

Common Ground: A Framework for Selecting Core Quality Measures for Mental Health and Substance Abuse Care

Richard C. Hermann, M.D., M.S.
R. Heather Palmer, M.B.,B.Ch., S.M.

There is widespread interest throughout the mental health system in routine quality assessment to facilitate quality improvement, oversight, purchasing, and consumer choice. In the absence of agreement on a limited number of meaningful and feasible quality measures, delivery systems, payers, managed care organizations, regulators, and accreditors have each implemented unique measures and specifications. The resulting heterogeneity among measures has increased the burden on providers, limited the comparability of results, and hindered efforts to focus limited resources on further development of the most promising measures. Policy makers have initiated efforts for stakeholders to reach consensus on a core set of measures for common use, but barriers to progress remain, including differences in stakeholder needs and trade-offs between prioritizing desirable attributes of measures and representing the mental health system broadly. The authors present a framework for the selection of a core set of measures, clarify divergent perspectives, and make recommendations for further development of core quality measures for mental health care. (*Psychiatric Services* 53:281–287, 2002)

In recent years, stakeholders in the mental health care system have moved to embrace routine assessment of the quality of care delivered. To a varying extent, health care facilities, delivery systems, and plans use data on quality to improve the care they provide (1–3). Government regulators, accreditors, and managed behavioral health care organizations provide care delivery organizations with comparative data on quality in efforts to drive improvement from above (4, 5). More than 50 stakeholder groups have proposed more than 300 process measures for quality assessment (6,7).

However, these measures vary widely in their evidence base, technical sophistication, and readiness for routine use. Moreover, the number and diversity of these measures increase the burden of data collection on providers and reduce the usefulness of the results.

Policy makers and other stakeholders have responded to the proliferation of measures by calling for the adoption of core measures (8–12), or “standardized performance measures that are selectively identified and limited in number . . . [and that] can be applied across programs . . . [with] pre-

cisely defined specifications . . . [and] standardized data collection protocols” (11). Core measures could reduce the burden on facilities and plans, which often must measure and report on different aspects of care for each agency, accreditor, and payer to which they are accountable. Common specifications would increase the comparability of data across facilities and plans. Core sets would also focus resources on the most promising measures for further development, testing, and case-mix adjustment. The concept of core measures is based on a number of assumptions: that quality measures meeting a broad range of criteria are available, that the same measures can be used for multiple purposes, and that diverse stakeholders can agree on a small number of measures.

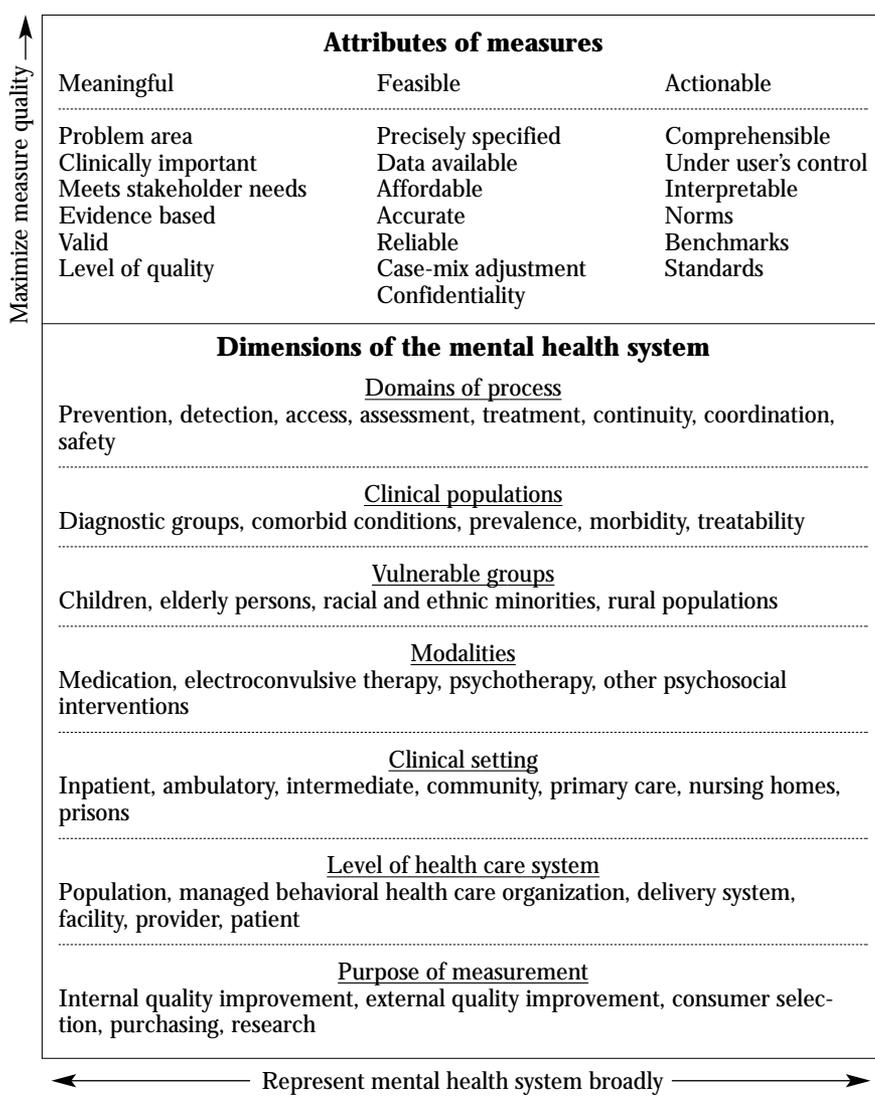
Despite the advantages it would offer, a broad-based set of core measures for the U.S. mental health system has proved elusive. Individual stakeholder groups have produced measures for use with their membership. These groups include the National Committee for Quality Assurance (NCQA) (13), the American Managed Behavioral Healthcare Association (AMBHA) (14), and the National Association of State Mental Health Program Directors (NASMHPD) (15). The Washington Circle Group has developed and begun pilot testing measures for substance abuse (16). An initiative led by the American College of Mental Health Administration has made progress in identifying potential areas to measure (17). These efforts represent important steps toward convergence.

