Relationship Between Medicare’s Hospital Compare Performance Measures and Mortality Rates

Rachel M. Werner, MD, PhD
Eric T. Bradlow, PhD

It is widely recognized that the quality of health care in the United States is uneven and often inadequate.\(^1\,2\) In the outpatient setting, quality of care varies across individuals depending on age, sex, race, and socioeconomic status.\(^3\) Overall, only half of US individuals receive recommended care.\(^3\) In hospitals, quality of care is also variable. Compliance with hospital performance measures varies not only across US hospitals but also across regions, conditions, and performance measures.\(^4\)

Because it is assumed that measuring quality of care is a key component in improving care, quality measurement is playing an increasingly prominent role in quality improvement. For example, quality is measured and in many cases reported for hospitals,\(^5\,10\) health plans,\(^11\) nursing homes,\(^12\) home health agencies,\(^13\) and physicians.\(^14\,16\) These efforts are designed to provide an incentive to improve the quality of care delivered and to influence consumer choice of providers.\(^17\,18\) These measures also are increasingly being used to determine clinicians’ reimbursement.\(^19\,21\)

Recently, the US Centers for Medicare & Medicaid Services (CMS), along with other health care organizations, began measuring hospital performance and reporting this performance on their Web site, Hospital Compare. It is unknown whether these process performance measures are related to hospital-level outcomes.

**Objective** To determine whether quality measured with the process measures used in Hospital Compare are correlated with and predictive of hospitals’ risk-adjusted mortality rates.

**Design, Setting, and Participants** Cross-sectional study of hospital care between January 1 and December 31, 2004, for acute myocardial infarction, heart failure, and pneumonia at acute care hospitals in the United States included on the Hospital Compare Web site. Ten process performance measures included in Hospital Compare were compared with hospital risk-adjusted mortality rates, which were measured using Medicare Part A claims data.

**Main Outcome Measures** Condition-specific inpatient, 30-day, and 1-year risk-adjusted mortality rates.

**Results** A total of 3657 acute care hospitals were included in the study based on their performance as reported in Hospital Compare. Across all acute myocardial infarction performance measures, the absolute reduction in risk-adjusted mortality rates between hospitals performing in the 25th percentile vs those performing in the 75th percentile was 0.005 for inpatient mortality, 0.006 for 30-day mortality, and 0.012 for 1-year mortality (P < .001 for each comparison). For the heart failure performance measures, the absolute mortality reduction was smaller, ranging from 0.001 for inpatient mortality (P = .03) to 0.002 for 1-year mortality (P = .08). For the pneumonia performance measures, the absolute reduction in mortality ranged from 0.001 for 30-day mortality (P = .05) to 0.005 for inpatient mortality (P < .001). Differences in mortality rates for hospitals performing in the 75th percentile on all measures within a condition vs those performing lower than the 25th percentile on all reported measures for acute myocardial infarction ranged between 0.008 (P = .06) and 0.018 (P = .008). For pneumonia, the effects ranged between 0.003 (P = .09) and 0.014 (P < .001); for heart failure, the effects ranged between −0.013 (P = .06) and −0.038 (P = .45).

**Conclusions** Hospital performance measures predict small differences in hospital risk-adjusted mortality rates. Efforts should be made to develop performance measures that are tightly linked to patient outcomes.

**Context** In response to concerns about the quality of care in US hospitals, the Centers for Medicare & Medicaid Services began measuring hospital performance and reporting this performance on their Web site, Hospital Compare. It is unknown whether these process performance measures are related to hospital-level outcomes.
tions, began participating in the Hospital Quality Alliance, a large-scale public-private collaboration that seeks to make performance information on all acute care nonfederal hospitals accessible to the public, payers, and providers of care. \(^{22,23}\) These performance measures evaluate hospital quality on certain processes of care for patients with acute myocardial infarction (AMI), heart failure, pneumonia, and for surgical infection prevention. The Joint Commission on Accreditation of Healthcare Organizations (JCAHO), another Hospital Quality Alliance participant, also uses these measures, requiring hospitals to report their performance as a part of accreditation for most US hospitals. \(^{24}\) This comparative quality information is now available to the public, in slightly different forms, through the CMS Web site Hospital Compare (http://www.hospitalcompare.hhs.gov/) and the JCAHO Web site Quality Check (http://www.qualitycheck.org).

