

Beyond the initial indicators: lessons from the OECD Health Care Quality Indicators Project and the US National Healthcare Quality Report

EDWARD T. KELLEY^{1,2}, IRMA ARISPE³ AND JULIA HOLMES³

¹Organization for Economic Cooperation and Development, Paris, France, ²US Department of Health and Human Services, Agency for Healthcare Research and Quality, Rockville, and ³US Department of Health and Human Services, Centers for Disease Control and Prevention, Hyattsville, MD, USA

Abstract

Interest in comparative quality measurement and evaluation has grown considerably over the past two decades because of factors such as the recognition of widespread variation in clinical practice, the increased availability of evidence about medical effectiveness, and increasing concern about the cost and quality of health care. This article describes and contrasts two current efforts to develop health performance reporting systems: one, an international initiative—the Health Care Quality Indicator (HCQI) Project, sponsored by the Organization for Economic Cooperation and Development (OECD); and the other, a national project—the National Healthcare Quality Report (NHQR), sponsored by the US Agency for Healthcare Quality and Research. There are a number of lessons learned from a comparison of the two efforts that are relevant for the future of each project and for other indicator-based reporting efforts in quality of health care. These lessons are discussed in the article and include:

1. Conceptual frameworks should be established to guide the selection of indicators.
2. Choices should be made early on in the process to focus on a wide range of clinical conditions or to report on a few priority areas.
3. Methods should be developed to add and subtract indicators while maintaining a stable set of indicators to track over time.
4. Resources should be allocated to communication strategies and how best to present data results to diverse audiences.
5. Mechanisms should be put in place to maintain project momentum.

Keywords: indicators, performance reporting, quality

Interest in comparative quality measurement and evaluation has grown considerably over the past two decades because of factors such as the recognition of widespread variation in clinical practice, the increased availability of evidence about medical effectiveness, and increasing concern about the cost and quality of health care. This interest has led to several international, as well as national, efforts to summarize what is known about clinical quality of care using science-based indicators generated from a variety of data sources [1–3]. Beginning with the World Health Organization's 2000 World Health Report that ranked the health system performance of 191 countries on the basis of five composite indicators, including disability-adjusted life expectancy, equity, financing, and system responsiveness, international comparisons have generated considerable discussion

and debate [4–7]. Concerns have focused on the conceptual and methodological difficulties inherent in comparing health system performance at any level, both within and across countries, as well as how the results of such comparative analyses are communicated and used by policymakers [5]. Dilemmas at the conceptual level relate to the need for a valid performance measurement system that separates determinants of health that lie outside the health system from those that are attributable to the delivery of health care [7]. Methodological issues, particularly in cross-national comparative analyses, focus on the comparability of indicators across countries with respect to data specifications and data availability, as well as the interpretability of health indicators in light of differences in cultural context and health care delivery systems [2,4–9].

The views expressed in this article are those of the authors and do not necessarily reflect those of the Organization for Economic Cooperation and Development, the Agency for Healthcare Research and Quality, the National Center for Health Statistics or the U.S. Department of Health and Human Services.

Address reprint requests to Edward T. Kelley, PhD, Head, Health Care Quality Indicators Project, Organization for Economic Cooperation and Development, 2 Rue André Pascal, 75775 Paris, Cedex 16, France. E-mail: edward.kelley@oecd.org

This article describes and contrasts two current efforts to develop health performance reporting systems: one, an international initiative—the Health Care Quality Indicator (HCQI) Project, sponsored by the Organization for Economic Cooperation and Development (OECD); and the other, a national project—the National Healthcare Quality Report (NHQR), sponsored by the US Agency for Healthcare Research and Quality (AHRQ). In the pages that follow, we examine issues related to the purpose of the reporting efforts, the development of a conceptual framework and indicator

set, approaches for managing and presenting the data, and strategies for sustaining project momentum in national and international quality measurement initiatives.

