Screening for Depression Using the PHQ-2: Changes over Time in Conjunction with Mental Health Treatment

John A. Fleishman and Samuel H. Zuvekas and Harold A. Pincus

Agency for Healthcare Research and Quality Working Paper No. 14002

October 2014

Suggested citation: Fleishman JA, Zuvekas SH, Pincus HA. Screening for Depression Using the PHQ-2: Changes over Time in Conjunction with Mental Health Treatment. 14002 October 2014, http://gold.ahrq.gov.

Screening for Depression Using the PHQ-2: Changes over Time in Conjunction with Mental Health Treatment

John A. Fleishman, Samuel H. Zuvekas, and Harold A. Pincus

Abstract

Objective: The two-item Patient Health Questionnaire (PHQ-2) meets the criteria for general screening of depression suggested by the U.S. Preventive Services Task Force. This study examined changes in the PHQ-2 over time, stratifying by receipt of mental health treatment, to help interpret screening results.

Methods: We used nationally representative samples of the U.S. civilian, noninstitutionalized population from the 2004-2006 panels of the Medical Expenditure Panel Study. Adult respondents (n=23,770) completed the PHQ-2 twice, 10 months apart. Purchase of antidepressant medication and receipt of ambulatory mental health care were assessed for two four-month periods preceding PHQ-2 administrations. We categorized respondents as above or below the PHQ-2 threshold for probable depression and as receiving or not receiving mental health treatment. We examined changes in combined depression-treatment status over time using multinomial logistic regression.

Results: At each time point, 83% scored below PHQ-2 depression threshold and had no mental health treatment; 8% scored below the threshold and reported some mental health treatment; 6% were above threshold but had no treatment, and 3% were above threshold and received some mental health treatment. Eighty-five percent remained in the same depression-treatment combination over time. In multivariate analysis, the strongest predictor of depression-treatment status at time 2 was status at time 1. Fifty-seven percent of those depressed and without treatment at Time 1 were not depressed at Time 2.

Conclusions: The PHQ-2 is useful as a depression screener, with prevalence rates comparable to diagnostic interviews. The high proportion of remission without treatment complicates interpretation.

John A. Fleishman
Center for Financing, Access, and Cost Trends
Agency for Healthcare Research and Quality
540 Gaither Road
Rockville, MD 20850
Phone: (301)427-1674
Email: John.fleishman@ahrq.hhs.gov

Screening for Depression Using the PHQ-2: Changes over Time in Conjunction with Mental Health Treatment

John A. Fleishman, Samuel H. Zuvekas, and Harold A. Pincus

Depression is a relatively common condition, associated with substantial impairment. Persons with 12-month major depressive disorder (MDD) report significantly higher functional and role impairment than those without this disorder.[1] Although definitions of "minor" or "subthreshold" depression vary considerably,[2] persons with subthreshold depression are more likely than asymptomatic individuals to report disability days and other indicators of poor social and work role functioning.[3-12]

The U.S. Preventive Services Task Force (USPSTF) recommends that primary care patients be screened for depression in clinical practices that have staff-assisted depression care support.[13] The USPSTF noted that asking two questions, pertaining to depressed mood and anhedonia, may be sufficient for screening purposes.[13,14] Intended as a depression screener for use in clinical practice, the Patient Health Questionnaire-2 (PHQ-2) comprises two questions that are consistent with the USPSTF screening recommendations.[15] The PHQ-2 had a sensitivity of 0.87 and a specificity of 0.78 for MDD; sensitivity and specificity for any depressive disorder were 0.79 and 0.86.[16]

The USPSTF noted the need for more research on depression screening that ascertains response and remission rates. In addition to examining concurrent false positive rates, longitudinal patterns of responses to a screening instrument can provide important information on the instrument's performance, such as the proportion of those who initially screen positive who are no longer positive at follow-up.

For both MDD and subthreshold depression, research documents both improvement over time and persistence.[17] Longitudinal studies of persons with MDD report persistence rates between 24-37% and recovery rates of 35-47% after one year.[18-21] Longitudinal studies of

persons with subthreshold depression report that from 37-70% were asymptomatic one year later.[3,8,18-22] However, many longitudinal investigations have been based on relatively small samples, typically less than 100 and often less than 50. Moreover, many studies did not stratify by whether treatment for depression was received, complicating interpretation of remission rates.

The current study examines responses to the PHQ-2 depression screener over time. Instead of using data from a small number of primary care practices, we use nationally representative data and a sample size larger than in many previous reports. In addition, analyses stratify respondents by whether mental health treatment was received, thereby enhancing interpretation of changes in screener status. This is a naturalistic observational study. By examining improvement and persistence in screener status over a one-year period, and associations with receipt of treatment, the study provides a context for interpreting brief depression screeners such as the PHQ-2.

Method

Sample

Data come from the Medical Expenditure Panel Survey (MEPS), a large, nationally-representative U.S. household survey conducted annually by the Agency for Healthcare Research and Quality. MEPS has been used extensively to track patterns of mental health treatment in the United States.[23-28] A new cohort (panel) is initiated each year and provides information for a 2-year reference period. MEPS conducts five in-person interviews, typically with one person per household, who reports for all household members. To obtain information that could be unreliable if reported by a proxy, including symptoms of depression, a self-administered questionnaire (SAQ) is administered each year to adults (age>17). Overall, response rates for the first SAQ among MEPS respondents were 91.9%, 91.3%, and 90.3% in 2004, 2005, and 2006, respectively. The SAQ sample was poststratified to the Census Bureau's Current Population

Survey and is representative of the U.S. civilian non-institutionalized population. The mean interval between SAQ administrations in the first and second panel years was 10.9 months (median=11).

