Title
Prediction of Hospital Acute Myocardial Infarction and Heart Failure 30-Day Mortality Rates Using Publicly Reported Performance Measures

Journal
Journal for Healthcare Quality, 35(2)

ISSN
1062-2551

Authors
Aaronson, David S
Bardach, Naomi S
Lin, Grace A
et al.

Publication Date
2013-03-01

DOI
10.1111/j.1945-1474.2011.00173.x

Peer reviewed

Published in final edited form as:
J Healthc Qual. 2013 ; 35(2): 15–23. doi:10.1111/j.1945-1474.2011.00173.x.

Prediction of Hospital Acute Myocardial Infarction and Heart Failure 30-Day Mortality Rates Using Publicly Reported Performance Measures

David S. Aaronson, MD, **Naomi S. Bardach, MD**, **Grace A. Lin, MD, MAS**, **Arpita Chattopadhyay, PhD**, **L. Elizabeth Goldman, MD, MCR**, and **R. Adams Dudley, MD, MBA**

Abstract

Objective—To identify an approach to summarizing publicly reported hospital performance data for acute myocardial infarction (AMI) or heart failure (HF) that best predicts current year hospital mortality rates.

Setting—A total of 1,868 U.S. hospitals reporting process and outcome measures for AMI and HF to the Centers for Medicare and Medicaid Services (CMS) from July 2005 to June 2006 (Year 0) and July 2006 to June 2007 (Year 1).

Design—Observational cohort study measuring the percentage variation in Year 1 hospital 30-day risk-adjusted mortality rate explained by denominator-based weighted composite scores summarizing hospital Year 0 performance.

Data Collection—Data were prospectively collected from hospitalcompare.gov.

Results—Percentage variation in Year 1 mortality was best explained by mortality rate alone in Year 0 over other composites including process performance. If only Year 0 mortality rates were reported, and consumers using hospitals in the highest decile of mortality instead chose hospitals in the lowest decile of mortality rate, the number of deaths at 30 days that potentially could have been avoided was 1.31 per 100 patients for AMI and 2.12 for HF ($p < .001$).

Conclusion—Public reports focused on 30-day risk-adjusted mortality rate may more directly address policymakers' goals of facilitating consumer identification of hospitals with better outcomes.

Keywords

outcomes; process measures; public reporting; quality improvement

Introduction

Public reporting of hospital performance has become a common strategy for monitoring and improving the quality of healthcare. Such reports are intended to stimulate quality improvement among providers by identifying areas where quality is lacking and by

For more information on this article, contact R. Adams Dudley at adams.dudley@ucsf.edu.

providing motivation for protection of reputation or market share (Berwick, James, & Coye, 2003; Hibbard, 2008a). In addition, public reporting may be used by consumers to select higher quality providers. Although there is some debate about whether or not consumers use hospital performance data to choose their site of care (Faber, Bosch, Wollersheim, Leatherman, & Grol, 2009; Romano & Zhou, 2004), guiding consumers' choices continues to be a widely stated rationale for public reporting (CMS, 2008; Hibbard, 2008b).

The Centers for Medicare and Medicaid Services (CMS) was mandated by the Deficit Reduction Act of 2005 to maintain a consumer-targeted website publicly reporting hospital performance on quality of care, which CMS has stated will "empower consumer choice" (CMS, 2008). The current version of the web-site presents hospital-specific mortality rates for acute myocardial infarction (AMI) and heart failure (HF) and data about several process measures for each condition. Thus, in order to use the information presented to choose a hospital, a consumer must create some composite assessment of overall hospital performance, although the consumer generally is not explicitly aware of doing this (Hibbard, 2008b).

However, research has shown that consumers' decision-making capacity is reduced as the number of variables for which data are presented increases (Hibbard, Slovic, & Jewett, 1997; Vaiana & McGlynn, 2002). Simplifying the comparative quality data for easy and meaningful interpretation by consumers has been suggested as a future goal of hospital performance reporting. CMS and (CMS, 2004) the Agency for Healthcare Research and Quality (AHRQ; Hibbard, 2008a) have advocated the strategy of reducing consumers' cognitive burden by explicitly calculating composite scores from the available information and presenting those, rather than leaving consumers to create their own implicit composites.

