Psychiatric Practice Variations in the Diagnosis and Treatment of Major Depression

Teresa L. Kramer, Ph.D.
Allen S. Daniels, Ed.D.
Gayle L. Zieman, Ph.D.
Carla Williams, B.A.
Naakesh A. Dewan, M.D.

Objective: Practice variations in the diagnosis, treatment, and outcomes of patients with major depression were examined within six psychiatric practices participating in a national outcomes-management project. _Methods:_ Six of 20 psychiatric clinics met selection criteria for this study and provided a database of 5,106 patients. Patients completed the BASIS-32, the Short-Form-36 Health Survey, and a Beginning Services Survey. Treatment information was also obtained directly from the clinician or through a medical record review. _Results:_ Although 73.1 to 77 percent of patients screened positive for a depressive disorder, only 18.5 to 36.8 percent were diagnosed with major depression ($p<.001$). Between 39 and 72 percent of patients received psychotropic medications, a significant difference across sites ($p<.001$). In addition, the number of psychotherapy sessions was significantly different across sites ($p<.001$). _Conclusions:_ Patient care varies considerably across psychiatric practices, a finding that is particularly relevant for developers of performance indicators and risk-adjustment strategies for mental health. (_Psychiatric Services_ 51:336–340, 2000)

Major depression afflicts more than nine million people in the United States in any six-month period. The illness incurs heavy financial and societal losses from reduced productivity and prolonged disability, increased medical costs, emotional suffering, and even death (1–8). Despite these high prevalence rates and severe consequences, many providers lack the skills, knowledge, and time necessary to assess, detect, and treat affective disorders. In some studies, one-third to one-half of the primary care providers failed to recognize the signs of major depression (9–11). Other studies have shown that depression is often undertreated even when it is recognized (12–16). Recognition of this gap has sparked the development and distribution of guidelines for the assessment and treatment of depression by professional organizations (17–19). These guidelines include recommendations for medication type, therapeutic dosages, counseling referrals, and monitoring of treatment over time. They are designed to maximize the effectiveness of care, reduce variability in treatment approaches, and allow for continuous monitoring of appropriate care delivery.

In recent years, researchers have begun to study whether these guidelines are being used in clinical settings and, if so, how outcomes are affected by their implementation. In addition, particularly in carve-out and capitated systems, numerous managed care organizations have required that providers in their networks use established guidelines to develop and implement clinical pathways and algorithms for various disorders, including major depression.

The recent emphasis on outcomes assessment allows us to monitor whether providers are practicing guideline-concordant care for various disorders in real-world settings. Systems that measure the process and outcomes of care provide data about the variability across providers and clinical sites, including number of psychotherapy sessions, type and amount of medications prescribed, and follow-up care. These findings can lead to improvements in care through a continuous quality improvement plan that links measurement to problem identification and solution (20–22). By creating a database that documents practice patterns in various regions, we begin to

establish treatment parameters against which providers can be held accountable.

We analyzed the practice variations in the treatment of patients with major depression within six psychiatric practice settings participating in a national outcomes-management project. We sought to determine whether the sites varied in the rates of clinician diagnosis of major depression, prescribing patterns for antidepressant medications, number of psychotherapy sessions for depressed patients, and clinical outcomes at the three-month follow-up.

Methods
Participants
Six sites were selected from a total of 20 psychiatric clinics participating in the national outcomes project. Selection was based on the following criteria: participation in the project since its implementation in February 1995; enrollment of at least 500; variability in geographical location; multidisciplinary providers in a staff model; at least 50 percent of the practice supported by carve-out, capitated contracts; and establishment of a quality improvement program and utilization review process.

At the time of analysis, more than 10,000 patients were enrolled in the six sites. However, a large number of patients were excluded due to missing data, resulting in a database of 5,106 patients.

Instruments
At the initial visit to the clinics, patients completed three instruments:

The 36-item Short-Form Health Survey (SF-36) is a widely used instrument that assesses physical and psychological functioning on eight subscales: limitations in physical activities because of health problems; limitations in social activities because of physical or emotional problems; limitations in usual role activities because of physical health problems; bodily pain; general mental health (psychological distress and well-being); limitations in usual role activities because of emotional problems; vitality (energy and fatigue); and general health perceptions (23).

