Chapter 3
Quality Measurement

Linda Harrington

Executive Summary

Quality measurement is a form of evaluation. It is fundamental for any attempt to improve the quality of healthcare. It is necessary for demonstrating current performance, setting goals, and validating achievement of those goals. Quality measurement doesn't have to be overly cumbersome, but it needs to be on target, measuring those things that matter, and make a difference in healthcare improvement.

The history of measurement of healthcare quality parallels the history of epidemiology and statistics in the late nineteenth and early twentieth centuries and is intertwined with the evolution of health services research in the late twentieth century. Quality measurement is now moving center stage in the twenty-first century as healthcare moves to a value-based system and the promises afforded through the massive amounts of electronic clinical data being accumulated. This chapter will focus on concepts necessary for the practical application of measurement in quality improvement (QI), highlighting challenges and opportunities related to the digital world.

Learning Objectives

Upon completion of this chapter, readers should be able to:

- Describe the historical evolution of the science of quality measurement
- Compare the characteristics of structure, process, and outcome measurements
- Construct appropriate measurements for QI projects
- Identify the necessary characteristics of quality measures, including reliability and validity

L. Harrington (✉)
Baylor College of Medicine, Houston, TX, USA

A. P. Giardino et al. (eds.), *Medical Quality Management*,
https://doi.org/10.1007/978-3-030-48080-6_3

- Evaluate the success of QI projects
- Describe challenges and solutions associated with electronic data
- Identify upcoming trends in the science of quality measurement

Introduction

Measurement is very familiar to physicians. They routinely interact with different types of measurement in research, in practice, and increasingly in quality improvement. With all three types of measurement, discovery occurs when data are collected and analyzed.

Measurement is performed in research to discover new knowledge. Similarly, clinical measurement is done to determine a patient's wellness or illness. The goal of quality measurement is to discover the structures and/or processes of care that have a demonstrated relationship to positive health outcomes and are under the control of the healthcare system [1].

The objectives of healthcare quality measurement that have passed the test of time are to:

- Provide data to inform quality improvement efforts
- Inspect and certify that an organization or individual meets previously established standards
- Compare groups for a variety of purposes, including selective contracting by purchasers and choice of providers and practitioners by individuals
- Inform patients, families, and employees about the healthcare decisions and choices they face
- Identify and possibly eliminate substandard performers—those whose performance is so far below an acceptable level that immediate actions are needed
- Highlight, reward, and disseminate best practices
- Monitor and report information about changes in quality over time
- Address the health needs of communities [2]

Quality measures must be consistent in how they perform and valid in measuring what is intended to be measured. Measurement in research must be reliable and valid, otherwise threatening the internal validity of a study and as a result generalizability of the findings. Similarly, clinical measurement must be reliable and valid to accurately diagnose patients and evaluate progress.

Quality Measurement Framework

The most influential framework for guiding quality measurement development is from the Institute of Medicine [3] and focuses on the following six aims:

Safe: Avoid harm to patients from the care that is intended to help.

Effective: Provide services based on scientific knowledge to all who could benefit and refrain from providing services to those who aren't likely to benefit, avoiding underuse and misuse, respectively.

Patient-centered: Provide care that is respectful of and responsive to individual patient preferences, needs, and values, ensuring that patient values guide all clinical decisions.

Timely: Reduce wait times and sometimes harmful delays for both those who receive and those who give care.

Efficient: Avoid waste, including waste of equipment, supplies, ideas, and energy.

Equitable: Provide care that does not vary in quality because of personal characteristics such as gender, ethnicity, geographic location, and socioeconomic status [3, 4].

Desirable Characteristics of Quality Measurement

The Agency for Healthcare Research and Quality (AHRQ) has put forth the following as characteristics of effective quality measurement [5]:

Standardization: Quality measures are standardized at the national level, which means that all healthcare providers are reporting the same data in the same way.

Availability: Data will be available for most healthcare organizations being profiled.

Timeliness: The quality measure allows for results to be available by distribution of a report when it is most needed by consumers.

Experience: Use of the quality measure demonstrates that it measures actual performance, not shortcomings in information systems.

Stability: The quality measure is not scheduled to be eliminated or removed from a measurement data in exchange for a better measure.

Evaluability: Quality measures allow for the results to be evaluated as either better or worse than other results, in contrast to descriptive information that merely demonstrate differences. An example is complication rate, where a lower rate is always better.

Distinguishability: Quality measures reveal significant differences among healthcare organizations or other comparators.

Credibility: The quality measure is audited or does not require an audit.

Relevance: The quality measure should be relevant to consumers, providers, clinicians, payers, and policy-makers and should be of interest or value to the stakeholders and the project at hand.

Evidence-Based: Quality measures, especially those related to clinical issues, should be based on sound scientific evidence. Measures should clearly link structure or process to outcomes.

