Published in final edited form as:

J Affect Disord. 2017 January 15; 208: 615–620. doi:10.1016/j.jad.2016.10.013.

Generating an Efficient Version of the Edinburgh Postnatal Depression Scale in an Urban Obstetrical Population

Jackie K Gollan, Ph.D.¹, Stephen R. Wisniewski, Ph.D.², James F. Luther, M.A.³, Heather F. Eng, B.A.⁴, John Louis Dills, M.S.⁵, Dorothy Sit, M.D.⁶, Jody D. Ciolino, Ph.D.⁷, and Katherine L. Wisner, M.D.⁸

¹Department of Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine, 676 North St Clair Street, Suite 1000, Chicago, IL, 60611 USA

²Epidemiology Data Center, Graduate School of Public Health, University of Pittsburgh, 127 Parran Hall, 130 DeSoto Street, Pittsburgh, Pennsylvania, 15261, USA

³Epidemiology Data Center, Graduate School of Public Health, University of Pittsburgh, 127 Parran Hall, 130 DeSoto Street, Pittsburgh, Pennsylvania, 15261, USA

⁴Epidemiology Data Center, Graduate School of Public Health, University of Pittsburgh, 127 Parran Hall, 130 DeSoto Street, Pittsburgh, Pennsylvania, 15261, USA

⁵Epidemiology Data Center, Graduate School of Public Health, University of Pittsburgh, 127 Parran Hall, 130 DeSoto Street, Pittsburgh, Pennsylvania, 15261, USA

⁶Department of Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine, 676 North St Clair Street, Suite 1000, Chicago, IL, 60611 USA

⁷Department of Preventive Medicine-Biostatistics, Northwestern University Feinberg School of Medicine, 680 N Lake Shore Drive, Suite 1400, Chicago IL 60611, Chicago, IL, 60611 USA

⁸Department of Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine, 676 North St Clair Street, Suite 1000, Chicago, IL, 60611 USA

Abstract

Background—Postpartum depression incurs significant burden and suffering.

Methods—We investigated the latent structure of the most commonly used screening measure, the Edinburgh Postnatal Depression Scale (EPDS) in women ($N = 15,172$) and tested its predictive validity for the diagnosis of depression as determined with a structured clinical interview. Exploratory and confirmatory factor analyses, Receiver Operating Characteristic curves, and logistic regression analyses were conducted.

Corresponding author and reprint requests Jackie K. Gollan, Ph.D., Department of Psychiatry and Behavioral Sciences, Northwestern University's Feinberg School of Medicine, 676 North St Clair Street, Suite 1000, Chicago, IL 60611, Tel.: 312-695-6121.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Results—A seven-item one factor scale (items 1, 2, 6, 7, 8, 9, 10) emerged with a Goodness of Index Fit Index (GFI) = 0.96, relative to the ten-item two factor version of the EPDS (GFI = 0.94). The seven-item EPDS achieved good sensitivity and specificity in predicting the 10-item EPDS, with a cut point score of 4 on the seven item EPDS to predict a 10-item EPDS score of 10 or more (sensitivity = 95%, specificity = 91%). The seven and 10-item EPDS showed a similar ability to predict a diagnoses of depression (area under the ROC curve = 0.795 for the 10-item, 0.770 for the seven-item EPDS). Logistic regression analyses showed similar predictive ability between the seven- and 10-item scales in predicting scores higher than 18 on the clinical interview.

Limitations—The sample represents women from one Midwest medical center and the EPDS was measured via phone.

Conclusion—The seven-item one factor version of the EPDS is an efficient and effective measure of depression severity on par with the two factor 10-item version of the EPDS.

Keywords

EDPS; pregnancy; factor structure; depression; anxiety

INTRODUCTION

An efficient postpartum depression (PPD) case identification and intervention strategy holds the potential to reduce maternal disability and to avert a new generation at risk. The most commonly used screening tool world-wide for postpartum depression is the 10-item Edinburgh Postnatal Depression Scale (EPDS; 1), which is designed to assess depression in the last seven days. Despite its design as being a measure of depression, several items on the EPDS refer to anxiety. Numerous studies have conducted factor analyses on the EPDS to test its operational efficiency towards the goals of making a precise assessment of postpartum depression. For example, in Table 1, a variety of studies support the EPDS as a one-dimensional measure, suggesting that three items do not contribute to the depression factor (2–6). However, some investigators have identified a two-factor structure in samples of pregnant (7) and postpartum women (4, 8–10), and others support a three-factor structure in samples of postpartum women (6, 11–13). Given the variety of results, new work is necessary to investigate the extent to which the EPDS reflects the dimensions of depression and anxiety, towards generating an evidence-based screening measure with increased internal factor consistency and internal reliability.

