

EVIDENCE-BASED CLINICAL MEDICINE

Efficiency of a Two-Item Pre-Screen to Reduce the Burden of Depression Screening in Pregnancy and Postpartum: An IMPLICIT Network Study

Ian M. Bennett, MD, PhD, Andrew Coco, MD, James C. Coyne, PhD,
Alex J. Mitchell, MBBS, James Nicholson, MD, MSCE, Ellen Johnson, MD,
Michael Horst, PhD, and Stephen Ratcliffe, MD, MSPH

Objective: Systems for efficient case finding of women with major depression during pregnancy and postpartum are needed. Here we assess the diagnostic accuracy of a modified 2-item patient health questionnaire (PHQ-2) as a pre-screen in assessing depression.

Methods: Cross-sectional assessments at 15 weeks' gestation (n = 414), 30 weeks' gestation (n = 334), and 6 to 16 weeks postpartum (n = 193) among women from a diverse set of races/ethnicities, participating in the IMPLICIT maternal care quality improvement network. The Edinburgh Postnatal Depression Scale score (≥ 13) was used as the criterion measure for the PHQ-2.

Results: A positive 2-item screen had sensitivity of 93%, 82%, and 80% and specificity of 75%, 80%, and 86% for Edinburgh Postnatal Depression Scale score of ≥ 13 for assessment at 15 and 30 weeks gestational age and postpartum, respectively. The positive/negative predictive values for the PHQ-2 were 44/98, 24/91, and 30/98 for each time point, respectively. Areas under the receiver operating characteristic curve analysis suggested that 2-item assessments at each time point had approximately equal diagnostic validity.

Conclusions: Two questions were efficient to rule out depression and reduced the need for further screening of approximately 60% to 80% of women, depending on the point in pregnancy or postpartum. A diagnostic interview follow-up of women screening positive is still required. (J Am Board Fam Med 2008;21:317-325.)

The identification of perinatal major depression occurring during pregnancy and the first year postpartum is considered a critical goal of the maternal care system but is poorly conducted.¹⁻⁴ Only a minority of women suffering from perinatal depression are identified by health care providers of either maternal or infant care.⁵⁻⁸ Although limitations of mental health service delivery play a role in this discrepancy, a major impediment to low rates of depression case finding is the difficulty in the

administration of depression screening in busy clinical settings.⁹ A range of instruments have been used during pregnancy and postpartum, but they generally take at least 5 minutes to complete or rely on adequate literacy of patients for self-administration. In the general primary care setting a 2-item pre-screen has been validated and has reasonable psychometric properties for ruling out patients at low risk for depression, allowing them to avoid further assessment.¹⁰⁻¹² This approach has obvious

This article was externally peer reviewed.

Submitted 23 February 2008; revised 28 April 2008; accepted 29 April 2008.

From the Department of Family Medicine and Community Health (IMB, JCC, JN), the Department of Psychiatry (JCC), and the Leonard Davis Institute of Health Economics (IMB), University of Pennsylvania School of Medicine, Philadelphia; the Lancaster General Hospital Department of Family and Community Medicine and Lousie von Hess Medical Research Institute (AC, MH, SR), Lancaster, PA; the Department of Liaison Psychiatry, Leicester General Hospital (AJM), Leicestershire, UK; and the Good Samaritan Family Practice Residency Program (EJ), Lebanon, PA.

Funding: This work was supported by a grant from the March of Dimes. IMB was supported by grants from the NICHD (1K23HD048915-01A2) and the NIMH (1R03MH074750-01).

Prior presentation: This work was presented in part at the 2007 North American Primary Care Research Group annual meeting, Vancouver, Canada, October 20 through October 23, 2007.

Conflict of interest: none declared.

Corresponding author: Ian M. Bennett, MD, PhD, Family Medicine and Community Health, University of Pennsylvania School of Medicine, 2nd Floor Gates Pavilion, 3400 Spruce Street, Philadelphia, PA 19104 (E-mail: ian.bennett@uphs.upenn.edu).

advantages during pregnancy and has been recommended, with the addition of a third item concerning a request for help, for use by maternal care providers in the British health system.^{13,14} However, to our knowledge this instrument has not been validated in the prenatal and postpartum settings where the normal symptoms of pregnancy can overlap with those of depression.

