Can We Effectively Use the Two-Item PHQ-2 to Screen for Postpartum Depression?

Sung Y. Chae, MD; Mark H. Chae, PhD; Alina Tyndall, MD; Maria Rochelle Ramirez, MD; Robin O. Winter, MD, MMM

BACKGROUND AND OBJECTIVES: Postpartum depression screening is widely advocated to identify and treat affected individuals given the significant impact of this disorder on patients and their families. An effective, efficient method is needed to improve compliance with screening, which has led to an increased interest in the use of the two-item Patient Health Questionnaire 2 (PHQ-2). The aim of this study was to determine the sensitivity and specificity of the PHQ-2 in screening for postpartum depression.

METHODS: A prospective convenience study was conducted among 200 postpartum women attending their postpartum or 4- and 6-month well-child visits at a multiethnic family medicine residency center. The sensitivity and specificity of the PHQ-2 was determined by using the well validated Edinburgh Postnatal Depression Scale (EPDS) as the gold standard. Positive responses to either scale led to further evaluation and referral.

RESULTS: The sensitivity of the PHQ-2 was 100%, and the specificity was 79.3% using the EPDS as the reference standard. In addition, the PHQ-2 identified an additional four/nine women who were subsequently diagnosed with postpartum depression based on follow up of their positive screens.

CONCLUSIONS: This study supports previous findings indicating that the PHQ-2 can be an effective tool in screening for postpartum depression.

(Fam Med 2012;44(10):698-703.)

Depression is considered one of the most prevalent disorders in primary care with far-reaching consequences. This has led the US Preventive Services Task Force to recommend screening all adults for depression.1 Postpartum depression has been increasingly recognized as a significant variant of this disorder causing morbidity not only for the mother but also for her children and family. There is a general consensus that it is important to screen for this disorder, with laws mandating screening in New Jersey and the recent passage of the Melanie Blocker Stokes Mothers Act in 2010 on the federal level to fund support programs, education, and research for postpartum depression.

There are multiple methods available to screen women for postpartum depression.2-6 The Edinburgh Postnatal Depression Scale (EPDS) is a frequently used 10-item instrument developed specifically to screen for postpartum depression with the intent to exclude somatic symptoms common in the puerperium.7 It has been translated into multiple languages and has been validated for use with different cultures.8-10 However, some studies have suggested significant heterogeneity and a wide variance in sensitivity and specificity across diverse groups. As such, the EPDS may not be culturally appropriate for use with patients from some ethnic backgrounds.11

In addition, the EPDS requires time and literacy for the patient to complete the questionnaire and for the provider to appropriately score the measure. Chaudron et al12 conducted a study in a busy, academic pediatric practice where an attempt was made to implement universal screening for postpartum depression by administering the EPDS during first-year well-child visits. Only 46% of visits had a documented screen. If one excluded the screens that were.
Given these conflicting results, and the relative dearth of research on the PHQ-2 as a screening measure for postpartum depression, the purpose of this study was to determine the sensitivity and specificity of the PHQ-2 as a screen for postpartum depression administered during both postpartum and well-child visits in a busy residency practice setting.

Methods

Study Design

The study participants consisted of a convenience sample of 200 women who were seen during their 6-week postpartum visits or with their infants at their 4- or 6-month well-child visits at a multiethnic family medicine residency practice in New Jersey between February and December 2008. This program delivers care to over 300 women annually. IRB approval was obtained, but informed consent was waived since the state of New Jersey mandates postpartum depression screening.

Participants received a written survey (English or Spanish), which included the EPDS and PHQ-2 and a demographics questionnaire. Participants who screened positive with either screen were interviewed and assessed using the DSM-IV criteria for depression.21 The clinicians reviewed the criteria with these participants to confirm the diagnosis of depression. The EPDS was considered positive if the participant scored ≥ 13 as recommended by the original validation study7 and the New Jersey Department of Health and Human Services website. A positive screen using the PHQ-2 was determined if the participant answered yes to either of the two questions. The investigators AT and MR contacted the participants by phone for a detailed interview if the clinician seeing them at the office visit did not document doing so. All the clinical providers were advised at the beginning of the study that if the last question in the EPDS was positive (pertaining to suicidality), the screen had to be addressed immediately. If a participant was diagnosed with depression, appropriate treatment and referral were provided.

The EPDS was considered the gold standard to determine the sensitivity and specificity of the PHQ-2. All statistical procedures were conducted using the Statistical Package for the Social Sciences, version 16.0 (SPSS 16.0).

