

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 Woodhouse¹ and Simon Gilbody^{1,7*}

¹Mental Health and Addiction Research Group, Department of Health Sciences, University of York, York, UK

²York Trials Unit, Department of Health Sciences, University of York, York, UK

³Epidemiology and Cancer Statistics Group, Department of Health Sciences, University of York, York, UK

⁴Northern Lincolnshire and Goole NHS Foundation Trust, Scunthorpe General Hospital, Scunthorpe, UK

⁵Leeds and York Partnership NHS Foundation Trust, Bootham Park Hospital, York, UK

⁶Department of Health Sciences, University of York, York, UK

⁷Hull York Medical School, York, UK

⁸Patient and public involvement representative, York, UK

⁹Harrogate and District NHS Foundation Trust, Harrogate District Hospital, Harrogate, UK

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

Declared competing interests of authors: Simon Gilbody is a Health Technology Assessment (HTA) Evidence Synthesis Board member and a HTA Efficient Study Designs Board member. Catherine Hewitt is a HTA Commissioning Board member.

Published February 2018

DOI: 10.3310/hsdr06060

Scientific summary

BaBY PaNDA

Health Services and Delivery Research 2018; Vol. 6: No. 6

DOI: 10.3310/hsdr06060

NIHR Journals Library www.journalslibrary.nihr.ac.uk

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 \times WTP – cost) between the two most cost-effective strategies in both prenatal and postnatal periods was relatively small ($< \text{£}15$) 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

Funding for this study was provided by the Health Services and Delivery Research programme of the National Institute for Health Research.

Health Services and Delivery Research

ISSN 2050-4349 (Print)

ISSN 2050-4357 (Online)

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HS&DR archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hsdr. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

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

The research reported in this issue of the journal was funded by the HS&DR programme or one of its preceding programmes as project number 11/2004/23. The contractual start date was in April 2013. The final report began editorial review in August 2016 and was accepted for publication in December 2016. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HS&DR editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HS&DR programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HS&DR programme or the Department of Health.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Littlewood *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Health Services and Delivery Research Editor-in-Chief

Professor Jo Rycroft-Malone Professor of Health Services and Implementation Research, Bangor University, UK

NIHR Journals Library Editor-in-Chief

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the EME Programme, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA and EME Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andrée Le May Chair of NIHR Journals Library Editorial Group (HS&DR, PGfAR, PHR journals)

Dr Martin Ashton-Key Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Dr Peter Davidson Director of the NIHR Dissemination Centre, University of Southampton, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Dr Catriona McDaid Senior Research Fellow, York Trials Unit, Department of Health Sciences, University of York, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Wellbeing Research, University of Winchester, UK

Professor John Norrie Chair in Medical Statistics, University of Edinburgh, UK

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Professor Martin Underwood Director, Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of members of the NIHR Journals Library Board:
www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: journals.library@nihr.ac.uk