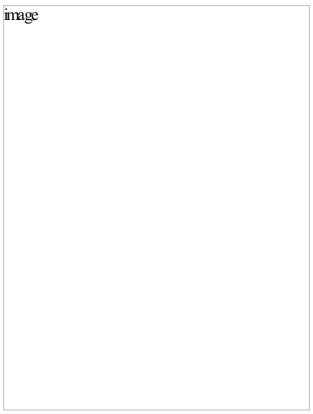


Postpartum Depression Screening Tools: A Review

Nneamaka Ukatu, Camille A. Clare, Maria Brulja

PII: S0033-3182(17)30225-6
DOI: https://doi.org/10.1016/j.psym.2017.11.005
Reference: PSYM836

To appear in: *Psychosomatics*

Cite this article as: Nneamaka Ukatu, Camille A. Clare and Maria Brulja, Postpartum Depression Screening Tools: A Review, *Psychosomatics,* doi:10.1016/j.psym.2017.11.005

Postpartum Depression Screening Tools: A Review

Nneamaka Ukatu B.A.,[a] Camille A. Clare, M.D., M.P.H.,[b] Maria Brulja D.O.[b]

[a]School of Medicine, New York Medical College Valhalla, NY USA

[b]Department of Obstetrics and Gynecology, New York Medical College Valhalla, NY USA

Corresponding author

Camille A. Clare, M.D., M.P.H.

New York City Health + Hospitals/Metropolitan

1901 First Avenue

Room 4B5

New York, NY 10029

Phone: 212-423-6796

Fax: 212-423-8121

Email: Camille_clare@nymc.edu

Keywords: postpartum screening, antenatal screening, maternal depression, postpartum depression

ABSTRACT

Purpose: The purpose of this study is to analyze the accuracy of screening tools in detecting postpartum depression (PPD).

Basic Procedures: A review of the literature was conducted using PubMed, Clinical Key and Google Scholar from the years 2001 to 2016 with a modified PRISMA method. The keywords, "postnatal depression screening," "antenatal depression screening," and "maternal depression" were used in the search. Sixty-eight articles were reviewed, and thirty-six further analyzed.

Main Findings: The accuracy of screening tools was dependent upon a number of factors. The studies reviewed differed in the types of screening tools tested; the combination of screening tools administered; the timing in which screening tools were administered; the geographic location of patients screened; and the reference standard(s) used.

Principal conclusions: No tool could be deemed best at accurately detecting PPD on the basis of sensitivity and specificity. Additionally, there was no recommended time period in which screening should be done. Thus, further research is needed to elucidate the accuracy of PPD screening tools, and the best criteria to determine this.

Keywords: postpartum screening, antenatal screening, maternal depression, postpartum depression

CURRENT KNOWLEDGE ON THE SUBJECT

- Postpartum depression is an under-recognized phenomena inflicting mothers.
- Several screening tools have been created and many have been tested among different patient populations.

WHAT THIS STUDY ADDS

- We provide a comprehensive analysis of many factors affecting the administration of screening tools.
- We also examine how the application of this analysis of screening tools for postpartum depression may be utilized for different patient populations.

INTRODUCTION

Postpartum depression (PPD) is characterized by depressive episodes occurring in the period after childbirth. It is estimated that the disease occurs in up to 20% of all women [1]. Although the exact causes of postpartum depression remain unknown, several risk factors have been identified. Women who have history of psychiatric illness [2], with limited partner support,

in abusive relationships and a history of substance abuse are at increased risk of [3]. Other special groups of patients at risk include adolescent mothers, immigrant women, those with low socioeconomic status, racial and ethnic minorities, and fathers [4].

Postpartum depression can cause grave consequences for both the mother and child. Women who suffer from postpartum depression have been found to have decreased maternal/neonatal bonding. Neonates born to women with PPD were more likely to be in the foster care system [5]. Furthermore, the children of depressed mothers have higher rates of mood disorders, and overall decreased general levels of functioning when compared to children born to non-depressed mothers. Additionally, women experiencing depression have poorer health outcomes and lower quality of life than non-depressed women [6].

Despite the severe consequences that PPD has on both the mother and the child, up to 50% of these cases will go undiagnosed [7]. This illustrates the need for effective screening methods to ensure that all women with PPD will be identified. There are several screening tools that have been developed to diagnose PPD. Those specific to detect maternal depression in the peripartum or postpartum period include the Edinburg Postpartum Depression Scale (EPDS), the Postpartum Depression Screening Scale (PDSS) and the Pregnancy Risk Questionnaire (PRQ). General depression screening tools have also been used to screen for PPD in new mothers. They include the Beck Depression Inventory-II (BDI-II), the General Health Questionnaire-12 (GHQ-12), the Center for Epidemiological Studies Depression (CES-D), and the Patient Health Questionnaire versions 2, 8 and 9. Additionally, these tools have been the most validated, and are commonly used in detecting PPD.