Measures

Edinburgh Postnatal Depression Scale (EPDS). The EPDS is the most widely used screening instrument for postpartum depression.11,22 It is a self-report Likert-type questionnaire with 10 items inquiring the degree and frequency to which the participant experienced emotional and cognitive symptoms of depression over the past week.7

Patient Health Questionnaire 2 (PHQ-2). The PHQ-2 consists of the first two items from the longer Patient Health Questionnaire-9, which consists of nine items that align with the DSM-IV criteria for major depression. These nine items comprise the first part of the full Patient Health Questionnaire that was developed as a self-report scale to screen for common psychiatric disorders in the primary care setting.24 The two-question screen assesses both sad mood and anhedonia and if used in a dichotomous yes/no fashion, a positive response to either item yields a positive result. Alternatively, the questions could be scored with a Likert-type scale of 0 to 3, with 0 being “not at all” to 3 being “nearly every day.” In this study the dichotomous scoring approach was used.

Results

The demographic variables of the study sample are listed in Figure 1. This sample consisted of a multiethnic participant pool with a relatively high proportion of individuals who classified themselves as Asian. A total of 46.5% indicated that they had...
<\$20,000 annual household income. However, 42.5% reported having earned a college education or higher. Forty-five of the 200 surveys (22.5%) were positive by the PHQ-2 screen. Among these, seven (3.5%) were also positive by the EPDS screen of ≥ 13. All those who screened positive on the EPDS also had a positive PHQ-2 screen. The sensitivity of the PHQ-2 was 100%, and the specificity was 79.3% using the EPDS as the reference standard. Figure 1 displays the ROC curve and shows that 0.902 of the area under the curve was accounted for by the PHQ-2 screen (P < .001).

Of all the participants who screened positive, the investigators were able to follow up on 42 of these 45 women (the others could not be contacted by phone). One of these participants was excluded as she was only 1 week postpartum (baby blues could not be excluded), and one other participant could not even remember filling out the screening instrument and denied depressive symptoms. Of the remaining 40/45 women who screened positive, nine fulfilled DSM IV criteria for depression, which led to 9/200 (4.5%) being ultimately diagnosed with postpartum depression. The EPDS screen (using the ≥ 13 screening cutoff) missed four/nine of the women subsequently diagnosed with postpartum depression by DSM IV criteria. If a lower cut-off of ≥ 10 was considered, two/nine of these women would be missed.

There were no significant correlations by nonparametric analyses of categorical demographic variables except a Spearman’s correlation of 0.479 between a personal and family history of depression (P = .001). Analyses of scale variables showed a significant Pearson’s correlation of -0.232 (P = .002) between the number of years in the United States and the EPDS score. Evaluation of the point biserial correlation between the EPDS score and depression by DSM IV criteria also showed significance at 0.593 (P = .001). This modified form of the Pearson correlation was used as it was the most appropriate method to examine the relationship between a dichotomous variable (depression by DSM IV) and a continuous variable (EPDS score).

**Discussion**

The results of this study indicate a sensitivity of 100% and specificity of 79.3% for the PHQ-2 compared to the EPDS as the reference standard, which is consistent with the findings of Bennett et al.19 whose study showed that the PHQ-2 compared well to the EPDS in its ability to screen for depression during the peripartum period. A subsequent study also indicated that the PHQ-2 displayed a high sensitivity of 100% for postpartum depression in comparison to the Structured Clinical Interview for the DSM-IV among women screened during their infants’ well-child visits.20 Cutler et al.20 attributed the poor sensitivity value of 43.5% of the PHQ-2 in their study to the lower socioeconomic, educational, and multiethnic nature of the participants in their study. However, Bennett et al.19 also included multiethnic women of lower socioeconomic status, and the higher sensitivity of the PHQ-2 appeared unrelated to the educational status of their participants. The difference in results may lie in the administration of the PHQ-2 in the dichotomous yes/no format in the Bennett study19 in contrast to the Likert-type approach used by Cutler et al.20 The Likert type scale may better address the responses of more highly educated samples as there was improved sensitivity (85.7%) of the PHQ-2 among the subset of women who had some college education or more in the Cutler study. Our study may more closely correspond to the results of Bennett et al.19 due to our use of the dichotomous yes/no approach, which may avoid this educational distinction. Lowering the cut-off of the Likert scale PHQ-2 from 3 to 2 also improved its sensitivity significantly from 43.5% to 78.3% in the Cutler study. This likely reflects the greater sensitivity afforded by the dichotomous approach, suggesting that this
may provide better utility from the PHQ-2 screen.

