Abstract

Introduction: Postpartum depression is a non-psychotic depressive episode with serious repercussions on the bond between the mother and her child, hence the importance of detecting it in a timely manner. **Objective:** To determine the accuracy of the Edinburgh Postnatal Depression Scale as a diagnostic test and to analyze the consequences of screening and the probability of depression after applying the test. **Method:** Screening of 411 women with the Edinburgh Postnatal Depression Scale during the postpartum period; Beck’s Depression Inventory was used as reference. **Results:** At a cutoff point of 12, a sensitivity of 70.4 %, specificity of 72.2 %, positive predictive value of 36.9 % and negative predictive value of 91.4 % were obtained with Edinburgh Postnatal Depression Scale, as well as an area under the curve of 0.729 and a p-value of 0.0003. Out of 49 women without treatment for postpartum depression, five were identified to require it. **Conclusions:** The Edinburgh Postnatal Depression Scale has moderate accuracy; its application is simple, accessible and should be routine. It is necessary for strategies to detect and treat postpartum depression to be implemented in Mexico.

**KEY WORDS:** Postpartum depression. Edinburgh Postnatal Depression Scale. Sensitivity. Specificity. Screening. Diagnostic accuracy.

Diagnostic accuracy of the Edinburgh Postnatal Depression Scale: consequences of screening in Mexican women

Emma del C. Macías-Cortés,¹* Virgilio Lima-Gómez¹ and Juan Asbun-Bojalil²

¹Secretaría de Salud, Hospital Juárez de México; ²Instituto Politécnico Nacional, Higher School of Medicine, Postgraduate Education Section. Mexico City, Mexico

**Correspondence:**

*Emma del C. Macías-Cortés

E-mail: emc2008@hotmail.es

0016-3813© 2019 Academia Nacional de Medicina de México, A.C.. Published by Permanyer. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Introduction

Postpartum depression (PPD) is a mild to moderate non-psychotic depressive episode that begins after delivery and extends throughout the first postpartum year. Its global prevalence is 10 to 20 %, and in Mexico, from 13 to 24 %. It usually resolves spontaneously in three to six months, although in 25 % of cases it can last more than one year. It is important for PPD to be diagnosed and treated due to the deep emotional impact on the mother-child bond; children of depressed mothers have been observed to have three times more risk to develop emotional and behavioral problems.

It is essential for health personnel to become aware of the importance of identifying PPD with rapid, inexpensive and simple tests. In 2016, Place et al. stated that there are few publications on the prevention, detection and treatment of PPD in Mexico. PPD is mentioned in various health policies, but protocols to routinely detect and treat it in obstetric units have not been implemented. Furthermore, even when screening tests are applied, the results are not always used for follow-up and treatment; only 23 to 60 % of women with a positive screening test follow an adequate treatment.

The most widely used, recommended and internationally validated test for PPD screening is the Edinburgh Postnatal Depression Scale (EPDS); Beck’s Depression Inventory (BDI) is an internationally validated scale to determine the severity of depression. Studies on the validity of EPDS as a diagnostic test show differences in sensitivity, specificity and predictive values, in the type of population, sample size, cutoff points and diagnostic instruments used as reference. In Mexico, there are only few reports on screening with EPDS. Alvarado et al. recommend a cutoff point of 11/12 (sensitivity of 75 % and specificity of 93 %) at four weeks postpartum. However, no reports have been published on the consequences of these results in clinical practice, and neither has an analysis of EPDS performance and diagnostic accuracy.

The main purpose of this investigation was to determine the accuracy of EPDS as a diagnostic test, with BDI being used as the reference method; sensitivity, specificity, predictive values, ROC curve and Youden index were calculated at different cutoff points, and the consequences of screening and the probability of PPD after applying the test were analyzed.

Method

Observational, cross-sectional, descriptive study, carried out in the outpatient clinic of Hospital Juárez de México, Ministry of Health, from February to November 2009, registered with number HJM1411/07.10.09.

Women aged between 15 and 45 years, with a minimum education of fifth year of primary school, whose delivery had occurred at between 15 days and eight weeks before the application of the EPDS, and who verbally granted their consent after being informed on the purposes and procedures of the research, were included. Women with known thyroid disease, and those whose children had congenital malformations or disease requiring medical attention were excluded. Women who met the selection criteria answered the EPDS questionnaire, which explores the mood of the woman in the previous seven days with 10 questions, each question with a score from 0 to 3 (highest total score of 30 points); a score ≥ 10 was considered to be indicative of PPD. Subsequently, those who obtained an EPDS score ≥ 10 were asked to answer the BDI questions which is a self-applied 21-question questionnaire whose 0-to-63 score indicates the severity of depression: from 5 to 9 points, minimal depression; 10 to 18 points, mild to moderate depression; from 19 to 29 points, moderate to severe depression; from 30 to 63 points, severe depression.