Purpose of the reports

A snapshot comparison of the OECD HCQI Project and the US NHQR is presented in Table 1. The OECD HCQI Project, initiated in 2001, is currently the only ongoing

Table 1 A snapshot comparison of the OECD Health Care Quality Indicators Project and the US National Healthcare Quality Report

Characteristic	OECD HCQI	US NHQR
Origin	Request by member countries	Legislative mandate to develop two Reports to Congress on health care quality and health disparities
Purpose	Identify measures appropriate for cross-national comparison Assess measure and data comparability Resolve methodological comparability issues if possible Disseminate information about the measures to member countries	Develop a set of indicators appropriate for profiling health care quality for the nation, including trends over time Examine differences at the subnational level Examine variations by socioeconomic status Disseminate information to a broad audience
Conditions covered	Cancer Vaccine preventable diseases/immunizations Asthma Heart disease (AMI/stroke) Waiting time for surgery Risk factors (smoking)	Cancer Diabetes ESRD Heart disease HIV/AIDS Maternal child health Mental health Respiratory disease Nursing home and home health care
Number of measures	13	148 in the first (2003) NHQR
Type of indicators	Process and outcome	Process and outcome
Development process	Technical experts with vetting by ministers of member countries	Interagency working group (technical experts) with vetting by private sector organizations and clearance by US government
Development of a framework	Examined frameworks of participating countries Built on existing international comparisons (Commonwealth and Nordic Ministers Council working group) Built on US Institute of Medicine framework	Examined other national efforts to develop consensus on measures (HP2010) Contracted with the US Institute of Medicine for development of a framework
Primary criteria for selecting measures	Importance of the indicator	Importance of the indicator
Dissemination	Scientific soundness Feasibility to generate data on an ongoing basis Technical working papers Inclusion of selected measures in OECD data base	Scientific soundness Feasibility to generate data on an ongoing basis Reports to Congress Measure specifications included in national quality measures clearinghouse Development of condition-specific and state-specific reports Data available on the web

international effort aimed at measuring quality of care across a range of health care conditions, such as cancer, diabetes, cardiovascular disease; across a range of dimensions of health care quality, such as effectiveness, patient safety and responsiveness or patient centeredness; and across a range of patient needs, such as preventive health, curative care, living with disability, and coping with end of life. The Project's work on a broad set of indicators in new disease areas, such as primary care and prevention, patient safety, and mental health, is groundbreaking in that, in some instances, it is the first effort aimed at developing consensus around what is essential to measure in health care quality for those conditions internationally. The effort at deriving an initial set of indicators for the HCQI Project has spanned 4 years, involving extensive consensus building across the 23 participating countries and detailed analysis on the comparability of data across different country data sources. The HCQI Project was developed originally under the OECD's Health Project, a multi-disciplinary effort begun at the OECD in 2001. The work on quality indicators was developed at the request of member countries interested in improving performance measurement of health system outputs that could be used in conjunction with the OECD's well developed database on health care spending, utilization, and mortality levels. The *HCQI Project Initial Indicators Report* and *HCQI Project Conceptual Framework Paper* were released in March 2006¹.

During that same time period, within the United States, the development process was well underway for the first-ever NHQR and its companion, the National Healthcare Disparities Report (NHDR). The mandate for these annual congressional reports was set forth in the 1999 enabling legislation for the AHRQ [10]. A wide-scale effort, the development of the NHQR engaged the technical and substantive input of nine federal agencies and organizational units within the Department of Health and Human Services (DHHS)² as well as a range of state public partners and private sector organizations. Now in its fourth year, the NHQR offers a well-established, consensus-based set of health care quality measures across four dimensions of quality—effectiveness, safety, timeliness, and patient centeredness. The report examines effectiveness of care across nine clinical condition areas—cancer, diabetes, end-stage renal disease, heart disease, HIV/AIDS, maternal and child health, mental health, respiratory diseases, and nursing home and home health care. In terms of the number of measures and dimensions of care reviewed, they were the broadest examination of quality ever completed in the United States or any other major industrialized country [11].

Both the HCQI and the NHQR had as their objective to develop a set of indicators that can be used as the basis for identifying variation and stimulating investigation to understand why differences exist and what can be done to reduce those differences. Both envisioned an indicator set that

would be scientifically sound, important at a clinical and policy level, and feasible to generate on a regular basis using standard specifications that result in comparable data. However, they differed in two important ways. Firstly, the NHQR was legislatively mandated as an annual report on trends in health care quality. Secondly, the purpose of the report is to summarize the current state of health care quality in terms that are understandable and relevant to a broad audience including providers, consumers, researchers, and policymakers. To that extent, the immediate focus of the project was systematic reporting on (i) trends and change over time, (ii) differences at the subnational level, specifically by state, and (iii) variations by selected socio-demographic characteristics. The report contained detailed tables and took on a variety of formats targeted toward a range of interested constituencies.