Further, item factor loading across studies reporting factor analyses of the EPDS have reflected a multidimensional view of the EPDS across divergent cultures and nationalities, and phases of peripartum, producing a construct of an ‘anxiety’ factor comprised of items 3, 4, and 5. Strong correlations between the EPDS total score and anxiety subscales in pregnant women (11) range from $r = 0.56$ on the anxiety subscale of the Symptom Checklist-90-R (14) and $r = 0.54$ on the State Trait Anxiety Inventory, State and Trait versions, although correlations between the STAI and EPDS lack full support in the literature (15, 16). However, anxiety and depression factors have shown strong associations at weeks 14 and 30 postpartum ($r = 0.73$; $r = 0.82$, respectively) (16), suggesting a shared dimensional substrate

of reactivity evident in pregnancy (17). Using a two-factor 10 item version of the EPDS reduce the efficiency of the EPDS and may conflate symptoms of depression with anxiety.

A second area of debate is the acceptability of the cut-point for identifying women with major depression during postpartum. The proposed cut-off by the developers is an EPDS > 13 producing a specificity = 78% and positive predictive value = 73% (1, 18)¹. The EPDS has been associated with scores on the Hamilton Rating Scale for Depression = 20 (19, 20). However, findings from other studies support a cut-point of EPDS = 15 during pregnancy (1) though these cut-off scores vary depending on the version of the EPDS (21, 22). Further, inconsistencies emerge with Cox and Holden's (2003) (23) suggestion that clinicians use a cutoff score of = 10 as an indicator of postpartum depression when ample resources to conduct post-screening clinical interviews are available, and a higher cutoff score (= 13) when limited capacity for post-screening evaluation exists.

A third, and clinically relevant, issue is the use of the EPDS as an efficient method to detect depression. A review of 37 studies validating the EPDS (24) determined the capacity of the EPDS to detect depression when compared to a structured diagnostic interview. Using the 10 criterion, the sensitivity ranged from 59%–100% and the specificity from 44%–97%, and the = 13 criterion resulted in sensitivity -and specificity spanning similar wide ranges from 34%–100% and 49%–100%, respectively. To the extent that there are two or more factors (representing depression, anxiety, anhedonia), the question remains to what extent would women who screen 'positive' would also be diagnosed with Major Depressive Disorder (MDD)?

Given the debates in the field, identifying an efficient version of the EPDS using a large urban sample in the United States may offer clinical value and permit clinicians to ascertain depression, anxiety, or a dimensional trait related to common mental disorders during postpartum. As a protocol for universal screening of postpartum depression has yet to be implemented, seeking to identify the extent to which the EPDS can capture 'true positive' cases has high public health significance. Specifically, it supports non-mental health clinicians in their effort to measure characteristics of depression that can be addressed with a treatment plan, while also educating patients of symptoms that may relate to a condition that requires self-monitoring and professional care. The study aims were to: 1) identify the specific latent structure of the EPDS in a large-scale community study of women using factor analyses; 2) investigate the extent to which the factorial structure derived in Aim 1 can be replicated with the remaining sample with a refined structural model; 3) identify the extent to which we can use this EPDS to establish new threshold scores for a shortened version of the EPDS that correspond to the standard EPDS threshold for clinical use; and 4) evaluate the extent to which the shortened EPDS predicts severity of depression using the Structured Interview Guide for the Hamilton Depression Rating Scale with Atypical Depression Supplement (SIGH-ADS) (19).

¹Specificity is defined as the ability of the EPDS to correctly classify a person as healthy; sensitivity is defined as the ability of the EPDS to correctly define a person as depressed; positive predictive value is defined as the percentage of participants with a positive EPDS who actually have depression.

METHODS

Study Description

This was a secondary analysis from a dataset entitled '*Identification and Therapy of Postpartum Depression Study*' (Wisner, PI; NCT 00282776, funding period 09/01/2006-07/31/2011). Basic information about this protocol is described elsewhere (25). This study involved case identification, diagnoses, and intervention of PPD at a large urban medical center with nearly 10,000 deliveries per year. This protocol was approved by the University of Pittsburgh Institutional Review Board (IRB). The screening protocol was the initial component of a comprehensive case identification, diagnosis, and intervention project for depression in postpartum women.

The investigation had three components. First, nurses or social workers approached women on the maternity ward at Magee-Womens Hospital of the University of Pittsburgh Medical Center (Pennsylvania) to offer information about postpartum depression and to invite them to participate in the screening phase of the study by telephone between four and six weeks post-delivery.² Eligible women signed a waiver approved by the University of Pittsburgh IRB, which allowed collection of contact information and later telephone screening. Second, between Weeks 4 and 6 postpartum, trained telephone screeners called enrolled women.³ Finally, women who endorsed EPDS scores > 10 on the phone were offered a home visit evaluation for psychiatric assessment, which was scheduled within two weeks of the screen. Women who declined the home visit were offered a telephone screen for MDD, but were not included in the analysis. Women with EPDS score > 20 or those who endorsed any response other than "none" on the EPDS self-harm question (item 10) were immediately interviewed by the supervising clinician for safety assessment and intervention planning.

