Implementing Standardized Assessments in Clinical Care: Now's the Time

Marcia Valenstein, M.D.
David A. Adler, M.D.
Jeffrey Berlant, M.D., Ph.D.
Lisa B. Dixon, M.D., M.P.H.
Rebecca A. Dulit, M.D.
Beth Goldman, M.D.
Ann Hackman, M.D.
David W. Oslin, M.D.
Samuel G. Siris, M.D.
William A. Sonis, M.D.

In this Open Forum the Committee on Psychopathology within the Group for the Advancement of Psychiatry (GAP) strongly encourages clinicians and health systems to implement standardized assessments of patients' outcomes for mental disorders, particularly disorders such as depression. The GAP committee describes how calls for the regular use of standardized scales in clinical settings naturally follow from the development and dissemination of treatment guidelines. It discusses the challenges involved in implementing routine outcome measures in clinical settings and explains why the advantages of measurement-based care make addressing these challenges worthwhile. Finally, the committee makes practical suggestions for clinicians and systems attempting to implement routine outcome measures in their clinics. (*Psychiatric Services* 60:1372–1375, 2009)

The authors are members of the Committee on Psychopathology of the Group for the Advancement of Psychiatry. Dr. Valenstein is affiliated with the Health Services Research and Development, Serious Mental Illness Treatment Research and Evaluation Center, Department of Veterans Affairs (VA), Box 130170, Ann Arbor, MI 48113-0170 (e-mail: marciav@med.umich.edu). Dr. Adler is with the Department of Psychiatry, Tufts University School of Medicine, Boston. Dr. Berlant is in private practice in Boise, Idaho, and Kentfield, California. Dr. Dixon is with the VA Capitol Health Care Network Mental Illness Research, Education, and Clinical Centers (MIRECC), Baltimore. Dr. Hackman is with the Department of Psychiatry, University of Maryland School of Medicine, Baltimore. Dr. Dulit is with the Weill Medical College of Cornell, New York City. Dr. Goldman is a medical consultant with Blue Cross Blue Shield of Michigan, Detroit. Dr. Oslin is with the Veterans Integrated Service Network 4, MIRECC, Philadelphia. Dr. Siris is with the Department of Psychiatry, Albert Einstein College of Medicine, New York City. Dr. Sonis is with the Department of Psychiatry, Drexel University College of Medicine, Philadelphia.

Getting good clinical outcomes is what the field of psychiatry is about. On this point, patients, providers, and payers all agree. Therefore, regular and systematic assessment of patient outcomes during treatment is an important component of good clinical work.

Payers and health systems took an early lead in efforts to use standardized scales to routinely measure patient outcomes, using aggregated data to compare the outcomes of clinicians' patient panels but also providing feedback to clinicians regarding services use and processes of care (for example, refill patterns) among individual patients so that care processes and patient outcomes might be improved. Many publicly funded substance abuse and mental health programs now require reporting of standardized measures of patient outcomes (1), and Blue Cross Blue Shield of Massachusetts has recently offered providers additional compensation if their patients complete symptom and functional scales during clinical sessions and the providers report these results (2).

Although many mental health clinicians may not share payers' interest in using aggregate data to compare the outcomes of clinicians' patient panels, most clinicians do share the goal of achieving good outcomes for their patients. Yet clinicians have often hesitated or resisted implementing routine monitoring of outcomes in clinical settings (1,2). Even providers who work in health systems where regular monitoring of outcomes is mandated have not always integrated these measures into their clinical decision making.

In this Open Forum the Committee on Psychopathology within the Group for the Advancement of Psychiatry (GAP) follows up on its previous editorial on outcomes monitoring published in 1994 (3). The GAP committee proposes that thoughtful supplementation of clinical interviews with standardized symptom and functional measures can improve patient care. However, the committee also cautions that standardized scales must always be used as supplements rather than in place of careful clinical interactions. This policy recommendation was made after ten GAP psychiatrist members attending two multiday meetings in March and Novem-

ber 2007 reached consensus on the advisability of routine outcomes monitoring. They drafted a statement, which was reviewed by the GAP publication board in 2008. Two revisions of the statement were subsequently made based on GAP board member input.

The call for outcomes measurement

Although there are likely many progenitors of the current call for measurement-based care in clinical settings (for example, payers' concerns about value received for health care dollars and the long-standing use of standardized scales in clinical research), calls for measurement-based care also follow naturally from the development and dissemination of mental health treatment guidelines.

An example of this progression is demonstrated in how the implementation of treatment guidelines for major depressive disorder led to the now frequent recommendations for measurement-based care for patients with major depression (4). The Agency for Health Care Policy and Research and the American Psychiatric Association published treatment guidelines for major depression in 1993 (5,6). Although these early guidelines for major depression did not specifically recommend the use of standardized scales, both recommended systematic reevaluation of patients with depression at approximately six and 12 weeks after treatment initiation and recommended that subsequent treatment steps be based on whether patients had had a "full, partial, or no" response at these junctures. Once treatment guidelines recommended systematic assessments and treatment decisions based on response categories (full, partial, and no response), the stage was set for further operationalization and quantification of treatment response.