One way that hospital process performance measures might lead to improvements in health care quality is by providing consumers with the information needed to choose a hospital in which they will receive higher-quality care. \(^{17,18}\) However, like most performance measures, Hospital Compare captures information in a select group of patients with a given condition. Additionally, process performance measures capture information about only a small portion of the overall care delivered during a hospital stay. While some research has documented an association between higher adherence to care guidelines and better outcomes of patients who receive that care, \(^{25,26}\) to date there has been limited evidence demonstrating that hospitals that perform better on process measures also have better overall quality for the average patient. Our objective in this study was to determine whether quality measured with the process measures used in the CMS's Hospital Compare are correlated with and predictive of hospitals' risk-adjusted mortality rates. In other words, from a consumer's perspective, can a hospital's performance on process measures be used to choose a hospital in which patients can expect to have better outcomes?

**METHODS**

**Hospital Compare**

Participation in Hospital Compare is voluntary and initially only one quarter of hospitals reported their performance data. \(^{27}\) However, the Medicare Modernization Act in 2003 introduced financial incentives for hospitals to report data on 10 performance measures to the CMS. Starting in 2004, hospitals that did not submit performance data for these measures experienced a reduction in the annual Medicare fee schedule update of 0.4 percentage points. After implementation of the Medicare Modernization Act, the proportion of hospitals reporting their performance on these measures grew to 98%. \(^{23,28}\) Seven additional measures are also included in Hospital Compare but are not tied to financial incentives and have significantly lower levels of reporting. \(^{28}\) The data from these 17 measures have been posted on the CMS's Hospital Compare Web site since April 2005 and are updated quarterly.

**Quality Measures**

**Process Performance Measures.** We evaluated hospital performance based on publicly available data from the CMS on the original 10 process measures included in Hospital Compare. These measures evaluate quality of care for AMI, heart failure, and pneumonia between January 1 and December 31, 2004. Five of the measures assess quality of care for AMI: aspirin use within 24 hours of arrival, \(\beta\)-blocker use within 24 hours of arrival, angiotensin-converting enzyme inhibitor use for left ventricular dysfunction, aspirin prescribed at discharge, and \(\beta\)-blocker prescribed at discharge. Two of the measures assess quality of care for heart failure: assessment of left ventricular function and the use of an angiotensin-converting enzyme inhibitor for left ventricular dysfunction. Three of the measures assess quality of care for pneumonia: the timing of initial antibiotics, pneumococcal vaccination, and assessment of oxygenation within 24 hours of admission.

This evaluation is limited to the original 10 measures because hospital reporting rates are nearly universal. Reporting rates are much lower for the remaining 7 measures, ranging between 18% and 83%. \(^{28}\) For these 7 measures, a hospital's decision to report its data may be nonrandom \(^{23,28}\) (eg, hospitals that perform poorly may be less likely to report their performance) and hence the significant missing data would cause additional and challenging analyses that would require the inclusion of unverifiable assumptions about the selection process. While the JCAHO requires reporting of these measures for accreditation, diminishing concerns about missing data, the data from Hospital Compare were chosen instead because these data are from all US hospitals rather than only accredited hospitals.

For each of the 10 measures, a hospital's performance is calculated as the proportion of patients who received the indicated care out of all the patients who were eligible for the indicated care. All US acute care hospitals that participated in Hospital Compare during 2004 were included. To ensure the stability of the measures, hospitals with fewer than 25 patients in the denominator of a measure were excluded. This is the same convention that the CMS uses to report performance.

Hospital performance also was measured using 2 condition-specific performance measures. First, a composite measure was calculated by aggregating individual measures within conditions using a weighted average of performance across all measures. \(^{29}\) For example, the number of times measured AMI care was received at a hospital was divided by the number of times patients were eligible for all AMI measures. This composite measure was included because it is a metric of hospital performance that is currently used by the JCAHO \(^{30}\) and may provide more meaningful information about hospi-
HOSPITAL PERFORMANCE MEASURES AND LOW MORTALITY RATES

...tal performance than individual measures do.

Second, an “all-or-none” measure was developed, which identified both hospitals that performed well on all measures within a condition set and hospitals that performed poorly on all measures within a condition set. To do this, we first identified high-performing and low-performing hospitals, or hospitals that performed above the 75th percentile on every measure they reported and hospitals that performed below the 25th percentile on every measure they reported. This all-or-none measure was included because hospitals that perform well on all measures within a condition set may be of higher quality than hospitals that perform well on average, and therefore the relationship between performance and mortality rates may be easier to detect, if it exists.

