



2000 L Street, NW
Washington, DC 20036
888-275-7585
www.ncqa.org

Putting Reliable Health Care Performance Measurement Systems into Practice

*By Joachim Roski, PhD MPH; Vice President, Performance Measurement
Greg Pawlson, MD MPH, Executive Vice President*

Contents

Background	2
Introduction	2
Choice of Measures	3
Consensus-Building Processes	3
Data Collection and Verification	4
Data analysis and Reporting	5
Measure Maintenance	5
Context and Conclusions	6

Background

Since 1990 the National Committee for Quality Assurance (NCQA) has been at the forefront in measuring and reporting on the quality of health care in America. Measurement and reporting are integral to NCQA's accreditation programs for managed care and preferred provider organizations and for its physician recognition programs.

The effectiveness of the NCQA approach - producing documented improvements in the care that millions of American's receive and the resulting increases in longevity and quality of life - has contributed to the interest and demand for measurement in health care. At the same time, the current emphasis on 'pay for performance' has also encouraged greater and appropriate scrutiny of how we define and reward health care quality.

Because it is not practical or feasible to measure every element that might define quality in health care, NCQA focuses its efforts on areas of that offer the greatest potential benefit for the largest number of people.

Once an area for measurement is determined to be both a high priority and to have an impact, then the process of developing, refining and implementing actual measures truly begins.

Introduction

Measuring and reporting health care quality requires a complex combination of scientific rigor and technical competence in identifying suitable performance metrics, data collection processes and statistical analysis. In addition, performance measures must ideally be consensus-based if they are to have optimal impact in improving health care quality. NCQA has developed and uses an evidence-based consensus process by which such measures are developed, implemented and updated.

Since NCQA's pioneering initial efforts, performance measurement has become more established and health plans, physician specialty groups and others have entered the measurement development arena. The National Quality Forum (NQF) and Ambulatory Care Quality Alliance (AQA) have launched efforts to create consensus around, and in the case of AQA, to implement, a standardized set of measures. In addition, more than 100 different programs around the country rely on performance measurement results as the basis for pay for performance. In 2006, The Institute of Medicine issued several reports related to

quality measurement and to pay for performance. Finally, the Centers for Medicare and Medicaid Services (CMS) and Congress are actively moving the Medicare program in the direction of greater public reporting and pay for performance.

The enhanced activity in measurement, as well as the use of measures in pay for performance and other efforts to expand accountability in health care, increase the need to implement measures in a precise and standardized manner that provides accurate and meaningful results.

As we will see, careful implementation of performance measurement systems is as important as the development of an initial set of performance measures.

This NCQA Issue Brief describes key elements that are critical to successful implementation of clinical measures including 1) choice of measures, 2) data collection and verification processes, 3) initial pilot testing, 4) data analysis and reporting, and 5) maintenance and up-dating of measure specifications. It also describes major political and clinical practice considerations including both the levers needed to drive for more widespread implementation of performance measurement systems, and the incentives for measurement implementation. Finally it touches on the lack of evidence- and consensus-based performance metrics for many aspects, and levels of the U.S. health care system.

Choice of Measures

While choosing measures would appear to be simple, there are several critical factors that need to be considered. First and foremost is how the measures are going to be used. Some of the key questions include:

1. How will the measures be used? This can range from quality improvement within a single practice to widespread public reporting. While standardized, reliable, valid, well-vetted and widely used measures are useful even for practice quality improvement efforts, they are absolutely essential for public reporting efforts or other uses that extend outside a single practice site.
2. What are the goals of measurement? Is the goal to measure adherence to a single specific guideline or standard or to try to determine some overarching concept of “quality” of care. The scope of what is being

measured and reported will affect many other aspects of measurement as well.

3. How many measures are needed? The number of measures required must take into consideration the goals of measurement (the broader the concept the more measures needed), the desired level of accuracy and reliability, and the technical performance characteristics of the measures that are available (the more definitively linked to patient outcomes the measures are, the fewer that will be needed).
4. What data sources are available, usable and affordable? Theoretically data can be obtained from a number of sources including electronic records (claims data, pharmacy dispensing data, laboratory data, medical records), paper records generated by all physicians and hospitals, survey data from patients and systems or structural information from an on-site audit. However, the costs associated with obtaining the data may limit use to one or a very few sources such as electronic claims data.