Our analytic sample included respondents ages 18 and older in the 2004-2006 MEPS panels who completed the SAQ themselves in both years of eligibility. We excluded 7,807 respondents (25% of those eligible for the first-year SAQ) who either were lost to follow-up, did not complete a second SAQ within the 24-month observation period, or had one or both SAQs completed by someone else. The final analytic sample comprised 23,770 respondents.

Depression Status

The Adult SAQ contains the 2-item Patient Health Questionnaire (PHQ-2) , which asks how often the respondent has been bothered over the last two weeks by problems of "feeling down, depressed, or hopeless" and "little interest or pleasure in doing things." Responses range from "not at all" (0) to "nearly every day" (3). Evidence for construct and criterion validity has been presented[15,16] A score of 3 or higher ("above threshold") is suggested as a cut-point for depression screening.[15] We imputed 503 missing values of the PHQ-2 at Time 1 and 475 at Time 2, using imputation by chained equations.[29]

Mental Health Care Utilization

We constructed person-level indicators for whether the person reported (1) any mental-health-related ambulatory visit and (2) obtaining any antidepressant medication. Mental health visits were identified from among all ambulatory visits reported for each MEPS respondent based on the following criteria: 1) the main reason for the visit was for "psychotherapy or mental health counseling"; 2) the visit was to a psychiatrist, psychologist or social worker; 3) mental health, alcohol or drug treatment was received; or 4) one or more of the conditions associated with the visit was consistent with DSM-IV/ICD-9 codes 291, 292, 295-316 or ICD-9 V codes for

screening or treatment. Data on prescription drug use in the MEPS were obtained directly from households as well as from pharmacies used by MEPS respondents.[30]

We examined utilization in two time periods oriented to the times of the SAQ administrations. Time 1 includes the month in which the first SAQ was completed and the three prior months. Time 2 is the analogous four-month period for the second SAQ. Household-reported date information was used to determine the timing of mental health visits relative to the dates of SAQ completion. Month of visit was almost always reported. Binary variables indicated whether a mental-health visit occurred in each 4-month period. We also constructed, in parallel fashion, Time 1 and Time 2 measures of any visit to a primary care provider, whether or not it was mental-health related. We constructed parallel measures of any antidepressant medication in Times 1 and 2. We also combined indicators of ambulatory mental health visits and antidepressant receipt to derive indicators of receipt of any outpatient mental health treatment (visit, antidepressant, or both), during Time 1 and during Time 2.

Analyses

Combined Depression/Treatment Status. Because mental health service use was assessed prior to measuring depressive symptoms, examining the effect of symptoms on service use is inappropriate. Consequently, at each time point, analyses combined the PHQ-2 depression threshold indicator with the indicator of any prior mental health treatment, forming a four-category "depression/treatment status" variable (i.e., below threshold and no treatment [noD/noTx], below threshold and any treatment [noD/Tx]; above threshold and no treatment [D/noTx], and above threshold/any treatment [D/Tx]). Additional analyses examined visits and medications separately.

Main analyses examined the association of depression/treatment status at Time 1 with depression/treatment status at Time 2. To adjust for the possible confounding effects of other covariates, Time 2 depression/treatment status was regressed on Time 1 status and several

sociodemographic variables using a multinomial logit model. To ascertain whether results were influenced by those cases near the PHQ-2 threshold, supplementary analyses examined the full 0-6 range of PHQ-2 scores. To account for differences between the analytic sample and those excluded from analyses, we use a propensity-score adjustment procedure to reweight our analytic sample.[31,32] All analyses incorporated these adjusted analytic weights and accounted for the effect of the complex MEPS sample design.

Results

At time 1, 83% scored below PHQ-2 depression threshold and had no mental health treatment (Table 1); 8% scored below the threshold but did report some mental health treatment; 6% were above threshold but had no treatment, and 3% were above threshold and received some mental health treatment. The proportion with any mental health treatment at Time 1 rose from 6% to 11%, 19%, 21%, 31%, 35%, and 47% for respondents with PHQ-2 scores of 0-6, respectively (results not shown). Thus, even among those with the highest PHQ-2 score (n=499), less than half received mental health care in a four-month period. Of those who received mental health treatment, 50% only purchased antidepressant medication, 25% had mental-health-related visits but took no antidepressant medication, and 25% had mental-health-related visits and also purchased antidepressant medication (results not shown).

Each of the sociodemographic variables was significantly associated contemporaneously with depression/treatment status (Table 1). Differences were consistent with prior epidemiological findings. (Appendix)

Change in Depression/Treatment Status

The distribution of depression/treatment status at Time 2 was similar to that at Time 1: 83% in noD/noTX, 8% in noD/Tx, 6% in D/noTx, and 3% in D/Tx groups (not shown). Among those who were below threshold and received no treatment at Time 1, 92% remained unchanged

at Time 2 (Table 2). Among the NoD/Tx group at Time 1, 58% remained unchanged; 30% were still below threshold but no longer in care. Among the D/noTx group at Time 1, the majority (57%) were still not in care but had dropped below threshold, while 33% remained unchanged. Finally, of those above threshold and in care at Time 1, 45%remained depressed despite treatment; another 30% remained in care but fell below threshold. Overall, the table shows relatively high proportions along the diagonal, indicating persistence in depression/treatment status over time; only 15% changed depression/treatment status. Excluding noD/noTX at Time 1, 47% remained unchanged.

