

Quality indicators for international benchmarking of mental health care

RICHARD C. HERMANN^{1,2}, SOEREN MATTKE³, DAVID SOMEKH⁴, HELENA SILFVERHJELM⁵, ELLIOT GOLDNER⁶, GYLES GLOVER⁷, JANE PIRKIS⁸, JAN MAINZ⁹ AND JEFFREY A. CHAN¹

¹Tufts-New England Medical Center, Center for Quality Assessment & Improvement in Mental Health (CQAIMH), ²Departments of Psychiatry and Medicine, Tufts University School of Medicine, Boston, MA, USA, ³Organisation for Economic Co-operation and Development, Health Policy Unit, Paris, France, ⁴Association of Quality in Health, London, UK, ⁵National Board of Health and Welfare, Stockholm, Sweden, ⁶University of British Columbia, Vancouver, Canada, ⁷University of Durham, Durham, UK, ⁸University of Melbourne School of Population Health, Program Evaluation Unit, Melbourne, Australia, and ⁹The National Indicator Project, Aarhus, Denmark

Abstract

Objective. To identify quality measures for international benchmarking of mental health care that assess important processes and outcomes of care, are scientifically sound, and are feasible to construct from preexisting data.

Design. An international expert panel employed a consensus development process to select important, sound, and feasible measures based on a framework that balances these priorities with the additional goal of assessing the breadth of mental health care across key dimensions.

Participants. Six countries and one international organization nominated seven panelists consisting of mental health administrators, clinicians, and services researchers with expertise in quality of care, epidemiology, public health, and public policy.

Measures. Measures with a final median score of at least 7.0 for both importance and soundness, and data availability rated as 'possible' or better in at least half of participating countries, were included in the final set. Measures with median scores ≤ 3.0 or data availability rated as 'unlikely' were excluded. Measures with intermediate scores were subject to further discussion by the panel, leading to their adoption or rejection on a case-by-case basis.

Results. From an initial set of 134 candidate measures, the panel identified 12 measures that achieved moderate to high scores on desired attributes.

Conclusions. Although limited, the proposed measure set provides a starting point for international benchmarking of mental health care. It addresses known quality problems and achieves some breadth across diverse dimensions of mental health care.

Keywords: benchmarking, consensus development, international, mental health, quality measures

Background

Mental health and substance-related disorders are prevalent, disabling, and costly. An estimated 450 million people worldwide are affected by mental, neurological, or behavioral problems at any given time, and studies indicate that approximately one in five will experience a psychiatric disorder within a given year [1]. A World Health Organization (WHO) study found these conditions to account for almost 11% of the global burden of disease in 1990. Among the 10 leading causes of disability worldwide, 5 are psychiatric conditions—depression, bipolar disorder, schizophrenia, obsessive-compulsive disorder, and alcohol abuse [2].

The percentage of annual health care expenditures spent on mental health care varies widely, exceeding 20% among

the highest spending nations [1]. Furthermore, mental health and substance-related disorders incur large indirect costs in utilization of other medical services, in lost work productivity, and in the burden on families and other caregivers. Effective medication and psychosocial treatments exist for many mental disorders [3]. However, research studies have documented wide variations in the quality of care including gaps between clinical practice and evidence-based guideline recommendations [4]. These findings have led to widespread attention to improving the quality of mental health care [5–9].

In this issue of the journal, Mattke *et al.* [10] describe the potential utility of measurement-based quality improvement and international comparisons of measure results. Data on key processes and outcomes of care can facilitate improvement

Address reprint requests to Richard C. Hermann, Tufts-New England Medical Center, Center for Quality Assessment & Improvement in Mental Health (CQAIMH), 750 Washington Street, NEMC# 345, Boston, MA, 02111, USA. E-mail: rhermann@cqaimh

within organizations delivering care, provide oversight of quality by public agencies and private payers, and provide insight into what levels of performance are feasible. These activities require robust measures that permit meaningful comparisons across providers, systems, or geographic regions. In many areas of health care, however, there is a lack of agreement on which measures should be used. In mental health care, the challenges are to some extent even greater—the diverse nature of the field and competing priorities among stakeholders have slowed consensus development on a core set of measures for common use [4]. Despite these limitations, several countries have implemented measures to evaluate mental health care [11–22].

