

The evolution of healthcare quality measurement in the United States

■ H. Burstin¹, S. Leatherman^{1,2} & D. Goldmann³

From the ¹National Quality Forum, Washington, DC; ²Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC; and ³Institute for Healthcare Improvement, Cambridge, MA, USA

Content List – Read more articles from the symposium

Abstract. Burstin H., Leatherman S., Goldmann D. (National Quality Forum, Washington, DC; University of North Carolina, Chapel Hill, NC; and Institute for Healthcare Improvement, Cambridge, MA, USA). Evaluating the quality of medical care (Review-Symposium). *J Intern Med* 2016; **279**:154–159.

Quality measurement is fundamental to systematic improvement of the healthcare system. Whilst the United States has made significant investments in healthcare quality measurement and improvement, progress has been somewhat limited. Public and private payers in the United States increasingly mandate measurement and reporting as part of pay-for-performance programmes. Numerous issues have limited improvement, including lack of alignment in the use of measures and improvement strategies, the fragmentation of the US healthcare system, and the lack of national electronic systems for measurement, reporting,

benchmarking and improvement. Here, we provide an overview of the evolution of US quality measurement efforts, including the role of the National Quality Forum. Important contextual changes such as the growing shift towards electronic data sources and clinical registries are discussed together with international comparisons. In future, the US healthcare system needs to focus greater attention on the development and use of measures that matter. The three-part aim of effective care, affordable care and healthy communities in the US National Quality Strategy focuses attention on population health and reduction in healthcare disparities. To make significant improvements in US health care, a closer connection between measurement and both evolving national data systems and evidence-based improvement strategies is needed.

Keywords: health care, quality, efficiency, disparities, international comparisons.

Introduction

Significant investments in healthcare quality measurement and improvement have been made in the United States. Many sectors of health care have viewed quality measurement as fundamental to systematically improving health system performance. Major efforts on the part of both public and private payers to drive quality improvement by mandating measurement and reporting and promoting quality improvement initiatives in the healthcare delivery system have resulted in some progress, although more is needed to achieve consistently high-quality care for all [1]. System-related issues, such as the complexity, lack of alignment and fragmentation of the US healthcare system, are frequently cited as reasons for these suboptimal results. Somewhat paradoxically, well-intentioned measurement efforts may be over-

whelming the healthcare delivery system rather than catalysing improvement.

Here, we provide an overview of the historical and current context for US healthcare measurement, including the growing shift towards electronic measures and clinical registries, and the changes that will be needed to bring about the desired changes in the healthcare system.

Historical context for US healthcare quality measurement

The history of quality measurement and the focus on patient outcomes in the United States can be traced back hundreds of years to pioneers such as Florence Nightingale and Ernest Codman. Notable events catalysed awareness of and investments in measurement and improvement. In 1966, the critical framework developed by Avedis Donabe-

dian provided a valuable conceptualization of structure, process and outcomes as essential domains for evaluating medical care quality [2]. The release of hospital-specific mortality data in 1987 by the Health Care Financing Administration stirred controversy about the wisdom of public release of outcomes data, but also resulted in serious efforts to improve measurement, such as case-mix adjustment models. In the early 1990s, the Health Employer Data Information Set (HEDIS) from the National Committee on Quality Assurance (NCQA) helped to standardize evaluation, allowing comparative reporting and promoting improvement initiatives within individual health plans across the United States.

The 1998 Advisory Commission on Consumer Protection and Quality in the Health Care Industry, established by President Clinton, called for a national effort to improve and sustain the quality of care in the United States. As part of its call to action, the Commission recommended the development of a public-private partnership that would help standardize measures across the healthcare enterprise. At the time, the science of quality measurement was still developing: measures were not widely available for many care settings and clinical conditions; measurement initiatives failed to routinely involve the healthcare professions or patients; and measures were frequently developed and promulgated by individual healthcare organizations, resulting in competing measurement sets and a lack of broad multistakeholder engagement.

