
4: fairly unsure	6 (1.5)	4 (1.0)
5: very unsure	1 (0.3)	0 (0.0)
Not answered/missing	1 (0.3)	0 (0.0)

a Prenatal acceptability survey: how sure do you feel about your answer for question 1 about feeling down? (Question 5)
b Prenatal acceptability survey: how sure do you feel about your answer for question 2 about feeling little interest? (Question 8)

TABLE 46 Acceptability survey question: confidence in answers for Whooley screening/case-finding questions 1 and 2 (postnatal assessment at 3 to 4 months)

Participant response to question	Whooley question 1,[a] n (%)	Whooley question 2,[b] n (%)
1: very sure	244 (70.7)	259 (75.1)
2: fairly sure	75 (21.7)	63 (18.3)
3: sure	24 (7.0)	20 (5.8)
4: fairly unsure	2 (0.6)	2 (0.6)
5: very unsure	0 (0.0)	0 (0.0)
Not answered/missing	0 (0.0)	1 (0.3)

a Postnatal acceptability survey: how sure do you feel about your answer for question 1 about feeling down? (Question 5)
b Postnatal acceptability survey: how sure do you feel about your answer for question 2 about feeling little interest? (Question 8)

TABLE 47 Acceptability survey question: confidence in answers for Whooley screening/case-finding question 3 (postnatal assessment at 3–4 months)

Participant response to question[a]	Prenatal (20 weeks), n (%)	Postnatal (3–4 months), n (%)
Total asked	110	95
1: very sure	61 (55.5)	57 (60.0)
2: fairly sure	29 (26.4)	20 (21.1)
3: sure	12 (10.9)	10 (10.5)
4: fairly unsure	2 (1.8)	3 (3.2)
5: very unsure	1 (0.9)	1 (1.1)
Not answered/missing	5 (4.5)	4 (4.2)

a Whooley question 3: is this something you feel you need or want help with?

TABLE 48 Acceptability survey question: confidence in answers for the 10 questions for the EPDS for pregnant women

Participant response to question	Prenatal (20 weeks),[a] n (%)	Postnatal (3–4 months),[a] n (%)
1: very sure	272 (69.6)	254 (73.6)
2: fairly sure	96 (24.6)	74 (21.4)
3: sure	18 (4.6)	14 (4.1)
4: fairly unsure	0 (0.0)	1 (0.3)
5: very unsure	1 (0.3)	0 (0.0)
Not answered/missing	4 (1.0)	2 (0.6)

a Prenatal and postnatal acceptability survey: how sure do you feel about your answers for the 10 questions for the EPDS? (Question 17)

TABLE 49 Acceptability survey question: confidence in answers for individual questions within the EPDS

EPDS question number	Prenatal (20 weeks,[a] n (%)	Postnatal (3–4 months),[a] n (%)
1	6 (1.5)	10 (2.9)
2	8 (2.0)	8 (2.3)
3	14 (3.6)	6 (1.7)
4	12 (3.1)	3 (0.9)
5	5 (1.3)	2 (0.6)
6	9 (2.3)	5 (1.4)
7	7 (1.8)	3 (0.9)
8	8 (2.0)	5 (1.4)
9	8 (2.0)	7 (2.0)
10	4 (1.0)	4 (1.2)

a Prenatal and postnatal acceptability survey: were there any questions (from 1 to 10) which you were particularly unsure about your answer? (Please cross the box if you were unsure about your answer to that question.) (Question 18)

TABLE 50 Acceptability survey question: how comfortable did you feel when you were completing the Whooley questions and the EPDS during pregnancy?

Participant response to question	Prenatal (20 weeks), n (%)	
	Whooley questions	EPDS 10 questions
1: very comfortable	292 (74.7)	311 (79.5)
2: fairly comfortable	64 (16.4)	60 (15.3)
3: comfortable	26 (6.6)	18 (4.6)
4: fairly uncomfortable	9 (2.3)	2 (0.5)
5: uncomfortable	0 (0.0)	0 (0.0)
Not answered/missing	0 (0.0)	0 (0.0)

TABLE 51 Acceptability survey question: how comfortable did you feel when you were completing the Whooley questions and the EPDS during the postnatal period?

	Postnatal (3–4 months), *n* (%)	
Participant response to question	Whooley questions	EPDS 10 questions
1: very comfortable	262 (75.9)	276 (80.0)
2: fairly comfortable	61 (17.7)	53 (15.4)
3: comfortable	17 (4.9)	13 (3.8)
4: fairly uncomfortable	5 (1.4)	2 (0.6)
5: uncomfortable	0 (0.0)	0 (0.0)
Not answered/missing	0 (0)	1 (0.3)

TABLE 52 Whooley questions: rating scales from acceptability survey (questions 2–9) showing averages[a]

	Prenatal (20 weeks)				Postnatal (3–4 months)			
Question	*n*	Mean	Median	SD	*n*	Mean	Median	SD
Question 2: comfortable regarding tool	391	1.37	1	0.71	345	1.32	1	0.64
Question 3: easy to understand question 1	391	1.32	1	0.70	345	1.28	1	0.68
Question 4: easy to remember for question 1	390	1.65	1	0.85	345	1.59	1	0.81
Question 5: sure regarding answer for question 1	390	1.41	1	0.71	345	1.37	1	0.64
Question 6: easy to understand question 2	390	1.38	1	0.73	343	1.28	1	0.61
Question 7: easy to remember for question 2	391	1.57	1	0.84	344	1.51	1	0.76
Question 8: sure regarding answer for question 2	391	1.41	1	0.70	344	1.32	1	0.61
Question 9: sure about answer for question 3	105	1.60	1	0.84	91	1.58	1	0.90

a Lower ratings towards 1 are more favourable.

