
Judith L Meijer
Chantal Beijers
Mariëlle G van Pampus
Tjitte Verbeek
Ronald P Stolk
Jeannette Milgrom
Claudi LH Bockting
Huibert Burger

BJOG. 2014 Dec;121(13):1604-10.

chapter 4
Abstract

Objective

To investigate whether the ten-item Edinburgh Postnatal Depression Scale (EPDS) administered antenatally in 1,620 women in the general population is accurate in predicting postpartum depressive symptoms (EPDS score ≥10), and whether a two-item EPDS has similar predictive accuracy.

Methods

Mean values, area under the receiver operating characteristics-curve (AUC), sensitivity, specificity and predictive values of antenatal EPDS for the likelihood of developing postpartum depressive symptoms were calculated. Analyses were repeated for each trimester, several cut-off values and a two-item EPDS (low mood and anhedonia).

Results

Mean EPDS scores were significantly higher during each trimester in women with postpartum depressive symptoms, compared to those without (p<0.001). Using the prevailing cut-off (≥10), the AUC was reasonable (0.74), sensitivity was 26.3%, positive predictive value 25.9%, specificity 93.0% and negative predictive value 93.1%. Using a lower cut-off value (≥5), sensitivity was 70.8% and specificity was 65.4%, but positive predictive value was low (15.9%). Negative predictive value was exceedingly high with 96.0%. Results were similar during the second and third trimester. The predictive accuracy of the two-item EPDS appeared inferior.

Conclusions

The EPDS was not sufficiently accurate in predicting risk of postpartum depressive symptoms. Nevertheless, when using the ≥5 cut-off value, it may be adequate for initial screening, followed by further assessments and possibly antenatal intervention when positive. Furthermore, when negative, women may be reassured that postpartum depressive symptoms are unlikely. A two-item version showed poor predictive accuracy.
Introduction

Depression affects 8-15% of women in the first year postpartum\(^4\). This poses a considerable burden to the women, their families and society\(^5,6\). Children born to women who experienced postpartum depression, are at risk of insecure attachment, which in turn is associated with cognitive, behavioral and emotional problems\(^7,8\). Moreover, postpartum depression is commonly preceded by antenatal depression\(^4,9,10\), which has been associated with unfavourable obstetric outcomes and impaired child neurodevelopment\(^10-12\), independent of effects due to postpartum depression\(^11\). Thus, timely detection of antenatal depressive symptoms is critical for mother and child. Correspondingly, the American College of Obstetricians and Gynecologists recommends antenatal screening\(^13\).

The ten-item Edinburgh Postnatal Depression Scale (EPDS)\(^14,15\) is commonly used for assessing depressive symptoms, and was validated as antenatal and postpartum screener for minor or major depression\(^14\). While depressive symptoms are known to be associated with the development of a depressive disorder\(^17\), it is remarkable that predicting the absolute risk of postpartum depression, based on antenatal depressive symptoms as assessed by the EPDS, has been studied scarcely\(^10,18,19\). In a large cohort, the relative importance of several antenatal risk factors for postpartum depression was studied\(^18\). Yet, the predictive accuracy of the antenatal EPDS was quantified exclusively for a single cut-off value and not specified for trimester of pregnancy. In another large cohort, the pattern of depressive symptoms during pregnancy and in the postpartum period was studied, but not the predictive accuracy of the EPDS\(^10\). One study observed that when using a higher than commonly used cut-off in the second trimester, the EPDS was accurate in predicting depression at six weeks postpartum\(^19\). However, their study suffered from a high loss to follow-up rate.

Consequently, the predictive accuracy of the EPDS for identifying an increased risk of postpartum depression is still unknown. At the same time, a brief, reliable self-report antenatal screener to predict the likelihood of minor or major postpartum depression would facilitate midwives and gynaecologists in daily clinical practice\(^20\), for example a screener that only includes the two key symptoms of depression, i.e. anhedonia and mood\(^21\).