The Organization for Economic and Community Development's Health Care Quality Indicators Project (OECD–HCQI) is the first effort of which we are aware to identify measures for international benchmarking of quality of mental health care. This article reports on the methods employed to develop consensus among participants along with the resulting measures, as well as challenges to be surmounted for further progress to be achieved.

Methods

We conducted a consensus development process with a panel of international experts drawing on established procedures and based on a framework for quality-measure selection. The consensus development process employed elements of the modified Delphi method, because it has been applied to selection of quality measures [23,24]. The panelists' evaluation was conducted over two phases. In the first phase, each panelist anonymously rated measures on numerical scales for importance, soundness, and feasibility. The results were used to identify measures with sufficient consensus for inclusion or exclusion. For measures where consensus was lacking, the panel conducted a second review, making decisions about inclusion/exclusion on a case-by-case basis.

Our framework [25] reflects the twin—and in some respects conflicting—goals of measure selection. In selecting quality measures, organizations typically seek to maximize desirable measure attributes, which the OECD–HCQI characterizes as: importance, scientific soundness, and feasibility. But organizations also typically seek measures that are representative of highly diverse health care systems. Dimensions of this diversity include domains of quality (e.g. prevention, access, assessment, treatment, continuity, coordination, safety, and outcomes). They also include breadth among clinical disorders (emphasizing conditions with high prevalence, morbidity, and treatability), treatment modalities (both medication and psychosocial variation), clinical settings across the continuum of care, and vulnerable subpopulations including children, the elderly, and racial/ethnic minorities. Complicating the process of measure selection are tensions among these goals. Clinically important, evidence-based measures typically require richer, more costly data sources [26]—thus, more important measures may be less feasible to implement. Measures of mental health care for children are at an earlier stage

of development and testing than measures for adults—thus, broader representativeness may conflict with scientific soundness. These examples illustrate trade-offs in measure selection that our framework helps to make explicit and addressable [25].

In response to a call for participation from the OECD Secretariat six countries (United Kingdom, Sweden, Canada, Australia, Denmark, and the United States) and one international organization, the European Society for Quality in Healthcare (ESQH) nominated seven panelists. Panel members included mental health administrators, clinicians, and services researchers with expertise in quality of care, epidemiology, public health, and public policy. Additionally, each panelist had experience in the development of national core sets of quality measures for mental health care.

The seven panelists participated in a multistage consensus development process (Figure 1). Firstly, panelists identified and reviewed 134 measures of care for mental health and substance-use disorders (the *candidate set*). Indicators were drawn from OECD member countries' initiatives, conducted by national health departments, payers, accreditors, researchers, and other stakeholder organizations. Specific sources included the Canadian Mental Health Advisory Network, the United Kingdom Department of Health, the Center for Quality Assessment and Improvement in Mental Health's National Inventory of Mental Health Quality Measures, numerous US stakeholder initiatives, and published research reports. Information about these sources is detailed in the report, *Selecting Indicators for the Quality of Mental Health Care at the Health Systems Level in OECD Countries* [27].

The 134 candidate measures were screened against the criteria established by the OECD Secretariat to identify measures appropriate for international benchmarking [28]. The criteria sought to establish a conceptual focus, standardized methods, and a preliminary, achievable goal that would serve as a foundation for future work. For inclusion in the candidate set, measures had to meet each of the following:

- Indicators focusing on quality (as opposed to cost or utilization).
- Indicators relevant to assessing quality at the system level.
- Indicators focusing on technical (rather than interpersonal) quality.
- Indicators constructed from pre-existing administrative data based on standardized coding.
- Single item indicators (rather than multi-item scales).

Of the candidate measures, 23 met screening criteria (the *screened set*). These measures, accompanied by information on their specifications and data sources, were reviewed and rated by the panel using a structured assessment process. Each measure was rated anonymously on 9-point Likert scales for indicator importance and scientific soundness. In rating importance, panelists were asked to consider import to policy, impact on health, and susceptibility to being influenced by the health care system. In rating scientific soundness, panelists were asked to consider face validity, content validity, and explicitness of the evidence base. Each panelist also rated the feasibility of data collection in his or her country as likely, possible, or unlikely for each measure. Ratings were obtained from five panelists (all but representatives from Denmark and