In this study, we assess the ability of the 2-item version of the patient health questionnaire (PHQ-2), with a dichotomous (yes/no) response, to increase the efficiency of screening for risk of major depression during pregnancy and postpartum. Women were evaluated at 2 times during pregnancy and once in the postpartum period using the Edinburgh Postnatal Depression Scale (EPDS) as the criterion. The EPDS is a widely used first step in identifying perinatal depression.^{15,16} The EPDS is most appropriately followed up with an in depth diagnostic interview of women screening positive, but it is efficient in ruling out further assessment for women screening negative; with that caveat in mind, Milgrom et al¹⁷ found that the EPDS had a positive predictive value (PPV) of 76% when measured against Diagnostic and Statistical Manual of Mental Disorders-IV criteria. The goal of the current study was to determine whether the 2-item screening measure could adequately identify women with elevated perinatal depressive symptoms so that a more extensive assessment of women unlikely to have major depression could be avoided.

Methods

Setting

The study sample included patients receiving prenatal and postpartum care from 11 family medicine residency clinical practices in the IMPLICIT (Interventions to Minimize Preterm and Low birth weight Infants through Continuous Improvement Techniques) continuous quality improvement network, which includes rural (n = 7) and urban (n = 4) sites in Pennsylvania (n = 9) and New York (n = 2); institutional review board approval was received from all participating institutions. Women were recruited sequentially to the study from November 2004 through February 2007. Eligibility criteria included English- or Spanish-speaking ability and a singleton intrauterine pregnancy. Women were enrolled in the study at any of 3 time points: twice during pregnancy (15 and 30 weeks' gestation) and

once postpartum (6 to 16 weeks after delivery). All eligible women who had complete depression screening data (scores for PHQ-2 and EPDS, see below) for a particular time point were assessed. Women were excluded who had missing depression screening data because of the requirements of the statistical analyses described below. At the 15 week prenatal assessment, 414 of the 785 eligible women had complete data and were used for analysis (52%), at 30 weeks' gestation 334 of 785 had complete data (58%), and at the postpartum assessment 193 of the 274 eligible women had complete data (70%). Women with missing depression screening data did not differ significantly from those included in the analysis by any demographic variables ($P > .05$). Women were assessed by physicians or nurses during routine maternal care. Data were then recorded using standardized forms and then entered, after being de-identified, into an Internet-based centralized database for the IMPLICIT network.

Measures of Depression

Depressive symptomatology was measured using the English and Spanish versions of the EPDS, a 10-item instrument developed for use in the postpartum period and validated for use during pregnancy, with scores ranging from 0 to 30.^{18,19} The EPDS has been used in a wide range of populations, including low-income women similar to those in the current study.^{2,20,21} A number of cut points have been used and proposed for the identification of perinatal depression, although there are considerable methodological inconsistencies and inadequacies in available studies, precluding a definitive choice among them.²² We have chosen to evaluate EPDS score ≥ 13 because it has the strongest support for identifying major depression.¹³ Although studies are limited, the pooled sensitivity for the identification of major depression in high-quality studies at this cut point both in pregnancy and postpartum has been found to be $>95\%$ with a specificity of $>80\%$.^{13,23} The potential pre-screening tool was the 2-item version of the PHQ-2, modified so that response choices were dichotomous (yes/no) rather than 4 ordinally related items.^{10,24,25} Responding yes to either of the items was considered a positive result: "During the past month have you often been bothered by feeling down, depressed, or hopeless?" and "During the past month have you often been bothered by little interest or pleasure in doing things?" Both the

PHQ-2 and EPDS were initially provided to the patients in printed form as self-response questionnaires. Physicians then orally reviewed the responses to these assessments with patients individually. History of depression was determined by abstracting documentation of previous depression from prenatal and postpartum clinical charts.

Descriptive Variables

Descriptive variables were selected to aid in the characterization of the sample because of their established associations with depression during pregnancy and postpartum. Demographic variables selected were age, race/ethnicity, educational level, marital status, parity (previous live births), and health insurance status.^{26–33} To differentiate between women who are still in school and those who did not complete school, we divided the <high school completion category based on age: ≤ 18 years old (potentially still in school) and > 18 years old (no longer eligible for high school). Self-reported lifetime history of depression is associated with depression during pregnancy and postpartum and was included for evaluation as a potential additional screening item for depression.^{34–36}

Statistical Analysis

Analyses for each of the study time points were conducted independently of the others and so are distinct cross sectional assessments. Bivariate associations between descriptive variables at each time period and a positive EPDS (≥ 13) were assessed using Student's *t* test and the χ^2 statistic, with appropriate extension when assessing variables with more than 2 (dichotomous) categories.