In the current prospective study, we examined to what degree the ten-item EPDS administered during the first, second and third trimester of pregnancy, using various cut-offs, is valid for the assessment of the absolute risk of postpartum depressive symptoms. Since a two-item version would be more convenient, we further investigated the predictive performance of a two-item EPDS including the key symptoms of depression only.
Methods

Sample

The present study was carried out within the ongoing population-based Pregnancy, Anxiety and Depression (PAD) Study. This prospective cohort study has been set up to investigate symptoms of and risk factors for anxious or depressive symptoms during pregnancy and the first half year postpartum. All pregnant women in their first trimester of pregnancy, visiting a total of 109 collaborating primary obstetric care centers and seven hospitals in the Netherlands, are invited to participate. Written informed consent is obtained from each participant. After baseline assessments at the end of the first trimester, follow-up assessments using online questionnaires take place at the end of the second and third trimesters of pregnancy, as well as at six months postpartum.

Data used for the present study was collected from May 2010 to September 2012. By the end of that period, 4,157 women had completed the baseline assessments. The eligible population for the present analysis consisted of women who were six months postpartum by the time of the database closure (N=1,620, 39%). Out of these, 1,276 responded by filling out the questionnaires yielding a response rate of 78.8%. Non-responders did not significantly differ from responders on parity, educational level or antenatal depressive symptoms. However, non-responders were significantly younger (30 vs. 29 years, p<0.02).

Measurements

Demographic and pregnancy related variables included in the present study were age, parity and educational level, and were assessed at baseline. Educational level was defined as the highest completed education, and was divided into five categories; elementary education, lower and higher tracts of secondary education, higher vocational education and university education.

The Dutch version of the ten-item Edinburgh Postnatal Depression Scale (EPDS) was used to measure symptoms of depression. This version of the EPDS has shown good internal validity with a Cronbach’s alpha of 0.8214. The baseline EPDS was administered as part of the initial antenatal screening and each follow-up assessment wave. Total EPDS scores of ≥10 indicate depressive symptoms at a level corresponding to an increased risk of minor or major depression. The outcome of this study, i.e. postpartum depressive symptoms at six months after giving birth, was defined accordingly.
Multiple imputation of missing data

Complete information on all variables was obtained on 1013 (62.5%) women. The percentage of missing data ranged from 1.5 (EPDS at baseline) to 21.1 (postpartum EPDS) for the variables of main interest. Under the assumption that missing values were missing at random (MAR) or missing completely at random, we imputed missing data using multiple imputation by chained equations. This was done to avoid the potential bias and decreased statistical power associated with complete case analysis23,24. We substantiated our assumption of the missing data mechanism being MAR by, for each variable separately, predicting missingness using logistic regression. Independent variables were all variables that were considered potential predictors of missingness24 i.e. antenatal anxiety, age at baseline, parity and maternal educational level. The explained variance (Nagelkerke R2) from these analyses ranged from 7.1% to 63.2%, suggesting that at least in part data are MAR and imputation of missing values would likely improve the validity of our analyses. Multiple data sets (N=5) were generated to account for the uncertainty in imputed data.

Data analyses

Descriptive statistics for demographic variables were calculated according to the likelihood of the presence of postpartum depressive symptoms. We tested whether the mean value of each antenatal EPDS differed between women experiencing postpartum depressive symptoms and those who did not, using an independent samples t-test. Additionally, we calculated the prevalence of antenatal depressive symptoms (EPDS ≥10) at each assessment.

Subsequently, we evaluated the predictive accuracy of each antenatal EPDS in terms of sensitivity, specificity, predictive values and overall discriminatory power. We first constructed receiver operating characteristic (ROC) curves, by plotting the sensitivity against the 1-specificity for all possible cut-off values of antenatal EPDS. Informed by these curves, we constructed crosstabs using all possible cut-off values between 5 and 10, and calculated the measures of predictive accuracy. The overall discriminatory power of the EPDS independent of cut-off value was quantified as the area under the ROC-curve (AUC). The AUC can be interpreted as the extent to which the EPDS separates women with postpartum depressive symptoms from those without. Values of 0.70 to 0.80 indicate a reasonable AUC, 0.80 to 0.90 a good one and values of 0.90 or above are interpreted as excellent25.