Study Population

A total sample of 15,172 of newly-delivered women were enrolled into one study as described in a previous report (25), of which the first one-third of the sample (n = 5,055) were used for the exploratory factor analysis, and the remaining two-thirds of the sample (N=10,117) were used for the confirmatory factor analysis. Women were eligible to participate if they delivered a live infant, were English speaking, 18 years of age or older, could be contacted by phone, and provided informed consent.

Assessments

Items on the Edinburgh Postnatal Depression Scale (1) are scored 0, 1, 2, or 3 according to increased severity of the symptom. Items 3, 5, 6, 7, 8, 9, and 10 are reverse scored (i.e., 3, 2, 1, and 0). The total score is determined by adding the scores for each of the 10 items, with a range of possible scores from 0 to 30. The split-half reliability of the scale was 0.88, and good internal reliability was shown with an alpha of 0.87 (1).

²The EPDS phone screen was generated between four and six weeks after birth to coincide with (a) the post-delivery OB/GYN evaluation and (b) increase in psychiatric evaluations and rapid onset, post-birth episodes of mental illness.

³Screeners ceased calls after two weeks. A postcard was sent after three days. Attempts to contact ended after six weeks.

To assess the screening quality of the seven- and 10-item EPDS to screen depression, we used a 29-item clinical interview, the SIGH-ADS (19). The SIGH-ADS is a measure of depression severity, comprised of the 21-item Hamilton Rating Scale for Depression (26) and an additional eight items to ascertain neurovegetative symptoms. These interviews were conducted during at-home visits to assign depression diagnoses. Scores ≤ 8 represent a euthymic mood state; scores ≥ 18 represent depression.

Statistical Plan

First, we conducted an unrestricted exploratory factor analysis using maximum likelihood (factor analytic method) using a polychoric correlation matrix to address the ordinal scale of the EPDS with one third of the sample ($n = 5,055$). This initial factor analysis was followed by extraction of the relevant number of factors to evaluate the final factor loading with the rest of the sample using SAS Version 9.3. Second, we conducted a confirmatory factor analysis of the selected factor model using structured equation modeling using maximum likelihood extraction (factor analysis) PROC CALIS using SAS Version 9.3 with the remaining two-thirds of the sample ($n = 10,117$). We used the following indices, including Goodness-of-Fit Index (GFI) to assess the fit of the factor structure, ranging between 0 and 1 with higher scores indicating a better fit; the Root Mean Square Error of Approximation estimate (RMSEA, (27)) to estimate good model fit; the Bentler-Bonnett Normed Fit Index (NFI (24) to estimate incremental fit, and the Parsimonious Goodness of Fit Index to evaluate room for an improvement of model. Third, we used Receiver Operating Characteristic (ROC) curve analyses to compare a shortened version of the EPDS (chosen as a result of the factor analyses inferences) with the full 10-item EPDS. The area under the ROC represents the discriminatory capacity of the seven item EPDS in comparison to the 10-item EPDS, with an area under the ROC of 85% representing an acceptable exchange of the two versions of EPDS. Finally, we used logistic regression to determine the extent to which a score > 10 on the one-factor seven item EPDS predicted clinical depression ascertained by a structured clinical interview.

RESULTS

On average, participants were in their late twenties (Mean = 29.6 ± 5.6 years), and the majority were Caucasian ($n = 12,129$, 80%) with a smaller percentage of African-Americans ($n = 2,220$, 14.6%) and a Multiracial group ($n = 810$, 5.3%) comprised of Native American, Middle Eastern, or Multiracial backgrounds. The mean EPDS score across participants was 5.2 ± 4.4 . Note that while differences in the distributions of race and medical insurance across the sub-samples are significant at $p < 0.10$ and $p < 0.05$ respectively, the sub-samples are near mirror images and statistical significance is driven by the extremely large sample sizes. Table 2 provides data on demographic characteristics by sample.

Exploratory Factor Analysis

Exploratory factor analysis (EFA) using maximum likelihood estimation extraction method (factor analytic method) and promax rotation was used to assess the dimensional structure of the 10-item EPDS. The Kaiser-Guttman rule, scree plot, and proportion criterion all support a two factor solution. Factor loading for the two-factor solution are presented in Table 3,

revealing a seven-item (i.e., 1, 2, 6–10) and a three-item (i.e., 3–5) factor (i.e., 3–5).⁴ These latter items correspond to a potential anxiety dimension of the EPDS, while the others correspond to depression, suggesting the possibility of a single-factor or ‘pure’ depression scale.

The scree plot suggests that there were two factors, along with the Kaiser-Guttman rule and the proportion criterion. Moreover, there are no items that load $> .4$ on both factors. Therefore, we performed Confirmatory Factor Analyses on both the 10-item single factor and the proposed two factor solution to see which best fit the data. Table 3 provides a summary of Exploratory Factor Analysis results for EPDS using Maximum Likelihood Estimation.

