



Published in final edited form as:

Can J Psychiatry. 2010 September ; 55(9): 549–557.

Challenges and Opportunities in Measuring the Quality of Mental Health Care

Amy M. Kilbourne, PhD¹, Donna Keyser, PhD², and Harold Alan Pincus, MD³

¹Associate Director, VA Ann Arbor National Serious Mental Illness Treatment Research and Evaluation Center; Investigator, HSR&D Center of Excellence, and Associate Professor of Psychiatry, University of Michigan Medical School, Ann Arbor, MI

²Senior Research Scientist, RAND Corporation, Pittsburgh, PA; Professor and Vice Chair, Department of Psychiatry and Director of Quality and Outcomes Research at New York Presbyterian Hospital, New York, NY

²Management Scientist, RAND Corporation and Associate Director, RAND-University of Pittsburgh Health Institute, Pittsburgh, PA

Abstract

The purpose of this paper is to delineate the barriers to mental health quality measurement, and identify strategies to enhance the development and use of quality measures by mental health providers, programs, payers, and other stakeholders in the service of improving outcomes for individuals with mental health and substance use disorders. Key reasons for the lag in mental health performance measurement include lack of sufficient evidence regarding appropriate mental health care, poorly defined quality measures, limited descriptions of mental health services from existing clinical data, and lack of linked electronic health information. We discuss strategies for overcoming these barriers that are being implemented in several countries, including the need to have quality improvement as part of standard clinical training curricula, refinement of technologies to promote adequate data capture of mental health services, use of incentives to promote provider accountability for improving care, and the need for mental health researchers to improve the evidence base for mental health treatment.

Keywords

mental disorders; quality of care; quality improvement; information systems; performance measurement

Background

Mental disorders (including serious mental illness, such as schizophrenia or bipolar disorder, depression, anxiety disorders, and substance use disorders) are leading causes of disability worldwide, are associated with substantial costs, and when left untreated, can lead to premature mortality.¹ Despite the proliferation of evidence-based guidelines for the treatment of mental disorders, the quality of care and subsequent outcomes remain suboptimal for those suffering from these illnesses.² A recent report by the U.S. Institute of Medicine (IOM) and others documented substantial gaps in evidence-based care for mental

disorders, citing poor quality in detection, treatment, and follow-up care.² The stigma of a mental disorder diagnosis may also hinder many from seeking care in the first place.

Within the past ten years, there has been a dramatic growth in the development of performance (or quality) measures to assess and redress gaps in evidence-based health care in general. Experts have recognized that quality measurement is a key driver in transforming the health care system, and routinely measuring quality using performance measures derived from evidence-based practice guidelines is an important step to this end.³ Notably, national and provincial governments as well as regulatory, accreditation and other non-governmental organizations around the globe have proposed and implemented performance measures to be used by different health plans and organizations for a broad range of services and conditions. In the U.S., these organizations have included the National Committee on Quality Assurance (NCQA) and the National Quality Forum (NQF). Performance measures have been increasingly used in health care to compare and benchmark processes of care in order to remediate gaps between evidence-based care and actual practice, and hold providers accountable for improving quality of care. Recently, in the U.S. and United Kingdom (UK), for example, pay-for-performance (P4P) initiatives and other financial incentives have been implemented as a means to promote quality improvement.⁴⁻⁵

Nonetheless, there is documented evidence (i.e., IOM) that worldwide, the mental health services sector lags behind in the development and implementation of performance measures and strategies for implementing them as tools to improve quality and outcomes.² There are three key reasons for this lag: lack of a sufficient evidence base through which to develop specific, valid and clearly defined measures,² inadequate infrastructure to develop and implement quality measures and capture elements of mental health services, and lack of a cohesive strategy to apply mental health quality measurement across different settings in the service of improving care. Measuring quality of mental health care is particularly challenging because a substantial amount of mental health services are delivered outside the health care sector (e.g., criminal justice, education, social services), and there is insufficient evidence for some mental health treatments as well.² Data elements necessary to measure quality of mental health care are incomplete or even missing in many settings, and, even when data collection does occur, it tends to be inconsistent across different organizations. Moreover, mental health programs and providers have not fully embraced quality measurement due to infrastructure and policy barriers intrinsic to mental health, including providers' concerns regarding patient privacy, "cookbook medicine," and the silos across different provider types and credentialing requirements. Finally, in many countries, the mental health sector is far behind the rest of health care in the use of health information technology.