Of note, follow up of the positive screens in this study indicated that the PHQ-2 identified an additional four women who were eventually diagnosed with postpartum depression who were missed by the EPDS. As indicated in other studies using the EPDS as a screening tool for postpartum depression,\textsuperscript{12,27} including cross-cultural studies,\textsuperscript{9,11} a lower cutoff of 9-10 may be more appropriate in some populations to improve the sensitivity of the EPDS. This study supports this assertion as a lower cutoff of 10 would have detected an additional two women subsequently diagnosed with postpartum depression. However, even with the lower cutoff, the PHQ-2 detected an additional two women who were missed by the EPDS.

There may be concerns for an instrument with increased sensitivity due to the higher likelihood of false positives.\textsuperscript{11} On the other hand, it is interesting that in this multi-ethnic population with a fairly large number of women who were immigrants (average number of years in the United States was 17.8 years, SD=1.1 for women whose average age was 26.7 years, SD=5.6), the percentage of those who were diagnosed with postpartum depression was relatively low at 4.5%. In such populations, it may be especially useful to have a more sensitive screening instrument.

As noted previously, a brief, reliable screening instrument may be particularly helpful in a busy practice setting to improve provider compliance with screening.\textsuperscript{3,12,13} A screening tool that can be administered verbally may also be helpful in improving patient compliance with screening, particularly in a multi-ethnic practice where literacy may be a significant issue. The PHQ-2 has been increasingly recognized as a useful tool in screening for postpartum depression. The National Institute for Health and Clinical Excellence, an independent organization that evaluates and provides health care guidelines in the United Kingdom, recommended the use of the PHQ-2 as an initial screen for postnatal depression.\textsuperscript{28}

Family physicians caring for women and their children are in a particularly good position to screen women for postpartum depression during their postpartum visits as well as during well-child visits. However, as noted by the breadth of research on postpartum depression screening in the pediatric literature\textsuperscript{12,13,17,23,26,27} and the obvious relevance to obstetric providers, it would be important for all maternal and child health providers to screen for postpartum depression. A brief, reliable measure would likely increase the probability that busy clinicians would add this screening to their regular repertoire.

Based on the results of this study, the PHQ-2 was added to the postpartum visit template in the electronic medical record of the study practice and continues to be effectively used as a screening measure.

### Table 1: Demographic Characteristics of Participants*

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>34</td>
<td>17</td>
</tr>
<tr>
<td>Asian</td>
<td>41</td>
<td>20.5</td>
</tr>
<tr>
<td>Caucasian</td>
<td>52</td>
<td>26</td>
</tr>
<tr>
<td>Hispanic</td>
<td>48</td>
<td>24</td>
</tr>
<tr>
<td>Other</td>
<td>18</td>
<td>9.0</td>
</tr>
<tr>
<td>Missing</td>
<td>7</td>
<td>3.5</td>
</tr>
<tr>
<td><strong>Income</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;$20,000</td>
<td>93</td>
<td>46.5</td>
</tr>
<tr>
<td>$20,000–$50,000</td>
<td>61</td>
<td>30.5</td>
</tr>
<tr>
<td>&gt;$50,000</td>
<td>8</td>
<td>4.0</td>
</tr>
<tr>
<td>Missing</td>
<td>38</td>
<td>19</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below high school</td>
<td>19</td>
<td>9.5</td>
</tr>
<tr>
<td>High school or GED</td>
<td>84</td>
<td>42</td>
</tr>
<tr>
<td>College or higher</td>
<td>85</td>
<td>42.5</td>
</tr>
<tr>
<td>Missing</td>
<td>12</td>
<td>6.0</td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never married</td>
<td>93</td>
<td>46.5</td>
</tr>
<tr>
<td>Married</td>
<td>86</td>
<td>43</td>
</tr>
<tr>
<td>No longer married</td>
<td>12</td>
<td>6.0</td>
</tr>
<tr>
<td>Missing</td>
<td>9</td>
<td>4.5</td>
</tr>
<tr>
<td><strong>Number of years living in United States</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 years or less</td>
<td>35</td>
<td>18</td>
</tr>
<tr>
<td>5+ to 10 years</td>
<td>18</td>
<td>9.0</td>
</tr>
<tr>
<td>10+ to 15 years</td>
<td>13</td>
<td>6.5</td>
</tr>
<tr>
<td>15+ to 20 years</td>
<td>34</td>
<td>17</td>
</tr>
<tr>
<td>More than 20 years</td>
<td>80</td>
<td>40</td>
</tr>
<tr>
<td>Missing</td>
<td>19</td>
<td>9.5</td>
</tr>
</tbody>
</table>

* Total (n=200)
for postpartum depression during the standard 6- to 8-week postpartum visit.

