California HealthCare Foundation

Do Hospital Characteristics Drive Clinical Performance?

An Analysis of Standardized Measures

Prepared for
California HealthCare Foundation

by
Bruce Spurlock, M.D.
Rebecca Abravanel, Ph.D.
Sarah Spurlock

Convergence Health Consulting, Inc.

December 2008

About the Authors

Convergence Health Consulting, Inc. is a multi-disciplinary team of professionals supporting health care organizations to advance patient safety, improve quality, strengthen leadership, promote sustainable change, and facilitate a solution-oriented dialogue among physicians, hospitals, health plans, purchasers, the public and other health care stakeholders. Bruce Spurlock, M.D., President and Chief Executive Officer of Convergence Health, is a national expert in quality and patient safety and serves as the Executive Director for the California Hospital Assessment and Reporting Taskforce (CHART) and as the Executive Director, Clinical Acceleration, for the Bay Area Patient Safety Collaborative (Beacon). Rebecca Abravanel, Ph.D., is a statistician, demographer, and sociologist with research clients in several industries, including health care. Sarah Spurlock is a research assistant and student of political science and history at the University of California, Davis.

About the Foundation

The **California HealthCare Foundation** is an independent philanthropy committed to improving the way health care is delivered and financed in California. By promoting innovations in care and broader access to information, our goal is to ensure that all Californians can get the care they need, when they need it, at a price they can afford. For more information, visit www.chcf.org.

Contents

Introduction

With the advent of the "transparency movement" in health care, a limited but nonetheless substantial amount of standardized information has become available to compare the clinical performance of physicians, hospitals, health plans, and skilled nursing facilities. National organizations such as the Hospital Quality Alliance (HQA), the AQA, the National Quality Forum (NQF), The Joint Commission (TJC), and the National Committee for Quality Assurance (NCQA) promulgate detailed measures, data collection rules, and recommendations for reporting results gleaned from clinical performance information.

As information on hospital performance accumulates, issues emerge that go beyond the basics of which hospitals perform better or worse for a given clinical condition. If it were understood which strategies or characteristics lead to higher clinical performance in some hospitals, other hospitals might adopt them in an effort to provide better care to their patients.

The purpose of this project was to understand which, if any, hospital characteristics were associated with hospital performance as measured by CHART. Since CHART has data on several aspects of clinical care across a range of hospital services, this project sought to identify any patterns that could describe high and low performers among the various measures. The project's goal was to ascertain quantitatively whether standardized, publicly disclosed hospital characteristics were independently associated with performance.

Few studies have quantitatively examined this question in depth: The relationship between measured performance and hospital size or teaching status is inconsistent.[1] Other structural characteristics associated with Leapfrog Group standards have a modest but significant relationship with clinical performance.[2] Ownership (tax-exempt, investor-owned) and rural/urban characteristics also have mixed results.[3] A number of studies have examined common quality process measures supported by the Centers for Medicare & Medicaid Services (CMS) and The Joint Commission (TJC), but these found only a tenuous connection between hospital performance on process measures and clinical outcomes.[4]

This project took a different approach. It used CHART data to assess the impact of various hospital characteristics, taking advantage of new measures of performance — especially clinical outcomes — to test hypotheses and identify patterns. Likewise, the project used operational, financial, and utilization information regarding California hospitals, available through the Office of Statewide Health Planning and Development (OSHPD), to provide additional variables in testing for patterns.

Data from CHART

In California, stakeholders from health plans, purchasers, hospitals, consumer advocates, physicians, and other groups have created a voluntary hospital performance measurement and reporting program called the California Hospital Assessment and Reporting Taskforce (CHART). Initially sponsored by the California HealthCare Foundation (CHCF), CHART produced its first public report on California hospital clinical performance in March 2007. This report captured data from several sources, much of which was publicly available but dispersed among a number of entities and Web sites. It also published previously unreported measures of care such as patient experience, mortality rates for patients in intensive care units, and rates of hospital-acquired pressure ulcers. As of June 2008, 220 hospitals representing 82 percent of the state's daily census (CHART does not include children's hospitals) participate in CHART.

The CHART Web site (www.calhospitalcompare.org) is designed for easy consumer access and use but also provides information appropriate for hospitals undertaking improvement activities, health plans assessing their networks, and purchasers considering health benefits design.

Methodology

Data Sources and Structure

CHART captures data from a number of sources (see Appendix A) and collates the information for each participating California hospital. This project used five data sets from CHART's initial launch in March 2007, plus four of CHART's quarterly data updates (see Appendix B). CHART measures were the dependent variables in the analysis.

The project also used publicly available data from OSHPD: the Hospital Annual Financial and rolling quarterly data sets. The OSHPD data—which include structural, financial, and operational measures—were selected to best match the time periods covered by the CHART data (see Appendix A, Table 2). Data also were gathered from other sources, including the Accreditation Council for Graduate Medical Education (ACGME) and researchers who developed a report on California hospitalists.[5] In all, 30 hospital attributes were evaluated against CHART results.

