

DESIGNING EFFECTIVE HEALTH CARE QUALITY TRANSPARENCY INITIATIVES

By Ha T. Tu and Johanna R. Lauer

Among the many health care quality transparency initiatives introduced in recent years, two state-based programs stand out for thoughtful design, implementation and usable, useful data: CalHospitalCompare, a report card for California hospitals, and Massachusetts Health Quality Partners, a report card for Massachusetts primary care physician groups. According to a new Center for Studying Health System Change (HSC) analysis, both programs share key elements that contribute to their effectiveness: engaging and collaborating with the provider community from the outset; paying particular attention to the caliber of the quality data reported; presenting the quality data to consumers in formats that are easy to understand and remember; and providing hospitals and physicians with detailed information on their own performance. Quality transparency initiatives that do not focus sufficiently on these key design and implementation elements are unlikely to influence quality improvement in a meaningful way.

Case Studies: CalHospitalCompare and Massachusetts Health Quality Partners

In recent years, federal and state governments, health plans and others have launched a plethora of quality transparency initiatives intended to help consumers compare the performance of doctors or hospitals and, ultimately, to improve the quality of care. These programs vary greatly in the thoughtfulness of their design and implementation, the usability of the data—how meaningful, accurate and reliable the data are—and the usefulness of data to consumers—how easy to understand and remember the information is for consumers with different levels of health literacy and numeracy.

This Issue Brief highlights two quality transparency initiatives that can be considered success stories in being thoughtfully designed and implemented and presenting usable and useful quality information (see Data Source). Key features that make these programs effective and useful will

be highlighted—features that other quality transparency programs may be able to draw from and replicate. The elements described are not intended to form a comprehensive list of desired program features. Rather, they represent some of the most salient and replicable characteristics of well-designed and well-implemented quality transparency programs.

The first program is CalHospital Compare, a Web site launched in March 2007 that rates California hospitals on more than 70 performance measures, encompassing process, outcome and patient experience measures. The Web site is the result of a partnership between the California Hospitals Assessment and Reporting Taskforce (CHART), the California HealthCare Foundation and the University of California at San Francisco Institute for Health Policy Studies. CHART was formed in 2004 with the objective of develop-

ing a standardized quality report card for California hospitals; CalHospitalCompare is its consumer Web site. Currently, more than 240 hospitals—representing 86 percent of the average daily inpatient census of California hospitals—participate in the program.

The second program profiled is Massachusetts Health Quality Partners (MHQP), which introduced a Web site in 2005 that compares the performance of primary care physician groups in Massachusetts using more than 30 process and patient experience measures.¹ MHQP was established in 1995 by a group of Massachusetts health care leaders. Currently, MHQP reports quality ratings for 150 medical groups that include 4,500 primary care physicians. The quality ratings are drawn from data collected by five participating health plans that collectively cover about half of all commercially insured Massachusetts residents.



Successful public quality transparency initiatives tend to build on a broad base of stakeholders—including providers, insurers, purchasers, consumer groups and policy makers—from the earliest stages of program design.

Engaging and Collaborating with Stakeholders

Successful public quality transparency initiatives tend to build on a broad base of stakeholders—including providers, insurers, purchasers, consumer groups and policy makers—from the earliest stages of program design. It is particularly important to include members of the provider community that will be assessed by the program. Engaging providers from the beginning increases participation in the program (in voluntary transparency initiatives), helps ensure clinical and practical relevance of the measures, and helps increase acceptance by providers of the program's measures and methods.

Both CHART and MHQP were developed and continue to be governed by a broad set of stakeholders that include central roles for providers. One of CHART's major stakeholders from the outset has been the California Hospital Association, and hospital representatives have always played an active role in the CHART steering committee that selects and develops performance measures and data methodology. Before new measures are reported on CalHospitalCompare, there is a pilot phase where providers and other stakeholders can review and raise any concerns about the preliminary data and methodology. The feedback received in this pilot phase sometimes results in changes to data collection or data modeling approaches to gain more widespread acceptance among the stakeholders.