Consensus-Building Processes

Perhaps the most critical step beyond choice of measures is the need to develop consensus around the selection of a set of performance measures. The need for consensus starts with the process of developing the measures and extends to implementation. Building support for a measurement program begins with choosing measures that have been developed and/or vetted via a transparent, consensus-based process. Transparency around measures is critical since there are frequent decisions involving trade-offs between accuracy and reliability and cost of data acquisition that are inherent in the development and specification for any performance measure. Some measure development and implementation processes have included only a narrow constituency (such as staff of a health plan or members of a medical specialty group) and often do not stand up to the scrutiny that is usually demanded in widespread use of a measure. Examples of major measure developers that include input and review by multiple groups and provide an opportunity for public comment include measures developed by NCQA, The Joint Commission and the AMA sponsored Physicians Consortium for Performance Improvement (PCPI). Relying on measures that have been endorsed by the NQF is also helpful, since NQF’s consensus process includes broad stakeholder input and comment. In the implementation phase, similar critical decisions must be made about sample size, number of measures, grouping of results, scoring and other parameters.

Critical Elements of a Robust Consensus Process

Some guidelines for gaining stakeholder input based on the processes in place at NQF and NCQA, and the recent set of criteria promulgated by the American Medical Association (AMA).

1. Oversight of measure development, consensus or implementation process by a committee or group including clinical and measurement experts along with informed stakeholders from the employer, public and private payer, consumer, health plan, physician, and health services research communities.
2. Opportunity for public comment (including from those likely to be measured at critical junctures).
3. Use of measures and implementation processes that are transparent (that is understandable).
4. Careful review and evaluation of public comments by both staff and oversight committees.

Data Collection and Verification

The next step in implementation is to ensure that the data collection process adheres to principles and procedures guaranteeing comparability of the data. This is most important for projects that extend across multiple entities. The basis, rationale and approach for verifying health care performance data need to be as reliable as the verification of financial information. Based on its experience and processes to ensure the ongoing integrity of the HEDIS[®] data set, NCQA's data collection and verification processes may serve as an illustration of a system ensuring not only the reliability and validity of data sources fueling performance measurement but also the data collection processes associated with a performance measurement implementation system.

1. **Pilot testing of measures:** Before implementing a new measure in any NCQA program, NCQA conducts a pilot test that provides an assessment of the feasibility of data collection. Such issues as the accuracy and completeness of data that is available from various sources, the sample size that can be obtained, and the ability of the data to be audited are essential to know prior to full

implementation. The testing is done in sites that are similar to those that will be included in the measurement program (health plan or physician office etc).

2. **Sample size determination:** Deciding on what sample size is optimal and/or possible is a technical as well as a practical task. While statisticians can provide estimates of what would be an adequate sample size (confidence intervals and risk of misclassification for various measures given different sample sizes etc), the final choice is usually a complex tradeoff between accuracy and the difficulty and cost of data collection. When performance measures are collected through a survey, an additional step is required to determine how many surveys must be mailed out to ensure a minimally acceptable number (response rate) for each survey item.
3. **Data transmission:** In order to minimize the possibility of errors in data transmission when compiling organizations' performance rates, standardized data transmission tools and protocols should be used. The reporting organizations should be given a chance to review and comment on their data well in advance of the reporting deadline. Electronic data submission tools, such as web-based systems, can ensure thorough automated logic checks (e.g., range check, minimum-maximum value check, etc.) reducing the opportunities for inadvertent errors in data transmission.
4. **Verification:** NCQA began to institute verification processes to ensure the comparability and integrity of reported data in 1995. The goals of the verification process are to ensure that the HEDIS production processes conform to the issued technical specifications of performance measures. During the verification process the capabilities of the information system and its ability to process medical, member, and practitioner information are also assessed.

It is also necessary to provide technical support to organizations that participate in a performance measurement system. This assistance can be provided via regular newsletters, telephone help-lines or virtual and in-person educational opportunities. In order to systematically track any specific issues and concerns around the implementation of a specific performance measure, a database should be established that compiles information on the number and types of questions for individual measures. Such information can subsequently be used to improve future publications of specification and data collection rules.

Separate quality control processes for survey-based performance measures are also important (see box).