Multivariate Analyses. Table 3 reports results of a multinomial logistic regression of the four-category depression/treatment status variable at Time 2. Entries are exponentiated coefficients, which can be interpreted as relative risk ratios (RRRs). Depression/treatment status at Time 1 was the strongest predictor of depression/treatment status at Time 2. Compared to the noD/noTx group, the noD/Tx group at Time 1 was significantly more likely to have the same status (RRR = 42.3) or to be in the depressed/treatment category (RRR=35.2) than in the noD/noTx category at Time 2. Those in the D/noTx group at Time 1 were more likely (compared with the reference noD/noTx group) at Time 2 to be in D/noTx (RRR=7.5) or in D/Tx (RRR=8.2). These results reflect substantial persistence of symptom status over time, even after adjusting for other factors. Finally, those in the D/TX group at Time 1 were more likely to be in the noD/Tx (RRR= 46.8), D/noTx (RRR=13.0), or D/Tx categories (RRR=246.5) at Time 2, compared to the noD/noTx group at Time 1. The multivariate analyses also revealed significant effects of gender, race/ethnicity, age, education, insurance, and poverty status, controlling for Time 1 depression/treatment status.

It is possible that some respondents could move from above threshold at Time 1 to below threshold at Time 2 (or vice versa) due to a one-point change in their PHQ-2 score. For such people, a change in depression classification could merely represent measurement unreliability.

However, only 205 respondents went from a PHQ-2 score of 3 at Time 1 to 2 at Time 2, and only 172 moved from 2 to 3. This small number of respondents is unlikely to have had a major impact on the results.

We examined the mean 0-6 PHQ-2 score for each combination of depression/treatment status at Times 1 and 2 (Table 4). Mean scores were virtually unchanged for those who remained in the same status in both time periods. Those who moved from D/noTx to D/Tx showed a small change in mean PHQ-2. Those who shifted from D/noTx to noD/noTx showed a reduction in mean PHQ-2 scores from 3.8 to 1.0. Those who switched from D/Tx to below threshold at Time 2 showed drops in PHQ-2 from over 4 to just above 1. Shifts across the PHQ-2 threshold represent substantial changes in symptom intensity.

As ancillary analyses (Table 5), we varied the definition of mental health treatment, examining any antidepressant use (ignoring visits), any mental health visit (ignoring medications), any visit to a mental health professional, and any visit to a non-specialist physician (without seeing a specialist). The general pattern of change was similar to that for any treatment. Among those above threshold and receiving care at Time 1, depression persistence was lowest when treatment was defined as antidepressants (57% -- bolded figures); when treatment at Time 1 comprised visits to a non-specialist the highest percentage (18% - italicized number) were above threshold and not in care at Time 2, compared with other treatments.

Discussion

Findings using the PHQ-2 as a screener are broadly consistent with prior results using other screening tools. At Times 1 and 2, 9% of respondents scored above the PHQ-2 threshold for possible depression. These proportions are greater than the estimated 12-month prevalence of MDD in the National Comorbidity Study Replication (6.6%).[1] Using a PHQ-8 scoring algorithm that reflects DSM-IV criteria, Kroenke et al. reported current prevalences of 4.27% for major

depression and 4.82% for "other" depression in a U.S. national sample.[33] A large fraction of those with PHQ-2 scores above 2 may have subthreshold depression.

The original PHQ-2 report, which was based on a sample of primary care patients, found 15.2% scoring 3 or higher.[15] Current PHQ-2 prevalence is, however, similar to proportions scoring above threshold in community-based samples using the related PHQ-8 and PHQ-9 scales, of which the PHQ-2 is a subset (e.g., 8.7% in a U.S. sample using the PHQ-8). [34,35] In contrast, 17.8% scored above the PHQ-9 threshold in a primary care sample.[36] The prevalence of depressive symptomatology is probably higher in a sample using health care services than in the general population.[37]

Examining PHQ-2 scores over time places the concern with false positives into a context of symptom persistence and remission. The rate of spontaneous remission among those without treatment is high. Among patients above threshold and not receiving treatment at Time 1, 57% (unadjusted) were subsequently below threshold and not receiving treatment. (Of this group, 96% reported no mental health treatment at all in the roughly one year between Times 1 and 2.) The likelihood of remission without treatment depended on the severity of symptoms at Time 1 but was not confined to those at the screener cutoff. Overall, for PHQ-2 scores from 1 through 6, the proportions in the noD/noTx category at Time 2 were 0.89, 0.80, 0.67, 0.60, 0.46, and 0.43, respectively, consistent with prior studies showing that remission is more likely for those with subthreshold depression than for those with MDD.[8,18] Among the D/noTx group at Time 1, the proportions in the noD/noTX group at Time 2 were 0.70, 0.58, 0.45, and 0.40 for those with PHQ-2 scores of 3-6, respectively (results not shown). The likelihood of remission without treatment was substantial for those with PHQ-2 scores of 5 or 6.