By contrast, although the purpose of the HCQI was to develop indicators for cross-national comparison, the project recognized early on the substantial and unique challenges of arriving at comparable data and thus turned to the difficult methodological work of detailing specifications, reconciling differing data sources, understanding differences in health care systems and their mandates, and separating differences in performance from differences in measurement. In addition, whereas the NHQR had the advantage of using a single data source for each measure, the HCQI Project team had the challenge of integrating disparate data sources for a single measure. The comparative goal of the project therefore evolved to a longer term objective, with the short-term goal being to generate information that will raise questions and stimulate further investigation about the measures themselves. The first phase of the project has resulted in technical papers detailing the measure specifications and the investigation into their comparability [12,13].

Development of a framework and selection of indicators

The formation of both the OECD HCQI Project and the NHQR drew on prior, related initiatives. The HCQI project has built on two pre-existing international collaborations organized by the Commonwealth Fund of New York (five countries) and The Nordic Minister Council Working Group on Quality Measurement (six countries).³ When the NHQR was initiated in 1999, no framework existed for selection of measures; consequently, two approaches were used to develop a framework. Firstly, to identify priority condition areas, the NHQR drew from Healthy People 2010, the United States' most comprehensive effort to build consensus between local, state, and national policymakers and the public health community on goals, objectives, and measures for the nation's health. In

¹These papers are available from the OECD at www.oecd.org/health.

²There are now 12 federal agencies and organization units within US DHHS working on the US National Reports.

³The Commonwealth Fund's International Working Group on Quality Indicators included the United States, the United Kingdom, Canada, Australia, and New Zealand. The Nordic Minister Council Working Group on Quality Measurement includes Greenland, Sweden, Norway, Finland, Iceland, and Denmark.

addition, AHRQ contracted with the National Academy of Sciences Institute of Medicine (IOM) to produce *Envisioning the National Health Care Quality Report* [14]. That report conceptualized a framework that encompassed dimensions of quality (effectiveness, safety, timeliness, patient centeredness, and for the NHDR, equity). It also put forth life stages (staying healthy, getting better, living with illness, and end of life) to help structure the measure set to ensure comprehensive coverage of the US population. Ultimately, however, both the HCQI Project and the NHQR drew from the framework put forth by the IOM.

The NHQR, because of its mandate to track quality of care for the 'nation' as a whole, has a mandate to present a broad picture of health care quality. However, the HCQI Project is not bound by such a mandate and could, alternately, choose to focus on particular areas. An area of consideration for the future of the HCQI Project's measure set is whether this international effort will concentrate on broadening its indicators to include a wide range of clinical conditions to provide a more comprehensive picture of health care quality at a 'system' level or whether the project should chose to focus its efforts on indicator development in a few priority areas. If these areas were tied closely to current country priorities, it would allow the project to track international data relevant to current policy directives within participating countries. The initial efforts at developing such indicators in 'priority areas' are presented in the articles in this journal. The fact that both the NHQR and the HCQI Project track progress on very basic, broad consensus indicators means that they track progress in caring for diseases which affect a broad area of any given country's population. However, the priorities for improving national performance in a given disease will vary from country to country and within countries across time as new administrations set new priorities. Understanding those priorities is an extra effort that must be addressed by staff associated with efforts such as the NHQR and the HCQI Project.

The NHQR experience illustrates the challenges of developing and using a broad-based, comprehensive set of indicators intended to profile the nation as a whole. Whereas the

addition of indicators is valuable to address the needs of emerging health issues, a proliferation of indicators can be at odds with the goal of providing concise messages about quality to a wide range of audiences. Additions to and deletions from an indicator set can also compromise the ability of the reporting institution to track quality over time using a stable set of indicators.

The solution arrived at for the US national report was to designate a set of core indicators which would be highlighted in the NHQR, whereas noncore indicators would appear only as tables in an appendix. Although indicators in the full measure set must have met criteria based on importance, scientific soundness, and feasibility, the Interagency Work Group established additional criteria for selecting the core report measures. Primary, secondary, and balancing criteria are summarized in Table 2. Primary criteria were given greater weight than secondary criteria. Balancing criteria were included to ensure that core report measures covered all conditions and sites of care included in the full measure sets. This process yielded 46 core report measures of health care quality that will remain stable over the years and can be tracked as part of composite measures and trend reporting.