One area of quality reporting that often meets with provider resistance is risk adjustment for outcomes measures, such as mortality rates following bypass surgery. Providers often question whether the particular risk-adjustment method used adequately captures differences in patient mix across providers, and sometimes providers advocate for risk-adjustment methods that exclude the outliers (the sickest, costliest cases). CHART dealt with this issue by calculating performance on outcome measures using different risk-adjustment models and demonstrating to the hospitals that their ratings relative to their peers generally did not change significantly under one risk-adjustment method versus others.

One measure of the extent of hospital buy in to the CHART program is the financial support that the program now receives from hospitals. As of 2009, each of the 240 hospitals participating in CHART makes voluntary financial contributions to the program.

Since its inception, the MHQP program has had active participation from the Massachusetts Medical Society, and each measure was selected through a collaborative process that included input from physicians. Before measures are reported publicly, MHQP sends datasets to physicians, allowing them to review the ratings and notify the program if any data appear inconsistent. For example, when MHQP circulated data to physicians on a new measure on sore-throat testing and treatment, many physicians notified MHQP of data inconsistencies that, upon investigation, resulted from coding changes. This feedback prompted MHQP to delay public reporting of the measure until the data errors were corrected.

Ensuring High-Caliber Data

How accurately data are abstracted, coded, aggregated, audited, validated and reported can profoundly affect the usefulness of performance ratings. If two quality transparency programs report the same measures, one can have a much greater positive impact by devoting resources to such activities as training vendors and staff at provider sites to collect data in an accurate, standardized manner; auditing sufficient samples of records; and validating datasets by checking for omissions, misclassifications and other errors.

Many of the performance measures collected by CHART and publicly reported by CalHospitalCompare are identical to measures reported by hospitals to the Joint Commission and the Centers for Medicare and Medicaid Services (CMS), but CHART appears unique in the steps it takes to improve data quality, including (1) providing training and certification to data vendors and hospital staff to ensure standardized data abstraction and coding within and among facilities; (2) validating datasets to identify problems such as missing data and misclassification errors; and

(3) thorough auditing.

When CHART began collecting performance data, hospitals sent CHART the same datasets they had been sending to the Joint Commission and CMS, yet CHART's validation tests on these datasets found major errors that had previously gone undetected by other organizations. These errors included missing data that should have been present (e.g., months of missing data for major domains in large acute care hospitals) and obvious misclassification errors (e.g., maternity performance results reported for hospitals not offering maternity services). CHART contacted the data vendors, which were able in many cases to trace the data anomalies to coding errors and to rectify the problems with relative ease.

In terms of data auditing, current CMS rules require only five patient records be audited per hospital per quarter, no matter the number of patients the hospital treats. CHART takes a more rigorous though flexible auditing approach. Instead of specifying a fixed number or proportion of patient records to be audited per hospital, CHART aims to develop measure-specific audit strategies thorough enough to convince program managers and stakeholders of the data's accuracy and reliability.

CHART varies the probability that a particular hospital will be audited based on the prior performance of that hospital on each specific measure, so that hospitals with "superior" or "poor" scores will be more likely to be audited than hospitals with "average" scores. In addition, CHART audits each measure independently, because program managers have observed that hospitals may collect very accurate data for some measures but misinterpret the data collection process for other measures. Before a new domain of measures is publicly reported, CHART conducts a pilot round of data collection followed by a thorough audit. For example, before intensive care unit (ICU) outcome and process measures were added, CHART audited 20 patient charts at about a quarter of the participating hospitals, with all high and low outliers chosen for audit, as well as a random selection of "average" hospitals.

The data reported by MHQP come primarily from the Healthcare Effectiveness Data and Information Set (HEDIS) col-

lected by health plans for the National Committee for Quality Assurance (NCQA), which already has standardized data collection methods and requires plans to undergo a compliance audit by an independent auditor. As a result, MHQP does not need to conduct the same training and auditing practices that CHART undertakes.