The purpose of this paper is to delineate the barriers to mental health quality measurement, and identify strategies to enhance the development and use of quality measures by mental health providers, programs, payers, and other stakeholders in the service of improving outcomes for individuals with mental health and substance use disorders.

Quality Measurement in Healthcare

Understanding the challenges of measuring quality of mental health care requires an understanding of the history of measuring quality. Over 30 years ago, Donabedian published a quality framework that is much used today.⁶ This framework incorporates three domains of quality measurement: structure, process, and outcomes. The rationale underlying the framework is that health care structures, including resources and policies, can inform processes of care provided by clinicians, which in turn can influence patient outcomes. The framework also provides a useful typology for distinguishing among different approaches to measurement. Specifically, health care structure measures evaluate characteristics of the

treatment setting's services, including program fidelity, staffing, and infrastructure (i.e. are quality services available?). Process measures examine interactions between consumers and the structural elements of the health care system (i.e. are consumers actually receiving high quality services in a way that conforms to the evidence base?). Outcome measures examine the results of these interactions for patients, including functioning, morbidity, mortality, quality of life,⁷ and patient satisfaction (i.e. is the care making a difference for individuals and society?).

Each type of quality measure has its strengths and limitations. Structure measures are relatively simple to ascertain through reports from program or clinic leaders, yet are subject to response bias (e.g., program leader tendency to over-report resources or over-idealize clinic operations) or vagueness in terminology or questions (e.g., defining program fidelity). Moreover, structure measures do not indicate whether good care happened, but only if the site has capacity to provide good care. Process measures have been widely used for performance measurement, and are more appealing to providers because they represent aspects of care over which providers have the most control. However, there are concerns that many process measures are overly dependent on a patient's care-seeking behavior, which is often not reported in data sources (e.g., were outpatient visits not completed because they were not scheduled by the provider, or did the patient miss them?). Alternatively, outcome measures are appealing because they actually assess whether a patient's status improves or not. However, use of these measures requires additional case mix adjustment in order to ensure that observed differences in outcomes are not due to clinical differences in severity of illness across patients. Currently, methods to adjust for case mix (risk adjustment) in mental health are limited, primarily due to incomplete data on psychiatric symptoms and other patient characteristics. Use of outcome measures to incentivize providers can also be problematic, as there is concern that providers may be reluctant to take on sicker patients in order to make patients panels look better overall.

In light of the strengths and limitations, some regulatory agencies have implemented criteria for evaluating and selecting quality measures, notably, the U.S. National Quality forum and National Committee on Quality Assurance. In the early 1990s, organizations such as the NCQA began to operationalize evidence-based care by establishing quality measures, notably process measures, through the Health Experience Data and Information Set (HEDIS), and hence, had to develop standards of performance measurement. These standards have their origins in the criteria developed by the RAND Corporation,⁸ which have been widely used to select and refine quality indicators, provide a good start for identifying and refining performance measures. These criteria include: 1) clinical importance (i.e., Does the measure represent a substantial deficit in mental health care? Can it result in actionable improvements in mental health care?), 2) validity (i.e., Is the measure scientifically sound? Is it associated with improved clinical care? Is it sensitive to risk adjustment and insensitive to being gamed?), and 3) feasibility (i.e., How easy is it to collect the information required for the quality measure?). The RAND method also recommends a panel of experts (usually 9) to rate each indicator based on these criteria using a scale (1–5), and dropping indicators that fall below a certain threshold (e.g., below 5).

The number of organizations using quality indicators to benchmark care across different practices has grown substantially since the development of quality indicators, with several countries are taking the lead to develop and implement mental health performance measures as part of their nationalized health systems.^{9–12} In a systematic review of governmental and non-governmental organizations in the U.S., Herbstman (2009) found 36 initiatives that applied quality indicators for mental disorders, ranging from federal and state governments, health plans, and non-governmental and professional organizations.¹³ Currently, one of the most extensive implementations of mental health quality measures is the International

Initiative for Mental Health Leadership Clinical project on mental health quality measurement, described in detail in this issue.