Receiver Operating Characteristic Curve Analysis

Receiver operating characteristic (ROC) curve analysis was used to evaluate the discrimination capacity of the predictive test (the PHQ-2 alone or with history of depression), or its ability to accurately identify those women in the dichotomized categories (EPDS positive or negative at the cut point of ≥ 13).³⁷ ROC curves plot sensitivity (true positive ratio) by 1-specificity (true negative ratio) for a series of thresholds or cut points established by responses to the instrument items. These thresholds provide information regarding the test characteristics, which can be used to determine the relative usefulness of the test and the specific threshold that maximizes the desired characteristics

for a specific clinical setting (emphasizing either the sensitivity or specificity of the instrument). For a 2-item screening instrument, requiring a positive response from only one of the items (cut point = 1 of 2) maximizes the likelihood of identifying cases (increased sensitivity) but also increases the likelihood of false positives (reduced specificity). Requiring both screening items to be positive to identify a case (cut point = 2) increases the likelihood of missing true positives but also reduces the likelihood of false positives. The area under the ROC curve represents an overall measurement of performance of the screening test, with 1.0 a perfect test and 0.5 representing a test with no discriminating capacity.³⁸ All analyses were conducted using SPSS version 15.0 (SPSS, Inc., Chicago, IL).

Results

Sample Characteristics

The description of the study sample is included in Table 1. Overall the sample was comprised primarily of young, low-income women with demographic characteristics consistent with that setting. More than half of the participants were between the ages of 19 and 24 years and were white, with approximately one quarter having less than high school education (only one quarter had any college or more); nearly half of the women in the study (45%) were having their first child, and nearly all with Medicaid insurance or self paid. The distribution of these demographic variables did not vary significantly between the 3 study periods except for race/ethnicity ($P > .05$); there was a smaller proportion of African-American women in the postpartum sample than in the 2 pregnancy groups.

The percentage of women in the 15- and 30-week prenatal and postpartum study periods who were positive by EPDS score for current depressive symptomatology was 17% (72 of 414), 13% (44 of 334), and 5% (10 of 193), respectively. Of the demographic variables, only Medicaid insurance status was associated with a positive screen with the EPDS. Approximately 30% of women at each of the 3 time points reported a history of depression. Depression history was significantly associated with a positive EPDS screen for the 15-week and 30-week assessments but not the postpartum assessment. The PHQ-2, both alone and when combined with history of depression, was highly associated with a positive EPDS at all 3 of the study assessment points.

Table 1. Characteristics of the Study Sample of Pregnant and Postpartum Women with Prevalence of Depressive Symptomatology*

Descriptive Variables	15 Week Pregnancy		30 Week Pregnancy		Postpartum	
	Total (n = 414)	EPDS ≥13 (n = 72)	Total (n = 334)	EPDS ≥13 (n = 44)	Total (n = 193)	EPDS ≥13 (n = 10)
Demographic variables						
Age						
14–18	53 (13)	11 (15)	45 (14)	8 (18)	21 (11)	0
19–24	220 (53)	39 (54)	172 (51)	23 (52)	103 (53)	5 (50)
25–29	95 (23)	18 (25)	76 (22)	10 (23)	43 (22)	3 (30)
30+	46 (11)	4 (6)	41 (12)	3 (7)	26 (14)	2 (20)
Race						
Hispanic	20 (5)	4 (6)	15 (5)	1 (2)	10 (5)	0
African-American	125 (30)	27 (38)	88 (26)	15 (34)	41 (21)	2 (20)
White	258 (62)	39 (54)	226 (68)	28 (64)	137 (71)	8 (80)
Other	11 (3)	2 (3)	5 (2)	0	5 (3)	0
Education						
<High school	94 (23)	20 (28)	76 (23)	16 (36)	43 (22)	1 (10)
Still in high school	42 (10)	6 (8)	36 (11)	6 (14)	17 (9)	0
High school or equivalent	172 (42)	29 (40)	136 (41)	14 (32)	89 (46)	7 (70)
Any college or more	106 (26)	17 (24)	86 (26)	8 (18)	44 (22)	2 (20)
Parity (previous live births)						
0	185 (45)	27 (38)	142 (43)	14 (32)	81 (42)	3 (30)
1	117 (28)	23 (32)	89 (27)	16 (36)	54 (28)	4 (40)
2–5	112 (27)	22 (31)	103 (31)	14 (32)	58 (30)	3 (30)
Insurance						
Medical assistance						
Self-pay	375 (91)	70 (97)	301 (90)	42 (96)	168 (87)	10 (100)
Private	39 (9)	2 (3)	33 (10)	2 (5)	25 (13)	0
Depression measures						
History of depression	122 (30)	45 (63)	97 (29)	23 (52)	54 (28)	5 (50)
Two-item screen positive	151 (37)	67 (93)	94 (28)	36 (82)	33 (17)	8 (80)
Two-item screen and history of depression	42 (50)	70 (97)	148 (44)	41 (93)	54 (37)	9 (90)