TABLE 53 Edinburgh Postnatal Depression Scale: rating scales from acceptability survey (questions 12–17) showing averages[a]

	Prenatal (20 weeks)				Postnatal (3–4 months)			
Survey question about EPDS in general	*n*	Mean	Median	SD	*n*	Mean	Median	SD
Question 12: comfortable	391	1.26	1	0.56	344	1.25	1	0.55
Question 13: easy to understand	391	1.28	1	0.61	343	1.22	1	0.50
Question 15: easy to remember	388	1.35	1	0.63	340	1.34	1	0.62
Question 17: sure about answers	387	1.35	1	0.59	343	1.31	1	0.56

a Lower ratings, towards 1, are more favourable.

Appendix 10 Examples of quotations

Theme 1: acceptability of screening/case-finding in the perinatal period

Taking depression seriously as it is an important issue which can affect anyone

. . . just to say it's nice to know that people are thinking seriously about this, 'cos I do think it's a serious issue.

Trust D 2919 (20 weeks antenatal)

I think it [screening] should be a mandatory thing for a certain length of time after pregnancy, after maybe six weeks . . . because you've got your baby blues straight after, the six week check, then that's it, I think maybe eight, between eight and ten weeks, doing it anyway and then the new mother would then feel, oh actually, yeah, let me think about this situation, it might pick up.

Trust A 1148 (3–4 months postnatal)

Well it's so different once you've had a baby, 'cos you could, like you could be all right now but once you've got a baby and you're on your own at home . . . you're not as much in control of your life as you were before.

Trust B 3171 (20 weeks antenatal)

. . . it's [screening is] really important 'cos I think you're a lot more mindful of your feelings 'cos sometimes they're so extreme, the way that you've been feeling and, and its feelings and emotions you've never felt before, or you'd never experienced, yeah, because having a child completely changes yer, so . . .

Trust B 3776 (3–4 months postnatal)

Good to identify women with depression and as early as possible

Yeah, I think so, yeah, I think so. I mean personally I, I feel like it's a good thing because then you wanna, you wanna catch anything earlier, really you wanna nip things in the bud before things develop too much, do you know what I mean?

Trust C 2273 (20 weeks antenatal)

Important to screen throughout perinatal period as things change or do not get picked up initially

. . . yeah, I would say it [screening] was, would be a really good thing, because you can, you don't know how you're going to react when you get a new baby, you know, everyone's different, and it might be that if someone copes completely fine during pregnancy and then suddenly everything changes. So I think it's worth doing sort of with everybody really. It's not just a certain type of person that could end up with postnatal depression. You don't know do you?

Trust A 2415 (20 weeks antenatal)

Theme 4: confidence in answers to the Whooley questions and Edinburgh Postnatal Depression Scale screening/case-finding questions

Discomfort for women to answer wording of Whooley question 1 honestly: illustrative quotations

. . . feeling down's sort of (pause) the basis really, isn't it?

Trust A 2699 (20 weeks antenatal and 12 months postnatal)

I understand the terms of down and depressed as being similar in meaning but useful as some may admit to feeling down but not depressed.

Trust A 3913 (20 weeks antenatal, 3–4 months postnatal and 12 months postnatal)

I resonate most strongly with 'feeling down'.

Trust C 1489 (3–4 months postnatal)

I feel the question is geared towards the severe end of depression.

Trust C 2273 (20 weeks antenatal)

. . . some days, you do, you just wanna cry and everything gets on top of you but you don't wanna be judged . . . I like the words 'feeling down' much better than, than depression.

Trust B 3776 (20 weeks antenatal)

I know I picked the one [answer] that doesn't sound so bad.

Trust B 1653 (3–4 months postnatal and 12 months postnatal)

If I'm honest now, I would, probably was feeling down at the time, but then I would have answered no.

Trust C 1489 (3–4 months postnatal)

. . . if you were feeling depressed, you might say, you know, in those [Whooley] questions, do you know what I mean, you might not want to say. It's hard, isn't it, how . . . or if, it says if you're feeling depressed, you might feel embarrassed to say yes, or I don't know, if someone really was feeling it, I don't know if they'd be truthful filling in a form.

Trust D 2919 (20 weeks antenatal)

It [honest answer] would depend on how the person feels on the day, do they have their front on today?

Trust A 2912 (20 weeks antenatal)

Reasons why women are less confident about their answer to the Whooley question 3 regarding help for depression

Lack of knowledge of what help is available

. . . but if it was yes, I don't know, I don't know what help I would be expecting. I'd sort of be thinking oh God, are the men in white coats going to come and take me away. I don't quite know, or maybe it even needs to be some choices.

Trust C 2273 (20 weeks antenatal)

Need to admit depression and need for help to self and others

I just feel that's quite a personal, big step in one question. I don't know, I just feel a bit, it makes me feel a bit uncomfortable. . . . If you're having these questions repeated, you kinda know what's coming; so maybe the next time you might think actually, you know what, I really do need some help.

Trust C 2273 (20 weeks antenatal)

. . . it's about being proud, I think, about want, the word like help and accepting help.