In the U.S., an important example of quality measurement in mental health is the work of the Department of Veterans Affairs (VA) Veterans Health Administration. The VA provides a useful comparator organization in the international sphere given that it is single-payer, national health care program for U.S. veterans.¹⁴ Currently, the Altarum Institute and the RAND Corporation are conducting a national evaluation of the VA's mental health services, using Donabedian's structure-process-outcomes model of quality measurement. Structure is evaluated by describing the continuum of care available to veterans with five targeted mental health diagnoses (i.e., schizophrenia, bipolar disorder, post-traumatic stress disorder [PTSD], major depressive disorder, and substance use disorder) using facility survey and administrative data. Process is evaluated by looking at what services veterans with the five targeted mental health diagnoses actually receive using administrative and medical record data. Outcomes measures consist of services received that made a difference to veterans' functioning and quality of life using client survey and medical record data. As part of this work, the Altarum-RAND team has developed and vetted over 200 quality indicators for the five targeted mental disorders. In addition, the VA Primary Care-Mental Health integration program has implemented performance measurement for depression, alcohol, and PTSD screening and management.¹⁵

Applying Quality Measures

Quality measures are vital for ensuring the uptake and delivery of evidence-based care. Quality of care cannot be improved without monitoring how such care is delivered, from its organization to processes and ultimately patient outcomes. Measuring quality using quality indicators derived from evidence-based practice guidelines is an important step towards implementation of evidence-based care and monitoring quality improvement efforts. Experts have strongly suggested that quality measurement is a key driver in transforming the health care system and clinicians at the front line need to be actively engaged in the process of improving quality at multiple levels (e.g., payment, licensure, organizational change), and payment systems should reward high quality of care). Standardized measures that allow results to be monitored and tracked uniformly over time are the foundation of performance improvement.

Moreover, the public reporting of performance is vital for holding health care organizations accountable for improving care. Despite the plethora of quality measures, some health care payers and organizations have realized that they would be ignored unless providers were held accountable for improving care. For example, in Scotland, there are national efforts to measure quality of mental health care through the use of "National Targets" on readmission rates, antidepressant prescribing, suicide prevention, and dementia screening, as well as benchmarking programs to compare healthcare providers on performance measures.¹² In the U.S., the National Quality Forum applies quality indicators to rate different private health plans.

Quality measures are also used to evaluate whether differences in care exist across regions or healthcare providers, and in the U.S. and U.K, they can also be used to reward or reprimand providers financially. Some countries have gone a step further and established financial incentives, such as pay-for-performance (P4P), for achieving standards based on quality measures.⁵ P4P involves financial incentives to providers or health care systems when certain target quality of care measures are achieved within providers' patient panels. Currently, a growing number of financial incentives programs focus on behavioral health in the U.S.¹⁶ and in England.⁵ In particular, England established P4P for general practitioners in 2004, comprising of several chronic disease composite indicators, including mental

health, notably the “percentage of patients with severe long-term mental health problems reviewed in the preceding 15 months, including a check on the accuracy of prescribed medication, a review of physical health, and a review of coordination arrangements with secondary care”.⁵

However, the use of performance-based financial incentives such as P4P has been criticized for failing to adequately address patient complexities. In mental health, current risk adjustment methods cannot fully account for these complexities, as these patients tend to be poorer and sicker. “Cherry picking” is a concern, as the sickest patients can be dropped from a panel in order to improve performance measures. Still, while this has been a concern with physical health performance measures, in England, there has been little evidence of cherry-picking (i.e., less than 1% of practices excluded large numbers of payments from the program who might not have scored well on the indicators).⁵

The cost of implementing P4P is also an issue. In the U.S., evidence suggests that financial incentives for improved performance need to be 10–15% of an individual provider’s salary to induce practice change.¹⁷ In some cases, P4P implementation has led to rising costs¹⁸ with little improvement in quality. In contrast to typical U.S. P4P programs, the UK’s program paid practitioners up to 25% of their salary, leading to rapidly rising healthcare costs. Hence, without a means to restructure existing payment systems, it is unclear whether P4P can be sustained over time.

Overall, while P4P holds some promise for incentivizing providers to achieve performance benchmarks, especially for single-payer health systems, cost is still a concern, and the metrics used to incentivize providers need to clearly represent evidence-based care and adequately account for patient case-mix. For P4P to be successfully implemented and accepted by mental health providers, the indicators that are used to benchmark care need to be refined and accepted by providers and other stakeholders. Moreover, for mental health services, improving the quality measures themselves is a vital step towards improving performance-based measurement and gaining widespread acceptance of programs designed to change provider behavior.