Confirmatory Factor Analysis

We conducted confirmatory factor analyses of the 10-item single factor model, the two-factor model, and the 7-item single factor model with the remaining two thirds of the sample ($n = 10,117$). We used the method of maximum likelihood estimation to compare the baseline models with the hypothesized models. Model statistics and indices of incremental fit for each of the models are presented in Table 4. As can be seen, all hypothetical models fit the data better than the baseline models with chi-square values substantially decreasing. However, given that the chi-square is highly sensitive to large sample size and considering the extremely large study sample, such large, statistically significant chi-square values are to be expected. (When N is set to 100 in the correlation matrix used for the CFAs, the chi-square test for the 10-item single factor model remains significant. In contrast, the chi-square tests for the two-factor and 7-item single factor models are not significant). For the 10-item single factor model, only the NFI suggests a good fit while the other indices suggest a poor fit. With values for the GFI and NFI > 0.9 and an improvement in the parsimonious GFI, the two-factor model is clearly a better fit to the data than the 10-item single factor model.

As for the proposed single factor 7-item model, modest increases in both the GFI and NFI suggest a slightly better fit to the data when compared to the two-factor 10-item model. However, the parsimonious GFI is lower even when compared to the 10-item single factor model. In sum, relative to the 10-item single-factor structure, results of the CFAs support a two factor structure (7-item and 3-item) using the 10-item EPDS and the single factor seven-item version of the EPDS. Table 4 provides a summary of the fit statistics from the one and two-factor models using 10 items and a one-factor model using seven items.

ROC and Logistics Regression Analyses

Results from the ROC analyses showed that the seven-item EPDS achieved good sensitivity and specificity in predicting the 10-item EPDS. A cut point score of 4 on the seven item EPDS may be used to predict a 10-item EPDS score of 10 or more (sensitivity = 95% and

⁴Notably, because the chi-square significance test is sensitive to sample size (we have a sample greater than 5,000), this test is highly likely to reject the null irrespective of the number of factors. This was supported with the chi-square results when the sample was reduced to 100.

specificity = 91%). In addition, a cut point score of 5 may be used to predict a 10-item EPDS score of 13 or more (sensitivity = 94% and specificity = 99%).

Of the participants who completed the SIGH-ADS, ($n = 211$), 151 endorsed SIGH-ADS > 18. Both seven- and 10-item EPDS exhibited associations with depressive status as measured by the SIGH-ADS score ($p < 0.001$, sample correlation coefficient = 0.53 for seven-item EPDS and 0.56 for the 10-item EPDS). In addition, we demonstrated that the ROC curves (using SIGH-ADS diagnosis as an outcome) for both the seven- and 10-item EPDS are similar and they show relatively strong predictive ability (area under the ROC curve is 0.795 for the 10-item EPDS and 0.770 for the seven-item EPDS). Based on these data, a potential cut point for seven-item EPDS is a score of 8 (sensitivity = 76% and specificity = 68%), and a potential cut point for the 10-item EPDS is 14 (sensitivity = 83% and specificity = 66%). Though not perfect in predicting SIGH-ADS score/diagnosis, they have similar properties with respect to predictive ability. Further, estimated 95% confidence intervals for the area under the ROC curves based on 1,000 bootstrapped samples were (0.718, 0.822) and (0.748, 0.841) for the seven- and 10-item EPDS, respectively. As these confidence intervals overlap substantially, we have further evidence that the seven-item and 10-item EPDS are similar in predictive ability for the SIGH-ADS diagnosis.

Furthermore, the parameter estimates from the logistic regression analyses suggest similarities between the seven- and 10-item EPDS in their ability to predict scores higher than SIGH-ADS 18 [*yes* ($n = 151$), *no* ($n = 59$)]. A one-unit increase in score on the seven-item EPDS corresponds to a model-estimated increased odds of depression (according to the SIGH-ADS > 18) of 1.38 (95% confidence interval: 1.23, 1.55), while a one-unit increase in score on the 10-item EPDS corresponds to a model-estimated increased odds of depression of 1.34 (95% confidence interval: 1.22, 1.47). Again, the substantial overlap of these confidence intervals indicate similar predictive ability of both the seven- and 10-item EPDS for depression as determined by the SIGH-ADS > 18 criterion. Figure 1 shows the ROC for scores higher than 9 and 12 on the seven-item EPDS. Figure 2 shows the ROC for the seven-item and the 10-item EPDS in predicting depression diagnosis using the SIGH-ADS.

DISCUSSION

The results of this study support a one-dimensional factor structure on a condensed (seven-item) version of the EPDS. This factor was identified as an overall depression measure (items 1, 2, 6, 7, 8, 9, 10), and items 3, 4, and 5 were excluded, suggesting a single-factor or 'pure' depression scale. This finding corresponds with the results of previous published reports of one-factor structure using large scale studies (4), which proposed a shortened version of the EPDS. This finding contrasts with results that reported a two-factor structure including Dutch postpartum women (29), Canadian women at six weeks postpartum (13), U.S. adolescent mothers in early postpartum (9), and Australian mothers on a mother-infant inpatient unit (30).