Trust B 1653 (3–4 months postnatal and 12 months postnatal)

Theme 5: comfort to answer the Whooley questions and Edinburgh Postnatal Depression Scale screening/case-finding questions

Social and cultural reasons causing discomfort for women to answer Whooley questions honestly

Social stigma associated with depression, especially for expectant and new mothers

I think it is good that it [perinatal screening for depression] is being looked at and being brought more to the public eye, not sort of, depression's a thing but that it is more normal, the many percentage of people have it and it's not something to be embarrassed or scared about.

Trust C 2288 (20 weeks antenatal)

. . . yeah, I think you wouldn't, I personally wouldn't wanna be sort of misunderstood or stigmatised or flag up something really bad from answering yes to that [Whooley question 1] if that makes sense?

Trust B 3776 (20 weeks antenatal and 3–4 months postnatal)

Fear of consequences of having depression

. . . when I saw a GP when I were pregnant and they suggested getting counselling. I think it was in the news that the woman had to have a forced caesarean and her child had been taken off her; so that [laughs] that really worried me and I thought oh no, I'd better not say that I'm feeling this way in case they take my baby off me.

Trust A 1851 (3–4 months postnatal and 12 months postnatal)

I think, I feel that the health visitors, this time round I have felt they've, I've, the underlying feeling has been they've come to check on me, whether I've been a good mum; whether that's because it's a bit of the postnatal depression or whether that's general, I don't know but . . . I think they should be, yeah, because that's what they're there for, and I suppose maybe explaining what their exact role is and why they're there might take some of that away.

Trust A 1148 (3–4 months postnatal and 12 months postnatal)

Especially at your booking-in appointment, because they ask you if your pregnancy's planned and how you're feeling about it as well, so you might feel quite happy about your pregnancy but then if you're feeling hopeless, they might think you don't want your baby. So I think there's maybe . . . it's extreme emotions that you, you'd probably think if you said yes [to Whooley question 1] then you'd be sorta whisked, whisked off to some sort of mental institution probably.

Trust A 2415 (20 weeks antenatal)

I didn't want people to think oh well, she don't love her baby, 'cos she's got postnatal depression.

Trust D 3486 (20 weeks, 3–4 months postnatal and 12 months postnatal)

Appendix 11 Women's experiences of the mental health-care pathway

Within the sample of 17 women who had experiences of the mental health-care pathway, three women (trust B 3629, trust A 1834, trust A 2943) who had experienced a historical case of depression and/or anxiety and were not experiencing a current episode had a self-reported positive experience of the care pathway for that episode. This included both a mild case and a severe case that required urgent psychiatric care. One of the women had a history of anxiety only (trust A 2943) and is now having additional checks for reassurance even though she does not have a currently identified episode of anxiety or depression.

Nine of the 17 women had a current case of depression identified either by their HP or as a result of the woman telling their HP as part of their routine care. All nine women had a history of depression with or without anxiety.

Two of the women reported having had a very positive care experience for their mental health problems (trust B 1653, trust C 2288). These experiences included one woman (trust B 1653) who was extremely resistant to any offer of help initially, particularly the antidepressants offered by her GP. The exemplary practice of her MWs and particularly her HV throughout the 12-month postnatal period resulted in the woman receiving regular additional home-based support and being personally accompanied to attend a mother and baby group for the first time. At the 12-month BaBY PaNDA study interview, the woman reported taking antidepressants and receiving support from a breastfeeding counsellor in addition to receiving ongoing visits from her HV. The other woman who reported a highly positive experience of her mental health-care pathway has been experiencing depression since age 11/12 years and had conceived her baby through IVF having tried for 5 years (trust C 2288). This woman has an excellent relationship with her female GP, who manages her symptoms very carefully with medication and monthly appointments. The woman also has a referral in place to give her direct access to community care services if and when she feels she needs them.

The remaining four women had a mixed experience of their treatment and management of their mental health problems (trust C 1489, trust A 1851, trust B 2912, trust A 2699). In all cases, the women reported the variation in experience in relation to different HPs and their personal approach, commitment and type of support they offer. Care from MWs and HVs appears to be preferred over the GP in many cases owing to the nature of the support, which tends to be focused more on talking, advice and links to support groups than on the use of antidepressants. Not all MWs and HVs are, in the views of some of the women, interested or supportive, however. In one case, it was a mental health worker who was the key HP to determining a positive and effective mental health-care pathway for one woman (trust A 1851). The woman noted the excellent relationship she had with this mental health worker as the reason why the counselling sessions were effective on this occasion.

Two of these four women had taken the decision to make an appointment and seek treatment or support for their symptoms at the request of their husband (trust B 2912, trust A 1851). One woman had a negative experience of counselling in the past and had stopped attending her sessions. She felt her emotional well-being was not being monitored as closely as she might need given her history of depression and anxiety, and the abnormal growth rate of her foetus resulting in a planned early induction:

> *I think it would be helpful to, you know, even if they weren't these exact questions, you know, just for midwives to kinda check in on how you're feeling. I mean they're probably a bit desensitised to it all because obviously they've heard everything possible, but I think, I think there's an emphasis on how the baby's doing which is, which yeah, it should be, but I think sometimes the mother's kind of forgotten a little bit so long as the baby's OK.*
>
> *Trust A 1851*

Both women ultimately felt the treatment was effective, although one woman believed it was the support from her friends while her husband was away that improved her symptoms, rather than the antidepressants from her GP (trust B 2912).