The SF-36 also contains a three-item depression screener, which assesses whether the patient is likely to have experienced a depressive disorder within the past year. Patients are asked, first, if they have had two weeks or more during which they felt sad, blue, or depressed or lost interest in things they usually cared about; second, if they have had two years or more of feeling depressed or sad most days; and, third, if they have felt depressed or sad much of the time in the past year. Patients are scored positive if they respond affirmatively to the first part or to the second two parts. Previous studies have shown that of those who screen positive, 83 to 94 percent of general medical patients and 89 to 93 percent of mental health care patients meet criteria for major depression (24).

The BASIS-32 assesses psychological symptoms on a scale ranging from 0 to 4 in five domains: interpersonal relationships, daily living–role functioning, depression-anxiety, impulsive-addictive behavior, and psychosis (25,26). A total mean score is also obtained. The scales developed for the BASIS-32 have demonstrated high internal consistency, whereas test-retest reliability averaged .76 for the five subscales. In this report, only differences in the depression-anxiety scale and total mean scores are considered.

The Beginning Services Survey (BSS) was developed specifically for this project by the first three authors. The BSS includes questions on patient demographics, access to care, and treatment expectations.

At the three-month follow-up, the patient completed both the SF-36 and BASIS-32. Six months after treatment began, a treatment events checklist was completed by the clinician or a medical review was conducted to obtain data on diagnosis, number of psychotherapy sessions, and medications prescribed.

All surveys were encoded onto scannable cards that were automatically read and entered into a database through a software program created specifically for this project. The clinical provider was thereby able to generate individual patient reports for immediate use in their assessment or triage.

Procedure
A training session was held with representatives from all 20 sites in January 1994. Sites were asked to obtain data from all patients or a random sample of at least 20 percent of patients (for example, every fifth patient).

Patients were asked to complete the BASIS-32, SF-36, and BSS at the outset of treatment as part of their routine clinical care. Patients and parents of minors also signed a consent form, acknowledging participation in this clinical project and agreeing to follow-up. They were informed that they could refuse to participate at any time without their treatment being affected and that all other information generated from the project would be aggregated anonymously, although their primary clinician would have access to their individual results.

Follow-up assessments were conducted by mail to ensure that patients who dropped out of treatment were included in the evaluation process. At the training session, sites were encouraged to maintain at least a 30 percent follow-up rate, which is significantly lower than most research standards, but was considered feasible without adding undue burden to the clinic sites.

Using an encrypted format, individual sites downloaded all data for this project from May 1995 through August 1997 and sent it to the University of Cincinnati, which currently houses the national database and conducts quarterly analyses of the aggregate data.

Results
Baseline information was available on 5,201 patients; follow-up was available on 1,252 depressed patients. Overall, patients were likely to be female, Caucasian, married, and well educated, with an age between 25 and 54 years. A significant difference was found across sites in racial background (χ^2=629.1, df=10, p≤.001), age (χ^2=366.2, df=30, p≤.001), education (χ^2=157.9, df=20, p≤.001), and marital status (χ^2=69.9, df=15, p<.001). More specifically, at two sites from the Southwest and Pacific Coast, both with large Hispanic pop-

ulations, more patients identified their racial background as other than Caucasian or African American. At another site, 20 percent of the treatment population were between ages 14 and 18 years and therefore more likely to be single and less educated. There was no significant difference in gender across sites.

Diagnostic information
At each practice site, 73.1 to 77 percent of patients screened positive for a depressive disorder on the SF-36 three-item screener, which was not a significant difference between sites (22). However, only 18.5 to 36.8 percent of all patients were diagnosed as having major depression by the treating clinician. This finding represented a significant difference across sites (χ^2=68.01, df=5, p≤.001), with sites 1 and 4 reporting the lowest rates (18.5 and 21.9 percent, respectively) and sites 5 and 6 reporting the highest rates (34 and 36.8 percent, respectively).

