Alternative Pay-for-Performance Scoring Methods
Implications for Quality Improvement and Patient Outcomes

Seth W. Glickman, MD, MBA,† William Boulding, PhD,‡ Jason M. T. Roos, BA,‡ Richard Staelin, PhD,‡ Eric D. Peterson, MD, MPH,§¶ and Kevin A. Schulman, MD*‡¶*

Downloaded from http://journals.lww.com/lww-medicalcare by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdgGj2MwlZLeI= on 11/18/2021

Background: Pay-for-performance programs typically rate hospitals using a composite summary score in which process measures are weighted by the total number of treatment opportunities. Alternative methods that weight process measures according to how hospitals organize care and the range for possible improvement may be more closely related to patient outcomes.

Objectives: To develop a hospital-level summary process measure adherence score that reflects how hospitals organize cardiac care and the range for possible improvement; and to compare associations of hospital adherence to this score and adherence to a composite score based on the Centers for Medicare and Medicaid Services scoring system with inpatient mortality.

Research Design and Subjects: Hospital-level analysis of 7 process measures for acute myocardial infarction (AMI) and 4 process measures for heart failure at 4226 hospitals, and inpatient mortality after AMI at 1351 hospitals in the United States. Data are from the Hospital Compare and Joint Commission Core Measures databases for October 2004 through September 2006.

Measures: Associations between composite scores based on Centers for Medicare and Medicaid Services methodology and alternative adherence scores with inpatient survival after AMI.

Results: In principal components analysis, hospital cardiac care varied between hospitals largely along the lines of "clinical" (ie, pharmacologic interventions) and "administrative" (ie, patient instructions or counseling) activities. A scoring system reflecting this organization was strongly associated with inpatient survival and fit the mortality data better than the composite score. Higher administrative activities scores, holding the clinical activities score fixed, were associated with lower survival.

Conclusions: In-hospital cardiac care is organized by clinical and administrative processes of care. Pay-for-performance schemes that incentivize hospitals to focus on administrative process measures may be associated with decreased adherence to clinical processes. A pay-for-performance scheme that acknowledges these factors may be associated with improved inpatient mortality.

Key Words: cardiology, guideline adherence, health policy, hospitals, outcome and process assessment (health care), quality assurance, health care, quality indicators, United States

(*Med Care* 2009;47: 1062–1068)

Pay for performance is gaining momentum as a mechanism for providing incentives for hospitals to improve the quality of clinical care.[1–4] A common incentive scheme and one used by the Centers for Medicare and Medicaid Services (CMS) is to base pay-for-performance payments on a single summary measure compiled from a number of performance indicators.[5,6] An important assumption underlying this combination is that the composite score provides a valid measure of the quality of care provided by the hospital and is strongly associated with patient outcomes.

This article explores an alternative method for developing such a summary measure. Our basic assumption is that a goal of hospital administrators, like designers of pay-for-performance schemes, is to improve the quality of patient care. Hospital administrators seek to achieve this goal in operating environments with resource and budget constraints. As a consequence, it is reasonable to assume that they have developed a set of cost-effective procedures that they believe are necessary to achieve good patient outcomes.

In our approach, we first identify how hospitals currently manage processes associated with cardiac care by examining the interrelationships between a number of different process measures for acute myocardial infarction (AMI) and heart failure. This examination allows us to infer the underlying dimensions of how hospitals operate. Then, having identified these underlying dimensions, we examine the implications of behaving according to these practices as opposed to following standards dictated by the current CMS pay-for-performance scoring method. We then ask the following questions: Are these 2 sets of standards in alignment? Is successful adherence to these 2 different standards associ-

From the *Center for Clinical and Genetic Economics, Duke Clinical Research Institute, Duke University School of Medicine, Durham, North Carolina; †Department of Emergency Medicine, University of North Carolina School of Medicine, Chapel Hill, North Carolina; ‡The Fuqua School of Business, Duke University, Durham, North Carolina; §Duke Clinical Research Institute, Duke University School of Medicine, Durham, North Carolina; and ¶Department of Medicine, Duke University School of Medicine, Durham, North Carolina.

Supported in part by a gift from the Douglas and Stefanie Kahn Charitable Gift Fund (to S.W.G.).

Reprints: Kevin A. Schulman, MD, Center for Clinical and Genetic Economics, Duke Clinical Research Institute, Duke University School of Medicine, PO Box 17969, Durham, NC 27715. E-mail: kevin.schulman@duke.edu.