The BaBY PaNDA assessment identified one woman with a history of depression as having a current episode of depression and anxiety at both 20 weeks prenatally and 12 months postnatally (trust A 2699). This woman participated in three in-depth study interviews to discuss her care pathway at 20 weeks prenatal and 3–4 and 12 months postnatal. The woman reports that the HP screens her informally at every check given her mental health history, family history and natural tendency to be self-doubting and anxious. She has also experienced two miscarriages between her first child and this fourth pregnancy, birth and motherhood:

> *I think with my history of depression and with like my mum and my grandma, they did, every time I went for a check-up, they did ask, you know, 'How's your mood? Do, have you ever been feeling low?' . . . I'm not a very confident person as well, which doesn't help. I'm not one that can walk up to somebody and start a conversation sometimes, it depends on the people I suppose really.*
> Trust A 2699 (20 weeks prenatally, 3–4 postnatally and 12 months postnatally)

The woman started to open up to her MW over time as they built up a relationship during the prenatal period. The HV also asked her how she was feeling from her first contact just after the birth of her baby but, as time went by, the woman felt the appointments were rushed and quick checks of the baby's health only. This led her to say everything was fine. The woman also felt that her consultant discussed her coming off antidepressants too soon. Again, she felt the consultant was in a rush and she felt pressured not to get into a discussion:

> *Sometimes I think, with some doctors, you can just walk into the room and you can feel, like you can just completely open up to them, whereas some doctors you can't, and, which is why I always try to stick to the same doctor, but I mean obviously sometimes you can't 'cos they're quite popular or what have you, but, yeah, I do sort of, I believe, in a way, that if it'd have been sort of like me usual doctor, I maybe would have said, well you know, I've had days where I've been like, I can't be bothered to do anything typa thing.*
> Trust A 2699

Another case (as detailed below) illustrates failures in the screening/case-finding and care pathway process, which resulted in a lack of care provision for a woman who was admitting to self-harming (trust A 1148). This woman, who has a history of depression and anxiety, was identified as having a current episode of moderate depression in her BaBY PaNDA assessments at 20 weeks prenatal and 3–4 months postnatal. The woman has one child already and found the pregnancy difficult with the condition of hyperemesis, persistent severe vomiting, ultimately requiring a hospital admission. She was worried about the effect this would have on her baby as well as distress at leaving her older child while in hospital. Although the woman was pleased her MW was focused on the baby throughout the pregnancy, she felt there was no assessment of whether or not she needed additional support during the pregnancy. The birth of her second child was induced resulting in a birth by forceps with an epidural. The baby experienced health issues following the birth culminating in an operation at 7 weeks to remove a lump caused by a blocked gland. Although the HV did frequently screen the woman for depression using the EPDS case-finding instrument, the woman felt the HV was only interested in the baby and did not offer any options for support.

> *I mean the questions were asked. I can't dismiss that, but nothing was acted on, just doing this ten point check, OK you're on borderline, we'll do it again next time, that's basically all that happened. There wasn't anything there to say, right, well this is how it is and this is how you're feeling, this is how, what things we can do and things like that, there, there's just no . . .*
> Trust A 1148

The mother reported in her in-depth interview for the BaBY PaNDA study at 3–4 months postnatally that she was self-harming, having hit the wall and scraped her knuckles on purpose. She told the HV she was not coping but, in the view of the woman, the HV appeared to be only focused on the score of the repeated EPDS case-finding assessments.

> *I told her, and I said, 'you know, I, I don't think I'm coping very well.' And she said, 'Oh you're borderline, I can offer you help but I don't need to bring on any, anything extra.' What does that mean, I don't know. So it was, it was left at that.*
>
> *Trust A 1148*

The woman's GP also carried out an informal screen for depression at the 6-week check and offered her CBT. The long waiting list for CBT and the need to follow it up herself when feeling so down with the depression resulted in no progress for this core treatment option.

> *. . . when I was at the doctor's, it was the six week check, we did touch on it [depression] and he did mention this other form of not counselling but another method of working, but it was going to be a long drawn out affair, waiting weeks for the appointment so to be fair, I didn't follow it up and they never asked again and, I just felt, I did ask for help but it wasn't there, so what's the point? . . . Depression's horrible. In one sense you feel alone, particularly when, you know, you, you feel as though you're asking for help and you're not getting it. It's weird, just being in that situation you just, you don't know what to do and you just want to hide away in a little hole and just want to go away and leave me alone, but then you want the help. It's a merry go round.*
>
> *Trust A 1148*

A further five women were identified as having a current episode of depression by the CIS-R reference standard within one of the BaBY PaNDA study assessments. None of these women were referred by their GP onto the mental health-care pathway for further assessment or treatment as a result of the BaBY PaNDA assessment (their GPs would have been sent a letter from the BaBY PaNDA team advising them of their depression outcome).

Appendix 12 Case studies to illustrate the overlap of symptoms for depression and motherhood (theme 7)

A: trust C 2273 (20 weeks prenatal BaBY PaNDA in-depth interview)

This pregnant woman was experiencing symptoms of being extremely tired and exhausted in this pregnancy. She had experienced three miscarriages and has another child already. She was assessed as having mild depression and anxiety with the CIS-R (as the diagnostic reference standard) at her prenatal BaBY PaNDA assessment at 20 weeks. No risk of depression was identified with either the Whooley questions or EPDS screening/case-finding instruments at any time point. The woman's case of depression was not identified by the routine care of a HP, but support was provided by a mental health worker via the GP following the BaBY PaNDA assessment.