An additional 16.3 to 38.4 percent of patients were diagnosed by clinicians as having adjustment disorder with depressed mood (χ^2=143.62, df=5, p≤.001), whereas 5.4 to 13.9 percent were diagnosed with dysthymia (χ^2=22.25, df=5, p≤.001).

When comparing patients with and without a clinical diagnosis of major depression, significant differences were found across sites on age (χ^2= 109.9, df=30, p<.001), race (χ^2= 180.6, df=10, p<.001), education (χ^2= 43.5, df=20, p<.001), and marital status (χ^2=35.9, df=15, p<.05). Compared with other locations, sites 1 and 4 had approximately twice as many depressed patients between the ages of 14 and 18 years, a difference that affected educational level and marital status. Furthermore, site 2 had a greater number of elderly (over 64) and married patients who were depressed. The sites with larger Hispanic populations also had a higher percentage of ethnic patients diagnosed with depression.

Multivariate analysis of variance (MANOVA) was used to examine differences on the BASIS-32. A significant difference was found across sites (F=4.8, df=30, 20,762, p<.001). Site effects for depression-anxiety (F=

Percentage of depressed patients in six psychiatric clinics who were receiving antidepressant medications, by type of depression
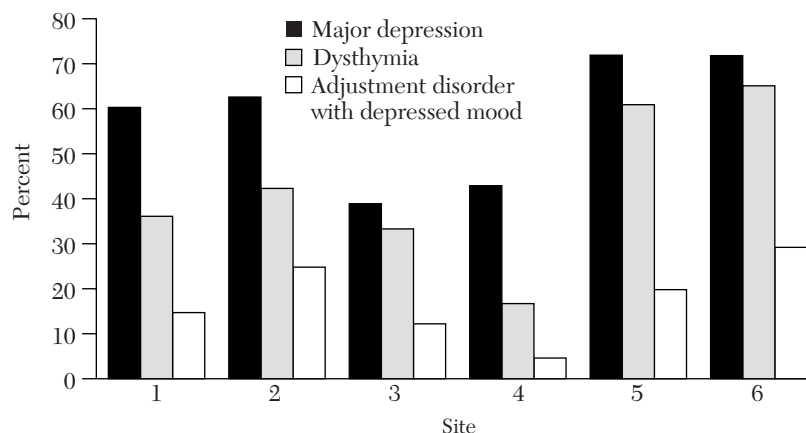
3.28, df=5, 5,195, p≤.001) and for the total measures of the BASIS 32 (F= 4.15, df=5,5,195, p<.001) were also significant. Subsequent comparisons of sites using Duncan's multiple range test showed that patients at site 1 had higher measures on the BASIS-32 than those at sites 3, 4, and 5. Patients at site 6 had higher total measures than those at site 5.

Differences on the SF-36 were also examined using a MANOVA. This analysis showed significant difference across sites (F=1.64, df=40, 22,595, p<.01). Site effects were found for the bodily pain scale (F=2.91, df=5, 5,190, p<.05) and the physical functioning scale (F=4.83, df=5, 5,190, p<.001). Results of Duncan's multiple range test on the bodily pain scale showed higher ratings, indicating more bodily pain, at site 1 than at sites 2 and 3, higher ratings at site 2 than sites 4 and 5, and higher ratings at site 3 than at site 5. A similar test on the physical functioning scale showed site 1 had higher ratings than sites 2 and 4, and site 2 had higher ratings than at sites 3, 4, 5, and 6.

Medication information
Figure 1 shows that of the patients diagnosed with major depression, 38.9 to 71.9 percent received psychotropic antidepressant medications. These findings represented a significant difference across sites (χ^2=81.2, df=5, p<.001). More specifically, sites 3 and 4 had fewer patients who received medication for their major depression

(38.9 and 42.9 percent, respectively).

Figure 1 also shows the percent of patients with an adjustment disorder or dysthymia receiving antidepressant medication. The range across sites was 4.6 to 29.2 percent for adjustment disorder with depressed mood (χ^2=32.5, df=5, p<.001) and 16.7 to 65.1 percent for dysthymia (χ^2=34.8, df=5, p<.001). As with major depression, sites 3 and 4 were less likely to prescribe medications.