Reliability: Reliability is the degree to which the quality measure is free from random error. Measurement indicators and data collection techniques must be stable enough to justify the use of the collected information to make a judgment about quality. The same measurement process using the same data should produce the same results when repeated over time.

Validity: Validity of a quality measure refers to the degree to which the measure is associated with what it purports to measure. A key question to be answered is whether the measures selected to indicate the presence or the absence of quality actually represent quality in patient care.

Feasibility: Quality measures should be realistic and practical to collect and analyze. Measures that require too much time, money, or effort to collect may not be feasible to use.

History

The science of quality measurement is commonly recognized to have originated in the work of Florence Nightingale and her reports to the British parliament on mortality rates in British field hospitals during the Crimean War. Her early efforts to quantify healthcare were coupled with the birth of modern concepts of infection control, giving credence to the idea that measurement is needed for improvement [6].

Several decades later, Ernest Codman linked the interest in mortality to invasive procedures [7]. His exploration of postsurgical mortality can be considered the start of investigations into hospital outcomes. Codman is also credited with the notions that hospitals should have organized medical staffs and records of patient care—essentially the birth of structural measures.

The next major development in the measurement of quality occurred outside of healthcare in the interval between the World Wars. Walter Shewhart developed a branch of new statistics called statistical process control while working on the manufacture of telephones [8]. Perhaps his most important contribution was a change in the focus of measurement from the quality of products themselves to the steps required to produce those products. His other major contribution was a method to identify shifts in the manufacturing process that were statistically meaningful, a method that was simple enough to be implemented easily by individuals who did not have advanced scientific training. These shifts took the form of batches of product that differed significantly from other batches.

Arguably, the most influential contribution to measurement of healthcare quality occurred in the early 1960s when Donabedian began to explicitly differentiate the quality measures related to structure, process, and outcome [9]. The relative importance of process and outcome measures is still a subject of discussion in the contemporary literature, and Donabedian's general framework remains the dominant paradigm. Together, the three measures provide the best and most complete picture of quality.

In the last decade of the twentieth century, considerable efforts were made to implement approaches to quality measurement from industries outside of healthcare. One such approach is to apply measures of process and outcome to multiple domains of value across the organization. Kaplan and Norton originally advocated this approach in the information technology industry [10]. Batalden and Nelson brought it to healthcare in the very practical form of scorecards and dashboards [11]. Embedded in measurement tools is the understanding that the consideration of individual measures alone can be misleading, if not dangerous, because healthcare is a complex system subject to unintended consequences and that multiple perspectives (e.g., patient, provider, payer) must be considered in the design of a useful measurement system.

One of the most significant changes in quality measurement that has occurred in more recent history is the movement of healthcare from a volume-based model to a value-based model. Physicians are now being evaluated and reimbursed for quality, elevating the importance of quality measurement. The simultaneous implementation of electronic health records (EHRs) affords the opportunity to improve the collection and reporting of quality measures. Yet, significant challenges exist.

While EHRs have created improvements, such as electronic prescribing and increased access to patient records from any wired location, they have also created new challenges relevant to quality measurement, specifically the data [12]. EHR data are being degraded with preconfigured data entry aids such as documentation templates, macros, smart phrases, default text, copy-paste and copy-forward functionality [13], as well as overuse of alerts resulting in alert fatigue and overrides. These impact the characteristics and value of quality data.

Prior to EHR data, claims data were often used in quality improvement. In 2007, Tang et al. published a study funded by the Centers for Medicare & Medicaid Services comparing methodologies calculating quality measures based on administrative data, sometimes referred to as claims data, versus clinical data derived from EHRs [14]. The researchers sought to understand whether there was a significant difference in identifying patients with diabetes mellitus and associated quality measures between the more granular or detailed clinical data and claims data that are aggregated from clinical data. Using a random sample of 125 patient records, the researchers found 75% of diabetic patients were identified by manual review of the EHR, considered the gold standard at the time, compared to 97% of diabetics identified using queries of the EHR. As a result, there was a statistically significant difference in the quality measures for frequency of HbA1c testing (97% vs. 68%, $p < 0.001$), control of blood pressure (61% vs. 45%, $p = 0.05$), frequency of testing for urine protein (85% vs. 55%, $p < 0.001$), and frequency of eye exams for diabetic patients (62% vs. 41%, $p < 0.03$) [14].

It is important to note that claims data today are largely derived from EHR data. In fact, many healthcare professionals suggest that EHRs were designed first and foremost to capture data for the purposes of submitting claims. As such, claims data derived from EHRs are subject to the same issues of data quality mentioned previously. Thus, the quality of data used in QI is an ongoing challenge.

While organizations struggle to ensure the quality of data and thereby quality measurement, there is an ongoing increase in the number of quality measures expected of physicians that are inefficient and imbalanced [15]. There are measures that are duplicative, such as multiple measures for the same condition on follow-up care that use different time reporting requirements. Some quality measures overlap one another, such as diabetes composite measure and separate hemoglobin A1c measure. The ease in capturing some data and difficulty with other data has resulted in an overrepresentation and underrepresentation of measures. For example, there are several measures related to childhood immunizations and few related to chronic care. Given the evolution in quality measurement and the challenges being faced, there are some fundamental concepts in the types of quality measures that have sustained.