**Limitations**

There are several limitations to this study. This study used the EPDS as the reference standard to determine the sensitivity and specificity of the PHQ-2 rather than a diagnostic measure. However, there have been multiple validation studies of the EPDS as an effective screening measure for postpartum depression, and the purpose of this study was to help assess the reliability of other studies that found relatively high sensitivities and specificities of the PHQ-2 compared with the EPDS, in contrast to others that did not. Of note, this study seemed to indicate that the PHQ-2 may actually perform better in some populations as it detected an additional four women subsequently diagnosed with postpartum depression.

Also, the demographics of the study population may not be representative of most practices. It included a very multiethnic population (see Table 1) and those with a high level of education of college or greater (42.5%) although this was not reflected in the socioeconomic status with 46.5% of women reporting <$20,000 in annual household income. In addition, many of the participants were not born in the United States. On the other hand, the multiethnic nature of this sample may be relevant given the changing racial and ethnic demographic of the United States. It is helpful to note the ability of the PHQ-2 to screen for postpartum depression in this setting.

This study was also conducted in a residency practice. This may limit the ability to generalize these findings to other practice sites. It is useful to note, however, that postpartum depression screening can be effectively implemented in a busy residency practice despite the inefficiencies inherent in a setting that espouses teaching as a priority.

**Future Directions**

Given the brevity of the PHQ-2, it would be interesting to determine if the broader acceptance of this measure as a screening tool improves the compliance of clinicians with mandated or recommended postpartum depression screening in New Jersey and elsewhere. Currently, the New Jersey Department of Health and Human Services encourages the use of the EPDS as the main screening tool for postpartum depression.

Yawn et al. showed a strong concordance between the PHQ-9 and the EPDS. Another avenue of research would be to investigate if the step-wise use of the PHQ-2 with subsequent administration of the remainder of the PHQ-9 in positive screens helps to maintain the high sensitivity of the PHQ-2 with an improvement in specificity as suggested by another recent validation study of the PHQ-2 and 9 in a primary care sample.

It is also noteworthy that an increasing body of literature indicates the importance of birth-related paternal depression and its potential impact on the well-being of the family. Multiple studies, including this one, support the effectiveness of postpartum depression screening during well-child visits. Given the brevity and reliability of the PHQ-2 screen, it would be helpful to study the expansion of depression screening to fathers or other important child care providers who bring children to their well-child visits. There are ethical, legal, and pragmatic issues that arise, particularly for pediatricians who do not directly provide care to the parents. However, as noted by Olson et al., pediatricians often found that screening mothers for depression allowed them to obtain information that was valuable to caring for their patients. Family physicians providing care to the entire family may be in a better position in this regard.

It is clear that postpartum depression screening is important to identify a widespread problem that has clear implications for the entire family. The PHQ-2 is an effective tool that may improve compliance with

---

<table>
<thead>
<tr>
<th>Test Result Variable(s)</th>
<th>Area</th>
<th>Standard Error*</th>
<th>Asymptotic Significance**</th>
<th>Asymptotic 95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHQ item 1 depression</td>
<td>.890</td>
<td>.078</td>
<td>.000</td>
<td>.737 - 1.042</td>
</tr>
<tr>
<td>PHQ item 2 anhedonia</td>
<td>.772</td>
<td>.101</td>
<td>.015</td>
<td>.574 - .969</td>
</tr>
<tr>
<td>PHQ total score</td>
<td>.902</td>
<td>.028</td>
<td>.000</td>
<td>.847 - .956</td>
</tr>
</tbody>
</table>

The test result variable(s): PHQ item 1 depression, PHQ item 2 anhedonia
EPDS—Edinburgh Postnatal Depression Scale
PHQ—Patient Health Questionnaire 2
* Under the nonparametric assumption
** Null hypothesis: true area = 0.5
recommended or mandated screening and could possibly be applied to other members of the family as well.

ACKNOWLEDGMENTS: Drs Tyndall and Ramirez were both residents at the JFK Family Medicine Residency Program during the data collection for this study.

This project was presented in a research forum at the 2009 Society of Teachers of Family Medicine Annual Spring Conference in Denver.

CORRESPONDING AUTHOR: Address correspondence to Dr Chae, JFK Family Medicine Center, 56 James Street, Edison, NJ 08818. 732-321-7494. Fax: 732-906-4986. schae@jfkhealth.org.

References


