

Validation and Utility of a Self-report Version of PRIME-MD

The PHQ Primary Care Study

Robert L. Spitzer, MD

Kurt Kroenke, MD

Janet B. W. Williams, DSW

and the Patient Health Questionnaire Primary Care Study Group

MENTAL DISORDERS IN primary care are common, disabling, costly, and treatable.¹⁻⁵ However, they are frequently unrecognized and therefore not treated.²⁻⁶ Although there have been many screening instruments developed,^{7,8} PRIME-MD (Primary Care Evaluation of Mental Disorders)⁵ was the first instrument designed for use in primary care that actually diagnoses specific disorders using diagnostic criteria from the *Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition*⁹ (DSM-III-R) and *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*¹⁰ (DSM-IV).

PRIME-MD is a 2-stage system in which the patient first completes a 26-item self-administered questionnaire that screens for 5 of the most common groups of disorders in primary care: depressive, anxiety, alcohol, somatoform, and eating disorders. In the original study,⁵ the average amount of time spent by the physician to administer the clinician evaluation guide to patients who scored positively on the patient questionnaire was 8.4 minutes. However, this is still a considerable amount of time in the primary care setting, where most visits are 15 min-

Context The Primary Care Evaluation of Mental Disorders (PRIME-MD) was developed as a screening instrument but its administration time has limited its clinical usefulness.

Objective To determine if the self-administered PRIME-MD Patient Health Questionnaire (PHQ) has validity and utility for diagnosing mental disorders in primary care comparable to the original clinician-administered PRIME-MD.

Design Criterion standard study undertaken between May 1997 and November 1998.

Setting Eight primary care clinics in the United States.

Participants Of a total of 3000 adult patients (selected by site-specific methods to avoid sampling bias) assessed by 62 primary care physicians (21 general internal medicine, 41 family practice), 585 patients had an interview with a mental health professional within 48 hours of completing the PHQ.

Main Outcome Measures Patient Health Questionnaire diagnoses compared with independent diagnoses made by mental health professionals; functional status measures; disability days; health care use; and treatment/referral decisions.

Results A total of 825 (28%) of the 3000 individuals and 170 (29%) of the 585 had a PHQ diagnosis. There was good agreement between PHQ diagnoses and those of independent mental health professionals (for the diagnosis of any 1 or more PHQ disorder, $\kappa = 0.65$; overall accuracy, 85%; sensitivity, 75%; specificity, 90%), similar to the original PRIME-MD. Patients with PHQ diagnoses had more functional impairment, disability days, and health care use than did patients without PHQ diagnoses (for all group main effects, $P < .001$). The average time required of the physician to review the PHQ was far less than to administer the original PRIME-MD (<3 minutes for 85% vs 16% of the cases). Although 80% of the physicians reported that routine use of the PHQ would be useful, new management actions were initiated or planned for only 117 (32%) of the 363 patients with 1 or more PHQ diagnoses not previously recognized.

Conclusion Our study suggests that the PHQ has diagnostic validity comparable to the original clinician-administered PRIME-MD, and is more efficient to use.

JAMA. 1999;282:1737-1744

www.jama.com

utes or less.¹¹ Therefore, although PRIME-MD has been widely used in clinical research,¹²⁻²⁸ its use in clinical

settings has apparently been limited. This article describes the development, validation, and utility of a fully

Author Affiliations: Biometrics Research Department, New York State Psychiatric Institute and Department of Psychiatry, Columbia University, New York (Drs Spitzer and Williams); Regenstrief Institute for Health Care and Department of Medicine, Indiana University, Indianapolis (Dr Kroenke). Members of the Patient Health Questionnaire (PHQ) Primary Care Study Group and participating sites are listed at the end of this article.

Financial Disclosure: Drs Spitzer and Williams receive honoraria and consulting money from Pfizer Inc, which has supported this work.

Corresponding Author: Robert L. Spitzer, MD, Biometrics Research Department, New York State Psychiatric Institute, Unit 60, 1051 Riverside Dr, New York, NY 10032 (e-mail: rls8@columbia.edu). For a complimentary copy of PHQ materials that can be reproduced, e-mail Dr Spitzer.

self-administered version of the original PRIME-MD, called the PRIME-MD Patient Health Questionnaire (henceforth referred to as the PHQ).

DESCRIPTION OF PRIME-MD PHQ

The 2 components of the original PRIME-MD, the patient questionnaire and the clinician evaluation guide, were combined into a single, 3-page questionnaire that can be entirely self-administered by the patient (it can also be read to the patient, if necessary). The clinician scans the completed questionnaire, verifies positive responses, and applies diagnostic algorithms that are abbreviated at the bottom of each page. In this study, the data from the questionnaire were entered into a computer program that applied the diagnostic algorithms (written in SPSS 8.0 for Windows [SPSS Inc, Chicago, Ill]). The computer program does not include the diagnosis of somatoform disorder, because this diagnosis requires a clinical judgment regarding the adequacy of a biological explanation for physical symptoms that the patient has noted.

A fourth page has been added to the PHQ that includes questions about menstruation, pregnancy and childbirth, and recent psychosocial stressors. This report covers only data from the diagnostic portion (first 3 pages) of the PHQ. Users of the PHQ have the choice of using the entire 4-page instrument, just the 3-page diagnostic portion, a 2-page version (Brief PHQ) that covers mood and panic disorders and the nondiagnostic information described above, or only the first page of the 2-page version (covering only mood and panic disorders) (FIGURE 1).