Analysis

The project's analysis initially examined the impact of various hospital attributes on clinical outcomes. Figures on coronary artery bypass graft (CABG) mortality, intensive care unit (ICU) mortality, hospital-acquired pressure ulcer (HAPU) rates, and patient experience were examined to determine if a relationship existed for any of the 30 hospital attributes. The project explored whether the relationship met basic statistical criteria and whether it was linear or non-linear (e.g., U-shaped, asymptotic). Such a relationship, or lack of relationship, did not determine the final effect within a regression model, where other factors are controlled to establish whether the relationship is independent of other variables. But it is important to understand patterns in the data to more accurately build the models. For correlations of individual hospital characteristics with clinical outcomes from this analysis, see Appendix E.

Once the overall structure of the data and the relationships between hospital attributes and clinical outcomes were understood, the analysis continued by building regression models to determine whether any variables that previously had shown no relationship now emerged (having been suppressed in the individual analysis by another variable) or variables that had demonstrated a relationship now disappeared (were not independently associated with the clinical outcome). Summaries of these regression models are presented in Appendix F.

Finally, once the impact of the variables on outcomes was understood, the project tested whether the variables had any impact on improvement. This part of the analysis examined the impact of hospital attributes on selected core measures and on two outcome measures: CABG mortality and patient experience. Individual hospital performance for each CHART measure was analyzed to determine if performance was improving or declining. Because the CHART data for each core measure and patient experience cover a rolling 12-month period, both the earliest and latest results were examined. For the specific time periods, see Appendix A. The project also examined the characteristics of the hospitals with the greatest improvement and the hospitals whose performance declined, using logistic regression to predict the likelihood of membership in each of these groups.

Key Findings

Of the 30 hospital characteristics the project examined against 11 primary performance and improvement measures, only four were suggestive—using presence in at least three statistical models as the threshold—as independent predictors of performance:

- Disproportionate Share Hospital (DSH) status (meaning a significant proportion of indigent care and Medi-Cal patients);
- Membership in a large system (more than ten hospitals);

- Percentage of gross revenue collected; and
- Initial starting value (in the improvement analyses).

DSH Status

DSH designation was a significant factor in five models: CABG mortality, ICU mortality, patient experience, and improvement in both heart failure and one surgical infection prevention measure (SIP 3). In all but one of these cases, being a DSH facility produced worse results or had a negative impact on improvement. The exception was observed with respect to one surgical infection prevention measure (SIP 3), where DSH status was significantly positively associated with improvement.

DSH status is a marker for high levels of indigent and Medi-Cal-funded care, and reflects a patient population with socio-demographics that differ from populations in non-DSH facilities. The patients who make up the DSH population may have a higher severity of illness, but presumably the risk-adjustment models control for a large portion of such clinical severity. DSH status also might be a marker regarding gross revenue and profitability, but these factors disappeared when controlled for in all of the models, as did DSH revenue as a percentage of gross revenue. In other words, DSH status as a source of revenue or as an indicator of financial health does not factor into clinical performance. Further, the proportion of minority discharges was not a factor in any of the 11 models, suggesting that race alone cannot account for the observed results. DSH status, then, is probably a marker for socio-demographic or clinical factors that are not represented by any of the 30 variables tested in this study.

The available data do not reveal whether the mostly negative performance and weaker improvement associated with DSH status is a result of poorer hospital care, more challenging and unmeasured patient characteristics, or both. The fact that DSH status was prominent in several performance domains suggests that disparities of care in the hospital setting are more widespread than previously thought. This finding requires more in-depth analysis using other data sets that include socio-demographic measures.

Membership in a Large System

Membership in a health system of more than ten hospitals was significantly related to the improvement measure for heart failure care and to two surgical improvement measures. In addition, membership in a medium-to-large system (four or more hospitals) was positively associated with improvement in the pneumonia quality measure. However, system size was not in the final model for either performance or improvement in any of the outcome measures. This suggests that large health systems may be better able to improve process measures over time, but that large system size has no effect on overall clinical outcomes. It is noteworthy that the profitability of large health systems was not a factor regarding either improvement or clinical outcomes.

The explanation for this modest relationship between large system size and a few improvement measures is not apparent from the data but could stem from several factors, including: 1) making improved performance a priority (several health systems include core measure performance in executive compensation formulas, but the data on this are not publicly available); 2) allocating corporate resources toward improvement at the expense of other priorities; 3) sharing effective improvement practices among system hospitals; and 4) using technology and corporate information systems to drive improvement, as well as other factors that were not publicly available. These are hypotheses only, however, and these and other possible factors require further study and analysis.

Percentage of Gross Revenue Collected

The percentage of gross revenue collected is a calculated variable in the OSHPD financial data set, reflecting the actual amount received from patients and third-party payers as a percentage of all patient revenue. Essentially, it represents the ability of hospitals to collect money for the services they provide. It comes as no surprise that the higher percentage of gross revenue hospitals collect, the better their performance on some quality indicators—in this study, ICU mortality, patient satisfaction, and improvement in patient satisfaction and one surgical infection prevention composite (SIP 3). Presumably, these hospitals are able to use the collected revenue to improve quality of care. One exception should be noted, however, with regard to the otherwise positive impact of percentage of gross revenue collected: This variable is negatively associated with improvement in one surgical infection prevention measure (SIP 1), but the reason is unclear.

It is noteworthy that while several other measures of income also were examined—gross inpatient revenue, pre-tax net income, and net income—the gross revenue collected measure had the only consistently significant effect in the models. That is, the percentage of gross revenue collected was important with respect to performance and performance improvement, but neither the level of revenue (gross or net) nor profitability was significantly correlated with clinical performance.

