Understanding Quality Scorecards: A Primer for Boards

by Caitlin Gillooley

Overview

The number of public quality scorecards for hospitals has increased exponentially in recent years as consumers take more interest in getting the most value for their health care dollar. These attempts at simplifying the complex hospital environment into laymen's terms often condense hospital performance on a select number of quality measures into a letter grade, star, or ranking—much easier for the average stakeholder to understand than risk-adjusted infection rates.

Hospitals have long supported transparency on quality information, and the emergence of the public report cards reflects consumers' keen interest in better understanding the quality of care in hospitals. At a time when hospitals, consumers and policymakers alike are focused on improving the value of care, quality data is crucial to ensuring health choices are not based on lower prices alone.

However, the producers of these scorecards and reports bear great responsibility in making judgment calls on behalf of consumers. The average patient won't know whether a certain scoring methodology is truly reflective of the quality of care provided at a specific hospital, or if the calculations behind the letter-grade are statistically flawed and don't show much about the provider at all. In addition, the easy-to-understand nature of the scores carries the inherent risk of oversimplifying the complexity of delivering quality care: each patient, doctor, and hospital is unique and operates in an environment that is only so comparable to every other patient, doctor, and hospital. Finally, the proliferation of scorecards means that hospitals often receive discordant ratings across different reports, even when the reports are based on some of the same measures.

Questions also remain about how consumers are using quality information, though available studies suggest consumer uptake has been somewhat limited. Studies from *Penn State University*, the *Harvard School of Public Health*, and *Health Affairs* have found that consumers are actually not likely to use these data: in 2008, while 30 percent of respondents reported seeing comparative quality data, only 14 percent said they actually used the information to choose a doctor, hospital, or insurance plan. The same survey found that the primary users of this information were white, college-educated, and over 45 years old; vulnerable populations, arguably those most in need of high-value care, were far less likely to read about provider quality. Another study showed that, while awareness of physician quality has increased modestly across the past five years, there has been virtually no increase in awareness of

hospital quality. Researchers cite information overload and a lack of personally relevant outcomes as reasons that consumers don't actively use quality data.

That said, there are other incentives for hospitals and health systems to pay attention to quality scorecards. Like a personal resume, a list of industry awards can make a hospital an appealing candidate for research grants, programs, or philanthropic awards.

So how should hospital boards consider these ubiquitous but often flawed score cards? The first step is to know the main organizations proffering these scores, the methodologies they use, and why a certain grade might not be what it seems.

Report Card 101

We have compiled a list of the most popular score cards available in the table following this article. In the table, you'll see summaries of the important aspects of each report.

Organization Bio: Even though they generally serve the same consumer/payer audience, scorecards are developed by several types of stakeholders. From government regulators and payers; to non-profit journals; to for-profit, general-interest magazines, the authors of various scorecards differ somewhat in their motivations to publish quality data. Some are looking to promote high-value purchases by consumers (e.g., CMS Star Ratings), while others use their background investigations into quality to call attention to providers they deem to be "high-risk" (e.g., ProPublica).

Data Source: An evaluation is only as good as the data on which it is based; however, gathering data can be expensive. Because of the cost and logistical difficulties, many organizations use publicly available data from the Centers for Medicare & Medicaid Services (CMS). Much of those data are collected from hospitals as part of CMS's inpatient and outpatient quality reporting programs, which tie Medicare payment updates to reporting quality data. While CMS data capture a significant number of hospital interactions, the data is not necessarily representative of patients outside of the Medicare population. In addition, many individual measures in programs have significant flaws. Thus, where CMS uses a sub-par methodology, the report card will, too.

Other organizations attempt to collect their own data, usually through surveys filled out by hospitals. However, self-reported surveys can be suspect: reporting differs by respondent (e.g., Who in the organization answered the survey? Where did that person get the information? How did he/she interpret the question?). Survey designs also are not always validated to ensure that they collect the intended information; surveyors struggle with incomplete or missing information. While there is no perfect source of data, some sources are more reliable than others.