There has not been a consensus among the medical community regarding which tool is most accurate for screening for PPD. An accurate screening tool is one that is able to distinguish between healthy and unhealthy patients [8]. Sensitivity is the ability of the tool to correctly identify women who are at risk of postpartum depression. Whereas, specificity is the ability of the tool to correctly identify women who are not at risk of postpartum depression. As such, there is no universal policy in place for when and how to screen women for postpartum depression. Thus, we conducted a review of the literature to examine the accuracy of the listed screening tools, and to determine which special considerations are needed to evaluate women for PPD.

METHODS

A comprehensive review of the literature was performed using PubMed, Clinical Key and Google Scholar from the years 2001 to 2016 to reflect the most up to date literature. The keywords, "postnatal depression screening," "antenatal depression screening," and "maternal depression" were utilized in the search. A modified PRISMA method was used [9]. The accuracy of current postnatal depression screening tools was reviewed. Specific screening tools were analyzed, and an assessment of the current literature concerning the methods of postpartum depression screening was performed.

A total of 140 articles were identified from the literature search. After the removal of duplicates, 119 articles remained for review. Articles were eliminated if they did not focus on the analysis of screening tools. Additionally, only papers that were written in English were selected. A total of sixty-eight articles were left for further screening. Further elimination of articles was done if the analysis was not primarily based on the sensitivity and specificity of the tools. From

the sixty-eight articles left for review, thirty-six were analyzed further. Figure 1 illustrates the modified PRISMA format which was utilized.

From the thirty-six items analyzed, sixteen articles were validated surveys using psychiatric diagnostic interviews. Six articles used surveys without psychiatric interviews, of which one article was a retrospective review. Two articles were randomized control trials. There were twelve review articles; two of which were retrospective reviews, while the remaining ten were systematic reviews. The focus of our analysis was limited to tools that have been widely validated, or greatly used among clinicians.

RESULTS

The screening tools tested included the Postpartum Depression Screening Scale (PDSS), Pregnancy Risk Questionnaire (PRQ), Beck Depression Inventory-II (BDI-II), Edinburg Postpartum Depression Scale (EPDS), General Health Questionnaire-12 (GHQ-12), Center for Epidemiological Studies Depression Scale (CES-D) and the Patient Health Questionnaire versions 2, 8, and 9. Of these, 37.5% tested only one screening tool, while 62.5% tested a combination of the listed tools. The combination of screening tools was provided to patients in a step-wise matter, or was included in a packet containing multiple questionnaires. Table 1 outlines the summary characteristics of the studies that tested screening tool accuracy.

Overall, the *sensitivity* and the *specificity* of the screening tools were determined. The PHQ-2 had a reported sensitivity range of 62% [10] to 100% [11]. It had the highest sensitivity of all the screening tools. The Edinburgh Postpartum Depression Scale was the most widely tested screening tool (70.8% of studies). The reported sensitivity values for the EPDS ranged from 22.2% [11] to 96% [12]. The EPDS had the lowest reported sensitivity of all the screening tools

analyzed. Other tools had similar variations in their sensitivity. The Beck Depression Inventory II had a reported specificity range of 45.3% [13] to 100% [7]. Of all the tools, it showed both the highest and lowest specificity. Other tools had less variation in the reported specificity. Figure 2 outlines the sensitivities and specificities of the screening tools.

There were notable differences intrinsic to the screening tools tested. The *categories* of questions included in the tools varied. Certain tools, such as PHQ-2 and EPDS, limited questions to feelings of sadness or anxiety. Other screening tools, such as BDI-II and PHQ-9, screened for physical symptoms, such as fatigue, energy loss and sleep changes.

The screening tools also differed in the *reference period* examined and the types of questions included. The PRAMS-6 questionnaire allowed patients to report symptoms felt since delivering, while PHQ-9 limited the timeframe to two weeks prior to giving birth [14]. Two studies noted that the examination of symptoms over the entire postpartum period may be important for the recognition and better identification of PPD [14, 15].

The studies analyzed varied in the *methodology* utilized to administer screening tools to study subjects. Twenty-four of the studies tested the validity of the screening tools directly. These studies differed in the types of screening tools tested; the combination of screening tools administered; the timing in which screening tools were administered; and the reference standard(s) used. The geographic locations of the patients selected also varied. Table 1 outlines the summary characteristics of the twenty-four studies that tested screening tool accuracy. The remaining twelve studies consisted of either systematic or retrospective reviews. They also displayed similar variability in methodology, which made a comparison across the studies difficult.

The *timing* of screening was different among the studies analyzed. Four studies screened patients during the antenatal period [10, 16-18]. This was done at the initial visit or at some point during the first, second, or third trimester. Eight studies screened patients during the antenatal and postnatal periods. Of these, four studies conducted the initial screening during the antenatal period, and followed up with the same patients after childbirth [13, 19-21]. The other four screened both antenatal and postnatal patients simultaneously [22-25]. Lastly, twelve studies screened patients only during the postpartum period [7, 11, 12, 18, 26-33]. The exact timing of screening varied from immediately after childbirth to fourteen months postpartum. However, 75% of the studies screening PPD in the postpartum period did so before six months postpartum.