Copyright © 2009 by Lippincott Williams & Wilkins

ISSN: 0025-7079/09/4710-1062

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

ated with different patient outcomes? Do certain types of process measures have stronger associations with patient outcomes than others? What are the potential implications of any differences for the design of future pay-for-performance scoring systems?

METHODS

Data Sources

We obtained data on 8 process measures for AMI and 4 process measures for heart failure and the total number of eligible and adhering cases for each process measure from the US Department of Health and Human Services Hospital Compare database for 4226 acute care hospitals for the reporting period October 2004 through September 2005.[7] We found that hospitals generally had a single reperfusion strategy for patients with ST-segment elevation myocardial infarction (STEMI), either thrombolytic therapy or percutaneous coronary intervention. Therefore, we used both variables to form a single measure that reflects the proportion of patients who received timely reperfusion therapy for STEMI, leaving us with 7 unique process measures for AMI. We obtained hospital-level structural variables, including academic affiliation, geographic location, population density, bed size, and presence of advanced specialty services (ie, percutaneous coronary intervention and cardiac surgery) from the database of the American Hospital Association (AHA).

We obtained our outcome measure, risk-adjusted inpatient mortality after AMI, from the Joint Commission's Core Measures database.[8] Use of this outcome measure relies on our assumption that risk-adjusted mortality is a good proxy for the overall quality of cardiac care.

All analyses used the largest set of hospitals possible given the available data. Our analysis of the interrelationships of process measures is based on the 4226 hospitals from the Hospital Compare database. The hospitals had to have at least 1 eligible patient in each of the process measure categories (ie, AMI and heart failure) and a minimum of 25 total treatment opportunities across all measures. Most hospitals greatly exceeded these minimum requirements, averaging 92 treatment opportunities per measure. We merged these data with records from the AHA and Joint Commission data sets using each hospital's Medicare provider number. We then selected all hospitals from our sample of 4226 hospitals that had an outcome measure in the Joint Commission data set and a valid AHA record. This approach yielded 1351 hospitals. We used this subset for the regression analyses. The institutional review board of the Duke University Health System determined that the study was exempt from formal review.

Adherence Scores

We calculated 2 sets of hospital-level process adherence scores. For the first set of scores, we used the same composite scoring method used by CMS to calculate hospital-level adherence scores in the Hospital Quality Incentive Demonstration pay-for-performance program.[9] That is, we calculated scores for both AMI and heart failure by summing the number of times each therapy was administered and dividing this amount by the sum of total eligible opportunities for all patients at the hospital.

For the second set of scores, we began with a statistical analysis of the processes associated with cardiac care. We examined the interrelationships between the 11 cardiac care performance measures and identified highly related subsets of these measures (defined so that if a hospital scored high on 1 measure, the hospital also tended to score high on the other measures in the subset). Using principal components analysis, we found that a 2-component solution adequately represented the covariance of the data. Then, using 2-component factor analysis, we partitioned the 11 measures into 2 groups based on their loadings and clinical judgment. Group 1 contained measures for which the loadings on factor 1 were greater than on factor 2, whereas group 2 contained measures for which the reverse was true.

We then calculated each hospital's adherence score for each group of measures. We started by weighting the adherence on a particular measure within the group by an amount proportional to its structural factor loading. We then adjusted these weights by multiplying them by the adherence measure's population SD, to reflect the possible range for improvement, and normalizing them within each group so they summed to 1. Weight w for adherence measure m in group g is

$$w_{mg} = \frac{(SFL_{mg})(SD_m)}{\sum_{\mu=1}^{11} (SFL_{\mu g})(SD_\mu)},$$

where *SFL* denotes the structural factor loading and SD the population SD. Finally, we multiplied these weights by the hospital's adherence score and summed over all the measures in that group.

Although the CMS composite method and the alternative method produce scores for each hospital, the methods differ in 2 important ways. First, the CMS composite method organizes process metrics into scores according to therapeutic area (AMI and heart failure), whereas the alternative method organizes process metrics according to the underlying dimensions of how hospitals operate. Second, the CMS composite method weights performance measures based on the number of available opportunities, allowing process measures with the greatest total number of cases to influence the adherence scores most. In contrast, the alternative method weights process measures based on their correlation with what the composite score is purported to measure. Furthermore, the alternative method weights process measures according to the range for possible improvement; measures with greater variance in adherence are weighted more heavily.

