

Positive Screens for Psychiatric Disorders in Primary Care: A Long-Term Follow-Up of Patients Who Were Not in Treatment

Myrna M. Weissman, Ph.D.
Yuval Neria, Ph.D.
Marc J. Gameroff, Ph.D.
Daniel J. Pilowsky, M.D., M.P.H.

Priya Wickramaratne, Ph.D.
Rafael Lantigua, M.D.
Steven Shea, M.D.
Mark Olfson, M.D., M.P.H.

Objective: Screening for psychiatric disorders has gained acceptance in some general medical settings, but critics argue about its value. The purpose of this study was to determine the clinical utility of screening by conducting a long-term follow-up of patients who screened positive for psychiatric disorders but who were initially not in treatment. **Methods:** A cohort of 519 low-income, adult primary care patients were screened for major depression and bipolar, anxiety, and substance use disorders and reassessed with the Structured Clinical Interview for DSM-IV after a mean of 3.7 years by a clinician blind to the initial screen. Data on treatment utilization was obtained through hospital records. The sample consisted of 348 patients who had not received psychiatric care in the year before screening. **Results:** Among 39 patients who screened positive for major depression, 62% (95% confidence interval=45.5%–77.6%) met criteria for current major depressive disorder at follow-up. Those who screened positive reported significantly poorer mental and social functioning and worse general health at follow-up than the screen-negative patients and were more likely to have visited the emergency department for psychiatric reasons (12.1% and 3.0%, odds ratio [OR]=6.4) and to have major depression (OR=7.6). Generally similar results were observed for patients who screened positive for other disorders. **Conclusions:** Commonly used screening methods identified patients with psychiatric disorders; about four years later, those not initially in treatment were likely to have enduring symptoms and to use emergency psychiatric services. Screening should be followed up by clinical diagnostic assessment in the context of available mental health treatment. (*Psychiatric Services* 61:151–159, 2010)

Screening for depression in primary care has a 40-year history (1–8). The rationale for screening follows from the observation that patients with depression are usually first seen in the primary care setting and that depression escapes clinical detection, causes substantial morbidity if untreated, and may complicate the course of chronic medical disease (9–11). Depression screening in U.S. general medical practice has been advocated for, and such screening is required in the U.S. Department of Veterans Affairs health care system (12). In 2002 the U.S. Preventive Services Task Force recommended screening in practices that have systems in place to assess and treat depression (13). A number of educational and treatment strategies have been proposed—and some have been tested—to improve management of depression after screening in primary care (14). However, critics have argued that screening is costly, has not been shown to improve outcome, and diverts clinical attention away from the mental health care of patients with recognized disorders and more serious symptoms (3,15–18).

Cross-sectional and short-term follow-up studies comparing patients with undetected depression to those with identified depression suggest that undetected depression tends to be milder and often resolves without intervention (19–21). An understanding of the long-term course of untreated patients who screen positive for depression would provide insight

Dr. Weissman, Dr. Neria, Dr. Gameroff, Dr. Pilowsky, Dr. Wickramaratne, and Dr. Olfson are affiliated with the Department of Psychiatry, College of Physicians and Surgeons, Columbia University, New York City. Dr. Weissman, Dr. Neria, and Dr. Pilowsky are also with the Department of Epidemiology, Mailman School of Public Health, Columbia University, where Dr. Shea is affiliated. Dr. Weissman, Dr. Gameroff, Dr. Wickramaratne, and Dr. Olfson are also with the Division of Epidemiology, New York State Psychiatric Institute, New York City. Dr. Shea is also with the Division of General Medicine, Department of Medicine, College of Physicians and Surgeons, Columbia University, where Dr. Lantigua is affiliated. Send correspondence to Dr. Weissman at New York State Psychiatric Institute, 1051 Riverside Dr., Unit 24, New York, NY 10032 (e-mail: mmw3@columbia.edu).

into the utility of screening. Few studies have followed patients screened for depression in primary care beyond one year to determine longer-term outcomes (22–25) or have focused specifically on untreated patients who screen positive for depression. Finally, although screens have been developed and disseminated for a range of other psychiatric disorders (26–30), most longitudinal assessments have been limited to depression and have not addressed other major psychiatric disorders with high prevalence rates in primary care settings, such as anxiety and substance use disorders.

This article reports results of a blind follow-up assessment of a consecutive sample of low-income primary care patients who screened positive for psychiatric disorders commonly seen in primary care and who were not receiving mental health care in the year before screening. The purpose of the study was not to test the validity of the screening methods, because they have been previously validated in primary care against independent clinical assessments (28), but rather to determine the clinical utility of the screens in identifying patients with a poor long-term outcome.

Methods

Baseline study

Design and participants. Baseline assessments and follow-up assessments were conducted in 2001–2003 and 2004–2007, respectively, at the faculty and resident group practice of the Division of General Medicine, Columbia University Medical Center in New York City. Each year 18,000 adult patients account for approximately 80,000 outpatient visits to the practice. Patients are from the surrounding communities in northern Manhattan, and most are Medicaid or Medicare beneficiaries. Baseline study methods have been described in detail elsewhere (31,32).

Briefly, patients were systematically approached on the basis of the position of the seat that they freely selected in the clinic waiting room. Eligible patients were between 18 and 70 years of age, had made at least one prior visit to the practice, could speak and understand Spanish or English,

were waiting for face-to-face contact with a primary care physician, and had no health problems that would prevent them from completing the survey. Of the 1,346 patients who met eligibility criteria, 1,156 consented to participate and completed the survey (86% response rate).

Information was collected on demographic characteristics. Social, family, and work functioning was assessed with the Sheehan Disability Scale (SDS) (33). The Medical Outcomes Study Short-Form Health Survey (SF-12) (34) and selected questions from the Social Adjustment Scale–Self Report (35) were also used. For this study, the work item on the SDS was excluded from total SDS scores because only one-fifth of the sample was gainfully employed. Information about mental health treatment history, including the types of psychotropic medications prescribed during the past month, was reported by the patient.