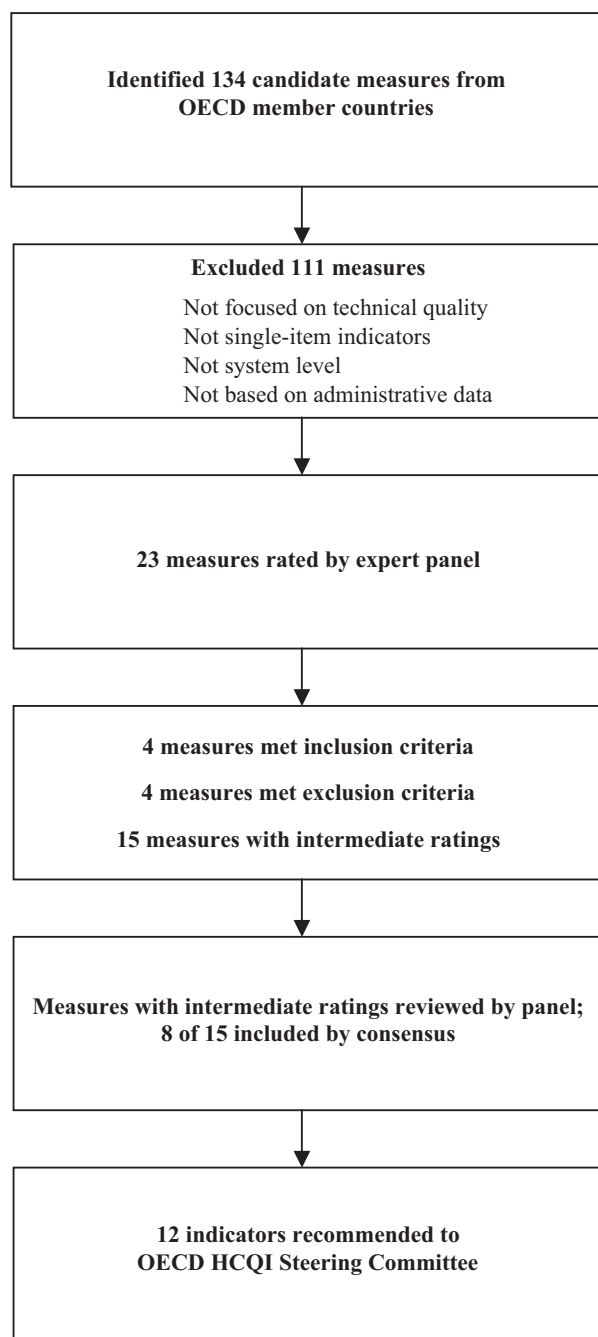


Figure 1 Measure selection process.

ESQH); all seven panelists participated in discussions preceding and subsequent to the rating process.

In developing decision rules for inclusion/exclusion of measures, we adopted the approach used in the RAND appropriateness method [29]: 9–7 indicating agreement, 6–4 neither agreement nor disagreement, and 3–1 disagreement. To derive a *preliminary set*, we included measures that [1] had a median score ≥ 7 for both importance and soundness and [2] more than half of participating panelists reported that data availability for the indicator was either ‘possible’ or ‘likely’. Excluded from this preliminary set were measures that had a

median score ≤ 3 for either importance or soundness or half or more participating panelists reported that data availability was ‘unlikely’.

The remaining measures—with importance and soundness scores between 4 and 7 as well as more than half of panelists reporting data availability to be at least ‘possible’—were returned to the panel for a second round of review. Each measure was discussed further with regard to their merit and their contribution to the breadth of topics addressed. Inclusion/exclusion decisions were then made on a case-by-case basis by consensus. The resulting *recommended set* was forwarded to the OECD–HCQI Steering Committee for their consideration.

Results

Of the initial 134 *candidate set* measures, only 23 (17%) met the specified screening criteria and were rated by the expert panel. Of the 23 measures meeting screening criteria, 4 received ratings meeting the inclusion criteria for the *preliminary set*, 4 met exclusion criteria, and 15 with intermediate ratings were forwarded to the panel for further review. Panelists discussed the strengths and weaknesses of the measures receiving intermediate ratings along with their relative contribution to a broad and balanced measure set. Consensus was achieved on including 8 of the 15 measures in the final, *recommended set*, for 12 measures.