Furthermore, consumers are generally more concerned about actual results of care (e.g., mortality) than technical processes of care (Lansky, 1998). Therefore, if the purpose is to guide consumer choices, the data presented should predict outcomes at individual hospitals during the time period in which the consumer is making a choice. However, CMS data necessarily are historical (i.e., from a period before the time when the consumer might be choosing a hospital), and it is not clear how predictive historical data are of current hospital performance. To our knowledge, no study has examined how well hospital performance data from the previous year on AMI and HF—whether as individual data elements or in composite form—predict mortality rates in the current year. Other work has focused on historical mortality (with or without incorporating volume) predicting "current" surgical or NICU mortality (Birkmeyer, Dimick, & Staiger, 2006; Dimick, Staiger, Baser, & Birkmeyer, 2009; Luft & Romano, 1993; Rogowski et al., 2004).

Our aim in this study was to identify the approach to summarizing the information available at the time a consumer would be making a choice that best predicts current year hospital mortality rates and would thus be most useful in achieving CMS' stated goal of guiding consumer choices. We performed this study for AMI and HF admissions because those are the only conditions for which process and mortality data were available on a national sample of hospitals for a 2-year period (pneumonia data were available for only 1 year). Although we realize consumers usually cannot choose their hospital during AMI or an acute

NIH-PA Author Manuscript

NIH-PA Author Manuscript

exacerbation of HF, they are often able to select a hospital for chronic management of their cardiac disease and that selection may affect the hospital to which they are admitted for the acute event. In addition, the lessons from studying AMI and HF may be applicable to other conditions in which choice is more frequently feasible.

Methods

Study Design and Sample

We performed an observational cohort study of all acute care hospitals reporting performance on process and outcome measures for AMI and HF to CMS from July 2005 to June 2007. The chosen time period provided the most current publicly available data for hospital performance on these measures. For this study, July 2005 to June 2006 is defined as "Year 0" or the prior year, and July 2006 to June 2007 is defined as "Year 1" or the current year. We excluded hospitals with less than 25 eligible patients for measurement in either year as CMS deems performance rates for these hospitals to be unstable (CMS, 2009). Details of CMS' methodology for measuring hospital performance are reported on http://www.hospitalcompare.hhs.gov.

Hospital descriptive characteristics, including bed size, teaching status, availability of cardiac intensive care unit, and proprietary status, were obtained from the 2006 American Hospital Association database.

Quality Measures

We included process measures that were reported for at least 80% of hospitals in order to calculate a composite score for as many hospitals as possible. Process measures for AMI included aspirin at arrival and at discharge; beta blocker at arrival and at discharge; smoking cessation counseling; and angiotensin-converting enzyme (ACE) inhibitor or angiotensin receptor blocker (ARB) for left ventricular systolic dysfunction. Excluded process measures for AMI were thrombolytic agent received within 30 min, and primary percutaneous coronary intervention received within 120 min (this has been reduced to 90 min for more recent years) of arrival (only available from 63% and 54% of hospitals, respectively). All HF process measures were included as follows: assessment of patients' left ventricular function, ACE inhibitor or ARB for left ventricular dysfunction, and smoking cessation counseling.

Hospital 30-day risk-adjusted mortality rates for AMI and HF were abstracted from CMS' website (CMS, 2009). The risk-adjustment models have been validated against other models of risk adjustment and found to have high correlation over time (Krumholz et al., 2006).

Calculating Composite Measures

There are many ways to construct a composite measure. For example, composite scores may be "compensatory" (a poor score on one individual measure can be compensated by a good score on another measure) or "conjunctive" (all measures must be met to achieve a good score). In addition, individual measures forming a composite score may be transformed and/or weighted differently.

We used the composite construction method recommended by CMS (CMS, 2004; Shwartz et al., 2008)—a compensatory composite calculated using denominator-based weights—to create two disease-specific composite measures for both AMI and HF: process performance only and overall performance (process plus mortality). The denominator-based approach divides the number of times the processes were followed in a hospital (e.g., for AMI care) by the total number of times patients were eligible for each AMI measure. This approach also is currently used by The Joint Commission to facilitate consumer interpretation (Commission, 2008).

To construct an overall composite score, we transformed mortality rates into survival rates (1— mortality rate) so that the outcome measure had the same directionality as the process measures (i.e., closer to 100% represented better performance). We then applied equal weighting based upon the number of process and outcome measures used to obtain the composite score, as recommended by CMS. For example, our AMI overall composite score was constructed of six process measures and one outcome measure, so a weight of six of seven was applied to the process measure composite and one of seven to mortality rate. Then the two scores were added to give an overall composite performance score. The performance of the composite scores using Year 0 data to predict Year 1 mortality rates for each hospital was compared to using Year 0 mortality rates alone.

