

Quality Measurement in Health Care: A Need for Leadership amid a New Federalism

Richard C. Hermann, MD, MS, and Caitlin K. Rollins

federalism, improvement, measurement, measures, medical necessity, practice guidelines, quality

Over the past decade, market forces have increasingly reshaped the American health care system. Public and private payers have hired for-profit managed care organizations to approve the utilization of health care services. Purchasers and their employees select among competing health plans. Health care facilities and systems vie for contracts to treat populations of patients.

The infusion of market forces into the health care system has outpaced the dissemination of data on the quality of care. Market forces can allocate resources efficiently in the presence of what economists term “perfect information.” But in the health care marketplace, not even imperfect information is available. While employees may receive information about health care benefits to guide their choices among plans, they know little about the intensity and types of care allowed by a plan’s utilization-review criteria. Similarly, consumers, purchasers, and payers lack sufficient information to compare providers based on their quality of the care.

Market economics suggest that when purchasers lack information that differentiates products on the basis of quality, they will maximize their utility by seeking the lowest price. If providers and plans are unable to demonstrate higher quality, they will have little financial incentive to expend resources to improve care. In fact, greater resource utiliza-

tion in the absence of evidence of better quality could be perceived as having lower value and could therefore result in reduced demand.

Health care organizations have sought to address the need for quality-of-care information by developing practice guidelines, measures that assess guideline conformance, and report cards that disseminate measure results. As is apparent from the text box, there are many active measurement initiatives in mental health care. Nevertheless, it is widely recognized that current efforts do not adequately inform market decisions.^{1,2} Multiple factors have impeded progress. Research evidence and practice guidelines are limited in terms of their specificity and comprehensiveness. Constructing meaningful and feasible measures is technically challenging. Resources to assess quality are limited. Comparing plans and providers requires case-mix adjustment (i.e., statistical adjustment to reduce the influence of patient characteristics on results that are unrelated to provider performance). However, these methods are at an early stage of development in the area of mental health care.³

Foremost among obstacles to meeting information needs on quality of care is the absence of leadership. A wide array of professional societies, industry groups, and government agencies have developed guidelines and measures reflecting diverse perspectives. In mental health care alone, more than 50 groups have proposed quality measures. These indicators vary widely in their evidence base, technical sophistication, and readiness for use.⁴ Some of the more highly regarded measures are described in Table 1, but few have been broadly adopted. Instead, overlapping, uncoordinated measurement requirements burden providers and deplete resources needed for improvement. The use of different measures (or varying specifications of the same measure) yields data that cannot be used to compare performance. The health care system needs leadership to develop consensus among

Reprint requests: Richard C. Hermann, MD, MS, Center for Quality Assessment and Improvement in Mental Health, Tufts-New England Medical Center, 750 Washington Street, NEMC #345, Boston, MA 02111. Email: rhermann@cqaimh.org

Harv Rev Psychiatry 2003;11:215–219.

© 2003 President and Fellows of Harvard College

DOI: 10.1080/10673220390235313

**QUALITY-MEASUREMENT INITIATIVES
IN MENTAL HEALTH CARE**

Benchmarking Indicators (National Association of Psychiatric Health Systems)
Consumer-Oriented Mental Health Report Card (Mental Health Statistics Improvement Program)
Core Performance Measures for Substance-related Conditions (Washington Circle Group)
Health Plan Employer Data and Information Set (HEDIS) (National Committee for Quality Assurance)
Indicators for Behavioral Health (American College of Mental Health Administration)
National Mental Health Program Performance Monitoring System (Department of Veterans Affairs)
ORYX Measure Initiative (Joint Commission on Accreditation of Healthcare Organizations)
Performance Measures for Managed Behavioral Healthcare Programs (American Managed Behavioral Healthcare Association)
Performance Measures for Mental Health Systems (National Association of State Mental Health Program Directors)
Physician Consortium for Performance Improvement (American Medical Association)
Quality Indicator Project (Maryland Hospital Association)
Quality Measures (Foundation for Accountability)
Task Force Report on Quality Indicators (American Psychiatric Association)

guidelines and standards for care, to select priority areas for measurement, to adopt core measures for common use, and to encourage dissemination of the results in a comparative format.

