Patient Health Questionnaire Depression Scale as a Suicide Screening Instrument in Depressed Primary Care Patients: A Cross-Sectional Study

Lisa A. Uebelacker, PhD; Nicole M. German, BA;
Brandon A. Gaudiano, PhD; and Ivan W. Miller, PhD

Objective: The aim of this study was to examine the sensitivity and specificity of the suicide item on the 9-item Patient Health Questionnaire (PHQ-9) when compared to a structured interview (the Structured Clinical Interview for *DSM-IV*; SCID-I mood module) in primary care patients with elevated depression symptoms.

Method: In this cross-sectional study, we analyzed data from 166 patients from 2 primary care clinics, 1 in Rhode Island and 1 in Massachusetts, who were enrolled in studies that focused on depression in primary care. Of the total participants, 101 were enrolled in the survey study, and 65 were screened for or enrolled in either an open trial or a pilot randomized controlled trial. Data were collected between May 2004 and May 2009.

Results: We found that the specificity of the PHQ-9 suicide screening item was 0.84 and sensitivity was 0.69 for the sample as a whole.

Conclusions: This study suggests that the routine use of the PHQ-9 may be useful in primary care practice in that it may identify individuals at risk for suicide who would not otherwise have been identified. However, denial of suicidality on the PHQ-9 should be probed further if there are other risk factors for suicide present.

Trial Registration: clinicaltrials.gov Identifier: NCT00541957

Prim Care Companion CNS Disord 2011;13(1):e1–e6

Current epidemiologic estimates of lifetime prevalence of suicidal ideation and suicide attempts in the United States are 5.6%–14.3% and 1.9%–8.7%, respectively.[1] Reported findings[2] show that a significant proportion of individuals who complete suicide—estimates range between 20% and 76%—had contact with a primary care provider in the month before the suicide. Thus, primary care providers have the opportunity to identify suicidality and implement preventative interventions. In order to do this, primary care providers must identify patients at risk in an efficient way.

Currently, primary care providers are likely to identify only a portion of individuals with suicide ideation. Data show that, despite increased contact with health professionals, only 22% of suicide completers verbally reported suicide ideation or intent to any physician during the 28 days before suicide completion.[3] Even in a practice undergoing quality improvement procedures for depression treatment, 60% of patients with current suicidal ideation reported that their primary care physician did not ask about self-harm during the relevant index visit.[4]

One way to increase detection is universal screening for suicide. Recently, the Joint Commission on the Accreditation of Healthcare Organizations included as one of its National Patient Safety Goals that the hospital "identify individuals at risk for suicide."[5(p9)] Specifically, the Joint Commission requires that patients being treated for emotional or behavioral disorders in general hospitals (as either inpatients or outpatients) undergo a suicide risk assessment and that the hospital provide for the patients' immediate safety. As a result of this requirement, hospital-based primary care clinics must implement procedures for conducting suicide risk assessments. Self-report screening instruments could be a useful first step in either a universal or targeted (ie, toward a high-risk group such as those with mental disorders) screening process.

Few studies have examined the association between self-reported suicide screening instruments and clinician reports of suicidality. According to the US Preventive Services Task Force report[6] on screening for suicide risk, no studies have evaluated the usefulness of screening high-risk groups in primary care, and it is unclear whether self-report items can reliably identify most cases of suicidality. In 1 study investigating a mental health self-report screen in primary care, a 2-item suicidal ideation subscale that was part of a larger general mental health screening questionnaire had low sensitivity (0.43–0.62, depending on the sample) but adequate specificity (>0.90) when compared with a structured clinician interview.[7] In contrast, in psychiatric inpatients,[8] a single suicide item (from the Beck Depression Inventory[9]) was very

CLINICAL POINTS
◆ There is increased focus on the need to identify patients at risk for suicide.
◆ Routine use of the 9-item Patient Health Questionnaire (PHQ-9) may be useful in primary care practice in that it may identify individuals at risk for suicide who would not otherwise have been identified.
◆ Denial of suicidality on the PHQ-9 should be probed further if there are other risk factors present for suicide.

sensitive to suicidality as assessed by clinician interview (sensitivity = 0.91). Similarly, several studies suggest that a 5-item subscale of the Geriatric Depression Inventory is reasonably sensitive (ranging from 0.71–0.81) to suicidality as assessed by clinician interview.[10–12]