In March 2001, the Substance

Dr. Hermann is assistant professor of psychiatry, health policy, and management at Harvard Medical School and School of Public Health and director of the Center for Quality Assessment and Improvement in Mental Health at the Cambridge Health Alliance. Dr. Palmer is professor of health policy and management and director of the Center for Quality of Care Research and Education at the Harvard School of Public Health. Send correspondence to Dr. Hermann at the Department of Psychiatry, Cambridge Hospital, 1493 Cambridge Street, Cambridge, Massachusetts 02139 (e-mail, richard_hermann@hms.harvard.edu).

Figure 1

A framework for selecting core quality measures



Abuse and Mental Health Services Administration (SAMHSA) hosted a summit at the Carter Center in Atlanta to pursue the development of systemwide core measures. Individuals from more than 75 mental health and substance abuse organizations attended, representing a wide range of perspectives. With this and other initiatives under way, a review of considerations underlying the development of a core set of measures is timely.

Overview of the framework

We present a framework, summarized in Figure 1, that integrates considerations of the attributes of measures—in a sense, the qualities of quality measures—and dimensions of the mental health system. In doing so, we highlight two challenges to the de-

velopment of a core set of measures. First, there are fundamental tensions between maximizing the quality of measures and broadly representing diverse features of the mental health system. Second, competing priorities among stakeholders become manifest in the process of selecting what to measure.

In presenting a framework, we seek to make more explicit these measure attributes and system dimensions—and trade-offs between them—in order to educate stakeholders and facilitate the process of developing a core set of measures. Our framework is built on the work of others. Most groups that have developed measures have based their initiatives on a framework, typically some variation of Donabedian's triad of structure, process,

and outcome (18). For example, the American Psychiatric Association categorized measures into four groups: access, quality, perceptions of care, and outcomes (19). Several organizations have also described desirable attributes of measures, such as meaningfulness or basis in evidence (9).

We focus on process measures, which examine interactions between patients and the health care system. Many of the concepts we discuss could be applied to other methods of assessing quality of care, including measurement of outcomes, patients' perspectives, and the fidelity of treatment programs to evidence-based protocols. Our focus is not intended to imply that process measures are superior; we have previously written about the importance of varied approaches to quality assessment (20, 21). Our focus here reflects the fact that process measures are already widely used in mental health care (6) and could be used more effectively.

An example of a simple process measure is the proportion of outpatients for whom an antidepressant is prescribed for major depression who remain on the medication for the 12-week acute treatment phase (an NCQA measure) (22). With adjustment for differences among patients—not a small matter—this measure could be applied to case-loads of individual providers, to clinics and group practices, and to facilities, health plans, or beneficiaries of public health insurance programs.