Initial Starting Level of Performance

During the analysis of improvement measures, the project evaluated the impact of the initial starting value as an independent predictor of improvement. For all of the change-over-time measures, this had the greatest impact on improvement. In several cases, this factor alone explained between 25 percent and 50 percent of the variance in improvement. While initial starting value is not, strictly speaking, a hospital attribute, it was the dominant element in all of this study's models exploring improvement, and has been described elsewhere.[6] This is not surprising, since the absolute opportunity to improve is the greatest with the weakest performers. Also, since the highest level of performance for the process measures is 100 percent, the ability of the strongest performers to improve is capped.

Non-factors in Hospital Performance

While this study found a few hospital attributes that appear modestly related to performance, the large majority of factors that never or rarely appear in the models may be more significant. In other words, the bigger story may be the many hospital characteristics that have no bearing on hospital performance or improvement.

Nine major domains (representing 15 variables) did not factor in any of the 11 hospital performance or improvement models. These hospital attributes are:

- Teaching status;
- Rural hospital status;
- Measures of hospital value (net property, plant, and equipment [PPE], net PPE/bed, and bad debt);
- Length of stay;
- Presence of hospitalists;
- CABG volume;
- Proportion of managed care revenue and managed care days;
- Proportion of minority discharges;
- Performance on the Leapfrog computerized provider order entry standard;
- Performance on the Leapfrog intensivist physician standard;
- System membership (two or more hospitals);
- Profitability (total margin, operating margin, and cost-per-charge ratio);
- Staffing measures;
- Occupancy measures; and
- Pre-tax net income.

Another five major hospital domains (representing 12 variables) were present only in one of the 11 models—a relationship much too tenuous to consider them performance predictors.

Three major hospital domains were present in two models but were inconsistent in the direction or type of measure:

- Region was a factor in HAPU rates and ICU mortality but different regions were involved in each and no discernable pattern existed with regional groups;
- Ownership status was a factor, but with strikingly different results: investor-owned hospitals had less improvement in one surgical improvement measure and tax-exempt hospitals were weakly associated with improved patient experience; and
- Gross system revenue was related to better performance in CABG mortality and CABG improvement but is not involved in any other clinical area.

These three hospital characteristics are only weak predictors of performance in a limited number of models and do not display any consistent patterns. Statistically, they have some role in each of the clinical conditions where they were present in the model but may well be markers for other, unreported aspects of hospital performance. In sum, it is difficult to make the case that they are important predictors, and they probably are not factors in the larger view of hospital performance.

Study Limitations

Although the study included a large number of hospital characteristics, the list of attributes it examined is not exhaustive. Most quality experts believe other attributes are also important. With regard to improvement, in particular, many experts suggest that leadership participation plays a role. In this regard, future study could include a survey instrument recently adopted by CMS to assess leadership practices associated with better performance and safety (the Healthcare Leadership and Quality Assessment Tool). Participation in an executive compensation program that rewards improved performance on items like CHART measures also may be a marker of executive engagement. Now that standardized measures are available, this latter issue is ripe for study to determine how compensation formulas contribute to clinical improvement.

The level of engagement and participation by hospital medical staff is generally recognized as playing a role in clinical performance. Anecdotal reports from hospital leadership routinely mention physician engagement as a strategic imperative when talking about transparency and public reporting, but this issue is much harder to assess since standardized tools and approaches are not available.

Hospitals consider staffing level in the quality department as a driver of improved clinical performance. As with all hospital staffing discussions, however, the exact makeup and responsibilities of the quality department are difficult to categorize. Some hospitals have begun to move quality department responsibility for specific clinical measures toward the operational unit involved. For example, surgical infection prevention is the responsibility of the operating room along with the surgery department, while ICU process measures reside with the ICU staff. The differences in these practices would make comparisons among hospitals difficult.

Specific departmental structures and practices within operational units at hospitals were not studied in this project (except for the Leapfrog group standard involving intensivist physicians). With regard to hospitalist programs, a hospital-wide structure this project evaluated, they did not appear to be independent drivers of the clinical areas examined. There are now standardized measures with which to examine specific practices and structures across a large group of hospitals, so other organizational structures and practices should be studied in the future.

Testing in this study was very limited regarding the impact of health information technology. Only one such proxy (from Leapfrog), related to computerized physician order entry, was studied, and it was not an independent driver of performance on the measures examined. Since health information technology is a major part of discussions in both the hospital market and health policy circles, it should be studied against actual performance.

Implications

The findings from the study's analysis of various hospital characteristics on a range of clinical performance measures may have significant implications regarding public reporting, performance improvement, and public policy.

Transparency and Improvement

Initially, the transparency movement in health care attempted simply to shine some light on clinical performance by hospitals, physicians, other providers, and health plans so that consumers and purchasers could make more informed decisions regarding their health care choices. Over time, however, the notion that publicly reported information might be widely used to make choices has waned, replaced by hope that it could be used to spur improvement in clinical performance. Some recent studies suggest that public reporting has had some limited effect on motivating work on quality improvement, but for the most part the drivers of improvement are poorly understood.[7] (It is also thought that publicly reported information may promote the development and deployment of new financing structures, such as pay-for-performance and value-based purchasing, but the scope of this project did not permit an evaluation of such structures.)