Scoring Methodology: The scoring methodology describes what areas of a hospital's performance are evaluated, for example, mortality or infection rates, and how those rates are translated into a resulting score, grade, or ranking. There is some overlap in the measures evaluated: hospitals are generally graded based on their ability to keep patients alive (mortality rates) and safe (hospital-acquired condition rates). A few score cards have begun to incorporate other aspects of quality, such as adherence to evidence-based processes and staffing ratios. Scorecards either determine a benchmark against which hospital performance will be judged—if performance is above the benchmark, the hospital gets a good grade—or rank hospitals in order of their outcomes.

The nature of this task—turning complex medical practice in multiple procedures, specialties, and facilities into a single score—is inherently fraught. Many scorecards sacrifice accuracy and specificity for the sake of simplicity; for example, a scorecard might base an overall quality score on infection rates for one or two procedures because that's the data that are accessible, or use a select list of measures as a proxy for overall quality. However, because hospitals vary significantly in what services they provide, those measures might not be relevant for a particular provider. In addition, a certain methodology might fail to control for certain factors that influence outcomes (for example, the sociodemographic characteristics of the community where the hospital operates) because it's difficult to develop an algorithm to account for those factors.

Advantages/Disadvantages: Finally, we've provided a summary of the strengths and weaknesses of these various scorecards. Strengths can be a unique point of view that highlights data that might be more relevant to patients seeking quality information, or a data gathering approach that doesn't pose an additional burden on hospitals. Weaknesses are often in the methodologies used to calculate grades or the limited scope of the report (e.g., leaves out certain kinds of providers or only looks at one type of safety indicator).

Board members should consider these characteristics when determining which score cards, if any, they should be incorporating into their strategies.

Hospitals have long supported transparency of quality data, but have urged that consumers view report cards as one tool among many to inform their health care decisions. Consumers must consider the advantages and disadvantages of the design of each report card. Furthermore, there is increasing concern that the growth in the number of report cards has created confusion for hospitals and consumers, and that not all report cards use sound methodologies. To promote a more rational approach, the AHA, along with other national hospital associations, has endorsed a set of *principles* for developing public report cards, calling for them to be well-defined in purpose, to use valid measures and to have transparent methodologies.

Scorecards aren't going away: consumers want help distinguishing one hospital from another, and these reports offer the promise of synthesizing complex information for worried patients. Notwithstanding their flaws, public reports often draw attention from consumers and the media. Boards can help their organizations by asking questions about how the scorecards relate to the hospital's broader approach to quality, and what response they may have. Here's where you can start:

- How does my board track my hospital's quality of care? Do we consider internal metrics? Public scorecards?
- Does our board have a clear understanding of the scorecards we are using to track performance: What the metrics are designed to measure? What the data indicate? What board members should be looking for when they review these reports and what types of questions they could ask, etc.?
- Does our perception of our hospital's quality (and our own internal data) match our scores/grades/rankings on public scorecards?
- Should we track how we perform on public scorecards as part of our organization's internal performance tracking?
- What do our scores/grades/rankings suggest we need to do—introduce more quality initiatives? Improve our marketing?
- If we have scores on public reports that we do not like, are we prepared to respond to media inquiries? How can we showcase the good work that we do?

Caitlin Gillooley is AHA associate director of policy. For more, please visit www.aha.org.