Types of Quality Measures

The gold standard for defining quality measurement remains Donabedian's three-element model of structure, process, and outcome [9].

Structural Measures

Structural measures relate to the ability of an organization to provide high-quality care associated with a healthcare setting, including its design, policies, and procedures. The underlying assumption is that healthcare organizations that have the necessary quantity and quality of human and material resources and other structural supports are best prepared to deliver quality healthcare. Examples of structural measures include the availability of appropriate equipment and supplies in a hospital setting and the education, certification, and experience of clinicians in an institution. Structure-focused measures often are easy to access. Healthcare organizations routinely maintain data on equipment and supply inventories, staffing, patient acuity, and staff qualifications.

Process Measures

Process measures are more often referred to today as *performance measures*. They are used to evaluate if appropriate actions were taken for an intended outcome and how well these actions were performed to achieve a given outcome. The underlying clinical assumption is that if the right things are done right, the best patient outcomes are more likely to occur [16]. An example of an evidence-based process measure to assess the quality of care for a patient with acute myocardial infarction is the proportion of patients admitted with this diagnosis (without beta-blocker contraindications) who received beta-blockers within 24 hours after hospital arrival.

Four criteria for successful process measures include the following:

- A strong evidence base demonstrating that the care process represented by the process measure leads to improved outcomes.
- The process measure accurately represents whether the evidence-based care process was provided.

- The measure addresses a process that has few intervening care processes that must occur before the improved outcome is realized.
- Deploying the process measure has little or no chance of introducing unintended adverse consequences.

Process measures that meet all four criteria are most likely to improve patient outcomes [17].

Example: Process Measure for Diabetes

Percentage of patients whose hemoglobin A1c level was measured twice in the past year

Outcome Measures

Outcome measures seek to capture changes in the health status of patients following the provision of a set of healthcare processes and include the costs of delivering the processes. The patient is the primary focus, and outcome measures should describe the patient's condition, behavior, and response to or satisfaction with care. Outcomes traditionally are considered results that occur as a consequence of providing healthcare and cannot be measured until the episode of care is completed. Episodes of care may include hospitalizations, physician office visits, or care provided in post-acute care settings. For example, to assess the quality of care for patients with acute myocardial infarction admitted to a coronary care unit, the outcome measures may be related to incidence of reinfarction and patient satisfaction with the care received in the unit.

Outcome measures provide an indirect measure of the overall quality of an organization and can provide trending and benchmarking opportunities to demonstrate progress. On the other hand, outcomes can be influenced by factors that are not measured or are beyond the control of clinicians, such as genomics, case mix, and socioeconomic or environmental influences. As a general rule, the more structure and process variables a QI project employs, the greater the reliability of outcome measures [18].

Historically, quality measurement has focused primarily on outcomes. Today, structure and process measures provide important insights, illuminating which areas to address to improve outcomes. Structure and process provide direct measures of quality and thus yield more sensitive measures of quality, which can direct clinicians to the most effective ways to improve patient care. To be valid, however, structure and process must be empirically related to outcomes and be able to detect genuine differences in patient care. To maintain validity, they also must continually be reviewed and updated in accordance with current science (i.e., evidence).

Example: Outcome Measure for Diabetes

Average hemoglobin A1c level for population of patients with diabetes in the past year

Bundled Measures

The Institute for Healthcare Improvement (IHI) developed the concept of "bundles" to assist healthcare providers in providing reliable care for patients undergoing specific treatments with known inherent risks [19]. A bundle is a package of interventions that must be applied for every patient every time. The power of bundles is related to the underlying evidence.

Example: Ventilator Bundle Measures

> *Process measure*: Percentage of intensive care patients on mechanical ventilation for whom all four of the ventilator bundle interventions are documented.
> *Outcome measure*: Number of ventilator-associated pneumonias per 1000 ventilator days.

Balancing Measures

Balancing measures are important for illustrating whether unintended consequences are introduced during quality improvement initiatives. For example, a balancing measure of hospital readmissions is important when the goal is to reduce length of stay. Another example of a balancing measure is the incidence of hyperglycemic episodes in intensive care when trying to reduce the number of hypoglycemic episodes.

Benchmarking

Quality improvement plans often include *benchmarking*, an effort to determine the current status of quality and compare it to the highest performers internal to an organization or external to the organization (e.g., comparing performance with competitors) [20]. An Achievable Benchmark of Care (ABC), as identified by Kiefe et al., is produced by benchmarks that (1) are measurable and attainable, (2) are based on the achievements of the highest performers, and (3) provide an appropriate number of cases for analysis [21].