The data from this study suggest that only a few out of a large number of mostly structural hospital attributes are associated with more than one or two performance measures. Somewhat surprisingly, a hospital's or health system's financial health is not broadly associated with either performance or improvement. While this study did not assess all structural hospital characteristics, the fact that so many had limited or no impact begins to challenge long-held assumptions about the role of such characteristics in quality improvement. Donabedian's famous quality triad—with structure, process, and outcomes in equipotent roles—does not seem to hold up based on the data analyzed in this study. Of course, certain structural variables that were not tested in this study (primarily because data for them are not available) may conceivably play significant roles, so this conclusion about Donabedian's triad remains a tentative one.

Based on this study's findings, however, it can be postulated that clinical performance and improvement are predominantly driven not by structural hospital attributes but by the reliable implementation of effective practices across a diverse group of patients with a given condition. This includes attention to execution, and an understanding of care-related work flow abetted by the growing field of improvement science.[8] Some organizational structures may yield higher or lower levels of reliable implementation, but it is likely these structures will be directly related to the actual practices rather than broad constructs such as hospital size, teaching status, ownership, hospital financial health, or the many other attributes evaluated in this study.

There are several implications of the project's analysis for public reporting and performance:

1. With regard to improvement measures, several are nearly "topped off" and so provide little information to consumers and other stakeholders. Also, since improvement is often most strongly predicted by a low initial starting value, by definition significant improvement will slow as starting value rises. As a result, public reporting entities may want to alter the use of topped-off measures such as "all-or-nothing" reporting (in which, for a given condition, if any individual measure is not performed, care is deemed to be unreliable and a score of zero—"nothing"—is rendered), or simply report those hospitals that perform below a chosen threshold.

2. Since this study found no relationship between most hospital attributes—such as teaching status, hospital size, ownership, and other characteristics—and clinical performance, stratifying hospitals by these categories tends to mislead the intended audience. DSH status may be an exception, but there should be no stratification by DSH status until it is clear that performance differences are due primarily to poorer care and not to unmeasured patient characteristics.

3. Because of the lack of correlation between structural characteristics and performance measures, hospitals should emphasize the execution of improvement strategies rather than major structural or transformational strategies.

4. All hospital types have the opportunity to improve, regardless of their organizational structures, finances, or initial level of performance. Different types of hospitals and hospital systems can share effective improvement strategies with a fair amount of confidence that the strategies are not unique to their facility.

Public Policy

Policy discussions regarding hospital quality often revolve around ownership type, region, or some other organizational structure. This study uncovered that only in a small number of models did these types of attributes drive performance. That is, they do not serve as broad proxies for overall quality of care at California hospitals. Consequently, policymakers should adjust their approach regarding matters of performance, ignoring these structural characteristics when creating legislation or regulation to improve California hospital quality.

In the specific case of DSH facilities, the findings of this study suggest an urgent need to explore why this status often correlates with poor performance. DSH facilities receive revenue from the federal government based on a complicated formula using matching funds sent by state and local governments. But this study determined that DSH status is not associated with poorer performance because of a weaker financial picture or insufficient funds overall for the hospitals.

A different explanation for poorer performance focuses on the patient population. DSH status is likely to reflect patients with particular socio-demographic features—patients who tend to have poor follow-up after discharge, who often experience delays in care for a variety of reasons, and who may not respond to treatments in the same ways as other patients. If the explanation for poorer outcomes and less improvement in DSH facilities is found to result mostly from such unmeasured patient characteristics, then most risk-adjustment models meant to normalize various patient factors will need modification. Also, public reporting programs might consider altering their stratification of DSH facilities in order to account for the differences in patient population.

There is also, however, a need to investigate whether and to what extent DSH facilities perform worse as the result of poorer care delivery. If additional research uncovers actionable deficiencies in care practices, specific policies will need to attend to those deficiencies. From a policy perspective, poorer care at DSH facilities needs to be addressed regardless of its sources, whether unmeasured differences in patient factors, deficient care practices, or other undetermined causes.

Endnotes

1. Kroch, E.A., M. Duan, S. Silow-Carroll, and J. Meyer. April 2007. "Hospital Performance Improvement: Trends in Quality and Efficiency." The Commonwealth Fund. www.commonwealthfund.org/publications/publications_show.htm?doc_id=471264; Thornlow, D.K., and G. Stukenborg. March 2006. "The Association Between Hospital Characteristics and Rates of Preventable Complications and Adverse Events." *Medical Care* 44(3): 265–269.

2. Jha, A.K. June 2008. "Does the Leapfrog Program Help Identify High-Quality Hospitals?" *Journal on Quality and Patient Safety* 34(6): 318–325.

3. Hines, S., and M. Joshi. June 2008. "Variation in Quality of Care Within Health Systems." *Journal on Quality and Patient Safety* 34(6): 326–332; Kahn, C. January/February 2006. "Snapshot of Hospital Quality Reporting and Pay-for-Performance Under Medicare." *Health Affairs* 25(1): 148–162.

4. Bradley, E.H., et al. July 2006. "Hospital Quality for Acute Myocardial Infarction: Correlation Among Process Measures and Relationship with Short-term Mortality." *Journal of the American Medical Association* 296: 72–78.