Sensitivity Analyses

Although CMS recommends the composite weighting approach described above, we recognized that a consumers’ own weighting preferences may vary. As a sensitivity analysis, therefore, we tested alternative combinations of weighting for the components comprising a composite score. We tested different weights by incrementally increasing the weight on the survival rate and decreasing the weight on processes to determine whether this affected the ability of the overall composite to predict Year 1 mortality rates.

We also analyzed whether the exclusion of the two process measures for AMI that were missing for many hospitals—thrombolytic agent received within 30 min, and primary percutaneous coronary intervention received within 120 min of arrival—altered our findings. We recreated our process and overall composite scores using the subset of hospitals that reported on one or both of these process measures, adjusting the composite weighting to account for the inclusion of additional quality measures.

Because there was minimal variation among hospitals for some of the included process measures, we hypothesized that identifying hospitals with very low composite scores might have more predictive power than using the scores as continuous variables, since hospitals with very low composite scores would have the most extreme deviations in process performance from average. Therefore, we also tested whether a model capturing performance as a binary variable indicating whether a hospital was in the lowest or highest decile of hospital composite scores better predicted Year 1 risk-adjusted hospital mortality rates than the models using composite scores as continuous variables. We also treated each of the 10 deciles of hospital performance as dummy variables to further test the predictive value of both extremes of the performance distribution.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Statistical Analysis

We calculated the mean performance rate with standard deviation for individual process measures, the process composite score, and the overall (processes and mortality) composite score, along with 30-day risk-adjusted hospital mortality rate for AMI and HF in Year 0. Because we wanted to know how useful the information on different measures was for informing patients of their mortality risk, we used linear regression to determine the strength of the relationship between disease-specific hospital 30-day risk-adjusted mortality rate in Year 1 and Year 0 disease-specific process composite score, overall composite score, and mortality rate alone. In these models, Year 1 mortality rate was the dependent variable, and Year 0 data were used to calculate the independent variables. Regression models reveal the relationship between two or more variables by estimating the line that best fits the observed data. A measure of the fit is provided by the coefficient of determination (*r*-squared), which reveals how well the variation in the independent variable explains the variation in the dependent variable. An *r*-square value of 0 implies that none of the variation in the dependent variable is explained by the independent variable, and an *r*-square value of 1 that means 100% of the variation in the dependent variable is explained by the independent variable.

To demonstrate the potential clinical impact of consumers using different components of the available performance data to choose a hospital, we calculated the number of potentially avoidable deaths at 30 days for AMI and HF if consumers were to choose hospitals in the top decile instead of the bottom decile, based on each approach to summarizing the available data—individual process scores, process composite score, overall composite score, and 30-day risk-adjusted mortality rate (Replogle & Johnson, 2007). The difference in arithmetic means was used (i.e., assigning equal weights to each hospital) because the hospital-specific rates were already adjusted for hospital level differences in patient characteristics, volume, etc. In this case, because we were comparing the means between two categories of hospitals (those in the top decile and those in the bottom decile, we used the analysis of variance (ANOVA) to determine statistical significance. ANOVA is a general technique that uses the *F*-statistics to test the hypothesis that the means among two or more populations are equal.

The University of California, San Francisco institutional review board determined that this study was exempt from review. Analyses were conducted using Stata/SE version 10 (Stata-Corp, College Station, TX). All reported *p* values are two-sided and considered significant at .05.

Results

Sample and Measures

The sample for our analyses included 190,637 (Year 0) and 178,586 (Year 1) patients with AMI and 326,627 (Year 0) and 303,338 (Year 1) patients with HF from 1,868 hospitals. Hospital characteristics in Year 0 are shown in Table 1.

Rates of hospital performance for individual processes, the process composite, and the overall composite measures as well as 30-day risk-adjusted mortality rate for AMI and HF are shown in Table 2. Mean hospital 30-day risk-adjusted mortality rate was 16.3 (*SD* 1.6) and 11.0 (*SD* 1.4) per 100 patients presenting with AMI and HF, respectively.

Relationship Between Year 0 Hospital Performance and Year 1 30-day Risk-Adjusted Mortality Rate

The hospital process composite and overall composite from Year 0 accounted for a small but statistically significant portion of the variation in 30-day risk-adjusted hospital mortality in the following year for AMI (3.6% and 3.9%, both $p < .001$), but not for HF (3.4% and 3.2%; both $p > .05$; Figure 1). In contrast, the explanatory power of 30-day risk-adjusted mortality in the current year was several-fold greater than both the process measure alone and the composite measures. Percentages of variation in the following year's 30-day risk-adjusted mortality explained by current year mortality was 8.6 ($p < .001$) for AMI and 17.9% for HF, respectively ($p < .001$; Figure 1).