Evolution of the National Quality Forum

The National Quality Forum (NQF) was a clear outgrowth of the Commission's work. Established in 1999, the NQF is a public-private partnership focused on setting standards for healthcare quality. At that time, little systematic information on the quality of health care was available and no commonly held 'rules of the road' for quality measurement and reporting existed. Without fair and transparent benchmarking, policy levers to improve care, including selective contracting and pay for performance, could not reasonably be contemplated. The Strategic Framework Board, comprising nine experts in quality of care, was established by the NQF in 1999 to (i) design a strategy for a national quality measurement and reporting system, (ii) articulate the guiding principles and priorities for a national system, and (iii) identify potential barriers to successful implementation and propose possible solutions'. A series of

reports in *Medical Care* published by members of the Strategic Framework Board in 2003 outlined an approach to the development of national goals and selection of a 'common set of measures and reporting strategies' [3]. Whilst the NQF adopted criteria for standardized measures as part of its multistakeholder endorsement process, many key recommendations from the Strategic Framework Board and from others regarding prioritization, reporting strategies and developing capacity for quality improvement nationwide have not been effectively implemented and remain important unfinished business.

In its first 15 years, the NQF has served as the consensus-based, standard-setting organization for quality measures in the United States. Using a rigorous evaluation and review process by multiple stakeholders, NQF-endorsed measures are considered the gold standard in measurement. The NQF's standing as a consensus-based, standard-setting organization allows the federal government to use endorsed measures preferentially in their programmes. In addition, each year since 2012, the NQF has advised the government on the selection of measures for use in more than 20 federal public reporting and pay-for-performance programmes. The NQF also helps to support measure alignment across public and private sectors. Increasingly, the NQF has taken on a leadership role in measurement science, with expert and consensus reports on complex and controversial issues in measurement, including adequacy of risk adjustment and strategies to link cost and quality.

Table 1 National Quality Forum measure evaluation criteria

Importance to measure and report (1st must-pass)
What is the level of evidence for the measures?
Is there an opportunity for improvement?
Scientific acceptability of the measurement properties (2nd must-pass)
What is the reliability and validity of the measure?
Usability and Use
Can potential audiences use performance results for both accountability and performance improvement?
Feasibility
Can the measure be implemented without undue burden, capture with electronic data/EHRs?
Assess related and competing measures
EHR, electronic health record.

The NQF uses standard evaluation criteria for the endorsement of measures that reflect desirable characteristics of quality measures and best practices in measure development and testing (Table 1). Over the years, the evaluation criteria have evolved into a hierarchy. Importance to measure and report reflects the greatest potential of driving improvement and resides at the top of the hierarchy. If a measure is not important, the other evaluation criteria are less important. This 'must-pass criterion' hones the evidence for the measure focus and gaps in care, with considerable variation or less-than-optimal performance demonstrated across providers and populations. The evidence criterion requires a systematic review or an assessment of the quality, quantity and consistency of the body of evidence for the focus. For these measures, the NQF requires a rationale that supports the relationship between the outcome and at least one process, intervention or service. The second most important criterion in the hierarchy is scientific acceptability of measure properties. The goal is to make valid conclusions about quality. If a measure is not reliable and valid, there may be a risk of misclassification and improper interpretation. The third criterion focuses on the usability and use of the measures. The goal is to use endorsed measures for decisions related to accountability and improvement. This criterion also considers whether the benefits of the measure outweigh evidence of unintended negative consequence to patients. Measure feasibility is also assessed to ensure that the measure causes as little burden as possible. If a measure is not deemed feasible, alternative approaches should be considered.

The NQF has a clearly stated preference for outcome measures, and the proportion of outcome measures has consistently grown to more than 30% of the NQF portfolio. However, measures in widespread use do not always address issues considered to have the greatest leverage for improving health and health care. For example, measures that focus on documenting in the medical record that clinical processes were performed (such as checking boxes on a discharge planning form) are less likely to result in meaningful improvement than measures that also capture whether these processes were performed appropriately. Patient-centred care frequently emphasizes 'nothing about me without me' as a core principle. Measures that only assess whether providers offered counselling or treatment choices without patient reflection on whether what was offered was understandable do not provide

meaningful information for improvement. Further, measures that demonstrate only that a clinician ordered an evidence-based treatment may not improve clinical outcomes if the patient does not adhere to the recommended therapy because there was no shared consideration of the patient's preferences and circumstances.

