CHAPTER ONE

Introduction

The quality of healthcare must improve, both in the United States and internationally. At its best, the capabilities of modern medicine and surgery would have been inconceivable a generation ago. However, despite quantum advances in diagnosis and treatment, the exciting promise of gene-based personalized medicine, and the widespread adoption of electronic health records, healthcare quality continues to be limited by its often low quality and reliability, which are well below those of other high-stakes professions [1] such as military and commercial aviation and nuclear power. Recognition of these deficiencies was a primary motivator for contemporary public report cards and value-based healthcare initiatives.

1.1 Excessive Variation in Healthcare Outcomes

The most obvious manifestation of healthcare's unreliability is its persistent, pervasive, and excessive variation [2]. This is evident in the widely varying rates of diagnostic and therapeutic procedures (i.e., underuse, overuse, or misuse); quality of care, as measured by the structure, process, and outcomes domains introduced by Professor Avedis Donabedian [3, 4]; safety, as measured by adverse events such as hospital-acquired infections; and costs. The healthcare system has also failed to reduce excessive variation in timely and equitable access to care (particularly for minority or other vulnerable populations) and in patient-centeredness, as measured by patient-reported outcomes and experience of care surveys. It is the accurate measurement and reporting of these and related types of healthcare variation, particularly at the extremes, that characterizes *provider profiling*, defined as the comparisons of a health care provider's performance with a normative or community standard [5].

Measurement and remediation of unwarranted variation have become key objectives of virtually all healthcare reform initiatives, as demonstrated by their emphasis on transparency and accountability. Healthcare quality report cards are now available from numerous sources, examples of which include:

- Centers for Medicare and Medicaid Services (CMS) [6, 7] and the CMS *Care Compare* website [8]

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 05 Nov 2026 at 10:37:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9780511843525.002

- Agency for Healthcare Research and Quality [9] measures, such as the Patient Safety Indicators (PSI), Inpatient Quality Indicators (IQI), and Maternal Health Indicators (MHI)
- Centers for Disease Control and Prevention (CDC): National Healthcare Safety Network (NHSN) infection tracking measures [10]
- State departments of health and their derivatives, such as those in New York [11], Pennsylvania [12], California [13], and Massachusetts [14]
- Healthcare facility endorsement and certification organizations such as the Joint Commission [15]
- Healthcare quality measure endorsement organizations such as the National Quality Forum (NQF) [16] and its successor Consensus-Based Entity [17], the Battelle Partnership for Quality Measurement [18]
- Payers and regulators, usually based on administrative claims data
- Healthcare performance improvement organizations, such as Vizient [19] and Press Ganey [20]
- Professional societies, typically based on clinical registry data (e.g., Society of Thoracic Surgeons) [21]
- US News and World Report [22] and other commercial media outlets
- Independent, national, nonprofit healthcare quality and safety organizations such as the Leapfrog Group [23] and the National Committee for Quality Assurance (NCQA) [24], the latter of which produces the HEDIS performance measures [25] for health plans
- Online internet sites (e.g., Healthgrades [26], RateMDs [27], Vitals [28], Web MD [29], Yelp [30], Consumer's Checkbook [31])

1.2 Excessive Variation in Healthcare Quality Measurement

With the rapid proliferation of healthcare report cards, it has become apparent that variation exists not only in healthcare quality but also in how that quality is measured and reported. Any organization can develop and publish a scorecard or public report, using whatever data source, measure concept, and methodology it chooses. Unlike the standards for accounting overseen by the Financial Accounting Standards Board (i.e., the Generally Accepted Accounting Principles, or GAAP), there are few if any prescribed standards or practices for healthcare quality measurement. The predictable result is that different report cards often yield widely differing assessments of the same healthcare provider, and these cannot all be valid [32–51].

Though unacceptable, this excessive variation in healthcare quality measurement in not entirely surprising given its inherent complexity. Consider the many major topics that must be addressed when designing and implementing a healthcare quality measure (see Box 1.1), all of which are discussed in this book.