We calculated the number of potentially avoidable deaths at 30 days in Year 1 associated with a patient choosing a hospital in the top versus bottom decile of Year 0 composite or mortality performance data to give a clinically relevant assessment of the relative advantage to using each type of measure for decision making (Table 3). For AMI, using Year 0 30-day risk-adjusted mortality rate alone produced a larger reduction in potentially avoidable deaths than using process composite or overall composite scores though all were significant (1.31 lives per 100 patients vs. 0.61 and 0.56 lives per 100 patients, respectively, $p < .001$). Consumer selection of a hospital in Year 1 based upon Year 0 hospital AMI 30-day risk-adjusted mortality rate alone would find a reduction in potentially avoidable deaths of 1.31 lives per 100 patients ($p < .001$). For HF, use of Year 0 hospital process (−0.03 lives per 100 patients, $p = .68$) or overall (−0.08 lives per 100 patients, $p = .63$) composite measures to select hospitals did not yield any significant benefit in the risk of 30-day risk-adjusted mortality in Year 1. However, using the Year 0 30-day mortality rate for HF yielded a significant reduction in Year 1 potentially avoidable deaths of 2.12 lives per 100 patients ($p < .001$).

Sensitivity Analyses

We varied the weights in the overall composite measures to reflect the fact that consumers might assign a wide range of weights to processes relative to mortality. However, no process composite weighting between 0.01 and 1.00, explained as much variation in Year 1 mortality as using the Year 0 mortality rate alone (equivalent to a process weight of zero).

We also recalculated the process and overall performance composite scores for AMI with the inclusion of the two measures for AMI that were excluded in our initial analysis (thrombolytic agent received within 30 min and primary percutaneous coronary intervention received within 120 min of arrival) for 618 hospitals that reported both measures. The weighted composites for those hospitals did not explain a significant percentage of the variation (5.2%, $p = .35$ and 5.5%, $p = .22$, respectively) in Year 1 mortality.

Finally, to test the hypothesis that extremes of performance might be most predictive when there was little overall variation in process measure performance, we explored the impact of using a model replacing the process and overall composites with a binary variable indicating whether a hospital was in the lowest or highest decile of hospital performance on these composites in Year 0 to predict Year 1 mortality. These models had no additional predictive power compared to models using the Year 0 mortality alone. For AMI, the model using the binary variable indicating bottom or top decile hospital performance on the process and overall composite explained only 3.2% (p = .03) and 3.1% (p = .03) of the variation in Year 1 mortality rates, respectively. For HF, the model using the binary variable indicating bottom or top hospital performance on the process and overall composite did not explain a significant portion of the variation in Year 1 mortality rates (1.1%, p = .58 and 1.0%, p = .91, respectively). A sensitivity analysis evaluating the predictive ability of the extremes in the hospital performance on the process and overall composites, which used dummy variables for each decile of performance, did not improve the predictive value of the data compared to hospital performance analyzed as a continuous variable.

Discussion

There is a growing body of evidence that consumers often find hospital quality reports difficult to interpret (Faber et al., 2009). In response, many organizations, including CMS, have advocated using composite scores to reduce the amount of data presented. This is the first study to assess whether a composite score created from *Hospital Compare* data would give consumers better information about their expected outcomes than the individual performance measures. We found that 30-day mortality alone had better, though still small, predictive ability for subsequent hospital mortality rates than tested composite measures for AMI and HF.

It is not a surprising finding that Year 0 process measures explain little of the variation in Year 1 mortality, given prior research that even same year process measure performance explains a relatively small portion of mortality (i.e., process performance in Year 0 does not explain much of mortality variation in Year 0; Bradley et al., 2006; Werner & Bradlow, 2006). Other approaches to creating composite ratings of hospitals that include variables other than mortality have also shown little ability to predict future mortality performance (Chen, Radford, Wang, Marciniak, & Krumholz, 1999; Krumholz, Rathore, Chen, Wang, & Radford, 2002; Wang et al., 2007). Furthermore, we found that consumer use of Year 0 30-day mortality rates alone to choose better performing hospitals would have yielded potentially better outcomes (in the form of greater number of deaths avoided) than using composite scores. Taken together, the prior research and our study suggest that to create a consumer-focused hospital performance report that is simpler and more meaningful (for CMS' stated purpose of guiding consumer choice [CMS, 2008]), including only historical outcomes such as 30-day risk-adjusted mortality rates in the report may be more effective than creating a composite of mortality and process data.