All assessment forms were translated from English to Spanish and back-translated by a bilingual team of mental health professionals. Records were linked with the hospital administrative database to obtain the frequency of outpatient, inpatient, and emergency service contacts during the 12 months before the follow-up assessment as well as the ICD-9-CM diagnoses for each contact. The institutional review boards of Columbia University Medical Center and the New York State Psychiatric Institute approved the study protocol, and all participants provided informed written consent.

Screens for psychiatric disorders. Patients were screened for major depressive disorder, generalized anxiety disorder, panic disorder, and alcohol use disorder, using the PRIME-MD Patient Health Questionnaire (PHQ) (28,29). A drug use disorder module patterned after the PHQ alcohol use disorder module was also administered. Because of low prevalence rates, two pairs of disorders were combined for analytic purposes: alcohol and drug use disorders, and generalized anxiety disorder and panic disorder.

The Mood Disorder Questionnaire (MDQ) (26,27), which has also been

validated in primary care (36), was used to screen for bipolar spectrum disorder. Although the time frame of the MDQ is not restricted to the present, we considered a positive MDQ screen indicative of bipolar spectrum disorder even if the patient was not in a current episode, because bipolar disorder is considered a lifetime disorder.

Treatment for psychiatric disorders. Patients were asked, “Have you ever been treated for an emotional or mental problem?” and if yes, “When was the last time you were treated for an emotional or mental problem?” Patients were also asked, “Have you ever taken a prescribed medication for an emotional or mental problem?” and if yes, “When was the last time that you took a prescribed medication for an emotional or mental problem?” Both questions about the timing of treatment had the following response choices: more than one year ago, less than one year ago but not in the last month, and in the last month. Patients who had been treated or took medication less than one year ago were defined as currently receiving mental health treatment.

Follow-up study

Design and response rate. The mean \pm SD time between baseline and follow-up interviews was $3.7 \pm .5$ years. Our follow-up assessment included an extended interview administered by clinically trained, bilingual and bicultural interviewers who were blind to the initial clinical data. Of the 1,156 participants at baseline, 825 (71%) consented at baseline to be re-contacted for a follow-up interview. Because of financial constraints, letters were sent to only the first 796 patients; thus 29 of the 825 consenting patients were not invited. Among those who were sent letters about the follow-up study, the follow-up clinical status of 70% ($N=561$) was determined. These patients were either interviewed at follow-up ($N=519$) or found to be deceased or unable to participate because of institutionalization or illness ($N=42$). The 561 follow-up participants represent 49% of the original sample of 1,156. Our interest was in 348 of the 561 patients who screened positive at baseline for

at least one of the mental disorders and had not received mental health treatment during the year before screening.

Clinically assessed psychiatric disorders. At follow-up, major sections of the Structured Clinical Interview for DSM-IV-TR (SCID) Axis I Disorders (37) were administered by experienced, clinically trained interviewers who prepared a narrative summary of the interview. Based on all available follow-up data, one of the authors (MMW) completed best-estimate diagnoses blind to the initial screening status (38). The diagnostic reviewer and the interviewer discussed any wide discrepancies. Current disorders were categorized as major depressive disorder, anxiety disorder (generalized anxiety disorder or panic disorder), substance-related disorder (including alcohol or substance abuse or dependence and alcohol- or substance-induced mood, anxiety, or psychotic disorder), and bipolar spectrum disorder (bipolar type I or II disorder or bipolar disorder not otherwise specified). Disorders were considered present if they were rated as “definite” or “probable,” except for substance-related disorders, which were also considered present if rated as “possible” because of a tendency for patients to underreport substance-related disorders (39).

Data analysis

Our primary analyses were restricted to the 348 patients who screened positive for a disorder at baseline and who reported that they had not received mental health treatment in the year before screening. For each of the screened disorders (major depressive disorder, generalized anxiety disorder or panic disorder, alcohol or drug use disorder, and bipolar disorder), we compared patients by screening status (positive or negative) on various measures that were assessed at follow-up, including clinical diagnoses based on the SCID and health and functioning. Using hospital administrative records for the 12 months before each follow-up assessment, we compared baseline screen-positive and screen-negative groups on total number of outpatient visits, emergency department visits,

and hospital admissions. We separated all visits and admissions into those that included at least one mental disorder diagnosis (ICD-9-CM codes 290–319) and those that included no mental disorder diagnosis. Because emergency department visits and hospital admissions were relatively rare, we coded each variable as any versus none. The Charlson Comorbidity Index (40) was used to compare the general medical status of the groups. In all analyses, we adjusted for the number of days between baseline screening and the follow-up interview.

Data weighting. We used the propensity weighting method (41) to adjust for potential attrition bias (unit nonresponse bias) after the baseline assessment. To the extent that correlates of nonresponse can be identified, the follow-up sample (N=519) was weighted to represent the baseline sample (N=1,156). Nonresponse patterns were examined at four sequential stages: refusal to be recontacted (N=331); absence of a follow-up letter (N=29); inability to locate (N=146); and refusal of the follow-up assessment, poor health status, or death (N=131).

For each of the four stages, we performed the following procedure. A binary logistic regression model was constructed with unit response versus nonresponse as the outcome, and variables from the baseline study were entered as candidate predictors by using stepwise selection criteria (to enter, $p=.20$; to stay, $p=.05$) and allowing for up to two-way interactions. At each stage, the weight for each patient was the reciprocal of the probability of the predicted response. Information about the baseline variables that significantly predicted response versus nonresponse at each of the four stages is available upon request.