Constructing a Measurement

Quality measures are constructed in several ways including proportions or percentages, ratios, means, medians, and counts [22]. The choice of measure depends on the goal trying to be achieved.

Proportion

Comparisons of quality measures within systems and across providers require standards for how quality measures are expressed. The generally accepted standard for the expression of quality measures involves a numerator and a denominator. The numerator describes the desired characteristics of care, and the denominator specifies the eligible sample. For example, in the treatment of heart failure patients, the numerator for one possible proven measure is the number of people who actually receive beta-blockers, and the denominator is the number of people who are eligible to receive beta-blockers. Together, the numerator and denominator provide a measure of insight into the quality of the treatment of heart failure with beta-blockers.

Ratio

Ratio measures are often structural measures denoting the capacity of a healthcare provider. Examples include the ratio of providers to patients, hospitalization per 10,000 residents of a targeted area, staff-to-patient ratio, or percent of heart failure patients tracked in a registry. Interpretation can be tricky as sometimes higher values are better, whereas at other times lower values are the goal.

Mean and Median

Means and medians are often used to measure processes. A commonly used example would be the median time an eligible patient arrives in the emergency department to the administration of fibrinolytic therapy. This measure captures timeliness of treatment when time is an important factor in the outcome of care. Like ratios, the interpretation of means and medians must be carefully examined as high values or low values may be good or bad depending on the context.

Count

Counts are quality measures often seen in surveillance such as the investigation of adverse outcomes. Examples include foreign body left in surgical patients or transfusion reactions. They cannot be used for comparison purposes across providers as they lack specification of the population.

Several factors should be considered when constructing a quality measure: the age of the persons included, the measurement period, the system or unit being examined, and whether the measure will be within a program of care, across an entire healthcare setting, or local, or national should be identified and taken into account [23].

Timing of Measurements

Baseline Measurement

Almost all quality improvement processes, projects, or programs begin with the measurement of quality in its current state, which is known as a *baseline measurement*. Baseline measurements use many types of quantitative and qualitative data as indicators and allow a supporting analysis and an eventual judgment to be made about the status of medical quality at that point in time.

In Table 3.1, a baseline assessment is shown for a group of patients with diabetes whose hemoglobin A1c levels were evaluated in Year 1. This evaluation was used to design a QI project and to determine the change in hemoglobin A1c levels after 1 year of intervention.

The drawback to baseline measures is that they provide snapshots of measured characteristics of structure, process, or outcomes at one point in time. Measurement at another time can only be interpreted as higher or as lower than baseline and does not indicate actual or sustained improvement. Measurement tools that allow for trending are discussed in the next section.

Table 3.1 Baseline assessment: hemoglobin A1c levels at baseline and after a 1-year intervention for patients with diabetes

Member interventions		
Applied program		
Stratification of diabetic population		
Special needs case management		
Outreach activities and education		
Referrals to employer program		
Provider interventions		
Contacted physician and coordinated information		
Sponsored a physician education program		
Member outcomes		
Improved diabetes control		
Lowered hemoglobin A1c		
Direct cost savings		
Reduced hospital readmission rate for diabetes		
	Hemoglobin A1c levels	
	Year 1	**Year 2**
N	212	212
Median	7.30%	7.10%
Average	7.62%	7.39%
% of patients with values <7.5%	54.7%	60.8%
% of patients with values >9.5%	16.5%	11.3%

Trending Measurements

Run Chart

A *run chart* is a quality tool used to identify trends by measuring changes in structure, processes, or outcomes over time. The run chart is created in an XY graph in which the x-axis represents time and the y-axis represents the aspect of the structure, process, or outcome being measured. A central line, if used, indicates the median of the data.

A *run* consists of consecutive points below or above the central line indicating a shift in the structure, process, or outcome measure being examined. A *trend* is a steady inclining or declining progression of data points representing a gradual change over time. Figure 3.1 provides an example of a run chart measuring length of stay over time.

This run chart shows a decreasing trend in length of stay, which suggests that interventions targeting a reduction in length of stay may be effective, assuming average daily census and patient acuity have remained similar over time. Run charts provide ready information on runs and trends in structure, process, and outcomes and are easy to construct and interpret. For more statistical power, control charts are preferred.

Control Chart

Control charts are most often used with process measures and are a more sensitive tool than run charts. The focus is on process variation. Additional features include a central line composed of the mean value of the data and upper control limits (UCL) and lower control limits (LCL) typically representing three standard deviations from the mean.