However, to ensure data quality, the program uses an independent auditor to check its own methods of aggregating data from the individual physician level to the medical group level. MHQP also developed a methodology to adjust administrative data to better align them with data from patient chart reviews.² MHQP also ensures that physicians are assigned to the correct medical group by seeking verification of physician information from each medical group.

Presenting Consumers with Meaningful Quality Information

Simply providing consumers with an abundance of quality data is insufficient to facilitate informed decision making. Instead, research suggests that the data must be "evaluable" and presented in a format that allows consumers to process the information and correctly interpret its meaning.³

Consumers find performance measures most useful when the information is presented to them as grades or ratings, conveyed in the form of words, stars or symbols.⁴ Presenting only numerical point estimates, confidence intervals or bar charts leaves many consumers confused about whether the differences across providers are significant. In addition, presenting ratings where almost all providers fit into the "average" category leaves consumers frustrated. An alternative—using multiple benchmarks to rank providers—helps to create meaningful categorizations of high and low performers that consumers find more useful.

CalHospitalCompare has dealt with these issues by (1) developing multiple benchmarks for each performance measure; and (2) developing a five-point scale for hospital performance on each measure, by comparing each hospital's performance to the benchmarks. The benchmarks are specific to each condition or domain, but for most measures except patient experience, the top 10 percent of national performance



Simply providing consumers with an abundance of quality data is insufficient to facilitate informed decision-making.

Instead, research suggests that the data must be "evaluable" and presented in a format that allows consumers to process the information and correctly interpret its meaning.



When consumers are presented with many separate performance measures, they may need help in aggregating these measures in a meaningful way.

is used as the high benchmark, the national average is used as the middle benchmark, and performance 10 percent below the national average is used as the low benchmark. For patient experience measures, national benchmarks do not yet exist, so the CalHospitalCompare hospitals are compared to one another, using the 10th, 50th and 90th percentiles as the three benchmarks.

For each measure, the rating for a hospital is determined based on where the confidence interval for the hospital's performance estimate falls relative to the benchmarks. For example, if a hospital's entire confidence interval falls below the low benchmark, the hospital is assessed as "poor," but if the hospital's confidence interval straddles the low and middle benchmarks, the hospital is rated "below average."⁵ By comparing hospitals' performance confidence intervals to the multiple benchmarks, CalHospitalCompare is able to provide consumers with ratings on a five-point scale, from "superior" to "poor." This approach is augmented by color-coded icons (e.g., green for superior, yellow for average, red for poor) that have been shown in consumer testing to be effective in reinforcing the ratings in consumers' minds.

In contrast, other public quality transparency Web sites—including CMS's Hospital Compare—display only point estimates of performance on process measures (e.g., percent of heart attack patients given a beta blocker), with no accompanying grades or ratings to interpret for consumers whether, for example, the 90 percent attained by Hospital A is different in any meaningful way from the 93 percent achieved by Hospital B.

In reporting outcome measures, such as mortality, Hospital Compare conforms to a strict rule of detecting and reporting a difference only if an estimate is at least two standard deviations away from the mean. Using this stringent approach to identify superior and inferior providers means that, typically, almost all providers (95%) will land in the average category, and only 2.5 percent (or 1 in 40) providers will be in each of the superior and inferior categories—an approach that consumers are likely to find frustrating and unhelpful in steering them toward or away from particular hospitals. Here again, CalHospitalCompare's use of multiple benchmarks and a five-point rating scale

helps to create enough distinct categories that consumers can identify superior and inferior performers for each measure.

Prior to the formation of CHART, conventional wisdom held that most providers would strongly resist performance ratings that failed to use the strict two-standard-deviation rule for detecting differences.⁶ However, because CHART's benchmarking and rating systems were developed with hospital input from the start, they have been widely accepted in California's hospital community. In addition, many hospitals, which would have been lumped into the average category with almost all of their peers under the conventional methodology, saw an opportunity to distinguish themselves with "superior" or "above-average" designations.