> Yeah, I just felt that by the end of, end of those questions, I thought hmm, that's gonna look like, you know, I felt like sort of saying its because I'm pregnant that I'm not sleeping and I normally sleep really well and I don't, I feel like it's hormonal rather than anything else. Basically I'm a wreck . . . It's definitely more challenging this time round.

B: trust B 1653 (3–4 months postnatal and 12 months postnatal BaBY PaNDA in-depth interviews)

One mother of a new baby in trust B experienced symptoms of low mood and crying a lot during pregnancy. She suffers from irritable bowel syndrome, which she reports is affected by stress. She has a history of anxiety and had a positive identification of mixed anxiety and depression at 3–4 months and 12 months postnatally with the CIS-R and the EPDS, but only at 12 months for the Whooley questions. No risk of depression or anxiety were identified at the prenatal (20-week) assessment. This woman was experiencing difficult life circumstances with her dad in hospital, her partner unable to work due to sickness and significant financial concerns. She reported having received a lot of support from MWs and HPs since pregnancy and started to feel better and started interacting again around 4–5 months postnatally. In this case, the symptoms of depression and anxiety were less likely to be actually caused by pregnancy or motherhood, despite the mother's claims that they were normal. Indeed, this woman delighted in the joys of motherhood during her 12-month postnatal interview and realised that her high expectations may have been part of the problem of her feeling down.

> But again that was just me being, it just felt nor [normal], like that was a normal thing, you know, just getting upset for no reason and just being silly [when pregnant] . . . I just feel like fine, I go to baby groups now, so it's just like I feel like, if you can say it, but like a proper mum, just like normal . . .

> . . . just having a baby, it's different, and having a family, you feel like a whole like complete, with partner and then dog and house, and everything just seems to be going really well, and it's lovely, like he's just started walking, he started walking yesterday, so it's just like the little milestones just like, oh it's just lovely . . .

C: trust B 3776 (20 weeks prenatal and 3–4 months postnatal BaBY PaNDA in-depth interviews)

One pregnant woman in trust B had two in-depth interviews about her views and experiences of screening/case-finding and completed all three BaBY PaNDA assessments throughout the perinatal period. Based on the CIS-R primary and secondary diagnoses, this woman was identified as having general anxiety disorder and mild mixed anxiety and depression during pregnancy and mild mixed anxiety and depression at 3–4 months postnatally. Both the Whooley questions and the EPDS had a positive identification of depression at both time points. The presence of depression or anxiety was no longer detected by any of the screening/case-finding instruments at 12 months postnatally. The BaBY PaNDA assessment reported the woman as having no history of depression or anxiety and it was her first pregnancy and first baby.

This woman's depression or anxiety was not detected by a HP during the perinatal period. A pivotal point in the self-treatment of her depression and anxiety was when she became aware of the problem following her BaBY PaNDA assessments, in-depth interviews and ultimately breaking down in a public place. Her newborn baby was admitted to the intensive care unit and she felt a lot of self-pressure to still breastfeed and be the successful mother she wanted to be. As her symptoms worsened, the new mother took the next proactive step of admitting her symptoms to, and seeking support from, her mother and husband. Together they used practical strategies such as yoga, relaxation and family holidays to help her through her the tiredness, depression and anxiety.

In summary, this woman correctly identified that her symptoms of depression were directly related to her pregnancy and experience of early motherhood, but was not identified as at risk in routine care despite being identified during her BaBY PaNDA assessments.

> *. . . all of the physical things you experience in pregnancy, like being tired and this, that and the other, which if I wasn't pregnant you think oh that's, that's a trigger for depression or something like that . . . it's just because your body's all over the place. I would say I'm normally quite steady [mood] and it's [mood] definitely quite changeable now [during pregnancy] . . . it's [low mood and worry is] made worse when you've got another life to care for and you've had no sleep.*

D: trust B 1296 (3–4 months postnatal BaBY PaNDA in-depth interview)

This new mother has one older child already and both were delivered as unplanned premature births. She has no history of depression or anxiety and no identification of a current risk by any assessment within BaBY PaNDA or routine care. This woman found it hard with her first child after the birth as the focus seemed to be on the premature baby with no support for the parent. She has gained in confidence from her first pregnancy and motherhood, however, and wanted to stick to her breastfeeding plan despite having another premature birth. This woman clearly experienced genuine difficulties related to the birth and managing the potentially conflicting expectations of parenting a premature baby. These did not result in the woman becoming depressed although she was clearly aware and possibly concerned that her anxieties about the difficulties she faced would be misunderstood.

> *. . . it needs to be realised that you, that you, you're all, you're allowed to be worried, that's a natural thing, natural part of being a parent, or a new parent, and that doesn't necessarily put you in a class of being depressed.*

The above cases studies illustrate that it may not be possible to unravel the complexity of perinatal, life circumstance and historical factors that may cause or exacerbate symptoms of depression for a pregnant woman or new mother. Despite this, women and HPs need the relevant knowledge, guidance and support to increase the likelihood of identifying new cases of women being at risk of depression in the perinatal period regardless of the cause. Women also need to be reassured by their HP that their symptoms can be

understood as 'normal' within the often challenging experience of being pregnant and/or a new mother and options such as jointly keeping a watchful eye on how the woman's symptoms progress will ensure they do not become problematic.