Risk-Adjusted Mortality Rates. Using the Medicare Provider Analysis and Review (MEDPAR) file, which contains all Medicare Part A claims for 2004, we calculated condition-specific hospital risk-adjusted mortality rates using the standard convention of the ratio of expected mortality rate to the observed mortality rate. For this calculation, each patient’s predicted probability of death was calculated using logistic regression and adjusted for 30 comorbidities defined by Elixhauser et al., and by age, race, ZIP-code level median income and education, sex, insurance status, and whether the admission was emergent or elective. Each hospital’s expected mortality rate was then calculated by summing the predicted probabilities of death for all of the patients divided by the total number of patients in that hospital. The risk-adjusted mortality rate was calculated by taking the ratio of the observed to expected mortality rates, standardized by the average nationwide mortality rate.

Condition-specific inpatient, 30-day, and 1-year risk-adjusted mortality rates were calculated for each hospital. For these calculations, 3 cohorts of patients were defined: all patients who had been admitted to the hospital with the principal diagnosis of AMI (International Classification of Diseases, Ninth Revision, Clinical Modification [ICD-9-CM] codes 410.0-410.9), heart failure (ICD-9-CM codes 402.01, 402.11, 402.91, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93, or 428.0-428.9), or pneumonia (ICD-9-CM codes 480.8, 480.9, 481, or 482.0-487.0). In the pneumonia cohort, patients who had a principal diagnosis of sepsis (ICD-9-CM codes 038.0-038.9) or respiratory failure (ICD-9-CM codes 518.81 or 518.84) and a diagnosis of pneumonia also were included. Finally, patients who did not have an admitting diagnosis of pneumonia were excluded. This was done to exclude patients who developed nosocomial pneumonia during their hospital stay.

Characteristics of Hospitals

The characteristics of each hospital were obtained from Medicare’s 2004 Provider of Service file. Characteristics included in the analyses included profit status, number of beds, teaching status, and whether a facility performs open heart surgery (as a measure of hospital technology). These measures were chosen because they are often used as implicit measures of hospital quality and are known to affect patient outcomes.

Statistical Analysis

For descriptive purposes, hospitals were grouped into thirds based on their average 1-year risk-adjusted mortality rate for AMI, heart failure, and pneumonia. Hospital characteristics were then aggregated within tertiles. The association between hospital characteristics and tertiles of risk-adjusted mortality rate were assessed.

A Bayesian approach was used to test the relationship between hospital performance on the 10 performance measures and composite measures with condition-specific inpatient, 30-day, and 1-year risk-adjusted mortality rates unadjusted for hospital characteristics. To do this, Bayesian “shrinkage” was applied to each hospital’s observed and expected condition-specific mortality rates. This process weights the hospital’s mortality rates based on the degree of uncertainty used to calculate those rates. Rates from hospitals with small caseloads have greater uncertainty and hence these rates are shrunken more toward the population mean. Similarly, hospitals with large caseloads have a relatively smaller amount of shrinkage and the shrunken estimate is closer to the hospital’s observed rate. Bayesian shrinkage helps account for expected regression to the mean in mortality rates. Mortality rate estimates that are unadjusted for regression to the mean are biased because the observed mortality rates will usually be farther from the population mean than the true mortality rate. We estimated the relationship between each hospital’s risk-adjusted mortality rate and performance, while controlling for other hospital characteristics, using a linear model. The relationship between each performance measure and condition-specific risk-adjusted mortality rates were modeled separately.

After estimating the relationship between hospital performance and mortality rates, we estimated the distance between Medicare beneficiaries and the closest high-performing hospital. Hospitals were classified as high-performing if they performed in the 75th percentile based on their condition-specific composite performance. Distances were calculated based on miles between a hospital’s ZIP-code centroid and a Medicare beneficiary’s ZIP-code centroid. The median distance between Medicare beneficiaries and the closest high-performing hospital and the proportion of beneficiaries living within 30 miles of a high-performing hospital were calculated. Thirty miles was chosen as a cut point because we were examining acute conditions that may benefit from minimizing the time between symptom onset and initiating treatment.

Finally, a sensitivity analysis was conducted due to concerns about the accuracy and completeness of the Hospital
Compare data. If poor-performing hospitals either submitted inaccurate data or excluded more patients from the denominator of the performance measures, this could attenuate the relationship between performance and mortality. Therefore, new data were simulated to try to reflect this possibility. To do this, it was first assumed that the 10% of hospitals with the lowest condition-specific composite performance had inflated their performance by 5%. Second, it was assumed that a more extreme case in which the 20% of hospitals with the lowest condition-specific composite performance inflated their performance by 10%. These assumed perturbations in the data were corrected for by systematically deflating the performance of affected hospitals and by rerunning the analyses.

The study protocol and waiver of informed consent was reviewed and approved by the institutional review board of the University of Pennsylvania. All Bayesian analyses were performed using WinBUGS version 1.4 (MRC, Cambridge, England). All other analyses were performed using Stata version 9.0 (StataCorp, College Station, Tex). A P value of less than .05 was considered statistically significant for all analyses.