It is important to note that no information on survival was explicitly used to establish the individual weights for either of the scoring methods. The weights in the alternative method are based on how processes are related (and implicitly on how hospitals organize to deliver these processes), whereas the CMS weights used in this study reflect the aggregate use of each process measure. CMS uses inpatient

© 2009 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

mortality directly in evaluation of hospital performance (10% overall weighting); however, we removed this measure from the composite scoring system in this study to permit its use as the dependent variable in our analyses.

Score Comparison

The overarching objective of both the CMS composite scoring scheme and our alternative scoring scheme is to capture the quality of hospital care. We used multivariable logistic regression to compare the association of each scoring system with the hospital's inpatient AMI survival rate in the subsequent year. (We replicated these analyses using 6-month rates and contemporaneous rates. We also used weighted least squares in which the dependent variable was the proportion of surviving AMI patients. In each case, we obtained similar results, though the association between hospital score and outcome was strongest with a 1-year lag.) We performed 2 separate sets of analyses. In the first set, the independent variables were adherence scores for AMI and heart failure calculated using the CMS composite scoring method. In the second set, the independent variables were adherence scores calculated for each of the 2 factors identified using the alternative method. In each model, the dependent variable was inpatient survival (1–mortality). The unit of analysis was the hospital, and hospitals with more outcome opportunities were weighted more heavily. We included the previously mentioned hospital-level covariates in the regression models to control for other hospital-level influences on survival. We compared the overall fit of the regression models for each scoring system using χ^2 tests, the Akaike information criterion (AIC), and the weighted least squares adjusted R^2 statistics.

We compared the association between changes in the adherence scores and inpatient AMI survival in 2 ways. First, we looked at the odds ratio (OR), assuming a 1 SD change (calculated around the mean) in the particular adherence score. Second, we calculated the predicted change in mortality rate for each of the bottom 3 quarters of the hospitals in our sample, assuming an adherence score was improved by 1 quartile. To assess the stability of the results when comparing differences in predicted mortality between adherence scores, we examined differences in predicted mortality in 5000 bootstrap samples. We used R version 2.5.0 (R Foundation for Statistical Computing, Vienna, Austria) for all analyses; $P < 0.05$ was considered statistically significant.

RESULTS

Table 1 shows the baseline characteristics of the hospitals, median scores on the performance measures, and patient outcomes. In principal components analysis of the 11 process measures, we found that most of the variation in adherence to processes of care could be explained through a 2-factor solution. We found a strong elbow at 2 factors, and the third eigenvalue was 1.017. The correlation of each of these process measures with each of the 2 factors is given in Table 2 and reveals that process adherence varies largely along the lines of clinical aspects of cardiac care (ie, pharmacologic and other interventions implemented by physicians) and administrative aspects of cardiac care (ie, coun-

TABLE 1. Hospital Characteristics, Process Measures, and AMI Mortality

Variable	All Hospitals
Hospital characteristics (N = 1351)	
Hospital bed size, median (IQR)	248 (159–370)
Facilities, n (%)	
Academic or affiliated facility	588 (43.5)
Cardiac intensive care	853 (63.1)
Diagnostic cardiac catheterization	777 (57.5)
Interventional cardiac catheterization	932 (69.0)
Geographic region, n (%)*	
East North Central	205 (15.2)
East South Central	95 (7.0)
Guam and Puerto Rico	2 (0.2)
Mid-Atlantic	171 (12.7)
Mountain	81 (6.0)
New England	87 (6.4)
Pacific	194 (14.4)
South Atlantic	259 (19.2)
West North Central	89 (6.6)
West South Central	168 (12.4)
Population density, n (%)	
Nonmetropolitan area	208 (15.4)
<100,000 population	25 (1.9)
100,000–250,000 population	165 (12.2)
250,000–500,000 population	164 (12.1)
500,000–1 million population	183 (14.6)
1 million–2.5 million population	319 (23.6)
>2.5 million population	287 (21.2)
Process measures and outcomes, median (IQR)	
AMI mortality	0.07 (0.05–0.11)
Acute myocardial infarction	
Aspirin at arrival	0.96 (0.93–0.98)
Aspirin at discharge	0.96 (0.91–0.98)
β-blocker at arrival	0.93 (0.88–0.97)
β-blocker at discharge	0.95 (0.91–0.98)
Fibrinolysis <30 min for ST-segment elevation myocardial infarction	0.27 (0.00–0.50)
Percutaneous coronary intervention <90 min for ST-segment elevation myocardial infarction	0.67 (0.52–0.79)
ACE inhibitor or ARB for left ventricular systolic dysfunction	0.85 (0.75–0.93)
Smoking cessation counseling	0.92 (0.80–0.98)
Heart failure	
Angiotensin-converting enzyme inhibitor or angiotensin receptor blocker for left ventricular systolic dysfunction	0.82 (0.76–0.90)
Smoking cessation counseling	0.85 (0.69–0.95)
Discharge instructions	0.58 (0.39–0.76)
Assessment of left ventricular function	0.92 (0.87–0.96)