Results of a MANOVA showed significant differences across sites on the BASIS-32 for depressed patients who received medication compared with those who did not (F=2.43, df= 30, 4,314, p<.001). Significant site differences were found for the depression-anxiety scale (F=4.67, df=5, 1,083, p< .001) and total mean scores (F=2.34, df=5,1,083, p<.05). Results of a Duncan's range test showed site 1 had higher ratings, indicating more severe symptoms, on all scales than did the other five sites, and site 2 had higher ratings on depression-anxiety than did sites 3, 4, 5, and 6. There were no significant differences across sites on the SF-36.

Psychotherapy attendance
The mean number of psychotherapy sessions for patients with major depression was significantly different across sites (F=11.58, df=5, 1,596, p< .001). The range was from 4.34 at site 3 to 9.17 at site 1. Site 1 had more sessions than sites 3, 5, and 6, and sites 3 and 6 had more sessions than site 4.

Clinical outcomes

Results of the MANOVA demonstrated significant differences across the sites on the BASIS-32 at the three-month follow-up (F=2.22, df=30, 4,902, p<.001). However, no significant differences were found on the depression-anxiety scale or total mean scores. Significant differences were found on the SF-36 (F=1.61, df=40, 5,390, p≤.01). Sites differed only on the vitality scale (F=2.52, df=5, 1,243, p<.05); a Duncan's range test showed site 2 had higher ratings, indicating higher vitality, than sites 3, 5, and 6.

Discussion

Although approximately 75 percent of the SF-36 screenings for a depressive disorder were positive, clinical diagnoses of major depression varied by site, from 19 to 37 percent. When combined with dysthymia, the rates ranged from 30 to 44.1 percent, which is similar to the rate of 32 percent found in the Medical Outcomes Study (27). More important than the variations in rates of detected depression were the findings pertinent to the appropriateness of treatment and subsequent outcomes. Of patients diagnosed as depressed by a clinician, 39 to 71.9 percent received psychotropic medications, depending on the site.

Two sites were particularly deviant in prescribing antidepressant medication: sites 3 and 4 had fewer than half of their depressed patients on medication, even though patients in these two groups reported levels of affective symptoms equivalent to those in other practices. It appears these two outliers were not providing care in a manner that was comparable to other psychiatric group practices or within the parameters established by the U.S. Agency for Health Care Policy and Research (17,18) and the American Psychiatric Association (19).

Although the mean number of psychotherapy sessions across all sites was low to moderate (4.34 to 9.17), there was no indication of "tradeoffs" between psychosocial and psychopharmacological interventions; that is, practices less likely to prescribe medications were not more likely to extend the length of psy-

chotherapy. There was also a trend for site 3 patients to report a higher level of symptoms on the BASIS-32 and lower functioning on the SF-36 at follow-up. Given this site's tendency to diagnose fewer depressed patients and prescribe fewer antidepressant medications, it would be important to monitor its progress over time to ascertain whether the variations in care are also associated with significantly poorer patient outcomes.

Practice variations in the diagnosis and treatment of depression may be attributed to differences in the clinical milieu, the patient population, or both. For example, if patients were sicker at one site, it would be expect-

This study showed that patient care may vary considerably across specialty care settings, even in psychiatric practices where one would anticipate more uniformity in diagnostic rates and treatment regimens.

ed that they would more likely meet criteria for major depression. Yet the data do not support this hypothesis. First, results of depression screenings, which were independent of clinicians' diagnoses, were fairly consistent across sites. For example, 74.8 percent of the patients screened positive at site 5, which was the site with the highest number of patients clinically diagnosed with depression (36.8 percent). By contrast, 75.8 percent of the patients at site 4 screened positive, but only 21.9 percent were diagnosed.

Second, scores on the BASIS-32

were fairly equivalent at baseline across sites with one exception: site 1 patients reported distress on the depression-anxiety scale, yet only 18.5 percent of them were diagnosed with depression by clinicians. Thus patient-reported symptoms appeared to be unrelated to the variations observed in clinician diagnosis.

Certainly, other patient characteristics inherent in a practice's treatment population, such as medical illness or comorbid substance abuse, may have increased a clinician's propensity to diagnose depression. These variables were beyond the scope of this study but warrant further investigation. Finally, significant differences were found across sites in pharmacological and psychosocial interventions, even when patients reported similar distress levels on the BASIS-32 and equivalent functioning on the SF-36.