1.3 Consequences of Undisciplined Quality Measurement

The undesirable consequences of excessive, chaotic variation in the measurement of healthcare quality are substantial, as these ratings impact provider reputations, credentialing, referrals, disciplinary actions, membership in payer networks, and reimbursement. Most important, patients use these data to guide their choice of providers, and conflicting or inaccurate information is both confusing and potentially misleading.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 05 Nov 2026 at 10:37:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9780511843525.002

Box 1.1 Considerations in healthcare quality measurement

- Selection of target population (e.g., diagnosis, procedure, demographic group)
- Existing gap from optimal outcomes
- Extent of inter-provider variation
- Potential impact (e.g., large at-risk population and/or high morbidity/mortality)
- Data source (e.g., claims, registry, electronic health records (EHR))
- Data collection burden
- Data quality and audit
- Specific measure type (e.g., process, structure, outcomes, appropriateness)
- Adequate sample size and endpoint frequency for analyses
- Appropriate patient to provider attribution
- Patient observation period (e.g., in-hospital; 30-day or longer fixed periods)
- Provider observation period (e.g., typically 1–3 years)
- Statistical approaches (e.g., risk standardization, inference)
- Measure validity, reliability
- Methods for outlier determination
- User-friendly presentation
- Time lag from data collection to measure availability
- Public reporting versus confidential feedback
- Unintended consequences (e.g., gaming, risk aversion)

Scientifically flawed, misleading, or conflicting report cards are not innocuous, and we disagree with the aphorism that "Any data are better than no data." In addition to confusing patients, inaccurate reports may misclassify doctors and hospitals, potentially leading patients to seek care from low-performing providers or to avoid providers that are average or excellent. Flawed performance reports may lead providers to inappropriately focus scarce resources on areas incorrectly identified as underperforming, or to be lulled into complacency regarding areas that may actually need improvement. Similarly, they may lead payers to penalize good performers or reward poor performers.

Not only do all stakeholders deserve the most accurate, reliable healthcare performance data, but data collection, analysis, and reporting are also resource-intensive processes, and healthcare is resource-constrained. Therefore, considerable selectivity must be used in choosing where to apply these limited resources, and efforts must be made to limit data collection and administrative burden for providers. Berwick [52] has argued that we have far too many performance measures, and he advocates for substantial decreases in their numbers and the associated data collection burden and costs (i.e., 50% reduction in 3 years, 75% in 6 years). Ideally, this would be achieved by selecting and utilizing only the most credible measures and report cards, those likely to provide the most accurate information to patients and other healthcare consumers.

1.4 The Need for Quality Measurement Standards

The best healthcare quality rating systems, whether used for confidential feedback or public reporting, are methodologically rigorous; based on the most credible data sources; risk-adjusted for patient severity; published in the peer-reviewed literature

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 05 Nov 2026 at 10:37:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9780511843525.002

(where the editorial review process often uncovers errors and makes constructive suggestions); and, ideally, externally vetted and endorsed by credible national organizations, as described in Section 1.5. Examples of current and past healthcare quality rating systems that have generally evinced these principles include:

- Centers for Medicare & Medicaid Services, many of whose measures are included on the CMS *Care Compare* website [6, 8, 53–55]
- Some states, such as New York [11, 56], Pennsylvania [12], and Massachusetts [14, 57–59]
- Professional organizations such as the Society of Thoracic Surgeons (STS) [21], the American College of Cardiology [60, 61], the American Heart Association [62, 63], and the American College of Surgeons [64] typically using clinical registry data
- Regional performance improvement collaboratives such as the Northern New England Cardiovascular Disease Study Group [65–67], which utilized its own rigorous clinical data registry and methodologies, and similar organizations in Michigan [68] and Virginia [69] which have used STS clinical registry data

Unfortunately, among the hundreds of report cards and rating systems now available, methodologically rigorous, peer-review published, externally validated examples such as these are the exception. Many contemporary report cards produced by commercial entities or so-called public service organizations are often conceptually and scientifically problematic [37, 70–74]. The details of their methodologies are either proprietary or difficult to access; they are rarely scrutinized through the peer-review publication process; and they lack external validation or endorsement. However, patients, providers, and other stakeholders are typically unaware of these issues, and their often conflicting results simply confuse and discredit the measurement process rather than inform patients and the public.