Evolution of healthcare quality measurement

The National Quality Strategy (NQS), introduced in 2011 by the Agency for Healthcare Research and Quality on behalf of the US Department of Health and Human Services, has set six clear improvement priorities to achieve better care, more affordable care and healthier communities [4]. This national strategy offers an organizing framework upon which to evolve healthcare measurement. The priorities include the following:

- 1** Making care safer by reducing harm caused during the delivery of care (e.g. central line-associated bloodstream infections);
- 2** Ensuring that all persons, and their family, are engaged as partners in their care (e.g. advanced care planning);
- 3** Promoting effective communication and coordination of care (e.g. patient experience of care);
- 4** Promoting the most effective prevention and treatment practices for the leading causes of mortality (e.g. primary percutaneous coronary intervention within 90 min of hospital arrival);
- 5** Working with communities to promote widespread use of best practices to enable healthy living (e.g. avoidable hospitalization for asthma);
- 6** Making quality care more affordable for individuals, families, employers and governments by developing and spreading new healthcare delivery models (e.g. total cost of care).

The National Healthcare Quality and Disparities Report highlights the NQS domains in annual chartbooks (Table 2). The most recent report highlighted the slow pace of improvement in these priority areas (Fig. 1). The rate of change varies across the domains, with a greater proportion of measures demonstrating improvement in person-centred care.

Table 2 Domains in the US National Healthcare Quality and Disparity Report, 2014

- **Access to Health Care** presents measures that cut across several priority areas and includes measures of health insurance, usual source of care, timeliness of care, and infrastructure to provide health care to minority and low-income populations.
- **Patient Safety** tracks safety within a variety of healthcare settings, including hospitals, nursing homes, home health settings and ambulatory care settings.
- **Person- and Family-Centered Care** examines individual experiences with care in an office or clinic setting, during a hospital stay and while receiving home health care. It tracks measures of perceptions of communication with providers and satisfaction with the provider–patient relationship.
- **Care Coordination** presents data to assess the performance of the US healthcare system in coordinating care across providers and services. It includes measures of transitions between healthcare settings and health information technologies that help to coordinate care.
- **Effective Treatment** is organized around care for the leading causes of mortality and morbidity in the United States: cancer, cardiovascular disease, chronic kidney disease, diabetes, HIV and AIDS, mental health and substance abuse, musculoskeletal diseases and respiratory diseases.
- **Healthy Living** examines healthcare services that typically cut across clinical conditions: maternal and child health, lifestyle modification, clinical preventive services, functional status preservation and rehabilitation, and supportive and palliative care.
- **Care Affordability** discusses costs of health care and tracks measures of financial barriers to care as well as misuse of healthcare services.
- **Priority Populations** summarizes quality and disparities in care for populations at elevated risk for receiving poor health care, including racial and ethnic minorities; low-income populations; children; older adults; residents of rural areas; lesbian, gay, bisexual and transgender populations; and individuals with disabilities, multiple chronic conditions or special healthcare need.

In recent years, the development and use of outcome measures, including patient experience and patient-reported outcomes (PROs), has grown. Using information about a patient's health, health condition or experience with care that comes directly from the patient, these outcomes can supplement clinically derived measures to provide a fuller picture of healthcare performance. In the United States, the family of Consumer Assessment of Healthcare Providers and Systems (CAHPS) survey tools has been effectively utilized across most healthcare settings. Increasingly, performance on these patient-experience-of-care surveys has been built into most federal payment and reporting programmes. Hospital CAHPS scores can be compared across US hospitals on a federal website, HospitalCompare [5], and the hospital value-based purchasing programme includes these scores in payment calculations. Some organizations are utilizing natural language processing and data-mining applications to gain actionable insights from free-text patient comments. Whilst there has been growth in the use of PRO measures, transition to the use of aggregated PRO data to assess quality of care by physicians, hospitals and

other providers remains limited. A federally supported measurement system, PROMIS, offers item banks of patient-reported surveys for adults and children. The NQF completed foundational work that outlined a pathway to move from PROs to PRO-based performance measures [6]. To date, there has been early uptake of primarily condition-specific (e.g. depression remission) and procedure-specific (e.g. change in physical functioning after total knee replacement) PRO-based performance measures.