Health professionals raised concerns that the Whooley questions, the recommended screening/case-finding instrument in routine practice, do not have the ability to distinguish between mental health and, for example, general tiredness. This was identified as a barrier to correctly identifying depression by a number of HPs. To overcome these issues, HPs emphasised the importance of the assessing of 'other factors' (e.g. women's body language). One HP described how they emphasise the word 'often' to ascertain whether or not women are 'just having a bad day' or are at risk of depression. The experience, clinical judgement and confidence of HPs are therefore potentially important facilitators for the correct identification of new cases of women genuinely at risk of depression in the perinatal period.

> *If they [women] answer, you know, feeling tired, having little energy every day, straightaway that's three points, well then already you're on the path to a mild/moderate depression you know. So yeah, I think that's the trouble with, again but it is a screening tool and, you know, screening tools haven't got the skill of the person asking the questions.*
>
> *HV, trust D, 03*

Our findings demonstrate that women and HPs are both aware of this difficult issue. However, the interview data do not suggest there is a consensus or agreed best practice on how best to manage this difficult issue to avoid a lack of detection of new cases of depression during pregnancy and new motherhood.

Appendix 13 Strategies to improve screening/case-finding and mental health-care pathway as suggested by health professional interviewees

Training

The varied training received and the amount of time HPs spend dealing with mental health, led to calls for regular, mandatory and standardised training on both mental health and the screening/case-finding instruments. Refresher courses were also suggested, as a number of HPs reported receiving no formal mental health training since their undergraduate degree or since screening/case-finding instruments were introduced a number of years ago. Given the lack of general mental health training, some suggested that training on depression and anxiety as well as other mental health conditions would be useful. The mental health specialists also called for more training on the screening/case-finding instruments, specifically role play, how to ask screening/case-finding questions and the importance of their wording. They also emphasised the importance of training for student HPs or those who are perhaps uncomfortable using the screening/case-finding instruments, to ensure that they have the confidence to identify mental health problems.

> *I think we need to do more about the screening questions, perhaps role playing, that kinda thing.*
>
> *HV, trust D, 03*

Less frequently, HPs suggested that more training around referral is required. In particular, knowing when to refer, the referral pathways and what happens post-referral. In response to perceptions that GPs lack awareness of referral pathways, one HP suggested that GPs should receive training on referral pathways for perinatal mental health, NICE guidance and the reasons for HVs referring women. Although considered by one HV to be outside their role, a couple of HPs felt that they would benefit from CBT and counselling skills training, as they are often required to provide support to women without formal training. Finally, a minority of HPs discussed the need for better access to mental health and screening/case-finding guidelines and the need for staff to be given dedicated time to undertake training.

> *Definitely I think counselling, because often we do counselling and we don't even know we're doing it, you know, and we may, we, we [laughs] you know, it would be far better if we had more, you know, education on how to do it correctly, 'cos at the moment you just, you're kind of like anecdotally think oh what worked well and what didn't and work it that way. You get the basics about listening, active listening and, you know, basic communication skills, but to actually take that further forward; I mean like the CBT would be great.*
>
> *HV, trust A, 04*

Organisation of care

Recommendations for organisations aimed at improving screening/case-finding and identifying depression and anxiety were also proposed. Most commonly, HPs made recommendations for improving referral and women's experience of the care pathway such as having psychiatric nurses in clinics to assist HPs; simpler, clearer referral pathways; formal communication systems (information technology systems) between HPs and GPs; and giving MWs and HVs the ability to directly refer into services outside of primary care, such as IAPT. Four HPs from three trusts also called for more funding and resources and longer appointment times.

This was particularly prevalent for HPs in two trusts following closures to acute mental health services, lengthy treatment waiting times and the aforementioned staffing issues in one trust:

> *They definitely need to put in, more money into it, don't they, and more time into it?*
>
> *MW, trust B, 07*

Dedicated appointment for mental health assessment

- Creating a patient-centred environment for screening/case-finding and referral through the provision of a dedicated core appointment(s) in both the prenatal and postnatal periods (trust C).
- Ensuring that women see the same MW and HV (trusts A and C).
- MW to see women in their own home, to ensure that HPs are able to take the time to discuss and consider other factors to identify depression and anxiety (trust C).

Women support groups through independent sector

- Voluntary organisations for mild cases that are similar to those for breastfeeding, which are aimed at providing women with the opportunity to talk to other mothers who have had, or who may be experiencing, mental health problems (trust B).

Joined-up screening/case-finding with children's centres

- Health professionals to work directly with children's centres (trust B).