Frequencies and percentages of demographic characteristics and and central tendency and dispersion measures of the scale scores were obtained. Prevalence, sensitivity, specificity, predictive values, and likelihood ratios (LR or post-test probabilities) were calculated with 95 % confidence intervals (CI), whereby the MedCalc program, version 19.0.3, was used. For the BDI, a score of 10 was used as reference. Sensitivity was defined as the percentage of individuals with the disease who have a positive test, and specificity, as the percentage of individuals without the disease who have a negative test. The positive predictive value (PPV) is the probability that a patient with a positive result (abnormal) has the disease; the negative predictive value (NPV) is the probability of not having the disease when the test result is negative (normal). According to reports in the literature, for the calculation of predictive values, a PPD prevalence (prior probability) of 18.7 % was considered.

LRs indicate how much the result of a diagnostic test will increase or decrease pretest probability. The following criteria were considered:

- higher than 10 or lower than 0.1, significant changes.
- 5-10 or 0.1-0.2, moderate changes.
Table 1. Demographic characteristics of women with suspected postpartum depression who obtained ≥ 10 points on the Edinburgh Postnatal Depression Scale

<table>
<thead>
<tr>
<th>Marital status</th>
<th>Level of education</th>
<th>Occupation</th>
<th>Deliveries</th>
<th>Addictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Married</td>
<td>Primary school</td>
<td>Homemaking</td>
<td>One</td>
<td>No</td>
</tr>
<tr>
<td>44.8 % (n = 37)</td>
<td>12.7 % (n = 10)</td>
<td>89.9 % (n = 71)</td>
<td>44.3 % (n = 35)</td>
<td>91.1 % (n = 72)</td>
</tr>
<tr>
<td>Cohabitating</td>
<td>Secondary school</td>
<td>Work away from home</td>
<td>Two</td>
<td>Yes, tobacco</td>
</tr>
<tr>
<td>40 % (n = 32)</td>
<td>60.8 % (n = 48)</td>
<td>7.6 % (n = 6)</td>
<td>30.4 % (n = 24)</td>
<td>3.8 % (n = 3)</td>
</tr>
<tr>
<td>Single</td>
<td>High school</td>
<td>Students</td>
<td>Three</td>
<td>Yes, alcohol</td>
</tr>
<tr>
<td>12.7 % (n = 10)</td>
<td>5.1 % (n = 4)</td>
<td>2.5 % (n = 2)</td>
<td>15.2 % (n = 12)</td>
<td>3.8 % (n = 3)</td>
</tr>
<tr>
<td>Widow/divorced</td>
<td>College degree</td>
<td></td>
<td>Four</td>
<td>Yes, drugs</td>
</tr>
<tr>
<td>0 % (n = 0)</td>
<td>0 % (n = 0)</td>
<td></td>
<td>3.8 % (3)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Postpartum depression severity in women with ≥ 10 points on Edinburgh Postpartum Depression Scale

<table>
<thead>
<tr>
<th>BDI score</th>
<th>Frequency (n = 79)</th>
<th>PPD severity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-9</td>
<td>22.8</td>
<td>18 Minimal</td>
</tr>
<tr>
<td>10-18</td>
<td>39.2</td>
<td>31 Mild to moderate</td>
</tr>
<tr>
<td>19-28</td>
<td>24</td>
<td>19 Moderate to severe</td>
</tr>
<tr>
<td>29-63</td>
<td>13.9</td>
<td>11 Severe</td>
</tr>
</tbody>
</table>

Beck’s Depression Inventory was used as reference method.

Results

The EPDS was applied to 441 women; 17.9 % (n = 79) obtained a score ≥ 10. Table 1 describes the sociodemographic characteristics. Average age was 26 ± 7 years, and most were married women (44.8 %), primigravids (44.3 %), had secondary education (60.8 %) and were exclusively dedicated to household chores (89.9 %). Average number of days after delivery in which the test was positive was 28 ± 13.

Average EPDS score was 5.7 ± 5; 82.1 % (n = 362) obtained less than 10 points. The 79 women with a score ≥ 10 had an average of 14.58 ± 4.08 points; in that group, BDI average score was 17 ± 11.1. Despite having an EPDS score of ≥ 10, 22.8 % (18/79) had no data consistent with PPD. Mild to moderate PPD was the most common (39.2 %), followed by moderate to severe (24 %), and only 13.9 % had data consistent with severe PPD (Table 2).