Our findings should not be interpreted to suggest that it is not worthwhile to collect process measure data. Many of the currently used process measures have been shown to affect mortality in randomized controlled trials, and the weak association with mortality in an observational study does not negate those findings. Rather, process measure data can and should be used for other important purposes, such as reporting to hospitals their own performance relative to benchmarks, with the expectation of hospitals using the data in internal quality improvement programs (Jencks, Huff, & Cuerdon, 2003). Additionally, if new process measures are developed and tested in clinical trials that show better prediction of outcomes, the usefulness of including these process measures in reports to consumers should be reevaluated.

We realize some may be uncomfortable with the idea of publicly reporting only risk-adjusted 30-day mortality rates. This discomfort may reflect continuing uncertainty about whether differences in risk-adjusted mortality rates truly reflect differences in hospital performance or are related to other variables, such as unmeasured patient clinical or socioeconomic status variables, access to care, or hospital characteristics. Additionally, the risk-adjustment model developed by CMS adjusts for the existence of comorbidities, but not their severity (Krumholz et al., 2006). Although researchers (and CMS) should continue to seek better methods of risk adjustment, we feel that it is unlikely that improved risk-adjustment methodologies would change our findings (in fact, it might increase the relative predictive power of Year 0 mortality, because it would be a truer measure of the hospital's performance in Year 0). Notwithstanding, we acknowledge that unmeasured confounders have the potential to affect results of any analysis.

Furthermore, the reality is that public reporting (including of risk-adjusted mortality) is already occurring at the national level and in many states as well, with the stated goal of directing patients to hospitals with better outcomes (CMS, 2008). Thus, the question of what subset of available performance measures will best predict future mortality rates remains relevant, as it is a major priority for many policymakers and can have an impact on all providers.

There are other limitations to this study. Our dataset consists of publicly available measurements of hospital quality and does not allow for the examination of other outcomes potentially affected by process performance and of interest to consumers, such as hospital inpatient and 1-year risk-adjusted mortality (Saposnik et al., 2008). However, these outcomes are not currently reported on *Hospital Compare* and thus are not likely currently used by consumers in their assessments of hospital quality. Thirty-day risk-adjusted readmission rates are just now being released by CMS and were not available for incorporation in this study. In addition, CMS has changed its reporting of outcomes data from 12- to 36-month aggregate hospital rates, adding another challenge for investigators and consumers alike as they try to understand their utility. The addition of a time lag in reporting of mortality data by CMS will not likely improve its use as a predictor of "current" mortality rates (Luft & Romano, 1993). Further research is needed to assess whether the introduction of a time lag in the reporting of outcomes data will diminish our findings. Another important caveat to our potentially avoidable deaths calculation is that it must be significantly discounted for feasibility issues, such as the fact that patients might not live near a top decile hospital. Further research that incorporates these feasibility considerations is needed for more precise estimates on the overall potential impact on mortality.

Lastly, as the length of time a set of process measures has been included in CMS initiatives increases, more hospitals reach a ceiling of process measure performance (Fonarow &

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Peterson, 2009; Normand, Wolf, & McNeil, 2008). As the variation in process measure performance declines, it becomes less likely that incorporating this information will improve our ability to predict hospital outcomes. This may contribute to the minimal predictive benefit of hospital process measure performance in our study. However, even using an indicator for falling into the bottom decile of hospital performance did not produce models that predict Year 1 mortality nearly as well as models using Year 0 mortality alone.

Despite the plethora of available public performance reports, public reporting of quality measures has not had a great impact on consumer behavior (Fung, Lim, Mattke, Damberg, & Shekelle, 2008). Our study demonstrates that a composite score created as CMS, AHRQ, and others recommend to improve comprehension of data has little predictive value to the consumer trying to identify a hospital that will improve his or her chances of surviving an AMI or an episode of HF. Mortality rate from the prior year alone is a stronger predictor of 30-day risk-adjusted mortality rate in the consumer's current year. Thus, public reports focused on 30-day risk-adjusted mortality rate would likely reduce the cognitive burden on consumers and may more directly address policymakers' goals of facilitating consumer identification of hospitals with better outcomes.

Acknowledgments

Dr. Dudley's work on this project was supported by an Investigator Award in Health Policy from the Robert Wood Johnson Foundation.