The lack of measurement standards, like so many other problems in healthcare quality, was presciently observed a century ago by Dr. Ernest Amory Codman at the Massachusetts General Hospital in his paper titled *Uniformity of hospital morbidity reports* [75]:

> The truth is, that each hospital publishes these statistics as a sort of boast that it is doing something; the tables look well in a report, and impress the trustees and subscribers; otherwise than this they are not used by anybody. Anyone who might wish to use them knows that they are too inaccurate and too diverse in plan and method of classification to be of service either for comparison one with another, or for large statistics made by adding them one to another.
>
> . . . there is no uniformity at present in the reports of our hospitals. . . if all our hospitals had a uniform report, no matter how simple it was,. . . the added figures would form statistical and scientific data of value.
>
> . . . I think that every one would admit that such a simple report, if given uniformly by every hospital, would be of far greater value than the elaborate reports on which our hospitals spend so much labor and money at present.

Codman anticipated the need for scientifically objective and accurate performance reports, based on samples of adequate size and using pre-specified methodologies:

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 05 Nov 2026 at 10:37:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9780511843525.002

> Great hospitals like the Massachusetts General Hospital have a duty to perform to medical and surgical science. By grouping cases into series large enough to favor comparative study and by observing definite previously-determined points a rational clinical science can be established [76].

Sadly, more than a century later, the goal of making provider quality measurement a standardized, rational clinical science has yet to be consistently achieved.

In response to the current chaotic landscape in healthcare performance measurement, numerous organizations and investigators have identified major areas of weakness and inconsistency, often suggesting approaches to promote greater rigor and standardization. See *Key references* Section 1.9 at the end of this chapter for a representative list of articles addressing this topic. Also, selected examples of these frameworks and recommendations are provided in the Appendix.

1.5 Consensus-Based Entity: *National Quality Forum* and its successor, the *Partnership for Quality Measurement*

A major step toward establishing standards for measures used by CMS was enactment of *The Medicare Improvements for Patients and Providers Act of 2008*, which required that the Department of Health and Human Services (HHS) contract with a Consensus-Based Entity (CBE) [17, 77] for multiple functions relating to performance measures used by the federal government. These functions include *priority setting*, as well as *endorsement* only of those measures found to be "evidence-based, reliable, valid, verifiable, relevant to enhanced health outcomes, actionable at the caregiver level, feasible to collect and report, and responsive to variations in patient characteristics, such as health status, language capabilities, race or ethnicity, and income level." The law also required that the CBE be responsible for periodic *maintenance* of such endorsement.

National Quality Forum was the inaugural CBE. Its successor and the current CBE is the Battelle *Partnership for Quality Measurement*[18, 78–81] (PQM). Major project areas for PQM Endorsement and Maintenance include Primary Prevention: Initial Recognition and Management; Management of Acute Events, Chronic Disease, Surgery, and Behavioral Health; Cost and Efficiency; and Advanced Illness and Post-Acute Care. As of August 2025, the PQM Measure Evaluation Rubric includes Importance; Closing Care Gaps; Scientific Acceptability; Feasibility; and Usability and Use. Detailed descriptions of the evaluation criteria for each of these domains may be found in pages 48–54 of the PQM Evaluation and Maintenance (E&M) Guidebook (August 2025 version) [78].

Though a rigorous and respected process for validating healthcare performance measures, CBE endorsement has most often been required and applied to measures used by CMS. Other than these, pursuit of PQM measure endorsement by other measure developers and stewards is laudable but voluntary. It demonstrates their desire to adhere to the most rigorous current standards for healthcare measures. Unfortunately, many measure developers do not pursue PQM endorsement, either because of the time and resources necessary to complete this process or because they recognize that their measure would likely fail to be endorsed.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 05 Nov 2026 at 10:37:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9780511843525.002