*Edinburgh Postnatal Depression Scale score ≥13.

All values provided a n (%). Bolded values indicate significant association ($P < .05$; Pearson's χ^2 or Fisher's exact test if cells contain ≤5).

Diagnostic Accuracy

The characteristics of the PHQ-2 for identifying positive EPDS cases, alone and in combination with a history of depression, are shown in Table 2 and Figure 1. The cut points for the combined PHQ-2 and history of depression screening instrument are labeled as 1/2 (positive for either the PHQ-2 or history of depression) and 2 (positive for both PHQ-2 and history of depression). The overall performance of the measure as determined by the area under the ROC curve were similar across the 3 study points and had similar point estimates and overlapping confidence intervals when the

PHQ-2 was used alone or in combination with the history of depression (Table 2, column 2).

(1) Screening at week 15 (estimated gestational age [EGA])

The PHQ-2 allowed the detection of 67 of 72 women with depression, defined by a score of 13 or above on the EPDS, with a sensitivity of 93.1%. Specificity was 75.1%, with a PPV of 44% and negative predictive value (NPV) 98%. When the PHQ-2 was combined with a history of depression, sensitivity was improved to 97.2% but at the expense of specificity (61.1%).

Table 2. Diagnostic Validity of the Two-Item Patient Health Questionnaire for the Edinburgh Postnatal Depression Scale*

Variable(s)	AUROC Curve (95% CI)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
EPDS ≥13					
15 weeks' EGA					
PHQ-2	0.80 (0.69–0.91)	93	75	44	98
PHQ-2 and history of depression	0.85 (0.76–0.94)				
Cut point 1/2		97	61	59	91
Cut point 2		58	92	21	99
30 weeks' EGA					
PHQ-2	0.81 (0.74–0.88)	82	80	24	91
PHQ-2 and history of depression	0.85 (0.76–0.94)				
Cut Point 1/2		93	63	42	91
Cut Point 2		41	91	22	98
Postpartum					
PHQ-2	0.83 (0.68–0.98)	80	86	30	98
PHQ-2 and history of depression	0.83 (0.70–0.97)				
Cut point 1/2		90	67	44	96
Cut point 2		60	96	11	99

*Edinburgh Postnatal Depression Scale (EPDS) score ≥13.

PHQ-2, 2-item Patient Health Questionnaire; AUROC, area under the receiving operator characteristic; PPV, positive predictive value; NPV, negative predictive value; EGA, estimated gestational age; Cut point 1/2, one of 2 screening items positive; Cut point 2, both screening items positive.

(2) Screening at week 30 (EGA).

The PHQ-2 facilitated the detection of 36 of 44 women with depression, defined by a score of 13 or above on the EPDS, with a sensitivity of 81.8%. Specificity was 80.0%, with a PPV of 38% and NPV 97%. When the PHQ was combined with a

history of depression, sensitivity was improved to 93.2% but at the expense of specificity (63.1%).

(3) Postpartum Screening (6 to 16 weeks after birth)

The PHQ-2 identified 8 of 10 women with postpartum depression, defined by a score of 13 or