Quality Control Best Practices for Survey Based Performance Measures

To ensure high quality surveys, processes must be in place to monitor and support adherence to survey protocols and the quality of survey administration:

1. Surveying organizations should receive detailed materials containing all survey specifications. In addition to written materials, survey administration expectations can be conveyed through annual training of organizations involved in survey data collection.
2. If multiple organizations are involved, a verification process should be established to ensure that all organizations are following technical specifications for fielding the survey in order to assure comparability.
3. Subsequently surveying organizations must ensure that staff:
 - Monitor practices for survey mailings (e.g., checking every 25th piece of mail before it goes out)
 - Establish other quality control processes (e.g., monitoring phone interviews for accuracy and tone).
 - Adhere to confidentiality practices (e.g., assigning unique identifiers, use of passwords and firewalls).
 - Provide appropriate customer support to organizations on whose behalf surveys are fielded
 - Assure appropriate data entry processes (e.g., 100% validation for key entry of data).
 - Edit and clean data (e.g. checking for out of range values, misuse of unique identifiers).

Data analysis and Reporting

Before verified and accurate data have reached the organization responsible for compiling the data, standard processes should be in place that ensure that the data analysis processes optimally support the intended use of the measures (e.g., public reporting, performance feedback). It is helpful to initially create mock tables that array sample data in the desired format. Several

reporting formats supporting various uses of the data may be desirable for the same data.

Following completion of descriptive analysis, additional analyses to better understand performance variation, trends and associations should be conducted. Such analysis can assist in identifying underlying performance trends and explanations (e.g., performance differences by region, patient age or socio-economic status, correlation of targeted clinical processes or outcomes with organizational or other variables) and might subsequently inform quality improvement activities tailored to the issues uncovered through exploratory data analysis.

Measure Maintenance

Once performance metrics have undergone development, testing, and consensus approval and have begun to be deployed, it is critical that they be maintained, updated and improved in an ongoing fashion. Critical elements include the updating and necessary refinements of data element specifications and collection rules in addition to the periodic in-depth, and ad-hoc evaluation of the merits of the performance measure. While a full review and reevaluation of a measure might take place only every few years (NCQA uses a three year cycle), some revisions, such as those involving drug or procedure codes, may be required on a much shorter time interval of weeks or months. NCQA maintains technical panels, including persons with expertise in pharmacy, laboratory and other fields, as well as a HEDIS users group, which provide constant feedback and input to measure maintenance activities.

Refinement of data elements and data collection rules: Over the course of the implementation of performance metrics it is likely that the required data elements for the measure as well as the data collection rules will need to be periodically updated between data collection cycles. These updates typically do not affect trendability of the performance data but offer the opportunity to measure performance with increased validity.

If performance measures rely on the use of administrative codes (ICD-9/10; CPT-4), or involve pharmacy data, the required and acceptable codes may need to be updated as new drugs or procedures become available or providers' coding practices change for the conditions tracked by a performance measure.

Clinical measure logic review: In addition to regularly reviewing and refining data elements and collection rules for each performance measure,

performance metrics should be completely reevaluated on a regular basis (e.g., every three years) and, if warranted, ad-hoc basis with respect to the overall evidence-base supporting the measure.

Newly available clinical evidence may continue to support the measure as specified, suggest minor or major changes to the specifications or suggest discontinuation of the measure and potential substitution with a different measure. Critical data to review also includes current and time-trended performance levels for the measure. Clinical evidence typically accumulates incrementally over time in support of particular clinical practices. However, at times important and compelling clinical evidence may be published altering clinical practice overnight and requiring major revision or even withdrawal of a measure.

Context and Conclusions

NCQA has created a successful model for both development and implementation of health care performance measurement using a consensus-driven and transparent process that actively engages those being measured. A good deal of attention has been focused on developing measures based on strong scientific evidence and in areas that have the greatest impact on the population in terms of both health costs and implications. The current issue brief focuses on the equally important issues of implementation of measures in a manner which will lead to the creation of information on health care that is accurate, reliable and valid. We especially emphasize the need for collaboration with all stakeholders including those being measured. Even the most carefully crafted and tested measures will yield data that is inaccurate and misleading, or a measurement program that is unsuccessful, if careful attention is not given to implementation.



National Committee for Quality Assurance

Washington, DC based NCQA is a private, non-profit organization dedicated to improving health care quality. NCQA accredits and certifies a wide range of health care organizations, recognizes physicians and physician groups in key clinical areas and manages the evolution of HEDIS®, the tool the nation's health plans use to measure and report on their performance. NCQA is committed to providing health care quality information through the Web, media and data licensing agreements in order to help consumers, employers and others make more informed health choices. For more information, visit www.ncqa.org.