References

Berwick DM, James B, Coye MJ. Connections between quality measurement and improvement. Medical Care. 2003; 41(1 Suppl):I30–I38. [PubMed: 12544814]

Birkmeyer JD, Dimick JB, Staiger DO. Operative mortality and procedure volume as predictors of subsequent hospital performance. Annals of Surgery. 2006; 243:411–417. [PubMed: 16495708]

Bradley EH, Herrin J, Elbel B, McNamara RL, Magid DJ, Nallamothu BK, et al. Hospital quality for acute myocardial infarction: Correlation among process measures and relationship with short-term mortality. Journal of the American Medical Association. 2006; 296:72–78. [PubMed: 16820549]

Chen J, Radford MJ, Wang Y, Marciniak TA, Krumholz HM. Do "America's Best Hospitals" perform better for acute myocardial infarction? New Eng-land Journal of Medicine. 1999; 340:286–292.

CMS. CMS HQI Demonstration Project: Composite quality score methodology overview. 2004. Retrieved November 26, 2009, from http://www.cms.hhs.gov/HospitalQual-ityInits/downloads/HospitalCompositeQualityScoreMethodologyOverview.pdf

CMS. Medicare enhances consumer information on hospital care. 2008. Retrieved November 26, 2009, from http://www.cms.hhs.gov/apps/media/press/release.asp?Counter=3244&intNumPerPage=10&checkDate=&checkKey=&srchType=1&numDays=3500&srchOpt=0&srchData=&keywordType=All&chkNewsType=1%2C+2%2C+3%2C+4%2C+5&intPage=&showAll=&pYear=&year=desc=&cboOrder=date&

CMS. Hospital Compare–-A quality tool provided by Medicare. 2009. Retrieved November 26, 2009, from http://www.hospitalcompare.hhs.gov/staticpages/help/hospital-resources.aspx

The Joint Commission. Improving America's Hospitals: The Joint Commission's Annual Report on Quality and Safety. Washington, DC: U.S. Government Printing Office; 2008. Retrieved from http://www.jointcommission.org/assets/1/6/2008_Annual_Report.pdf

Dimick JB, Staiger DO, Baser O, Birkmeyer JD. Composite measures for predicting surgical mortality in the hospital. Health Affairs (Millwood). 2009; 28:1189–1198.

Faber M, Bosch M, Wollersheim H, Leatherman S, Grol R. Public reporting in health care: How do consumers use quality-of-care information? A systematic review. Medical Care. 2009; 47:1–8. [PubMed: 19106724]

Fonarow GC, Peterson ED. Heart failure performance measures and outcomes: Real or illusory gains. Journal of the American Medical Association. 2009; 302:792–794. [PubMed: 19690314]

Fung CH, Lim YW, Mattke S, Damberg C, Shekelle PG. Systematic review: The evidence that publishing patient care performance data improves quality of care. Annals of Internal Medicine. 2008; 148:111–123. [PubMed: 18195336]

Agency for Healthcare Research and Quality. Best practices in public reporting. Washington, DC: Hibbard, J; 2008a. (publication No. HHSA290200710022T)Retrieved from http://www.ahrq.gov/qual/pubrptguide1.htm

Hibbard JH. What can we say about the impact of public reporting? Inconsistent execution yields variable results. Annals of Internal Medicine. 2008b; 148:160–161. [PubMed: 18195340]

Hibbard JH, Slovic P, Jewett JJ. Informing consumer decisions in health care: Implications from decision-making research. Milbank Quarterly. 1997; 75:395–414. [PubMed: 9290635]

Jencks SF, Huff ED, Cuerdon T. Change in the quality of care delivered to Medicare beneficiaries, 1998–1999 to 2000–2001. Journal of the American Medical Association. 2003; 289:305–312. [PubMed: 12525231]

Krumholz HM, Rathore SS, Chen J, Wang Y, Radford MJ. Evaluation of a consumer-oriented internet health care report card: The risk of quality ratings based on mortality data. Journal of the American Medical Association. 2002; 287:1277–1287. [PubMed: 11886319]

Krumholz HM, Wang Y, Mattera JA, Han LF, Ingber MJ, Roman S, et al. An administrative claims model suitable for profiling hospital performance based on 30-day mortality rates among patients with heart failure. Circulation. 2006; 113:1693–1701. [PubMed: 16549636]

Lansky D. Measuring what matters to the public. Health Affairs (Millwood). 1998; 17:40–41.