The original PRIME-MD assessed 18 current mental disorders. By grouping several specific mood, anxiety, and somatoform categories into larger rubrics, the PHQ greatly simplifies the differential diagnosis by assessing only 8 disorders. Like the original PRIME-MD, these disorders are divided into threshold disorders (corresponding to specific *DSM-IV* diagnoses, such as major depressive disorder, panic disorder,

other anxiety disorder, and bulimia nervosa) and subthreshold disorders (in which the criteria for disorders encompass fewer symptoms than are required for any specific *DSM-IV* diagnoses: other depressive disorder, probable alcohol abuse or dependence, and somatoform and binge eating disorders).

One important modification was made in the response categories for depressive and somatoform symptoms that, in the original PRIME-MD, were dichotomous (yes/no). In the PHQ, response categories are expanded. Patients indicate for each of the 9 depressive symptoms whether, during the previous 2 weeks, the symptom has bothered them "not at all," "several days," "more than half the days," or "nearly every day." This change allows the PHQ to be not only a diagnostic instrument but also to yield a measure of depression severity that can be of aid in initial treatment decisions as well as in monitoring outcomes over time. Patients indicate for each of the 13 physical symptoms whether, during the previous month, they have been "not bothered," "bothered a little," or "bothered a lot" by the symptom. Because physical symptoms are so common in primary care, the original PRIME-MD dichotomous-response categories often led patients to endorse physical symptoms that were not clinically significant.

An item was added to the end of the diagnostic portion of the PHQ asking the patient if he or she had checked off any problems on the questionnaire: "How difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?" As with the original PRIME-MD, before making a final diagnosis, the clinician is expected to rule out physical causes of depression, anxiety and physical symptoms, and, in the case of depression, normal bereavement and history of a manic episode.

STUDY PURPOSE

Our major purpose was to test the validity and utility of the PHQ in a multisite sample of family practice and general internal medicine patients by answering the following questions:

1. Are diagnoses made by the PHQ as accurate as diagnoses made by the original PRIME-MD, using independent diagnoses made by mental health professionals (MHPs) as the criterion standard?

2. Are the frequencies of mental disorders found by the PHQ comparable to those obtained in other primary care studies?

3. Is the construct validity of the PHQ comparable to the original PRIME-MD in terms of functional impairment and health care use?

4. Is the PHQ as effective as the original PRIME-MD in increasing the recognition of mental disorders in primary care patients?

5. How valuable do primary care physicians find the diagnostic information in the PHQ?

6. How comfortable are patients in answering the questions on the PHQ, and how often do they believe that their answers will be helpful to their physicians in understanding and treating their problems?

METHODS

Sites and Selection of Subjects

The study was conducted at 8 primary care sites (5 general internal medicine and 3 family practice). The institutional review board of each site approved the study protocol.

From May 1997 to November 1998, 3890 patients, 18 years or older, were invited to participate in the study. There were 190 who declined to participate, 266 who started but did not complete the questionnaire, and 434 whose questionnaires were not entered into the data set because either more than 1 page was not completed or there were inadequate data to rule in or out 1 or more PHQ diagnoses. This resulted in the 3000 cases reported here (1578 family practice, 1422 general internal medicine). All sites used 1 of 2 subject selection methods to minimize sampling bias: selection of either consecutive patients for a given clinic session or every *n*th patient until the intended quota for that session was achieved.

Figure 1. First Page of Primary Care Evaluation of Mental Disorders Brief Patient Health Questionnaire

Brief Patient Health Questionnaire

This questionnaire is an important part of providing you with the best health care possible. Your answers will help in understanding problems that you may have.

Name _____ Age _____ Sex: Female Male Today's Date _____

1. Over the last 2 weeks, how often have you been bothered by any of the following problems?
- | | Not at all | Several days | More than half the days | Nearly every day |
|---|--------------------------|--------------------------|--------------------------|--------------------------|
| a. Little interest or pleasure in doing things | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| b. Feeling down, depressed, or hopeless | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| c. Trouble falling or staying asleep, or sleeping too much | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| d. Feeling tired or having little energy | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| e. Poor appetite or overeating | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| f. Feeling bad about yourself — or that you are a failure or have let yourself or your family down | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| g. Trouble concentrating on things, such as reading the newspaper or watching television | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| h. Moving or speaking so slowly that other people could have noticed? Or the opposite — being so fidgety or restless that you have been moving around a lot more than usual | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| i. Thoughts that you would be better off dead or of hurting yourself in some way | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

2. Questions about anxiety.
- | | | |
|--|--------------------------|--------------------------|
| a. In the <u>last 4 weeks</u> , have you had an anxiety attack—suddenly feeling fear or panic? | NO | YES |
| | <input type="checkbox"/> | <input type="checkbox"/> |
| If you checked "NO", go to question #3. | | |
| b. Has this ever happened before? | <input type="checkbox"/> | <input type="checkbox"/> |
| c. Do some of these attacks come <u>suddenly out of the blue</u> — that is, in situations where you don't expect to be nervous or uncomfortable? | <input type="checkbox"/> | <input type="checkbox"/> |
| d. Do these attacks bother you a lot or are you worried about having another attack? | <input type="checkbox"/> | <input type="checkbox"/> |
| e. During your last bad anxiety attack, did you have symptoms like shortness of breath, sweating, your heart racing or pounding, dizziness or faintness, tingling or numbness, or nausea or upset stomach? | <input type="checkbox"/> | <input type="checkbox"/> |

3. If you checked off any problems on this questionnaire so far, how difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?

- | | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| Not difficult at all | Somewhat difficult | Very difficult | Extremely difficult |
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

Copyright held by Pfizer Inc, but may be photocopied ad libitum. For office coding, see the end of the article.

Table 1. Operating Characteristics of the Self-administered PRIME-MD PHQ (n = 585) Compared With the Original Clinician-Administered PRIME-MD (n = 431)*

	Sensitivity, % (95% CI)		Specificity, % (95% CI)		Overall Accuracy (95% CI)		κ (95% CI)	
	PHQ	Original	PHQ	Original	PHQ	Original	PHQ	Original
Any PRIME-MD psychiatric diagnosis	75 (69-81)	83 (78-88)	90 (87-93)	88 (84-92)	85 (82-88)	86 (83-89)	0.65 (0.58-0.72)	0.71 (0.64-0.78)
Any mood disorder	61 (52-70)	67 (59-75)	94 (92-96)	92 (89-95)	88 (85-91)	84 (81-87)	0.58 (0.47-0.67)	0.61 (0.53-0.69)
Major depressive disorder	73 (59-87)	57 (45-69)	98 (96-100)	94 (93-95)	93 (91-95)	92 (89-95)	0.54 (0.42-0.66)	0.61 (0.50-0.72)
Any anxiety disorder	63 (53-73)	69 (59-79)	97 (95-99)	90 (87-93)†	91 (89-93)	86 (83-89)	0.65 (0.56-0.74)	0.55 (0.45-0.65)
Panic disorder	81 (69-93)	57 (37-97)	99 (98-100)	99 (98-100)	98 (97-99)	96 (94-98)	0.84 (0.75-0.93)	0.60 (0.42-0.78)†
Probable alcohol abuse/dependence	62 (48-76)	81 (62-100)	97 (95-99)	98 (97-99)	95 (93-97)	98 (97-99)	0.60 (0.47-0.73)	0.71 (0.54-0.88)
Any eating disorder	89 (77-100)	73 (54-92)	96 (95-97)	99 (98-100)†	96 (94-98)	98 (96-100)	0.61 (0.48-0.74)	0.73 (0.58-0.88)

*All data are reported using mental health professionals' diagnoses as the criterion standard. PRIME-MD indicates Primary Care Evaluation of Mental Disorders; PHQ, Patient Health Questionnaire; and CI, confidence interval.

†The only differences are specificity original < PHQ for any anxiety disorder ($P < .001$), specificity PHQ < original for any eating disorder ($P = .014$), and κ PHQ > original for panic disorder ($P = .019$).

Data Collection

Patients. Before seeing their physicians, all patients completed PHQs. Additionally, they completed items regarding physician visits and disability days during the previous 3 months, their comfort with answering the PHQ questions, and how valuable they believed the PHQ would be to their physicians in understanding and treating the problems they were having. In addition, each patient completed the Medical Outcomes Study Short-Form General Health Survey (SF-20),²⁹ which measures functional status in 6 dimensions.

Physicians. A total of 62 physicians participated in the study (21 general internal medicine, 41 family practice [19 of whom were family practice residents]). Their mean (SD) age was 37 (6.5) years, and 63% were male.

After evaluating each patient but before reviewing the PHQ, the physician noted whether the patient was new or established, the physician's knowledge of any current mental disorders, and types of current physical disorders (hypertension, heart disease, diabetes, liver disease, renal disease, arthritis, pulmonary disease, cancer, or other). The physician then reviewed the PHQ and asked any additional questions necessary to clarify responses on the questionnaire. Also noted were any treatments or referrals for mental disorders that were being initiated or planned during that particular visit. Midway

through the study, physicians noted how long it took them to review each PHQ and ask clarifying questions. Finally, at the conclusion of the study, all physicians completed confidential questionnaires asking them about the value and usefulness of the PHQ.

Mental Health Professionals. To determine the agreement of PHQ diagnoses with those of MHPs, midway through the study an MHP (a PhD clinical psychologist or 1 of 3 senior psychiatric social workers) attempted to interview by telephone all subsequently entered subjects who had a telephone, agreed to be interviewed, and could be contacted within 48 hours. All except 1 site participated in these validation interviews. The MHP was blinded to the results of the PHQ. The rationale and additional details of the MHP telephone interview, which used the overview from the Structured Clinical Interview for *DSM-III-R*³⁰ and diagnostic questions from the original PRIME-MD, are described in the original PRIME-MD report.⁵

RESULTS

Description of Patients

The mean (SD) age of the patients was 46 (17.2) years, with a range of 18 to 99 years; 66% were female; 79% were white (not Hispanic), 13% were African American, 4% were Hispanic; 48% were married, 12% were divorced, 23% were never married; and 25% were college gradu-

ates. There was considerable site variability (site ranges: mean age, 40-62 years; female, 54%-89%; college graduate, 2%-50%). Of the total sample, 80% were established clinic patients, and the remainder were being seen for the first time. The most common types of physical disorders were hypertension (25%), arthritis (11%), diabetes (8%), and pulmonary disease (7%).