The 9-item Patient Health Questionnaire (PHQ-9) is a short depression screening instrument commonly used by primary care physicians. The PHQ-9 has been shown to be a valid and feasible measure for detecting depression in many groups and contains 1 item that assesses suicidal ideation.[13,14] However, we are not aware of any studies that examined the association between endorsements of suicidal ideation on the PHQ-9 and in an interview with a health professional. If the PHQ-9 suicide item proved to have good concordance with assessments of suicidality obtained through an interview, it could be used as a universal screening tool to identify individuals at risk for suicide (and depression) for further assessment and intervention in primary care. Alternatively, it could also be used as a targeted screening tool for populations at risk for suicide, ie, those with suspected depression. Therefore, the primary aim of this study is to examine the concordance between the PHQ-9 suicide item and the suicide item on the mood module of the Structured Clinical Interview for *DSM-IV* Axis I Disorders[15] (SCID-I). The suicide item in the mood module of this interview has specific subitems for indicating when patients endorse (1) thoughts of death, (2) wish to die, (3) suicide plans, and (4) suicide attempts.

METHOD

Overview

Data for these analyses are drawn from 3 separate studies with similar recruitment methods. Studies included a survey study of depressive symptoms in primary care,[16] an open trial of behavior therapy for depression in primary care,[17] and a pilot randomized, controlled trial of behavior therapy for depression in primary care (L.A.U.; R. B. Weisberg, MD; I.W.M.; unpublished data; 2010; clinicaltrials.gov Identifier: NCT00541957). All studies received institutional review board approval. Relevant methods for these studies were similar; differences are noted below.

Participants

Participants were 166 patients from 2 primary care clinics, 1 in Rhode Island and 1 in Massachusetts. One clinic is located in a general hospital in an urban area and is staffed by family medicine physicians and residents. The second clinic is a free-standing family medicine practice in a suburban setting staffed by family medicine physicians and a nurse practitioner. Of the total participants, 101 were enrolled in the survey study and 65 were screened for or enrolled in either the open trial or the randomized controlled trial. Data were collected between May 2004 and May 2009. Table 1 includes a description of demographics by group.

Procedures

In the survey study, potential participants were approached in the waiting area and asked if they were interested in a study on "depression, stress, or fatigue." Research assistants attempted to approach all patients in the waiting area during specific clinic sessions. If a person was interested, spoke English, and was not pregnant, he/she completed a brief consent for screening. Following consent, participants completed the PHQ-9. If they scored ≥ 10, they were invited to complete the second phase of the study, a telephone interview that included informed consent and oral assessment of demographics and the SCID-I. Participants were paid $50 for this interview. The telephone interview occurred within 1 month after completion of the PHQ-9.

During the open trial and the pilot randomized controlled trial, participants were recruited through the waiting room screening procedure described above, through passive recruitment processes (ie, the participant picked up a brochure in the waiting area and called the study telephone number), and through physician referral. The waiting room screening procedure was identical to that described above, with the exception that, in order to move on to the second phase (which involved an in-person assessment), participants had to score ≥ 10 on the PHQ-9 and also (a) be taking an antidepressant medication and (b) not currently be in psychotherapy. If participants met these criteria, they were scheduled for an in-person interview (within 1 month) to determine eligibility for the trial. This interview included informed

consent followed by assessment of demographics and the SCID-I. Participants were paid $35 for this interview.

If participants were referred because clinicians thought they would benefit from depression treatment, or participants called because they saw a brochure advertising the study, procedures were identical except that the participant completed consent for screening and the PHQ-9 over the telephone rather than in person.

Assessment Instruments

Demographics. We assessed demographics, including gender, race, ethnicity (Latino/non-Latino), marital status (married or cohabiting vs not), age, education, and household income via self-report. Because most respondents who reported on race were white, we collapsed the race variable into 2 options: white or minority. Many Latinos choose not to respond to the race question.

SCID-I mood module. Trained bachelor's-level raters administered the mood disorder module of the SCID-I[15] in order to assess for current major depression, lifetime mood disorders, length of depressive episode, and dysthymia. Consistent with *DSM-IV* criteria for major depressive disorder, raters ask participants if they have experienced sad mood or anhedonia most of the day, nearly every day, for at least 2 weeks in the past month. If the participant endorses at least 1 of these 2 symptoms, the rater goes on to inquire about other *DSM-IV* depressive symptoms during that 2-week period, including suicidality. The participant is coded as endorsing suicidality if he/she has "recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide." If the participant endorses any type of suicidality, raters also separately code each of the following subitems as present (or not): (1) thoughts of own death, (2) suicidal ideation, (3) specific plan, and (4) a suicide attempt.