Summarizing and presenting data

One key lesson from the US report experience is that thought needs to be given to reporting approaches as the presentation of data is perhaps the key factor in whether the data are understood and used by the report target audience. The HCQI Project's initial report on indicators focuses on methodological issues and solutions in comparing countries internationally on quality of care [12]. Data are presented in tabular format for countries for which comparable data are available, and discussion of the issues related to each indicator is featured prominently. This approach was appropriate given the analytic issues that have been raised during the development of the initial indicator set. For the NHQR, by contrast, a greater effort was put into the final presentation of data, so that it would be useful for the reports' target audiences: state

Table 2 Criteria for selecting core report measures

Primary criteria	Importance/clinical significance/prevalence Data reliable Able to be tracked for multiple groups and at multiple levels/number of comparisons possible Sensitive to change/evidence-based process measures favored over outcomes Easy to interpret and understand/methodological simplicity High utility for directing public policy
Secondary criteria	Applicable to general population rather than unique to select population Data available regularly/data available recently Data source supports multivariate modeling Linkable to established indicator sets (i.e. Healthy People 2010 targets)
Balancing criteria	Balance across health conditions Balance across sites of care At least some State data At least some multivariate models

and national policymakers. In 2004, the availability of state data was improved in the measure set, and state data, in the form of maps, were used in many sections to highlight state and regional differences in quality of care. Moreover, efforts were made at analyzing states by quintile to allow for easily understandable maps. Maps have also been used to highlight examples of excellence where there may be lessons for other states to study, as in the example below from the 2005 report. The use of maps in the 2004 report and the release of 'state snapshots' which summarized the report data for each state in easily digestible web-based pages resulted in major national attention to the reports. Articles on the report appeared in most major metropolitan newspapers and local radio as each state's media sought to examine where the state led, and fell behind, other states in the region. The possibility of using such presentation techniques and providing country summaries is a definite possibility for the HCQI Project and one that would build on country-specific datum summaries done previously in health and other sectors of the OECD's work.

The first NHQR was primarily textual; in subsequent reports, a significant effort was made at improving the presentation of the written report with the adoption of a 'chart-book' format and bulleted summaries of the prevalence, mortality, morbidity, and cost of the different diseases. The analytical approach in the NHQR has remained relatively constant and simple by statistical standards, primarily using bivariate comparisons between national numbers and states

or comparing age or sex groups. As the HCQI Project moves into analyzing differences between countries, developing a straightforward method for presenting differences between countries as well as for analyzing the interaction of different factors in international quality differences will be important.

After the issuance of the first NHQR, efforts were made to summarize change in quality, drawing from methodologies developed for use in HealthyPeople 2010 [15]. Among the issues considered were methods for creating summary measures (inclusion criteria and weighting of component measures) and methods for assessing change over time. The HCQI Project will need to examine similar methodological issues in determining how to achieve a streamlined presentation of country differences and methods of summarizing across countries on different quality indicators (Figure 1).

Maintaining the momentum

The credibility of reports such as those produced by the OECD and by the AHRQ-led effort to develop the NHQR depends as much on the process used to produce the report as on the final product(s). Production of major analytical reports such as the NHQR and the HCQI Project depends heavily on advisory group input and support. In the case of the HCQI Project, its country Expert Group was essential in the production of the report, acting as the data suppliers for