A question slowing the maturation of quality assessment and improvement in health care is: *who* will lead? A logical candidate is the federal government, which plays numerous roles in the health care system, including payer, provider, employer, and regulator. The need for quality-of-care data has emerged, however, during an era described as the New Federalism, a two-decade devolution of responsibility for social welfare and services from the federal government to states, localities, and the private sector.⁵ While this movement contributed to the rise of marketplace medicine—leading to the need for information about quality—it has also constrained the federal government from leading the way in providing that information.

Many nonfederal organizations play important roles in developing quality-measurement systems. These organizations include the Joint Commission on Accreditation of Healthcare Organizations (representing hospitals), the National Committee for Quality Assurance (representing health plans), and the Leapfrog Group (representing large

TABLE 1. Illustrative Quality Measures for Mental Health Care

Quality measure*	Source
Prevention	
Depression screening for primary care patients during 12-month period	Veterans Administration ¹⁶
Access	
Denials for mental health or substance-related services per number of requests in 12-month period	Washington Circle Group ¹⁷
Assessment	
Assessment of drug and alcohol use at initial evaluation for psychiatric disorder	American Psychiatric Association ¹⁸
Treatment	
≥12-week continuation after initiation of antidepressant drug for major depression	National Committee for Quality Assurance ¹⁹
Psychotherapy visit for individuals after hospitalization or ER visit for borderline personality disorder	American Psychiatric Association ¹⁸
Safety	
Number of inpatient injuries per patient day in 3-month period	New York State Office of Mental Health ²⁰
Assessment of suicidal ideation at initial evaluation for patients with major depression	Joint Commission on Accreditation of Healthcare Organizations ²¹
Continuity	
Outpatient visit within 7 days of hospital discharge for psychiatric disorder	National Committee for Quality Assurance ¹⁹
Coordination	
Enrollment in intensive case management for patients with 4 ER visits or 2 hospitalizations for schizophrenia in 12-month period	American Psychiatric Association ¹⁸
Outcome	
Schizophrenia Outcomes Module	University of Arkansas ²²
Behavior and Symptom Identification Scale (BASIS-32)	Eisen et al., 1986 ²³

*From the National Inventory of Mental Health Quality Measures (<http://www.cqaimh.org/quality.html>).

employers). Leading organizations in mental health care include the National Association of State Mental Health Program Directors, the National Association of Psychiatric Health Systems, the American Managed Behavioral Healthcare Association, and the American Psychiatric Association. But these groups represent different sectors of a complex and competitive system that has not coalesced to arrive at a single, consensus-based approach.

In order for quality measurement to mature and serve the information needs of the market, federal leadership may be needed to bring together disparate groups with conflicting agendas. In the current political environment, since federal mandates are unlikely, the government's role should primarily be as facilitator, funder, and participant in the health care market. In some cases, however, regulatory or legislative approaches may be necessary. The following are some specific objectives, along with suggestions for achieving them.

AN OBJECTIVE AND RIGOROUS PROCESS FOR DEVELOPING PRACTICE GUIDELINES

Clinical practice guidelines provide the foundation for quality measurement. They synthesize research evidence and consensus among experts into specific recommendations for treating common conditions. These recommendations often serve as a basis for quality measures of clinical practice. Through the mid-1990s, the Agency for Healthcare Research and Quality (AHRQ) developed practice guidelines for dozens of conditions using a rigorous method that included a thorough review of research evidence and, where evidence was lacking or inconclusive, a formal consensus-development process. Consensus panels included a broad spectrum of participants with expertise in the relevant areas. For example, the panel that developed AHRQ's widely respected guidelines for depression brought together psychiatrists, psychologists, social workers, primary care clinicians, consumers, and clinical researchers.⁶