The weights from each of these four models were multiplied to produce an overall weight for each of the 519 follow-up participants. To minimize undue effects from excessively high or low weights, we performed a quantile split of the overall weights to produce six weighting classes. Each member of a weighting class was assigned the mean weight for that class

(41). Final weights ranged from 1.486 to 3.610, and summed to 1,156, the size of the baseline sample. All analyses were weighted and were conducted with SAS, version 9, specifically the SURVEYFREQ, SURVEYMEANS, SURVEYREG, and SURVEYLOGISTIC procedures. The level of significance was set at .05 (two-tailed).

Results

Baseline characteristics of untreated patients

Among the 348 patients who reported no mental health treatment at baseline, rates of positive screens varied by disorder: major depressive disorder, 11% (95% confidence interval [CI]=7.6–14.5); alcohol or drug use disorder, 9% (95% CI=5.9%–12.6%); bipolar disorder, 7% (95% CI=3.9%–9.4%); and generalized anxiety disorder or panic disorder, 6% (95% CI=3.9%–8.9%). As shown in Table 1, few significant sociodemographic differences were found by screening status (positive or negative) within the disorder groups. Compared with their screen-negative counterparts, patients who screened positive for major depressive or bipolar disorder had lower educational achievement, patients who screened positive for an alcohol or drug use disorder were more likely to be male and non-Hispanic, and the group screening positive for bipolar disorder had a higher proportion of males.

Diagnoses and functional status from the baseline screening for each disorder group are shown in Table 2. High rates of comorbidity were found among the baseline screening diagnoses. Among patients who screened positive for major depressive disorder, for example, a substantial percentage also screened positive for generalized anxiety disorder or panic disorder (26%) and for bipolar disorder (19%). Compared with patients who screened negative for major depressive disorder, generalized anxiety disorder or panic disorder, and bipolar disorder, those who screened positive for each of these disorders had significantly lower mental functioning and those who screened positive for depressive disorder and generalized anxiety disorder or panic disorder had significantly worse social functioning.

Table 1

Demographic characteristics of 348 primary care patients screened for psychiatric disorders who reported no psychiatric treatment in the year before screening, by disorder group at baseline (in weighted percentages)^a

| Characteristic | Major depressive disorder | | | Generalized anxiety or panic disorder | | | Alcohol or drug use disorder | | | Bipolar disorder | | |
|-------------------------|---------------------------|------------|-----|---------------------------------------|------------|-----|------------------------------|------------|-------|------------------|------------|------|
| | Yes (N=39) | No (N=306) | p | Yes (N=27) | No (N=319) | p | Yes (N=29) | No (N=305) | p | Yes (N=24) | No (N=324) | p |
| Gender | | | .19 | | | .26 | | | <.001 | | | .050 |
| Female | 79 | 68 | | 79 | 68 | | 20 | 74 | | 49 | 70 | |
| Male | 21 | 32 | | 21 | 32 | | 80 | 26 | | 51 | 30 | |
| Age | | | .27 | | | .74 | | | .99 | | | .59 |
| 18–44 | 29 | 33 | | 28 | 32 | | 31 | 33 | | 41 | 32 | |
| 45–54 | 42 | 26 | | 36 | 27 | | 31 | 28 | | 31 | 28 | |
| 55–64 | 19 | 24 | | 26 | 23 | | 23 | 23 | | 13 | 24 | |
| 65–70 | 11 | 17 | | 10 | 17 | | 15 | 16 | | 16 | 17 | |
| Race-ethnicity | | | .28 | | | .75 | | | .02 | | | .87 |
| Hispanic | 90 | 83 | | 82 | 84 | | 67 | 86 | | 82 | 84 | |
| Non-Hispanic | 10 | 17 | | 18 | 16 | | 33 | 15 | | 18 | 16 | |
| Immigrant | | | .66 | | | .72 | | | .34 | | | .59 |
| Yes | 78 | 75 | | 72 | 75 | | 67 | 76 | | 80 | 75 | |
| No | 22 | 25 | | 28 | 25 | | 33 | 24 | | 21 | 25 | |
| Marital status | | | .3 | | | .24 | | | .27 | | | .12 |
| Married or cohabiting | 26 | 35 | | 23 | 35 | | 45 | 34 | | 50 | 33 | |
| Other | 74 | 65 | | 77 | 65 | | 56 | 66 | | 50 | 67 | |
| High school graduate | | | .04 | | | .75 | | | .33 | | | .052 |
| No | 71 | 52 | | 57 | 54 | | 44 | 54 | | 74 | 52 | |
| Yes | 29 | 48 | | 43 | 47 | | 56 | 46 | | 26 | 48 | |
| Annual household income | | | .86 | | | .14 | | | .72 | | | .92 |
| <\$6,000 | 39 | 40 | | 55 | 39 | | 38 | 41 | | 41 | 40 | |
| ≥\$6,000 | 61 | 60 | | 45 | 61 | | 63 | 59 | | 59 | 60 | |
| Paid worker | | | .9 | | | .22 | | | .2 | | | .91 |
| Yes | 26 | 27 | | 16 | 27 | | 15 | 29 | | 28 | 27 | |
| No | 74 | 73 | | 84 | 73 | | 85 | 72 | | 72 | 73 | |

^a The percentages are weighted to account for attrition at follow-up. The p values were generated from a logistic regression (for binary outcomes) or a linear regression (for continuous outcomes) using weighted data. Ns for some screening diagnoses sum to less than 348 because of missing data.

However, similar significant differences were not observed with respect to alcohol or drug use disorder. A positive screen for major depressive disorder was consistently associated with poor physical health and functioning.