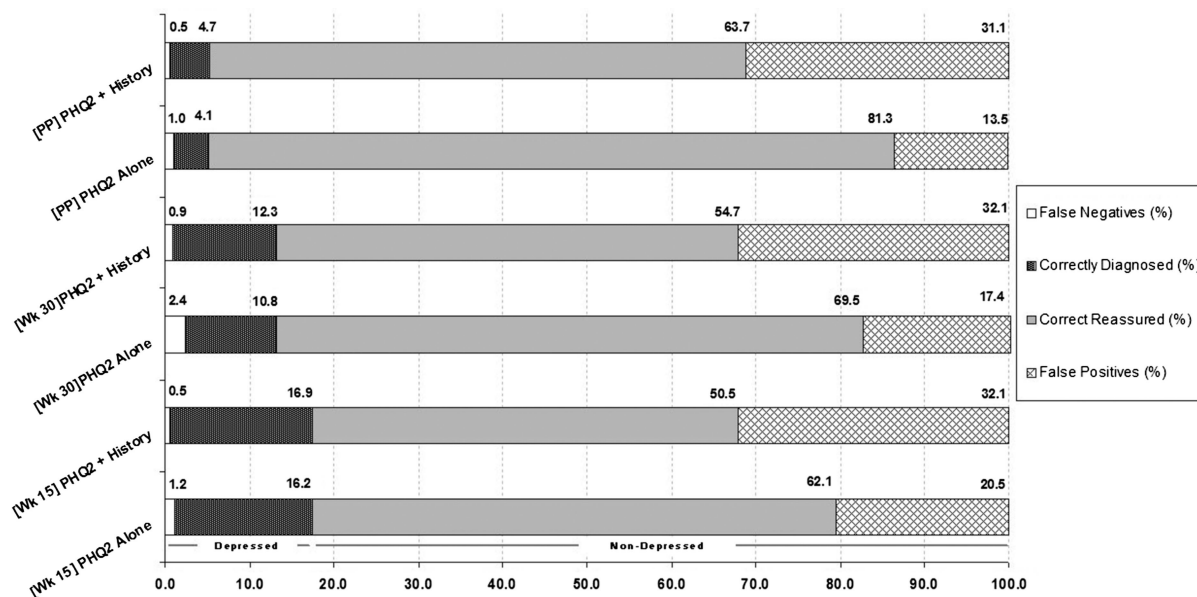


Figure 1. Proportionate accuracy of the PHQ-2 versus EPDS (12 vs 13) per 100 unselected patients.

above on the EPDS, with a sensitivity of 80.0%. Specificity was 85.8%, with a PPV of 24% and NPV 99%. When the PHQ was combined with a history of depression sensitivity was improved to 90.0% but at the expense of specificity (67.2%).

Discussion

In this analysis of women during pregnancy and postpartum we determined that pre-screening with the 2-item PHQ generally had good sensitivity and specificity for identifying women with positive scores ≥ 13 on the Edinburgh Postnatal Depression Scale. The addition of a question regarding history of depression improved the sensitivity, at the expense of specificity, and the PPV of this screening. Out of 100 unselected cases seen at 15 weeks, the PHQ-2 alone would help correctly detect 16 out of 17 depressed patients (missing only one patient) and correctly reassure 62 out of 83 patients, falsely diagnosing 21. Similarly, out of 100 unselected patients seen postpartum, the PHQ-2 alone would help correctly identify 4 out of 5 depressed patients (the prevalence is 5% in this case) and correctly reassure 81 out of 95 patients, misidentifying 13. This demonstrates that when the PHQ-2 is negative it has a 97% to 99% chance of accurately ruling out depression with very few false negatives. Thus, this rapid screening procedure may be useful as part of a multistage case finding strategy to rule out women who would otherwise require further time-consuming assessment for the presence of maternal depression.

Based on the rate of positive 2-item screens in the current study, 60% to 80% of the women being assessed (depending on the particular point in pregnancy and postpartum) can be ruled out with high certainty of not having depression. When the history of depression item is added, the number of women quickly ruled out is slightly higher but at a cost of approximately 10% in case-finding accuracy. It is important to keep in mind that even when the PHQ-2 is negative and the NPV is 90% or higher, it is possible that these women actually have major depression. For this reason we suggest that a negative screen should not overrule clinical concern and that anyone judged to be symptomatic clinically but with a negative pre-screen or screen for depression should have a further concurrent or follow-up examination. It may be somewhat surprising that adding history of depression questions

did not substantially improve the performance of prescreening, particularly given the recognition that depression is a recurrent episodic disorder.³⁹ One explanation is that the women in our sample were young enough that they were still experiencing their first episode of depression. A second reason might be that the PHQ-2 alone has reached a ceiling regarding rule-out accuracy. A third possible reason questions the validity of simple self-report questions for assessing history without a basis for explaining the questions or probing the responses that an interview provides, a finding that has been reported elsewhere.⁴⁰

Importantly, we have shown that the modified PHQ-2 used here functions well as a first pre-screening step in early and late pregnancy and postpartum. By efficiently ruling out those who do not have depression, resources can be targeted at the women at highest risk of a major depressive disorder. However, the value of this strategy depends on diagnostic follow-up of the women registering positive with both the pre-screening and the EPDS. A positive score on the EPDS is by itself an inadequate basis for clinical decision making and, if used as such, scarce resources might be used on the treatment of women who are not truly depressed and in need of help. There is some evidence that busy primary care clinicians are inclined not to follow up on the results of screening questionnaires, instead accepting a positive score on a screening questionnaire as a sufficient basis for diagnosis.⁴¹ It would be unfortunate if the introduction of a pre-screen further encouraged this practice; before recommending this pre-screening strategy clinically it would be important to determine that it would not be used inappropriately.