Under pressure from congressional representatives who believed the government should "stay out of medical practice," AHRQ halted guideline development in 1999, with the intent that professional societies, managed care organizations, and other private organizations would continue the work.⁷ As hoped, guidelines have proliferated, but few have achieved the rigor, consensus, and credibility of AHRQ's work. Many organizations lack sufficient methodological expertise or resources, and it is common for organizations to have vested interests in guideline content. For example, pharmaceutical companies that fund guideline development for depression may have an incentive to promote newer, more costly antidepressants over traditional treatments. In contrast, managed care organizations and health plans have incentives to control costs by restricting the use of these drugs. Professional organizations come under pressure by clinicians to ensure that the organization's guidelines support types of psychotherapy practiced by its members. Each of these groups has extensive knowledge to contribute to guideline development, but objective leadership and a rigorous development process are needed to ensure a balanced result that is based on research evidence and consensus among stakeholders.

Perhaps there is a middle ground between the "intrusive hand" of government and a "hands-off" approach. AHRQ currently funds a dozen or so evidence-based practice centers (EPCs) at universities and other nonprofit institutions. These organizations currently assess research evidence on clinical interventions and conditions, but lack the mandate to take the next step and derive practice guidelines from evidence syntheses. Several of the EPCs would be well prepared to serve in the role of "honest broker" and lead stakeholders in the development of guidelines. Even a limited expansion of their mission would require broad support from Congress, but this represents just the type of public-private partnership that needs to be pursued.

DISCLOSURE OF HEALTH PLAN CRITERIA FOR DETERMINING MEDICALLY NECESSARY CARE

However rigorous their development, the influence of guidelines on clinical practice is limited because they are usually not integrated into the review, financing, and delivery of care. Instead, health plans employ utilization management (UM) criteria to guide staff clinicians and to approve or deny external requests for treatment. UM criteria directly influence access to care and its quality for thousands of Americans each day. For example, such criteria determine how long an individual experiencing an acute exacerbation of schizophrenia remains in the hospital, and whether an individual with substance dependence receives inpatient, outpatient, or intermediate care. Although UM criteria differ from practice guidelines in a number of ways, in theory they should be similarly informed by clinical research and expert opinion. We rarely know the specific content of these criteria, however, because they are typically proprietary, and few health plans disclose them. Public disclosure would allow consumers and purchasers to make informed decisions when selecting health plans. Prompted by law suits filed by consumer advocates, a large California health plan recently announced that it would make its internal guidelines available.⁸ In the interest of informing participants in the marketplace, Congress might encourage other plans to do the same.

A CORE SET OF MEANINGFUL AND FEASIBLE QUALITY MEASURES

The major stakeholder groups in the health care system—namely, accreditors, state and local health agencies, consumer advocacy groups, managed care and provider organizations, insurers, employer purchasers, labor groups—need to come together around common measures of quality.

Several federally funded umbrella organizations are working to achieve consensus among the various players. The Institute of Medicine produced a framework that included priorities for quality assessment and improvement.² The National Quality Forum has convened representatives from major stakeholder organizations to adopt consensus-based measures. AHRQ is compiling a national quality report from existing public health and quality-assessment databases.² In mental health and substance-related care, stakeholders have worked together on consensus measures under the auspices of the Forum on Performance Measures and the Washington Circle Group.⁹

Although each of these groups has made progress, broad agreement on meaningful, feasible, and well-specified measures has been achieved in very few areas of health care. Accreditors and other groups are moving toward the common adoption of diabetes measures. The Washington Circle Group's measures for substance-related care are receiving support from a number of groups, including the U.S. Substance Abuse and Mental Health Services Administration and the National Committee for Quality Assurance.

There are many impediments to further progress. Especially important is that participating stakeholder groups often have conflicting priorities that limit well-intended efforts to produce consensus.¹⁰ For example, while many groups advocate that quality measures reflect evidence-based standards of care, such measures typically require rich data sources and are costly to implement. Organizations representing facilities and plans (which are responsible for collecting measurement data) express concern about this burden. In addition, even when an umbrella organization of stakeholders agrees on common measures, it may lack the authority to implement them.