Accuracy: Agreement With MHPs

The 585 subjects who had an MHP interview within 48 hours of completing the PHQ were, within each site, similar to patients not reinterviewed in terms of demographic profile, functional status, and frequency of psychiatric diagnoses. One modification from the original PRIME-MD algorithm was necessary. The number of symptoms required for diagnosing major depressive disorder could remain the same as in *DSM-IV*, ie, 5 of 9 during the previous 2 weeks. However, because the PHQ response set was expanded from the simple yes/no in the original PRIME-MD to 4 frequency levels as described above, lowering the PHQ threshold from "nearly every day" to "more than half the days" considerably improved sensitivity (from 37% to 73%) while maintaining high specificity (94%).

The operating characteristics of the PHQ are generally satisfactory and comparable to those obtained in the original PRIME-MD study (TABLE 1). Of

note, the sensitivity of the PHQ for major depressive disorder was somewhat higher (73% vs 57%). As in the original study, the prevalences for PHQ diagnoses and MHP diagnoses were nearly identical, indicating that the PHQ did not have a systematic tendency to overdiagnose or underdiagnose any psychiatric disorder.

We also examined agreement between the PHQ results and the MHP on a computer-derived index of depression symptom severity (the sum of the scores for the 9 PHQ- or MHP-recorded depressive symptoms; possible range, 0-27). The correlation between the PHQ and MHP for this index was 0.84.

Diagnostic Results of PHQ Evaluations

Overall, 28% of the subjects had a PHQ diagnosis, of which 15% had a threshold diagnosis and 13% a subthreshold diagnosis only (TABLE 2). The overall prevalence of psychiatric disorder was somewhat lower than in the original study (28% vs 39%). The proportion of patients with a psychiatric disorder who had more than 1 disorder was also somewhat lower (36% vs 56%). As in the original PRIME-MD study, the prevalence varied considerably across sites, which, at least in part, is undoubtedly attributable to significant differences in patient demographic variables across the sites.

Physician Time Reviewing PHQ

The physician time required to review the PHQ ($n = 1527$) was less than 1 minute for 42% of the subjects, 1 to 2 minutes for 43%, 3 to 5 minutes for 13%, and more than 5 minutes for only 3%. Thus, the time required of the physician to review the PHQ is far less than the time to administer the clinician-administered PRIME-MD (less than 3 minutes for 85% of the subjects given the PHQ vs 16% of the subjects given the PRIME-MD in our original study).

Relationship of PHQ Results to Functional Status, Health Care Use, and Disability Days

FIGURE 2 shows the mean scores on the 6 scales of the SF-20 for 4 groups of sub-

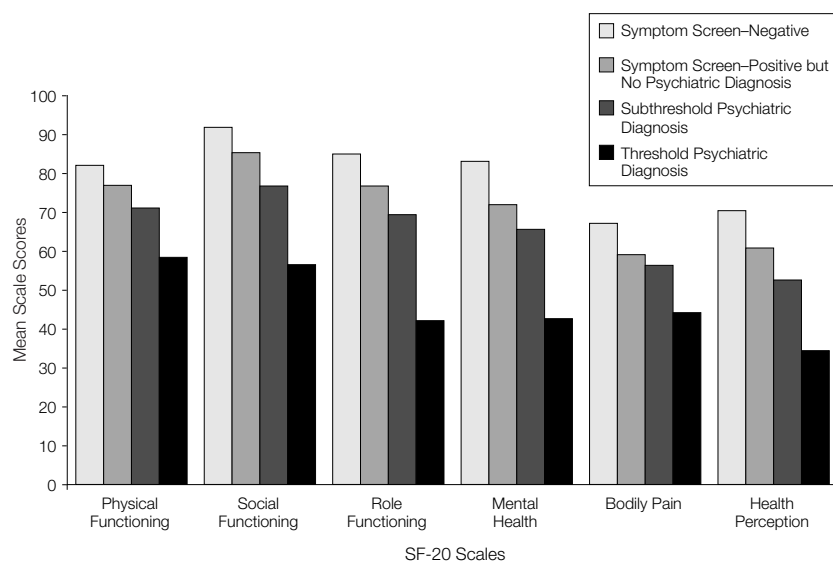
jects. Each of the disorders (except for alcohol abuse) has 1, 2, or 3 symptoms that must be present for the diagnosis to be made (eg, depressed mood or loss of interest for major depressive disorder). Patients who did not endorse any of these required symptoms on the PHQ were considered to be symptom-screen negative. Patients who had 1 or more of these required symptoms but did not qualify for a subthreshold or threshold diagnosis were considered to be symptom-screen positive but to have no psychiatric diagnosis. The third and fourth groups met criteria for subthreshold and any threshold diagnoses, respectively. Scores were adjusted by analysis of covariance for number of physical disorders, sex, age, minority status, education level, and site. As with the original PRIME-MD study, the symptom-screen negative group had the highest level of functioning on all of the scales, followed by the symptom-screen positive group, the subthreshold group, and, finally, the