Results suggest comparability of seven- and 10-item EPDS in predicting depression severity according to the SIGH-ADS, and clinicians could consider using a threshold score of 8 on the seven-item scale to screen for depression. Translating these features into a clinical

context, one meaningful EPDS factor may be considered. Relying on the anxiety items in clinic may be premature given that research shows weaker correlations between anxiety measures (STAI) and the EPDS anxiety subscale, which may imply weaker construct validity (11), though studies continue to represent. On the other hand, Matthey (2008) (31) demonstrated that the EPDS anxiety subscale detected an anxiety disorder with a sensitivity of 66.7% and specificity of 88.2%, suggesting that the EPDS may offer an efficient screen for anxiety symptoms during postpartum. Because comorbid anxiety with depression increases the likelihood of treatment refractoriness, the identification of anxiety can inform treatment planning and enhances disease management for women with anxiety (32–34).

Clinically, the emergent one-factor structure for a condensed version of the EPDS generates an efficient measure that assists in clinical care in several ways. First, the use of the shorter screen permits clinicians to discern whether mothers show a clinical exacerbation of depression beyond normative postpartum adjustment. Scores of 8 or higher on the seven-item EPDS scale facilitates monitoring of the mental health of mothers during their postpartum visits. Second, use of the shorter scale permits clinicians to collect information that is critical to detecting emergent psychiatric illness. Third, the use of the seven-item versus 10-item EPDS offers efficiency with minimal cost and burden to clinician and patient. The seven-item EPDS is a shorter screening tool that offers a practical approach to measuring unrecognized or pre-symptomatic postpartum depression for women in the US. Relative to our findings, epidemiologic studies (18, 35) have shown that 91.1% of screen-positive postpartum women were diagnosed with primary mood disorders, predominantly MDD. In sum, our findings suggest two options for clinical use: first, the EPDS can be administered in its present format where the clinician may more closely inspect the second factor when looking for anxious features of postpartum depression; and second, the EPDS can be administered as a reduced 7-item “pure” depression scale in consideration of using a stronger screening measure, improved time management, and reduced patient burden. Though one could argue that a patient’s time in completing three additional items is only an incremental advantage, such a modest reduction saves time across thousands of patients. Moreover, the benefit in limiting the EPDS to seven items permits the patient to differentiate depression from anxiety.

The main strength of this study was its careful enrollment of a sample of women who were seeking services from a large obstetrical hospital, the sample size, the adherence to rigorous systematic screening, efforts to representation of diverse racial groups and cultures (36–37). Further strengths include and the attention to administering the EPDS between four to six weeks postpartum, which reflects standard of care for many women (e.g., visit OB/GYN during this time), and the delineation of extent to which scores on the two versions of the EPDS identify clinical depression.

The limitations of the study include the two phase design of the study in which the clinical interview (SIGH-ADS) was designated as the gold standard in identifying depression. Also, our use of a phone-based administration of the EPDS may have constrained reporter reliability. Finally, the study included clinical interviews focused on depression severity and excluded anxiety severity, which limited our ability to investigate the anxiety subscale.

With a large sample, we found that the seven-item, one factor version of the EPDS was as efficient as its original ten-item version. As guidance in clinical detection of postpartum varies across the US, and in other countries (38), and other shorter methods of detection have been carefully proposed (e.g., the two Whooley questions)(38), our findings have identified how to use a shorter version of an existing measure that represents the main features of depression. We also identified cutoff scores of the seven-item EPDS that permits clinicians to extrapolate the likelihood of their patients having clinical depression. Transitioning towards an evaluation protocol that uses the seven-item score has clinical meaning as it encourages clinicians to use the threshold scores to support their decision to conduct a comprehensive clinical assessment for postpartum depression.