In its most recent report, the Institute of Medicine made a persuasive call for the federal government to support consensus development through its multiple roles in health care.¹¹ In addition to acting as a lawmaker and regulator, the government provides health care directly through the Veterans Administration, the armed forces, and the Indian Health Service. Government programs reimburse providers for services through Medicare and, in partnership with the states, through Medicaid and the Children's Health Insurance Program. As an employer, the federal government is among the largest purchasers of health coverage in the country. Additionally, federal funds make up a large share of budgets for medical education, training, and research. Federal agencies charged with each of these areas have plentiful opportunities to foster consensus by encouraging participation in common umbrella activities, by adopting the measures and specifications that they produce, and by using the results of those measures results to guide purchasing and decisions.

AN ITERATIVE PROCESS OF TESTING AND IMPLEMENTATION

The evaluation and improvement of measure validity and reliability, the development of case-mix models, and the assessment of methods to avoid gaming and other undesired consequences of measurement are ongoing, iterative processes. Providers justifiably fear that premature implementation of imperfect measures might penalize them unfairly. But perfection can also serve as the enemy of the good. In order to conduct the necessary testing and refinement, we need hands-on experience applying measures in diverse settings and generating sizable data sets that include large populations spanning systems of care. In order to lessen resistance and foster consensus, groups implementing measures should begin by using the results to foster both internal quality improvement and collaborative benchmarking activities. Only after testing and refinement is completed should measures be linked to incentives or sanctions. In the longer term, we will need to learn not only which measurable areas offer the best prospects for improving care, but which are the most cost-effective areas for quality improvement, as defined in terms of units of clinical outcome per unit of resources expended. For example, if a local mental health care system has a limited budget for quality improvement, should it seek to improve assessment and treatment of substance abuse among patients with schizophrenia? Should it address discontinuities in care experienced by patients discharged from hospitals who do not successfully transition into outpatient care? Or should these funds be used to expand access to existing services rather than attempting to improve them? Methods of analyzing cost-effectiveness are well established for assessing clinical interventions.¹² Applying them to quality-improvement interventions will be feasible once data are available for a broad array of conditions, modalities, and common quality problems.

AN EXPANDED INFRASTRUCTURE FOR QUALITY ASSESSMENT

Much has been made of the promise of technology—e.g., electronic medical records and computerized systems for entering drug orders—to improve the quality and safety of care. While this potential is real, it may not be achieved quickly, particularly for smaller hospitals and independently practicing clinicians. In the meantime, there is a need to develop less glamorous, but nonetheless essential, infrastructure—such as a minimal data set of information on patient characteristics, treatment utilization, and outcomes that can be collected accurately and reliably. Much of this information is already routinely collected in the course of providing and paying for health care, but in the absence of uniform

standards, the information exists in diverse forms that resist collection, analysis, and comparative use. A standardized data set would build on standards established under the Health Insurance Portability and Accountability Act for diagnostic codes (ICD-9 and DSM-IV), procedures (CPT-IV), and drugs (NDC), but expand beyond them to include information that cannot be gleaned from administrative data sets. Illustrating the potential for this approach is the Center for Medicare and Medicaid Services' Minimal Data Set for skilled nursing facilities¹³ and its consumer-oriented Web site (<http://www.medicare.gov/nhcompare/home.asp>). Such activities are in keeping with the federal government's traditional role in setting standards with the goal of enabling markets to function.

CONCLUSION

Early investigations into the quality of health care resembled archeological digs: large-scale studies were conducted at substantial expense and complexity in order to answer individual questions—for example, the impact of prospective payment on inpatient care.^{14,15} In order to inform everyday decisions of participants in the health care marketplace, we need to function less like archeologists and more like meteorologists, drawing on data that are routinely collected by a matrix of sensors distributed throughout the health care system. Early stages of quality-measure development, consisting of local experimentation using diverse methods, were consistent with a New Federalism emphasis on decentralization and local control. These initiatives have produced many promising approaches; they have not, however, led to adequate information on quality. Within constraints imposed by the political context, federal action may be needed to bring quality measurement to maturity.