The SCID-I was administered by telephone in the survey study and in person in the open and randomized controlled trials. Although it is possible that telephone administration results in reduced detection of certain symptoms, previous data suggest[18] that telephone administration produces results very similar to in-person assessment. We include type of study (survey vs open and randomized controlled trials) as a covariate in analyses below.

All SCID-I raters undergo extensive training. This training includes reviewing the instrument and the *DSM-IV*, listening to audiorecordings of experienced interviewers, practicing using role plays, observing experienced interviewers in person, and conducting interviews with a supervisor present until the interviewer is deemed competent to conduct interviews alone. SCID-I raters have ongoing (ie, approximately monthly)

training to reduce rater drift. SCID-I raters also review all interviews with an experienced clinical psychologist. To document reliability for the purposes of this study, a second trained rater listened to audio-recordings of 23% (n = 39) of SCID-I interviews. Of these, the interviewer asked the suicide item in 29 interviews (17% of total). Interrater reliability was good for presence of any type of suicidality (raw agreement = 86%; κ = 0.73), endorsement of the subitem "thoughts of own death" (raw agreement = 83%; κ = 0.71), endorsement of the subitem "suicidal ideation" (raw agreement = 93%; κ = 0.83), and "specific plan" (raw agreement = 100%; κ = 1.0). No participants reported a recent suicide attempt.

PHQ-9. The PHQ-9 is 9-item measure of depression with documentation of adequate reliability and validity.[13,19] The PHQ-9 was used to screen patients for elevated depressive symptoms in the previous 2 weeks; a PHQ-9 score ≥ 10 was a requirement for all studies. The ninth item of this measure specifically asks respondents if, in the past 2 weeks, they have been bothered by "thoughts that you would be better off dead or of hurting yourself in some way." Respondents are prompted to choose "not at all," "several days," "more than half the days," or "nearly every day." For the current study, suicidality on the PHQ-9 was defined as an endorsement of "several days" or more to the item. We include method of administration (in-person vs telephone) as a covariate in the analyses reported below.

RESULTS

Demographics and Clinical Characteristics of the 2 Samples

We present demographics and clinical characteristics in Table 1. As can be seen in this table, the survey study sample and the trial sample differed in gender, age, education, and presence of a current major depressive episode and did not differ in other demographic or clinical variables.

Sensitivity and Specificity in Detecting any Suicidality

We analyzed data using SPSS Statistics, version 17.0 (IBM Corporation, Somers, New York). The mean length of time between the administration of the PHQ-9 and the SCID-I was 9.3 days (SD = 7.9). Overall, on the PHQ-9, 57 participants (34.3%) endorsed some amount of suicidality in the past 2 weeks. On the SCID-I, 58 participants (34.9%) endorsed some degree of suicidality in the past month. The rate of agreement between the PHQ-9 and SCID-I reports was 78.9%, which indicated that some individuals endorsed suicidality on the PHQ-9 but not on the SCID-I and vice versa. In the sample as a whole, sensitivity was 0.84 and specificity was 0.69. We present percent of true positives, true negatives, false positives, and false negatives and sensitivity and specificity for the overall group and for specific subgroups in Table 2.

Table 1. Sample Demographics and Clinical Characteristics[a]

Variable	Survey Study (n = 101)		Open and Randomized Controlled Trials (n = 65)		Statistic	
	n	%	n	%	χ²	df
Gender					4.40*	1
Women	77	76.2	58	89.2		
Men	24	23.8	7	10.8		
Race					3.71	1
White	72	83.7	52	94.5		
Minority	14	16.3	3	5.5		
Ethnicity					0.88	1
Latino	15	15.5	6	10.2		
Non-Latino	82	84.5	53	89.8		
Marital status					1.83	1
Married/cohabitating	41	42.3	33	51.5		
Not cohabitating	56	57.7	29	48.5		
Education					12.67*	1
No college	54	58.7	16	28.6		
At least some college	38	41.3	40	71.4		
Household income					4.90	2
$0–25,000	50	53.2	25	44.6		
> $25,000–$50,000	27	28.7	12	21.4		
> $50,000	17	18.1	19	33.9		
Current major depressive episode					6.49*	1
Absent	52	52.5	21	32.3		
Present	47	47.5	44	67.7		
Lifetime diagnosis					0.90	3
No mood diagnosis	4	4.1	2	3.2		
Bipolar mood disorder	10	10.2	4	6.3		
Unipolar mood disorder	82	83.7	56	88.9		
Schizoaffective disorder	2	2.0	1	1.6		
Endorsed suicidality on the PHQ-9					2.46	1
No	71	70.3	38	58.5		
Yes	30	29.7	27	41.5		
Endorsed suicidality on the SCID-I					0.58	1
No	68	67.3	40	61.5		
Yes	33	32.7	25	38.5		
	Mean	SD	Mean	SD	t	df
Age, y	35.7	10.7	43.8	9.7	−4.85*	160
PHQ-9 score	15.9	4.1	16.8	4.7	−1.41	164