References

1. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *The British Journal of Psychiatry: The Journal of Mental Science*. 1987; 150:782–786. [PubMed: 3651732]
2. Berle JO, Aarre TF, Mykletun A, Dahl AA, Holsten F. Screening for postnatal depression. Validation of the Norwegian version of the Edinburgh Postnatal Depression Scale, and assessment of risk factors for postnatal depression. *Journal of Affective Disorders*. 2003; 76(1–3):151–6. [PubMed: 12943945]
3. Des Rivieres-Pigeon C, Seguin L, Brodeur JM, Perreault M, Boyer G, Colin C, et al. The Edinburgh Postnatal Depression Scale: the validity of its Quebec version for a population low socioeconomic status mothers. *Canadian Journal of Community Mental Health = Revue Canadienne de Sante Mentale Communautaire*. 2000; 19(1):201–14. [PubMed: 12152176]
4. Pallant JF, Miller RL, Tennant A. Evaluation of the Edinburgh Post Natal Depression Scale using Rasch analysis. *BMC Psychiatry*. 2006; 6:28. [PubMed: 16768803]
5. Reichenheim ME, Moraes CL, Oliveira AS, Lobato G. Revisiting the dimensional structure of the Edinburgh Postnatal Depression Scale (EPDS): empirical evidence for a general factor. *BMC Medical Research Methodology*. 2011; 11:93. [PubMed: 21689442]
6. Chabrol H, Teissedre F. Relation between Edinburgh Postnatal Depression Scale scores at 2–3 days and 4–6 weeks postpartum. *Journal of Reproductive and Infant Psychology*. 2004; 22:33–9.
7. Adouard F, Glangeaud-Freudenthal NM, Golse B. Validation of the Edinburgh postnatal depression scale (EPDS) in a sample of women with high-risk pregnancies in France. *Archives of Women's Mental Health*. 2005; 8(2):89–95.
8. Astbury J, Brown S, Lumley J, Small R. Birth events, birth experiences and social differences in postnatal depression. *Australian Journal of Public Health*. 2009; 18(2):176–84. 1994.
9. Logsdon MC, Usui WM, Nering M. Validation of Edinburgh postnatal depression scale for adolescent mothers. *Archives of Women's Mental Health*. 12(6):433–40.
10. Matthey S, Henshaw C, Elliott S, Barnett B. Variability in use of cut-off scores and formats on the Edinburgh Postnatal Depression Scale: implications for clinical and research practice. *Archives of Women's Mental Health*. 2006; 9(6):309–15.
11. Brouwers EP, van Baar AL, Pop VJ. Does the Edinburgh Postnatal Depression Scale measure anxiety? *Journal of Psychosomatic Research*. 2001; 51(5):659–63. [PubMed: 11728506]
12. Jomeen J, Martin CR. Replicability and stability of the multidimensional model of the Edinburgh Postnatal Depression Scale in late pregnancy. *Journal of Psychiatric and Mental Health Nursing*. 2007; 14(3):319–24. [PubMed: 17430456]
13. Ross LE, Gilbert Evans SE, Sellers EM, Romach MK. Measurement issues in postpartum depression part 1: anxiety as a feature of postpartum depression. *Archives of Women's Mental Health*. 2003; 6(1):51–7.
14. Derogatis, LR. Symptom Checklist-90-Revised: Administration, scoring and procedures manual. 3rd. Minneapolis, MN: National Computer Systems, Inc; 1994. p. 55440

15. Spielberger, C., Gorsuch, R., Lushene, R. Manual for the State-Trait Inventory. Palo Alto, CA: Consulting Psychologists; 1970.
16. Stuart S, Couser G, Schilder K, O'Hara MW, Gorman L. Postpartum anxiety and depression: onset and comorbidity in a community sample. *The Journal of Nervous and Mental Disease*. 1998; 186(7):420–4. [PubMed: 9680043]
17. Jomeen J, Martin CR. Confirmation of an occluded anxiety component within the Edinburgh Postnatal Depression Scale (EPDS) during early pregnancy. *Journal of Reproductive and Infant Psychology*. 2005; 23:143–54.
18. O'Hara MW, Swain AM. Rates and risks of postpartum depression - a meta-analysis. *International Review of Psychiatry*. 1996; 8(1):37–54.
19. Williams, JBW., Terman, M. Structured Interview Guide for the Hamilton Depression Rating Scale with Atypical Depression Supplement (SIGH-ADS). New York: New York State Psychiatric Institute; 2003.
20. Peindl K, Wisner K, Hanusa B. Identifying depression in the first postpartum year: Guidelines for office-based screening and referral. *Journal of Affective Disorders*. 2004; 80(1):37–44. [PubMed: 15094256]
21. Adouard F, Glangeaud-Freudenthal NM, Golse B. Validation of the Edinburgh postnatal depression scale (EPDS) in a sample of women with high-risk pregnancies in France. *Archives of Women's Mental Health*. 2005; 8(2):89–95.
22. Felice E, Saliba J, Grech V, Cox J. Validation of the Maltese version of the Edinburgh Postnatal Depression Scale. *Archives of Women's Mental Health*. 2003; 9(2):75–80. 2006.
23. Cox J, Holden J. Perinatal mental health: a guide to the Edinburgh Postnatal Depression Scale (EPDS): Royal College of Psychiatrists.
24. Gibson J, McKenzie-McHarg K, Shakespeare J, Price J, Gray R. A systematic review of studies validating the Edinburgh Postnatal Depression Scale in antepartum and postpartum women. *Acta psychiatrica Scandinavica*. 2009; 119(5):350–64. [PubMed: 19298573]
25. 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(5):490–8. [PubMed: 23487258]
26. Hamilton M. A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*. 1960; 23:56–62.
27. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*. 1996; 1:130–49.
28. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*. 1980; 88:588–606.
29. Pop VJ, Komproe IH, van Son MJ. Characteristics of the Edinburgh Post Natal Depression Scale in the Netherlands. *Journal of Affective Disorders*. 1992; 26(2):105–10. [PubMed: 1447427]
30. 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; 118(1–3):101–12. [PubMed: 19275960]
31. Matthey S. Using the Edinburgh Postnatal Depression Scale to screen for anxiety disorders. *Depression and Anxiety*. 2008; 25(11):926–31. [PubMed: 18041072]
32. Austin MP, Hadzi-Pavlovic D, Priest SR, Reilly N, Wilhelm K, Saint K, et al. Depressive and anxiety disorders in the postpartum period: how prevalent are they and can we improve their detection? *Archives of Women's Mental Health*. 2010; 13(5):395–401.
33. Tuohy A, McVey C. Subscales measuring symptoms of non-specific depression, anhedonia, and anxiety in the Edinburgh Postnatal Depression Scale. *The British Journal of Clinical Psychology/the British Psychological Society*. 2008; 47(Pt 2):153–69.
34. Bowen A, Bowen R, Maslany G, Muhajarine N. Anxiety in a socially high-risk sample of pregnant women in Canada. *Canadian Journal of Psychiatry Revue*. 2008; 53(7):435–40.
35. Munk-Olsen T, Laursen TM, Pedersen CB, Mors O, Mortensen PB. New parents and mental disorders: a population-based register study. *JAMA: The Journal of the American Medical Association*. 2009; 296(21):2582–9.