Despite our best efforts to control for certain structural components of these practices, other unmeasured variables within the cultural milieu may have influenced clinical decisions. They include the clinician's discipline, philosophical approaches to diagnostic nosology, training in the treatment of depression, or other more subtle issues such as a tendency to judge one or two symptoms as meeting the threshold or giving more weight to a coexistent stressor. Geographic variations in treatment patterns have been found in general medicine (28,29) and, more specifically, in the treatment of depression in community settings (30) and inpatient settings (31), as well as among different types of interventions, such as electroconvulsive therapy (32).

This study showed that patient care may vary considerably across specialty care settings, even in psychiatric practices where one would anticipate more uniformity in diagnostic rates and treatment regimens. These findings have particular relevance for developers of performance indicators and risk-adjustment strategies for mental health. For example, the National Committee for Quality Assurance has recently proposed that behavioral health care organizations report the percentage of depressed patients receiving antidepressant medication as part of the Healthplan

Employers Data Information Set (HEDIS 3.0) (33). Such information is critical for determining the industry's current practices and monitoring improvements over time as a result of such policy decisions.

In addition, Wennberg (34,35) has suggested that such geographic variation in certain markets may reflect a misuse of care. Although this study was not able to determine provider competencies, the possibility that clinicians in certain practices were not as well informed could not be ruled out by the current findings. Comparative data on such factors as provider-patient ratios, availability of psychiatric consultation, guidelines on continuing medical education, and affiliation with a teaching hospital might elucidate the reasons for these differences; however, it was not feasible to publish such data without compromising the anonymity of the sites.

Because this investigation was not designed as a research study, the findings must be interpreted cautiously. Incomplete medical records, unvalidated clinical diagnoses, low follow-up rates, and nonrandomization of patients to the various treatment settings limit the generalizability of this work. These groups represent the "best of the best" in their commitment to outcomes monitoring and benchmarking, however. Future work should focus on enhancing this national outcomes project to improve medical record documentation, increase follow-up rates, validate diagnoses across sites, and increase the number of participating sites. In light of cost-containment policies, increasing oversight by accrediting agencies, and scrutiny by multiple stakeholders, it is imperative that we monitor our own performance and initiate action to improve the quality of care when irregularities are observed. ♦

References

1. Diagnostic and Statistical Manual of Mental Disorders, 4th ed. Washington, DC, American Psychiatric Association, 1994

2. Greenberg PE, Stiglin LE, Finkelstein SN, et al: The economic burden of depression in 1990. Journal of Clinical Psychiatry 54:405–418, 1993

3. Greenberg PE, Stiglin LE, Finkelstein SN, et al: Depression: a neglected major illness. Journal of Clinical Psychiatry 54:419–424, 1993

4. Kessler RC, Frank RG: The impact of psychiatric disorders on work loss days. Psychological Medicine 27:861–873, 1997

5. Kouzis AC, Eaton WW: Emotional disability days: prevalence and predictors. American Journal of Public Health 84:1304–1307, 1994

6. Mintz J, Mintz LI, Arruda MJ, et al: Treatments of depression and the functional capacity to work. Archives of General Psychiatry 49:761–768, 1992

7. Von Korff M, Ormel J, Katon W, et al: Disability and depression among high utilizers of health care: a longitudinal analysis. Archives of General Psychiatry 49:91–100, 1992

8. Wells KB, Stewart A, Hays RD, et al: The functioning and well-being of depressed patients: results from the Medical Outcomes Study. JAMA 262:914–919, 1989

9. Coyne JC, Schwenk TL, Fechner-Bates S: Nondetection of depression by primary care physicians reconsidered. General Hospital Psychiatry 17:3–12, 1995

10. Kessler RC, Amick R, Thompson J: Factors influencing the diagnosis of mental disorders among primary care patients. Medical Care 23:50–62, 1985

11. Rost K, Zhang M, Fortney J, et al: Persistently poor outcomes of undetected major depression in primary care. General Hospital Psychiatry 20:12–20, 1998.