The above analyses were repeated for the EPDS scores recorded in the second and third trimester. In addition, we quantified the prevalence and predictive accuracy of exceeding the
≥10 cut-off at all three antenatal assessments.

Subsequently, we examined the predictive accuracy of the two EPDS items that refer to the key symptoms of depression. These are items 2 for anhedonia and 8 for depressed mood, according to the Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV)\textsuperscript{21}. We summed the ratings of these items to a total score. This two-item version ranged from 0-6, the first cut-off value we tested was ≥3. Based on the results, the ROC-curve and the fact that, according to the DSM-IV, a diagnosis of depression needs at least one of these two items to be answered positively, we subsequently tested a cut-off value of ≥1.

Measures of predictive accuracy were supplied with 95% confidence intervals. Results were pooled using Rubin’s method for multiple imputation inference\textsuperscript{26}. Multiple imputation and statistical analyses were carried out using IBM SPSS Statistics 20.

**Results**

**Descriptives**

Mean gestational age of the study population (N=1,620) was 14, 23 and 34 weeks at the baseline, second and third trimester measurement waves, respectively. Women were five (range 4-7) months postpartum at the final measurement. At baseline, 139 women (8.6%) experienced depressive symptoms according to the cut-off value of ≥10. These figures were 154 (9.5%) and 164 (10.1%) at the second and third trimester measurements, respectively. One hundred and thirty-seven women (8.5%) experienced postpartum depressive symptoms, which is equal to their a priori risk.

Table 1 shows that demographic factors were equally distributed among women with and without postpartum depressive symptoms. At each instance of administration, scores on the EPDS were statistically significantly higher in the group with postpartum depressive symptoms than in the group without postpartum symptoms of depression (p<0.001).

**Crosstabs**

Table 2 shows parameters of predictive accuracy of the ten-item EPDS in the first trimester using cut-off values ≥10 and ≥5, as well as those for the two-item version using cut-off values ≥3 and ≥1. Although for the ten-item EPDS we tested all cut-off levels between those presented in table 2, we decided to show only the most clinically relevant results.
Table 1  Demographic characteristics of the study population (n=1,620) according to the presence of postpartum depressive symptoms

<table>
<thead>
<tr>
<th></th>
<th>No postpartum depressive symptoms * (n=1,483)</th>
<th>Postpartum depressive symptoms * (n=137)</th>
<th>Fraction of missing data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (min-max)</td>
<td>30.0 (17-46)</td>
<td>29.7 (18-43)</td>
<td>4.1%</td>
</tr>
<tr>
<td>Primiparae, N (%)</td>
<td>625 (41%)</td>
<td>52 (38%)</td>
<td>29.8%</td>
</tr>
<tr>
<td>Educational level, N (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary education</td>
<td>7 (&lt;1%)</td>
<td>1 (&lt;1%)</td>
<td></td>
</tr>
<tr>
<td>Lower tracts of secondary education</td>
<td>139 (9%)</td>
<td>15 (11%)</td>
<td></td>
</tr>
<tr>
<td>Higher tracts of secondary education</td>
<td>508 (34%)</td>
<td>46 (34%)</td>
<td></td>
</tr>
<tr>
<td>Higher vocational education</td>
<td>596 (40%)</td>
<td>57 (42%)</td>
<td></td>
</tr>
<tr>
<td>University education</td>
<td>233 (16%)</td>
<td>18 (13%)</td>
<td></td>
</tr>
<tr>
<td>Depression level first trimester b, mean (SD)</td>
<td>3.91 (3.32)</td>
<td>7.48 (4.86)</td>
<td>1.5%</td>
</tr>
<tr>
<td>Depression level second trimester b, mean (SD)</td>
<td>4.62 (3.14)</td>
<td>8.79 (4.93)</td>
<td>18.2%</td>
</tr>
<tr>
<td>Depression level third trimester b, mean (SD)</td>
<td>4.67 (3.17)</td>
<td>8.83 (4.58)</td>
<td>16.7%</td>
</tr>
<tr>
<td>Depression level postpartum b, mean (SD)</td>
<td>3.81 (2.52)</td>
<td>13.36 (3.75)</td>
<td>21.2%</td>
</tr>
</tbody>
</table>