Like CalHospitalCompare, MHQP also presents provider ratings in a format easy for consumers to understand. MHQP assesses each medical group or practice site on a scale of one to four stars, using three benchmarks—the national 50th percentile, the national 90th percentile and the MHQP Massachusetts statewide rate. The majority of the patient experience measures use cut-points at the 15th, 50th and 85th percentiles among all physician groups surveyed. For some measures, MHQP adds a fifth star indicating whether the medical group reached a target score that was set for the group prior to the measurement year. That target was the score that the top 25 percent of all Massachusetts medical groups had reached or exceeded in the previous year.

When consumers are presented with many separate performance measures, they may need help in aggregating these measures in a meaningful way.⁷ CalHospitalCompare combines related measures into composite measures to ease interpretation of results for consumers, though the program steers clear of providing overall scores, as provider performance can vary substantially across domains.

For example, CalHospitalCompare combines all the separate patient experience measures into one composite patient experience rating for the hospital. (However, CalHospitalCompare reports patient experience ratings separately for medical, surgical and maternity patients, because these patients' experiences are considered too different from one another

to be grouped together meaningfully.) If CalHospitalCompare users are interested in greater detail, they can click on a button on the screen to view results for each individual patient experience measure. In contrast, CMS's Hospital Compare provides no composite ratings for any performance domain; rather, it reports point estimates for each of the 10 patient experience measures separately, with no score or rating attached for any measure. Such a disaggregated approach risks overwhelming consumers with too much information and providing too little guidance about how to interpret sometimes contradictory results.

Influencing Providers to Improve

Quality transparency initiatives tend to view all consumers as their target audiences, but the true consumer audience for any given program is likely to be limited to those consumers who need and use the providers whose performance is reported by the program. For example, a program that reports performance for inpatient services is unlikely to attract consumers who don't have an imminent need for such services, and national data indicate that only 8 percent of Americans are admitted to hospitals on an inpatient basis annually.⁸ In addition, even when consumers are part of the true target audience because they are in the market for a given service, many do not believe that quality differs enough across providers for these differences to have concrete, serious—even life-or-death—consequences.⁹ As a result, many consumers may see little or no need to use a quality transparency program, even when they are aware of a program that rates providers relevant to their needs.

Given such challenges, it may not be realistic to expect that consumer use of provider quality comparisons will move enough market share to motivate providers to improve their quality. But despite any shortcomings in the consumer-choice model of quality improvement, quality transparency initiatives can still have a powerful impact on quality through the “sunshine effect,” by which providers—seeing their quality metrics publicly compared to their competitors—are motivated to improve quality to protect their public and professional reputations and to adhere to professional norms.¹⁰

Recognizing that the sunshine effect can

be a powerful driver of quality improvement, but that providers seeking to improve quality need access to more granular data than the information publicly reported on CalHospitalCompare, CHART provides participating hospitals with patient-level spreadsheets for all of the performance measures, including performance on measures still under development and deemed not yet ready for public release.

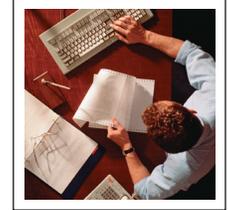
Similarly, MHQP provides individual physicians and medical groups with detailed data, including performance on preliminary measures under development. According to CHART program managers, the program has had a pronounced effect on hospitals' quality improvement initiatives, even though consumer awareness and use of CalHospitalCompare remain modest.

Implications for Other Initiatives

The approaches used by CalHospitalCompare and MQHP can be replicated by other quality transparency initiatives. Some would be more time-consuming and costly to adopt than others. Achieving provider buy in and collaboration, for example, can be an unwieldy and time-intensive process not only at the inception of the program, but also on an ongoing basis, as new performance measures and data methods are considered. Similarly, thorough data auditing to ensure data quality may require a greater commitment of resources than many quality transparency programs are able or willing to make.