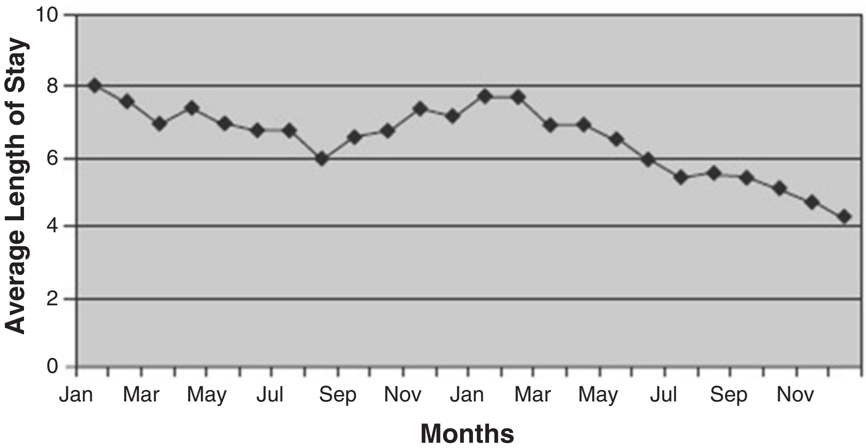

Fig. 3.1 Example of a run chart for average length of stay

A *statistical control chart* is a graph that represents the continuous application of a particular statistical decision rule to distinguish between normal and abnormal variations. Figure 3.2 shows a statistical control chart for the number of visits per day for a provider organization and covers each day. The threshold is the point at which intensive evaluation or action is taken.

Case Study • • •

Cardiac Services: Dartmouth-Hitchcock Medical Center

The cardiac services unit at Dartmouth is one of the pioneers in contemporary approaches to measurement and improvement of healthcare quality. In their work, measurement has been used as a central tool for tracking and improving care [24]. They have argued persuasively that measurement of clinical process and outcome must be controlled by the clinicians delivering care. Several key principles defined their approach.

Clinicians were involved in the design of a panel of measures that were both useful to them in their daily practice and useful to administrators and external stakeholders. This panel encompassed the entire process of care and contained a balanced set of cost and quality measures. Patient-centered measures (e.g., satisfaction, functional status) were incorporated along with other traditional measures of process and outcome (e.g., mortality, morbidity). Details concerning variations were presented, as were aggregate measures over time. In addition, current variation was evaluated against historical performance using statistical control charts.

Data for the project were obtained by chart abstraction in the perioperative period (i.e., at 3 weeks after surgery for satisfaction, at 6 months after surgery for functional status). Process variables were obtained in real time. The SF-36 indices of physical functioning, role functioning, bodily pain, and general health were used for the functional status measures. Among the measures of the surgical process were pump time, percent returning to pump, percent reexplored for bleeding, and internal mammary artery usage. Control charts were used with the surgical process data.

Control charts also were used for early detection of quality issues, allowing for near real-time correction. For example, the team was able to detect an increase in sternal wound infections by using a technique called a "successes between failures" chart to identify infrequent events and differentiate them from chance occurrences. This control chart allowed the team to decide if the increase in infections was due to random variation or a process shift. Because they used real-time data, they were able to quickly identify the process change related to this increase in infections and to correct it. Conventional methods usually result in delayed identification and more adverse events before solutions are found.

The results from this initiative are striking, although they cannot be attributed to measurement alone. Coronary artery bypass graft-related mortality dropped from 5.7% to 2.7% in a 2-year period; the average total intubation time decreased from 22 hours to 14 hours; and the number of patients discharged in fewer than 6 days increased from 20% to 40%.