Screening/case-finding for fathers and partner

- Routine screening/case-finding for fathers and partners (trust D) as well as mothers.
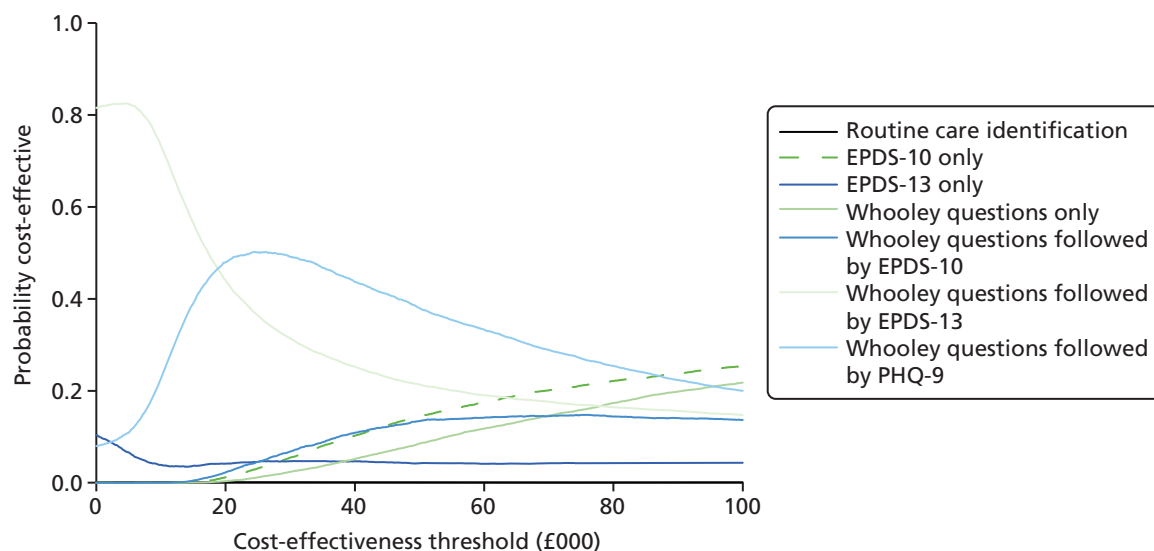
Appendix 14 Cost-effectiveness analysis

FIGURE 17 Cost-effectiveness acceptability curves for screening/case-finding strategies for prenatal model.
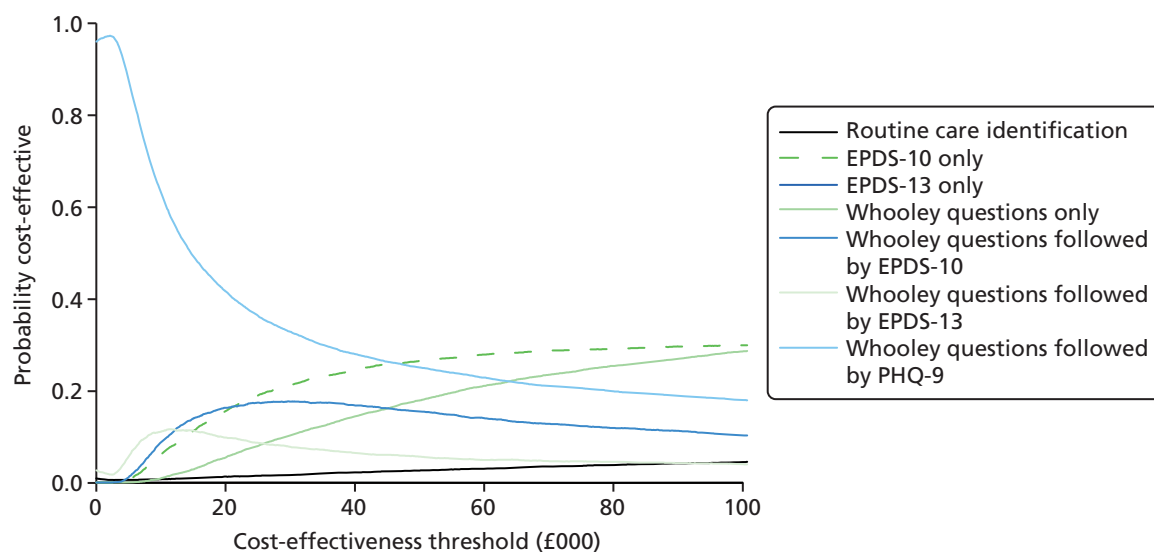