Other approaches outlined are simpler and less costly for existing quality transparency programs to incorporate. For example, some of the data validation measures taken by CHART, such as performing logical checks for missing or misclassified data, can be adopted by other programs at relatively modest cost. In addition, the use of multiple benchmarks and the development of four- or five-point rating scales based on those benchmarks are features that can be adopted by programs such as Hospital Compare to simplify data presentation and make the information more useful for consumers.

It may make sense for other quality transparency initiatives to adopt simpler, low-cost measures first, before tackling more difficult, resource-intensive ways to improve the effectiveness of the programs. Ultimately,



According to CHART program managers, the program has had a pronounced effect on hospitals' quality improvement initiatives, even though consumer awareness and use of CalHospitalCompare remain modest.



Data Source

HSC researchers conducted a literature review and examined several health care quality transparency initiatives, including the two profiled in this Issue Brief. To learn more about CalHospitalCompare, researchers reviewed documentation at www.cal-hospitalcompare.org and the Web site of the California Hospitals Assessment and Reporting Taskforce at chart.ucsf.edu. For Massachusetts Health Quality Partners, researchers reviewed documentation at www.mhqp.org. Researchers also interviewed representatives from both programs. A two-person research team conducted each interview, and notes were transcribed and jointly reviewed for quality and validation purposes.

ISSUE BRIEFS are published by the Center for Studying Health System Change.

600 Maryland Avenue, SW, Suite 550
Washington, DC 20024-2512
Tel: (202) 484-5261
Fax: (202) 484-9258
www.hschange.org

President: Paul B. Ginsburg



however, quality transparency programs are unlikely to have substantial influence on quality improvement unless they gain widespread stakeholder acceptance—especially from the providers being rated—and seriously commit to improving the caliber of the quality data reported. ■

Notes

1. MHQP also includes several outcomes measures but currently reports only statewide averages rather than physician group ratings for these measures because of small sample sizes.
2. As a requirement for NCQA accreditation, health plans conduct chart reviews for a sample of enrollees to determine the accuracy of claims data. MHQP obtains data from each health plan on the differential between patient chart data and claims data, and applies this differential as an adjustment factor for each performance measure.
3. Hibbard, Judith H., and Ellen Peters, “Supporting Informed Consumer Health Care Decisions: Data Presentation Approaches that Facilitate the Use of Information in Choice,” *Annual Review of Public Health*, Vol. 24 (2003).
4. Gerteis, Margaret, et al., “Testing Consumers’ Comprehension of Quality Measures Using Alternative Reporting Formats,” *Health Care Financing Review*, Vol. 28, No. 3 (Spring 2007).
5. This methodology is explained in detail at the CalHospitalCompare Web site. See <http://www.calhospitalcompare.org/Resources-and-Tools/Choosing-a-Hospital/About-the-Ratings.aspx>.
6. Dudley, R. Adams, Diane Rittenhouse and Richard Bae, *Creating a Statewide Hospital Quality Reporting System*, California HealthCare Foundation, The Quality Initiative (February 2002).
7. Interview with Judith Hibbard.
8. Adams, Patricia F., Patricia M. Barnes and Jackline L. Vickerie, “Summary Health Statistics for the U.S. Population: National Health Interview Survey, 2007,” National Center for Health Statistics,

Vital Health Statistics, Series 10, No. 238 (November 2008).

9. Hibbard, Judith H., and L. Gregory Pawlson, “Why Not Give Consumers a Framework for Understanding Quality?” *Joint Commission Journal on Quality and Safety*, Vol. 30, No. 6 (June 2004).
10. Marshall, Martin N., et al., “The Public Release of Performance Data: What Do We Expect to Gain? A Review of the Evidence,” *Journal of the American Medical Association*, Vol. 283, No. 14 (April 12, 2000).

Funding Acknowledgement: This work was supported by the Robert Wood Johnson Foundation.