Although we assessed the PHQ-2 as a pre-screen it has been recommended for use alone as a screening instrument in general primary care practice.¹⁰ It is possible that this instrument, without the use of the EPDS, has sufficient accuracy to identify women at risk (or those not at risk) for depression during pregnancy and postpartum. Currently, the EPDS is well validated and highly accepted in perinatal settings, but little work has justified the use of the PHQ-2 to date. A 3-question approach (essentially the PHQ-2 plus a help question) has been recommended by the British health system's National Institute for Clinical Excellence guidelines but has never been formally evaluated in perinatal settings.¹² In one study, Olson et al⁴²

found that 17% of 1398 mothers answered positively to at least one question of the PHQ-2, although validity was not reported. Recently, Dubowitz et al⁴³ adapted the PHQ-2 into the “Parent Screening Questionnaire.” The authors evaluated this questionnaire among 216 mothers in a primary care clinic compared with the Beck Depression Inventory II completed 2 months later. When a positive response to *either* or *both* of the 2 questions was considered, the sensitivity was 74%, the specificity was 80%, and the NPV was 95%, but the PPV was only 36%.

An unresolved issue is how many of those who screen positive actually are willing to accept professional help. In one study,⁴⁴ only 30% of those who screened positive agreed to be contacted for further help. In a second study,⁴⁵ only 23% of high scorers on the Beck Depression Inventory took up the offer for psychological therapy and 10% agreed to medication. This highlights that screening alone is rarely sufficient for improvements in quality of care, and enhanced detection should be paired with enhanced treatment and follow-up.⁴⁶

A number of limitations of the current study should be reviewed. First, we use as a criterion measure the EPDS screening instrument rather than the clinical diagnosis of depression. However, the goal of the current analysis was to reduce the burden of screening for depression through a pre-screen step, not to diagnose major depressive disorder. Because the formal diagnosis of depression requires significant diagnostic expertise and the investment of valuable clinical time, the rapid identification of women at risk for depression (and not at risk) is critical to the efficient delivery of maternal care. Another limitation of the current study is that the women included were primarily low income as indicated by the high rate of Medicaid insurance. The findings from this study may not be generalizable to higher-income and other less vulnerable populations. However, women with low incomes have higher risk for depressive symptomatology and are less likely to receive mental health services and so are a particularly important group to target for screening.^{4,47} One recent study⁴⁸ found that the sensitivity of the PHQ-2 for a positive EPDS score was lower for women with a high school education or less than for those who went on beyond high school. We did not find any difference in sensitivity across levels of educational attainment in our sample. This difference may be because of

our use of a simpler dichotomous yes/no response choice for the PHQ-2 and a follow-up oral review of the items, which reduces concerns of low education- and literacy-related inaccuracy of a written and more complex instrument.¹¹ Finally, although we had complete data for more than 50% of the women eligible for analysis at each of the time points, it is possible that missing data were related to depression risk status. For example, women with more obvious symptoms may have been either more or less likely to have both the pre-screen (PHQ-2) and screening (EPDS) conducted (both were required to be included in the analysis). Because of the protocol-driven system of data collection in this study, we believe that it is unlikely that any such bias was widespread enough to greatly influence our findings.

Based on the results of these analyses it is reasonable for maternal care providers who are currently using or are planning to use the EPDS to use the PHQ-2 pre-screen, with simplified response items (yes/no), as an initial component of a multi-step case-finding strategy for depression during pregnancy and postpartum. Although this approach should not overrule clinical concern, women who answer “no” to both of the questions in the PHQ-2 can, with high confidence, avoid the longer EPDS and a diagnostic clinical interview. Women who answer “yes” to either of these items should proceed to the EPDS, followed by an in-depth clinical interview for women who have a positive EPDS (≥ 13). Although further research is needed, we recommend that the screening process make use of a direct oral interview rather than relying entirely on a written questionnaire to elicit or review screening questions to avoid error caused by low education and literacy.