*East North Central includes Illinois, Indiana, Michigan, Ohio, and Wisconsin. East South Central includes Alabama, Kentucky, and Mississippi. Mid-Atlantic includes New Jersey, New York, and Pennsylvania. Mountain includes Arizona, Colorado, Idaho, Montana, Nebraska, New Mexico, Utah, and Wyoming. New England includes Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont. Pacific includes Alaska, California, Hawaii, Oregon, and Washington. South Atlantic includes Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, and Virginia. West North Central includes Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, and South Dakota. West South Central includes Arkansas, Louisiana, Oklahoma, and Texas.

IQR indicates interquartile range.

© 2009 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

seling and patient instructions likely implemented by hospital staff involved in discharge planning). We interpret these 2 underlying factors as reflecting how the delivery of cardiac care is naturally organized in the 4226 hospitals we studied.

TABLE 2. Categories and Structural Factor Loadings for Performance Measurement

	Structural Factor Loading*	
Process Measure	**Clinical Cardiac Activities**	**Administrative Cardiac Activities**
Acute myocardial infarction		
Aspirin at arrival	0.72	0.15
Aspirin at discharge	0.82	0.13
β-blocker at arrival	0.77	0.18
β-blocker at discharge	0.88	0.21
Timely reperfusion therapy for STEMI	0.36	0.19
ACE inhibitor or ARB for left ventricular systolic dysfunction	0.48	0.23
Smoking cessation counseling or advice	0.27	0.65
Heart failure		
ACE inhibitor or ARB for left ventricular systolic dysfunction	0.54	0.29
Smoking cessation counseling or advice	0.11	0.97
Discharge instructions	0.23	0.68
Assessment of left ventricular function	0.58	0.39

*Composite clinical cardiac activities and administrative cardiac activities scores were calculated by weighting the adherence on a particular measure within the group by an amount proportional to its factor loading and then summing over all the measures in that group. Structural factor loadings are Pearson correlations between adherence measures and OLS factor scores.

We hereafter refer to the first factor as "clinical cardiac activities" (CCAs) and to the second factor as "administrative cardiac activities" (ACAs). This categorization contrasts with the CMS composite scoring system, which organizes process metrics into scores according to therapeutic area.

Table 3 provides a comparison of the relative contributions of each process measure to the overall composite scores for the 2 scoring systems. These weights reflect the fact that, as shown in Table 1, the range of possible scores differs substantially across the different measures (eg, the interquartile range for aspirin on arrival was 0.93–0.98, whereas the range was 0.75–0.93 for prescription of an angiotensin-converting enzyme [ACE] inhibitor or angiotensin receptor blocker [ARB] for left ventricular systolic dysfunction [LVSD]). Thus, although the factor weights shown in Table 2 are substantially greater for aspirin on arrival than for ACE inhibitor or ARB for LVSD, the effective weights after adjustment for each measure's variation are almost equal because the variation in observed performance was much greater for the latter measures.

As can be seen by comparing these different sets of weights, the 2 scoring systems emphasize different performance measures and thus different aspects of cardiac care. The CMS adherence score for AMI places more weight on aspirin and β-blocker at arrival and discharge (which have compliance rates exceeding 93%) and smoking cessation counseling, whereas the CCA measure places more weight on early reperfusion for STEMI and prescription of an ACE inhibitor or ARB for LVSD.