In the sections that follow we ask three questions related to principles of core measure selection. What characteristics of process measures are desirable? What features of the mental health system should these measures evaluate? What trade-offs are necessary to achieve balance between the attributes of high-quality measures and the diverse priorities of stakeholders?

Desirable attributes of measures

Meaningfulness

Embedded in the construct of the meaningfulness of a quality measure are several concepts, some inherently subjective and others based on more objective information. The face validity of a measure is subjective. Individu-

als who are familiar with a process can be asked whether the process is clinically important to measure, whether there is a gap between optimal and actual practice, and whether closing the gap would improve patients' outcomes. An individual's response to these questions may vary with his or her stake in the health care system.

In some cases, empirical data are available to inform these judgments. Some measures have been derived from research evidence establishing the effectiveness of a clinical process. For instance, the previously described measure of antidepressant use is based on randomized controlled studies establishing that an acute episode of depression is more likely to remit if a patient takes an antidepressant medication for at least 12 weeks. In contrast, there is no evidence that a single outpatient visit within 30 days of inpatient care for depression has an impact on outcome—a continuity-of-care measurement from the same developer (22).

Research data may also address whether a measure reflects a potential quality problem by identifying gaps between actual and optimal practice. NCQA data for 2000 show that 55.6 to 62.6 percent of individuals enrolled in participating health plans who initiate antidepressant treatment for major depression discontinue the medication before 12 weeks (5).

After measurement of a particular process has been implemented, predictive validity can be assessed. Predictive validity characterizes the association within a treated sample between conformance to the measure and clinical outcome. In such a retrospective analysis, Melfi and colleagues (23) found that adherence to antidepressant treatment guidelines was associated with a lower probability of relapse or recurrence.

Other aspects of meaningfulness of measures are less well established, such as the threshold of quality assessed by a measure. Does a given measure reflect minimally acceptable care, average care, or best practice? The continuity measure—the proportion of patients who make an outpatient visit within 30 days of hospital discharge—seemingly reflects a minimal level of care. In contrast, the measure of the proportion of patients

remaining on an antidepressant medication for 12 weeks or more is based on a guideline recommendation. An organization can further calibrate a measure's threshold to clinical circumstances by setting a performance standard—for example, that 80 percent of patients continue an antidepressant medication for 12 weeks.

Feasibility

Although the concept proposed for a measure may be simple, constructing the measure's specifications is often more complex. Each measurable component—clinical process, population, and data source—must be defined operationally. Inclusion criteria, procedure codes, and time frames must be precisely specified. Data collection protocols, abstraction forms, and programming specifications need to be developed. The measure should then be tested in a variety of settings and health systems, with particular attention given to the availability and accuracy of the data and the reliability of the collection process.

One of the most important challenges in selecting core measures is the affordability of data collection—another aspect of feasibility. Typically, use of administrative or claims data, which are collected routinely in the course of administering or billing for care, is the least burdensome. However, resources are required to access, link, and analyze these data. Abstracting data from medical records or collecting data directly from clinicians and patients is more labor-intensive. The collective burden of gathering data for a core set of measures must be balanced against the resources available for measurement. To date, there is no consensus among facilities, payers, and regulators about appropriate costs for measurement or who should pay them.

Other challenges to the feasibility of routine process measurement remain. An operational balance is needed between having access to data for measurement and safeguarding patient confidentiality. This issue is likely to receive greater attention as federal privacy regulations in the Health Insurance Portability and Accountability Act (HIPAA) are implemented.

Comparing results across clini-

cians, facilities, or plans may require adjustment for differences in patient populations that are unrelated to quality of care. Such case-mix adjustment is often said to be less necessary for process measures than for outcome measures, because a clinical process, such as the assessment of a patient's mental status, may be fully under the control of the clinician. However, many process measures rely on utilization data to determine whether the duration or intensity of treatment is appropriate. For psychiatric care, utilization data are problematic because they reflect the actions of both clinicians and patients. Clinicians can influence patients' compliance by providing education, scheduling follow-up visits, using outreach resources, and addressing medication nonresponse and side effects. Nonetheless, persons with severe mental illness have no-show rates for scheduled appointments as high as 50 percent (24). Differences in patient populations, such as the rate of comorbid illness and substance abuse, also influence compliance. Thus statistical adjustment for underlying differences in patient populations may be needed to provide fair comparisons among clinicians or facilities.