36. Guedeney N, Fermanian J. Validation study of the French version of the Edinburgh Postnatal Depression Scale (EPDS): new results about use and psychometric properties. *European Psychiatry: the Journal of the Association of European Psychiatrists*. 1998; 13(2):83–9. [PubMed: 19698604]
37. Montazeri A, Torkan B, Omidvari S. The Edinburgh Postnatal Depression Scale (EPDS): translation and validation study of the Iranian version. *BMC Psychiatry*. 2007; 7:11. [PubMed: 17408479]
38. Bosanquet K, Bailey D, Gilbody S, Harden M, Manea L, Nutbrown S, McMillan D. Diagnostic accuracy of the Whooley questions for the identification of depression: a diagnostic meta-analysis. *British Medical Journal Open*. 2015; 5:e008913.

Highlights

- Investigating the extent to which the Edinburgh Postnatal Depression Scale (EPDS), a common self-report screen of postpartum depression, reflects the dimensions of depression and anxiety.
- Examining the internal factor consistency and internal reliability of the EPDS, our results shows that a seven-item one factor scale emerged, performing comparably to a ten-item two factor version of the EPDS in predicting scores of 18 or higher on a clinical interview.
- Results suggest that the seven-item one factor version of the EPDS is an efficient and effective measure of depression severity on par with the two factor ten-item version of the EPDS.

Figure 1.
Receiver Operating Curves for Scores Higher than 9 and 12 on the Seven-item EPDS.

Figure 2. Receiver Operating Curves for the Seven-Item and the 10-Item EPDS in Predicting Depression Diagnosis Using the SIGH-ADS.

Table 1

Item factor loadings of studies reporting factor analyses results of the EPDS

Study	Number	Factors	Sample	Factor 1 Depression	Factor 2 Anxiety	Factor 3 Anhedonia
Adouard et al., 2005	2		28–34 weeks pregnant, $n = 60$	1, 2, 7, 8, 9	3, 4, 5, 6, 10	
Asbury et al., 1994	2		English, 6–9 months postpartum, $n = 771$	1, 2, 5, 7, 8, 9, 10	3, 4, 5	
Berle et al., 2003	1		Norwegian, 6–12 weeks postpartum, $n = 411$	1–10		
Bowen et al., 2008	3		1 st prenatal visit, $n = 400$	1, 2, 8	3, 4, 5	10
Brouwers et al., 2001	3		Second trimester, $n = 197$	1, 2, 8	3, 4, 5	10
Chabrol & Teissedre, 2004	3		French, 2–3 days postpartum, $n = 299$	8, 9, 10	3, 4, 5, 6, 7	1, 2
Des Rivieres-Pigeon et al., 2000	1		French Canadian, 3–5 weeks postpartum, $n = 224$	1, 2, 6, 7, 8, 9, 10	3, 4, 5	
Guedeney & Fermanian, 1998	2		4 months postpartum, $n = 87$	3, 4, 5, 6, 7, 9	1, 2, 8, 10	
Jomeen & Martin, 2005	3		Late 1 st trimester/early 2 nd trimester, $n = 101$	1, 2, 6, 7, 8, 9	3, 4, 5	10
Jomeen & Martin 2007	3		English, early 3 rd trimester, $n = 148$	1, 2, 8	3, 4, 5	10
Lloyd-Williams et al., 2000	3		Terminal cancer patients	1, 2, 5	3, 4, 7	6, 8, 9, 10
Logsdon et al., 2009	2		Adolescents, postpartum 4–6 weeks, $n = 149$	1, 2, 8, 9, 10	3, 4, 5, 6, 7	
Matthey, 2008	2		6 weeks postpartum, $n = 238$	1, 2, 6, 7, 8, 9, 10	3, 4, 5	
Montazeri et al., 2007	3		Persian, 6–8 weeks postpartum, $n = 100$	3, 4, 5, 8	6, 7, 8, 9, 10	1, 2
Pallant, Miller, Tennant, 2006	2		Australian, 6 wks to 6 months postpartum, $n = 324$	1, 2, 3, 4, 6, 7, 9, 10	5, 8	
Phillips et al., 2009	2		Australian, inpatient, newly delivered to 12 months postnatal $n = 309$	1, 2, 6, 7, 8, 9, 10	3, 4, 5	
Pop et al., 1992	2		Dutch, 4 weeks postpartum, $n = 293$	1, 2, 7, 8, 9, 10	3, 4, 5, 6	
Reichenheim et al., 2011	1		Brazilian, five months or less postpartum, $n = 811$	1, 2, 6, 3, 4, 5, 6, 7, 8, 9, 10		
Ross et al., 2003	3		Canadian, 6 weeks postnatal, $n = 150$	1, 2, 8, 9	3, 4, 5	10
Small et al., 2007	3		Australian, 6–7 months postpartum, $n = 313$	3, 4, 5, 6, 7, 8, 9	1, 2, 10	
Teissedre & Chabrol, 2004	1		French, 4–6 weeks postpartum, $n = 722$	1–10		
Tuohy & Vey, 2008	3		3–9 months postpartum, $n = 400$	7, 8, 9, 10	1, 2	3, 4, 5