Patients screened for PPD were from *diverse backgrounds*. Geographically, in twelve studies, patients lived in the United States. The remaining studies screened patients located in Canada, France, Vietnam, Spain, Turkey, China, Australia, the Netherlands and the United Kingdom. Patients also differed in their socioeconomic status; educational level; income; age; marital status; and gender. Certain studies focused on patients representing only one of these demographics. One study sought to validate current screening tools among adolescent women [28]. Another study focused on the accuracy of screening men in the postpartum period [29]. One study screened only low income African American patients [22]. Lastly, four review articles focused on the validity of screening across socioeconomic groups, immigrant women and those of diverse cultural backgrounds, respectively [34-37].

Other disparities in the methodology may account for the reported sensitivities and specificities of the tools. Most studies utilized a *reference standard* to validate the results of the screening tools. The most commonly used standard was the Structured Clinical Interview for DSM-IV (SCID), which was employed in 54.16% of studies. The World Health Organization

Composite International Diagnostic Interview (CIDI) was another standard. It validated the tested tools in 16.6% of the studies. Two studies used the Mini-International Neuropsychiatric Interview (M.I.N.I.) in conjunction with other reference standards [18, 33]. Another study used the EPDS to validate the PHQ-2 [11]. Five studies did not use any reference standard for validation [12, 13, 24, 25 27].

Another methodological distinction was the use of *different combination of tools*. Pairing tools in the studies may have altered the accuracy of PPD detection. Certain studies provided subjects with combinations of screening tools, while other studies did not. One study noted that the combined use of EPDS and PDSS increased the validity in the antenatal period [18]. Thus, a comparison of the screening methods was difficult to assess.

Lastly, differences in *scoring methods* has also been reported in the literature. The PHQ-9 has been widely quoted to have a sensitivity and specificity of 88% on the basis of a validation study on 6000 non-pregnant patients [23]. However, all of the subsequent postpartum studies reviewed showed a sensitivity that was consistently lower. The sensitivity of the tool differed on the basis of how the question items were scored. One study compared simple scoring to a more complex scoring method using PHQ-9 [26]. The complex scoring method was based on DSM-IV diagnosis criteria for major depressive disorder. The simple scoring method yielded a sensitivity of 82%, while complex scoring resulted in a measured sensitivity of 67% [26]. One study used a dichotomous yes/no scoring method on the PHQ-2, and achieved 100% sensitivity [11]. Scoring the PHQ-2 using a Likert scale resulted in a sensitivity of 62% [10]. It has been noted that the use of dichotomous scoring may be better to avoid educational distinction, as Likert scoring may be biased towards more highly educated patients [11].

In addition to the scoring method, differences in the *cut-off scores* also resulted in disparate findings, as highlighted in Table 2. Before comparing different screening tools to discern the best one at detecting PPD, more validation studies may be needed to determine the optimum cut-off score for individual tools. Furthermore, large population-based studies can address whether certain cut-off scores must be applied to different populations [38], or to varying time periods, that is, the antenatal versus the postnatal periods.

DISCUSSION

The diagnostic performance of the available postpartum depression screening tools fluctuates depending on many factors. This discrepancy may be due to differences in the methodology of the studies such as the screening tools used; the reference standards; or a combination of tools, and the cut off scores. Results may also be influenced by the time period analyzed by the tools; the timing of screening; and the patient populations studied. These disparities in published findings make it difficult to properly compare the available tools. Consequently, no recommendation can be made about the most effective tool for detecting PPD.

Given the increased prevalence rates of PPD among low socioeconomic patients, obese patients, adolescent mothers, racial and ethnic minorities and immigrant women [4], it is important to expand research efforts to these special populations. The accuracy of screening methods must be analyzed within the context of this diversity. Additionally, clarification is needed regarding current definitions of PPD [34], and how this may change between different patient populations.

Research has shown that the initial symptoms of PPD may not be sadness, but rather insomnia, anxiety, irritability and confusion [39]. Inclusion of these items in a screening tool

may allow it to be more sensitive in recognizing PPD; however, it may not allow for the proper differentiation between the normal physical changes associated with the postpartum period, and abnormal symptoms [36]. Thus, the revision of PPD specific tools, like EPDS, to include such items may be beneficial. This may allow for sensitivity in detection while appropriately discerning normal and abnormal symptoms.

Overall, the timing of screening varied from the initial prenatal visit to fourteen months postpartum. With such a wide span of time of almost two years, there may be physical and psychological variety in the experiences of the women tested. Although there is no PPD diagnosis in DSM-IV or V, there is a modifier to major depressive disorder. Thus, screening during the first four weeks postpartum is in accordance with this modifier. However, screening during the first two weeks of this period may result in more false positives, as it may fail to differentiate the presence of "baby blues." It has been noted that use of screening tools in the first postpartum month may result in lower sensitivity and specificity versus later months [6]. Moreover, screening immediately postpartum may miss patients with a slower onset of PPD [40].