For other patient characteristics, such as race or ethnicity, adjustment would not serve the purpose of quality assessment. In this context, the goal would be to identify groups receiving substandard care and intervene to narrow these disparities. Case-mix adjustment methods can range from stratification to more complex multivariate analyses, but such methods are relatively underdeveloped for mental health measures.

Actionability

Although a measure may be well defined and may address a quality problem, the results may not be "actionable"—that is, users of the measure may not be able to act on them to improve care. Highly technical specifications may yield a result not easily comprehensible to users. Results may reveal a serious problem whose solution is not under the user's control. For example, the nursing staff of an inpatient unit initiated a quality improvement project to reduce the high rate of med-

ication errors detected on their unit, only to learn that labeling errors at the pharmacy—a process over which they had little authority—caused the problem. They were able to refocus the project to improve detection of errors; however, in many cases a solution is not so readily available.

More data are needed to facilitate interpretation of quality measurement results. For many measures, 100 percent conformance is not a reasonable expectation. Despite a clinicians' best efforts, some patients will discontinue an antidepressant prematurely. Consequently, a facility that receives a performance rate of 60 percent on this measure may wonder whether better results might have been attainable.

Several types of comparative data can enhance interpretability. Norms, such as those used in NCQA's Quality Compass (5), reflect average results for population-based samples. Benchmarks reflect the results attained by the best-performing plans and providers (25). Different points of comparison may have different effects. Using norms may reinforce a status quo, whereas using benchmarks may motivate participants to improve. Currently, few quality measures in mental health have established benchmarks (6). In the absence of benchmarks, some health care organizations prescribe standards, which are thresholds that they believe represent an acceptable and achievable level of care.

Mental health system representativeness

Selecting core measures to assess the quality of mental health care is an ambitious undertaking. The mental health care "system," which was described as a "de facto system" first by Regier and colleagues (26) and then more broadly by Fox and colleagues (27), includes a sprawling array of services, settings, populations, and modalities. Each of these dimensions, elaborated on below and depicted in Figure 1, is a potential area for core measures. Which dimensions, and which areas within them, should be given priority? As with decisions about attributes of measures, these decisions can be informed by data but may also differ by stakeholder perspective.

Domains of process

Donabedian defined process as "a set of activities that go on within and between practitioners and patients" (18). Accordingly, process measurement is not limited to treatment but can also examine processes of care before treatment (such as detection, access, and assessment), after phases of treatment (such as continuity across levels of care), and across sources of treatment (such as coordination among treaters, including a patient's prescribing clinician and therapist). Process measurement can also assess primary prevention in healthy populations. A seventh domain, patient safety, has been subject to increasing attention in recent years as research on the impact of medication errors and other safety concerns has accumulated (28). Measures have been developed for each of these domains in behavioral health, and each domain is highly important to one or more stakeholder groups (6).

Clinical populations

Some process measures can be applied broadly to populations, across diagnostic categories and demographic groups. Others are condition specific. In the selection of core measures, a wealth of data describing mental illness in populations is available for setting priorities among clinical disorders. National epidemiological studies have identified highly prevalent conditions, such as major depression, substance abuse, and several anxiety disorders (29). Bipolar disorder, schizophrenia, and obsessive-compulsive disorder are less prevalent but warrant consideration because of their impact on functioning—the three conditions rank among the ten leading causes of disability worldwide, as measured in years lived with a disability (30).

Although measurement often focuses on acute illness, Wagner and associates (31) have pointed out the need to assess and improve management of chronic conditions. The selection of measures should focus first on conditions that have clearly effective treatments and for which the quality of care can be decisive. Medical and substance use problems commonly co-occur with mental illness

and should also be a priority for measurement and quality improvement. Comorbid illnesses among persons who have primary psychiatric conditions are frequently undetected, inadequately treated, and associated with poorer outcomes (32–35).

Vulnerable groups

Diagnosis alone does not adequately capture population characteristics that are relevant in identifying what to measure. Research studies have highlighted quality-of-care problems experienced by groups who may have a limited ability to advocate for themselves and to navigate an increasingly complex health care system. These more vulnerable population groups include children (36), elderly persons (37,38), and persons with severe and persistent mental illness (39). Moreover, an expanding body of literature has documented disparities in health care among racial and ethnic minorities (40,41). Geographic variation in the concentration of mental health clinicians has also led to problems with access and quality in rural areas (42,43).

Modalities

Amid the diverse array of interventions, a core set of measures should at a minimum include both biological and psychosocial treatments. Measures can be based on findings from research studies of overuse and underuse of treatments, from analyses of utilization patterns and expenditures (44), and from population-based needs assessment. In developing measures, one challenge is to advance beyond determining whether or not a patient received a treatment to assessing the adequacy of the treatment's content, intensity, and duration. To date, this limitation has been more effectively addressed for biological than for psychosocial interventions (6).