Table 3 shows how EPDS validity and safety were modified according to different cutoff points:

- Sensitivity: from 93.4% at a cutoff point > 10, it decreased to 70.4 % at a cutoff point > 12 (95 % CI = 84.1-98.2 and 57.4-81.5, respectively).
- Specificity: from 16.6 % with a cutoff point of 10, it increased to 94.4 % when the cutoff point was 17; specificity was 72.2 % when the cutoff point was > 12 (95 % CI = 72.7-99.9 and 46.5-90.3, respectively).
- PPV: from 20.5 % when the score was > 10, it increased to 36.9 % with a cutoff point > 12 (95 % CI = 17.2-24.3 and 21.5–55.6, respectively).
- NPV: from 91.7 % when the score was < 10, it went to 91.4 % when it was < 12 (95 % CI = 73.1-97.8 and 86.8-94.5, respectively).

Considering a cutoff point of 12 for EPDS and 10 for BDI, positive LR was 2.5 and negative LR 0.4 (95 % CI = 1.2-5.4 and 0.3-0.7, respectively), which represent a small, but significant change.
Table 3. Validity and safety values and likelihood ratios according to the cutoff point on Edinburgh Postnatal Depression Scale and Beck’s Depression Inventory as reference method

<table>
<thead>
<tr>
<th>EPDS cutoff point</th>
<th>Sensitivity % (95 % CI)</th>
<th>Specificity % (95 % CI)</th>
<th>Positive likelihood ratio (95 % CI)</th>
<th>Negative likelihood ratio (95 % CI)</th>
<th>PPV % (95 % CI)</th>
<th>NPV % (95 % CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥ 10</td>
<td>100 (94.1-100)</td>
<td>0.00 (0.0-18.5)</td>
<td>1.0 (1.0-1.0)</td>
<td>—</td>
<td>18.7 (18.7-18.7)</td>
<td>—</td>
</tr>
<tr>
<td>&gt; 10</td>
<td>93.4 (84.1-98.2)</td>
<td>16.6 (3.6-41.4)</td>
<td>1.1 (0.9-1.4)</td>
<td>0.3 (0.1-1.6)</td>
<td>20.5 (17.2-24.3)</td>
<td>91.7 (73.1-97.8)</td>
</tr>
<tr>
<td>&gt; 11</td>
<td>80.3 (68.2-89.4)</td>
<td>38.8 (17.3-64.3)</td>
<td>1.3 (0.9-1.9)</td>
<td>0.5 (0.2-1.1)</td>
<td>23.2 (17.0-30.8)</td>
<td>89.6 (79.9-94.9)</td>
</tr>
<tr>
<td>&gt; 12</td>
<td>70.4 (57.4-81.5)</td>
<td>72.2 (46.5-90.3)</td>
<td>2.5 (12.5-4.4)</td>
<td>0.4 (0.3-0.7)</td>
<td>36.9 (21.4-55.6)</td>
<td>91.4 (86.8-94.5)</td>
</tr>
<tr>
<td>&gt; 13</td>
<td>57.3 (44.1-70)</td>
<td>83.3 (58.6-96.4)</td>
<td>3.4 (1.2-9.9)</td>
<td>0.5 (0.4-0.7)</td>
<td>44.2 (21.6-69.5)</td>
<td>89.5 (85.6-92.4)</td>
</tr>
<tr>
<td>&gt; 14</td>
<td>47.5 (34.6-60.7)</td>
<td>83.3 (58.6-96.4)</td>
<td>2.8 (1.0-8.3)</td>
<td>0.6 (0.5-0.9)</td>
<td>39.6 (18.4-65.6)</td>
<td>87.4 (83.4-90.5)</td>
</tr>
<tr>
<td>&gt; 15</td>
<td>37.7 (25.6-51)</td>
<td>88.8 (65.3-98.6)</td>
<td>3.3 (0.9-13.0)</td>
<td>0.7 (0.5-0.9)</td>
<td>43.8 (16.9-75.0)</td>
<td>86.1 (82.8-88.9)</td>
</tr>
<tr>
<td>&gt; 16</td>
<td>31.1 (19.9-44.3)</td>
<td>88.8 (65.3-98.6)</td>
<td>2.8 (0.7-10.9)</td>
<td>0.7 (0.6-1.0)</td>
<td>39.2 (14.2-71.5)</td>
<td>84.9 (81.6-87.7)</td>
</tr>
<tr>
<td>&gt; 17</td>
<td>24.5 (14.5-37.3)</td>
<td>94.4 (72.7-99.9)</td>
<td>4.4 (0.6-31.3)</td>
<td>0.8 (0.7-1.0)</td>
<td>50.4 (12.6-87.8)</td>
<td>84.5 (81.9-86.7)</td>
</tr>
<tr>
<td>&gt; 18</td>
<td>21.3 (11.9-33.7)</td>
<td>100 (81.5-100)</td>
<td>—</td>
<td>0.7 (0.7-0.9)</td>
<td>100 (82.9-88.3)</td>
<td>—</td>
</tr>
<tr>
<td>&gt; 19</td>
<td>13.1 (5.8-24.2)</td>
<td>100 (81.5-100)</td>
<td>—</td>
<td>0.8 (0.8-1.0)</td>
<td>100 (81.9-84.7)</td>
<td>—</td>
</tr>
<tr>
<td>&gt; 20</td>
<td>11.4 (4.7-22.2)</td>
<td>100 (81.5-100)</td>
<td>—</td>
<td>0.8 (0.8-1.0)</td>
<td>100 (81.8-84.3)</td>
<td>—</td>
</tr>
<tr>
<td>&gt; 21</td>
<td>9.8 (3.7-20.2)</td>
<td>100 (81.5-100)</td>
<td>—</td>
<td>0.9 (0.8-1.0)</td>
<td>100 (81.6-84.0)</td>
<td>—</td>
</tr>
<tr>
<td>&gt; 22</td>
<td>6.5 (1.8-15.9)</td>
<td>100 (81.5-100)</td>
<td>—</td>
<td>0.9 (0.9-1.0)</td>
<td>100 (81.3-83.3)</td>
<td>—</td>
</tr>
<tr>
<td>&gt; 23</td>
<td>4.9 (1.0-13.7)</td>
<td>100 (81.5-100)</td>
<td>—</td>
<td>0.9 (0.9-1.0)</td>
<td>100 (81.2-82.9)</td>
<td>—</td>
</tr>
<tr>
<td>&gt; 26</td>
<td>1.6 (0.0-4.8)</td>
<td>100 (81.5-100)</td>
<td>—</td>
<td>0.9 (1.0-1.0)</td>
<td>100 (81.8-82.0)</td>
<td>—</td>
</tr>
<tr>
<td>&gt; 29</td>
<td>0.0 (0.0-5.9)</td>
<td>100 (81.5-100)</td>
<td>—</td>
<td>1.0 (1.0-1.0)</td>
<td>—</td>
<td>81.3 (81.3-81.3)</td>
</tr>
</tbody>
</table>