RESULTS
Of the 4048 hospitals in the Hospital Compare database, 21 did not report performance on any of the 10 performance measures. A total of 284 hospitals were excluded because all of the measures they reported were based on fewer than 25 patients. An additional 86 hospitals included in Hospital Compare were not identified in the 2004 MEDPAR file and were dropped from the analyses. A total of 3657 hospitals were included in the final analyses.

Hospital Characteristics
The characteristics of the hospitals are summarized in TABLE 1. After grouping hospitals into tertiles based on their average 1-year risk-adjusted mortality rates by AMI, heart failure, and pneumonia, the average risk-adjusted mortality rate ranged from 0.27 to 0.40. Among the high-mortality hospitals, a smaller proportion of hospitals were large, were for profit, were teaching hospitals, or had open heart surgery capabilities (P < .001 for all comparisons). Hospital performance for each of the 10 individual performance measures and the condition-specific composite measures appear in TABLE 2.

©2006 American Medical Association. All rights reserved.
Relationship Between Hospital Performance and Risk-Adjusted Mortality Rates

After adjusting for hospital characteristics, the adjusted difference in risk-adjusted mortality rates for hospitals whose performance measures were in the 25th percentile compared with those in the 75th percentile were estimated, which we found to be a plausible range of hospital performance within individual health care markets (Table 3). For example, among hospitals performing in the 25th percentile for the measure of aspirin at admission for AMI, the risk-adjusted inpatient mortality rate at these hospitals is 0.074. Among hospitals performing in the 75th percentile on this measure, the predicted risk-adjusted inpatient mortality rate at these hospitals is 0.068. This resulted in an absolute reduction in mortality rates of 0.006.

Across all AMI performance measures, the absolute reduction in risk-adjusted mortality rates between hospitals performing in the 25th percentile vs those performing in the 75th percentile was 0.005 for inpatient mortality, 0.006 for 30-day mortality, and 0.012 for 1-year mortality. For the heart failure performance measures, the absolute mortality reduction was smaller, ranging from 0.001 for inpatient mortality to 0.002 for 1-year mortality. For the pneumonia performance measures, the absolute reduction in mortality ranged from 0.001 for 30-day mortality to 0.005 for inpatient mortality.

The adjusted difference in risk-adjusted mortality rates was also estimated using the all-or-none quality measure (Table 4). Between 8% and 14% of hospitals qualified as high-performing using this measure. Differences in mortality rates for hospitals performing above the 75th percentile

Table 3. Risk-Adjusted Mortality Rates for Hospitals Performing in the 25th and 75th Percentiles and Absolute Reduction in Risk-Adjusted Mortality Rates Between Hospitals in the 25th and 75th Percentiles for Each Measure*