Settings

As the focus of clinical care shifts from the inpatient setting—where quality assurance efforts originated—to ambulatory, intermediate, and community levels of care, quality measures must assess care at a range of levels and in various settings. Research studies have documented problems with the quality of mental

health care in the general medical sector (45) and in nontraditional clinical settings with high rates of mental illness, such as homeless shelters, prisons, and nursing homes.

System levels

Comparing quality of care across health plans can be thought of as assessment across horizontal components of the health care system. A more complete picture would include processes taking place within the various vertical strata of the system: patients, clinicians, facilities, delivery systems, managed behavioral health care organizations or plans, and populations. Methods of evaluating the treatment of individual patients, as opposed to aggregate groups, have become more systematic, building on traditional methods such as peer review and morbidity and mortality conferences to incorporate structured methods of assessment (46,47). Assessment at the level of the managed behavioral health care organization might compare organizations' denial and appeal rates in utilization management. A population-based measure might assess the rate of detection or appropriate treatment of depression on the basis of the prevalence of the disorder in communities.

Measurement purposes

Although process measurement in health care systems has been used primarily for internal and external quality improvement activities (6), it has other potential uses. As health care has become increasingly market driven, employer coalitions have sought measures of quality to inform purchasing decisions, which are otherwise based principally on cost. Similarly, consumer groups have sought measures that members and their families can use in the selection of providers and plans. Mental health services researchers also use measures of quality as one means of assessing the impact of organizational and financial changes in the delivery of care.

Balancing measure and system considerations

The challenges inherent in reaching consensus among stakeholders on core measures can be seen by viewing

Figure 1 not as a two-dimensional diagram but as a multidimensional matrix. Each cell in the matrix represents the intersection of measurement principles, potential subjects of measures, and conflicting stakeholder perspectives that must be prioritized internally and relative to other cells in the selection of measures for common use. Prioritizing and integrating these components was an implicit goal at policy-making forums such as the stakeholder meetings held by the American College of Mental Health Administration and the recent summit at the Carter Center. The prioritization process itself can also be made explicit, as it was in a consensus development process conducted by the Center for Quality Assessment and Improvement in Mental Health at Harvard Medical School. A diverse panel of stakeholders used a modified Delphi process to rate the attributes of measures; rating scores were then used to select a dimensionally balanced set of candidate core measures (Hermann R, Palmer R, Shwartz M, et al, unpublished manuscript, 2001).

Some fundamental tensions have emerged in both implicit and explicit processes. In many cases, assigning priority to one measure comes at the expense of another. Highly meaningful measures of evidence-based practices often require data from medical records and other sources rich in clinical information. However, these data are more costly to collect than more commonly used claims data, and therefore the measures may be less feasible. Detailed specifications produce an accurate and reliable measure that has high feasibility, but potential users may find results from complex measures difficult to comprehend, resulting in low actionability.

Fortunately, the attributes of measures can sometimes be improved. Upcoming federal standards under HIPAA may improve the accuracy and comparability of claims data by specifying data elements and definitions (48). Influential payers such as Medicare are beginning to add clinically important variables to the administrative data that providers submit for reimbursement, facilitating evidence-based quality measurement without chart review (49,50). Even

chart review itself—opposed by some groups because of its cost—can become more efficient. Nearly every health care facility does some chart review in response to external requirements. Developing a consensus on the most important chart-based variables would improve the use of limited resources for chart review.

A basic tension arises in the development of a core set of measures between seeking to maximize the quality of measures, depicted by the vertical arrow in Figure 1, and representing the breadth and diversity of the mental health care system, depicted by the horizontal arrow. For example, a principle of measure selection is a strong basis in research evidence. However, selecting measures that are supported by well-controlled research leads to a preponderance of medication measures that pertain to relatively few conditions (6)—hardly a broad representation of mental health care. Although some psychosocial interventions, including assertive community treatment, have a rigorous research foundation, the complexities of these interventions necessitate a more extensive evaluation of fidelity to the empirically tested model. A simple process measurement does not suffice. In contrast, the results of drug trials can be more easily assembled into simple measures of diagnosis, drug selection, dosage, and duration. Well-controlled trials exist for specific types of psychotherapy, such as interpersonal or cognitive-behavioral therapy, but neither administrative claims nor medical records typically document the type of therapy provided.