Longitudinal studies that examine remission of depression have used different measures of depression, samples (patients versus community), and time frames.[3,8,18,38-40] Estimates of the proportion of those with MDD at baseline who are asymptomatic at follow-up range from 35% to

62%; estimates of the proportion of remitters among those with subthreshold depression are higher. Most studies, however, do not stratify by whether treatment was received or examine only those who received treatment. This study's estimate of remission without treatment in a large community sample thus provides an important baseline for future research.

The best predictor of being in treatment at Time 2 was being in treatment at Time 1. Among those in the D/Tx group at Time 1, 14% were still depressed at Time 2 but not receiving treatment. These respondents could be considered to have dropped out of care prematurely. The dropout rate was higher for mental health visits (24%) than for any antidepressant use (14%), but it was lower for patients with more severe symptoms. These results are similar to the estimated overall dropout rate of 22.4% in the NCS-R, which used a somewhat more restrictive definition of dropout .[41,42] Patients who dropped out showed little improvement, as measured by changes in PHQ-2 scores. It is not clear if dropout was motivated by lack of improvement, or if symptom resolution was impaired by insufficient clinical intervention due to dropout.

It is important to note that a large proportion of individuals above threshold and receiving treatment at Time 1 were above threshold at Time 2 (45%). Persistence despite treatment might occur because these individuals had more severe or treatment-resistant forms of the disease, but it also may reflect failure to apply evidence-based treatment (e.g. psychotherapy, medication management) with sufficient fidelity, or lack of effectiveness of available treatments for a large number of individuals.

At both time points, 8% of respondents fell below threshold but received some form of mental health treatment. Druss et al. found that most users of mental health services who did not have a 12-month diagnosis either had a lifetime diagnosis, a subthreshold diagnosis, or a recent major stressful event.[43] Those in treatment but below threshold may be receiving maintenance therapy.

There are several limitations to this analysis. The observation period spans only two years; the pattern of results might change over a longer time period. In addition, the PHQ-2 was administered only twice during the observation period. The extent to which symptoms fluctuated in the interval between measurements is unknown. Respondents who appeared to drop out of treatment despite persistent depression may actually have resolved the original episode and then experienced the onset of a new episode. Third, issues of minimizing respondent burden precluded administering other measures of depression, to provide further calibration of the PHQ-2. Nevertheless, the large sample size and the nationally representative nature of the MEPS sample are substantial strengths.

Conclusions

The results reinforce the utility and validity of the PHQ-2. However, the results also point to broader issues in screening for depression, as recommended by the USPSTF. On the one hand, given the association of subthreshold depression with reduced quality of life,[5] one can argue that persons with subthreshold depression need to be considered for therapeutic intervention. On the other hand, substantial minorities of people at all PHQ-2 score levels improved without treatment, with improvement rates higher among those below threshold. MDD and subthreshold depression exhibit both persistence and remission, which complicates clinical decision-making. A strategy of measurement-based care (including watchful waiting for subthreshold depression), incorporating observation of trends over time in repeated systematic assessments, could prove useful in guiding whether, and when, to initiate care.[44]

References

1. Kessler RC, Berglund P, Demler O, et al. The epidemiology of major depressive disorder: Results from the national Comorbidity Survey Replication (NCS-R). Journal of the American Medical Association 289: 3095-3105, 2003.

2. Pincus HA, Davis WW, McQueen LE. "Subthreshold" mental disorders: A review and synthesis of studies on minor depression and other "brand names". British Journal of Psychiatry 174: 288-296, 1999.

3. Broadhead WE, Blazer DG, George LK, et al. Depression, disability days, and days lost from work in a prospective epidemiologic survey. Journal of the American Medical Association 264: 2524-2528, 1990.

4. Johnson J, Weissman MW, Klerman GL. Service utilization and social morbidity associated with depressive symptoms in the community. Journal of the American Medical Association 267: 1478-1483, 1992.

5. Wells KB, Stewart A, Hays RD, et al. The functioning and well-being of depressed patients. Journal of the American Medical Association 262: 914-919, 1989.

6. Olfson M, Broadhead WE, Weissman MW, et al. Subthreshold psychiatric symptoms in primary care group practice. Archives of General Psychiatry 53: 880-886, 1996.

7. Jaffe A, Froom J, Galambos N. Minor depression and functional impairment. Archives of Family Medicine 3: 1081-1086, 1994.

8. 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.

9. Backenstrass M, Frank A, Joest K, et al. A comparative study of nonspecific depressive symptoms and minor depression regarding functional impairment and associated characteristics in primary care. Comprehensive Psychiatry 47: 35-41, 2006.

10. Judd LJ, Paulis MP, Wells KB, et al. Socioeconomic burden of subsyndromal depressive symptoms and major depression in a sample of the general population. American Journal of Psychiatry 153: 1411-1417, 1996.

11. Goldney RD, Fisher LJ, Dal Grande E, et al. Subsyndromal depression: prevalence, use of health services and quality of life in an Australian population. Social Psychiatry and Psychiatric Epidemiology 39: 293-298, 2004.