Table 1 demonstrates the influence of each stage of the selection process on the diversity of the resulting measure sets. In proceeding from the candidate measure set to the criteria-based preliminary set, the number of quality domains represented by the resulting measure set decreased from eight to two, diagnostic groups categories from six to three, treatment modalities from five to one, and vulnerable populations from five to one. In the subsequent phase of the selection process, panelists noted what they regarded as important gaps among the subjects of these measures and identified what they considered ‘good enough’ measures among those with intermediate scores to fill these gaps. This phase resulted in an expansion in the number of measures within in each dimension in the final, recommended measure set to four quality domains, three diagnostic groups, four treatment modalities, and four vulnerable populations.

Table 2 describes the measures in the recommended set and their rating scores. The mean importance score for the set was 6.66, and the mean score for scientific soundness was 6.21. Data availability was assessed as ‘likely’ or ‘possible’ by all five panelists for seven measures and by four of the five panelists for the other five measures. Minor revisions were made to the recommended measures to refine specifications and eliminate redundancy.

Discussion

The 12 indicators recommended by the Mental Health Panel assess clinically important processes and outcomes of care

Table 1 Characteristics of initial, screened, and recommended indicators

	Indicator set			
	Candidate set (<i>n</i> = 134)	Screened set (<i>n</i> = 23)	Preliminary set (<i>n</i> = 4)	Recommended set (<i>n</i> = 12)
Quality domain				
<i>Process</i>	113	18	4	11
Treatment	43	12	2	6
Continuity	11	4	2	4
Coordination	4	2	0	1
Safety	23	0	0	0
Assessment	18	0	0	0
Prevention	0	0	0	0
Access	0	0	0	0
Other	14	0	0	0
<i>Outcome</i>	21	5	0	1
Clinical conditions				
Across diagnoses	84	13	2	7
Depressive disorders	38	6	1	3
Schizophrenia/Other psychotic disorders	5	1	0	0
Substance-related disorders	3	1	1	2
Bipolar disorder	1	1	0	0
Borderline personality disorder	1	1	0	0
Treatment modalities				
Medication	20	5	1	3
Case management	3	2	0	1
Psychotherapy	2	1	0	1
Electroconvulsive Therapy (ECT)	2	0	0	0
Other	2	2	0	1
Vulnerable populations				
Elderly	18	3	0	1
Severe persistent mental illness	17	2	0	1
Children/adolescents	2	0	0	0
Racial/ethnic minorities	1	1	0	1
Dual diagnosis MH/SA	2	1	1	1
Settings				
Outpatient	44	13	3	8
In-patient	42	4	1	1
Unspecified	48	6	0	3

where there is known variation in the quality of clinical practice. The measures, selected by an international expert panel using a structured process, represent progress toward the goal of identifying consensus-based measures for international benchmarking of mental health care. This is only a step in a longer-term process, however, as noted by the measures' moderate ratings on importance, scientific soundness, and feasibility. Information on these measures—including their clinical rationale, basis in research evidence, specifications, prior results, and testing—can be found in the OECD report and detailed review of measures [4,27].

The selection process achieved some degree of success toward the objective of selecting a measure set reflecting the diversity of mental health systems. The majority of the measures are applicable across diagnostic categories, whereas five

evaluate care specific to depression and substance-use disorders—conditions of high prevalence, morbidity, and treatability. The measures evaluate several domains of quality, including treatment, continuity, coordination, and outcome. They assess several modalities of treatment, including medication management, psychotherapy, and case management. They evaluate care in both in-patient and outpatient settings, as well as care for vulnerable subgroups such as elderly patients and racial/ethnic minorities.

Limited to measures previously implemented within OECD member countries and constructible from pre-existing administrative data sets, the resulting measure set also has significant gaps. Hundreds of quality measures have been proposed for mental health care; however, a US study found wide variability in their evidence base, operational development,

Table 2 Ratings for recommended mental health and substance-related quality indicators