Luft HS, Romano PS. Chance, continuity, and change in hospital mortality rates. Coronary artery bypass graft patients in California hospitals, 1983 to 1989. Journal of the American Medical Association. 1993; 270:331–337. [PubMed: 8315777]

Normand SL, Wolf RE, McNeil BJ. Discriminating quality of hospital care in the United States. Medical Decision Making. 2008; 28:308–322. [PubMed: 18310529]

Replogle WH, Johnson WD. Interpretation of absolute measures of disease risk in comparative research. Family Medicine. 2007; 39:432–435. [PubMed: 17549653]

Rogowski JA, Horbar JD, Staiger DO, Kenny M, Carpenter J, Geppert J. Indirect vs direct hospital quality indicators for very low-birth-weight infants. Journal of the American Medical Association. 2004; 291:202–209. [PubMed: 14722146]

Romano PS, Zhou H. Do well-publicized risk-adjusted outcomes reports affect hospital volume? Medical Care. 2004; 42:367–377. [PubMed: 15076814]

Saposnik G, Fang J, O'Donnell M, Hachinski V, Kapral MK, Hill MD. Escalating levels of access to in-hospital care and stroke mortality. Stroke. 2008; 39:2522–2530. [PubMed: 18617667]

Shwartz M, Ren J, Pekoz EA, Wang X, Cohen AB, Restuccia JD. Estimating a composite measure of hospital quality from the Hospital Compare Database: Differences when using a Bayesian hierarchical latent variable model versus denominator-based weights. Medical Care. 2008; 46:778–785. [PubMed: 18665057]

Vaiana ME, McGlynn EA. What cognitive science tells us about the design of reports for consumers. Medical Care Research and Review. 2002; 59:3–35. [PubMed: 11877877]

Wang OJ, Wang Y, Lichtman JH, Bradley EH, Normand SL, Krumholz HM. "America's Best Hospitals" in the treatment of acute myocardial infarction. Archives of Internal Medicine. 2007; 167:1345–1351. [PubMed: 17620526]

Werner RM, Bradlow ET. Relationship between Medicare's hospital compare performance measures and mortality rates. Journal of the American Medical Association. 2006; 296:2694–2702. [PubMed: 17164455]

Biographies

David S. Aaronson, MD, University of California, San Francisco, is a practicing Urologist at Kaiser Permanente in Oakland. His major research interests include measurements of the appropriateness, quality, and safety of care

Naomi S. Bardach, MD, University of California, San Francisco, is a pediatrician at the Pediatric Urgent Care Clinic and San Francisco General Hospital pediatrics clinic. She is also an Adjunct Assistant Professor in Pediatrics at the UCSF School of Medicine. Her research focuses on improving hospital care for children by applying advanced biostatistics, study design, qualitative methods, and surveys to assess interventions that prevent readmissions

Grace A. Lin, MD, MAS, University of California, San Francisco, is a primary care internist and Assistant Adjunct Professor of Medicine at the UCSF School of Medicine. Her research focuses on the appropriateness of medical care provided to patients, shared decision making, and assessing the quality of decision making by doctors and patients

Arpita Chattopadhyay, PhD, University of California, San Francisco, is an Assistant Professor at the UCSF School of Medicine. Her doctoral training was in Demography and her broad research interest is to study demographic attributes associated with inequities in access, utilization or quality of care. Her major research interests include population health and aging, socioeconomic disparities in disability, long-term care, and public policy

L. Elizabeth Goldman, MD, MCR, University of California, San Francisco, is an Assistant Professor of Medicine and affiliated faculty with UCSF's Philip R. Lee Institute for Health Policy Studies. She is also an inpatient attending at Moffitt-Long Hospital and San Francisco General Hospital. Her major research interests include measuring and assessing quality of care, and the impact of financial incentives and public reporting on safety-net facilities

R. Adams Dudley, MD, MBA, University of California, San Francisco, is Professor of Medicine and Health Policy and Associate Director for Research at UCSF's Philip R. Lee Institute for Health Policy Studies. His major research interests include developing measures of quality, appropriateness, and efficiency of care, assessing the impact of strategies intended to improve performance, and financial risk adjustment