Follow-up status of patients untreated at baseline

There was substantial stability between the baseline screening diagnosis and the clinician's diagnosis at follow-up (Table 3). Among patients who screened positive for major depressive disorder at baseline, 62% (CI=45.5%–77.6%) were diagnosed by the research clinician as having current major depressive disorder at follow-up (odds ratio [OR]=9.7). Significant correlations between baseline screen status and the same follow-up diagnosis were evident for generalized

anxiety disorder or panic disorder (OR=3.6), alcohol or drug use disorder (OR=31.9), and bipolar disorder (OR=8.0). All of the screen diagnoses at baseline were related significantly to at least one clinical diagnosis at follow-up, either the same diagnosis or one of the other screening diagnoses (Table 3, any of the above).

Compared with patients who screened negative, the patients who screened positive for major depressive disorder or generalized anxiety disorder or panic disorder consistently reported poorer mental functioning at follow-up, as indicated by expected mean differences on the SF-12 mental component summary score, and poorer social functioning at follow-up, as indicated by expected mean differences in scores on the SDS. However, screen status (posi-

tive or negative) for major depressive disorder and generalized anxiety disorder or panic disorder was not associated with the SF-12 physical component summary score or the Charlson Comorbidity Index score at follow-up. Screen status was also not significantly related to use of outpatient, emergency department, or inpatient services for general medical reasons as assessed by hospital administrative records, with one exception. Patients who screened positive for an alcohol or drug use disorder were significantly more likely to make at least one medical emergency department visit during the year before the follow-up assessment (OR=2.6).

Compared with patients who screened negative, untreated patients who screened positive for major depressive disorder, generalized anxiety

Table 2

Clinical characteristics of 348 primary care patients screened for psychiatric disorders who reported no psychiatric treatment in the year before screening, by disorder group at baseline (in weighted percentages or means)^a

| Characteristic | Major depressive disorder | | | Generalized anxiety or panic disorder | | | Alcohol or drug use disorder | | | Bipolar disorder | | |
|---------------------------------------|---------------------------|------------|-------|---------------------------------------|------------|-------|------------------------------|------------|------|------------------|------------|------|
| | Yes (N=39) | No (N=306) | p | Yes (N=27) | No (N=319) | p | Yes (N=29) | No (N=305) | p | Yes (N=24) | No (N=324) | p |
| Screen diagnosis | | | | | | | | | | | | |
| Major depressive disorder | | | — | | | <.001 | | | .37 | | | .004 |
| Yes | — | — | | 45 | 9 | | 15 | 10 | | 31 | 10 | |
| No | — | — | | 55 | 91 | | 85 | 90 | | 69 | 90 | |
| Generalized anxiety or panic disorder | | | <.001 | | | — | | | .57 | | | .37 |
| Yes | 26 | 4 | | — | — | | 9 | 6 | | 11 | 6 | |
| No | 74 | 96 | | — | — | | 92 | 94 | | 89 | 94 | |
| Alcohol or drug use disorder | | | .37 | | | .57 | | | — | | | .008 |
| Yes | 14 | 9 | | 13 | 9 | | — | — | | 28 | 8 | |
| No | 86 | 92 | | 87 | 91 | | — | — | | 72 | 92 | |
| Lifetime bipolar disorder | | | .004 | | | .37 | | | .008 | | | — |
| Yes | 19 | 5 | | 11 | 6 | | 19 | 5 | | — | — | |
| No | 81 | 95 | | 89 | 94 | | 81 | 95 | | — | — | |
| Mental and social functioning | | | | | | | | | | | | |
| SF-12 MCS score (M±SE) ^b | 34.9±1.6 | 50.9±.6 | <.001 | 36.1±2.0 | 50.1±.6 | <.001 | 46.1±2.3 | 49.5±.6 | .15 | 41.7±2.6 | 49.7±.6 | .002 |
| SDS score (M±SE) ^c | 9.1±1.0 | 2.7±.2 | <.001 | 8.0±1.1 | 3.1±.2 | <.001 | 3.7±1.0 | 3.2±.3 | .66 | 4.9±1.1 | 3.2±.3 | .13 |
| Poor or fair emotional health | | | <.001 | | | .01 | | | .16 | | | .65 |
| Yes | 71 | 16 | | 43 | 20 | | 32 | 20 | | 26 | 22 | |
| No | 29 | 84 | | 57 | 80 | | 68 | 80 | | 74 | 78 | |
| Physical functioning | | | | | | | | | | | | |
| SF-12 PCS score (M±SE) ^b | 37.0±1.7 | 40.6±.7 | .046 | 35.2±2.1 | 40.5±.7 | .01 | 40.0±2.1 | 40.4±.7 | .87 | 39.0±2.2 | 40.4±.7 | .54 |
| Poor or fair general health | | | <.001 | | | .09 | | | .99 | | | .36 |
| Yes | 93 | 54 | | 76 | 57 | | 58 | 57 | | 49 | 59 | |
| No | 7 | 46 | | 25 | 43 | | 43 | 43 | | 52 | 41 | |

^a The percentages are weighted to account for attrition at follow-up. The p values were generated from a logistic regression (for binary outcomes) or a linear regression (for continuous outcomes) using weighted data. Ns for some screening diagnoses sum to less than 348 because of missing data.

^b The Medical Outcomes Study Short-Form Health Survey (SF-12) mental component summary (MCS) and physical component summary (PCS) scores are expressed as T scores in relation to national norms (mean±SD=50±10), with higher scores indicating better functioning.

^c Sheehan Disability Scale. Possible scores range from 0 to 20, with higher scores indicating greater functional impairment.

disorder or panic disorder, or alcohol or drug use disorder at baseline were all significantly more likely to have made one or more psychiatric emergency department visits during the year before the follow-up assessment (major depressive disorder, OR=6.4; generalized anxiety disorder or panic disorder, OR=8.5; and alcohol or drug use disorder, OR=8.0).