CI = confidence interval, NPV = negative predictive value, PPV = positive predictive value.

under the curve = 0.729 (95 % CI = 0.617-0.823), p = 0.0003, with a standard error of 0.0634. Youden index was 0.4271 and a cutoff point at 12 was suggested.

Discussion

The present study is the first to assess EPDS performance as a diagnostic test in Mexican women whose screening result was positive for PPD. The frequency of positivity in EPDS (17.9 %) agrees with previously published results14 with different cutoff points, ranging from 9 to 12 depending on the study. EPDS positivity was found at 28 ± 13.2 days post-partum, suggesting the convenience of routinely screening from the second week after delivery and thereafter, when the women attend for examination of
the newborn. Although our objective was not to study risk factors for PPD, it is important to highlight that their detection would provide useful information to implement prevention and treatment strategies.

*Hospital Juárez de México* serves low-income populations without social security, which probably have risk factors such as domestic violence; according to the National Institute of Statistics and Geography, 43.9 % of Mexican women have had at least one episode of violence generated by their partners. Therefore, Mexico needs to implement national health policies to look after the emotional development of families; specifically, timely care of PPD should be a priority.

It is important to note that, of those women with positive EPDS, 22.7 % had no PPD according to BDI criteria, i.e., the result was false positive; in addition, 22.8 % of the women who had an EPDS score ≥ 10 had no depression. The highest percentage, 39.2 %, had mild to moderate depression, which is relevant because depression identification at this stage will allow early intervention in order to avoid an increase in severity and chronicity of the disorder.

Diagnosis is a dynamic process the purpose of which is to support correct decision making. Although having absolute diagnostic certainty is not necessary to adopt the correct therapeutic decision, diagnostic tests seek to reduce uncertainty. In this study, post-test probability (predictive value, PV) decreased uncertainty, since the probability of PPD increased with positive results or decreased with negative results.