REPORT CARD	ORGANIZATION BIO	DATA SOURCE	SCORING METHODOLOGY	ADVANTAGES	DISADVANTAGES
CMS Hospital Star Ratings (Hospital Compare)	Published by the Centers for Medicare & Medicaid Services (CMS)	CMS's Hospital Inpatient Quality Reporting (IQR) and Outpatient Quality Reporting (OQR) programs	1-5 stars based on performance on 57 measures	Draws on measures from required reporting programs (so no additional reporting required) Star rating may be easier to understand than individual measure scores Medicare is the single largest payer of health care services and many programs tie a significant amount of dollars to quality	Recent studies have raised serious questions on validity of methodology The list of selected measures may not be fully representative of hospital quality, providing a misleading picture
Leapfrog Hospital Survey	Non-profit organization representing employers and insurance purchasers	Unvalidated survey data reported by hospitals; additional data from secondary sources including AHA Health IT (HIT) supplement and annual survey, Hospital Compare, HAC Reduction program, AHRQ Patient Safety Indicators (PSI)	Grade A-F based on composite score from evaluation on performance in ensuring "Freedom from harm": Process/structure (how often a hospital gives patients recommended treatment for given condition/procedure), and Outcomes	Measures focused on patient safety issues, which are a key priority for hospitals	Use of deeply flawed claims-based safety measures in methodology Measure data may be up to three years old and not show more recent improvements in care VA, critical access hospitals, specialty, children's, mental health hospitals not included; arbitrary weighting of measures in composite score
US News & World Report Best Hospitals	For-profit company	AHA Annual Survey (volume), Medicare Provider Analysis and Review (MedPAR) (mortality); Medicare Standard Analytic File (SAF); survey of physician specialists	Ranking by specialty and by state based on performance on structure (volume, staffing, other resources), process (reputation among physicians, patient safety indicators), outcomes (mortality)	No application or data submission required Assesses multiple aspects of care	Reputational data alone may not fully reflect quality of care
Truven Top 100 Hospitals	For-profit health care research and consulting firm	MedPAR, Medicare hospital cost reports (all-payer), CMS Hospital Compare	List of 100 hospitals with highest achievement in scores on 11 measures including inpatient outcomes, process of care, extended outcomes, process efficiency, cost efficiency, financial health, and patient experience	Variety of types of measures provides more nuanced picture of quality than just mortality or infection Compares hospitals in groups with similar characteristics (bed size, teaching status, extent of residency/fellowship program)	Bases risk-adjustment model on proprietary methodology that projects discharge data, so results of scoring are not replicable and internal methodology is speculative VA, critical access hospitals, specialty, children's, mental health hospitals not included

REPORT CARD	ORGANIZATION BIO	DATA SOURCE	SCORING METHODOLOGY	ADVANTAGES	DISADVANTAGES
Consumer Reports Hospital Safety Ratings	Non-profit organization supported by subscriptions	Hospital Compare, Leapfrog, specialty societies, AHA annual survey	Score between 1-100 (higher is better) based on Performance on outcomes (infections, mortality, readmissions, adverse surgical events); Experience (communication about discharge, drug information); Practices (appropriate use of scanning, avoiding C-sections)	Numerical score may be easier for consumers to understand	Overall score not fully reflective of overall hospital quality Some underlying measures (CT imaging, mortality, readmission) have reliability and validity problems; uses unvalidated Leapfrog survey data
Healthgrades	For-profit company providing information to consumers	MedPAR, all-payer state data	List of top 50 (top 1%) and top 100 (top 2%) performers on mortality and in-hospital complications by procedure who have received the Healthgrades Distinguished Hospital Award for Clinical Excellence for a specific number of consecutive years	Rewards consistent, year-over-year quality Listing hospitals that have reached performance threshold avoids confusing, arbitrary grading or rating system What does "list system" mean?	Limited measures used to calculate scores; only risk-adjusted for comorbid diagnoses, age and gender, and source of admission; Inaccuracy of claims data. 23 states have no hospital receiving award; to be eligible for Distinguished Hospital Award for Clinical Excellence, hospital has to have evaluations in at least 21 of the 32 Healthgrades procedures and conditions using Medicare inpatient data
ProPublica Surgeon Scorecard	Independent, non-profit newsroom that produces investigative journalism in the public interest.	Medicare Standard Analytic File (SAF) Please spell out SAF	Low, medium, or high "Adjusted complication rate": hospital readmissions for conditions plausibly related to surgery and mortality within 30 days for eight surgical procedures; exclamation point symbol shown with rate for hospitals with at least one surgeon with a high adjusted complication rate	Uniquely focuses on surgeons, provides insight on specific specialties that might be more relevant for patients interested in those procedures	Masks hospital-to-hospital performance differences; questionable accuracy of methodology (doesn't include complications beyond those accompanied with 30-day readmission, patient risk doesn't affect score); Claims data are notoriously inaccurate in individual provider assignments