In a separate series of models (not shown), each of which controlled for all four baseline screens and patient demographic characteristics, a positive screen for depression independently predicted major depressive disorder (OR=7.6, 95% CI=3.0–19.4), alcohol or drug use disorder (OR=153.0, 95% CI=10.5±∞), poor or fair emotional health (OR=4.5, 95% CI=1.7–11.8), poor or fair general health

(OR=5.6, 95% CI=1.5–21.5), and a lower SF-12 mental component summary score (β =−14.2, 95% CI=−19.7 to −8.7) at follow-up. A positive screen for an anxiety disorder predicted fair or poor emotional health (OR=4.7, 95% CI=1.8–12.4) and a lower SF-12 mental component summary score (β =−4.5, 95% CI=−8.8 to −.2). A positive substance use disorder screen predicted major depressive disorder (OR=3.4, 95% CI=1.1–10.7) and lifetime bipolar disorder (OR=3.9, 95% CI=1.1–13.5). A positive bipolar disorder screen predicted only lifetime bipolar disorder at follow-up (OR=6.4, 95% CI=1.9–21.1).

Discussion

In a large urban university-affiliated practice serving low-income patients,

many patients who were not receiving psychiatric treatment in the year before screening screened positive for psychiatric disorders. Nearly four years later, these patients were significantly more likely than those who screened negative to be clinically diagnosed as having the original screen diagnosis, to be impaired, and to have made emergency department visits for psychiatric reasons in the year before the follow-up interview; they were not more likely than those who screened negative to have used general medical services in the past year. These findings were consistent for three of the screened disorders—major depressive disorder, generalized anxiety disorder or panic disorder, and alcohol or drug use disorder.

Comorbidity of the screened disorder

Table 3

Follow-up diagnoses and outcome measures for 348 primary care patients screened at baseline for psychiatric disorders and who reported no psychiatric treatment in the year before screening, by disorder group at baseline (in weighted percentages)^a

| Follow-up diagnosis or outcome | Major depressive disorder | | | | Generalized anxiety or panic disorder | | | | Alcohol or drug use disorder | | | | Bipolar disorder | | | |
|---|---------------------------|---------|--------------------|----------------|---------------------------------------|---------|-------------------|---------------|------------------------------|---------|-------------------|--------------|------------------|---------|-------------------|--------------|
| | Yes | No | AOR | 95% CI | Yes | No | AOR | 95% CI | Yes | No | AOR | 95% CI | Yes | No | AOR | 95% CI |
| Diagnosis^b | | | | | | | | | | | | | | | | |
| Major depressive disorder | 62 | 13 | 9.7 | 4.2 to 22.4 | 47 | 16 | 3.7 | 1.6 to 8.6 | 25 | 17 | 3.9 | 1.3 to 12.3 | 23 | 18 | 1.3 | .4 to 4.0 |
| Generalized anxiety or panic disorder | 21 | 5 | 3.9 | 1.4 to 11.4 | 17 | 6 | 3.6 | 1.1 to 12.2 | 15 | 6 | 4.5 | 1.1 to 19.2 | 13 | 7 | 1.4 | .3 to 5.8 |
| Substance-related disorder | 2 | 1 | 3.3 | .2 to 46.3 | 0 | 1 | — ^c | | 9 | 1 | 31.9 | 2.4 to 417.6 | 6 | 1 | 26.6 | .7 to 972.7 |
| Lifetime bipolar spectrum disorder | 22 | 7 | 2.9 | 1.0 to 8.1 | 12 | 8 | 1.6 | .4 to 6.2 | 25 | 7 | 5.4 | 1.7 to 17.1 | 40 | 6 | 8.0 | 2.7 to 23.6 |
| Any of the above | 69 | 23 | 6.1 | 2.7 to 13.9 | 58 | 26 | 3.3 | 1.4 to 7.8 | 40 | 27 | 3.3 | 1.2 to 8.6 | 52 | 26 | 2.9 | 1.1 to 7.4 |
| Mental functioning | | | | | | | | | | | | | | | | |
| SF-12 MCS ^d score (M±SE) | 35.8±2.4 | 50.3±.6 | -14.8 ^e | -19.7 to -10.0 | 41.4±2.2 | 49.3±.6 | -8.1 ^e | -12.2 to -4.0 | 46.8±2.3 | 48.9±.7 | -4.6 ^e | -9.2 to -.01 | 43.2±3.4 | 49.1±.6 | -5.1 ^e | -11.3 to 1.2 |
| SDS score ^f (M±SE) | 7.9±1.0 | 3.0±.3 | 5.1 ^e | 3.1 to 7.0 | 6.6±1.2 | 3.3±.3 | 3.6 ^e | 1.4 to 5.8 | 5.2±1.0 | 3.3±.3 | 3.0 ^e | 1.0 to 5.1 | 4.7±1.3 | 3.4±.3 | .8 ^e | -1.6 to 3.2 |
| Poor or fair emotional health | 63 | 24 | 6.2 | 2.7 to 14.4 | 64 | 26 | 6.1 | 2.4 to 15.5 | 44 | 27 | 2.6 | .97 to 6.8 | 51 | 27 | 2.6 | .9 to 7.1 |
| Physical functioning | | | | | | | | | | | | | | | | |
| SF-12 PCS ^d score (M±SE) | 36.6±2.2 | 39.8±.7 | -2.8 ^e | -7.1 to 1.5 | 36.4±2.3 | 39.7±.7 | -3.1 ^e | -6.8 to .6 | 40.3±2.1 | 39.8±.8 | .6 ^e | -3.6 to 4.8 | 37.8±2.9 | 39.7±.7 | -.9 ^e | -6.5 to 4.8 |
| Poor or fair general health | 86 | 52 | 8.2 | 2.3 to 28.9 | 74 | 54 | 3.0 | 1.0 to 8.9 | 51 | 55 | 1.2 | .5 to 2.9 | 57 | 55 | 1.1 | .4 to 3.2 |
| Medical comorbidity score ^g (M±SE) | 1.7±.4 | 1.2±.1 | .5 ^e | -.3 to 1.3 | 1.9±.4 | 1.2±.1 | .3 ^e | -.5 to 1.0 | 2.1±.6 | 1.1±.1 | .9 ^e | -.2 to 2.0 | 1.0±.2 | 1.3±.1 | -.1 ^e | -.6 to .5 |
| Health care utilization^h | | | | | | | | | | | | | | | | |
| Medical only Outpatient visits (M±SE) | 8.2±1.3 | 7.7±.5 | -.8 ^e | -3.5 to 1.9 | 8.3±1.8 | 7.8±.5 | -.9 ^e | -3.9 to 2.0 | 5.5±1.3 | 7.8±.6 | -1.4 ^e | -4.6 to 1.8 | 7.7±1.6 | 7.8±.5 | -.03 ^e | -3.2 to 3.1 |
| Any ER visits | 39 | 28 | 1.4 | .6 to 3.1 | 41 | 28 | 1.7 | .7 to 4.2 | 39 | 27 | 2.6 | 1.01 to 6.7 | 35 | 28 | 1.3 | .5 to 3.7 |
| Any hospital admissions | 9 | 12 | .4 | .1 to 2.0 | 11 | 12 | .4 | .1 to 3.1 | 19 | 11 | 2.0 | .6 to 6.2 | 7 | 12 | .7 | .1 to 3.4 |
| Psychiatric treatment | | | | | | | | | | | | | | | | |
| Outpatient visits (M±SE) | 2.4±1.0 | 1.2±.4 | .4 ^e | -2.3 to 3.1 | 2.9±1.4 | 1.2±.4 | .6 ^e | -1.9 to 3.2 | .4±.2 | 1.4±.4 | -.3 ^e | -1.3 to .7 | 1.7±1.0 | 1.3±.4 | .3 ^e | -2.6 to 3.2 |