12. Cuijpers P, de Graaf R, van Dorsselaer S. Minor depression: Risk profiles, functional disability, health care use and risk of developing major depression. Journal of Affective Disorders 79: 71-79, 2004.

13. U.S. Preventive Services Task Force. Screening for depression in adults: U.S. Preventive Services Task Force recommendation statement. Annals of Internal Medicine 151: 784-792, 2009.

14. Whooley MA, Avins AL, Miranda J, et al. Case-finding instruments for depression: Two questions are as good as many. Journal of General Internal Medicine 12: 439-445, 1997.

15. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: Validity of a two-item depression screener. Medical Care 41: 1284-1292, 2003.

16. Löwe B, Kroenke K, Gräfe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). Journal of Psychosomatic Research 58: 163-171, 2005.

17. Judd LL, Akiskal HS, Maser JD, et al. A prospective 12-year study of subsyndromal and syndromal depressive symptoms in unipolar major depressive disorder. Archives of General Psychiatry 55: 694-700, 1998.

18. Ormel J, Oldehinkel MA, Brilman E, vanden Brink W. Outcome of depression and anxiety in primary care. Archives of General Psychiatry 50: 759-766, 2003.

19. Pini S, Perkonnig A, Tansells M, Wittchen H-U. Prevalence and 12-month outcome of threshold and subthreshold mental disorders in primary care. Journal of Affective Disorders 56: 37-48, 1999.

20. Katon W, Lin E, Von Korff M, et al. The predictors of persistence of depression in primary care. Journal of Affective Disorders 31: 81-90, 1994.

21. Forsell Y. A three-year follow-up of major depression, dysthymia, minor depression and subsyndromal depression: Results from a population-based study. Depression and Anxiety 24: 62-65, 2007.

22. Wells KB, Burnam MA, Rogers W, et al. The course of depression in adult outpatients: Results from the Medical Outcomes Study. Archives of General Psychiatry 49: 788-794, 1992.

23. Chen J, Rizzo J: Racial and ethnic disparities in use of psychotherapy: evidence from U.S. national survey data. Psychiatric Services 61:364-372, 2010.

24. Olfson M, Marcus SC. National patterns in antidepressant medication treatment. Archives of General Psychiatry 66:848-856, 2009.

25. Cook BL, McGuire T, Miranda J. Measuring trends in mental health care disparities, 2000 - 2004. Psychiatric Services 58:1533-1540, 2007.

26. Olfson M, Marcus SC, Tedeschi M, et al. Continuity of antidepressant treatment for adults with depression in the United States. Archives of General Psychiatry 60:1236-1242, 2003.

27. Olfson M, Marcus SC, Druss B, et al. National trends in the use of outpatient psychotherapy. American Journal of Psychiatry 159:1914-1920, 2002.

28. Olfson M, Marcus SC, Druss B, et al. National trends in the outpatient treatment of depression. Journal of the American Medical Association 287:203-209, 2002.

29. Royston P. Multiple imputation of missing values: Update of ice. Stata Journal 5: 527-536, 2005.

30. Moeller, J. F., Stagnitti, M. N., Horan, E., Ward, Kieffer, N., Hock, E. Methodology Report #12: Outpatient Prescription Drugs: Data Collection and Editing in the 1996 MEPS (HC-010A). Agency for Healthcare Research and Quality, Rockville, MD, June 2001. http://www.meps.ahrq.gov/data_files/publications/mr12/mr12.shtml

31. Little RJA, and Rubin DB. Statistical Analysis with Missing Data, 2nd Edition. New York: Wiley, 2002.

32. Weissman MW, Neria Y, Gameroff MJ, et al. Positive screens for psychiatric disorders in primary care: A long-term follow-up of patients who were not in treatment. Psychiatric Services 61: 151-159, 2010.

33. Kroenke K, Strine TW, Spitzer RL, et al. The PHQ-8 as a measure of current depression in the general population. Journal of Affective Disorders 114: 163-173, 2009.

34. Strine TW, Mokdad AH, Balluz LS, et al. Depression and anxiety in the United States: Findings from the 2006 Behavioral Risk Factor Surveillance System. Psychiatric Services 59: 1383-1390, 2008.

35. Rief W, Nanke A, Klaiberg A, et al. Base rates for panic and depression according to the Brief Patient Health Questionnaire: A population-based study. Journal of Affective Disorders 82: 271-276, 2004.

36. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure. Journal of General Internal Medicine 16: 608-613, 2001.

37. Frank E, Rucci P, Katon W, et al. Correlates of remission in primary care patients treated for minor depression. General Hospital Psychiatry 24: 12-19, 2002.

38. Simon GE and VonKorff M. Recognition, management, and outcomes of depression in primary care. Archives of Family Medicine 4: 99-105, 1995.

39. Sargent JK, Bruce ML, Florio LP, et al. Factors associated with 1-year outcome of major depression in the community. Archives of General Psychiatry 47: 519-526, 1990.

40. Patten SB and Schopflocher D. Longitudinal epidemiology of major depression as assessed by the Brief Patient Health Questionnaire (PHQ-9). Comprehensive Psychiatry 50: 26-33, 2009.

41. Olfson M, Mojtabai R, Sampson NA, et al. Dropout from outpatient mental health care in the United States. Psychiatric Services 60: 898-907, 2009.