<table>
<thead>
<tr>
<th></th>
<th>25th Percentile</th>
<th>75th Percentile</th>
<th>Value</th>
<th>25th Percentile</th>
<th>75th Percentile</th>
<th>Value</th>
<th>25th Percentile</th>
<th>75th Percentile</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acute myocardial infarction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At admission</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspirin</td>
<td>0.074</td>
<td>0.068</td>
<td>&lt;.001</td>
<td>0.127</td>
<td>0.121</td>
<td>.006</td>
<td>0.001</td>
<td>0.269</td>
<td>0.254</td>
</tr>
<tr>
<td>β-Blocker</td>
<td>0.074</td>
<td>0.069</td>
<td>&lt;.001</td>
<td>0.127</td>
<td>0.121</td>
<td>.006</td>
<td>&lt;.001</td>
<td>0.268</td>
<td>0.257</td>
</tr>
<tr>
<td>At discharge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspirin</td>
<td>0.075</td>
<td>0.072</td>
<td>&lt;.001</td>
<td>0.126</td>
<td>0.122</td>
<td>.004</td>
<td>.001</td>
<td>0.265</td>
<td>0.255</td>
</tr>
<tr>
<td>β-Blocker</td>
<td>0.071</td>
<td>0.067</td>
<td>&lt;.001</td>
<td>0.126</td>
<td>0.121</td>
<td>.005</td>
<td>&lt;.001</td>
<td>0.265</td>
<td>0.255</td>
</tr>
<tr>
<td><strong>ACE inhibitor for LV dysfunction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composite score</td>
<td>0.079</td>
<td>0.076</td>
<td>.03</td>
<td>0.123</td>
<td>0.119</td>
<td>.004</td>
<td>.04</td>
<td>0.258</td>
<td>0.251</td>
</tr>
<tr>
<td><strong>Heart failure</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment of LV function</td>
<td>0.029</td>
<td>0.027</td>
<td>.02</td>
<td>0.079</td>
<td>0.077</td>
<td>.002</td>
<td>.01</td>
<td>0.313</td>
<td>0.311</td>
</tr>
<tr>
<td>ACE inhibitor for LV dysfunction</td>
<td>0.031</td>
<td>0.031</td>
<td>.44</td>
<td>0.080</td>
<td>0.082</td>
<td>-.002</td>
<td>.03</td>
<td>0.316</td>
<td>0.320</td>
</tr>
<tr>
<td>Composite score</td>
<td>0.029</td>
<td>0.027</td>
<td>.03</td>
<td>0.079</td>
<td>0.078</td>
<td>.001</td>
<td>.02</td>
<td>0.313</td>
<td>0.311</td>
</tr>
<tr>
<td><strong>Pneumonia</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxygenation assessment</td>
<td>0.042</td>
<td>0.041</td>
<td>&lt;.001</td>
<td>0.098</td>
<td>0.098</td>
<td>0</td>
<td>.37</td>
<td>0.291</td>
<td>0.290</td>
</tr>
<tr>
<td>Pneumococcal vaccination</td>
<td>0.044</td>
<td>0.042</td>
<td>.05</td>
<td>0.099</td>
<td>0.098</td>
<td>.001</td>
<td>.20</td>
<td>0.292</td>
<td>0.287</td>
</tr>
<tr>
<td>Timing of initial antibiotic therapy</td>
<td>0.043</td>
<td>0.035</td>
<td>&lt;.001</td>
<td>0.098</td>
<td>0.093</td>
<td>.005</td>
<td>&lt;.001</td>
<td>0.291</td>
<td>0.281</td>
</tr>
<tr>
<td>Composite score</td>
<td>0.042</td>
<td>0.038</td>
<td>.05</td>
<td>0.098</td>
<td>0.097</td>
<td>.001</td>
<td>.05</td>
<td>0.291</td>
<td>0.288</td>
</tr>
</tbody>
</table>

Abbreviations: ACE, angiotensin-converting enzyme; ARR, absolute risk reduction; LV, left ventricular.

*Mortality rates are adjusted for hospital characteristics.

Table 4. Proportion of Hospitals Performing in the 75th Percentile or Above and Absolute Reduction in Risk-Adjusted Mortality Rates Between Hospitals Performing Above the 75th Percentile and Hospitals Performing Below the 25th Percentile on All Reported Measures

<table>
<thead>
<tr>
<th></th>
<th>Hospitals Performing in ≥75th Percentile, %</th>
<th>Hospitals Performing in the 75th Percentile or Above vs Hospitals Performing Below the 25th Percentile*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inpatient</td>
<td>30 d</td>
</tr>
<tr>
<td></td>
<td>ARR</td>
<td>P Value</td>
</tr>
<tr>
<td>AMI</td>
<td>8.5</td>
<td>.008</td>
</tr>
<tr>
<td>Heart failure</td>
<td>14.0</td>
<td>.013</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>8.1</td>
<td>.014</td>
</tr>
</tbody>
</table>

Abbreviations: AMI, acute myocardial infarction; ARR, absolute risk reduction.
*Adjusted for hospital characteristics.
for all measures within a condition vs those performing lower than the 25th percentile on all reported measures ranged between 0.008 and 0.018 for AMI and between 0.003 and 0.014 for pneumonia. Differences in heart failure–related mortality between high-performing and low-performing hospitals were not statistically significant.

**Distance to High-Performing Hospitals**

The median distance between Medicare beneficiaries and the closest high-performing hospitals ranged from 26.1 miles (41.8 km) to 29.7 miles (47.5 km) by AMI, heart failure, and pneumonia (Table 5). Distances were longer in the Midwest, South, and West compared with the Northeast, where the median distance to a high-performing hospital was less than 15 miles (24 km). Approximately half of Medicare beneficiaries live within 30 miles (48 km) of a high-performing hospital.

**Sensitivity Analysis**

To test whether our results were affected by the possibility that low-performing hospitals systematically alter their measured performance by submitting inaccurate data or excluding more patients from the denominator of the performance measures, we assumed that the lowest-performing hospitals reported performance that was better than it actually was. After simulating data to adjust for this possible perturbation and rerunning the analyses, these changes only minimally changed the magnitude of the effect we found and did not change the overall interpretation of our results (available from author on request).

**COMMENT**

Hospital performance on the Hospital Compare measures were modestly correlated with condition-specific risk-adjusted mortality rates. Although hospital performance predicted differences in risk-adjusted mortality rates that were statistically significant in some cases, these differences were small. Based on these results, the ability of performance measures to detect clinically meaningful differences in quality across hospitals is questionable.