Newly developed measures are increasingly taking advantage of the best data available for measurement, creating hybrid measures of clinical data from electronic health records, clinical registries and claims. This is an important step away from the historical US reliance on billing data that may fail to capture clinically relevant and patient-centred information. Inclusion of standardized data from laboratory results, imaging results and patient vital signs could lead to improved risk adjustment. Linkages are needed between emerging electronic health records and patient-reported data to ensure that measures are built on the best

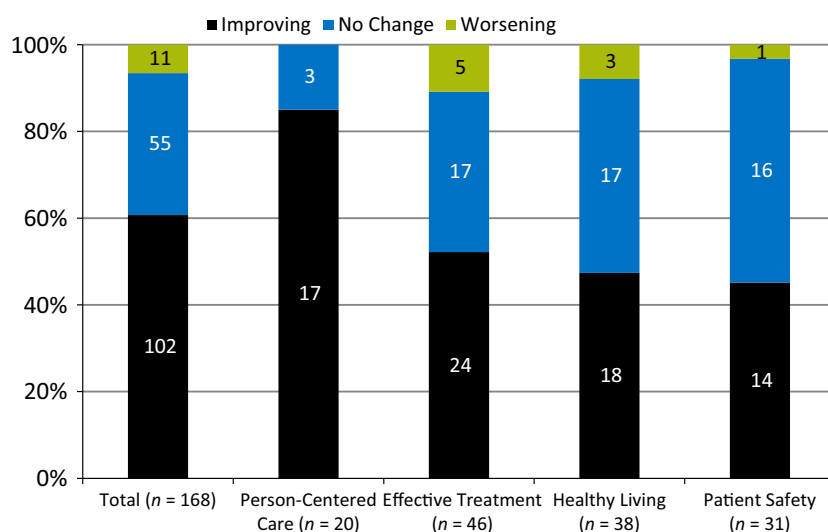


Fig. 1 Number and percentage of all US quality measures improving, not changing, or worsening through 2012, overall and by National Quality Strategy priority. Source: <http://www.ahrq.gov/research/findings/nhqrdr/nhqrdr14/key2.html>. Accessed 18 September 2015.

possible data sources. Advanced health systems have built innovative measurement approaches into their electronic systems. Unfortunately, these innovative measures have not been widely shared and few as yet have been adapted as national standards. An incomplete array of well-tested, practical measures exists to inform large-scale improvement efforts.

Increased use of robust national clinical registries provides great promise for improving patient outcomes, especially for high-volume surgical procedures. The Society of Thoracic Surgeons (STS) has developed comprehensive registries that have led to standardized measurement of clinical outcomes and contributed to marked reductions in cardiac surgery mortality and related complications. For example, the STS registry with feedback to cardiothoracic surgeons has reduced postoperative complications such as stroke and renal failure [7]. Similarly, the American College of Surgeons' National Surgery Quality Improvement Program and the American Heart Association's Get with the Guidelines programme have produced meaningful improvements through standardized measurement and benchmarking. In 2013, federal legislation authorized a Qualified Clinical Data Registry pathway for specialty societies to meet physician accountability requirements. The approved registries are required to provide timely feedback to participants to foster quality improvement. Whilst tremendous growth has been seen in the use of registries across a wide range of specialties, there have been accompanying challenges. Although many would prefer the use of high-quality data for

measurement, there has been limited transparency to date. With growing pressure from multiple stakeholders, including consumers and purchasers, some registries have moved towards public reporting of results [8]. There have also been concerns regarding the limited connectivity between clinical data registries and electronic health records and the potential for parallel development of these critical data sources. With a growing population of older patients with multiple chronic conditions, the ability to track patients across multiple registries is another important concern.

However, compared to other countries, in the United States the registry movement remains relatively undeveloped [9, 10]. Whilst the Get with the Guidelines initiative is a joint partnership between the US Centers for Disease Control and Prevention and the American Heart Association, the vast majority of registries are linked to physician specialty societies. There is nothing comparable in the United States to Denmark's 60 publicly financed nationwide clinical registries to track patient safety and quality. Further, the United States does not have a tradition of population-based registries. In fact, the United States has been late to embrace population health. A unique patient identifier has offered significant advantages in Denmark; however, highly politicized concerns regarding privacy have been a rate-limiting step in the United States. Although matching algorithms on multiple identifiers can achieve fairly high levels of patient matching, a unique patient identifier would offer many advantages. Support of national registries for quality measurement and bench-

marking has been limited to date, although recent efforts at the Patient-Centered Outcomes Research Institute (PCORI) have driven significant research collaboration and linked research databases (PCORNet). In addition, PCORI's Patient-Powered Research Networks are disease-specific research networks in comparative effectiveness comprised of activated patients, families and caregivers [11].