The Texas Implementation of Medication Algorithms (TIMA) project, which was designed to improve the treatment of major depression in publicly funded settings, took the next logical step. TIMA not only specified critical decision points when patients' symptoms should be reevaluated, but it also specified the percentage of change in depressive symptoms that should prompt alternative treatment interventions (<25% response, 25%–50%, 50%–75%, or a >75% symptom response), essentially basing treatment recommendations on a quantitative version of the "full, partial, or no response" categories outlined in earlier guidelines (7). Determining patients' treatment response with this degree of precision clearly required the use of standardized scales in addition to clinical assessments. When TIMA 's concrete and precise measurement-based algorithms were implemented in usual care, patients with major depression showed improved outcomes, demonstrating the clinical value of precisely measuring symptoms and following an evidence-based sequence in decision making (7).

Editor's Note: This Open Forum is the first contribution to an occasional series in which the Group for the Advancement of Psychiatry (GAP) (www.ourgap.org) will present ideas to further the understanding of mental illness and improve access to care and quality of treatment for persons with mental disorders. Since its beginnings in the post–World War II era of providing modern psychiatric care, GAP has continued to be a think tank operating through its committee structure of national experts to present reports and position statements that are disseminated nationally and internationally.

Challenges in implementing outcomes measurement

There are many important challenges in implementing standardized measurements in clinical settings. Many clinicians may not be familiar with standardized scales, other than the Global Assessment of Functioning, which is used as the clinician's overall assessment of a patient's level of functioning for axis V of the *DSM-IV* multiaxial assessment. These clinicians will require assistance in choosing standardized measures for their patients in addition to training in administering the chosen measures, interpreting measure scores, and integrating measurement data into their clinical work. Some clinicians may also need assistance with computer technology that will allow them to track patients' scores over time.

Clinician and patient concerns regarding the relevance and utility of standardized scales may also need to be addressed. For example, standardized scales for depression may not capture an individual patient's characteristic symptoms of depression (for example, some patients experience angry rather than sad mood, but several depression scales do not include items for angry mood). Conversely, standardized scales may include items that are not germane to individual patients (for example, a loss of libido is a common item on depression scales, but individual patients may not experience this symptom). Patients also must complete the same scale at several time points to meaningfully measure change. Standardized scales may also be less germane in specialized practices, such as psychoanalytic practices. All such factors may diminish clinician and patient willingness to participate in regular symptom measurement.

Addressing the challenges and benefits

Despite these real challenges, the advantages of moving to measurement-based care in clinical settings outweigh the drawbacks. Supplementing the clinical interview with rating scales allows the clinician to follow the lead of the patient during the clinical interview while using standardized scales to ask the same questions in the same manner to more reliably assess how treatment has had an impact on the key symptoms of the patient's psychiatric condition.

Incorporating standardized measures into the clinical workflow may facilitate a more balanced clinical interaction where time is spent systematically assessing both the disorder affecting the patient and the patient who has the disorder. As Sir William Osler, a physician who has been

called the Father of Modern Medicine, observed, "The good physician treats the disease; the great physician treats the patient who has the disease." Traditionally, mental health clinicians have excelled at focusing on the whole person but have sometimes failed to routinely and systematically assess core symptoms of their patients' psychiatric illnesses or functional capacities. Medical and psychiatric comorbidities, such as concurrent substance use, may remain unrecognized even though they can reduce treatment effectiveness or require the addition of alternative services. Clinicians may also fail to routinely ask about suicidal ideation, even though it can emerge unexpectedly and requires immediate attention. If these critical clinical factors are routinely assessed with standardized scales and detected in a timely manner, patients may be safer and clinicians' medical-legal liability may be reduced.

The use of patient self-rating scales may also give patients a new role in the medical encounter as their reports of symptom levels are acknowledged and recorded. Patients who regularly complete standardized assessments may become more knowledgeable about their disorders and feel more comfortable assessing their own treatment progress, a key aspect of disease self-management. Assessments over time may also allow patients to gain a longer-term perspective of their symptom trajectory and recognize periods of better and poorer functioning. Indeed, the regular use of rating scales has been considered integral to some forms of psychotherapy, such as cognitive-behavioral therapy.

Fortunately, given a modest but growing body of literature suggesting that regular monitoring with feedback may improve patient outcomes, particularly if the monitoring is embedded within evidence-based algorithms, most of the challenges involved in implementing standardized scales in clinical practice can be successfully met.

For example, clinicians can use recently published books to familiarize themselves with a variety of standardized scales. The American Psychiatric Association's *Handbook of Psychiatric Measures* outlines the advantages and disadvantages of rating instruments for a variety of common mental health disorders and provides copies of well-validated instruments in the public domain (8). For clinicians who do not work in environments with information technology support, simple personal computer programs such as Excel or Access might be adapted for outcomes measures, or Web-based outcomes programs can be used (9). Clinician time limitations might be addressed by emphasizing patient self-report scales rather than clinician-rated scales. Many self-report scales have been validated, correlate with clinician-rated scales, show change in response to treatment, and can be completed in less than ten minutes (8).