To assess which scoring system better captures the overall quality of inpatient care, we next compared which scoring system was more associated with subsequent in-hospital AMI mortality. We did this by regressing the hospital's inpatient AMI survival rate in the subsequent year on the current year's adherence scores derived from the 2 scoring methods and our set of control variables (performing 3

TABLE 3. Relative Contribution of Each Performance Measure*

	Alternative Method		Composite Method	
Process Measure	**Clinical Cardiac Activities**	**Administrative Cardiac Activities**	**Acute Myocardial Infarction**	**Heart Failure**
Acute myocardial infarction				
Aspirin at arrival	10%	—	24%	—
Aspirin at discharge	14%	—	21%	—
β-blocker at arrival	15%	—	20%	—
β-blocker at discharge	17%	—	21%	—
Timely reperfusion therapy for STEMI	9%	—	4%	—
ACE inhibitor or ARB for left ventricular systolic dysfunction	11%	—	4%	—
Smoking cessation counseling or advice	—	27%	7%	—
Heart failure				
ACE inhibitor or ARB for left ventricular systolic dysfunction	10%	—	—	12%
Smoking cessation counseling or advice	—	42%	—	7%
Discharge instructions	—	31%	—	35%
Assessment of left ventricular function	13%	—	—	47%

*Reported weights may not sum to 100% due to rounding.

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

TABLE 4. Odds of Inpatient AMI Survival According to Hospital Quality Scores Generated by the CMS Composite Scoring Method Versus an Alternative Method*

	Alternative Method			Composite Method		
	Model 1a (CCA Only)	Model 1b (ACA Only)	Model 1c (CCA and ACA)	Model 2a (AMI Only)	Model 2b (HF Only)	Model 2c (AMI and HF)
Hospital scores						
Clinical cardiac activities	1.11 (1.09–1.13)†	—	1.13 (1.11–1.16)†	—	—	—
Administrative cardiac activities	—	1.01 (0.99–1.03)†	0.96 (0.94–0.98)†	—	—	—
Acute myocardial infarction	—	—	—	1.08 (1.06–1.10)†	—	1.09 (1.07–1.12)†
Heart failure	—	—	—	—	1.02 (1.01–1.04)‡	0.98 (0.96–1.00)§
Akaike information criterion	8568	8688	8550	8623	8682	8620
Pearson χ^2	3161	3817	3105	3412	3792	3380
Adjusted R^2	0.31	0.28	0.31	0.29	0.27	0.29

*Values are expressed as odds ratio (95% confidence interval). Odds ratios indicate the likelihood of survival in hospitals with a 1 SD increase in adherence score compared to a hospital with no increase in the adherence score.
†$P < 0.001$.
‡$P < 0.01$.
§$P < 0.05$.

regressions for each scoring system). Table 4 shows the OR for each of the regressions, where the ORs are based on a 1 SD improvement in each adherence score. Higher CCA hospital scores (model 1a) were associated with a greater likelihood of inpatient survival (OR: 1.11; 95% confidence interval [CI]: 1.09–1.13); however, there was no significant association between ACA hospital scores (model 1b) and survival (OR: 1.01; 95% CI: 0.99–1.03). Higher CMS composite adherence scores for AMI (OR: 1.08; 95% CI: 1.06–1.10) and heart failure (OR: 1.02; 95% CI: 1.01–1.04) were associated with a greater likelihood of inpatient survival (models 2a and 2b). Of note, higher CCA scores were associated with greater improvements in patient outcomes than higher AMI scores (OR: 1.11 vs. 1.08; *t* statistic = 2.40; $P < 0.01$).[10] The model that included the CCA scores fit the mortality data better than the model using the CMS AMI scores (Pearson $\chi^2 = 3161$ vs. 3412; AIC = 8568 vs. 8623; $R^2 = 0.31$ vs. 0.29).

When CCA and ACA scores were included in the same regression model (model 1c), the CCA score had an even stronger association with survival (OR: 1.13; 95% CI: 1.11–1.16), whereas a higher ACA score had an inverse relationship with survival (OR: 0.96; 95% CI: 0.94–0.98). This inverse relationship was also observed for heart failure when we included the CMS composite AMI score (OR: 1.09; 95% CI: 1.07–1.12) and the CMS composite heart failure score (OR: 0.98; 95% CI: 0.96–1.00) in the same regression model (model 2c). Again, the model including both the CCA and ACA scores fit the mortality data better than the model using the CMS composite scores for AMI and heart failure (Pearson $\chi^2 = 3105$ vs. 3380; AIC = 8550 vs. 8620; $R^2 = 0.31$ vs. 0.29).