Although health and healthcare disparities are frequently raised as a national issue, little has been done to systematically include assessment of disparities in quality measurement. Whilst some evidence exists that large-scale improvements can result in improvements for all, the inability of routine measurement systems to capture differences by race, ethnicity, language, health literacy or insurance has hampered efforts to improve equity. In addition, large-scale efforts to improve quality may lead to improvement for underrepresented minorities, and the gap between groups often widens. To identify and reduce disparities, data should be stratified to highlight these differences and drive targeted improvement. Coupled with the need to better measure disparities is the growing concern about the potential unintended consequences of measurement and incentive-related efforts on safety-net providers.

The path forward for US healthcare quality measurement

As we look to the next decade, greater attention needs to be given to the development of measures that matter. The focus on the triple aim of effective care, affordable care and healthy communities in the NQS should help drive improvement for individuals and populations. With this expanded focus on the community, a set of core measures should be identified that can link to population health and can be considered important in achieving national priorities for health [12]. This focus on population health should also sharpen our efforts to identify and reduce disparities. To make significant progress, the US healthcare system needs to increasingly use data and information to change and improve actions and outputs over time. A recent report from the President's Commission on Science and Technology called for greater use of systems engineering approaches to drive further improvement in the healthcare system [13]. The IOM report, *Better care at lower cost. The path to continuously learning health care in America? But is this needed, perhaps: An Institute of Medicine Committee report also made a strong case for continuous learning and*

system improvement [4]. The US healthcare system needs to invest in rapid feedback on quality measures that would enable more continuous learning. Whilst measurement has been a powerful force in the US system, a closer connection to evolving national data systems and evidence-based improvement strategies is needed.

Conflict of interest statement

No conflict of interests to declare.

References

- 1 National Healthcare Quality and Disparities Report. 2014. <http://www.ahrq.gov/research/findings/nhqdr/nhqdr14/2014nhqdr.pdf> Accessed June 17, 2015.
- 2 Donabedian A. Evaluating the quality of medical care. *Milbank Mem Fund Q* 1966; **44**: 166–206.
- 3 McGlynn EA. Introduction and overview of the conceptual framework for a National Quality Measurement and Reporting System. *Med Care* 2003; **41**(Suppl 1): 1–7.
- 4 Report to Congress: National Strategy for Quality Improvement in Healthcare Improvement. Agency for Healthcare Research and Quality, March 2011.
- 5 Measures displayed on Hospital Compare. <https://www.medicare.gov/hospitalcompare/Data/Measures-Displayed.html>. Accessed September 10, 2015.
- 6 Patient-reported outcomes (PROs) in performance measurement. National Quality Forum, January 10, 2013. Washington DC.
- 7 El Bardissi AW, Aranki SF, Shubin S *et al.* Trends in isolated coronary artery bypass: an analysis of the Society of Thoracic Surgeons Cardiac Surgery Database. *J Thoracic Cardiovasc Surg* 2012; **143**: 273–81.
- 8 Consumer Reports and STS Public Reporting. <http://www.sts.org/news/consumer-reports-and-sts-public-reporting>. Accessed October 1, 2015.
- 9 Norgaard M, Johnsen SP. How can the research potential of clinical quality databases be maximized? The Danish experience. *J Intern Med* 2016; **279**: 132–40.
- 10 Emilsson L, Lindahl B, Koster M, Lambe M, Ludvigsson JF. Review of 103 Swedish healthcare quality registries. *J Intern Med* 2015; **277**: 94–136.
- 11 Fleurence RL, Beale AC, Sheridan SE, Johnson L, Selby JV. Patient Powered Research Networks aim to improve patient care and health research. *Health Aff* 2014; **33**: 1212–9.
- 12 Institute of Medicine Committee on core metrics for better health at lower cost. <http://www.iom.edu/coremetrics>. Accessed July 17, 2015.
- 13 Cassel CK, Saunders RS. Engineering a better healthcare system: a report from the President's Council of Advisors on Science and Technology. *JAMA* 2014; **312**: 787–8.

Correspondence: Helen Burstin, MD, MPH, Chief Scientific Officer, National Quality Forum, 1030 15th Street NW, Washington, DC 20005, USA.
(fax: 202-783-3434; e-mail: hburstin@qualityforum.org). ■