Table 2. Prevalence of Psychiatric Disorders Detected by PRIME-MD PHQ in 3000 Primary Care Patients*

Mental Disorder	Total Sample, No. (%)	Site Range, %
Any psychiatric diagnosis	825 (28)	16-38
Any threshold diagnosis	448 (15)	7-22
Subthreshold only	377 (13)	9-16
Any mood disorder	476 (16)	11-28
Major depressive disorder	292 (10)	5-13
Other depressive disorder	184 (6)	5-16
Any anxiety disorder	317 (11)	4-16
Panic disorder	165 (6)	2-9
Other anxiety disorder	221 (7)	3-10
Probable alcohol abuse/dependence	201 (7)	3-10
Any eating disorder	211 (7)	2-11
Binge eating disorder	187 (6)	2-9
Bulimia nervosa	24 (1)	0-2

*PRIME-MD PHQ indicates Primary Care Evaluation of Mental Disorders Patient Health Questionnaire. Ninety-five percent confidence intervals around the prevalence estimates are $\pm 2\%$.

Figure 2. Relationship of PRIME-MD PHQ Results to Functional Status



Because of missing data for some patients, the range of numbers of patients across scales was as follows: symptom screen-negative, 1044 to 1095; symptom screen-positive but no psychiatric diagnosis, 862 to 892; subthreshold psychiatric diagnosis, 331 to 337; and threshold psychiatric diagnosis, 393 to 409. All paired comparisons among the 4 groups were significant at $P < .05$ using Bonferroni correction for type I errors, with the exception of the difference between the symptom screen-positive patients but no psychiatric diagnoses and patients with subthreshold psychiatric diagnosis on the bodily pain scale. SF-20 indicates Short-Form General Health Survey; PRIME-MD PHQ, Primary Care Evaluation of Mental Disorders Patient Health Questionnaire.

Table 3. Self-reported Health Care Use and Disability Days by PRIME-MD PHQ Diagnostic Results*

	Symptom Screen-Negative, Group 1	Symptom Screen-Positive but No Psychiatric Diagnosis, Group 2	Subthreshold Psychiatric Diagnosis, Group 3	Threshold Psychiatric Diagnosis, Group 4	P Value†
Visits to a physician in past 3 mo, mean No.	0.98 [0.82-1.14] (n = 1057)	1.37 [1.19-1.55] (n = 870)	1.70 [1.43-1.97] (n = 326)	2.48 [2.22-2.74] (n = 391)	<.001 (All paired comparisons <.005 except for group 2 vs group 3; P = .29)
Days kept from usual activities because of not feeling well, mean	2.24 [1.38-3.10] (n = 1048)	4.83 [3.89-5.77] (n = 852)	6.64 [5.11-8.17] (n = 328)	16.96 [15.49-18.43] (n = 358)	<.001 (All paired comparisons <.001 except for group 2 vs group 3; P = .29)

*Scores are adjusted by analysis of covariance for number of physical disorders, sex, age, minority status, educational level, and site. PRIME-MD PHQ indicates Primary Care Evaluation of Mental Disorders Patient Health Questionnaire. Numbers in brackets are 95% confidence intervals.

†Nominal significance levels are reported for pairwise differences. Type I errors are controlled for by Bonferroni criteria.

Table 4. Frequency of Psychiatric Diagnoses Unrecognized by Physician*

	PRIME-MD			
	PHQ (n = 2901)†		Original (n = 731)	
	No. of Patients With Psychiatric Diagnosis	No. (%) of Patients Unrecognized by Physician	No. of Patients With Psychiatric Diagnosis	No. (%) of Patients Unrecognized by Physician
Any PRIME-MD psychiatric diagnosis	803	368 (46) [42-49]	287	138 (48) [42-54]
Any mood disorder	464	225 (49) [44-53]	191	127 (67) [60-73]
Any anxiety disorder	310	176 (57) [51-62]	137	83 (61) [52-69]
Probable alcohol abuse/dependence	194	151 (78) [72-84]	36	15 (42) [26-58]
Any eating disorder	209	187 (90) [85-94]	24	19 (79) [63-95]

*PRIME-MD indicates Primary Care Evaluation of Mental Disorders; PHQ, Patient Health Questionnaire. Numbers in brackets are 95% confidence intervals.

†Not 3000 because 99 patients were missing data for physician diagnosis.

threshold group. The group main effects were all significant ($P < .001$). TABLE 3 presents the mean values on 1 index of health care use and 1 of disability in the same 4 groups, with initial scores again adjusted for the variables just noted. As in the original PRIME-MD study, the same pattern of increasing use and disability is seen from the symptom-screen negative group to the threshold psychiatric diagnoses group, and the group main effects were all significant ($P < .001$).

We also examined how the probability of a subthreshold or threshold PHQ diagnosis varied depending on responses to the question: "How difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?" The percentage of subjects with a PHQ diagnosis varied significantly ($P < .001$), ranging from 17% of

the subjects who responded "not difficult at all," to 38% who responded "somewhat difficult," to 69% who responded "very difficult," to 91% who responded "extremely difficult." This question was also associated with functional impairment: the mean correlation of this item with each of the 6 SF-20 scales was 0.38, ranging from 0.27 for pain to 0.53 for mental health. We also found that the computer-derived index of depression severity had high correlations with the SF-20 scales (mean correlation was 0.49, ranging from 0.33 for pain to 0.73 for mental health).