1.6 The CMS Universal Foundation

Another important step toward standardizing healthcare performance measures used by more than 20 different CMS quality rating and value-based care programs (many of which are also used in commercial insurance programs) was the publication in March 2023 of the CMS Universal Foundation [82]. Recognizing the need to mitigate quality measure proliferation and to focus provider attention on measures that are most meaningful to patients and most likely to drive quality improvement and care transformation, CMS plans to align quality measures across its various programs with the following goals:

- Focus providers' attention on measures that are meaningful for the health of broad segments of the population
- Reduce provider burden by streamlining and aligning measures
- Advance equity with the use of measures that will help CMS recognize and track disparities in care among and within populations
- Aid the transition from manual reporting of quality measures to seamless, automatic digital reporting
- Permit comparisons among various quality and value-based care programs, to help the agency better understand what drives quality improvement and what does not
- Prioritize measures most likely to achieve these goals and have minimal unintended consequences

1.7 What Is in This Book

The goal of this book is to identify methodological considerations in the measurement of healthcare quality and to provide recommendations for implementing current best practices and avoiding common errors. These issues are equally applicable both in the United States, the source for many of the examples in this book, and internationally.

Healthcare quality metrics must meet the highest methodological standards, thus assuring the accurate, unbiased assessments of provider performance. Our hope is that developers will find the principles described in this text useful as they design measures, and that those evaluating the scientific properties of publicly reported measures will also find these principles to be a valuable reference standard. Ultimately, the overarching goal of all these activities is to improve the quality and safety of healthcare and reduce unwarranted variation, both in the United States and internationally.

This book also contains extensive information regarding health policy issues that are directly related to quality measurement, such as various strategic and tactical approaches for improving healthcare quality. We include detailed discussions of the evolution, controversies, and current status of public reporting, consumer choice initiatives, value-based purchasing, and direct-to-consumer dissemination of provider information. Although we introduce the concepts of cost variation, financial incentives, and penalties, all of which are important aspects of healthcare reform, these often involve specific methodological approaches that are beyond the goals of this text.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 05 Nov 2026 at 10:37:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9780511843525.002

We adopt a *data science* approach that includes both substantive and statistical considerations. The book is divided into two parts. In Volume 1, measurement and related health policy topics are presented comprehensively but at a level that does not require advanced statistical training. Extensive references are provided for each chapter, and several chapters also include a separate *Key References* section at the end. In addition to the references cited within the chapter text, these Key References are intended to highlight articles that may be especially useful for readers wishing to further pursue particular topics.

Volume 1, Chapter 2 (*Conceptual Models*) provides a brief introduction to the concept of provider profiling and selection of the appropriate level of provider attribution, followed by examples of performance measures encompassing continuous, binary, and mixed (non-commensurate) endpoints; single versus multiple measures; and different levels of attribution. Building on these examples, a conceptual model is presented for the design of healthcare measures. Volume 1, Chapter 3 (*History and Evolution of Healthcare Performance Measurement*) presents an evolutionary history of performance measurement in healthcare and its applications in both accountability and payment programs, all of which were motivated in large part by the excessive variability in quality between providers. The advantages and disadvantages of early indices of variability are described.

Volume 1, Chapters 4 and 5, discuss major types of performance measures and data sources, respectively. Based on the work of Professor Avedis Donabedian, Volume 1, Chapter 4 (*Types of Performance Measures*) expands upon his original quality taxonomy of structure, process, and outcomes measures [3, 4, 83, 84], with particular emphasis on the latter, as they are both the most important and methodologically challenging metrics. Volume 1, Chapter 5 (*Data Sources for Healthcare Quality Measurement*) describes the advantages and disadvantages of commonly used healthcare data sources including claims data, clinical registries, and EHR's, as well as the added utility of linking these various data sources.