FIGURE 18 Cost-effectiveness acceptability curves for screening/case-finding strategies for postnatal model.
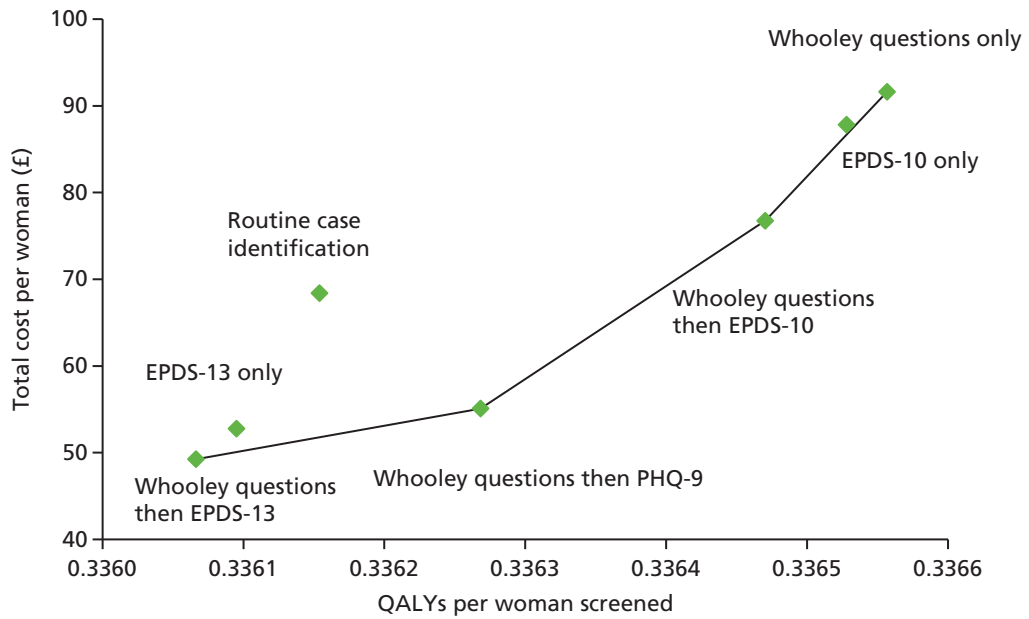
FIGURE 19 Cost-effectiveness efficiency frontier of screening/case-finding strategies for prenatal model.
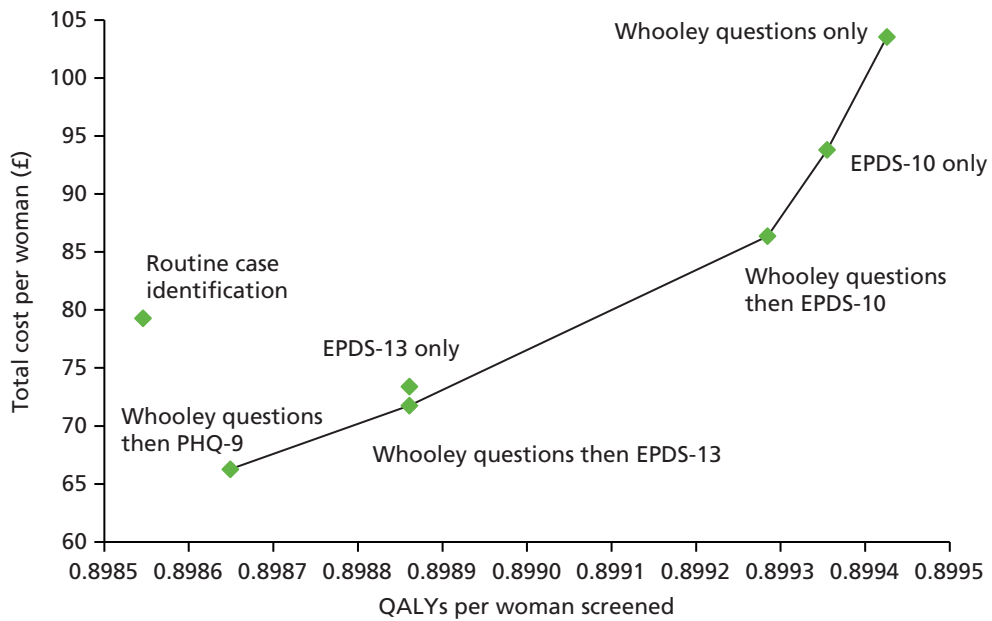

FIGURE 20 Cost-effectiveness efficiency frontier of screening/case-finding strategies for postnatal model.

TABLE 54 Net monetary benefit (QALYs × WTP threshold – costs) for all screening/case-finding strategies for perinatal depression

| Strategy | Net monetary benefit (mean estimates) (£) | | | |
| | Prenatal | | Postnatal | |
	WTP = £20,000	WTP = £30,000	WTP = £20,000	WTP = £30,000
Whooley questions followed by PHQ-9	6675.50	10,041.30	17,906.70	26,893.20
Whooley questions followed by EPDS (≥ 13-point cut-off point)	6674.90	10,037.70	17,905.50	26,894.10
EPDS (≥ 13-point cut-off point) only	6672.40	10,035.80	17,903.80	26,892.40
Routine care case identification	6659.00	10,023.30	17,891.60	26,877.10
Whooley questions followed by EPDS (≥ 10-point cut-off point)	6659.40	10,028.30	17,899.30	26,892.20
EPDS (≥ 10-point cut-off point) only	6649.40	10,019.30	17,893.30	26,886.90
Whooley questions only	6646.10	10,016.50	17,885.00	26,879.20

Appendix 15 National Screening Committee criteria

Taken from Hewitt et al.[76]

BOX 2 National Screening Committee criteria

The condition

1. The condition should be an important health problem.
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, latent period or early symptomatic stage.
3. All the cost-effective primary prevention interventions should have been implemented as far as practicable.
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.

The test

5. There should be a simple, safe, precise and validated screening test.
6. The distribution of test values in the target population should be known and a suitable cut-off level defined and agreed.
7. The test should be acceptable to the population.
8. There should be an agreed policy on the further diagnostic investigation of individuals with a positive test result and on the choices available to those individuals.
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.

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.
11. There should be agreed evidence-based policies covering which individuals should be offered treatment and the appropriate treatment to be offered.
12. Clinical management of the condition and patient outcomes should be optimised in all health-care providers prior to participation in a screening programme.

The screening programme

13. There should be evidence from high-quality randomised controlled trials that the screening programme is effective in reducing mortality or morbidity. When screening is aimed solely at providing information to allow the person being screened to make an 'informed choice' (e.g. Down's syndrome, cystic fibrosis carrier 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.
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.

BOX 2 National Screening Committee criteria (*continued*)

15. The benefit from the screening programme should outweigh the physical and psychological harm (caused by the test, diagnostic procedures and treatment).
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).
17. There should be a plan for managing and monitoring the screening programme and an agreed set of quality assurance standards.
18. Adequate staffing and facilities for testing, diagnosis, treatment and programme management should be available prior to the commencement of the screening programme.
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.
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 choice.
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.
22. If screening is for a mutation the programme should be acceptable to people identified as carriers and to family members.