Thanks to all the contributing members of the IMPLICIT network and the Family Medicine Educational Consortium for contributions that made this work possible.

References

1. Buist A, Bilszta J, Barnett B, et al. Recognition and management of perinatal depression in general practice—a survey of GPs and postnatal women. *Aust Fam Physician* 2005;34:787–90.
2. Chaudron LH, Szilagyi PG, Kitzman HJ, Wadkins HI, Conwell Y. Detection of postpartum depressive symptoms by screening at well-child visits. *Pediatrics* Mar 2004;113(3 Pt 1):551–8.
3. Cohen LS, Nonacs R, Viguera AC, Reminick A.

- Diagnosis and treatment of depression during pregnancy. *CNS Spectr* 2004;9:209–16.
4. Miranda J, Azocar F, Komaromy M, Golding JM. Unmet mental health needs of women in public-sector gynecologic clinics. *Am J Obstet Gynecol* 1998;178:212–7.
 5. Healthy People 2010: Understanding and improving health. Washington DC: Department of Health and Human Services; 2000.
 6. Bagedahl-Strindlund M, Borjesson KM. Postnatal depression: a hidden illness. *Acta Psychiatr Scand* 1998;98:272–5.
 7. Coates AO, Schaefer CA, Alexander JL. Detection of postpartum depression and anxiety in a large health plan. *J Behav Health Serv Res* 2004;31:117–33.
 8. Thio IM, Browne MAO, Coverdale JH, Argyle N. Postnatal depressive symptoms go largely untreated: a probability study in urban New Zealand. *Soc Psychiatry Psychiatr Epidemiol* 2006;41:814–8.
 9. Seehusen DA, Baldwin LM, Runkle GP, Clark G. Are family physicians appropriately screening for postpartum depression? *J Am Board Fam Med* 2005; 18:104–12.
 10. Berg AO, Allan JD, Frame PS, et al. Screening for depression: USPSTF recommendations and rationale. *Ann Intern Med* 2002;136:760–4.
 11. Arroll B, Khin N, Kerse N. Screening for depression in primary care with two verbally asked questions: cross sectional study. *BMJ* 2003;327:1144–6.
 12. Mitchell AJ, Coyne JC. Do ultra-short screening instruments accurately detect depression in primary care? A pooled analysis and meta-analysis of 22 studies. *Br J Gen Pract* 2007;57:144–51.
 13. Gaynes BN, Gavin N, Meltzer-Brody S, et al. Perinatal depression: prevalence, screening accuracy, and screening outcomes. Rockville (MD): Agency for Healthcare Research and Quality; 2005:119.
 14. Antenatal and postnatal mental health clinical management and service guidance (NICE clinical guideline #45). United Kingdom: National Health Service; 2007.
 15. Boyd RC, Le HN, Somberg R. Review of screening instruments for postpartum depression. *Arch Womens Ment Health* 2005;8:141–53.
 16. Eberhard-Gran M, Eskild A, Tambs K, Opjordsmoen S, Samuelsen SO. Review of validation studies of the Edinburgh postnatal depression scale. *Acta Psychiatr Scand* 2001;104:243–9.
 17. Milgrom J, Ericksen J, Negri L, Gemmill AW, Shakespeare J. Evaluation of screening for postnatal depression against the NSC handbook criteria. *Aust N Z J Psychiatry* 2005;39:833–9.
 18. Cox JL, Chapman G, Murray D, Jones P. Validation of the Edinburgh Postnatal Depression Scale (EPDS) in non-postnatal women. *J Affect Disord* 1996;39:185–9.
 19. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry* 1987;150:782–6.
 20. Areias ME, Kumar R, Barros H, Figueiredo E. Correlates of postnatal depression in mothers and fathers. *Br J Psychiatry* 1996;169:36–41.
 21. Bennett I, Switzer J, Aguirre A, Evans K, Barg F. “Breaking it down”: patient-clinician communication and prenatal care among African-American women of low and higher literacy. *Ann Fam Med* 2006;4:334–40.
 22. 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. *Arch Womens Ment Health* 2006;9:309–15.
 23. Murray D, Cox JL. Screening for depression during pregnancy with the Edinburgh Depression Scale (EPDS). *J Reprod Infant Psychol* 1990;8:99–107.
 24. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: validity of a two-item depression screener. *Med Care* 2003;41:1284–92.
 25. Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression. Two questions are as good as many. *J Gen Int Med* 1997;12: 439–45.
 26. Altshuler LL, Hendrick V, Cohen LS. Course of mood and anxiety disorders during pregnancy and the postpartum period. *J Clin Psychiatry* 1998; 59(Suppl 2):29–33.
 27. Berenson AB, Breitkopf CR, Wu ZH. Reproductive correlates of depressive symptoms among low-income minority women. *Obstet Gynecol* 2003;102: 1310–7.
 28. Gotlib IH, Whiffen VE, Mount JH, Milne K, Cordy NI. Prevalence rates and demographic characteristics associated with depression in pregnancy and the postpartum. *J Consult Clin Psychol* 1989;57:269–74.
 29. Hobfoll SE, Ritter C, Lavin J, Hulsizer MR, Cameron RP. Depression prevalence and incidence among inner-city pregnant and postpartum women. *J Consult Clin Psychol* 1995;63:445–3.
 30. Howell E, Mora P, Leventhal H. Correlates of early postpartum depressive symptoms. *Matern Child Health J* 2006;10:149–57.
 31. Howell EA, Mora PA, Horowitz CR, Leventhal H. Racial and ethnic differences in factors associated with early postpartum depressive symptoms. *Obstet Gynecol* 2005;105:1442–50.
 32. Kuo WH, Wilson TE, Holman S, Fuentes-Afflick E, O’Sullivan MJ, Minkoff H. Depressive symptoms in the immediate postpartum period among Hispanic women in three US cities. *J Immigr Health* 2004;6: 145–53.
 33. Larsson D, Lager I, Nilsson PM. Socio-economic characteristics and quality of life in diabetes mellitus—relation to metabolic control. *Scand J Public Health* 1999;27:101–5.