42. Pinto-Meza A, Fernandez A, Bruffaerts R, et al. Dropping out of mental health treatment among patients with depression and anxiety by the type of provider: Results of the European Study of the Epidemiology of Mental Disorders. Social Psychiatry and Psychiatric Epidemiology 46: 273-280, 2011.

43. Druss BG, Wang PS, Sampson NA, et al. Understanding mental health treatment in persons without mental diagnoses. Archives of General Psychiatry 64: 1196-1203, 2007.

44. Harding K, Rush AJ, Arbuckle M, et al. Measurement-based care in psychiatric practice: A policy framework for implementation. Journal of Clinical Psychiatry, 2011.

Table 1. Sample Characteristics and Association with Time 1 Depression/Treatment Status

	Overall Proportion	Time 1 Depression/Treatment Status			
		None/None	None/tx	Dep/ no tx	Dep/tx
Time 1 depression/tx					
None/none	0.83				
None/ treatment	0.08				
Dep/ no tx	0.06				
Dep/ tx	0.03				
Gender*					
Male	0.48	0.87	0.05	0.06	0.02
Female	0.52	0.79	0.10	0.07	0.03
Race/Ethnicity*					
White	0.70	0.82	0.09	0.05	0.03
Black	0.11	0.84	0.04	0.10	0.03
Hispanic	0.13	0.86	0.03	0.08	0.02
Other	0.07	0.86	0.03	0.07	0.03
Age*					
18-40	0.41	0.86	0.06	0.06	0.02
41-50	0.21	0.80	0.09	0.06	0.04
51-60	0.17	0.80	0.10	0.06	0.04
61-70	0.11	0.82	0.09	0.06	0.03
71-80	0.07	0.84	0.08	0.06	0.02
81+	0.03	0.82	0.07	0.08	0.03
Education*					
< HS	0.18	0.79	0.06	0.12	0.04
HS	0.32	0.82	0.07	0.07	0.04
Some College	0.50	0.85	0.09	0.04	0.02
Self-rated health (T1)*					
Excellent	0.19	0.94	0.04	0.02	0.01
Very good	0.36	0.89	0.07	0.03	0.01
Good	0.31	0.82	0.09	0.07	0.03
Fair	0.12	0.62	0.12	0.17	0.09
Poor	0.03	0.37	0.11	0.29	0.24
Insurance (T1)*					
Full-year private	0.54	0.85	0.09	0.04	0.02
Medicare	0.16	0.83	0.08	0.07	0.02
Full-Year public	0.06	0.58	0.13	0.15	0.15
Full-year uninsured	0.14	0.85	0.03	0.10	0.02
Part-year uninsured	0.10	0.81	0.06	0.10	0.04
Primary care MD visit*					
No	0.66	0.86	0.06	0.06	0.02
Yes	0.34	0.79	0.11	0.06	0.04
Poverty category*					
Poor	0.11	0.73	0.07	0.14	0.07
Near poor	0.04	0.78	0.07	0.11	0.05
Low	0.13	0.81	0.07	0.10	0.03
Middle	0.31	0.84	0.07	0.06	0.03
High	0.40	0.87	0.09	0.03	0.02

Table 1. (continued) Sample Characteristics and Association with Time 1 Depression/Treatment Status

	Overall Proportion	Time 1 Depression/Treatment Status			
		None/None	None/tx	Dep/ no tx	Dep/tx
Census Region†					
Northeast	0.18	0.84	0.08	0.05	0.03
Midwest	0.23	0.84	0.08	0.05	0.03
South	0.36	0.82	0.07	0.07	0.03
West	0.23	0.84	0.07	0.07	0.02

Note: * p-value for bivariate association was less than 0.0001.

† p-value for bivariate association was less than 0.

Table 2. Observed Proportions of Time 2 Depression/ Mental Health Treatment Statuses, by Time 1 Statuses

	Time 2 Depression/ Any MH Tx				
Time 1 Depression/ MH Tx	NoD/noTx	NoD/Tx	D/noTx	D/Tx	Row N (Unweighted)
NoD/noTx	0.924 (0.002)	0.033 (0.002)	0.038 (0.002)	0.006 (0.001)	19,415
NoD/ Tx	0.300 (0.013)	0.583 (0.013)	0.030 (0.006)	0.087 (0.007)	1,794
D/no Tx	0.573 (0.016)	0.039 (0.006)	0.331 (0.014)	0.056 (0.007)	1,781
D/ Tx	0.127 (0.014)	0.294 (0.019)	0.133 (0.016)	0.446 (0.023)	780

Note: "D" refers to PHQ-2 score of 3 or higher; "Tx" refers to any mental health treatment (either mental health visit or antidepressant purchase). Standard errors in parentheses.