The fundamental measurement attributes of bias, validity, reliability, and power (and the relation of the latter two, in particular, to sample size issues) are presented in Volume 1, Chapter 6 (*Healthcare Performance Measure Attributes*). Key decisions in the design of measures are introduced in Volume 1, Chapter 7 (*Target Conditions, Patient & Provider Observation Periods, and Attribution*), including guidelines for selecting appropriate conditions and procedures to measure; appropriate patient and provider observation periods; and levels of attribution for various conditions, procedures, and purposes. Risk model development and validation are presented in Volume 1, Chapter 8 (*Risk Modeling and Validation*), followed by an entire chapter (Volume 1, Chapter 9 (*Social Risk Factors*)) on the controversial topic of social risk factors, including race and ethnicity, in risk models used for provider profiling. Volume 1, Chapter 10 (*Performance Classification and Outliers*) discusses a critically important function of healthcare quality measurement – identifying high- and low-performing providers, an often controversial process with both statistical and policy implications. Having thus described fundamental considerations focusing on individual measures, composite measures are introduced in Volume 1, Chapter 11 (*Composite Performance Measures*), including the various types of composites and fundamentals of their construction, the latter being generally more complicated than for individual measures.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 05 Nov 2026 at 10:37:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9780511843525.002

In Volume 1, Chapter 12 (*Data Presentation Strategies*), strategies to accurately and effectively communicate healthcare performance information to the public are discussed, with an emphasis on maximizing comprehension by the intended audiences, who often lack quantitative literacy. Public reporting is introduced in Volume 1, Chapter 13 (*Public Reporting: Introduction, Ethical Considerations, and Outcomes*), including the ethical rationale for reporting and evidence for its value in improving healthcare quality. The role of public reporting in consumer choice and its impact on market share are then discussed in Volume 1, Chapter 14 (*Public Reporting: Consumer Behavior and Market Impact*). Unintended consequences of public reporting are introduced in Volume 1, Chapter 15 (*Unintended Consequences*), including two of the most widely recognized and feared – *gaming* and *risk aversion*. Volume 1, Chapter 16 (*Web-Based Public Reporting*) discusses the rapid growth of online, consumer report cards, types of data and narratives presented, and their validation against other sources of hospital and physician information. Finally, Volume 1, Chapter 17 (*Graphical Approaches to Performance Assessments*) presents a wide variety of graphical approaches to visually display provider performance information.

Volume 2 of the book focuses on technical concepts and approaches that benefit from an intermediate knowledge of statistics. Basic notation is introduced and the key statistical considerations as they relate to the conceptual model of assessing provider performance are discussed. Because the focus of inference is on the performance of healthcare providers, *risk-adjustment* to make causal inference is the most critical consideration, and Volume 2, Chapter 19 (*A Causal Inference Framework*) links risk-adjustment to causal inference. In the causal inference setting, "providers" are the "treatments" to which patients are assigned. Several approaches to risk-adjustment are described. Volume 2, Chapter 20 (*Introduction to Hierarchical Linear Models*) and Chapter 21 (*Hierarchical Generalized Linear Models*) provide a statistical treatment of hierarchical linear models (HLM) and hierarchical generalized linear models (HGLM), respectively. The assumptions made are detailed and the link between modeling simultaneously discrete and non-discrete outcomes and hierarchical models is made. HGLMs involve more technical details but are direct extensions of the HLM. In Volume 2, Chapters 20 and 21, we provide both frequentist and Bayesian modeling perspectives.

Volume 2, Chapter 22 (*Composite Measures and Multiple Outcomes*) gives technical approaches to creating composite quality measures and analyzing more than one quality metric, including when the metrics are on different measurement scales. Volume 2, Chapter 23 (*Computational Approaches*) describes how to undertake the complex computations necessary to arrive at parameter estimates in hierarchical models. Volume 2, Chapter 24 (*Inference for Hierarchical Generalized Linear Models*) discusses how to make inferences with hierarchical models, including approaches when adopting either a frequentist or a Bayesian framework. Volume 2, Chapter 25 (*Model Assessment*) provides strategies for assessing model fit. Because multiple levels of clustering may exist, checks in addition to those used in non-hierarchical regression modeling are necessary. Finally, Volume 2 concludes with Chapter 26 (*Future Directions*), which discusses methods for profiling providers when parametric models may not be suitable for risk-adjustment. This may arise when many risk factors are required to avoid unmeasured confounding.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 05 Nov 2026 at 10:37:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9780511843525.002

Approaches that relax parametric assumptions, first introduced in Volume 2, Chapter 19, are described using tools from the causal inference literature.