To ascertain the robustness of the results (given that we included heart failure process measures but used inpatient AMI mortality as the primary outcome measure), we partitioned the 4 performance measures that made up the heart failure adherence score into 2 new adherence scores, still using the CMS method of combining the individual performance measures. The first adherence score was made up of the 2 heart failure measures included in the CCA adherence score. The second adherence score included the 2 heart failure performance measures found in the ACA adherence score. We then regressed survival against the 2 new adherence scores and found that the coefficients for AMI (OR: 1.08; 95% CI: 1.05–1.10) and the new heart failure CCA measure (OR: 1.04; 95% CI: 1.01–1.06) were both positive, whereas the coefficient for the new heart failure ACA measure (OR: 0.95; 95% CI: 0.94–0.97) was negative.

Finally, using the logistic regression coefficients, we compared the average expected difference in mortality rate associated with (*a*) a CCA score that was 1 quartile higher, holding fixed the ACA, and (*b*) an AMI score that was 1 quartile higher, holding fixed the heart failure score. In the former case, the improvement in mortality was 8.7% (95% CI: 7.2–10.0); in the latter case, it was 6.0% (95% CI: 4.5–7.7). The improvement in mortality was greater with the CCA score compared with the AMI score in 98.9% of the 5000 bootstrap samples.

DISCUSSION

We based our analysis on the premise that (*a*) hospitals organize to deliver effective, high-quality patient care and (*b*) there is interhospital variation in their ability to do so. Using the relationship among the 11 AMI and heart failure process measures, it seems that hospitals generally organize the delivery of cardiac care by clinical (ie, pharmacologic and procedural interventions) and administrative (ie, patient instructions and counseling) functions, rather than strictly according to therapeutic area (ie, AMI or heart failure processes of care). Perhaps more importantly, we found that patients treated at hospitals that better perform these clinical functions tend to also have higher inpatient survival after AMI. Specifically, we found that a scoring system based on this organizational structure had a stronger association with future inpatient survival than the composite scoring method cur-

 © 2009 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

rently used by CMS in its pay-for-performance program. This finding is not based on establishing a link between a particular performance measure for a patient and a given outcome for that patient but instead on establishing a link between the different quality-of-care measures for a hospital and subsequent outcomes for patients treated at that hospital.

Our analysis of the change in the expected mortality rates associated with CCA and AMI adherence scores that are higher by 1 quartile indicates that the alternative scoring system was associated with greater relative improvements in outcomes than the CMS composite scoring system. This insight may help in the design of incentives that are more closely aligned with processes that drive the quality of cardiac care and thus are better predictors of subsequent patient outcomes, potentially leading to improvements in outcomes beyond those achieved with current systems.

We also found that greater proficiency in certain "administrative" processes of care found in the AMI and heart failure composite scores (such as discharge instructions and smoking cessation counseling) was associated with higher rates of in-hospital mortality for patients with AMI once we accounted for the CCA score. We interpret this finding in the context of resource constraints. Hospitals are under increasing financial stress, and financial commitment to quality processes has important implications for quality of care.[11,12] There are many ways hospitals can administer care to patients with AMI. However, if resources are scarce, it appears that, at the margin, resources expended on administrative activities associated with cardiac care may result in fewer resources that could be devoted to other (ie, CCA) activities. Our findings also suggest that hospitals that allocate their incremental efforts to better adhere to activities associated with the CCA measure are more likely to obtain better patient outcomes (at least as measured by inpatient AMI survival rates), compared with hospitals that allocate these scarce resources to ACA activities. The findings highlight the importance of the trade-offs that must be considered by hospital administrators in managing quality improvement.

At a macro level, policy makers and payers also can affect the quality of care through the selection of performance measures. Our findings underscore the need for the development and selection of performance measures that have a robust evidence base supporting their use. Previous studies have found that there is no association between some processes of care currently used in federal pay-for-performance programs and patient outcomes. For example, Fonarow et al[13] found that the current heart failure performance measures, aside from prescription of ACE inhibitors or ARBs at discharge, have little relation to mortality and combined mortality/rehospitalization in the first 60 to 90 days after discharge, while a recent study showed that smoking cessation counseling at discharge did not correlate with successful smoking cessation at 6 months and 1 year.[14]