Table 3*Continued from previous page*

| Follow-up diagnosis or outcome | Major depressive disorder | | | | Generalized anxiety or panic disorder | | | | Alcohol or drug use disorder | | | | Bipolar disorder | | | |
|--------------------------------|---------------------------|----|-----|-------------|---------------------------------------|----|-----|-------------|------------------------------|----|-----|-------------|------------------|----|-----|------------|
| | Yes | No | AOR | 95% CI | Yes | No | AOR | 95% CI | Yes | No | AOR | 95% CI | Yes | No | AOR | 95% CI |
| Any ER visits | 12 | 3 | 6.4 | 1.1 to 36.0 | 21 | 3 | 8.5 | 2.0 to 36.9 | 8 | 3 | 8.0 | 1.3 to 49.8 | 7 | 4 | 3.0 | .3 to 27.1 |
| Any hospital admissions | 15 | 4 | 3.8 | .8 to 19.0 | 21 | 4 | 4.2 | .9 to 19.3 | 5 | 4 | 2.0 | .3 to 12.6 | 10 | 4 | 3.1 | .7 to 13.7 |

^a The percentages are weighted to account for attrition at follow-up. The adjusted odds ratios (AORs) are for the group with a positive screen; they were generated from a logistic regression that used weighted data and controlled for gender, age group, ethnicity (Hispanic versus non-Hispanic), education, income, and time between baseline and follow-up interview. Ns for some screening diagnoses sum to less than 348 because of missing data.

^b Based on the Structured Clinical Interview for DSM-IV

^c The test could not be calculated because of a zero cell.

^d The Medical Outcomes Study Short-Form Health Survey (SF-12) mental component summary (MCS) and physical component summary (PCS) scores are expressed as T scores in relation to national norms (mean±SD=50±10), with higher scores indicating better functioning.

^e Expected score difference (with 95% confidence interval) for the group with the positive screen; generated from a linear regression that used weighted data and controlled for gender, age group, ethnicity (Hispanic versus non-Hispanic), education, income, and time between baseline and follow-up interview.

^f Sheehan Disability Scale. Possible scores range from 0 to 20, with higher scores indicating greater functional impairment.

^g Charlson Comorbidity Index score. Higher scores indicate greater medical comorbidity. Based on ICD-9 codes in patients' records during the year before follow-up. Sample range, 0–11

^h Data from hospital administrative records during the year before follow-up. ER, emergency room

ders was high. Follow-up diagnoses were predicted by the corresponding screening diagnosis at baseline. A positive screen for major depressive disorder also emerged as a powerful predictor of other follow-up diagnoses and of poor mental and physical health. These findings suggest that if resource constraints limit screening to a single mental disorder, a depression screen may be more likely to identify the most vulnerable patients over the long term.

One-year longitudinal primary care studies suggest that patients who screen positive for depressive symptoms or for a diagnosis of depression tend to have poor mental health outcomes. A study conducted in six culturally diverse countries (Israel, Brazil, Australia, Spain, Russian, and the United States) found that high depressive symptom scores were associated with poor health and functional status and increased use of health care independent of patient demographic characteristics (6). Other one-year follow-up studies of patients who screened positive for depression have had similar results (42–45). Adverse outcomes among patients who screen positive for de-

pression have been documented in a 45-month follow-up study of older primary care patients (20), a five-year follow-up study at Walter Reed Army Medical Center (24), and a seven-year follow-up study from Finland (25). However, some short-term follow-up studies of six to 12 months have reported that undetected mild depression resolves without intervention (20). None of these studies conducted separate analyses for primary care patients who were not in treatment at the time of initial screening, and only the Finnish study (25) used a diagnostic assessment at follow-up.