5. Auerbach, A., and E. Vasilevskis. University of California, San Francisco. Personal communication.

6. Kroch, E.A. et al. "Hospital Performance Improvement."

7. Fung, C.H., et al. January 2008. "Systematic Review: The Evidence That Publishing Patient Care Performance Data Improves Quality of Care." *Annals of Internal Medicine* 148(2): 111–123; Hibbard, J. January 2008. "What Can We Say About the Impact of Public Reporting? Inconsistent Execution Yields Variable Results." *Annals of Internal Medicine* 148(2): 160–161.

8. Berwick, D.M. March 2008. "The Science of Improvement." *Journal of the American Medical Association* 299(10): 1182–1184.

Appendix A: Data Sources

This study relied upon data from CHART concerning hospital performance across various measures. CHART collects data from a number of sources:

1. Core measures (CMS and TJC) are submitted from a hospital's data vendor who submits the same measures to national program Web sites.
2. Clinical outcomes regarding CABG and pneumonia mortality are derived from OSHPD reports.
3. ICU mortality variables and process measures are submitted by hospitals.
4. Patient experience data are submitted from a hospital's data vendor, specific to that measure domain, through the National CAHPS Benchmarking Database.
5. Hospital-acquired pressure ulcer (HAPU) results are submitted to the California Nursing Outcomes Coalition (CalNOC).
6. Leapfrog performance standards for participating California hospitals are obtained from the Leapfrog Group (www.leapfroggroup.org).

Performance data from CHART are refreshed quarterly; this project used five data sets, from March 2007 through May 2008. Some measures, such as HAPU rates and ICU mortality, were added in January 2008. Table 1 shows the sample period for the various measures in CHART.

Table 1. Project Sample Periods for CHART Data

DATA REFRESH SET	MARCH 2007	JULY 2007	OCTOBER 2007	JANUARY 2008	MAY 2008
CABG	2003	2003	2004	2005	2005
Maternity	2004	2004	2004	2006	2006
Patient Experience	December 2005 through February 2006	March through December 2006	March through December 2006	July 2006 through June 2007	July 2006 through June 2007
TJC/CMS Core Measures	Q3 2005 through Q1 2006	Q1 through Q4 2006	Q2 2006 through Q1 2007	Q3 2006 through Q2 2007	Q4 2006 through Q3 2007
Leapfrog	June 2006	June 2006	June 2007	Not directly reported	Not directly reported
OSHPD Pneumonia		2002 to 2004	2002 to 2004	2002 to 2004	2002 to 2004
ICU Process and Outcomes				Q1 through Q3 2007	Q1 through Q4 2007
CalNOC HAPU				Q1 through Q3 2007	Q1 through Q4 2007

The study also used financial, operational, and structural data on California hospitals from OHSPD. Other measures are taken from OSHPD reports, some of which are updated annually, others less frequently. Table 2 describes the time periods for OSHPD data that were used as a portion of the independent variables in the analysis.

Table 2. Project Sample Periods for OSHPD Data

	DATA SOURCE AND TIME PERIOD	
PERFORMANCE MEASURE	CHART	OSHPD
CABG	Q1 through Q4 2005	2005 annual financial data
Patient Experience	Q3 2006 through Q2 2007	Rolling quarterly financial data covering Q3 2006 through Q2 2007
ICU	Q1 through Q3 2007	Rolling quarterly financial data covering Q4 2006 through Q3 2007
HAPU	Q1 through Q3 2007	Rolling quarterly financial data covering Q4 2006 through Q3 2007

The project augmented OSHPD data on hospital characteristics with information from other sources:

- System size (number of hospitals) from the Web sites of each health system with hospitals in California;
- Teaching program designation from the Accreditation Council for Graduate Medical Education (ACGME—www.acgme.org/adspublic) list of residency programs in California (a more extensive list than the OSHPD teaching hospital designation and the Council of Teaching Hospitals [COTH] lists, which include only major teaching programs);
- System revenue for tax-exempt organizations from each organization's IRS Form 990 for 2006, found at www.guidestar.org; and
- Identification of hospitalist programs in California generated in the report "The Rise of the Hospitalist in California," Vasilevskis et al., July 2007, published by the California HealthCare Foundation (www.chcf.org/documents/policy/RiseHospitalistCalifornia.pdf). Additional conversations with the report's authors validated information regarding hospitalist programs in California in 2007.

Appendix B: Measures

CHART Measures (Dependent Variables)

The hospital performance measures reported by CHART can be found at www.calhospitalcompare.org, with detailed data definitions at chart.ucsf.edu. Briefly, CHART produces measures for the following:

- ICU mortality, using the MPM II risk-adjustment methodology (S. Lemeshow, D. Teres, J. Klar, J. S. Avrunin, S. H. Gehlbach, and J. Rapoport. November 1993. "Mortality Probability Models [MPM II] Based on an International Cohort of Intensive Care Unit Patients." *Journal of the American Medical Association* 270[20]: 2478–2486.) A sample of hospitals participates in an external audit. ICU mortality is reported on a rolling 12-month basis.
- Hospital-acquired pressure ulcers (HAPU), using a prevalence methodology (www.calnoc.org). Hospitals receive training and certification to conduct the study. HAPU data are reported on a rolling 12-month basis.
- CABG mortality, using a customized risk-adjustment methodology with medical record-based clinical variables calculated by OSHPD (www.oshpd.ca.gov/HID/Products/Clinical_Data/CABG/2005/CCORP2005_Web.pdf). CABG mortality is reported annually by OSHPD for calendar year periods.
- Patient experience, using a modified form of the HCAHPS® tool with four additional experience questions. Patient experience results are reported on a rolling 12-month basis.
- ICU process measures, which are self-reported, related to reducing ventilator-associated pneumonia. A subset of measures is audited by an external firm in a sample of hospitals. ICU process measures are reported on a rolling 12-month basis.
- Core measures, including acute myocardial infarction, heart failure, surgical improvement, and pneumonia, using CMS/TJC specifications. A weighted composite of the individual results for each domain is created (chart.ucsf.edu). Core measures are reported on a rolling 12-month basis.

Hospital Characteristics (Independent Variables)

Table 3 groups the hospital characteristics used as independent variables in this study into five basic categories.

Table 3. Hospital Characteristics by Category

INCOME	VALUE	PROFITABILITY	STRUCTURE	OTHER
Pre-tax net income	Net PPE (property, plant, and equipment [PPE] and construction-in-progress)	Total margin	Size	Region
Net income	Net PPE per bed	Operating margin	Teaching status	Rural
Gross inpatient revenue	Bad debt	Cost-per-charge ratio	Ownership (city/county, district, investor, non-profit)	Total nursing level
Percentage of gross revenue collected			System membership	Staffed or licensed bed occupancy
Disproportionate Share Hospital (DSH) status			System size (number of hospitals)	Adjusted length of stay
System gross revenue			Leapfrog computerized physician order entry standard	Presence of hospitalists
Proportion of managed care revenue			Leapfrog intensivist physician standard	Proportion of… • minority discharges • managed care days CABG volume

Appendix C: Hospital Performance by Condition

Coronary Artery Bypass Graft Mortality

The study included CABG performance data for both CHART hospitals (N=100) and non-CHART hospitals (N=20). While no significant differences were found between the CHART and non-CHART groups, there was significant variation in CABG mortality (range 0 to 11.5 percent) among individual hospitals.

Large hospitals fare worse when it comes to CABG mortality: Hospitals with annual discharges of 18,000 or more tend to have CABG mortality rates about 1.2 percent higher than hospitals with fewer discharges. System revenue also has a significant, though small, effect on CABG mortality: The larger the system revenue, the lower the CABG mortality. The decrease in mortality is 0.067 percent for each additional billion dollars in system revenue.

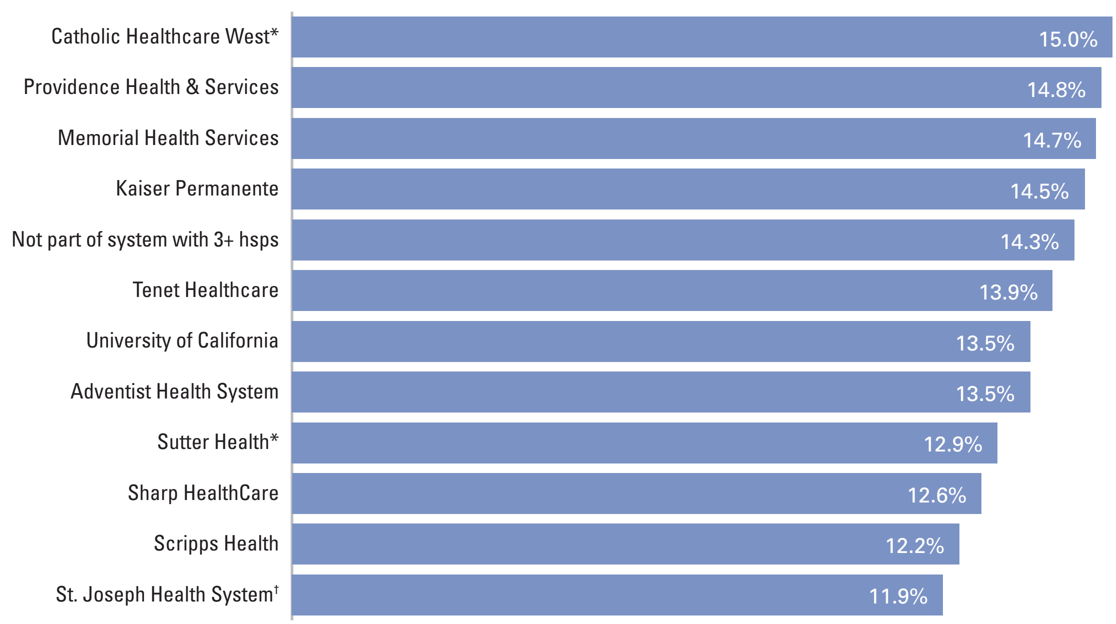
DSH facilities tend to have an almost 1 percent higher level of CABG mortality. This is due to DSH status alone: This difference remains when various financial factors are controlled for, regardless of hospital size (number of discharges).

Total nurse staffing levels also significantly affect CABG mortality: On average, the higher the nursing level, the lower the CABG mortality. More specifically, for each additional paid registered nurse or licensed vocational nurse hour per patient day, CABG mortality decreases by about 0.34 percent.