The CMS scoring method weights performance measures based on the number of available opportunities, allowing process measures with the greatest total number of cases to influence the composite score the most. This is particularly relevant to the discharge instructions for heart failure and AMI, where few patients are excluded from eligibility for these measures. In contrast, patients with STEMI make up only a subset of all patients with AMI and are given much less weight in a composite score than other measures, despite the robust association between early STEMI care and patient outcomes and wide variability in hospital performance nationally.[15,16] In addition, we found that a scoring system that weighted process measures with greater variance more heavily (eg, early reperfusion for STEMI vs. aspirin at arrival) was associated with greater improvements in inpatient mortality. Concerns have been raised about the cost-effectiveness of quality improvement interventions that target measures with "ceiling effects."[17] Accordingly, future pay-for-performance programs should consider using differential weighting for process measures according to the magnitude of association between the measure and the underlying construct being measured and the range of possible improvements.

At the time of the study data (2004 through 2005), many hospitals may have recognized the underlying importance of these individual process measures in their performance. For example, the median adherence rate for heart failure discharge instructions was 58%, whereas the median adherence rates for β-blockers at arrival and departure were 94% and 93% for patients with AMI, respectively. During the past several years, hospital performance for administrative measures such as smoking cessation counseling have increased dramatically, exceeding 95% mean adherence in 2006.[1] Although this has been welcomed by many as a positive development, this trend also raises some potential concerns. If certain measures are not associated with patient outcomes, then quality-improvement interventions targeting these measures may not be cost-effective. Concerns have also been raised regarding the potential adverse consequences of financial incentives, including undue provider attention to certain aspects of care to the detriment of others.[18,19] The results of this study suggest a possible trade-off between certain processes of care.

Our finding that cardiac care is organized largely along clinical and administrative aspects of care is consistent with current knowledge of hospital organizational structure.[20,21] Many total quality management projects, such as those that have likely developed in response to large federal pay-for-performance programs, generally originate and are implemented at the level of hospital management.[22] However, clinical care is delivered at a service-line level by teams of health care professionals. Many pay-for-performance scoring systems are organized strictly by therapeutic area, each of which incorporates clinical and administrative aspects of care. Unless management and hospital systems evolve to a more decentralized structure in which clinical care and quality improvement efforts are administered at the level of the service line, external performance measurement systems that are aligned with the natural organization of care delivery in hospitals should facilitate their implementation and help to improve patient outcomes.

Limitations

Our study has several potential limitations. First, the study was cross-sectional, which limited our ability to infer

© 2009 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.

causality between the quality-of-care scores and future in-hospital survival. However, the results were consistent across a wide range of baseline hospital scores and lag structures, and the hospital process measures temporally preceded the hospital outcome measure. Second, the study was limited to the care of patients with cardiac disease, and additional studies are needed to assess other therapeutic areas. Third, although the study tested process measures for both heart failure and AMI, the outcome measure was inpatient mortality secondary to AMI. Heart failure outcome data were not readily available to the research team. However, AMI deaths comprise most of inpatient deaths resulting from cardiac admissions and, in aggregate, should be a relatively reliable measure of overall hospital cardiac outcomes.[23] Thus, the inclusion of 1 or more heart failure outcome measures should not change the results or add much additional information. This conclusion is supported by the finding that when we looked at the association of high-quality care in heart failure as captured by the 2 clinical heart failure measures included in the CCA adherence score, these measures were positively associated with better AMI outcomes. Thus, good-quality hospital care for heart failure patients seems to be good evidence that AMI patients will receive good care at the same hospital.

A fourth potential limitation is the use of just 1 in-hospital outcome measure and not a collection of outcome measures, such as longer-term mortality and rehospitalization rates. The issue, then, is whether 1 outcome measure is a good proxy for all relevant cardiac patient outcomes. Currently, there is less publicly available CMS data on other outcome measures. As these data become available, future research can determine if in-hospital mortality is correlated with other outcome measures.

CONCLUSIONS

Hospitals, in their quest to delivery high-quality care, seem to organize care for patients with cardiac disease by clinical and administrative systems of care delivery. They also face resource constraints such that attention to administrative cardiac process measures may come at the expense of clinical measures and, therefore, may provide lower quality inpatient care. Our findings imply that a pay-for-performance incentive scheme that acknowledges these facts might lead to better cardiac patient outcomes, at least as measured by inpatient survival rates. As more performance and outcome data become available, it will be critical to reevaluate the design of pay-for-performance program to achieve the desired results of high-quality, cost-effective clinical care.