34. Bernazzani O, Saucier JF, David H, Borgeat F. Psychosocial predictors of depressive symptomatology level in postpartum women. *J Affect Disord* 1997;46:39–49.
35. Bernazzani O, Saucier JF, David H, Borgeat F. Psychosocial factors related to emotional disturbances during pregnancy. *J Psychos Res* 1997;42:391–402.
36. O'Hara MW, Schlechte JA, Lewis DA, Wright EJ. Prospective study of postpartum blues. Biologic and psychosocial factors. *Arch Gen Psychiatry* 1991;48:801–6.
37. Justice A, Covinsky K, Berlin J. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;130:515–24.
38. Hanley J, McNeil B. The meaning and use of the area under a receiver operating characteristic curve. *Radiology* 1982;143:29–36.
39. Klinkman MS, Schwenk TL, Coyne JC. Depression in primary care - more like asthma than appendicitis: The Michigan Depression Project. *Can J Psychiatry* 1997;42:966–73.
40. Coyne JC, Thompson R, Racioppo MW. Validity and efficiency of screening for history of depression by self-report. *Psychol Assess* 2001;13:163–70.
41. Valenstein M, Dalack G, Blow F, Figueroa S, Standiford C, Douglass A. Screening for psychiatric illness with a combined screening and diagnostic instrument. *J Gen Int Med* 1997;12:679–85.
42. Olson AL, Dietrich AJ, Prazar G, Hurley J. Brief maternal depression screening at well-child visits. *Pediatrics* 2006;118:207–16.
43. Dubowitz H, Feigelman S, Lane W, et al. Screening for depression in an urban pediatric primary care clinic. *Pediatrics* 2007;119:435–43.
44. Carter FA, Carter JD, Luty SE, Wilson DA, Framp-ton CMA, Joyce PR. Screening and treatment for depression during pregnancy: a cautionary note. *Aust N Z J Psychiatry* 2005;39:255–61.
45. Horowitz JA, Cousins A. Postpartum depression treatment rates for at-risk women. *Nurs Res* 2006;55(2 Suppl):S23–S27.
46. Gilbody SD, Sheldon TD, House ADM. Screening and case-finding instruments for depression: a meta-analysis. *CMAJ* 2008;178:997–1003.
47. Miranda J, Chung JY, Green BL, et al. Treating depression in predominantly low-income young minority women: a randomized controlled trial. *JAMA* 2003;290:57–65.
48. Cutler CB, Legano LA, Dreyer BP, et al. Screening for maternal depression in a low education population using a two item questionnaire. *Arch Womens Ment Health* 2007;10:277–83.