One way to interpret the magnitude of these results is in terms of the number of lives at risk from patients receiving care at low-performing hospitals instead of high-performing hospitals. For example, approximately 750,000 patients are hospitalized annually for AMI. If one third of those patients who received care at the lowest-performing hospitals received care at high-performing hospitals instead, in which mortality rates are 1.2 percentage points lower, approximately 3000 more lives could have been saved. When put in the context of the number of individuals who have access to high-performing hospitals, the number of lives at risk becomes smaller.

While a large majority of Medicare beneficiaries live within 30 miles of a high-performing hospital in the Northeast, only one third of Medicare beneficiaries do in the South. For those with commercial insurance, the number of patients with access to these hospitals may be even further diminished due to selective contracting between hospitals and insurers. Hospital crowding may further diminish access to high-performing hospitals. Despite our findings of a statistically significant relationship between hospital performance and mortality, the clinical significance of differences in hospital performance may be small due to the small effect size and the constraints limiting patient selection of high-performing hospitals.

There are plausible reasons for the modest effect of hospital performance on risk-adjusted mortality rates. First, Hospital Compare measures discrete aspects of care delivery rather than assessing global quality. While these measures have been tightly linked to patient outcomes in clinical trials and are included in clinical practice guidelines, risk-adjusted mortality rates are likely influenced by many factors that are independent of these measures, such as whether a hospital uses electronic health records, staffing levels, activities of quality oversight committees, or other clinical strategies that may impact the clinical outcomes of a hospital. Additionally, the measures

<table>
<thead>
<tr>
<th>Table 5. Distance to the Closest High-Performing Hospital</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance to Closest High-Performing Hospital, Median (IQR), Miles</td>
</tr>
<tr>
<td>---------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Acute myocardial infarction</strong></td>
</tr>
<tr>
<td><strong>Census regions</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Heart failure</strong></td>
</tr>
<tr>
<td><strong>Census regions</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Pneumonia</strong></td>
</tr>
<tr>
<td><strong>Census regions</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: IQR, interquartile range.
Conversion factor: To convert mile to kilometer, multiply by 1.6.
HOSPITAL PERFORMANCE MEASURES AND LOW MORTALITY RATES

used in this study include only a subset of measurable activities at hospitals. It is possible that a larger set of performance measures would represent a more complete picture of performance and would be more tightly linked to mortality. Second, there is relatively little variation in some performance measures across hospitals. For example, in the case of giving aspirin at admission for AMI, hospitals performing in the 25th percentile give aspirin 91% of the time, only 7 percentage points below hospitals performing in the 75th percentile. There is strong evidence to suggest that all eligible patients should receive aspirin for AMI. However, the small amount of variation across hospitals makes it difficult to detect the effect of aspirin. Third, in the case of discharge performance measures, such as prescribing β-blockers at hospital discharge after AMI, the care of patients after hospitalization may be beyond the control of the hospital. Past research has suggested that the persistence of long-term pharmacological therapies is low, with a substantial number of patients stopping treatment within the first 6 months.

It is also possible that our finding of a modest relationship between hospital performance and mortality rates is an artifact due to possible low reliability of the data used to calculate performance. The CMS audits the data that is reported to it for performance measurement, and others have found that the accuracy of these data is high. However, there is no mechanism to monitor whether the data are complete, or whether hospitals include all eligible patients in performance measurement.

Differences in how hospitals identify patients to include in performance measurement could attenuate the relationship between performance and mortality. For example, poor-performing hospitals may try to game the system by selecting patients who receive the processes that are being measured and excluding those who do not, thereby artificially inflating their measured performance. This strategy would make it more difficult to detect a relationship between performance and mortality if it exists. However, our simulations to control for this type of gaming revealed that even if the worst 20% of hospitals were inflating their performance by 10%, the effect of hospital performance on mortality would remain modest. This suggests that our findings do not result from systematic differences in how hospitals report data for performance measurement.

Our results have important implications. Some may judge the value of performance measures as low because they predict small differences in mortality rates. While the information in Hospital Compare may be minimally informative for consumers wishing to choose high-quality hospitals, consumers may nonetheless judge the revealed differences in risk as important. They may also value the data for other reasons, such as the reassurance it provides regarding the quality of medical care. In addition, public reporting may motivate hospitals to improve the quality of care they deliver. Numerous studies have found that measuring performance and providing feedback to clinicians leads to improvement in performance. Hospitals that receive feedback on their quality also improve their performance. Whether improvements in mortality will follow improvements in performance remains to be seen. The small relationship between performance and mortality documented herein suggests that such improvements in mortality also may be small.