REFERENCES

1. Glickman SW, Fang-Shu Ou, Delong ER, et al. Pay for performance, quality of care, and outcomes in acute myocardial infarction. *JAMA.* 2007;297:2373–2380.
2. Campbell S, Reeves D, Kontopantelis E, et al. Quality of primary care in England with the introduction of pay for performance. *N Engl J Med*. 2007;357:181–190.
3. Lindenauer PK, Remus D, Roman S, et al. Public reporting and pay for performance in hospital quality improvement. *N Engl J Med*. 2007;356:486–496.
4. Rosenthal MB, Landon BE, Normand SL, et al. Pay for performance in commercial HMOs. *N Engl J Med*. 2006;355:1895–1902.
5. Institute of Medicine. *Performance Measurement: Accelerating Improvement*. Washington, DC: National Academies Press; 2006.
6. O'Brien SM, DeLong ER, Dokholyan RS, et al. Exploring the behavior of hospital composite performance measures: an example from coronary artery bypass surgery. *Circulation*. 2007;116:2969–2975.
7. US Department of Health and Human Services. Hospital Compare Web site. Available at: http://www.hospitalcompare.hhs.gov/. Accessed August 13, 2007.
8. The Joint Commission. Health Care Quality Data. Available at: http://www.healthcarequalitydata.org/. Accessed August 13, 2007.
9. Premier Inc. Centers for Medicare & Medicaid Services (CMS)/Premier Hospital Quality Improvement Demonstration (HQID) project: findings from year two. Available at: http://www.premierinc.com/quality-safety/tools-services/p4p/hqi/resources/hqi-whitepaper-year2.pdf. Accessed November 18, 2007.
10. Marsaglia G. Ratios of normal variables and ratios of sums of uniform variables. *J Am Stat Assoc*. 1965;60:193–204.
11. Bazzoli GJ, Clement JP, Lindrooth RC, et al. Hospital financial condition and operational decisions related to the quality of hospital care. *Med Care Res Rev*. 2007;64:148–168.
12. Volpp KG, Ketcham JD, Epstein AJ, et al. The effects of price competition and reduced subsidies for uncompensated care on hospital mortality. *Health Serv Res*. 2005;40:1056–1077.
13. Fonarow GC, Abraham WT, Albert NM, et al. Association between performance measures and clinical outcomes for patients hospitalized with heart failure. *JAMA*. 2007;297:61–70.
14. Reeves GR, Wang TY, Reid KJ, et al. Dissociation between hospital performance of the smoking cessation counseling quality metric and cessation outcomes after myocardial infarction. *Arch Intern Med*. 2008;168:2111–2117.
15. McNamara RL, Wang Y, Herrin J, et al. Effect of door-to-balloon time on mortality in patients with ST-segment elevation myocardial infarction. *J Am Coll Cardiol*. 2006;47:2180–2186.
16. Lieu TA, Gurley RJ, Lundstrom RJ, et al. Primary angioplasty and thrombolysis for acute myocardial infarction: an evidence summary. *J Am Coll Cardiol*. 1996;27:737–760.
17. Glickman SW, Schulman KA, Peterson ED, et al. Evidence based perspectives on pay for performance and quality of patient care and outcomes in emergency Medicine. *Ann Emerg Med*. 2008;51:622–631.
18. Rosenthal MB, Fernandopulle R, Song HR. Paying for quality: providers' incentives for quality improvement. *Health Aff (Millwood)*. 2004;23:127–141.
19. Doran T, Fullwood C, Gravelle H, et al. Pay for performance programs in family practices in the United Kingdom. *N Engl J Med*. 2006;355:375–384.
20. Glickman SW, Baggett KA, Krubert CG, et al. Promoting quality: the health-care organization from a management perspective. *Int J Qual Health Care*. 2007;19:341–348.
21. Lega F, DePietro C. Converging patterns in hospital organization: beyond the professional bureaucracy. *Health Policy*. 2005;74:261–281.
22. Parker VA, Wubbenhorst WH, Young GJ, et al. Implementing quality improvement in hospitals: the role of leadership and culture. *Am J Med Qual*. 1999;14:64–69.
23. American Heart Association. Heart Disease and Stroke Statistics: 2008 Update-at-a-Glance. Available at: http://www.americanheart.org/downloadable/heart/1200078608862HS_Stats%202008.final.pdf. Accessed February 7, 2008.

© 2009 Lippincott Williams & Wilkins

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.