* According to Edinburgh Postnatal Depression Scale (EPDS), cut-off EPDS ≥10, min-max=0-30

b Based on EPDS at approximately 14 weeks pregnancy, min-max=0-30. Scores were statistically significantly higher in the group with postpartum depressive symptoms than in the group without postpartum symptoms of depression, with p ≤0.001

The ten-item EPDS using the ≥10 cut-off showed low sensitivity (26.3%) and high specificity (93.0%). However, the AUC of 0.74 was reasonable (figure 1). When we decreased the cut-off value to ≥5, the sensitivity increased to 70.8% and the specificity decreased to 65.4%. Positive predictive values were low at all considered cut-off values, whereas all negative predictive values were high. Although the positive predictive value for a score higher than 10 in the first trimester was only 25.9%, this risk is three times higher as the prior risk of postpartum depressive symptoms, which was 8.5%. Lowering the cut-off for the two-item version also showed higher sensitivity (59.1% vs. 12.4%) at the expense of the specificity and the positive predictive value (table 3). The AUC of the two-item version, i.e. 0.63, was considerably lower than that of the complete ten-item version.
Figure 1  Area Under the ROC Curve for baseline score Edinburgh Postnatal Depression Scale (EPDS) in the prediction of minor or major postpartum depression

When test positive was defined as a score of ≥10 at all three measurements (N=50; 3.1%), the positive predictive value increased to 40.0%, while the sensitivity decreased to 15.0%. The negative predictive value and the specificity remained high (92.5% and 97.9%, respectively).

All of the above analyses were repeated for the EPDS scores recorded in the second and third trimester. Results were essentially similar to those obtained using the first trimester scores. However, AUC’s were slightly higher for the ten-item EPDS with 0.77 and 0.78 at the second and third trimester, respectively. For the two-item version the AUC was 0.68 at the second trimester and 0.71 at the third.
<table>
<thead>
<tr>
<th>Cut-off ≥13</th>
<th>(n=55, 3.4%)</th>
<th>Cut-off ≥10</th>
<th>(n=137, 8.5%)</th>
<th>Cut-off ≥5</th>
<th>(n=610, 37.7%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity, % (95% CI)</td>
<td>16.8 (11.0-24.1)</td>
<td>26.3 (19.1-34.5)</td>
<td>70.8 (62.4-78.3)</td>
<td>14.6 (9.2-21.6)</td>
<td></td>
</tr>
<tr>
<td>Specificity, % (95% CI)</td>
<td>97.8 (97.0-98.5)</td>
<td>93.0 (91.6-94.2)</td>
<td>65.4 (62.9-67.8)</td>
<td>97.9 (97.0-98.6)</td>
<td></td>
</tr>
<tr>
<td>Positive Predictive Value, % (95% CI)</td>
<td>41.8 (28.7-55.9)</td>
<td>25.9 (18.8-34.0)</td>
<td>15.9 (13.1-19.0)</td>
<td>39.2 (25.8-53.9)</td>
<td></td>
</tr>
<tr>
<td>Negative Predictive Value, % (95% CI)</td>
<td>92.7 (91.3-94.0)</td>
<td>93.1 (91.7-94.3)</td>
<td>96.0 (94.6-97.2)</td>
<td>92.5 (91.1-93.8)</td>
<td></td>
</tr>
<tr>
<td>Area Under the Curve</td>
<td>0.74--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Test characteristics for predicting postpartum depressive symptoms from antenatal depressive symptoms.

Number and percentage of women scoring above the used cut-off at all antenatal measurements of the EPDS.