Limitations of Mental Health Quality Measures

The overall development and implementation of valid quality indicators and their application to promote accountability of care for mental disorders has not kept pace with physical health care. While several measures have been developed for mental disorders across different organizations, it is uncertain whether they have been validated, or at least vetted and rated by experts. Notably, only 16 of almost 600 NQF measures and only a fraction of U.S. HEDIS measures address mental health, and those that do are limited by too narrow a focus on clinical activity or a particular diagnosis, failing to encourage coordination, and not being linked to quality improvement activities.¹⁹ Important reasons for the lag in mental health performance measurement include:

1. Lack of sufficient evidence regarding appropriate care. Mental health, unlike general health sectors, lacks sufficient evidence for many conditions. For example, there is limited information on evidence-based care for adolescents, including use of atypical antipsychotics²⁰ or the use of combinations of antipsychotics in adults.²¹ The guidelines are also inconclusive in regards to the use of antidepressants in bipolar spectrum disorders, and medication treatment for PTSD.²²
2. Poorly defined parameters (e.g., definitions of treatment “engagement,” “psychotherapy,” “recovery,” or intervention “fidelity” are often not well-described)²³. Existing data sources often lack sufficient information to establish

accurate numerators and denominators for quality indicators (e.g., ICD-9 codes for mental health diagnoses), and some quality measures use strict exclusion criteria that do not apply to the majority of patients at risk (e.g., indicators for newly diagnosed depression only).

3. Limited descriptions of mental health care based on available data sources preclude their utility in routine care. Data on appropriate pharmacotherapy and psychotherapy are only available when pulled from multiple sources (e.g., billing claims, outpatient, and pharmacy records). Most performance measures are based on medical chart review or electronic claims data. Manual chart review is slow, labor intensive, expensive, constrained to small samples, and prone to inter-reviewer variation. Administrative claims data, while appealing because the sample size is larger and data pulls can be done electronically, are less meaningful given the wide variability in diagnostic coding of mental health conditions and lack of complete data on types of visits or treatments (e.g., specific treatment modalities). Administrative data may also be incomplete within mental health programs, as many lack access to data from general medical providers, labs, or prescriptions. There are also multiple definitions used to define specific performance measures (e.g., post-hospitalization follow-up: number of days), resulting in growing burden and confusion over data capture and reporting processes as providers try to implement quality measurement program requirements.
4. Lack of electronic health information to facilitate capture of the required health information. In a recent study, only 12% of U.S. hospitals had electronic medical records (EMRs).²⁴ In other countries such as Canada, national electronic medical records exist²⁵ but may lack full information on services needs of individuals with mental disorders.²⁶ Moreover, a substantial amount of mental health services are delivered outside the health care sector (e.g., criminal justice, social services), making integration of electronic information across sites challenging.

Improving Quality: Towards a Culture of Measurement-Based Care in Mental Health

To address these barriers to mental health quality measurement, the U.S. Institute of Medicine (IOM) described an approach for developing, testing, and validating mental health quality measures (2). Seven stages were identified: 1) improved conceptualization of what mental health services or outcomes are to be measured (either through clarity in guidelines or additional empirical studies), 2) operationalizing guidelines into quality measures, 3) pilot- testing the performance measure in routine care to assess validity, reliability, feasibility, and cost (including provider and auditor burden), 4) proposing the validated performance measure to an organization responsible for quality measurement and benchmarking, 5) further evaluation of performance measure implementation to ensure that it is performing as intended, 6) calculating performance based on the measure and summarizing results in a salient way for different stakeholders and audiences, and 7) maintaining the effectiveness and integrity of the performance measure over time (e.g., monitor for gaming, ceiling effects).²

This approach for measure development is an essential component of a larger framework to improve the quality of care that also requires active participation of providers, consumers, payers, and policy makers (Figure 1). To overcome challenges in measuring mental health care quality and achieve the stages outlined in this framework, we propose three strategies: 1) mental health quality measures themselves need to be refined and further validated, 2) information technology in mental health care needs to be improved to facilitate not only electronic data capture but improved performance tracking, and 3) providers, consumers,

professional organizations, health care organizations, and other stakeholders need to be more involved in the quality measurement, implementation and improvement process.