Understanding these tensions helps to bring into focus some of the trade-offs involved in the production of a set of core measures. Differences in priorities among stakeholders complicate the selection process. There has been little systematic study of the attitudes of mental health stakeholders toward quality measures. However, a review of stakeholders' published reports (6) and observations at national meetings suggests that differences are marked, rational, and deeply held.

Representatives of groups who carry out the data collection for quality measurement—managed behavioral

health care organizations, delivery systems, and accreditors—have frequently expressed concern about measurement burden. AMBHA's Performance Measures for Managed Behavioral Healthcare Programs (PERMS 2.0) (14) contains a preponderance of utilization measures, appropriate for a trade association for managed behavioral health care organizations, which principally manage resource utilization. Physicians are trained intensively in technical clinical processes, such as diagnosis and treatment interventions. Thus it is not surprising that the American Psychiatric Association's set of measures emphasizes technical quality of care. Lacking such clinical training, consumers and families often rely on personal experience to identify problems with quality; the consumer-focused measures of the Mental Health Statistics Improvement Project (MHSIP) focus on interpersonal experiences (51). More subtle differences emerge even among groups allied in many areas. Family advocates have encouraged evidence-based measures of medication use and assertive community treatment outreach, while consumers have advocated for measures that promote autonomy, generally by emphasizing such topic areas as recovery, peer support, and housing assistance.

Several opportunities to identify preliminary core measures for mental health care will present themselves over the next few years. The multi-stakeholder process begun last year at the Carter Center continues to pursue common process and outcome measures for adults and children with mental health and substance use disorders. The American Medical Association's Physician Consortium for Performance Improvement plans to join with other organizations to produce a consensus set of measures for depression. The Agency for Healthcare Research and Quality has funded a National Quality Measure Clearinghouse to disseminate detailed information about selected measures in each area of medicine, including mental health.

The challenge to policy makers, stakeholder group leaders, and measurement methodologists is to work together to select a set of measures

that reflects each of their priorities enough to warrant individual participation but that also covers enough common ground to justify broad use. There is widespread anxiety among stakeholder groups that, once selected, these measures—and the underlying treatment processes—will rapidly receive disproportionate attention. Policy leaders might gain stakeholder confidence by emphasizing gradual implementation of preliminary measures designed initially to provide individuals and institutions with confidential feedback for internal improvement activities. Experience with such measures will allow users to improve their feasibility and will generate data to evaluate validity and develop benchmarks. After iterative cycles of measure development, use of the measures for oversight activities and public disclosure may be better received. Fears of undue focus can be assuaged in part by rotating some of the measures in a core set over successive measurement cycles. However, continuous use of some measures allows for an examination of trends over time.

Process measures are most useful when combined with complementary methods of quality assessment. Clinicians can use outcome measures to compare the progress of their patients with that of patients treated at similar facilities. Adding process measures can then highlight areas for improvement when outcomes are lagging (20,21). Simple population-based process measures can indicate whether a patient is receiving an appropriate form of treatment. For example, such a measure will reveal whether an individual with unstable symptoms of schizophrenia and multiple hospitalizations is enrolled in an assertive community treatment program. More complex fidelity measures can be used to "drill down" to evaluate the quality of an individual assertive community treatment program in greater detail (52).

Assessment of patients' perceptions of care provides information about interpersonal process to complement the technical detail of process measurement (53). Another potential linkage is between measures that identify gaps in quality and the availability of

effective models for improvement. For example, health plans that score poorly on the NCQA's evidence-based depression measures may be motivated to consider quality improvement models for depression such as those developed by Wells and associates (54) and Katon and colleagues (55).

Further experience with core measures will in turn allow for further refinement of the concept of a single core set from a societal—that is, a broadly based—perspective. It may be that a "core menu" of measures is a more useful vision, with subsets used for individual purposes (quality improvement versus purchasing), settings (specialty versus primary care), and populations (competitively employed versus psychiatrically disabled groups). Despite the obstacles inherent in selecting core measures, we should not lose sight of the advantages: lowering measurement burden, focusing resources for development and testing, and improving the interpretability of results. Each of these steps will advance the mission of identifying quality problems and improving care for mental and substance use disorders. ♦

Acknowledgments

This work was supported by grant K08-MH001477 from the National Institute of Mental Health, by grant R01-HS-10303 from the Agency for Healthcare Research and Quality, and by the Substance Abuse and Mental Health Services Administration.