Diversity in the patient populations screened may alter the accuracy of the tool used. Patients of low socioeconomic status with a significant burden of stressful events may display altered scores when screened [34]. Additionally, the manifestation of PPD symptoms may differ along a cultural spectrum. Depressive symptoms in non-Western cultures lean towards somatization, while Western cultures report feelings of sadness [36]. Thus, diagnostic criteria based heavily on Western notions of PPD may not be appropriate transculturally [36]. A retrospective review of PPD screening in immigrant women revealed that the number of women identified as at risk was lower than estimated prevalence levels in this subgroup of patients [35].

This could be due, in part, to the creation of tools to match Western cultural practices and understanding of childbirth.

Furthermore, the use of EPDS among adolescent mothers demonstrated that standard cut-off scores may not be adequate in this patient population, and that optimum cut-off scores may be one to four points lower [28]. One retrospective review found that PPD symptoms in adolescent mothers were more influenced by prior depression and social support compared to adult mothers [15]. Thus, current predictive models of PPD could not be accurately applied to this patient population [15]. Lastly, while the use of EPDS among men in the postpartum period has been indicated to be effective [29], the sample size tested was too small to be generalizable.

The strength of this review is the comprehensive analysis of the various factors presented, all of which can alter the accuracy of screening tools. To date, no other study has done a widespread examination of all these factors. This can help guide clinicians on the best screening tools to use for patients meeting certain criteria. This can also guide further research and evaluation of the role that each factor contributes in altering screening tool accuracy.

Limitations of this study include the selection bias of the papers that were chosen for review. Articles were selected on the basis of being written in English and within a defined time period. Articles that did not fit these two criteria may contain more information that had been analyzed in this manuscript. Expansion of the period of study starting in 1980 or at the origins of the field of perinatal psychiatry might also reduce the selection bias noted, and to possibly allow for improved sensitivity and specificity of the screening tools analyzed. A study of the EPDS from 1987 to 2008 similarly acknowledged a heterogeneity of findings because of differences in methodologic approaches, language, and diagnostic criteria used [41]. Another limitation was the

decision to limit analysis to only ten screening tools. Further evaluation of other screening tools may yield more information. Lastly, the use of different methodologies employed by the studies reviewed was another limitation. This made reviewing the articles in a standardized matter difficult, and thus, may have altered the way individual articles were reviewed.

In conclusion, no screening tool is best at accurately detecting PPD. There is also no recommended time period in which to screen patients. In spite of this, the American College of Obstetricians and Gynecologists' Committee Opinion on screening for perinatal depression recommends the utility of screening tools that are shorter in length and that take less time to complete, such as the Edinburgh Postnatal Depression Scale and the Patient Health Questionnaire-9 [42]. Clinicians might consider this when choosing a screening tool that best fits into the scope of their practices. A consideration for the constitutional symptoms of depression will also reduce the specificity of the PHQ-9, as well as the Beck Depression Inventory and Center for Epidemiologic Studies Depression Scale [42]. Due to the fact that there has been no difference in clinical outcomes with the use of various screening tools, what becomes more important is not only using these screening modalities, but also having a collaborative approach to patients, who have access to needed resources and follow-up with psychiatric care providers.

Further research must be done to assess the optimum cut off score of individual tools; the best scoring method; the best time to screen; and the best combination of tools. By continuously improving our understanding of PPD, and the psychosocial context in which it occurs, we may adequately create methods that allow for the effective identification of at-risk women.

ACKNOWLEDMENTS

There are no acknowledgments to report.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

DECLARATION OF INTEREST

The authors report no conflicts of interest.

REFERENCES

1. Vieira T. When Joy Becomes Grief. Nursing for Women's Health. 2002 [accessed 2016 Jul 1]; 6(6):506–513.

2. Miller LJ. Postpartum depression. JAMA. 2002 [accessed 2016 Jul 1]; 287(6):762–765.

3. Bobo W, Yawn B. Concise Review for Physicians and Other Clinicians: Postpartum Depression. Mayo Clinic proceedings. 2014 [accessed 2016 Jul 1]; 89(6):835–844.

4. Clare C, Yeh J. Postpartum Depression in Special Populations. Obstet & Gynecol Surv. 2012; 67(5):313-323.

5. Brulja, M, Clare, CA. Patient Evaluation for the Development of Postpartum Depression, Steiner Young Investigator Award, The North American Society for Psychosocial Obstetrics and Gynecology, Poster Presentation, New York Academy of Medicine, New York, New York. April 9, 2016.

6. Owora A, Carabin H, Reese J, Garwe T. Diagnostic performance of major depression disorder case-finding instruments used among mothers of young children in the United States: A systematic review. Journal of Affect Disorders. 2016 [accessed 2016 Sep 25]; 201:185–193.

7. Beck C, Gable R. Comparative Analysis of the Performance of the Postpartum Depression Screening Scale with Two Other Depression Instruments. Nursing Research. 2001 [accessed 2016 Jun 21]; 50(4):242–250.