Improve Mental Health Measures over Time

To improve mental health quality measurement, the measures themselves need to be refined, particularly for processes and outcomes of care using the aforementioned criteria. These measures should also be vetted by multiple stakeholders who have a vested interest in improving mental health care (e.g., payers, consumers, providers, policymakers) in order to maximize the chances of their acceptance and use. Consumers can be a powerful force in the refinement and implementation of performance measures, and involving them in the decision-making process can facilitate consumer acceptance of quality measurement. Providers also need additional information on quality measurement, especially within the context of evidence-based care; and they need to be actively engaged in the process of improving quality at multiple levels (e.g., payment, licensure, organization/information technology). Evidence suggests that only a fraction of social workers receive training in evidence based care, and little if any on quality improvement.²⁷ In the past, community-based mental health providers have rarely been given the opportunity to provide substantive input on the development and implementation of mental health quality indicators in general. As the ultimate “end-users” of quality indicators, frontline providers have important insight into their clinical meaningfulness and feasibility for application to routine patient care. Hence, their acceptance of indicators is crucial if they are to be applied to monitor uptake of evidence-based care, for policy-level incentives that reward high quality of care such as P4P, as well as for sustaining clinical quality improvement over time.

In addition, a balanced portfolio of structure, process, and outcomes measures should be used in mental health. Such a strategy is useful for several reasons. First, it prevents organizations from “gaming,” by considering the structures in place to support evidence-based care along with patient outcome and clinical processes. Second, process measures are most subject to the provider control. Finally, outcomes can provide valuable information on quality of care, but require sophisticated risk adjustment approaches that are still in development.

Efforts to achieve consensus on a core set of quality measures that are meaningful and feasible to multiple stakeholders, as well as broadly representative of the mental health care system are beginning to show some promise. In 2004, Hermann and colleagues identified 28 measures for assessing mental health treatment that ranged from access, assessment of mental health symptoms, continuity of care, coordination, prevention, and safety.²⁸ These measures were selected from among over a hundred previously identified mental health care process measures by a 12-member panel of stakeholders from national organizations using a 2-stage modified Delphi consensus development process. Panelists rated each measure on 7 domains using a 9-point scale (1=best). Measures were then mapped to a framework of system dimensions to identify a core set with the highest ratings for system characteristics within each dimension. Overall, mean ratings for meaningfulness were: clinical importance (2.29); perceived gap between actual and optimal care (2.59); and association between improved performance and outcome (2.61). For feasibility, mean ratings were clarity of specifications (3.39); acceptability of data collection burden (4.77); and adequacy of case mix adjustment (4.20). About half of the indicators lacked supporting research evidence; many were rated infeasible, primarily due to lack of risk adjustment (case mix adjustment) data; and some were likely to achieve a ceiling effect (e.g., antipsychotic prescription for schizophrenia, >1 mental health visit in 12 months, assessment of alcohol or drug abuse). Based on these findings, 12 of the 28 measures, including 10 that could be constructed using administrative data and two that could be dichotomized to evaluate mental disorders and

substance use disorders separately, were then used to develop statistical benchmarks for quality of care for mental and substance use disorders.²⁸

Related efforts of an international expert panel to select indicators for the quality of mental health care at the health systems level in countries involved in the Organization for Economic Cooperation and Development (OECD) have resulted in 12 indicators, out of candidate set of 134, that cover four key areas of treatment (Table 1).²⁹ The measures were selected based on a structured review process that included considerations of scientific soundness, policy importance, and feasibility to construct from preexisting data. More recently, Hermann and colleagues applied a method for developing statistical benchmarks for 12 of the mental health measures identified through the process described above.³⁰

These efforts notwithstanding, further refinement of mental health performance measures is needed. Treatment-based process measures that reflect coordination or continuity of care rather than disease-specific measures are preferred, especially given the instability in mental health diagnostic data from administrative sources. Continuity of care is often used to benchmark high-quality mental health services. A Canadian study found a strong association between measures of continuity of care and outcomes among persons with serious mental illness.³¹ Other examples of measures include receipt of lab monitoring for mood stabilizers or monitoring of cardio-metabolic disease risk factors among users of second-generation antipsychotics, and timely follow up after hospitalization. These measures not only represent continuity of care, but also shared accountability between mental and general health providers. Another method is to construct so-called “tightly linked” measures, which assess desired outcomes and processes of care simultaneously. Kerr and Hayward (2003) developed a tightly-linked measure, which was defined as, among patients with diabetes, achieving an LDL level of <100 or whether there is documentation in the medical record that the provider attempted to assist the patient in lowering cholesterol (e.g., statin prescription, recommendation of lower cholesterol diet, etc).³² This indicator was associated with improved outcomes among people with diabetes. The challenge in implementing these tightly-linked measures is the need for manual implicit chart review, which can be costly and has the potential for low reliability.