This study has several limitations. Because it was conducted in a single urban primary care practice that serves a predominantly low-income immigrant population, the findings may not generalize to primary care clinics that serve other socioeconomic groups, where the rates of psychiatric disorders may be lower. However, low-income patients represent a large and vulnerable patient group who usually receive most of their psychiatric care in primary care (46). The exclusion of patients over 70 years of age prevents generalization to older primary care populations. The initial

screens were based on self-report, although they have been widely used and tested in primary care settings (28,36). Although self-report entails less interviewer involvement, we do not know the extent to which patients' perceptions of interviewers might affect their responses. Incomplete participation at follow-up raises a concern about selection bias; however, we incorporated nonresponse-adjusted weights that were based on a detailed examination of attrition at various stages. Although we focused on patients who reported no use of psychiatric treatment in the year before the initial screening, we were unable to determine with certainty whether these patients had mental health problems that had been recognized by a health practitioner before the baseline screening. We also do not know the extent to which treatment during the follow-up period may have affected follow-up outcome. In general, outpatient psychiatric treatment was low for all groups.

Conclusions

Screening for depression in primary care has a long history (13). Our study has shown that commonly used

screening methods identified patients who were likely four years later to be symptomatic and impaired and to have sought emergency department psychiatric services in the year before follow-up, even when the analysis excluded screen-positive patients who were receiving mental health treatment at baseline. Data from community samples have shown that the likelihood of receiving treatment for depression increases with the severity and number of symptoms (47). This has raised questions about the value of treating mild symptoms. Our findings indicate that simple screening procedures, especially for major depression, can identify primary care patients with milder symptoms that are nevertheless persistent and enduring mental health problems. As recommended by the U.S. Preventive Services Task Force, positive screens should be followed by a diagnostic assessment, which is critical in determining the specific diagnosis and guiding treatment.

There is considerable overlap between major depressive disorder and anxiety disorders in diagnosis and treatment, and screening separately for these disorders may yield marginal clinical benefit. However, because of pharmacological considerations, screening may be important in differentiating bipolar depression from major depression (31,32). Screening may also be important for identifying substance use disorders, for which the poorest portion of our sample were unlikely to be receiving treatment.

Our findings support screening, especially in primary care practices, where the prevalence of mental disorders is generally high. Screening-related improvement in long-term outcomes will require that detection be followed by effective treatment. As noted by others (48,49), screening should be considered only as part of a package of enhanced care. Several clinical trials have demonstrated the efficacy of different models for delivering mental health services in primary care (5,48,49).

Acknowledgments and disclosures

This study was supported by grant 1RO1-MHO72833 (Dr. Neria, principal investigator) from the National Institute of Mental Health

and by grants from GlaxoSmithKline and Eli Lilly & Company (Dr. Weissman, principal investigator). Funding sources played no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of manuscript.

Dr. Pilowsky has received speaker fees from AstraZeneca. Dr. Olfson has received support from or has served as a consultant for AstraZeneca, Bristol-Myers Squibb, Eli Lilly and Company, Janssen, McNeil Pharmaceuticals, and Pfizer. The other authors report no competing interests.

References

1. Coulehan JL, Schulberg HC, Block MR: The efficiency of depression questionnaires for case finding in primary medical care. *Journal of General Internal Medicine* 4:541-547, 1989
2. Coyne JC, Fechner-Bates S, Schwenk TL: Prevalence, nature, and comorbidity of depressive disorders in primary care. *General Hospital Psychiatry* 16:267-276, 1994
3. Coyne JC, Palmer SC, Sullivan PA: Screening for depression in adults. *Annals of Internal Medicine* 138:767-768, 2003
4. Gilbody SM, House AO, Sheldon TA: Routinely administered questionnaires for depression and anxiety: systematic review. *BMJ* 322:406-409, 2001
5. Goldberg DP, Blackwell B: Psychiatric illness in general practice: a detailed study using a new method of case identification. *British Medical Journal* 1:439-443, 1970
6. Herrman H, Patrick DL, Diehr P, et al: Longitudinal investigation of depression outcomes in primary care in six countries, the LIDO study: functional status, health service use and treatment of people with depressive symptoms. *Psychological Medicine* 32:889-902, 2002
7. Hoepfer EW, Nycz GR, Cleary PD, et al: Estimated prevalence of RDC mental disorder in primary care. *International Journal of Mental Health* 8:6-15, 1979
8. Shepherd M, Cooper B, Brown AC, et al: *Psychiatric Illness in General Practice*. London, Oxford University Press, 1996
9. Bush DE, Ziegelstein RC, Tayback M, et al: Even minimal symptoms of depression increase mortality risk after acute myocardial infarction. *American Journal of Cardiology* 88:337-341, 2001
10. Clouse RE, Lustman PJ, Freedland KE, et al: Depression and coronary heart disease in women with diabetes. *Psychosomatic Medicine* 65:376-383, 2003
11. Gross R, Olfson M, Gameroff MJ, et al: Depression and glycemic control in Hispanic primary care patients with diabetes. *Journal of General Internal Medicine* 20:460-466, 2005
12. Kirkcaldy RD, Tynes LL: Depression screening in a VA primary care clinic. *Psychiatric Services* 57:1694-1696, 2006
13. Pignone MP, Gaynes BN, Rushton JL, et al:

Screening for depression in adults: a summary of the evidence for the US Preventive Services Task Force. *Annals of Internal Medicine* 136:765-776, 2002