8. Baratloo A, Hosseini M, Negida A, Ashal G. Part 1: Simple Definition and Calculation of Accuracy, Sensitivity and Specificity. Emergency. 2015 [accessed 2016 Jun 20]; 3(2):48–49.

9. PRISMA. [Accessed 2017 Jul 12]. http://www.prisma statement.org/PRISMAStatement/CitingAndUsingPRISMA.aspx

10. Smith M, Gotman N, Lin H, Yonkers K. Do the PHQ-8 and the PHQ-2 accurately screen for depressive disorders in a sample of pregnant women? General Hospital Psychiatry. 2010 [accessed 2016 Jun 20]; 32(5):544–548.

11. Chae S, Chae M, Tyndall A, Ramirez M, Winter R. Can We Effectively Use the Two-Item PHQ-2 to Screen for Postpartum depression? Family Medicine. 2012 [accessed 2016 Jun 20]; 44(10):698–703.

12. Yawn B, Pace W, Wollan P, Bertram S, Kurland M, Graham D et al. Concordance of Edinburgh Postnatal Depression Scale (EPDS) and Patient Health Questionnaire (PHQ-9) to Assess Increased Risk of Depression among Postpartum Women. Journal of the American Board of Family Medicine. 2009 [accessed 2016 Jun 20]; 22(5):483-491.

13. Batmaz G, Dane B, Sarioglu A, Kayaoglu Z, Dane C. Can we predict postpartum depression in pregnant women? Clinical and Experimental Obstetrics and Gynecology. 2015 [accessed 2016 Jun 20]; 42(5):605–609.

14. Davis K, Pearlstein T, Stuart S, O'Hara M, Zlotnick C. Analysis of brief screening tools for the detection of postpartum depression: Comparisons of the PRAMS 6-item instrument, PHQ-9, and structured interviews. Archives of Women's Mental Health. 2013 [accessed 2016 Jun 20]; 16(4):271–277.

15. Nunes A, Phipps M. Postpartum Depression in Adolescent and Adult Mothers: Comparing Prenatal Risk Factors and Predictive Models. Maternal and Child Health Journal. 2013 [accessed 2016 Jun 20]; 17(6):1071–1079.

16. Bergink V, Kooistra L, Lambregtse-van den Berg M, Wijnen H, Bunevicius R, van Baar A. Validation of the Edinburgh Depression Scale during pregnancy. Journal of Psychosomatic Research. 2011 [accessed 2016 Jun 15]; 70(4):385–389.

17. Sidebottom A, Harrison P, Godecker A, Kim H. Validation of the Patient Health Questionnaire (PHQ)-9 for prenatal depression screening. Archives of Women's Mental Health. 2012 [accessed 2016 Jun 15]; 15(5):367–374.

18. Zhao Y, Kane I, Wang J, Shen B, Luo J, Shi S. Combined use of the postpartum depression screening scale (PDSS) and Edinburgh postnatal depression scale (EPDS) to identify antenatal depression among Chinese pregnant women with obstetric complications. 2015 [accessed 2016 Jun 14]; 226(1):113–119.

19. Austin MP, Hadzi-Pavlovic D, Saint K, Parker G. Antenatal screening for the prediction of postnatal depression: validation of a psychosocial Pregnancy Risk Questionnaire. Acta Psychiatrica Scandinavica. 2005 [accessed 2016 Jun 15]; 112(4):310–317.

20. McDonald S, Wall J, Forbes K, Kingston D, Kehler H, Vekved M, Tough S. Development of a Prenatal Psychosocial Screening Tool for Post-Partum Depression and Anxiety. Paediatric and Perinatal Epidemiology. 2012 [accessed 2016 Jun 15]; 26(4):316–327.

21. Ji S, Long Q, Jeffrey Newport D, Na H, Knight B, Zach E et al. Validity of depression rating scales during pregnancy and the postpartum period: Impact of trimester and parity. Journal of Psychiatric Research. 2011 [accessed 2016 Jun 15]; 45(2):213-219.

22. Tandon S, Cluxton-Keller F, Leis J, Le H, Perry D. A comparison of three screening tools to identify perinatal depression among low-income African American women. Journal of Affective Disorders. 2012 [accessed 2016 Jun 15]; 136(1-2):155-162.

23. Kroenke K, Spitzer R, Williams J. The PHQ-9. Journal of General Internal Medicine. 2001 [accessed 2016 Jun 15]; 16(9):606-613.

24. Simpson W, Glazer M, Michalski N, Steiner M, Frey B. Comparative efficacy of the generalized anxiety disorder 7-item scale and the Edinburgh Postnatal Depression Scale as screening tools for generalized anxiety disorder in pregnancy and the postpartum period. The Canadian Journal of Psychiatry. 2014 [accessed 2016 Jun 15]; 58(8):434-440.

25. Tran T, Tran T, La B, Lee D, Rosenthal D, Fisher J. Screening for perinatal common mental disorders in women in the north of Vietnam: A comparison of three psychometric instruments. Journal of Affective Disorders. 2011 [accessed 2016 Jun 20]; 133(1-2):281-293.