Outcome measures should also be refined by making them applicable to a wide range of patients, sensitive to change, and focused on functioning, symptoms, as well as other issues including suicidal ideation and substance use. One method to implement outcomes measures when risk adjustment data are limited is to assess incremental improvements in outcomes rather than absolute value. This process has been used in some U.S. educational systems to rate students’ performance over time rather than absolute test scores in order to evaluate teachers.³³ Given more advanced statistical methods available, the same approach could be used for health care providers. Monitoring outcomes can also help providers make better decisions regarding medication choice (e.g., switching antidepressants if no response to a symptom assessment), through cueing them on patient progress and facilitating treatment collaboration. In the U.S., the National Quality Forum has recently established a Mental Health Patient Outcomes project to identify, evaluate and endorse measures that address mental health outcomes- defined as changes (desired or undesired) in individuals or populations.

Improve Mental Health Information Technology

The current U.S. Administration is also supporting incentives to facilitate the implementation of electronic health records as a step towards improving performance measurement. Enhancing the uptake of performance measures not only requires improved information technology, but a refinement of data definitions to facilitate electronic data capture of mental health treatment. For example, there have been recent initiatives to refine

coded medical data (i.e. ICD 10/11 and DSM 5) to provide more clinically precise information, particularly for mental health treatment. In recognition of this need, multiple stakeholders in the U.S and around the world have begun to engage in standardization activity around mental health quality measurement, particularly to define an “ontology” to better classify what information is needed and required to develop clinically meaningful quality indicators. In computer science, an ontology is a rigorous organization of a knowledge domain (e.g., evidence-based mental health care) that is usually hierarchical and contains all the relevant entities, their definitions, and information sources. Measurement ontologies are increasingly being used, notably in ICD-11, to facilitate the uptake of quality measurement across practices with varying degrees of information technology implementation. For example, WHO’s ICD 11 process is applying an overall informatics framework that incorporates quality and patient safety concepts.

Involving Stakeholders: Sustaining Mental Health Measurement-based Care

We are not interested in measurement for measurement sake. Ultimately, we want measures to be used by multiple stakeholders at multiple levels to actually improve care. In order to achieve this goal, stakeholders such as health care providers, consumers payers, and policymakers must be involved from the start. Overall, these proposed improvements in the quality measurement process can facilitate a culture of “measurement-based care”, which is defined as enhanced precision and consistent use in disease assessment, tracking, and treatment to achieve optimal outcomes. Measurement-based care can not only improve documentation and subsequent reimbursement of services, but can also help patients become more aware of changes in their mental health symptoms, along with medication side effects to improve shared decision-making regarding treatment choice.

Adopting a culture of measurement-based care requires that stakeholders work together to develop and implement valid outcome measures to foster accountability to ensure practice change. Two examples of this process are described below. First, in the U.S., the Depression Improvement Across Minnesota, Offering a New Direction (DIAMOND), encourages the implementation of collaborative care models for depression in medical groups and health plans across the state by rewarding providers for using information technology to track patient progress through registries. In addition, the VA Primary Care Mental Health Integration program, which has implemented a set of outcome measures related to the most common mental health conditions seen in primary care, including depressive symptoms (PHQ-9), alcohol misuse (AUDIT-C), risk assessment for suicidality, and PTSD symptoms, and these assessments are now part of the VA’s computerized medical record system. Tools based on positive screens are also embedded in the VA’s computerized medical record system, and performance measures are in place to track whether providers are following up on depression screens and providing patient-specific mental health care management.

Conclusion

Together, the refinement of quality measures, investment in information technology, and fostering a culture of measurement-based care can enhance the quality and ultimately, outcomes of mental health services. Sustaining this effort will require a rethinking of how quality measurement is used to promote the uptake of evidence-based mental health care across systems of care that extend to the national (and international level). Methods for quality improvement should be taught and accountability maintained across different mental health provider groups, including psychiatrists, social workers, nurses, and other health professionals. Because the de facto health care setting for persons with mental disorders can encompass other settings outside of mental health (e.g., from primary care to criminal justice), measurement systems should cut across mental, physical, and substance use disorders. For example, both primary care and mental health providers should be held

accountable for performance measures that cut across co-occurring conditions (e.g., metabolic management from side effects of atypical antipsychotics). Ultimately, a measurement-based care culture should lead to greater accountability for mental health services regardless of where the consumer enters the system. In addition, we need to improve strategies to foster accountability in performance improvement. For example, performance incentives such as pay-for-performance should be given at the group and not individual provider level, and should reward based on incremental changes rather than attaining absolute benchmarks. This would reduce costs of performance incentives and maximize the potential for addressing system-level deficiencies in care. Finally, mental health researchers need to improve the evidence base for mental health treatment, so that quality measures are further refined for their ultimate application and acceptance as powerful tools to close the quality chasm in mental health.