Our results should be interpreted in light of how they inform efforts to tie performance to financial incentives. The CMS is currently experimenting with using measured performance to determine hospital reimbursement. Such incentive programs will have to be structured to maximize their impact. This might be done by focusing incentive programs on the lowest performing hospitals rather than trying to improve quality of care among all hospitals because the differences in performance and outcomes are small between hospitals in the 25th percentile and the 75th percentile.

There are potential limitations to our study. First, there are limitations inherent to a cross-sectional study. It is impossible to assess causation of improved performance leading to improved mortality. It is possible that the demonstrated relationship between process and outcomes is confounded by patient or hospital factors. Patient and hospital characteristics were controlled for in our analyses to reduce the likelihood of this sort of bias. Additionally, sensitivity analyses were performed to limit the analyses to academic hospitals and to different regions of the country (results not reported herein). These sensitivity analyses have not changed the documented relationship between process and outcomes. Nonetheless, it is possible that some of the differences in mortality rates between hospitals are attributable to unobserved patient and hospital characteristics. In particular, the persistence of mortality differences 1 year after hospitalization may suggest that the modest effects found in this analysis are due in part to factors other than hospital performance that are unobserved in our data.

Second, our risk-adjustment models were based on administrative data in which comorbidities may be underreported. However, other research has suggested that underreporting in administrative data does not impact the accuracy of hospital quality rankings based on risk-adjusted mortality rates. Furthermore, a well-validated risk-adjustment model designed specifically for use with administrative data was used, and this risk model had high predictive power in our testing.

Third, our estimates of hospital risk-adjusted mortality rates were based only on Medicare beneficiaries rather than all payers. However, Medicare beneficiaries make up more than 50% of hospital admissions for the conditions studied herein and health care dollars spent on Medicare beneficiaries significantly exceeds that of other payers.
addition, hospital quality research based on Medicare data may be generalizable to the broader US population.61

Fourth, mortality as an outcome was assessed. It is possible that hospital performance is an important predictor of other outcomes, such as length of stay, failure to rescue, or readmission rates. The CMS has initiated performance measurement not only in hospitals but also in nursing homes,13 and home health agencies,13 and efforts are under way to extend measurement to physicians.12 The use of performance measurement will be expanded in 2007 with the Deficit Reduction Act and efforts are under way to link performance to payment.19,62 These efforts have been made under the assumption that improvements in processes of care will lead to improvements in patient outcomes.63 Our study suggests that in the case of hospital performance, the CMS’s current set of performance measures are not tightly linked to patient outcomes. These findings should not undermine current efforts to improve health care quality through performance measurement and reporting. However, attention should be focused on finding measures of health care quality that are more tightly linked to patient outcomes. Only then will performance measurement live up to expectations for improving health care quality.

Author Contributions: Dr Werner had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Werner, Bradlow.

Acquisition of data: Werner.

Analysis and interpretation of data: Werner, Bradlow. Drafting of the manuscript: Werner.

Critical revision of the manuscript for important intellectual content: Werner, Bradlow.

Statistical analysis: Werner, Bradlow.

Obtained funding: Werner.

Administrative, technical, or material support: Werner.

Study supervision: Werner, Bradlow.

Financial Disclosures: None reported.

Funding/Support: Dr Werner was supported by a career development award from the Department of Veterans Affairs.

Role of the Sponsor: The Department of Veterans Affairs had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

Acknowledgment: We thank David Asch, MD, MBA (Center for Health Equity Research and Promotion, Philadelphia Veterans Affairs Medical Center Division of General Internal Medicine, University of Pennsylvania School of Medicine, and Leonard Davis Institute of Health Economics), for comments on drafts of this article. Dr Asch was not compensated for his review.

REFERENCES


First I write one sentence: then I write another. That’s how I write. But I have a feeling writing ought to be like running through a field.
—Lytton Strachey (1880-1932)
sicians and patients do not follow standard practice guidelines. For example, a recent survey of gastroenterologists found that 50% reported routinely recommending colonoscopy at intervals more frequent than those suggested by published guidelines. Above and beyond the particular patient's preference, we do not know if there are other reasons for not following standard guidelines in this example. But we share Dr. Sack's concern. An important goal of the Clinical Crossroads series is to encourage readers to consider evidence-based guidelines and the rationale for following them when making clinical decisions in their own practices.

Risa B. Burns, MD
Tom Delbanco, MD
tdelbanc@bidmc.harvard.edu
Beth Israel Deaconess Medical Center
Boston, Mass

Financial Disclosures: None reported.