Recognition of Mental Disorders

Of the 803 patients with a PHQ diagnosis, 46% (n = 368) had not been recognized by their physicians as having any diagnosis included in the PHQ system after being clinically evaluated but before physician review of the PHQ

(TABLE 4). The nonrecognition rate was even higher for specific diagnostic categories. The ability of the PHQ to detect a substantial number of unrecognized cases is comparable to that of the original PRIME-MD.

Perceived Value of PHQ to Physicians

At the conclusion of the study, most physicians reported that the diagnostic information provided by the PHQ was "very" (46%) or "somewhat" (41%) useful in management and treatment planning. The majority (80%) also reported that if they were able to have the clerical staff in their setting give the questionnaire to patients, it would be helpful if given routinely to all new patients, all patients who had not received a questionnaire in the last year, and to any patient for whom it seemed indicated at the time of the visit. Prior to the study, nearly half (48%) of the physicians acknowledged that they only "occasionally" asked their patients about many of the diagnostic symptoms included in the PHQ when their chief complaint did not suggest a mental disorder.

Impact of PHQ on Physician Management Decisions

Although our study was designed to test the diagnostic validity of the PHQ and not to influence treatment decisions, we still asked the physicians, for each case, what treatment or referral actions they initiated or planned to initiate for any problems reported in the PHQ. There were 363 patients with 1 or more PHQ

diagnoses who had not been previously determined by their physicians to have any PHQ diagnoses and for whom information on treatment decisions was available. For only 32% (n = 117) did physicians note that any new management actions were initiated or planned. For only 16% (n = 58) were follow-up visits planned; for only 3% (n = 11) were antidepressants prescribed; and for only 3% (n = 11) were referrals to MHPs provided. Of the 74 patients with PHQ diagnoses of major depression that had not previously been recognized, follow-up visits, antidepressant prescriptions, or mental health referrals were noted for only 22%, 10%, and 5% of the cases, respectively.

Reaction of Patients to PHQ

The majority (88%) of patients said they were “very” or “somewhat” comfortable answering the questions on the PHQ. Likewise, 89% believed that the questions were “very” or “somewhat” helpful in getting their physicians to better understand or treat the problems they were having.

COMMENT

The self-administered PHQ has diagnostic validity comparable to that of the original clinician-administered instrument. This was demonstrated both by agreement with an independent MHP interview (criterion validity) as well as by the strong association of PHQ diagnoses with indices of functional impairment and health care use (construct validity). As with the original PRIME-MD, most patients were comfortable answering questions and judged the information to be valuable to their physicians. Most physicians also found it useful and thought it would be helpful if used routinely. The PHQ is efficient, requiring much less of a clinician’s time than the original PRIME-MD.

In addition to its value in yielding provisional psychiatric diagnoses, the PHQ yields an index of depressive symptom severity. This index, which had a remarkably high correlation with an MHP assessment of the same dimension, may be useful in initial management deci-

sions as well as monitoring treatment outcome in depressed patients.

Previous self-report instruments used in primary care for case finding or screening yield indices of severity rather than categorical psychiatric diagnoses.^{6,7} The PHQ is the first entirely self-administered diagnostic instrument designed for use in primary care. We found that agreement between the PHQ diagnosis of major depressive disorder and the MHP diagnosis was maximized when the frequency threshold for the individual major depression items of the PHQ was lowered from the *DSM-IV* requirement of “nearly every day” to “more than half the days.” This finding may be due to the fact that in most structured diagnostic interviews, such as the Structured Clinical Interview for *DSM-III-R* and the interview used by our MHPs, patients who acknowledge depressive symptoms are asked, for each symptom, whether it has been present “nearly every day.” Given only this dichotomous choice, many patients whose symptoms have been present only “more than half the days” during the time period being considered may respond affirmatively, believing that to be their only opportunity to indicate that they frequently have had the symptom in question. Many patients with major depressive disorder diagnosed by the PHQ who reported on the questionnaire that the symptoms were present only “more than half the days” did answer affirmatively to the MHP when asked if the symptom had been present “nearly every day.” The implication of this finding is that a well-designed self-report instrument, which allows the subject to consider a range of frequency responses for symptoms, may yield a more accurate assessment of frequency than a clinician-administered structured interview, which, for efficiency of administration, presents a dichotomous choice of a single frequency taken from a diagnostic criterion.

An especially interesting finding in our study is the strong predictive value of the question at the end of the PHQ: “How difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?”

This global self-assessment of the degree of impairment associated with the patient’s psychological symptoms is a potent indicator of the likelihood of a psychiatric diagnosis and of functional impairment.

Several possible limitations in the study should be noted. The PHQ was scored with a computer program to ensure that the diagnostic algorithms were applied correctly. In pilot testing the questionnaire, we found that clinicians who had little training in applying the algorithms sometimes made errors. Our study does not indicate how much training is necessary in different primary care settings to ensure that the diagnostic algorithms are applied with minimal errors. Also, as with any diagnostic test, the PHQ does not detect all cases of mental disorders. Therefore, clinicians should ask additional questions of patients who they feel may have “false-negative” PHQ results.