References

1. Hermann R, Regner J, Yang D, et al: Developing a quality management system for behavioral healthcare: the Cambridge Health Alliance experience. *Harvard Review of Psychiatry* 8:251–260, 2000
2. Carpinello S, Felton C, Pease E, et al: Designing a system for managing the performance of mental health managed care: an example from New York State's prepaid mental health plan. *Journal of Behavioral Health Services Research* 25:269–278, 1998
3. McFarland D, Harmann L, Lhotak C, et al: The quest for TQM in a community mental health center: using the Baldrige criteria as a framework. *Joint Commission Journal on Quality Improvement* 22:37–47, 1996
4. Suber M, Martin M, Jones F, et al: The quality information management program database in South Carolina. *Joint Commission Journal on Quality Improvement* 22: 629–639, 1996
5. *Quality Compass 2000*. Washington, DC, National Committee for Quality Assurance, 2000

6. Hermann R, Leff H, Palmer R, et al: Quality measures for mental health care: results from a national inventory. *Medical Care Research and Review* 57(suppl 2):135-153, 2000
7. Hermann RC, Palmer RH, Finnerty M, et al: Process measures for the assessment and improvement of care for schizophrenia. *Schizophrenia Bulletin*, in press
8. Institute of Medicine: *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC, National Academy Press, 2001
9. Institute of Medicine: *Envisioning the National Health Care Quality Report*. Washington, DC, National Academy Press, 2001
10. *World Mental Health Report 2001: Mental Health: New Understanding, New Hope*. Geneva, World Health Organization, 2001
11. Manderscheid R, Henderson M: The field needs to agree on quality measures. *Behavioral Health Accreditation and Accountability Alert*, Feb 2001, pp 4-5
12. Oryx: *The Next Evolution in Accreditation*. Oakbrook Terrace, Ill, Joint Commission on Accreditation of Healthcare Organizations, 2000
13. *Health Plan Data and Information Set (HEDIS 2000)*. Washington, DC, National Committee for Quality Assurance, 1999
14. *Performance measures for managed behavioral healthcare programs (PERMS 2.0)*. Washington, DC, American Managed Behavioral Health Association, 1998
15. *Report of the Technical Workgroup of the NASMHPD President's Task Force on Performance Indicators*. Alexandria, Va, National Association of State Mental Health Program Directors, 1997. Available at www.nasmhpd.org/nri
16. McCorry F, Garnick D, Bartlett J, et al: Developing performance measures for alcohol and other drug services in managed care plans. *Joint Commission Journal on Quality Improvement* 26:633-643, 2000
17. *A Proposed Consensus Set of Indicators for Behavioral Health*. Pittsburgh, American College of Mental Health Administration, 2001
18. Donabedian A: *Explorations in Quality Assessment and Monitoring: The Definition of Quality and Approaches to Its Assessment*. Ann Arbor, Mich, Health Administration Press, 1980
19. *Report of the American Psychiatric Association Task Force on Quality Indicators*. Washington, DC, American Psychiatric Association, 1999
20. Palmer R: Using health outcomes data to compare plans, networks, and providers. *International Journal for Quality in Health Care* 10:477-483, 1998
21. Hermann R: Linking outcome measurement with process measurement for quality improvement. In *Outcome Measurement in Psychiatry: A Critical Review*. Edited by Ishak B, Sederer LI. Washington, DC, American Psychiatric Press, in press
22. *Health Plan Employer Data and Information Set (HEDIS 3.0)*. Washington, DC, National Committee on Quality Assurance, 1996
23. Melfi C, Chawla A, Croghan T, et al: The effects of adherence to antidepressant treatment guidelines on relapse and recurrence of depression. *Archives of General Psychiatry* 55:1128-1132, 1998
24. Smoller J, McLean R, Otto M, et al: How do clinicians respond to patients who miss appointments? *Journal of Clinical Psychiatry* 59:330-338, 1998
25. Kiefe C, Weissman N, Allison J, et al: Identifying achievable benchmarks of care: concepts and methodology. *International Journal for Quality in Health Care* 10:443-447, 1998
26. Regier D, Narrow W, Rae D, et al: The de facto US mental and addictive disorders service system: Epidemiologic Catchment Area prospective 1-year prevalence rates of disorders and services. *Archives of General Psychiatry* 50:85-94, 1993
27. Fox J, Merwin E, Blank M: De facto mental health services in the rural south. *Journal of Health Care for the Poor and Underserved* 6:434-468, 1995