[a]Totals do not always add up to 101 (for survey) or 65 (for open and randomized controlled trials) due to missing data or the participant declining to respond.
*P < .05.
Abbreviations: PHQ-9 = 9-item Patient Health Questionnaire, SCID-I = Structured Clinical Interview for *DSM-IV* Axis I Disorders.

Sensitivity and Specificity in Detecting Specific Aspects of Suicidality (SCID-I subitems)

We also examined the specific SCID-I suicide subitems. Of the 58 participants who endorsed the suicide item on the SCID-I, all reported thoughts of own death, 30 (53% of those with suicidality and 18% of the total sample) reported having suicidal ideation on the SCID-I,* 7 (12% of those with suicidality and 4% of the total sample) reported having a specific plan on the SCID-I,* and none reported a recent attempt. We next looked at

_____

*One participant's data regarding a specific plan are missing.

the sensitivity and specificity of the PHQ-9 in detecting SCID-I subitems "thoughts of own death," "suicidal ideation," or "suicide plan" (Table 3). Note that we would expect there to be "false positives" and were therefore primarily interested in the false negative rate, ie, whether the PHQ-9 failed to detect individuals with suicidal ideation or a suicide plan. The PHQ-9 did fail to detect 11 people with suicidal ideation on the SCID-I and 2 people who endorsed a suicide plan on the SCID-I.

Logistic Regressions

We conducted a logistic regression to determine whether process factors such as method of PHQ-9 delivery (in-person or telephone), type of trial (survey study or open or randomized controlled trial trial), or days between the administration of the PHQ-9 and SCID-I had an impact on the association between the PHQ-9 and SCID-I suicide items. Specifically, we entered the following independent variables into the model: (1) PHQ-9 endorsement of suicidality (yes or no), (2) the 3 process variables, and (3) the 3 interaction terms representing the interaction between endorsement of suicidality on the PHQ-9 and each of the 3 process variables. SCID-I suicidality endorsement (yes or no) served as the dependent variable. The PHQ-9 dichotomous variable is the only one that accounted for significant variance in SCID-I suicidality endorsement (B = 2.23, SE = 0.51, Wald χ^2_1 = 19.44, P < .001; P values for all other parameters were > .20).

Next, we conducted a logistic regression to determine whether participant total PHQ-9 score (not including the suicide item) could improve our ability to predict response to the SCID-I suicide item. In the first step, we entered endorsement of PHQ-9 suicidality (yes or no); in the second step, we entered the modified total score as an independent variable in the model. As expected, endorsement of PHQ-9 suicidality was significantly associated with response to the SCID-I suicide item in the first step (B = 2.48, SE = 0.39, Wald χ^2_1 = 40.78, P < .001) but adding the modified PHQ-9 total score into the model did not significantly impact our ability to predict SCID-I–rated suicidality (for the modified PHQ-9 total score: B = 0.09, SE = 0.05, Wald χ^2_1 = 3.15, P = .08.)

DISCUSSION

The PHQ-9 has previously demonstrated adequate reliability and validity as a depression screening instrument among primary care patients. The current study compared endorsement of a single item reflecting suicidality on the PHQ-9 to endorsement of suicidality in a structured interview. The PHQ-9 suicide item showed a specificity of 0.84 and sensitivity of 0.69. We failed to find evidence that how instruments were administered