26. 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. The Annals of Family Medicine. 2009 [accessed 2016 Jun 20]; 7(1):63-70.

27. Dennis C. Can we identify mothers at risk for postpartum depression in the immediate postpartum period using the Edinburgh Postnatal Depression Scale? Journal of Affective Disorders. 2004 [accessed 2016 Jun 20]; 78(2):163-169.

28. Venkatesh K, Zlotnick C, Triche E, Ware C, Phipps M. Accuracy of Brief Screening Tools for Identifying Postpartum Depression Among Adolescent Mothers. Pediatrics. 2013 [accessed 2016 Jun 20]; 133(1):e45-e53.

29. Edmondson O, Psychogiou L, Vlachos H, Netsi E, Ramchandani P. Depression in fathers in the postnatal period: Assessment of the Edinburgh Postnatal Depression Scale as a screening measure. Journal of Affective Disorders. 2010 [accessed 2016 Jun 20]; 125(1-3):365-368.

30. Chaudron L, Szilagyi P, Tang W, Anson E, Talbot N, Wadkins H et al. Accuracy of Depression Screening Tools for Identifying Postpartum Depression Among Urban Mothers. Pediatrics. 2010 [accessed 2016 Jun 20]; 125(3):e609-e617.

31. Phillips J, Charles M, Sharpe L, Matthey S. Validation of the subscales of the Edinburgh Postnatal Depression Scale in a sample of women with unsettled infants. Journal of Affective Disorders. 2009 [accessed 2016 Jun 20]; 118(1-3):101-112.

32. Navarro P, Ascaso C, Garcia-Esteve L, Aguado J, Torres A, Martín-Santos R. Postnatal psychiatric morbidity: a validation study of the GHQ-12 and the EPDS as screening tools. General Hospital Psychiatry. 2007 [accessed 2016 Jun 15]; 29(1):1-7.

33. Teissedre F, Chabrol H. A study of the Edinburgh Postnatal Depression Scale (EPDS) on 859 mothers: detection of mothers at risk for postpartum depression. Encephale. 2004 [accessed 2016 Jun 16]; 30(4):376-381.

34. King P. Validity of Postpartum Depression Screening Across Socioeconomic Groups: A Review of the Construct and Common Screening Tools. Journal of Health Care for the Poor and Underserved. 2012 [accessed 2016 Jun 15]; 23(4):1431-1456.

35. Tobin C, Di Napoli P, Wood-Gauthier M. Recognition of Risk Factors for Postpartum Depression in Refugee and Immigrant Women: Are Current Screening Practices Adequate? Journal of Immigrant and Minority Health. 2014 [accessed 2016 Jun 15]; 17(4):1019-1024.

36. Zubaran C, Schumacher M, Roxo M, Foresti K. Screening tools for postpartum depression: validity and cultural dimensions. African Journal of Psychiatry. 2011 [accessed 2016 Jun 7]; 13(5): 357-365.

37. Downe S, Butler E, Hinder S. Screening tools for depressed mood after childbirth in UK-based South Asian women: a systematic review. Journal of Advanced Nursing. 2007 [accessed 2016 Jun 21]; 57(6):565-583.

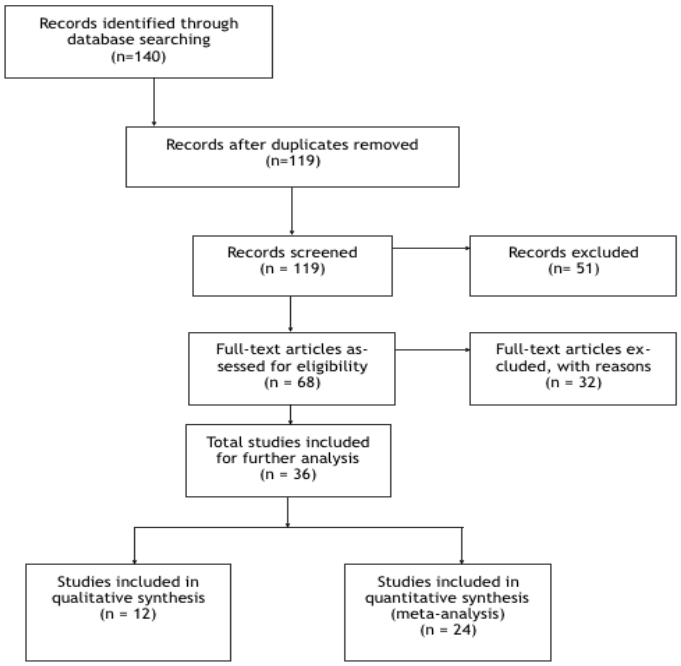
38. Austin M, Lumley J. Antenatal screening for postnatal depression: a systematic review. Acta Psychiatrica Scandinavica. 2003 [accessed 2016 Jun 21]; 107(1):10-17.