28. Institute of Medicine: *To Err Is Human: Building a Safer Health System*. Washington, DC, National Academy Press, 1999
29. Kessler R, McGonagle K, Zhao S, et al: Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: results from the National Comorbidity Survey. *Archives of General Psychiatry* 5:8-19, 1992
30. Murray C, Lopez A (eds): *Summary: The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability From Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020*. Cambridge, Mass, Harvard University Press, 1996
31. Wagner E, Austin B, Von Korff M: Organizing care for patients with chronic illness. *Milbank Quarterly* 74:511-544, 1996
32. Drake R, Mueser K, Clark R, et al: The course, treatment, and outcome of substance disorder in persons with severe mental illness. *American Journal of Orthopsychiatry* 66:42-51, 1996
33. Rabinowitz J, Mark M, Popper M, et al: Physical illness among all discharged psychiatric inpatients in a national case register. *Journal of Mental Health Administration* 24:82-89, 1997
34. Karasu T, Waltzman S, Lindenmayer J, et al: The medical care of patients with psychiatric illness. *Hospital and Community Psychiatry* 31:463-472, 1980
35. Maricle R, Hoffman W, Bloom J, et al: The prevalence and significance of medical illness among chronically mentally ill patients. *Community Mental Health Journal* 23:81-90, 1987
36. McGlynn E, Halfon N: Overview of issues in improving quality of care for children. *Health Services Research* 33:977-1000, 1998
37. Blank K, Cohen C, Cohen G, et al: Failure to adequately detect suicidal intent in elderly patients in the primary care setting. *Clinical Geriatrics* 9:26-36, 2001
38. Cohen C, Cohen G, Blank K, et al: Schizophrenia and older adults, an overview: directions for research and policy. *American Journal of Geriatric Psychiatry* 8:19-28, 2000
39. Wang P, Berglund P, Kessler R: Recent care of common mental disorders in the United States: prevalence and conformance with evidence-based recommendations. *Journal of General Internal Medicine* 15:284-292, 2000
40. Melfi C, Croghan T, Hanna M, et al: Racial variation in antidepressant treatment in a Medicaid population. *Journal of Clinical Psychiatry* 61:16-21, 2000
41. Adebimpe V: Race, racism, and epidemiological surveys. *Hospital and Community Psychiatry* 45:27-31, 1994
42. Knesper D, Wheeler J, Pagnucco D: Mental health services providers' distribution across counties in the United States. *American Psychologist* 39:1424-1434, 1984
43. Rost K, Williams C, Wherry J, et al: The process and outcomes of care for major depression in rural family practice settings. *Journal of Rural Health* 11:114-121, 1995
44. Manderscheid R, Henderson M (eds): *Mental Health, United States, 1998*. Rockville, Md, Center for Mental Health Services, 1998
45. Joseph R, Hermann R: Screening for psychiatric disorders in primary care settings. *Harvard Review of Psychiatry* 6:165-170, 1998
46. Rubenstein L, Kahn K, Reinisch E, et al: Changes in quality of care for five diseases measured by implicit review, 1981-1986. *JAMA* 264:1974-1979, 1990
47. Palmer R, Strain R, Maurer J, et al: A method for evaluating performance of ambulatory pediatric tasks. *Pediatrics* 73:269-277, 1984
48. Reeder L: Special report: HIPAA electronic data standards released. *Healthcare Leadership Management Report* 8(8):1-5, 2000
49. Palmer R: Process-based measures of quality: the need for detailed clinical data in large health care databases. *Annals of Internal Medicine* 127:733-738, 1997
50. Cleves M, Weiner J, Cohen W, et al: Assessing HCFA's health care quality improvement program. *Joint Commission Journal of Quality Improvement* 23:550-560, 1997
51. Teague G, Ganju V, Hornik J: The MHSIP mental health report card: a consumer-oriented approach to monitoring the quality of mental health plans. *Evaluation Review* 21:330-341, 1997
52. Teague G, Bond G, Drake R: Program fidelity in assertive community treatment: development and use of measure. *American Journal of Orthopsychiatry* 68:216-232, 1998
53. Cleary P, McNeil B: Patient satisfaction as an indicator of quality of care. *Inquiry* 25:25-36, 1988
54. Wells K, Sherbourne C, Schoenbaum M, et al: Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. *JAMA* 283:212-220, 2000
55. Katon W, Von Korff M, Lin E, et al: Collaborative management to achieve treatment guidelines: impact on depression in primary care. *JAMA* 273:1026-1031, 1995