Name	Description	Importance median	Soundness median	Data availability (<i>n</i> = 5 panelists)		
				Unlikely	Possible	Likely
<i>Treatment</i>						
Visits during acute phase treatment of depression	% of persons with a new diagnosis of major depression who receive at least three medication visits or at least eight psychotherapy visits in a 12-week period [34].	7.00	7.50	1	2	2
Hospital readmissions for psychiatric patients	% of discharges from psychiatric in-patient care during a 12-month reporting period readmitted to psychiatric in-patient care that occurred within 7 and 30 days [12].	7.00	7.00	0	0	5
Length of treatment for substance-related disorders	% of persons initiating treatment for a substance-related disorder with treatment lasting at least 90 days [35].	6.00	6.50	0	3	2
Use of anticholinergic antidepressant drugs among elderly patients	% of persons age 65+ years prescribed antidepressants using an anticholinergic anti-depressant drug [36].	6.00	6.00	1	2	2
Continuous antidepressant medication treatment in acute phase	% of persons age ≥18 years who are diagnosed with a new episode of depression and treated with antidepressant medication, with an 84-day (12-week acute treatment phase) treatment with antidepressant medication [15,37].	6.00	4.00	1	2	2
Continuous antidepressant medication treatment in continuation phase	% of persons age ≥18 years who are diagnosed with a new episode of depression and treated with antidepressant medication, with a 180-day treatment of antidepressant medication [15,37].	6.00	4.00	1	2	2
<i>Continuity</i>						
Timely ambulatory follow-up after mental health hospitalization	% of persons hospitalized for primary mental health diagnoses with an ambulatory mental health encounter with a mental health practitioner within 7 and 30 days of discharge [12].	8.00	7.00	0	1	4
Continuity of visits after hospitalization for dual psychiatric/ substance-related conditions	% of persons discharged with a dual diagnosis of psychiatric disorder and substance abuse with at least four psychiatric and at least four substance abuse visits within the 12 months after discharge [38].	7.00	7.00	0	3	2
Racial/ethnic disparities in mental health follow-up rates	% of persons with a mental health-related visit receiving at least one visit in 12 months after initial visit stratified by race/ethnicity [18].	7.00	6.50	0	3	2

continued

Table 2 *continued*

Continuity of visits after mental health-related hospitalization	% of persons hospitalized for psychiatric or substance-related disorder with at least one visit per month for 6 months after hospitalization [39].	6.00	6.00	0	1	4
<i>Coordination</i>						
Case management for severe psychiatric disorders	% of persons with a specified severe psychiatric disorder in contact with the health care system who receive case management (all types) [12].	7.00	6.50	0	4	1
<i>Outcome</i>						
Mortality for persons with severe psychiatric disorders	Standardized mortality rate for % of persons in total population with specified severe psychiatric disorders [12].	7.00	6.50	1	2	2

validity and reliability, and adequacy of case mix adjustment [4,26]. Table 1 demonstrates the implications of these findings for this project—namely the loss of diversity among measure topics as standards for importance, soundness and feasibility are applied. The recommended measure set lacks indicators of prevention, access, assessment, and safety of care. Although several measures could be meaningfully stratified by age, the set does not include measures specific to children’s mental health services. Common comorbidities of mental illness—including substance-use disorders and medical conditions—are only indirectly addressed. Measures were not available to assess emergent care or services at intermediate levels of care, such as residential or partial programs. The medication measures examine appropriateness of care more effectively than the measures of psychosocial interventions such as therapy or case management. A contributing factor is that administrative data provide more detailed information about medication management (e.g. dose, duration, and intensity) than about the content of psychosocial treatments. These challenges have been encountered in other initiatives that have sought to identify measures based on administrative data [30]. They are intensified here by the further restriction to data available on a comparable basis across several countries.

Limitations to this study include the number of participants (seven individuals from six countries and one international organization). Panelists did have a unique depth of experience, each previously having participated in one or more national or international consensus development processes on this topic. The generalizability of the results may be limited to developed countries, because developing countries struggle with the adequacy of their data systems in addition to their systems of mental health care. The WHO has developed the WHO Assessment Instrument for Mental Health Systems (WHO-AIMS) to support quality assessment and planning in low- and middle-income countries lacking robust systems for administrative data [31].

It should additionally be noted that measures of rate-based processes and outcomes represent a subset of a broader range of approaches to quality assessment in mental health care.

Other methods providing essential insights include: (i) evaluation of patient perceptions of care, (ii) measurement of clinical outcomes, such as change in symptoms, functioning, or quality of life, and (iii) assessment of the fidelity of evidence-based interventions to their empirically proven models. There has been little standardized implementation of such instruments either within or between nations [32].