Table 2

Demographic Characteristics by Sample

Measure	Sample			Analysis	
	All	Exploratory (<i>n</i> = 5,055)	Confirmatory (<i>n</i> = 10,117)	Test statistic	<i>p</i>
Age	29.6 ± 5.55	29.7 ± 5.54	29.6 ± 5.56	$\chi^2(1) = 1.38$	0.1680
Race				$\chi^2(2) = 5.42$	0.0664
White	12,129 (80.0)	4,040 (80.0)	8,089 (80.0)		
Black	2,220 (14.6)	714 (14.1)	1,506 (14.9)		
Other	810 (5.3)	297 (5.9)	513 (5.1)		
Hispanic	317 (2.2)	115 (2.4)	202 (2.1)	$\chi^2(1) = 1.33$	0.2484
Education (level)				$\chi^2(4) = 1.48$	0.8308
<High school	638 (4.2)	222 (4.4)	416 (4.1)		
High school	2,228 (14.7)	723 (14.3)	1,505 (14.9)		
Some college	3,845 (25.4)	1,279 (25.3)	2,566 (25.4)		
College	4,680 (30.9)	1,566 (31.0)	3,114 (30.8)		
Graduate school	3,751 (24.8)	1,256 (24.9)	2,495 (24.7)		
Medical insurance				$\chi^2(2) = 8.15$	0.0170
Private	10,870 (72.1)	3,692 (73.5)	7,178 (71.4)		
Public	4,115 (27.3)	1,302 (25.9)	2,813 (28.0)		
None	94 (0.6)	27 (0.5)	67 (0.7)		
Marital status				TP(0) < 0.01	0.1260
Single	4,415 (29.8)	1,425 (28.9)	2,990 (30.2)		
Married/cohabiting	10,271 (69.3)	3,464 (70.1)	6,807 (68.8)		
Divorced/separated	137 (0.9)	49 (1.0)	88 (0.9)		
Widowed	5 (0.0)	49 (1.0)	5 (0.1)		
EPDS	5.18 ± 4.37	5.20 ± 4.36	5.16 ± 4.37	$\chi^2(1) = 0.24$	0.6227

Note: EPDS Edinburgh Postnatal Depression Scale; TP Fisher's exact test table probability

Table 3

Summary of Exploratory Factor Analysis Results for EPDS Measure Using Maximum Likelihood Estimation ($N = 5,055$)

#	Item	Factor Loadings	
		“Depression”	“Anxiety”
1	Unable to laugh	.84	–.01
2	Lack of enjoyment	.83	–.02
3	Blame myself unnecessarily	.15	.55
4	Been anxious or worried	.02	.85
5	Been scared or panicky	.09	.74
6	Things have been getting on top of me	.61	.17
7	Difficulty sleeping	.61	.23
8	Felt sad or miserable	.78	.11
9	So unhappy that I have been crying	.68	.15
10	Thought of harming self	.50	.32

Note: EPDS = Edinburgh Postnatal Depression Scale. Factor; Factor loadings over .4 appear in bold.

Table 4

Summary of the Fit Statistics from the One and Two-Factor Models Using 10 Items and a One-Factor Model Using Seven Items.

Statistic	10 Items		7 Items
	One Factor	Two Factors	One Factor
Null Model			
Factor Loadings χ^2	67571	67571	48854
Factor Loadings df	45	45	21
Factor Loadings p	<.0001	<.0001	<.0001
Hypothetical Model			
Factor Loadings χ^2	6037	2959	1340
Factor Loadings df	35	34	14
Factor Loadings p	<.0001	<.0001	<.0001
Standardized Root Mean	.0517	.0409	.0234
Square Residual			
Goodness of Fit Index (GFI)	.8875	.9448	.9627
Parsimonious GFI	.6902	.7138	.6418
Root Mean Square Error of Approximation	.1302	.0922	.0968
Bentler-Bonett Normed Fit Index	.9107	.9562	.9726