Although the PHQ is clearly more efficient for clinicians to use than the original PRIME-MD, our study indicates that it may also be easier for clinicians to ignore. In our study, the PHQ had less impact on physician therapeutic actions than did the original PRIME-MD, which requires that clinicians gather much of the diagnostic information through direct interview. This may lead to more intimate and detailed awareness and understanding of their patients’ symptoms than simply reviewing a self-report questionnaire. This, in turn, might increase the likelihood of initiating some form of treatment or referral. Our study confirms what has been demonstrated in numerous other studies³¹⁻³³: merely providing clinicians with information about psychiatric diagnoses has only a moderate impact on their behavior. The relatively weak effect of information alone on producing changes in clinical practice is not unique to mental disorders but is also true for medical conditions in general.³⁴ Aside from improved detection, what is also required to improve the management and outcomes of patients with mental disorders in primary care are system changes, such as longer or more frequent visits, collaborative support from MHPs,

and better reimbursement for psychosocial evaluation and treatment.³⁵⁻³⁹

Ideally, the PHQ would be administered in clinical practice to all new patients, patients in whom a psychiatric diagnosis is suspected, and established patients on a periodic basis (eg, annually), as is done with other screening procedures. In contrast, because of the length of time required to administer the original PRIME-MD, it has been used primarily as a research tool; clinicians have predominantly used it only with patients in whom they already suspected a psychiatric diagnosis—rather than using it routinely to detect unrecognized cases.

The original PRIME-MD study⁵ demonstrated that primary care physicians could make valid psychiatric diagnoses with the aid of brief, structured interviews. From the perspective of psychiatric as-

essment in primary care, our current study demonstrates that a well-designed self-report questionnaire can also provide comparably valid diagnoses. The original PRIME-MD has been widely used in primary care research, and we expect that the PHQ may offer an advantage in future studies because of its comparable validity but greater efficiency. With proper integration into primary care practice, accompanied by other system changes, the PHQ could become a useful clinical adjunct to improve the recognition and management of mental disorders.

PHQ Office Coding Algorithm

Major depressive syndrome is indicated if answers to #1a or b and five or more of #1a-i are at least "More than half the days" (count #1i if present at all); other depressive syndrome, if #1a or b and two, three,

or four of #1a-i are at least "More than half the days" (count #1i if present at all); panic syndrome, if all of #2a-e are "YES."

Funding/Support: The development of the PHQ was underwritten by an educational grant from Pfizer US Pharmaceuticals Inc, New York, NY. PRIME-MD is a trademark of Pfizer Inc. Copyright held by Pfizer Inc. **Members of the PHQ Primary Care Study Group** coordinated the study at each of the participating sites: Raymond Hornyak, PhD, Conemaugh Memorial Health System, Johnstown, Pa; Robert Joseph, MD, Cambridge Hospital, Cambridge, Mass; Michael Roy, MD, MPH, Walter Reed Army Medical Center General Medicine Clinic, Washington, DC; Lawson Wulsin, MD, Franciscan University of Cincinnati Family Practice Center, Cincinnati, Ohio; Julia McMurray, MD, Women's Health Center, Madison, Wis; Mark Linzer, MD, University of Wisconsin General Internal Medicine Clinic, Madison; Joseph Kovaz, MD, Mazomanie Community Clinic, Mazomanie, Wis; and Kim Seeger, MD, Columbia-Presbyterian Family Health Center, New York, NY.

Acknowledgment: Mark Linzer, MD; Frank Verloin de Gruy III, MD; Steven R. Hahn, MD; and David Brody, MD, helped develop the original PRIME-MD. Miriam Gibbon, MSW; Jeffrey G. Johnson, Wis; Marcus Kraeber, MD; Robert Yaskanich, PhD; and Gregory J. Rys, MA, assisted in data collection. Mark Davies assisted in statistical analysis.