Progress from this point forward is likely to be incremental and iterative. First, this measure set will need refinement as the OECD further investigates the availability of required data elements among countries participating in the benchmarking initiative. Secondly, pilot implementation of these measures would provide data allowing for measure testing and development of case mix adjustment. Thirdly, the measure of life expectancy among individuals with severe mental illness illustrates the potential for linkage among data sources to contribute to the development of meaningful quality measures. The indicator reflects research findings that individuals with severe mental illness die at a younger age than members of the general population and that better detection and general medical care for these individuals could contribute to narrowing this gap [33]. The OECD plans to further explore the ability of member countries to link national health care and mortality data sets. Analogous linkages may present opportunities to examine educational or criminal-justice outcomes of mental health care. Fourthly, these initial efforts at consensus development should be built upon by identifying key clinical variables that, if added to existing administrative data systems, would allow for measurement of important clinical processes. A more extensive but crucial undertaking is developing consensus nationally and internationally on standardized methods for structured assessment of clinical diagnoses, symptom severity, and functional impairment as well as on tools to evaluate patient experiences and treatment fidelity.

In the meantime, the measure set recommended herein provides a starting point for international benchmarking of mental health care. It covers several relevant dimensions of care and addresses known variations in important processes

and outcomes. Although much work will be needed to refine, specify, implement, and augment these measures, they provide a foundation for further progress.

Acknowledgements

The authors thank Peter Hussey for his contributions to the indicator selection process, Elizabeth Cote and Leighna Kim for providing research support, and Victoria Braithwaite and Orla Kilcullen for their help in preparing this manuscript. The comments of the members of OECD Expert Group on the Health Care Indicators Project and the Ad Hoc Group on the OECD Health Project on an earlier version of this manuscript are greatly appreciated. The authors also thank John Martin, Martine Durand, Peter Scherer, and Jeremy Hurst for review and comments. This paper reflects the opinion of the authors and not an official position of the OECD, its Member countries or institutions participating in the project.

References

- World Health Organization. *Mental Health in the WHO European Region*, Fact sheet EURO/03/03, Copenhagen, Vienna, 2003.
- 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, MA: Harvard School of Public Health, on behalf of the World Health Organization and the World Bank, Harvard University Press, 1996.
- Gabbard GO, ed. *Treatments of Psychiatric Disorders*, third edition. Washington, DC: American Psychiatric Press, Inc., 2001.
- Hermann RC. *Improving Mental Healthcare: A Guide to Measurement-Based Quality Improvement*. Washington, DC: American Psychiatric Press, Inc., 2005.
- Committee on Quality Health Care in America, Institute of Medicine. *Crossing the Quality Chasm: Adaptation to Mental Health and Addictive Disorders*. Washington, DC: Institute of Medicine.
- World Health Organization. *The World Health Report 2001 – Mental Health: New Understanding, New Hope*. Geneva, Switzerland: WHO, 2001.
- Collaborative Working Group on Shared Mental Health Care. *Shared Mental Health Care in Canada – Current Status, Commentary and Recommendations*. Ottawa, Ontario, Canada: Canadian Psychiatric Association and the College of Family Physicians of Canada Collaborative Working Group on Shared Mental Health Care, 2000.
- The Australian Council for Safety and Quality in Health Care. *Achieving Safety and Quality Improvements in Health Care – Sixth Report to the Australian Health Ministers' Conference*. Canberra, ACT: Commonwealth of Australia, 2005.
- United Kingdom Department of Health. *National Service Framework for Mental Health: modern standards and service models*. London, England: UK National Health Service (NHS), 1999.
- Mattke S, Epstein A, Leatherman S. The OECD Health Care Quality Indicators Project: History and Background. *Int J Qual Health Care* (in press).
- Pirkis J, Burgess P, Dunt D, Henry L. *Measuring Quality in Australian Mental Health Services*, Melbourne, Australia: Centre for Health Program Evaluation, 1999 December.
- McEwan K, Goldner EM. *Accountability & Performance Indicators for Mental Health Services & Supports, A Resource Kit*. Ottawa, Canada: Federal/Provincial/Territorial Advisory Network on Mental Health, 2001.
- National Health Service Information Authority. *Mental Health Minimum Data Set, Data Manual*. London, England: United Kingdom National Health Service (NHS), 2002.
- The Commonwealth Fund. *First Report and Recommendations of the Commonwealth Fund's International Working Group on Quality Indicators*, 2004 June.
- National Committee for Quality Assurance. *Health Plan Employer Data and Information Set (HEDIS) 2006*. Washington DC: National Committee for Quality Assurance (NCQA), 2005.
- NASMHPD Research Institute. *Implementation of the NASMHPD Framework of Mental Health Performance Measures by States to Measure Community Performance: 2001*. Alexandria, VA: NASMHPD Research Institute, 2002.
- Garnick D, Lee M, Chalk M *et al*. Establishing the feasibility of performance measures for alcohol and other drugs. *J Subst Abuse Treat* 2002; **23**: 375–385.
- Center for Mental Health Services. *The Final Report of the Mental Health Statistics Improvement Project (MHSIP) Task Force on a Consumer-Oriented Mental Health Report Card*. Rockville, MD: Center for Mental Health Services, 1996.
- Shield T, Campbell S, Rogers A, Worrall A, Chew-Graham C, Gask L. Quality indicators for primary care mental health services. *Qual Saf Health Care* 2003; **12** (2): 100–106.
- National Mental Health Working Group Information Strategy Committee Performance Indicator Drafting Group. *Key Performance Indicators for Australian Public Mental Health Services. ISC Discussion Paper No. 6*. Canberra, Australia: Australian Government Department of Health and Ageing, 2005.
- Mainz J, Krog BR, Bjornshave B, Bartels P. Nationwide continuous quality improvement using clinical indicators: the Danish National Indicator Project. *Int J Qual Health Care* 2004; **16** (Suppl. 1): i45–i50.
- Glover GR. A comprehensive clinical database for mental health care in England. *Soc Psychiatry Psychiatr Epidemiol* 2000; **35** (11): 523–529.
- Hermann RC, Palmer RH, Leff HS *et al*. Achieving consensus across diverse stakeholders on quality measures for mental healthcare. *Med Care* 2004; **42** (12): 1246–1253.
- Schuster M, Asch S, McGlynn E, Kerr E, Hardy A, Gifford D. Development of a quality of care measurement system for children and adolescents. *Arch Pediatr Adolesc Med* 1997; **151**: 1085–1092.
- Hermann RC, Palmer RH. Common ground: a framework for selecting core quality measures. *Psychiatr Serv* 2002; **53** 3: 281–287.
- 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): 135–154.