ICU Mortality

With regard to performance regarding ICU mortality, there is a fair amount of variation among all hospitals (range 9.5 percent to 25 percent), and there are significant differences among hospital systems, as shown in Figure 1.

Figure 1. ICU Mortality by Health System



*Difference between this system and all other hospitals statistically significant at the 95 percent level of confidence.

†Difference between this system and all other hospitals statistically significant at the 90 percent level of confidence.

When other variables are controlled using a regression model, five factors play a role in the mortality rate associated with ICU care. (The high performance noted in Figure 1 for St. Joseph Health System and Sutter Health disappears in the regression model and so is not independently associated with better ICU mortality performance.) Each of the five ICU mortality factors has a nearly equal role in the model, but the model only weakly predicts mortality (R^2=0.126). These five factors are:

- Central region (predominantly the Central Valley) hospitals have worse performance;
- DSH hospitals have worse performance;
- Hospitals that do not submit any information to the Leapfrog Group on their implementation of the intensivist physician standard are associated with worse performance; and
- Hospitals in the CHW and Kaiser Permanente systems as a whole have worse performance (although individual hospitals may have average or better-than-average performance).

The study saw no financial or other structural hospital characteristics independently associated with ICU mortality rates.

It is unclear why the factors noted above are important, except that DSH status might reflect unmeasured clinical socio-demographic or other factors, poorer care for a more challenging population, or a combination of both (as discussed in the body of this report). Clearly, poorer performance for CHW and Kaiser systems is worth addressing by the particular hospitals and corporate structures involved.

Another important issue is that ICU mortality is almost certainly affected by pre-ICU care, perhaps even care provided before the patient is seen in the hospital. This might explain why the statistical relationship between the model and mortality rates is so weak: Overall actual care plays a much greater role in determining the outcome than hospital infrastructure or finance.

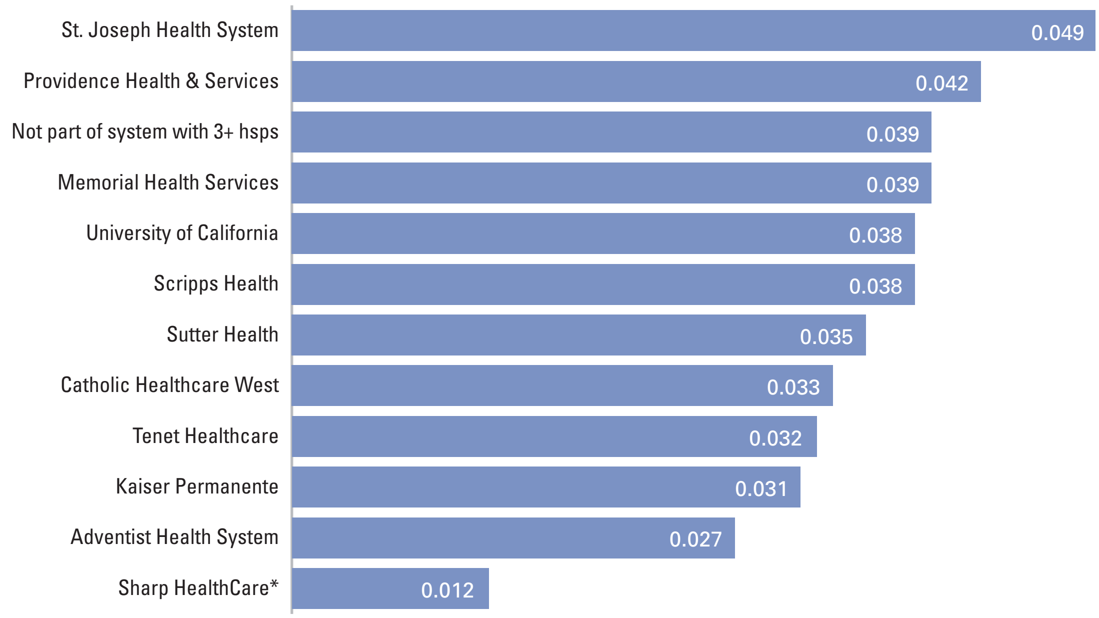
Hospital-Acquired Pressure Ulcers

Analysis of HAPU performance showed that HAPU rates differ by region, with northern California hospitals in rural areas and in the extended Sacramento area having much lower—better—rates. This relationship did not hold for the San Francisco Bay Area, or for central or southern California regions. The reason for this relationship is unknown—it is not explained by membership in a health system in general or in a particular health system, or by being a rural hospital. The results may reflect the fact that these hospitals have addressed this subject longer than other hospitals: Some of these hospitals have participated for years in a program with CalNOC to measure and internally report HAPU rates, so perhaps a greater number of hospitals in this region are CalNOC participants.

Hospital size also is associated with HAPU rates: Small hospitals (fewer than 6,500 discharges per year) consistently have better rates. Better performance in small hospitals was not correlated with being a rural hospital, since this factor was not in the final model. And while smaller hospitals are often financially challenged because of high fixed costs and the lack of economies of scale found in larger hospitals, it turned out that income, profitability, and value were not associated with HAPU rates. This might be because smaller hospitals may be able to provide better attention to fewer patients or to fewer higher-risk patients. For example, these hospitals may tend to transfer long term ICU patients to other settings.