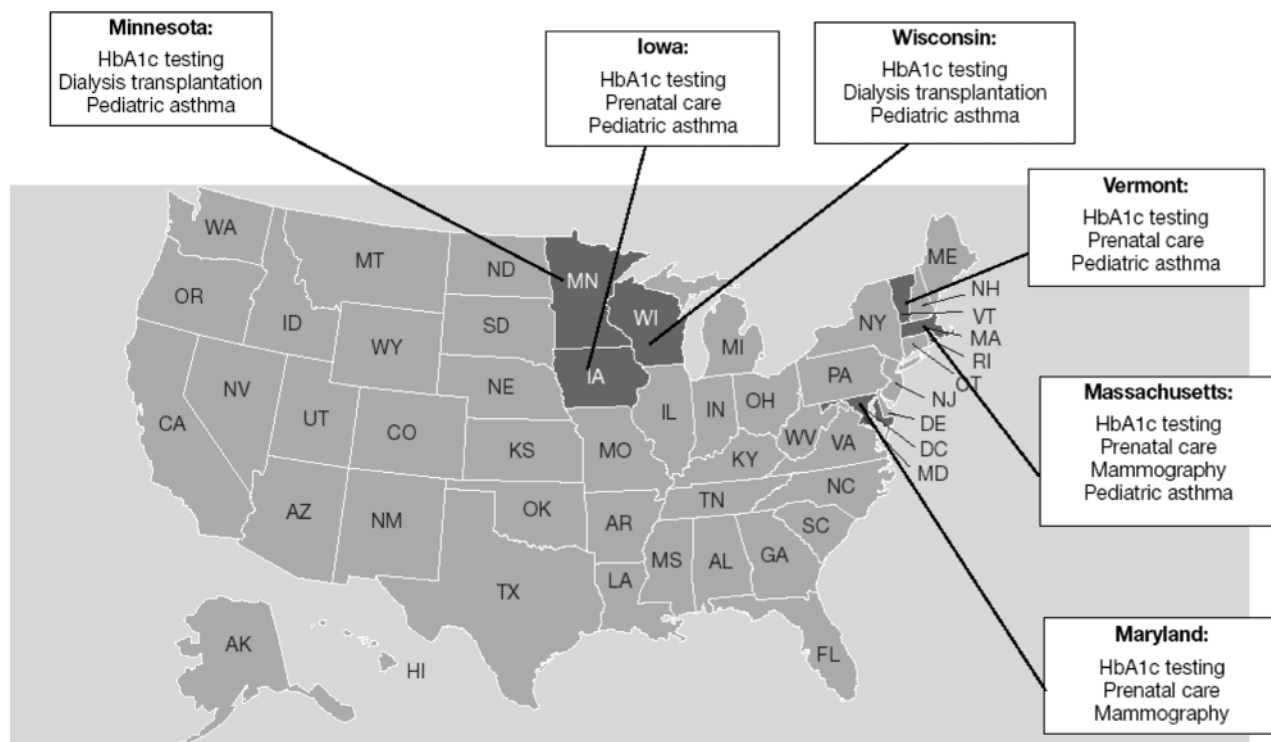


Figure 1 State-level summaries in the US National Reports: highlighting high performers.

Sources: HbA1c testing: Centers for Disease Control and Prevention, Behavioral Risk Factor Surveillance System, 2001–2003. Dialysis transplantation: University of Michigan Kidney Epidemiology and Cost Center, 1999–2002. Pediatric asthma: Agency for Healthcare Research and Quality, HCUP State Inpatient Databases, 2000–2002. Prenatal care: National Vital Statistics System-Nativity, 1999–2002. Mammography: Centers for Disease Control and Prevention, Behavioral Risk Factor Surveillance System, 2000–2002.

their respective countries and as co-authors on the first HCQI indicator report. In the case of the US reports, an Interagency Workgroup (IWG) was formed. In some cases, the IWG members supplied data for the reports, in others they were merely expert advisors. In other cases, the advice of outside national and international experts was sought. However, the NHQR experience highlights the importance of leadership to keep the energy and commitment of the advisory groups (as well as communication between the group members) intact. Although the HCQI Project has yet to set the periodicity with which its indicators will be updated, the task of maintaining interest and involvement of expert advisors become even more difficult in the case of an annual effort. The HCQI Project will have the added difficulty of relying on the good will and donated time of its experts to a greater degree than the US National Report, where the reports are officially a joint responsibility for the entire US DHHS.