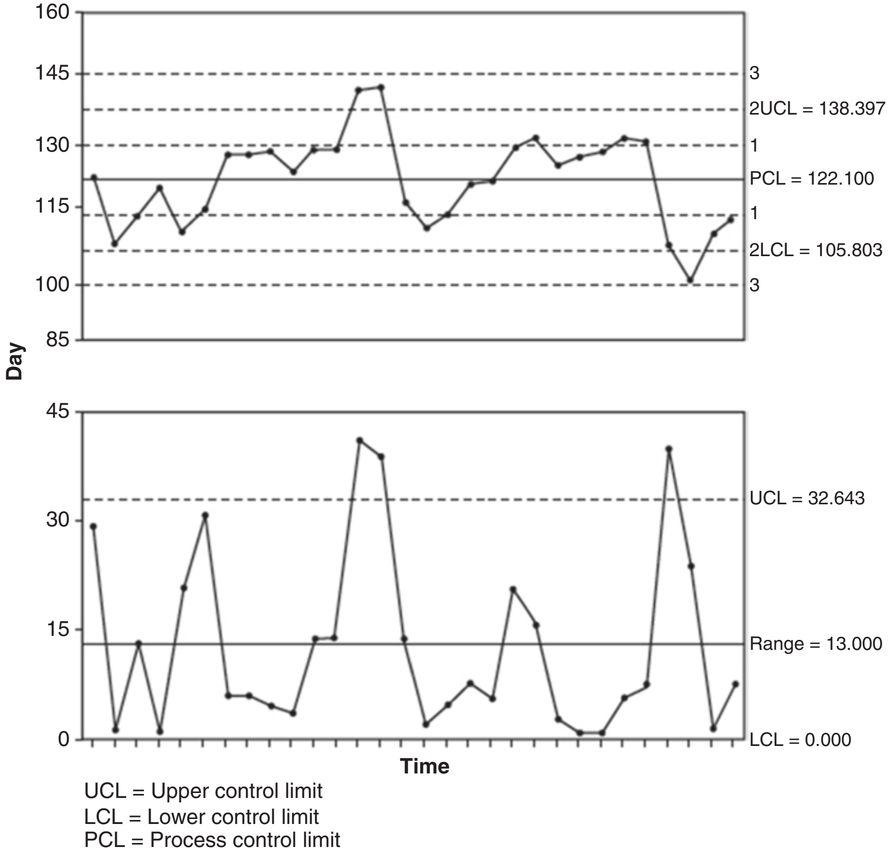

Fig. 3.2 Example of statistical control charts showing visits per day for a physician group

Interpreting Quality Measures

Appropriateness Model

There are many ways to interpret quality measures. The IHI advocates an "all-or-none" approach or *Appropriateness Model* to generate composite scores [22]. For example, if a patient with diabetes is expected to have a laboratory test, an eye exam, and a foot exam, failure to do any of these would result in failure of the composite measure of preventive diabetes care. The score reported reflects the proportion of patients who receive all the care recommended for them.

The AHRQ uses the Appropriateness Model to arrive at composite scores and to produce a comprehensive overview of the quality of care delivered in the United States. Composite measures based on this model are an increasingly large component of the report. AHRQ has chosen the Appropriateness Model because it reflects the philosophy that all citizens must receive all of the care that meets a high standard of evidence.

The 70% Standard Model

A variation of this method sets the threshold at less than 100%, usually at 70% (*the 70% Standard)*. Although the all-or-none approach of the Appropriateness Model strives for perfection (and consequently results in lower scores than another method using the same dataset), this approach and the 70% standard are sensitive to the number of indicators included in the composite.

Opportunity Model

Another common approach is the *Opportunity Model* where the number of opportunities to deliver care is summed to create the denominator and the number of cases in which indicated care is delivered is summed to create the numerator. The resulting percentage reflects the rate at which indicated care is delivered without penalizing some appropriate activities for the omission of others. This approach has been adopted by the CMS to reward hospitals for high performance (via pay for performance) in the Premier Hospital Quality Incentive Demonstration Project and internally in the Department of Veterans Affairs.

Program Evaluation

Program evaluation is necessary to measure the overall success of QI programs or projects and is usually conducted using two methods: formative evaluations and summative evaluations.

Formative Evaluations

Formative evaluations involve routine examination of data on program activities and provide ongoing feedback about components of the program that work and those that require intervention. *Dashboards* and *scorecards* are tools used in formative evaluations to track and trend quality improvement activities on a monthly basis. They highlight key quality improvement initiatives and identify successful progress, thereby allowing for timely intervention as necessary. For example, the use of a dashboard for critical care may report monthly compliance with a ventilator-associated pneumonia bundle.

Summative Evaluations

Summative evaluations are more formal and occur less often than formative evaluations, typically annually. Their focus is on measuring and determining the outcome or the effectiveness of the quality improvement program. The information evaluated is used to make decisions about the program, such as the need for more resources or education or perhaps better communication.

Effective program evaluations, whether formative or summative, are those that provide actionable information to program participants and management. Synthesis and use of information gleaned from program evaluations promote the continuous development of the quality improvement program.

One aspect of program evaluation that should not be overlooked is the ongoing assessment of the costs of quality measurement [25]. These include both direct costs associated with quality measurement operations and variable costs related to specific measures. Understanding the costs of quality measurement enables organizations to prioritize measures, understand the magnitude of costs, and ideally spur innovation in cost-effective quality measurement.

Quality Measurement in the Digital Age

Innovations in the digital age will continue to change quality data, measurement, and analyses. Data are largely digitized, albeit with an abundance of manual data entry, especially as it pertains to EHRs. Types of quality measures, specifically Donabedian's three-element model of structure, process, and outcome, will conceptually continue but will evolve operationally as they become more technologically based [26].

Quality measurement will become increasingly applied in real time, affording more rapid analyses and improvement. This will be enabled by the expanding use of sensors and data captured electronically [27]. Advances in analytics will afford faster and more powerful analyses of digitized data that are big, dark, and deep [28].

Big Data

Big data refers to "data whose scale, diversity, and complexity require new architecture, techniques, algorithms, and analytics to manage it and extract value and hidden knowledge from it" [29]. The goal of big data is to gain new insights and improve decision-making [30].

Big data involves what is commonly referred to as the 4Vs, i.e., volume, variety, velocity, and veracity [31]. Volume is defined as the vast amount of data generated every second. How much data constitutes big data is currently undefined. Variety refers to the range of data types and sources. This can include structured and unstructured text, images, numbers, and signals such as those from sensors. Velocity is defined as the speed at which data are generated. Lastly, veracity is the accuracy and reliability of data [31].