Table 3. Multinomial Logit Regression of Depression/Treatment Status in Year 2

	No D/ Tx	D/ No Tx	D Tx
Time 1 depression/Tx			
None/none	----	----	----
None/Tx	42.33 (35.93, 49.87)	2.23 (1.49, 3.34)	35.16 (26.46, 46.72)
Dep/ no Tx	1.71 (1.25, 2.33)	7.49 (6.33, 8.85)	8.17 (5.77, 11.55)
Dep/ Tx	46.76 (34.29, 63.76)	12.97 (9.49, 19.78)	246.54 (170.41, 356.70)
Gender			
Male	----	----	----
Female	1.60 (1.38, 1.85)	1.13 (0.99, 1.28)	1.31 (1.06, 1.62)
Race/Ethnicity			
White	----	----	----
Black	0.43 (0.32, 0.58)	1.07 (0.89, 1.29)	0.53 (0.36, 0.77)
Hispanic	0.63 (0.50, 0.79)	1.00 (0.80, 1.26)	0.73 (0.51, 1.04)
Other	0.54 (0.35, 0.83)	1.82 (1.38, 2.41)	0.76 (0.47, 1.22)
Age			
18-40	----	----	----
41-50	1.55 (1.27, 1.89)	1.18 (0.95, 1.48)	1.37 (1.04, 1.80)
51-60	1.63 (1.35, 1.96)	1.09 (0.88, 1.36)	1.56 (1.18, 2.06)
61-70	1.55 (1.16, 2.06)	1.00 (0.71, 1.40)	0.86 (0.56, 1.33)
71-80	1.42 (0.91, 2.21)	1.06 (0.64, 1.75)	0.54 (0.26, 1.14)
81+	1.50 (0.83, 2.70)	1.96 (1.12, 3.43)	0.62 (0.24, 1.58)
Education			
< HS	----	----	----
HS	1.04 (0.82, 1.32)	0.84 (0.71, 1.00)	0.83 (0.64, 1.07)
Some College	1.44 (1.16, 1.79)	0.68 (0.55, 0.83)	0.88 (0.66, 1.17)
Self-rated health (T1)			
Excellent	----	----	----
Very good	1.12 (0.87, 1.44)	1.43 (1.08, 1.89)	2.21 (1.18, 4.12)
Good	1.54 (1.19, 2.00)	2.15 (1.59, 2.91)	3.93 (2.22, 6.94)
Fair	1.76 (1.32, 2.36)	3.71 (2.76, 5.00)	6.81 (3.87, 11.96)
Poor	2.51 (1.58, 4.01)	7.58 (4.97, 11.57)	14.09 (7.33, 27.09)
Insurance (T1)			
Full-year private	----	----	----
Medicare	0.81 (0.57, 1.15)	1.06 (0.69, 1.61)	1.42 (0.79, 2.56)
Full-Year public	1.52 (1.18, 1.95)	1.94 (1.53, 2.46)	2.30 (1.65, 3.22)
Uninsured	0.76 (0.59, 0.98)	1.44 (1.13, 1.85)	1.04 (0.72, 1.50)
Part-year Ununsured	0.84 (0.63, 1.11)	1.60 (1.30, 1.97)	0.89 (0.59, 1.34)
Primary care MD visit			
No	----	----	----
Yes	1.20 (1.05, 1.37)	1.14 (0.98, 1.32)	1.22 (0.97, 1.52)
Poverty Category			
Poor	----	----	----
Near poor	0.82 (0.55, 1.24)	0.98 (0.76, 1.26)	0.71 (0.42, 1.22)
Low	0.94 (0.70, 1.28)	0.88 (0.70, 1.12)	0.90 (0.63, 1.28)

Table 3 (continued). Multinomial Logit Regression of Depression/Treatment Status in Year 2

	No D/ Tx	D/ No Tx	D Tx
Poverty Category (continued)			
Middle	0.83 (0.64, 1.08)	0.82 (0.65, 1.03)	0.78 (0.54, 1.12)
High	0.78 (0.58, 1.04)	0.60 (0.44, 0.77)	0.47 (0.31, 0.73)
Region			
Northeast	0.93 (0.73, 1.18)	1.01 (0.79, 1.28)	1.16 (0.80, 1.68)
Midwest	0.97 (0.79, 1.19)	0.91 (0.74, 1.12)	0.76 (0.53, 1.08)
South	1.04 (0.86, 1.25)	1.07 (0.89, 1.27)	1.26 (0.92, 1.71)
West	----	----	----
Interval between SAQs	0.99 (0.96, 1.03)	0.98 (0.95, 1.01)	1.01 (0.95, 1.07)

Note: “D” refers to PHQ-2 score of 3 or higher; “Tx” refers to any mental health treatment (either mental health visit or antidepressant purchase).

Table 4. PHQ-2 Score by Depression/Treatment Statuses at Each Time Period

	Time 2 Depression/Treatment Status			
Time 1 Depression/ Treatment Status	NoD/noTx	NoD/Tx	D/noTx	D/Tx
NoD/noTx	0.330 (0.006) 0.304 (0.006)	0.692 (0.042) 0.670 (0.038)	0.969 (0.040) 3.876 (0.042)	0.834 (0.088) 4.068 (0.114)
NoD/ Tx	0.668 (0.039) 0.621 (0.040)	0.726 (0.029) 0.723 (0.029)	1.377 (0.103) 4.053 (0.174)	1.320 (0.085) 4.231 (0.102)
D/ noTx	3.848 (0.042) 0.955 (0.038)	3.954 (0.174) 1.146 (0.140)	4.258 (0.049) 4.320 (0.054)	4.535 (0.162) 4.359 (0.142)
D/ Tx	4.148 (0.136) 1.253 (0.114)	4.216 (0.077) 1.321 (0.071)	4.636 (0.154) 4.568 (0.100)	4.736 (0.066) 4.758 (0.071)

Entries in each cell are mean PHQ-2 score at Time 1 and mean PHQ-2 score at Time 2.