**Figure 1.** Consequences of performing the screening with Edinburgh Postnatal Depression Scale in woman during the postpartum period. EPDS = Edinburgh Postnatal Depression Scale, PPD = postpartum depression, BDI = Beck’s Depression Inventory, PPV = positive predictive value, NPV = negative predictive value. Adapted from Leeflang MMG. Systematic reviews and meta-analyses of diagnostic test accuracy. Clin Microbiol Infect. 2014;20:105-113.
results when EPDS was applied. Given that PPD pretest probability (prevalence) (18.7 %) is not so high or close to 100 % to cross the therapeutic threshold and start treatment, or so low (0 %) as to cross the diagnostic threshold and rule out PPD, applying a simple, accessible and accurate test is required. Hence the importance of screening with EPDS, since these thresholds depend on each clinical scenario, on PPD severity and on the cost of false positive and false negative cases, both in clinical (risk of delaying or not establishing the diagnosis, with consequences on the mother, the family and the development of the child, due to undiagnosed depression) and economic terms (if not diagnosed, inability of the mother to work or expenses due to neglect of the child owing to maternal depression).

The present study contributes to know EPDS validity and accuracy, since it allowed knowing the agreement it had with BDI to correctly classify the presence or absence of PPD. The higher the sensitivity and specificity values, the more valid the test is known to be. With a cutoff point of 12, the best sensitivity (70.4 %) and specificity (72.2 %) values were obtained for EPDS, with these data reflecting the intrinsic value of the test before knowing the result.

Little has been published about EPDS in Mexico. Alvarado et al. reported similar sensitivity (75 %), but higher specificity (93 %) with a cutoff point of 11/12:25 PPV (36.9 %) in our study was lower than that reported by Alvarado et al. (50 %), whereas NPV was similar (91.4 %). It is important to remember that predictive values depend on the prevalence of PPD (18.7 %) and, therefore, are not directly applicable to scenarios with different pretest probabilities; regarding PPD, the reports differ depending on the type of study, place of study or instrument used as reference method. Therefore, to determine the probability of PPD (post-test) in scenarios with different pre-test probabilities, LRIs were calculated, which constitute sensitivity and specificity summary indices that are independent of pretest probability. Thus, in our study, the probability of finding a positive result was 2.5 times higher, which constituted a small change, although it was important due to the psychological, economic and social repercussions PPD has. There was also a small change (0.4) in the probability of finding a negative result in the absence of PPD after applying the EPDS.

The present study has the strength that the analysis goes beyond the knowledge of EPDS sensitivity, specificity and predictive values. To the best of our knowledge, and taking into account that the diagnostic process does not end with the calculation of post-test probability, in no previous study in Mexico have the consequences of screening with EPDS been discussed. Furthermore, the profitability of a diagnostic test does not only depend on its validity, but also on its performance in diagnostic-therapeutic decision-making and its cost.23 EPDS result does change initial clinical attitude and influences on the management of women with PPD; in addition, EPDS application is simple, inexpensive and accessible in routine clinical practice, which does not compromise its profitability. In the performance of a diagnostic test, the benefits women will experience with its application should also be considered: better bond with the child, better physical and emotional development of the children, more family harmony and reincorporation of the woman to work, among other.

A woman with an EPDS score > 12 should be evaluated and, if necessary, opportunely treated. According to the results of this study, out of 49 women who were considered healthy (and therefore did not receive treatment), five should have been referred for PPD care, which means that almost 10 % were without treatment despite the screening test.

Finally, the ROC curve and the area under the curve (0.729, 95% CI = 0.617-0.823) show that, although statistically significant, EPDS has a moderate accuracy; therefore, it will be better complementing it with another instrument or psychiatric assessment of suspected cases, in order to have a more accurate diagnosis.
Conclusion

It is important to raise awareness among health personnel and the general population about the relevance of PPD and of routinely screening mothers for PPD with EPDS during newborn assessment visits. When any woman has a score > 12 points, an additional evaluation will be needed with some other instrument and, sometimes, referring her to the specialist, who will determine the treatment. EPDS results are positive in 70% of women with PPD and negative in 72% of those without PPD. A woman with an EPDS positive result has a 37% probability of having PPD; if her result is negative, she has 91% probability of having PPD. Therefore, we can claim that EPDS is a diagnostic test with moderate accuracy.

Conflict of interests

None.

Ethical disclosure

Protection of human and animal subjects. The authors declare that no experiments were performed on humans or animals for this study.

Confidentiality of data. The authors declare that they have followed the protocols of their work center on the publication of patient data.

Right to privacy and informed consent. The authors have obtained the written informed consent of the patients or subjects mentioned in the article. The corresponding author is in possession of this document.

References