Based on ten-item Edinburgh Postnatal Depression Scale (EPDS) at approximately 14 weeks pregnancy.
Table 3  Test characteristics for predicting postpartum depressive symptoms from a two-item version for antenatal depressive symptoms

<table>
<thead>
<tr>
<th></th>
<th>Cut-off ≥3</th>
<th>Cut-off ≥1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=65, 4%)</td>
<td>(n=586, 36.2%)</td>
</tr>
<tr>
<td>Sensitivity, % (95% CI)</td>
<td>12.4 (7.4-19.1)</td>
<td>59.1 (50.4-67.4)</td>
</tr>
<tr>
<td>Specificity, % (95% CI)</td>
<td>96.8 (95.7-97.6)</td>
<td>65.9 (63.5-68.4)</td>
</tr>
<tr>
<td>Positive Predictive Value, % (95% CI)</td>
<td>26.2 (16.0-38.5)</td>
<td>13.8 (11.1-16.9)</td>
</tr>
<tr>
<td>Negative Predictive Value, % (95% CI)</td>
<td>92.3 (90.8-93.6)</td>
<td>94.6 (93.0-95.9)</td>
</tr>
<tr>
<td>Area Under the Curve</td>
<td>0.63</td>
<td>-</td>
</tr>
</tbody>
</table>

a  Based on items 2 and 8 of the Edinburgh Postnatal Depression Scale (EPDS) at approximately 14 weeks pregnancy, min-max=0-6
b  Number and percentage of women scoring above the used cut-off at baseline EPDS

Discussion

Main Findings

In this prospective cohort study, we found that women who developed postpartum depressive symptoms had statistically significantly higher antenatal EPDS scores. Irrespective of trimester, the ten-item EPDS had reasonable overall discriminatory power, but low sensitivity at the prevailing cut-off (≥10). When using a lower cut-off (≥5), both sensitivity and specificity were around 70%. However, for both cut-off values, positive predictive values were low, whereas negative predictive values were exceedingly high. A two-item version of the EPDS, consisting of merely the two key symptoms of depression, demonstrated a poor predictive performance. In addition, the EPDS was accurate in detecting current depressive symptoms. In the context of potential hazardous child effects of antenatal depressive symptoms, it is important that these symptoms are acknowledged by screening health care professionals.

Strengths and limitations

The present study is not without limitations. First, the EPDS is a self-report questionnaire and even though it is commonly used, based on the EPDS score we are merely able to identify women with a higher risk of depressive symptoms. However, almost 70% of women exceeding the threshold of 10 satisfy the criteria for a diagnosis of depression. Second, anxiety and depression are known to be co-morbid and several studies have shown that...
antenatal anxiety is a predictor of postpartum depression\textsuperscript{32,33}. However, because the only aim was to test the predictive accuracy of the EPDS, anxiety is considered beyond the scope of the current paper, and should therefore be addressed in future research. Finally, although our antenatal prevalence rates (9-10\%) are comparable with other studies\textsuperscript{1,2}, women experiencing depressive symptoms in their first trimester might have been more inclined to participate than women who are not familiar with depressive symptoms. In addition, we found a significant difference in age between the respondents and non-respondents. However, in absolute numbers the age difference is less than a year; therefore we do not think it affected our prevalence rates. Likewise, the prevalence of postpartum depressive symptoms in our study was relatively low (8\%). Previous studies reported prevalence rates of 8-15\%\textsuperscript{6,13}. However, in these studies, the time at which postpartum depressive symptoms were assessed ranged from a few days after delivery to two years postpartum, which may have caused the inclusion of maternity blues, thereby falsely increasing the prevalence of postpartum depression.

Our study also has several strengths. The present study is the first to date to examine the EPDS as a tool for antenatal assessment of the risk of postpartum depressive symptoms, during each trimester of pregnancy, evaluating several cut-off values for the sum score, as well as the possibility of item reduction. In addition, the present sample is a large prospective cohort, including 1,620 pregnant women in their first trimester, with a response rate of 78.8\%.