Acknowledgments

Funding and Acknowledgements

This research was supported by the Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service (IIR 02-283-2; A. Kilbourne, PI) and the National Institutes of Health (R34 MH74509; A. Kilbourne, PI). Funding was also received by the Irving Institute for Clinical and Translational Research at Columbia University (UL1 RR024156) and University of Pittsburgh's Clinical and Translational Science Institute (UL1 RR024153) both from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH), and the Mental Health Therapeutics CERT at Rutgers, the State University of New Jersey funded by the Agency for Healthcare Research and Quality (AHRQ) (5 U18 HS016097). Additional funding came from government and non-government organizations of the countries participating in the IIMHL project (Australia, Canada, Denmark, England, Germany, Ireland, Japan, Netherlands, New Zealand, Norway, Scotland, Taiwan, and the US).

The funding source had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs.

References

1. Lopez AD, Murray CC. The global burden of disease, 1990–2020. *Nat Med.* 1998; 4:1241–1243. [PubMed: 9809543]
2. Institute of Medicine. Washington D.C: National Academy Press; 2006. *Improving Quality of Health Care for Mental and Substance Use Conditions.*
3. Conway PH, Clancy C. Transformation of health care at the front line. *JAMA.* 2009; 301:763–765. [PubMed: 19224753]
4. Rosenthal MB, Landon BE, et al. Pay for performance in commercial HMOs. *N Engl J Med.* 2006; 355:1895–1902. [PubMed: 17079763]
5. Doran T, Fullwood C, Reeves D, et al. Exclusion of patients from pay-for-performance targets by English physicians. *N Engl J Med.* 2008; 359:274–284. [PubMed: 18635432]
6. Donabedian, A. *Explorations in Quality Assessment and Monitoring: The Definition of Quality and Approaches to its Assessment.* Ann Arbor, MI: Health Administration Press; 1980.
7. Lim KL, Jacobs P, Ohinmaa A, et al. A new population-based measure of the economic burden of mental illness in Canada. *Chronic Dis Can.* 2008; 28:92–98. [PubMed: 18341763]
8. Brook, RH. The RAND/UCLA Appropriateness Method. In: McCormack, KA.; Moore, ST.; Siegel, RA., editors. *Methodology Perspectives: US Department of Health Services.* p. 59-70.
9. Pinus HA, Naberb D. International efforts to measure and improve the quality of mental healthcare. *Current Opinion in Psychiatry.* 2009; 22:609. [PubMed: 19773655]
10. Gabecl W, Janssen B, Zielasek J. Current Mental health quality, outcome measurement, and improvement in Germany. *Opinion in Psychiatry.* 2009; 22:636–642.
11. Brown P, Pirkis J. Mental health quality and outcome measurement and improvement in Australia. *Current Opinion in Psychiatry.* 2009; 22:610–618. [PubMed: 19730372]