Table 2. Sensitivity and Specificity of the PHQ-9 Suicide Item in the Overall Sample and in Specific Subgroups[a]

	n	True Positives	False Positives	True Negatives	False Negatives	Specificity	Sensitivity
Overall	166	40 (24.1)	17 (10.2)	91 (54.8)	18 (10.8)	0.84	0.69
Mode of PHQ-9 administration							
By telephone	47	15 (31.9)	5 (10.6)	24 (51.1)	3 (6.4)	0.83	0.83
In person	119	25 (21.0)	12 (10.1)	67 (56.3)	15 (12.6)	0.85	0.63
Group/mode of SCID-I administration							
Open and randomized controlled trials (SCID-I in person)	65	20 (30.8)	7 (10.8)	33 (50.8)	5 (7.7)	0.83	0.80
Nonclinical trial (SCID-I by telephone)	101	20 (19.8)	10 (9.9)	58 (57.4)	13 (12.9)	0.85	0.61

[a]Data are presented as n (%) unless otherwise specified.
Abbreviations: PHQ-9 = 9-item Patient Health Questionnaire, SCID-I = Structured Clinical Interview for *DSM-IV* Axis I Disorders.

Table 3. Sensitivity and Specificity of the PHQ-9 Suicide Item in Comparison to Specific SCID-I Subitems[a]

Item	True Positives	False Positives	True Negatives	False Negatives	Specificity	Sensitivity
Thoughts of own death[b]	40 (24.1)	17 (10.2)	91 (54.8)	18 (10.8)	0.84	0.69
Suicidal ideation	19 (11.5)	37 (22.4)	98 (59.4)	11 (6.7)	0.63	0.73
Suicide plan	5 (3.0)	52 (31.3)	107 (64.5)	2 (1.2)	0.71	0.67
Suicide attempt[c]	0	57	109	0		

[a]Data are presented as n (%) unless otherwise specified. Total sample size: N = 166.
[b]These statistics are identical to those of the SCID-I suicide item as there is a 1–1 correspondence between endorsement of the SCID-I suicide item and this subitem.
[c]Number of patients who reported an attempt.
Abbreviations: PHQ-9 = 9-item Patient Health Questionnaire, SCID-I = Structured Clinical Interview for *DSM-IV* Axis I Disorders.

(ie, on the telephone or in person) had a significant impact on our ability to predict SCID-I–rated suicidality.

Our results for specificity and sensitivity for the sample as a whole are roughly equivalent to results[7] obtained with Broadhead's primary care screening tool, the Symptom Driven Diagnostic System for Primary Care.[7] This is true even though our samples were somewhat different: we included only those who screened positive for depression (ie, PHQ-9 score ≥ 10), whereas the Broadhead study included all primary care patients. Our results are also similar to those obtained with a 5-item suicide screening scale targeting an elderly population.[10–12] Given the value of parsimony, note that the 1-item screener performed as well as the 5-item screener.

Before making a recommendation about clinical use, we would like to point out the limitations of the instruments that we used as well as those of our study design. First, although the SCID-I does provide the opportunity for the interviewer to ask in detail about suicidality, it also relies on patient self-report of suicidality as well as the judgment of the interviewer. Thus, it is impossible for our "gold standard" (ie, the SCID-I) to measure suicidality without error. Although our interrater reliability on the SCID-I was acceptable, the inability to measure suicidality without error necessarily places an upper limit on the

sensitivity and specificity obtainable in our screening instrument (ie, the PHQ-9). This limitation would be common to all studies of screening instruments for suicidality given the nature of the construct.

The method of interview in our study may also be different from a community primary care practice setting. In this study, self-report and structured interviews were both administered by research staff in primary care clinics, and all participants gave written informed consent to answer questions about mood. Both of these procedures, which are specific to research, may affect a patient's willingness to disclose suicidal ideation (in either a positive or negative fashion). Also, in the open and randomized controlled trials, participants had to be taking an antidepressant medication, which means that they had some experience with mental health treatment. Further, we required that all participants have an elevated level of depression symptoms (ie, total PHQ-9 score ≥ 10), meaning that our results are primarily generalizable to the use of the PHQ-9 in a targeted (depressed) population. Finally, the SCID-I and PHQ-9 were not conducted on the same day. However, the time period focused on by these questionnaires (past month for the SCID-I and past 2 weeks for PHQ-9) had substantive overlap.

Clinically, the PHQ-9 appears to be as good as any other screener in detecting suicidal ideation and thus

can give primary care providers the opportunity to quickly screen for suicidal ideation. However, sensitivity is important when evaluating the usefulness of such a diagnostic screen because of the potentially severe consequences of false-negative results. The sensitivity found in this study (0.69) suggests that the PHQ-9 could be a useful screening instrument and is likely better than nothing, but is not perfect. Overall, the false-negative rate was approximately 11%. Of particular concern, the PHQ-9 missed 2 participants who had suicidal ideation with a plan. For these 2 participants, the SCID-I was conducted 1 day and 15 days after the PHQ-9 was completed. Although it is possible that in both cases the suicidal ideation with a plan occurred during a time period covered by the SCID-I but not the PHQ-9, it is concerning that these patients did not screen positive with any suicidality on the PHQ-9.