REFERENCES

1. Depression Guideline Panel. *Depression in Primary Care: Volume 1, Detection and Diagnosis*. Rockville, Md: US Dept of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research; 1993. Clinical Practice Guideline, No. 5. AHCPR publication No. 93-0550.
2. Goldman LS, Wise TN, Brody DS, eds. *Handbook of Psychiatry in Primary Care*. Chicago, Ill: American Medical Association; 1998.
3. Miranda J, Hohmann AA, Attkisson CC, Larson DB, eds. *Mental Disorders in Primary Care*. San Francisco, Calif: Jossey-Bass; 1994.
4. Ormel J, VonKorff M, Ustun TB, Pini S, Korten A, Oldehinkel T. Common mental disorders and disability across cultures. *JAMA*. 1994;272:1741-1748.
5. Spitzer RL, Williams JBW, Kroenke K, et al. Utility of a new procedure for diagnosing mental disorders in primary care: the PRIME-MD 1000 study. *JAMA*. 1994;272:1749-1756.
6. Üstün TB, Sartorius N, eds. *Mental Illness in General Health Care: An International Study*. New York, NY: John Wiley & Sons; 1995.
7. Mulrow CD, Williams JW Jr, Gerety MB, Ramirez G, Montiel OM, Kerber C. Case-finding instruments for depression in primary care settings. *Ann Intern Med*. 1995;122:913-921.
8. Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression. *J Gen Intern Med*. 1997;12:439-445.
9. *Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition*. Washington, DC: American Psychiatric Association; 1987.
10. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*. Washington, DC: American Psychiatric Association; 1994.
11. Schappert SM. National Ambulatory Medical Care Survey. *Vital Health Stat 13*. 1992;110:1-80.
12. Hahn SR, Kroenke K, Spitzer RL, Williams JBW. The PRIME-MD instrument. In: Maruish ME, ed. *The Use of Psychological Testing for Treatment Planning and Outcome Assessment*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1999.
13. Philbrick JT, Connolly JE, Wofford AB. The prevalence of mental disorders in rural office practice. *J Gen Intern Med*. 1996;11:9-15.
14. Beck AT, Guth D, Steer RA, Ball R. Screening for major depression disorders in medical inpatients with the Beck Depression Inventory for Primary Care. *Behav Res Ther*. 1997;35:785-791.
15. Parker T, May PA, Maviglia MA, Petrakis S, Sunde S, Gloyd SV. PRIME-MD: its utility in detecting mental disorders in American Indians. *Int J Psychiatry Med*. 1997;27:107-128.
16. Kobak KA, Taylor LH, Dotti SL, et al. A computer-administered telephone interview to identify mental disorders. *JAMA*. 1997;278:905-910.
17. The Iowa Persian Gulf Study Group. Self-reported illness and health status among Gulf War veterans: a population-based study. *JAMA*. 1997;277:238-245.
18. Kroenke K, Jackson JL, Chamberlin J. Depressive and anxiety disorders in patients presenting with physical complaints. *Am J Med*. 1997;103:339-347.
19. O'Malley PG, Wong PW, Kroenke K, Roy MJ, Wong RK. The value of screening for psychiatric disorders prior to upper endoscopy. *J Psychosom Res*. 1998;44:279-287.
20. O'Malley PG, Jackson JL, Kroenke K, Yoon K, Hornstein E, Dennis GJ. The value of screening for psychiatric disorders in rheumatology referrals. *Arch Intern Med*. 1998;158:2357-2362.
21. 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 Intern Med*. 1997;12:679-685.
22. Valenstein M, Kales H, Mellow A, et al. Psychiatric diagnosis and intervention in older and younger patients in a primary care clinic. *J Am Geriatr Soc*. 1998;46:1499-1505.
23. Miranda J, Azocar F, Komaromy M, Golding JM. Unmet mental health needs of women in public-sector gynecologic clinics. *Am J Obstet Gynecol*. 1998;178:212-217.
24. Nease DE Jr, Volk RJ, Cass AR. Investigation of a severity-based classification of mood and anxiety symptoms in primary care patients. *J Am Board Fam Pract*. 1999;12:21-31.
25. Jackson JL, O'Malley PG, Kroenke K. A psychometric comparison of military and civilian medical practices. *Mil Med*. 1999;164:112-115.
26. Leopold KA, Ahles TA, Walch S, et al. Prevalence of mood disorders and utility of the PRIME-MD in patients undergoing radiation therapy. *Int J Radiat Oncol Biol Phys*. 1998;42:1105-1112.
27. Bensenor IM, Pereira AC, Tannuri AC, et al. Systemic arterial hypertension and psychiatric morbidity in the outpatient care setting of a tertiary hospital. *Arg Neuropsiquiatr*. 1998;56:406-411.
28. Schmidt VM. "PRIME-MD"—a quick method for diagnosis of mental disease in general practice. *Ugeskr Laeger*. 1998;160:5370-5371.
29. Stewart AL, Hays RD, Ware JE Jr. The MOS Short-Form General Health Survey. *Med Care*. 1988;26:724-735.
30. Spitzer RL, Williams JB, Gibbon M, First MB. The Structured Clinical Interview for DSM-III-R (SCID), I. *Arch Gen Psychiatry*. 1992;49:624-629.
31. Higgins ES. A review of unrecognized mental illness in primary care. *Arch Fam Med*. 1994;3:908-917.
32. Katon W, Gonzales J. A review of randomized trials of psychiatric consultation-liaison studies in primary care. *Psychosomatics*. 1994;35:268-278.
33. Kroenke K, Taylor-Vaisey A, Dietrich AJ, Oxman TE. Interventions to improve provider diagnosis and treatment of mental disorders in primary care: a critical review of the literature. *Psychosomatics*. In press.
34. Greco PJ, Eisenberg JM. Changing physicians' practices. *N Engl J Med*. 1993;329:1271-1273.
35. Kroenke K. Discovering depression in medical patients: reasonable expectations. *Ann Intern Med*. 1997;126:463-465.
36. Katon W, Von Korff M, Lin E, et al. Collaborative management to achieve treatment guidelines. *JAMA*. 1995;273:1026-1031.
37. Katon W, Robinson P, Von Korff M, et al. A multifaceted intervention to improve treatment of depression in primary care. *Arch Gen Psychiatry*. 1996;53:924-932.
38. Klinkman MS. Competing demands in psychosocial care. *Gen Hosp Psychiatry*. 1997;19:98-111.
39. Williams JW Jr. Competing demands: does care for depression fit in primary care? *J Gen Intern Med*. 1998;13:137-139.