REFERENCES

1. Institute of Medicine. *Crossing the quality chasm: a new health system for the 21st century*. Washington, DC: National Academy Press, 2001.
2. Institute of Medicine. *Envisioning the National Health Care Quality Report*. Washington, DC: National Academy Press, 2001.
3. Hermann R. Risk adjustment for mental health care. In: Iezzoni L, ed. *Risk adjustment for measuring health care outcomes*. 3rd ed. Ann Arbor: Health Administration Press, in press.
4. Hermann RC, Leff HS, Palmer RH, et al. Quality measures for mental health care: results from a national inventory. *Med Care Res Rev* 2000;57(suppl 2):135–53.
5. Kondratas A, Weil A, Goldstein N. Assessing the new federalism: an introduction. *Health Aff (Millwood)* 1998;17(3):17–24.
6. Agency for Health Care Policy and Research Depression Guideline Panel. *Depression in primary care: Vol. 2. Treatment of major depression clinical practice guideline number 5*. Rockville, MD: U.S. Department of Health and Human Services, April 1993 (AHCPR pub. no. 93-0551).
7. Levin-Epstein M. Deathwatch for AHCPR ends as funding mandate changes. *Manag Care* 1999;8(9):14, 17.
8. Freudenheim M. Large H.M.O. to make treatment guidelines public. *New York Times*, 24 January 2003, p. 3C.
9. McCorry F, Garnick D, Bartlett J, Cotter F, Chalk M. *Improving performance measurement for alcohol and other drug services: report of the Washington Circle Group*. Rockville, MD: Washington Circle Group and the Center for Substance Abuse Treatment, 2000.
10. Hermann R, Palmer R. Common ground: a framework for selecting core quality measures. *Psychiatr Serv* 2002;53:281–7.
11. Corrigan J, Eden J, Smith B, eds. *Leadership by example: coordinating government roles in improving healthcare quality*. Washington, DC: Institute of Medicine, 2002.
12. Neumann P, Hermann R, Berenbaum P, Weinstein M. Methods of cost-effectiveness analysis in the assessment of new drugs for Alzheimer's disease. *Psychiatr Serv* 1997;48:1440–4.
13. Center for Medicare and Medicaid Services. *Minimum Data Set (MDS): Version 2.0 for nursing home resident assessment and care screenings* (September 2000).
14. Wells KB, Rogers WH, Davis LM, Kahn K, Norquist G, Keeler E, et al. Quality of care for hospitalized depressed elderly patients before and after the implementation of Medicare prospective payment system. *Am J Psychiatry* 1993;150:1799–805.
15. Kahn KL, Rubenstein LV, Draper D, Kosecoff J, Rogers WH, Keeler EB, et al. The effects of the DRG-based Prospective Payment System on the quality of care for hospitalized Medicare patients. *JAMA* 1990;264:1953–5.
16. Veterans Health Administration/U.S. Department of Defense. *VHA/DOD Performance Measures for the Management of Major Depressive Disorder in Adults, Version 2.0*. Washington, DC: Veterans Health Administration/Department of Defense, 2000.
17. Washington Circle Group. *Improving performance measurement by managed care plans for substance abuse: year 1 report of the Washington Circle Group*. Rockville, MD: Washington Circle Group, 1999.
18. American Psychiatric Association. *Report of the American Psychiatric Association Task Force on Quality Indicators*. Washington, DC: APA, 1999.
19. National Committee for Quality Assurance. *Health Plan Data and Information Set (HEDIS 2000)*. Washington, DC: NCQA, 1999.
20. New York State Office of Mental Health. *Report of the Performance Measurement/Management Workgroup Project*. Albany, NY: New York State Office of Mental Health, 1998.
21. Joint Commission on Accreditation of Healthcare Organizations. *National Library of Healthcare Indicators*. Oakbrook Terrace, IL: JCAHO, 1997.
22. Cuffel BJ, Fischer EP, Owen RR Jr, Smith GR Jr. An instrument for measurement of outcomes of care for schizophrenia: issues in development and implementation. *Eval Health Prof* 1997;20:96–108.
23. Eisen SV, Grob MC, Klein AA. BASIS: the development of a self-report measure for psychiatric inpatient evaluation. *Psychiatr Hosp* 1986;17:165–71.