39. Dalton K Holton W. Depression after childbirth. New York: Oxford University Press; 2001.

40. Myers, E. R., Aubuchon-Endsley, N., Bastian, L., Gierisch, J., Kemper, A., Swamy, G., Sanders, G Efficacy and Safety of Screening for Postpartum Depression. United States: Agency for Healthcare Research Quality U.S. Department of Health and Human Services; 2013 [accessed 2016 Jun 13].

41. Gibson J, McKenzie-McHerg, K, Shakespeare, J, Price, J, & Gray, R. A systematic review of studies validating the Edinburgh Postnatal Depression Scale in antepartum and postpartum women. Acta Psychiatr Scand 2009 May; 119(5): 350-64.

42. American College of Obstetricians and Gynecologists. Committee Opinion: Screening for perinatal depression. Number 630, May 2015.



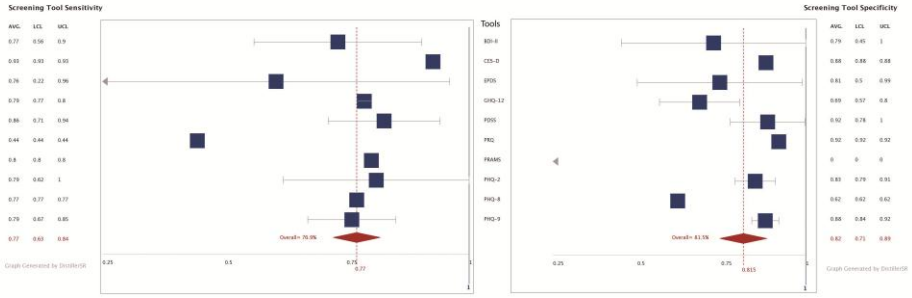


Table 1. Summary characteristics of the 24 screening studies

Author, Year, Country	Study Source	Sample Size	Period of assessment	Screening tools	Reference Standard
Kroenke et al. 2001 USA	CC	6000	Study not limited to pregnant or postpartum patients	PHQ-9	SCID
Gjerdingen et al. 2009 USA	CC	508	PN (0 - 1, 2, 4, 6, & 9-months)	PHQ-9, 2QS	SCID
Smith et al. 2010 USA	HC & CC	218	AN (<17 weeks)	PHQ-2, PHQ-8	CIDI
Bergink et al. 2011 Netherlands	CC	845	AN (1st, 2nd, 3rd trimester)	EPDS	CIDI
Chae et al. 2012 USA	CC	200	PN (4 and 6 months)	PHQ-2	EPDS
Yawn et al. 2008 USA	CC	500	PN (5 – 12 weeks)	EPDS, PHQ-9	None reported
Davis et al. 2013 USA	HC & CC	1392	PN (0 – 12 months)	PHQ-9, PRAMS-6	CIDI
Beck, Gable 2001 US	HC	150	PN (2 – 12 weeks)	BDI-II, EPDS, PDSS	SCID
Sidebottom et al. 2012 USA	CC	745	AN (initial visit)	PHQ-9	SCID
Dennis 2002 Canada	Not reported	594	PN (1, 4 and 8 weeks)	EPDS	None reported
Batmaz et al. 2014 Turkey	HC	285	AN (exact time not reported) and PN (24th week)	BDI-II, EPDS	None reported
Austin et al. 2005	HC	1296	AN (3rd trimester) and	PRQ	CIDI

Australia			PN (2 and 4 months)		
Tandon et al. 2011 USA	HV	95	AN & PN (exact times of screening not reported)	BDI-II, CES-D, EPDS	SCID
Zhao et al. 2015 China	HC	843	AN (various gestational stages)	EPDS, PDSS	SCID, MINI
Venkatesh et al. 2014 USA	CC	106	PN (6 weeks, 3- & 6-months)	EPDS	SCID
Edmondson et al. 2010 UK	HC	1192	PN (7 weeks)	EPDS	SCID
Simpson et al. 2014 Canada	HC	240	AN and PN (exact periods unspecified)	EPDS, GAD	None reported
McDonald et al. 2012 Canada	CC	1578	AN (24, 34 – 36 weeks) and PN (4 months)	EPDS, PSS, STAI	None reported
Tran et al. 2011 Vietnam	CC	364	AN(exact period unspecified) and PN (4 – 6 weeks)	EPDS,_GHQ-12, Zung SAS	SCID
Shuang et al. 2011 USA	HC	534	AN_(1^{st}, 2^{nd}, 3^{rd} trimester) and PN (up to 26 weeks)	EPDS, BDI-II, $HRSD_{15,17}$	SCID
Chaudron et al. 2010 USA	CC	198	PN (0 – 14 months)	EPDS, BDI-II, PDSS	SCID
Phillips et al. 2009 Australia	CC	309	PN (0 – 12 months)	EPDS, BDI-II, BAI	SCID
Navarro et al. 2006 Spain	HC	1453	PN (up to 6 weeks)	EPDS, GHQ-12	SCID
Teissedre et al. France	CC	859	PN (3 days, 6 weeks)	EPDS	MINI, BDI, SIGH-D