Dark Data

Dark data include all the unstructured data gathered in healthcare. These include text, images, audio recordings, as well as signals for wearable sensors, biometric data, retinal scans, and more. Dark data represents 80% of all data generated and is predicted to increase to 93% by 2020 [32]. Advances in computing power allow for increasing use of dark or unstructured healthcare data.

Deep Data

Deep data involve large amounts of data collected per patient [33]. These data provide a more complex view of patients. Value is derived through the time stamp and context of the deep data.

EHRs today do not offer big, dark, or deep data. While they have volume, EHRs lack the variety and velocity, slowed largely by manual data entry. Veracity is also an issue [34, 35]. Big data in healthcare will be uncovered with the addition to EHR data of genomic data and patient-generated data that is not episodic interactions with the current delivery system but instead involves lifestyle data. It is not until healthcare moves away from transactional data to a more dimensional, non-transactional data model will we receive better information on performance and better support for decision-making and quality improvement.

Future Trends

We believe that quality measurement as a science will be the future [36]. A convergence of factors supports the need for increased rigor in quality measurement, including ongoing issues in the delivery of quality patient care, pay for performance, and growing consumer awareness. The desire to improve the rigor of measurement parallels the need to improve quality and safety in patient care. Timely acquisition and analysis of sound data through the increasing use of information systems and the use of reliable and valid measurement tools are essential. Rigorous quality measurement promotes the generalizability of findings in quality improvement initiatives, expanding their usefulness to the larger patient population.

CMS's pay-for-performance reimbursement strategy uses quality measurement to reward providers and practitioners for complying with evidence-based standards for providing patient care. By rewarding quality, the hope is that compliance with new efficacious treatments will increase and clinical outcomes will improve. Chapter 9 will provide more details on CMS's pay-for-performance strategies.

We believe that in addition to payers, consumers will drive improvements in quality measurement. Consumers are increasingly interested in healthcare delivery, especially as they assume greater responsibility for the cost of care, through increasingly higher premiums, deductibles, and co-pays and become more active in monitoring their own health through telehealth, mHealth (mobile health), and other approaches.

Anything can be measured. How well something is measured is another issue. The caveat to the adage "Any data are better than no data" is the reality that bad data are worse than no data [37]. Bad data misclassify physicians and hospitals, provide misinformation to healthcare consumers, and waste time and resources.

The goal is to measure quality by focusing on the right structure, process, and outcome measures that are relevant, meaningful, important, evidence-based, reliable, valid, and feasible. Improving the quality of data being input into EHRs while increasing automation of data capture is critical to more timely analyses, prevention, and earlier intervention.

Lastly, increasing reporting requirements from multiple agencies has resulted in quality measurement becoming untamed [38]. According to Don Berwick, we need to reduce excessive measurement by 50% [38]. The goal is clear—make quality measurement meaningful.

Acknowledgments Harry Pigman, MD, MSHP, for his key contributions to the prior edition of this work.

References

1. Agency for Healthcare Research and Quality (2012) Understanding quality measurement.. https://www.ahrq.gov/professionals/quality-patient-safety/quality-resources/tools/chtoolbx/understand/index.html. Accessed 15 May 2018
2. Advisory Commission on Health Consumer Protection and Quality in the Health Care Industry (1998) Quality first: better health care for all Americans. U.S. Government Printing Office, Washington D.C.
3. Institute of Medicine (IOM) (2001) Crossing the quality chasm: a new health system for the 21st century. National Academy Press, Washington D.C
4. Donabedian A (1989) The end results of health care: Ernest Codman's contribution to quality assessment and beyond. Milbank Qtrly 67(2):233–256
5. Agency for Healthcare Research and Quality (2015) Key questions when choosing health care quality measures.. https://www.ahrq.gov/professionals/quality-patient-safety/talkingquality/create/gather/index.html. Accessed 15 May 2018
6. Scobie S, Thomson R, McNeil J et al (2006) Measurement of the safety and quality of health care. MJA 184(suppl 10):S51–S55
7. Codman EA (1990) The end result idea and the product of a hospital: a commentary. Arch Pathol Lab Med 114(11):1105
8. Shewart W (1981) Economic control of quality of manufactured product. American Society for Quality, Milwaukee
9. Donabedian A (1980) Explorations in quality assessment and monitoring. In: The definition of quality and approaches to its assessment, vol 1. Health Administration Press, Chicago