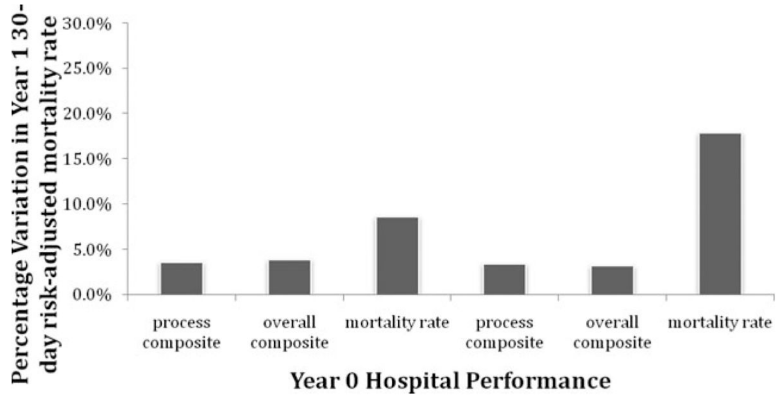

Figure 1. Percentage of Variation in Year 1 AMI and HF Mortality Explained by Year 0 Hospital Performance on Composite Scores (Process Composite and Overall Composite) and Mortality Rate*

*Percentage variation calculated with linear regression adjusting for disease-specific hospital patient volume. Overall composite is composed of process plus mortality measures.

Table 1

Hospital Characteristics in Year 0

	No. (%) of Hospitals *N* = 1,868
Bed size	
<100	169 (9.1)
100–400	1,305 (69.9)
>400	394 (21.1)
Ownership	
Private	271 (14.5)
Not for profit	1,405 (75.2)
Public	192 (10.3)
Teaching facility	601 (32.2)
Has cardiac intensive care beds	896 (47.9)

Table 2

Average Hospital Performance on Individual Process Measures, the Process and Overall Composite Measures, and 30-Day Risk-Adjusted Mortality for Acute Myocardial Infarction and Heart Failure in Year 0

	No. of Hospitals	Average Performance Rate (*SD*)
Acute myocardial infarction		
Aspirin at arrival	1,867	0.96 (0.04)
Aspirin at discharge	1,867	0.94 (0.07)
ACE-I or ARB for LV dysfunction	1,864	0.83 (0.13)
Smoking cessation advice/counseling	1,758	0.90 (0.16)
Beta blocker at discharge	1,867	0.94 (0.07)
Beta blocker at arrival	1,867	0.92 (0.07)
Process composite performance[a]	1,757	0.94 (0.05)
Overall composite performance[b]	1,733	0.92 (0.04)
30-day risk-adjusted mortality rate[c]	1,868	16.3 (1.6)
Heart failure		
Discharge instructions	1,749	0.63 (0.22)
Left ventricular function assessment	1,865	0.92 (0.08)
ACE-I or ARB for LV dysfunction	1,865	0.84 (0.10)
Smoking cessation advice/counseling	1,768	0.87 (0.15)
Process composite performance[a]	1,747	0.80 (0.11)
Overall composite performance[b]	1,733	0.82 (0.09)
30-day risk-adjusted mortality rate[c]	1,868	11.0 (1.4)

Note. ACE, angiotensin-converting enzyme; ARB, angiotensin receptor blocker; *SD*, standard deviation; LV, left ventricular.

[a]Process composite score reflects hospital performance on all included disease-specific individual process measures. Possible range 0–1.

[b]Overall composite score reflects hospital performance on all included disease-specific individual process and survival rates. Possible range 0–1.

[c]Mortality rate is expressed per 100 patients presenting with the disease.

Table 3

Average 30-Day Risk-Adjusted Mortality Rates for Hospitals Performing in Bottom vs. Top Deciles by Condition and Different Methods of Summarizing Year 0 Performance, with Estimated Number of Potentially Avoidable Deaths Per 100 Patients in Year 1 if Patients Using Bottom Decile Hospitals Could Switch to Top Decile Hospitals

Acute Myocardial Infarction				
Method of Summarizing Year 0 Performance	**Average Mortality Rate for Hospitals in the Bottom Decile of Performance**	**Average Mortality Rate for Hospitals in the Top Decile of Performance**	**Number of Deaths Potentially Avoided/100 Patients by Switching from a Bottom to a Top Decile Hospital**	***p*-Value**
Process composite performance	16.35	15.74	0.61	<.001
Overall composite performance	16.33	15.77	0.56	<.001
30-day risk-adjusted hospital mortality rate	16.55	15.24	1.31	<.001
Heart failure				
Process composite performance	11.03	11.06	−0.03	.68
Overall composite performance	11.01	11.09	−0.08	.63
30-day risk-adjusted hospital mortality rate	12.17	10.05	2.12	<.001