Note: "D" refers to PHQ-2 score of 3 or higher; "Tx" refers to any mental health treatment (either mental health visit or antidepressant purchase). Standard errors in parentheses.

Table 5. Observed Proportions of Time 2 Depression/ Mental Health Treatment Statuses, by Time 1 Statuses with Different Definitions of Treatment

	Time 2 Depression/ Any MH Tx				
Time 1 Depression/ MH Tx	NoD/noTx	NoD/Tx	D/noTx	D/Tx	Row N (Unweighted)
NoD/no antidepressant	0.916 (0.002)	0.039 (0.002)	0.038 (0.002)	0.007 (0.001)	19,811
NoD/ antidepressant	0.229 (0.013)	0.651 (0.014)	0.026 (0.006)	0.094 (0.008)	1,398
D/ no antidepressant	0.540 (0.015)	0.052 (0.015)	0.322 (0.012)	0.085 (0.008)	1,957
D/antidepressant	0.100 (0.013)	0.328 (0.023)	0.**104** (0.017)	0.**469** (0.026)	604
NoD/no visit	0.891 (0.003)	0.062 (0.002)	0.038 (0.002)	0.009 (0.001)	20,396
NoD/any visit	0.366 (0.022)	0.506 (0.023)	0.030 (0.008)	0.098 (0.011)	813
D/ no visit	0.506 (0.015)	0.086 (0.007)	0.301 (0.012)	0.107 (0.010)	2,077
D/ any visit	0.121 (0.017)	0.262 (0.024)	0.133 (0.019)	0.484 (0.027)	484
NoD/no specialist visit	0.883 (0.003)	0.069 (0.002)	0.037 (0.001)	0.011 (0.001)	20,804
NoD/specialist visit	0.289 (0.028)	0.578 (0.029)	0.029 (0.010)	0.103 (0.016)	405
D/ no specialist visit	0.476 (0.014)	0.100 (0.007)	0.291 (0.011)	0.133 (0.010)	2,274
D/ specialist visit	0.107 (0.022)	0.260 (0.031)	0.108 (0.025)	0.525 (0.034)	287
NoD/no other MD visit	0.879 (0.003)	0.073 (0.002)	0.038 (0.001)	0.011 (0.001)	20,844
NoD/other MD visit	0.422 (0.035)	0.453 (0.035)	0.023 (0.009)	0.102 (0.016)	365
D/ no other MD visit	0.460 (0.014)	0.108 (0.007)	0.277 (0.011)	0.154 (0.011)	2,346
D/ other MD visit	0.140 (0.026)	0.234 (0.036)	0.*181* (0.034)	0.445 (0.044)	215

Note: Standard errors in parentheses.

Appendix Sociodemographic Variables and Associated Results

The MEPS interview collected data on sociodemographic characteristics of each household member. Education was categorized as less than high school, high school graduate, or at least some college. Age was categorized as 18-40, 41-49, 51-60, 61-70, 71-80, or 81 and older. Race/ethnicity was coded as white (non-Hispanic), Black (non-Hispanic), Hispanic, and other. Hispanic ethnicity took precedence in assigning respondents to categories. People reporting multiple races were coded as "other". Family income was coded as poor (family income below 100 percent of the federal poverty line [FPL]), near poor (100-125 percent of FPL), low income (125-200 percent of FPL), medium income (200-400 percent of FPL), and high income (above 400 percent of FPL). Binary indicators represented these categories, female gender, and Census region (North, South, Midwest, and West). Health insurance coverage was represented by mutually exclusive indicators for whether the person had private health insurance all year, Medicare (among those aged 65 or older), full year public coverage other than Medicare, no insurance during the year, or no insurance for part of the year (with the remaining time on either private or public coverage).

Table 1 shows the unadjusted associations of independent variables with depression/treatment status at Time 1. Each variable was significantly associated contemporaneously with depression/treatment status. Demographic differences were consistent with prior epidemiological findings: Proportions above threshold were higher among women (versus men), minorities (versus white non-Hispanic respondents), those with less than high-school education (versus college), in fair or poor self-rated health (versus excellent), with full-year public insurance (versus full-year private), and living in poverty.

The multivariate analysis (Table 3) also revealed significant effects of sociodemographic variables, controlling for Time 1 depression/treatment status. Women were more likely than men to be in treatment at Time 2. Black respondents were less likely than whites to receive treatment at Time 2. People with some college education were more likely than the least-educated group to be below threshold but treated at Time 2, but were less likely to be in the D/noTx category. Respondents aged 41-60 were more likely than those aged 18-39 to receive treatment at Time 2; respondents aged 61 or older were less likely than the youngest group to be above threshold and treated at Time 2. Compared with those in excellent self-rated health, respondents with lower self-rated health were more likely to be above threshold and/or receiving treatment at Time 2. Respondents with full-year public insurance at Time 1 were more likely to report depressive symptoms at Time 2, both treated and untreated, than those with full-year private coverage at Time 1. Respondents uninsured for a full year were more likely than the privately insured to be in the D/noTx category at Time 2, and less likely to be in the noD/Tx category. Finally, those in the high income group in the first year were less likely to be above threshold at Time 2, compared to those in poverty.