14. Gilbody S, Whitty P, Grimshaw J, et al: Educational and organizational interventions to improve the management of depression in primary care: a systematic review. *JAMA* 289:3145-3151, 2003
15. Valenstein M, Vijan S, Zeber JE, et al: The cost-utility of screening for depression in primary care. *Annals of Internal Medicine* 134:345-360, 2001
16. Campbell TL: Is screening for mental health problems worthwhile in family practice? An opposing view. *Journal of Family Practice* 25:184-187, 1987
17. Katon W: Will improving detection of depression in primary care lead to improved depressive outcomes? *General Hospital Psychiatry* 17:1-2, 1995
18. Schwenk TL: Screening for depression in primary care: a disease in search of a test. *Journal of General Internal Medicine* 11:437-439, 1996
19. McQuaid JR, Stein MB, Laffaye C, et al: Depression in a primary care clinic: the prevalence and impact of an unrecognized disorder. *Journal of Affective Disorders* 55:1-10, 1999
20. Simon GE, VonKorff M: Recognition, management, and outcomes of depression in primary care. *Archives of Family Medicine* 4:99-105, 1995
21. Klinkman MS, Coyne JC, Gallo S, et al: False positives, false negatives, and the validity of the diagnosis of major depression in primary care. *Archives of Family Medicine* 7:451-461, 1998
22. Conradi HJ, de Jonge P, Ormel J: Prediction of the three-year course of recurrent depression in primary care patients: different risk factors for different outcomes. *Journal of Affective Disorders* 105:267-271, 2008
23. Callahan CM, Wolinsky FD, Stump TE, et al: Mortality, symptoms, and functional impairment in late-life depression. *Journal of General Internal Medicine* 13:746-752, 1998
24. Jackson JL, Passamonti M, Kroenke K: Outcome and impact of mental disorders in primary care at 5 years. *Psychosomatic Medicine* 69:270-276, 2007
25. Poutanen O, Mattila A, Seppala NH, et al: Seven-year outcome of depression in primary and psychiatric outpatient care: results of the TADEP (Tampere Depression) II Study. *Nordic Journal of Psychiatry* 61:62-70, 2007
26. Hirschfeld RM, Holzer C, Calabrese JR, et al: Validity of the mood disorder questionnaire: a general population study. *American Journal of Psychiatry* 160:178-180, 2003
27. Hirschfeld RM, Williams JB, Spitzer RL, et al: Development and validation of a screening instrument for bipolar spectrum disorder: the Mood Disorder Questionnaire.

- American Journal of Psychiatry 157:1873–1875, 2000
28. Spitzer RL, Kroenke K, Williams JB: Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study: Primary Care Evaluation of Mental Disorders, Patient Health Questionnaire. *JAMA* 282:1737–1744, 1999
 29. Spitzer RL, Williams JB, Kroenke K, et al: Utility of a new procedure for diagnosing mental disorders in primary care: the PRIME-MD 1000 study. *JAMA* 272:1749–1756, 1994
 30. Zimmerman M, Mattia JI: A self-report scale to help make psychiatric diagnoses: the Psychiatric Diagnostic Screening Questionnaire. *Archives of General Psychiatry* 58:787–794, 2001
 31. Das AK, Olfson M, Gameroff MJ, et al: Screening for bipolar disorder in a primary care practice. *JAMA* 293:956–963, 2005
 32. Olfson M, Das AK, Gameroff MJ, et al: Bipolar depression in a low-income primary care clinic. *American Journal of Psychiatry* 162:2146–2151, 2005
 33. Leon AC, Shear MK, Portera L, et al: Assessing impairment in patients with panic disorder: the Sheehan Disability Scale. *Social Psychiatry and Psychiatric Epidemiology* 27:78–82, 1992
 34. Ware J Jr, Kosinski M, Keller SD: A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Medical Care* 34:220–233, 1996
 35. Weissman MM, Olfson M, Gameroff MJ, et al: A comparison of three scales for assessing social functioning in primary care. *American Journal of Psychiatry* 158:460–466, 2001
 36. Hirschfeld RM, Cass AR, Holt DC, et al: Screening for bipolar disorder in patients treated for depression in a family medicine clinic. *Journal of the American Board of Family Practice* 18:233–239, 2005
 37. First MB, Spitzer RL, Gibbon M, et al: Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P). New York, New York State Psychiatric Institute, Biometrics Research, Nov 2002
 38. Leckman JF, Sholomskas D, Thompson WD, et al: Best estimate of lifetime psychiatric diagnosis: a methodological study. *Archives of General Psychiatry* 39:879–883, 1982
 39. Johnson T, Fendrich M: Modeling sources of self-report bias in a survey of drug use epidemiology. *Annals of Epidemiology* 15:381–389, 2005
 40. Deyo RA, Cherkin DC, Ciol MA: Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *Journal of Clinical Epidemiology* 45:613–619, 1992
 41. Little RJA, Rubin DB: Statistical Analysis With Missing Data, 2nd ed. Hoboken, NJ, Wiley, 2002
 42. Bodlund O, Andersson SO, Mallon L: Effects of consulting psychiatrist in primary care: 1-year follow-up of diagnosing and treating anxiety and depression. *Scandinavian Journal of Primary Health Care* 17:153–157, 1999
 43. Ormel J, Vonkorff M, Oldehinkel AJ, et al: Onset of disability in depressed and non-depressed primary care patients. *Psychological Medicine* 29:847–853, 1999
 44. Pini S, Perkonig A, Tansella M, et al: Prevalence and 12-month outcome of threshold and subthreshold mental disorders in primary care. *Journal of Affective Disorders* 56:37–48, 1999
 45. Wagner HR, Burns BJ, Broadhead WE, et al: Minor depression in family practice: functional morbidity, co-morbidity, service utilization and outcomes. *Psychological Medicine* 30:1377–1390, 2000
 46. Gresenz CR, Sturm R, Tang L: Income and mental health: unraveling community and individual level relationships. *Journal of Mental Health Policy and Economics* 4:197–203, 2001
 47. Olfson M, Klerman GL: Depressive symptoms and mental health service utilization in a community sample. *Social Psychiatry and Psychiatric Epidemiology* 27:161–167, 1992
 48. Gilbody S, Bower P, Fletcher J, et al: Collaborative care for depression: a cumulative meta-analysis and review of longer-term outcomes. *Archives of Internal Medicine* 166:2314–2321, 2006
 49. Gilbody S, Sheldon T, Wessely S: Should we screen for depression? *BMJ* 332:1027–1030, 2006