1.8 Who Should Read This Book

Volume 1 of this book provides an in-depth, comprehensive, but less technical presentation of the history and current practice of healthcare quality measurement. This would be appropriate for undergraduate and graduate students, health policy researchers, regulators, public health specialists, attending physicians and trainees, and quality leaders at the hospital, state, and national levels.

Volume 2 of this book is primarily aimed at readers who have quantitative training in measurement science including standard regression modeling such as linear regression, logistic regression, and Poisson regression. The textbook by Harrell provides a comprehensive treatment for these types of models [85]. Experience with multilevel modeling [86, 87] is helpful but not critical. However, the material contained in Gelman and Hill (2007) [86] is particularly noteworthy, and the reader is encouraged to consult this text. Finally, while this book focuses on reporting of public health outcomes, the methodological approaches are suitable for other applied areas such as education and small area analysis.

1.9 Key References

1.9.1 Measurement Chaos, Inconsistent and Flawed Ratings

- Auffenberg, G.B., Ghani, K.R., Ye, Z., et al. Comparing publicly reported surgical outcomes with quality measures from a statewide improvement collaborative. *The Journal of the American Medical Association: Surgery*, 151(7):680–682, 2016.
- Austin, J.M., Jha, A.K., Romano, P.S., et al. National hospital ratings systems share few common scores and may generate confusion instead of clarity. *Health Affairs (Millwood)*, 34(3):423–30, 2015.
- Ban, K.A., Cohen, M.E., Ko, C.Y., et al. Evaluation of the ProPublica surgeon scorecard "adjusted complication rate" measure specifications. *Annals of Surgery*, 264(4):566–74, 2016.
- Friedberg, M.W., Pronovost, P.J., Shahian, D.M., et al. A methodological critique of the ProPublica surgeon scorecard. *RAND Health Quarterly*, 5(4):1, 2016.
- Friedberg, M.W., Bilimoria, K.Y., Pronovost, P.J., et al. Response to ProPublica's rebuttal of our critique of the Surgeon Scorecard. *RAND Health Quarterly*, 6(1):4, 2016.
- HANYS Quality Institute. Report on report cards; understanding publicly reported hospital quality measures, 2013. www.hanys.org/quality/data/report_cards/2013/docs/2013_hanys_report_card_book.pdf.
- HANYS Quality Institute. Report on report cards, 2019. www.hanys.org/communications/publications/report_cards/docs/hanys_report_on_report_cards_2019.pdf.
- HANYS Quality Institute. Measures that matter: Moving from measure madness to measures that matter, 2016. www.hanys.org/quality/clinical_operational_oversight/measures_that_matter/docs/mtm_report.pdf.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 05 Nov 2026 at 10:37:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9780511843525.002

- Hwang, W., Derk, J., LaClair, M., and Paz, H. Finding order in chaos: A review of hospital ratings. *American Journal of Medical Quality*, 31(2):147–155, 2016.
- Leonardi, M.J., McGory, M.L., and Ko, C.Y. Publicly available hospital comparison websites: Determination of useful, valid, and appropriate information for comparing surgical quality. *Archives of Surgery*, 142(9):863–868, 2007.
- Raghuram, A.C., Dasari, T.K., Chou, B., et al. Confusion instead of clarity: Publicly reported cardiac surgery ratings for coronary artery bypass grafting and aortic valve replacement. *Journal of the American College of Surgeons*, 228(2):180–187, 2019.
- Rothberg, M.B., Morsi, E., Benjamin, E.M.,et al. Choosing the best hospital: The limitations of public quality reporting. *Health Affairs (Millwood)*, 27:1680–1687, 2008.
- Shahian, D.M. Measuring and reporting cardiac surgery quality: A continuing evolution. *The Journal of Thoracic and Cardiovascular Surgery*, 166(3):819–825, 2023.
- Sukul, D., Bhatt, D.L., Seth, M., et al. Appropriateness and outcomes of percutaneous coronary intervention at top-ranked and nonranked hospitals in the United States. *JACC Cardiovascular Interventions*. 11(4):342–50, 2018 (Commentary by Dehmer, G.J. Seeking Quality Cardiac Care: Is Public Reporting the Answer? *JACC: Cardiovascular Interventions*, 11(4):351–353, 2018.)