12. Coia D, Glassborow R. Mental health quality and outcome measurement and improvement in Scotland. *Current Opinion in Psychiatry*. 2009; 22:643–647. [PubMed: 19773656]
13. Herbstman B, Pincus HA. Measuring mental healthcare quality in the United States: a review of initiatives. *Current Opinion in Psychiatry*. 2009; 22:623–630. [PubMed: 19745741]
14. Asch SM, McGlynn EA, Hogan MM, et al. Comparison of quality of care for patients in the Veterans Health Administration and patients in a national sample. *Ann Intern Med*. 2004; 141:938–945. [PubMed: 15611491]
15. Post EP, Van Stone WW. Veterans Health Administration primary care-mental health integration initiative. *N C Med J*. 2008; 69:49–52. [PubMed: 18429567]
16. Bremer RW, Scholle SH, Keyser D, et al. Pay for performance in behavioral health. *Psychiatr Serv*. 2008; 59:1419–1429. [PubMed: 19033169]
17. Hillman AL, Ripley K, Goldfarb N, et al. The use of physician financial incentives and feedback to improve pediatric preventive care in Medicaid managed care. *Pediatrics*. 1999; 104:931–935. [PubMed: 10506237]
18. Robinson JC, Williams T, Yanagihara D. Measurement of and Reward for Efficiency In California's Pay-For-Performance Program. *Health Aff*. 2009; 28:1438–1447.
19. Kilbourne AM, Fullerton C, Dausey D, et al. A Framework for Measuring Quality and Promoting Accountability across Silos: The Case of Mental Disorders and Co-occurring Conditions. *Quality and Safety in Health Care*. 2010 in press.
20. Zito JM, Safer DJ, Berg LT, et al. A three-country comparison of psychotropic medication prevalence in youth. *Child Adolesc Psychiatry Ment Health*. 2008; 2:26. [PubMed: 18817536]
21. Miller AL, Craig CS. Combination antipsychotics: pros, cons, and questions. *Schizophr Bull*. 2002; 28:105–109. [PubMed: 12047009]
22. Nierenberg AA. Lessons from STEP-BD for the treatment of bipolar depression. *Depress Anxiety*. 2009; 26(2):106–109. [PubMed: 19177467]
23. First MB, Pincus HA, Schoenbaum M. Issues for DSM-V: adding problem codes to facilitate assessment of quality of care. *Am J Psychiatry*. 2009; 166:11–13. [PubMed: 19122012]
24. Jha AK, Desroches CM, Campbell EG, et al. Use of Electronic Health Records in U.S. Hospitals. *N Engl J Med*. 2009; 360:1–11. [PubMed: 19118299]
25. Health Canada: Electronic Health Information. Available at: http://www.hc-sc.gc.ca/hcs-sss/pubs/kdec/nf_eval/nf_eval3-eng.php
26. Booth RH. Using electronic patient records in mental health care to capture housing and homelessness information of psychiatric consumers. *Issues Ment Health Nurs*. 2006; 27:1067–1077. [PubMed: 17050339]
27. Insel TR. Translating scientific opportunity into public health impact: a strategic plan for research on mental illness. *Arch Gen Psychiatry*. 2009; 66:128–133. [PubMed: 19188534]
28. Hermann RC, Palmer H, Leff S, et al. Achieving consensus across diverse stakeholders on quality measures for mental healthcare. *Med Care*. 2004; 42:1246–1253. [PubMed: 15550805]
29. Hermann RC, Mattke S, the Members of the OECD Mental Health Care Panel. Selecting indicators for the quality of mental health care at the health systems level in OECD countries. *OECD Health Technical Papers No. 17*. 2004 October 28.
30. Hermann RC, Chan JA, Provost SE, et al. Statistical benchmarks for process measures of quality of care for mental and substance use disorders. *Psychiatr Serv*. 2006; 57:1461–1467. [PubMed: 17035566]
31. Adair CE, McDougall GM, Mitton CR, et al. Continuity of care and health outcomes among persons with severe mental illness. *Psychiatr Serv*. 2005; 56:1061–1069. [PubMed: 16148318]
32. Kerr EA, Smith DM, Hogan MM, et al. Building a better quality measure: are some patients with 'poor quality' actually getting good care? *Med Care*. 2003; 41:1173–1182. [PubMed: 14515113]
33. Felch J, Song J. Superintendent spreads the gospel of 'value-added' teacher evaluations. *Los Angeles Times*. Available at: www.latimes.com/news/local/la-me-teacher-eval18-2009oct18,0,4471467.story.

Table 1

Organization for Economic Cooperation and Development (OECD) Indicators for Mental Health Care

Area	Indicator Name
Continuity of Care	<ul style="list-style-type: none"> ▪ Timely ambulatory follow-up after mental health hospitalization ▪ Continuity of visits after hospitalization for dual psychiatric/substance related conditions ▪ Racial/ethnic disparities in mental health follow-up rates ▪ Continuity of visits after mental health-related hospitalization
Coordination of Care	<ul style="list-style-type: none"> ▪ Case management for severe psychiatric disorders
Treatment	<ul style="list-style-type: none"> ▪ Visits during acute phase treatment of depression ▪ Hospital readmissions for psychiatric patients ▪ Length of treatment for substance-related disorders ▪ Use of anti-cholinergic anti-depressant drugs among elderly patients ▪ Continuous anti-depressant medication treatment in acute phase ▪ Continuous anti-depressant medication treatment in continuation phase
Outcomes	<ul style="list-style-type: none"> ▪ Mortality for persons with severe psychiatric disorders ▪ Quality of life