12. Borus JF, Howes MJ, Devins NP, et al: Primary health care providers' recognition and diagnosis of mental disorders in their patients. General Hospital Psychiatry 10:317–321, 1998

13. Priest RH: Psychotherapy in general practice: expectations and experiences. British Journal of Medical Psychology 43:298–299, 1970

14. Wells KB, Katon W, Rogers B, et al: Use of minor tranquilizers and antidepressant medications by depressed outpatients: results from the Medical Outcomes Study. American Journal of Psychiatry 151:694–700, 1994

15. Tinsley JA, Shadid GE, Li H, et al: A survey of family physicians and psychiatrists: psychotropic prescribing practices and educational needs. General Hospital Psychiatry 20:360–367, 1998

16. Koenig HG, George LK, Meador KG: Use of antidepressants by nonpsychiatrists in the treatment of medically ill hospitalized depressed elderly patients. American Journal of Psychiatry 154:1369–1375, 1997

17. Depression Guideline Panel: Depression in Primary Care: Vol 2: Treatment of Major Depression. Clinical Practice Guideline 5. AHCPR pub 93-0551. Rockville, Md, Agency for Health Care Policy and Research, 1993

18. Depression in Primary Care: Vol 1: Detection and Diagnosis. Clinical Practice Guideline 5. AHCPR pub 93-0550. Rockville, Md, Agency for Health Care Policy and Research, 1993

19. American Psychiatric Association: Practice guideline for major depressive disorder in adults. American Journal of Psychiatry 150 (suppl 4):1–26, 1993

20. Rost K, Smith GR, Burnam MA, et al: Measuring the outcomes of care for mental health problems: the case of depressive disorders. Medical Care 20:MS266–MS273, 1992

21. Smith GR, Ross RL, Rost KM: Psychiatric outcomes module: depression (DOM), in Outcomes Assessment in Clinical Practice. Edited by Sederer LI, Dickey B. Baltimore, Williams & Wilkins, 1996

22. Smith GR, Rost KM, Fischer EP et al: Assessing the effectiveness of mental health care in routine clinical practice: characteristics, development, and uses of patient outcomes modules. Evaluation and the Health Professions 20: 65–80, 1997

23. Ware JE Jr, Sherbourne CD: The MOS 36-Item Short-Form Health Survey (SF-36): I. conceptual framework and item selection. Medical Care 30:473–483, 1992

24. Rost K, Burnam MA, Smith GR: Development of screeners for depressive disorders and substance disorder history. Medical Care 31:189–200, 1993

25. Eisen SV, Grob MC, Klein AA: BASIS: the development of a self-report measure for psychiatric inpatient evaluation. Psychiatric Hospital 17:165–171, 1986

26. Eisen SV, Dill DD, Grob MC: BASIS-32: A Self-Report Measure for Evaluating Patient Progress. Report No 77. Belmont, Mass, McLean Hospital Evaluative Service Unit, 1989

27. 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

28. Ashton CM, Peterson NJ, Souchek J, et al: Geographic variations in utilization rates in Veterans Affairs hospital and clinics. New England Journal of Medicine 340:32–39, 1999

29. Welch HG, Wennberg DE, Welch WP: The use of Medicare home health care services. New England Journal of Medicine 335:324–329, 1996

30. Keller MB, Klerman GL, Lavori PW, et al: Treatment received by depressed patients. JAMA 248:1848–1855, 1982

31. Fortney JC, Booth BM, Smith GR: Variation among VA hospitals in length of stay for treatment of depression. Psychiatric Services 47:608–613, 1996

32. Hermann RC, Dorwart RA, Hoover CW, et al: Variation in ECT use in the United States. American Journal of Psychiatry 152:869–875, 1995

33. Thompson JW, Bost J, Ahmed F, et al: The NCQA's quality compass: evaluating managed care in the United States. Health Affairs 17(1):152–158, 1998

34. Wennberg JE: Dealing with medical practice variations: a proposal for action. Health Affairs 3(2):6–32, 1984

35. Wennberg JE: Variations in medical practice and hospital costs. Connecticut Medicine 49:444–453, 1985