**Interpretation**

Despite reasonable overall discriminatory power, the EPDS appears to have limited predictive accuracy for absolute risk stratification, regardless of the cut-off value used or the trimester of administration, partly contradicting the findings of Lau et al.\textsuperscript{19}. They concluded that when using a cut-off of ≥15, the EPDS during the second trimester of pregnancy appeared to be an accurate predictor of depression at six weeks postpartum, with parameters of predictive accuracy ranging from 63\% to 85\%, and an overall AUC of 0.76, the latter being in line with our findings. However, they suffered from a high loss to follow-up rate (72\%), making the results susceptible to serious bias. From a more general point of view, Austin & Lumley concluded that no currently available antenatal screener is sufficient in predicting postpartum depression, due to the evidence that multiple risk factors are postpartum, which clearly cannot be included in antenatal screening\textsuperscript{27}. Additionally, Milgrom et al. showed that adding prior history of depression and low partner support to antenatal EPDS yielded a
more accurate overall prediction, using a cut-off of >12, mainly administered during the third trimester\(^{18}\). They observed a low positive predictive value of 29\%, which seems compatible with our findings; although the absolute risk of postpartum depressive symptoms for women scoring ≥5 was two times higher than the prior risk, it was still low (15.9\%). In order to allocate interventions as much as possible to women who have a high risk of postpartum depressive symptoms, a two-stage screening may therefore be advocated, as previously suggested\(^{20}\). Women with an EPDS score of ≥5 during the first trimester can be screened more elaborately on well known risk factors, such as history of depression and low partner support, to decide whether an antenatal intervention is necessary.

In suggesting a two-stage screening, the exceedingly high negative predictive value of the ten-item EPDS seems important. Women who scored <5 at the antenatal screening had a very low risk of postpartum depressive symptoms (4.0\%). As previously suggested\(^{29}\), it could therefore be used to exclude women who are not at risk for depressive symptoms. However, due to the lowered cut-off, considerably more women were identified as having a high risk of postpartum depressive symptoms, comparable to the number of women identified with the two-item version. In that context, it could be argued that, as an initial screener, the two-item version would be more facilitating in daily clinical practice. The National Institute for Health and Clinical Excellence (NICE) recommends a two-item dichotomous screening tool based on the key symptoms of depression\(^{30}\). However, this tool was developed with the only purpose of detecting current depressive symptoms\(^{20}\), in which it appears to be accurate\(^{20,28,29}\). In the present study, the two-item version also showed satisfactory predictive accuracy in the identification of current depressive symptoms, but the overall test parameters for predicting postpartum depressive symptoms are considerably lower. Therefore, in screening for high risk on postpartum depressive symptoms, we would suggest the use of the ten-item EPDS with a cut-off of ≥5.

Our study results do not suggest that timing of administering the EPDS during pregnancy affects its predictive accuracy. This is an important notion in view of the potential hazardous child effects of antenatal depressive symptoms, as this means that postponing screening for depressive symptoms cannot be justified by the argument that the predictive accuracy increases with duration of pregnancy.
Conclusion

In conclusion, we found that in no trimester of pregnancy the EPDS is a sufficiently accurate instrument to identify women with a high risk of symptoms of postpartum depression. However, when using a lower than commonly used cut-off value, the test characteristics were such that a stepwise approach using the EPDS as a first screening step could be recommended to clinicians, midwives and gynaecologists. Women with a high risk (i.e. EPDS score of ≥5 during the first trimester) can subsequently be screened more elaborately on well known risk factors, such as a history of depression and low partner support, to decide whether an antenatal intervention is necessary. Women scoring <5 on the antenatal ten-item EPDS can be reassured that it is very unlikely that they will develop a postpartum depressive symptoms. Although it would be convenient to use a brief screener in daily clinical practice, a two-item version showed poor predictive accuracy. Because the predictive accuracy of the ten-item EPDS was similar across the three antenatal administrations, there seems to be no reason to postpone screening and possible intervention until after the first trimester.
References

27. Austin MP, Lumley J. Antenatal screening for