In conclusion, the PHQ-9 is commonly used as a diagnostic instrument in primary care and is certainly more useful than not screening for suicide at all. It could be used as a universal screening tool—to screen for depression and suicide—or as a targeted screening tool—for patients who may be depressed or have a history of depression or other mental health problems. However, due to the imperfect sensitivity of the suicide item, clinicians should use caution in interpreting negative diagnostic results. If a patient denies suicide on the PHQ-9, but there are other significant risk factors for suicide present (eg, a recent history of suicidal behavior), it will be important for a clinician to probe further. However, the routine use of the PHQ-9 as a screening instrument in primary care may identify patients with suicidal thoughts who would not otherwise have been identified and thus allow the opportunity for intervention to reduce suicidal ideation and prevent suicide.

REFERENCES

1. Nock MK, Borges G, Bromet EJ, et al. Suicide and suicidal behavior. *Epidemiol Rev*. 2008;30(1):133–154.
2. Luoma JB, Martin CE, Pearson JL. Contact with mental health and primary care providers before suicide: a review of the evidence. *Am J Psychiatry*. 2002;159(6):909–916.
3. Isometsä ET, Heikkinen ME, Marttunen MJ, et al. The last appointment before suicide: is suicide intent communicated? *Am J Psychiatry*. 1995;152(6):919–922.
4. Nutting PA, Dickinson LM, Rubenstein LV, et al. Improving detection of suicidal ideation among depressed patients in primary care. *Ann Fam Med*. 2005;3(6):529–536.
5. The Joint Commission. 2011 national patient safety goals. Behavioral health care. http://www.JointCommission.org/assets/1/6/2011_NPSGs_BHC.pdf. Accessed January 4, 2011.
6. US Preventive Services Task Force. Screening for suicide risk: recommendation and rationale. *Ann Intern Med*. 2004;140(10):820–821.
7. Broadhead WE, Leon AC, Weissman MM, et al. Development and validation of the SDDS-PC screen for multiple mental disorders in primary care. *Arch Fam Med*. 1995;4(3):211–219.
8. Yigletu H, Tucker S, Harris M, et al. Assessing suicide ideation: comparing self-report versus clinician report. *J Am Psychiatr Nurses Assoc*. 2004;10(1):9–15.
9. Beck AT, Steer RA, Carbin MG. Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clin Psychol Rev*. 1988;8(1):77–100.
10. Fujisawa D, Tanaka E, Sakamoto S, et al. The development of a brief screening instrument for depression and suicidal ideation for elderly: the Depression and Suicide Screen. *Psychiatry Clin Neurosci*. 2005;59(6):634–638.
11. Heisel MJ, Duberstein PR, Lyness JM, et al. Screening for suicide ideation among older primary care patients. *J Am Board Fam Med*. 2010;23(2):260–269.
12. Heisel MJ, Flett GL, Duberstein PR, et al. Does the Geriatric Depression Scale (GDS) distinguish between older adults with high versus low levels of suicidal ideation? *Am J Geriatr Psychiatry*. 2005;13(10):876–883.
13. Martin A, Rief W, Klaiberg A, et al. Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. *Gen Hosp Psychiatry*. 2006;28(1):71–77.
14. Huang FY, Chung H, Kroenke K, et al. Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. *J Gen Intern Med*. 2006;21(6):547–552.
15. First MB, Spitzer RL, Gibbon M, et al. *Structured Clinical Interview for DSM-IV-TR Axis I Disorders*, Research Version, Patient Edition. New York, NY: Biometrics Research, New York State Psychiatric Institute; 2002: A1-A46.
16. Uebelacker LA, Smith M, Lewis AW, et al. Treatment of depression in a low-income primary care setting with colocated mental health care. *Fam Syst Health*. 2009;27(2):161–171.
17. Uebelacker LA, Weisberg RB, Haggarty R, et al. Adapted behavior therapy for persistently depressed primary care patients: an open trial. *Behav Modif*. 2009;33(3):374–395.
18. Simon GE, Revicki D, VonKorff M. Telephone assessment of depression severity. *J Psychiatr Res*. 1993;27(3):247–252.
19. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606–613.