Key: CC – community based clinic, HC – hospital based clinic, HV- home visitation program AN – antenatal, PN – postnatal
EPDS- Edinburg Postpartum Depression Scale, PDSS - Postpartum Depression Screening Scale, PRQ - Pregnancy Risk Questionnaire, BDI - Beck Depression Inventory, BDI-II - Beck Depression Inventory-II, CESD - the Center for Epidemiological Studies Depression, PHQ - Patient Health Questionnaire (versions 2, 8, & 9), 2QS - Two question screen, PSS - Cohen Perceived Stress Scale, STAI –State-Trait Anxiety Inventory, Zung Sas - $HRSD_{15,17}$ – Hamilton Rating Scale for Depression (15 and 17 items), BAI – Beck Anxiety Inventory, SIGH-D (Structured Interview Guide for the Hamilton Depression Scale

Table 2. Summary characteristics of the 24 screening studies

Author, Year, Country	Study Source	Sample Size	Period of assessment	Screening tools	Reference Standard
Kroenke et al. 2001 USA	CC	6000	Study not limited to pregnant or postpartum patients	PHQ-9	SCID
Gjerdingen et al. 2009 USA	CC	508	PN (0 - 1, 2, 4, 6, & 9-months)	PHQ-9, 2QS	SCID
Smith et al. 2010 USA	HC & CC	218	AN (<17 weeks)	PHQ-2, PHQ-8	CIDI
Bergink et al. 2011 Netherlands	CC	845	AN (1^{st}, 2^{nd}, 3^{rd} trimester)	EPDS	CIDI
Chae et al. 2012 USA	CC	200	PN (4 and 6 months)	PHQ-2	EPDS
Yawn et al. 2008 USA	CC	500	PN (5 - 12 weeks)	EPDS, PHQ-9	None reported
Davis et al. 2013 USA	HC & CC	1392	PN (0 - 12 months)	PHQ-9, PRAMS-6	CIDI
Beck, Gable 2001 US	HC	150	PN (2 - 12 weeks)	BDI-II, EPDS, PDSS	SCID
Sidebottom et al. 2012 USA	CC	745	AN (initial visit)	PHQ-9	SCID
Dennis 2002 Canada	Not reported	594	PN (1, 4 and 8 weeks)	EPDS	None reported
Batmaz et al. 2014 Turkey	HC	285	AN (exact time not reported) and PN (24^{th} week)	BDI-II, EPDS	None reported
Austin et al. 2005 Australia	HC	1296	AN (3^{rd} trimester) and PN (2 and 4 months)	PRQ	CIDI
Tandon et al. 2011 USA	HV	95	AN & PN (exact times of screening not reported)	BDI-II, CES-D, EPDS	SCID
Zhao et al. 2015 China	HC	843	AN (various gestational stages)	EPDS, PDSS	SCID, MINI
Venkatesh et al. 2014 USA	CC	106	PN (6 weeks, 3- & 6-months)	EPDS	SCID
Edmondson et al. 2010 UK	HC	1192	PN (7 weeks)	EPDS	SCID
Simpson et al. 2014 Canada	HC	240	AN and PN (exact periods unspecified)	EPDS, GAD	None reported
McDonald et al. 2012 Canada	CC	1578	AN (24, 34 - 36 weeks) and PN (4 months)	EPDS, PSS, STAI	None reported
Tran et al. 2011 Vietnam	CC	364	AN(exact period unspecified) and PN (4 - 6 weeks)	EPDS,_GHQ-12, Zung SAS	SCID
Shuang et al. 2011 USA	HC	534	AN_(1^{st}, 2^{nd}, 3^{rd} trimester) and PN (up to 26 weeks)	EPDS,|$BDI_{-n,}$ $HRSD_{15,17}$	SCID
Chaudron et al. 2010 USA	CC	198	PN (0 - 14 months)	EPDS, BDI-II, PDSS	SCID
Phillips et al. 2009 Australia	CC	309	PN (0 - 12 months)	EPDS, BDI-II, BAI	SCID
Navarro et al. 2006 Spain	HC	1453	PN (up to 6 weeks)	EPDS, GHQ-12	SCID
Teissedre et al. France	CC	859	PN (3 days, 6 weeks)	EPDS	MINI, BDI, SIGH-D

Key: CC - community based clinic, HC - hospital based clinic, HV- home visitation program AN - antenatal, PN - postnatal EPDS- Edinburg Postpartum Depression Scale, PDSS - Postpartum Depression Screening Scale, PRQ - Pregnancy Risk Questionnaire, BDI -Beck Depression Inventory, BDI-II - Beck Depression Inventory-II, CESD - the Center for Epidemiological Studies Depression, PHQ - Patient Health Questionnaire (versions 2, 8, & 9), 2QS - Two question screen, PSS - Cohen Perceived Stress Scale, STAI -State-Trait Anxiety Inventory, Zung Sas - $HRSD_{15,17}$ - Hamilton Rating Scale for Depression (15 and 17 items), BAI - Beck Anxiety Inventory, SIGH-D (Structured Interview Guide for the Hamilton Depression Scale