Well-informed clinicians can also explain the rationale for scale completion to patients. Regularly sharing patients' progress with them and incorporating information from outcome measures into clinical decision making is likely to increase patients' willingness to regularly complete standardized measures. Advances in the field, such as computer-adapted testing and item banks for common mental health symptoms, are being actively pursued by the National Institutes of Health and may soon decrease patient burden by reducing the number of questions needed to place patients along a continuum of symptom severity (10).

Two scales that measure depression severity, the nine-item Patient Health Questionnaire (PHQ-9) and the Quick Inventory of Depressive Symptomatology (QIDS), may be relatively easy to use for clinicians and health systems testing measurement-based care. Both scales measure depression severity and the PHQ-9 maps directly onto the diagnostic criteria for major depressive disorder (11). These scales have gained wider acceptance by clinicians than many other self-report measures, have been validated in clinical populations, and are sensitive to change (11,12). Neither scale requires extensive training to administer or interpret, nor do they impose substantial time burdens on clinicians or patients.

Conclusions

In summary the GAP committee notes that regular outcomes assessment represents a structural change in health care delivery that may alter the patient-provider interaction but may also improve patient outcomes and facilitate the systematic application of step-based care—that is, care in which the "next treatment step" is based on response of patient's symptoms to treatment and sequenced evidence-based treatments. The effects of measurement- and algorithm-based approaches to care continue to need careful assessment, and the cost-effectiveness of these approaches is not yet known. Although early studies suggest that more patients have positive outcomes with the use of routine standardized measures (7,13) and the face validity for measuring patient outcomes is high, the evidence for and barriers to routine outcomes measurement may differ for the various mental health disorders. However, the GAP committee strongly encourages clinicians and health systems to implement standardized assessments of patients' outcomes for mental health disorders, such as depression, for which there is evidence for benefit.

Acknowledgments and disclosures

The authors report no competing interests.

References

1. Mazade NA, Glover RW: Critical priorities confronting state mental health agencies. Psychiatric Services 58:1148–1150, 2007
2. Cooper D: Blue Cross announces fee increase and program to measure psychiatric outcomes. Vital Signs: Member Publication of the Massachusetts Medical Society, Aug 2007. Available at www.massmed.org
3. Committee on Psychopathology, Group for the Advancement of Psychiatry: Outcomes assessment and psychiatric services. Hospital and Community Psychiatry 45:1165, 1994
4. Trivedi MH, Daly EJ: Measurement-based care for refractory depression: a clinical decision support model for clinical research and practice. Drug and Alcohol Dependence 88(suppl 2):S61–S71, 2007
5. Depression Guideline Panel. Depression in Primary Care: Vol 2. Treatment of Major Depression. Clinical Practice Guideline, No 5. AHCPR pub no 93-0551. Rockville, Md, US Department of Health and Human Services, Public Health Service, Agency for

Health Care Policy and Research, 1993

6. Practice guideline for major depressive disorder in adults. American Psychiatric Association. American Journal of Psychiatry 150 (suppl 4):1–26, 1993
7. Trivedi MH, Rush AJ, Crismon ML, et al: Clinical results for patients with major depressive disorder in the Texas Medication Algorithm Project. Archives of General Psychiatry 61:669–680, 2004
8. Rush AJ Jr, First MB, Blacker D (eds): Handbook of Psychiatric Measures, 2nd ed. Washington, DC, American Psychiatric Association, 2007
9. NetOutcomes for Improving Clinical Care. Little Rock, Ark, Physician Assessment Services and Systems, 1998. Available at www.netoutcomes.net
10. Cella D, Yount S, Rothrock N, et al: The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. Medical Care 45(5 suppl 1):S3–S11, 2007
11. Lowe B, Unützer J, Callahan CM, et al: Monitoring depression treatment outcomes with the Patient Health Questionnaire–9. Medical Care 42:1194–1201, 2004
12. Rush AJ, Trivedi MH, Ibrahim HM, et al: The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biological Psychiatry 54:573–583, 2003
13. Lambert MJ, Harmon C, Slade K, et al: Providing feedback to psychotherapists on their patients' progress: clinical results and practice suggestions. Journal of Clinical Psychology 61:165–174, 2005

Submissions for Datapoints Column Invited

Submissions to the journal's Datapoints column are invited. Datapoints encourages the rapid dissemination of relevant and timely findings related to clinical and policy issues in psychiatry. National data are preferred. Areas of interest include diagnosis and practice patterns, treatment modalities, treatment sites, patient characteristics, and payment sources. The analyses should be straightforward, so that the figure or figures tell the story. The text should follow the standard research format to include a brief introduction, description of the methods and data set, description of the results, and comments on the implications or meanings of the findings.

Datapoints columns, which have a one-page format, are typically 350 to 400 words of text with one or two figures. The maximum total word count—including the title, author names, affiliations, references, and acknowledgments—is 500. Because of space constraints, submissions with multiple authors are discouraged; submissions with more than four authors should include justification for additional authors.

Inquiries or submissions should be directed to column editors Amy M. Kilbourne, Ph.D., M.P.H. (amy.kilbourne@va.gov), or Tami L. Mark, Ph.D. (tami.mark@thomsonreuters.com).