Table 3. Risk-Adjusted Mortality Rates for Hospitals Performing in the 25th and 75th Percentiles and Absolute Reduction in Risk-Adjusted Mortality Rates Between Hospitals in the 25th and 75th Percentiles for Each Measure

<table>
<thead>
<tr>
<th></th>
<th>25th Percentile</th>
<th>75th Percentile</th>
<th>Absolute Reduction</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acute MI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At admission</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspirin</td>
<td>0.074</td>
<td>0.068</td>
<td>0.006</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>β-Blocker</td>
<td>0.074</td>
<td>0.068</td>
<td>0.005</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>At discharge</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspirin</td>
<td>0.075</td>
<td>0.072</td>
<td>0.003</td>
<td>.02</td>
</tr>
<tr>
<td>β-Blocker</td>
<td>0.071</td>
<td>0.067</td>
<td>0.004</td>
<td>.01</td>
</tr>
<tr>
<td>ACE inhibitor for LV dysfunction</td>
<td>0.079</td>
<td>0.076</td>
<td>0.003</td>
<td>.05</td>
</tr>
<tr>
<td>Composite score</td>
<td>0.074</td>
<td>0.069</td>
<td>0.003</td>
<td>.03</td>
</tr>
<tr>
<td><strong>Heart failure</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment of LV function</td>
<td>0.029</td>
<td>0.027</td>
<td>0.001</td>
<td>.05</td>
</tr>
<tr>
<td>ACE inhibitor for LV dysfunction</td>
<td>0.031</td>
<td>0.031</td>
<td>0.001</td>
<td>.40</td>
</tr>
<tr>
<td>Composite score</td>
<td>0.029</td>
<td>0.027</td>
<td>0.001</td>
<td>.08</td>
</tr>
<tr>
<td><strong>Pneumonia</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxygenation assessment</td>
<td>0.042</td>
<td>0.041</td>
<td>0.001</td>
<td>.23</td>
</tr>
<tr>
<td>Pneumococcal vaccination</td>
<td>0.044</td>
<td>0.042</td>
<td>0.002</td>
<td>.15</td>
</tr>
<tr>
<td>Timing of initial antibiotic therapy</td>
<td>0.043</td>
<td>0.034</td>
<td>0.008</td>
<td>.03</td>
</tr>
<tr>
<td>Composite score</td>
<td>0.042</td>
<td>0.038</td>
<td>0.005</td>
<td>.01</td>
</tr>
</tbody>
</table>

Abbreviations: ACE, angiotensin-converting enzyme; ARR, absolute risk reduction; LV, left ventricular.

*Mortality rates are adjusted for hospital characteristics.

Recalculation of Data: In the Original Contribution entitled "Relationship Between Medicare's Hospital Compare Performance Measures and Mortality Rates" published in the December 13, 2006, issue of JAMA (2006;296:2694-2702), a normal distribution of the expected mortality rate was assumed during the creation of the statistical model that estimated the relationship between the expected mortality rate and performance and a linear model was used; however, the expected mortality rate should have been transformed using a logit transformation, which makes the transformed expected mortality rate normally distributed and does not violate the assumptions of using a linear model. In the abstract on page 2694, the second, third, and fourth sentences under Results should be: "Across all acute myocardial infarction performance measures, the absolute reduction in risk-adjusted mortality rates between hospitals performing in the 25th percentile vs those performing in the 75th percentile was 0.005 for inpatient mortality (P=.003), 0.007 for 30-day mortality (P=.002), and 0.012 for 1-year mortality (P<.001). For the heart failure performance measures, the absolute mortality reduction was smaller, ranging from 0.001 for inpatient mortality (P=.08) to 0.002 for 1-year mortality (P=.23). For the pneumonia performance measures, the absolute reduction in mortality ranged from 0.003 for 30-day mortality (P=.15) to 0.007 for 1-year mortality (P=.02)." On page 2696, second column, first paragraph, the second to last sentence should be: "We estimated the relationship between each hospital's risk-adjusted mortality rate and performance, while controlling for other hospital characteristics." On page 2698, second column, first full paragraph, the first and last sentences should be: "Across all AMI performance measures, the absolute reduction in risk-adjusted mortality rates between hospitals performing in the 25th percentile vs those performing in the 75th percentile was 0.005 for inpatient mortality, 0.007 for 30-day mortality, and 0.012 for 1-year mortality. For the pneumonia performance measures, the absolute reduction in mortality ranged from 0.003 for 30-day mortality to 0.007 for 1-year mortality." Also on that page, Table 3 should appear as shown and should replace the previously published table.