27. Hermann RC, Mattke S, 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*. Paris: Organization for Economic Co-operation and Development (OECD), 2004. <http://www.oecd.org/dataoecd/28/32/33865630.pdf>.
28. Kelley E, Hurst J. *Health Care Quality Indicators Project Conceptual Framework Paper. OECD Health Working Paper No. 23*. Paris, France: Organisation for Economic Co-operation and Development, 2006. <http://www.oecd.org/dataoecd/1/36/36262363.pdf>.
29. Brook R, Chassin M, Fink A. A method for the detailed assessment of the appropriateness of medical technologies. *Int J Technol Assess Health Care* 1986; **2**: 53–63.
30. Garnick D, Hendricks A, Comstock C. Measuring quality of care: fundamental information from administrative data sets. *Int J for Quality Health Care* 1994; **6 (2)**: 163–177.
31. World Health Organization (WHO). *World Health Organization Assessment Instrument for Mental Health Systems (WHO-AIMS)*. Geneva: World Health Organization, 2005.
32. Glover G. The audit of mental health services. *Qual Assur Health Care* 1990; **2**: 181–188.
33. Harris EC, Barraclough B. Excess mortality of mental disorder. *Br J Psychiatry* 1998; **173**: 11–53.
34. Veterans Health Administration/Department of Defense. *VHA/DOD Performance Measures for the Management of Major Depressive Disorder in Adults*, Version 2.0. Washington, DC, 2000.
35. New York State Office of Alcoholism and Substance Abuse Services. *IPMES/Workscope Objective Attainment System: FY2000 and 2000–2001*. Albany, NY: New York State Office of Alcoholism and Substance Abuse Services, 2000.
36. Wells K, Norquist G, Benjamin B, Rogers W, Kahn K, Brook R. Quality of antidepressant medications prescribed at discharge to depressed elderly patients in general medical hospitals before and after prospective payment system. *Gen Hosp Psychiatry* 1994; **16**: 4–15.
37. Agency for Healthcare Research and Quality. *The National Healthcare Quality Report*. Rockville, MD: U.S. Department of Health and Human Services, 2003.
38. Rosenheck R, Cichetti D. *A Mental Health Program Performance Monitoring System for the Department of Veterans Affairs*. West Haven, CT: Northeast Program Evaluation Center, 1995.
39. Leslie D, Rosenheck R. Comparing quality of mental health care for public-sector and privately insured populations. *Psychiatr Serv* 2000; **51 (5)**: 650–655.

Accepted for publication 16 June 2006