10. Kaplan R, Norton D (1996) The balanced scorecard: translating strategy into action. Harvard Business School Press, Boston
11. Nelson E, Mohr J, Batalden P et al (1996) Improving health care part 1: the clinical value compass. Jt Comm J Qual Improv 22(4):243–258
12. Hochman M (2018) Electronic health records: a "quadruple win," a "quadruple failure," or simply a time for reboot. J Gen Intern Med 33(4):397–399. https://doi.org/10.1007/s11606-018-4337-6
13. Harrington L (2017) Copy-forward in electronic health records: lipstick on a pig. Jt Comm J Qual Patient Saf 43:371–374
14. Tang P, Ralston M, Arrigotti M et al (2007) Comparison of methodologies for calculating quality measures based on administrative versus clinical data from electronic health record system: implications for performance measures. J Am Med Inform Assoc 14:10–15
15. Schuster M, Onorato S, Meltzer D (2017) Measuring the cost of quality measurement: a missing link in quality strategy. JAMA 318(13):1219–1220
16. Ranson S, Maulik S, Nash D (2005) The healthcare quality book: vision, strategy and tools. Health Administration Press, Chicago
17. Chassin M, Loeb J, Schmaltz S et al (2010) Accountability measures–using measurement to promote quality improvement. NEJM 363(7):683–688
18. Mant J (2001) Process versus outcome indicators in the assessment of quality in health care. Intl J Qual Health Care 13(6):475–480
19. Institute for Healthcare Improvement (2017) What is a Bundle?. http://www.ihi.org/resources/Pages/ImprovementStories/WhatIsaBundle.aspx. Accessed 17 Jan 2017
20. Mohr J, Mahoney C, Nelson E et al (1996) Improving health care, part 3: clinical benchmarking for best patient care. Jt Com J Qual Improv 22(9):599–616
21. Kiefe C, Weissman N, Allison J et al (1998) Identifying achievable benchmarks of care: concepts and methodology. Int J Qual Health Care 10(5):443–447
22. Nolan T, Berwick D (2006) All-or-none measurement raises the bar on performance. JAMA 295(10):1168–1170
23. Agency for Healthcare Research and Quality (2017) Evaluation of the AHRQ QI program. https://www.ahrq.gov/research/findings/final-reports/qualityindicators/chapter4.html
24. Nugent W, Schultz W, Plume S et al (1994) Design an instrument panel to monitor and improve coronary artery bypass grafting. JCOM 1(2):57–64
25. Schuster M, Onorato S, Meltzer D (2017) Measuring the cost of quality measurement: a missing link in quality strategy. JAMA 318(13):1219–1220
26. Donabedian A (1996) Evaluating the quality of medical care. Milbank Qrtly 44:166–206
27. Aitken M (2018) The growing value of digital health: evidence and impact on human health and the healthcare system. IQVIA Institute for Human Data Science, In. https://www.iqvia.com/institute/reports/the-growing-value-of-digital-health
28. Harrington L (2017) New data of the digital age: big, dark and deep. AACN Adv Crit Care 28(3):239–242
29. Smitha T, Kumar V (2013) Applications of big data in data mining. Int J Emerging Technol Adv Eng 3(7):390–393. http://www.ijetae.com/files/Volume3Issue7/IJETAE_0713_65.pdf
30. Demchenko Y, Ngo C, Membrey P (2013) Architecture framework and components for the big data ecosystem. Universiteit van Amsterdam System and Network Engineering Group. http://uazone.org/demch/worksinprogress/sne-2013-02-techreport-bdaf-draft02.pdf
31. Sicular S (2013) Gartner's big data definition consists of three parts, not to be confused with three "V"s. Forbes Magazine, In. https://www.forbes.com/sites/gartnergroup/2013/03/27/gartners-big-data-definition-consists-of-three-parts-not-to-be-confused-with-three-vs/#5459367642f6
32. Trice A (2015) The future of cognitive computing. The IBM cloud blog, In. https://www.ibm.com/blogs/bluemix/2015/11/future-of-cognitive-computing/. Accessed 15 May 2018
33. Poucke S, Thomeer M, Heath J et al (2016) Are randomized controlled trials the (g)old standard? From clinical intelligence to prescriptive analytics. J Med Internet Res 18(7):e185. https://doi.org/10.2196/jmir.5549

34. Valikodath N, Newman-Casey PA, Lee P et al (2017) Agreement of ocular symptom reporting between patient-reported outcomes and medical records. JAMA Ophthalmol 135(3):225–231. https://doi.org/10.1001/jamaophthalmol.2016.5551
35. Echaiz J, Cass C, Henderson J et al (2015) Low correlation between self-report and medical record documentation of urinary tract infection symptoms. Am J Infect Control 43(9):983–986
36. Harrington L, White S (2008) Interview with a quality leader: Mark Chassin, new president of the Joint Commission. J Healthc Qual 30(1):25–29
37. Shahian D, Mormand S, Friedberg M et al (2016) Rating the raters: the inconsistent quality of health care performance measurement. Ann Surg 264:36–38
38. Stempniak M (2015) Don Berwick offers health care 9 steps to end era of 'complex incentives' and 'excessive measurement'. In: Hospitals & Health Networks.. http://www.hhnmag.com/articles/6798-don-berwick-offers-health-care-9-steps-to-transform-health-care. Accessed 21 Feb 2017