Conclusion

It is clear that the initial work undertaken by the OECD in the HCQI Project and the United States in developing the NHQR was ambitious and has been groundbreaking in their breadth and methodological detail. Looking across the US Report and the HCQI efforts, there are a number of lessons learned that are relevant for health care quality reporting efforts at both the national and international levels. These include:

1. Conceptual frameworks should be established to guide the selection of indicators—while it is tempting to skip the step of laying out a framework, this exercise in both the US Report and the HCQI Project helped define the limits of the project and is the foundation for all future work in indicator development and reporting.
2. Choices should be made early on in the process to focus on a wide range of clinical conditions or to report on a few priority areas—the OECD HCQI work has been criticized for its lack of breadth, choosing initially to report on a limited number of clinical conditions such as cancer and diabetes. However, the United States attempted to provide a comprehensive overview of quality at the system level, including multiple clinical conditions, as well as the patient's perspective on care.
3. Methods should be developed to add and subtract indicators while maintaining a stable set of indicators to track over time—while the addition and deletion of indicators are necessary to remain responsive to changes in the evidence base for quality measurement, without a stable indicator set, it is difficult, if not impossible, to track quality over time. In addition, expanding the number of quality indicators is at odds with the goal of providing concise messages about quality to policy and consumer-based audiences.
4. Resources should be allocated to communication strategies and how best to present data results to diverse audiences—too often, indicator-based efforts such as these use all the product development time in indicator

specification discussions and data analysis, whereas the final presentation format is left as an afterthought. The US Reports illustrate the utility of using geographic maps and clear comparative charts using principles of benchmarking across states, whereas the HCQI Initial Report describes the many reasons why 'similar' indicators may differ in their collection and reporting across states, regions, or countries and why these differences need to be acknowledged.

5. Mechanisms should be put into place to maintain project momentum—it is difficult for consensus-driven projects such as the HCQI Project or the US National Reports, which depend on expert groups from multiple organizations or even countries, to maintain momentum when members change frequently. Clear communication on project progress and clear plans for updating the indicator sets are essential in keeping such indicator work moving forward.

Through a focus on a purpose-driven report—a consensus-based process for the selection of measures, and careful consideration of the presentation of data—comparative performance measurement projects such as these establish a firm foundation that will facilitate their success. However, this alone is not sufficient for success. Each project faces challenges such as the choice among using indicators for which there are available data versus developing new indicators; balancing the priorities of 28 countries along with those of other major international actors—such as the European Commission and the World Health Organization; and making the project useful for quality improvement efforts of national and international scope. The lessons of broad national efforts such as these show that once the report has been completed, the work has only begun.

References

1. World Health Organization. *The World Health Report 2000. Health Systems: Improving Performance*. Geneva: WHO, 2000.
2. Blendon RJ, Kim M, Benson JM. The public versus the World Health Organization on Health System Performance. *Health Aff* 2001; **20** (3): 10–20.
3. Richardson J, Wildman J, Robertson IK. A critique of the World Health Organization's evaluation of health system performance. *Health Econ* 2003; **12**: 355–366.
4. Hussy PS, Anderson GF, Osborn R et al. How does the quality of care compare in five countries? *Health Aff* 2004; **23** (3): 89–99.
5. Walshe K. International comparisons of the quality of health care: what do they tell us? *Qual Saf Health Care* 2003; **12**: 4–5.
6. Marshall MN, Shekelle EA, McGlynn EA, Campbell S, Brook RH, Rolond MO. Can health care quality indicators be transferred between countries? *Qual Saf Health Care* 2003; **12**: 8–12.
7. Nolte E, McKee M. Measuring the health of nations: analysis of mortality amenable to health care. *BMJ* 2003; **327**: 1129–1133.
8. McGlynn E. Perspective: there is no perfect health system. *Health Aff* 2004; **23** (3): 100–102.

9. Loeb JM. The current state of performance measurement in health care. *Int J Qual Health Care* 2004; **16 (suppl. 1)**: i5–i9.
10. US Congress. Healthcare Research and Quality Act 1999. Statutes at Large. Vol. 113, Sec 1653 (Public Law 106–129; Section 913(b)(2)) 1999.
11. US Department of Health and Human Services. *The US National Healthcare Quality and Disparities Report*. Rockville, MD: The Agency for Healthcare Research and Quality, 2006.
12. OECD Working Paper #22. *Health Care Quality Indicators Project Initial Report*. 10 March 2006.
13. OECD Working Paper #23. *Health Care Quality Indicators Project Conceptual Framework Paper*. 10 March 2006.
14. Institute of Medicine. *Envisioning the National Health Care Quality Report*. Washington, DC: National Academies Press, 2001.
15. Keppel K, Pamuk E, Lynch J *et al*. Methodological issues in measuring health disparities. *Vital Health Stat* 2005; **141**: 1–16.

Accepted for publication 4 July 2006