1.9.2 Recommendations

- Berenson, R.A., Pronovost, P.J., and Krumholz, H.M. Achieving the potential of health care performance measures. The Robert Wood Johnson Foundation, 2013. www.urban.org/sites/default/files/publication/23636/412823-achieving-the-potential-of-health-care-performance-measures.pdf.
- Bilimoria, K.Y., Birkemeyer, J.D., Burstin, H., et al. Rating the raters: An evaluation of publicly reported hospital quality rating systems, 2019. https://catalyst.nejm.org/evaluation-hospital-quality-rating-systems/.
- Collins, G.S., Reitsma, J.B., Altman, D.G., and Moons, K.G. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Annals of Internal Medicine*, 162(1):55–63, 2015.
- Drozda, J.P., Jr., Hagan, E.P., Mirro, M.J., et al. ACCF 2008 health policy statement on principles for public reporting of physician performance data: A report of the American College of Cardiology Foundation Writing Committee to develop principles for public reporting of physician performance data. *Journal of the American College of Cardiology*, 51(20):1993–2001, 2008.
- HANYS Quality Institute. Measures that matter: Moving from measure madness to measures that matter, 2016. www.hanys.org/quality/clinical_operational_oversight/measures_that_matter/docs/mtm_report.pdf.
- HANYS Quality Institute. Report on report cards; understanding publicly reported hospital quality measures, 2013. www.hanys.org/quality/data/report_cards/2013/docs/2013_hanys_report_card_book.pdf.
- HANYS Quality Institute. Report on report cards, 2019. www.hanys.org/communications/publications/report_cards/docs/hanys_report_on_report_cards_2019.pdf.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 05 Nov 2026 at 10:37:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9780511843525.002

- Friedberg, M.W. and Damberg, C.L. A five-point checklist to help performance reports incentivize improvement and effectively guide patients. *Health Affairs (Millwood)*, 31(3): 612–18, 2012.
- Krumholz, H.M., Brindis, R.G., Brush, J.E., et al. Standards for statistical models used for public reporting of health outcomes: An American Heart Association scientific statement from the quality of care and outcomes research interdisciplinary writing group: cosponsored by the Council on Epidemiology and Prevention and the Stroke Council. Endorsed by the American College of Cardiology Foundation. *Circulation*, 113:456–462, 2006.
- MacLean, C.H., Kerr, E.A., and Qaseem, A. Time out – charting a path for improving performance measurement. *The New England Journal of Medicine*, 378(19):1757–1761, 2018.
- National Quality Forum. Measure evaluation criteria and guidance for evaluating measures for endorsement. National Quality Forum, 2021. www.qualityforum.org/Show_Content.aspx?id=322.
- Porter ME, Larsson S, Lee TH. Standardizing patient outcomes measurement. *N Engl J Med*. 2016;374(6):504–506.
- Pronovost, P.J., Austin, J.M., Cassel, C.K., et al. Fostering transparency in outcomes, quality, safety, and costs: A vital direction for health and health care. *National Academy of Medicine: Perspectives*, 2016. https://nam.edu/wp-content/uploads/2016/09/Fostering-Transparency-in-Outcomes-Quality-Safety-and-Costs.pdf.
- Rothberg, M.B., Benjamin, E.M., and Lindenauer, P.K. Public reporting of hospital quality: Recommendations to benefit patients and hospitals. *Journal of Hospital Medicine*, 4(9):541–45, 2009.
- Shahian, D.M., Mort, E.A., and Pronovost, P.J. The quality measurement crisis: An urgent need for methodological standards and transparency. *The Joint Commission Journal on Quality and Patient Safety*, 42(10):435–438, 2016.
- Shahian, D.M., Normand, S.L.T., Friedberg, M.W., et al. Rating the raters: The inconsistent quality of health care performance measurement. *Annals of Surgery*, 264(1):36–38, 2016.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 05 Nov 2026 at 10:37:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9780511843525.002