A higher proportion of hospital days paid by managed care organizations is associated with worse HAPU performance. This was the only financial variable involved in the regression model. Why the proportion of managed care hospital days is a feature in only this one model is unclear.

Figure 2. Hospital-Acquired Pressure Ulcer Rates by Health System



*Difference between this system and all other hospitals statistically significant at the 95 percent level of confidence.

It is worth noting that when specific health systems with more than three hospitals were examined (see Figure 2, above), Sharp HealthCare stood out as a high performer and independent predictor of success regarding HAPU rates. The practices Sharp HealthCare has employed to lower HAPU rates probably should be studied, including the feasibility of implementing them at other hospitals.

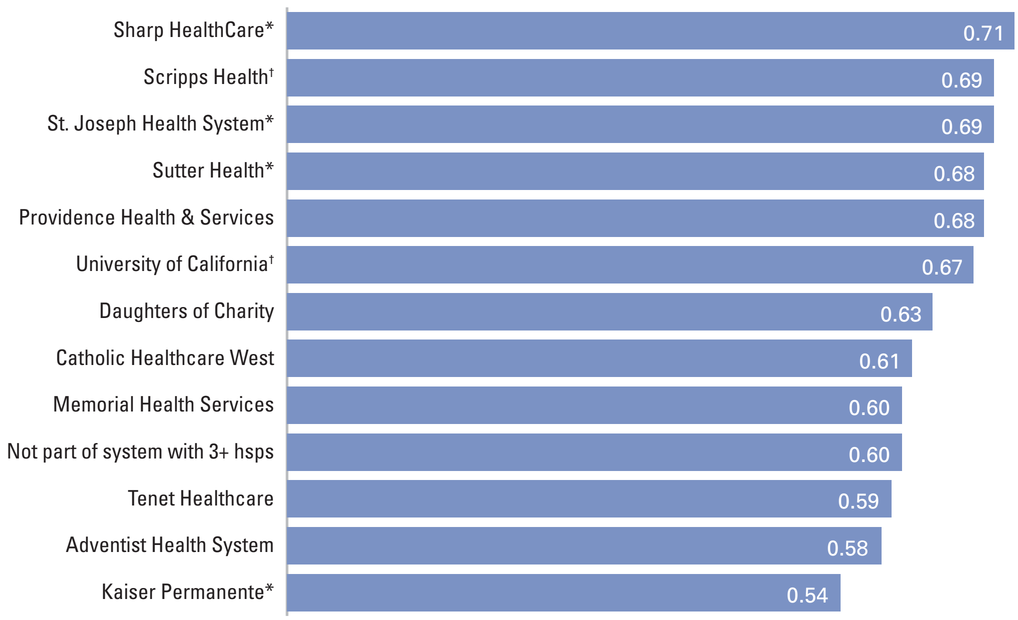
Patient Experience

The initial model confirms many of the findings from the individual variable evaluations. For example, medium-size hospitals (between 4,000 and 16,500 annual discharges) tend to have lower rates of patient satisfaction than other hospitals—an effect not traceable to other factors such as nonprofit status, DSH status, hospital financial health, or occupancy rates. The effect on patient experience scores, however, is small (4 percent). A higher licensed bed occupancy rate also has a negative impact on patient satisfaction, independent of the other variables included in the model, with a modest effect (9 percent).

DSH facilities also tend to have slightly lower (4 percent) rates of patient satisfaction, an effect not related to financial or other factors in the model. The study explored whether these differences might be driven by the racial diversity of the DSH patient population, but such differences did not explain this finding.

Hospitals that have a higher percentage of gross revenue collected, have higher total margins, or are nonprofit tend to have higher rates of patient satisfaction. An increase in the percentage of gross revenue collected yields a very small (0.2 percent) increase in patient satisfaction, while a 1 percent increase in total margin yields the largest effect on performance, a 20 percent increase in patient satisfaction scores. On average, nonprofit hospitals score 3 percent higher than other hospitals. As seen in Figure 3 on the next page, however, this nonprofit factor disappears when specific systems with four or more hospitals are included, suggesting that nonprofit status is a proxy for specific, but not all, tax-exempt health systems.

Figure 3. Patient Satisfaction Rates by Health System



*Difference between this system and all other hospitals statistically significant at the 95 percent level of confidence.

†Difference between this system and all other hospitals statistically significant at the 90 percent level of confidence.

Initial findings of system performance are shown in Figure 3.

When individual systems with four or more hospitals were added to the initial model, four health systems retained independent associations with patient satisfaction (Sharp HealthCare, Sutter Health, St. Joseph Health System, and Scripps Health). Status as a tax-exempt (not-for-profit) hospital disappeared in the final model, which strongly suggests that it is these four health systems that actually explain performance, rather than simply the fact of being a nonprofit. Sharp HealthCare had the highest effect, with a 10 percent higher level of performance.

Appendix D: Hospital Improvement Over Time

CABG Mortality Improvement

The study examined CABG mortality change in performance between 2003 and 2005 as reported by OSHPD. There was no overall improvement in mortality during this time, and fewer than 20 percent of hospitals continuously improved from 2003 through 2005.

The likelihood of improvement depended heavily on the initial starting level of performance. Figure 4, below, demonstrates improvement in mortality, in relation to initial level of mortality, between 2003 and 2005.

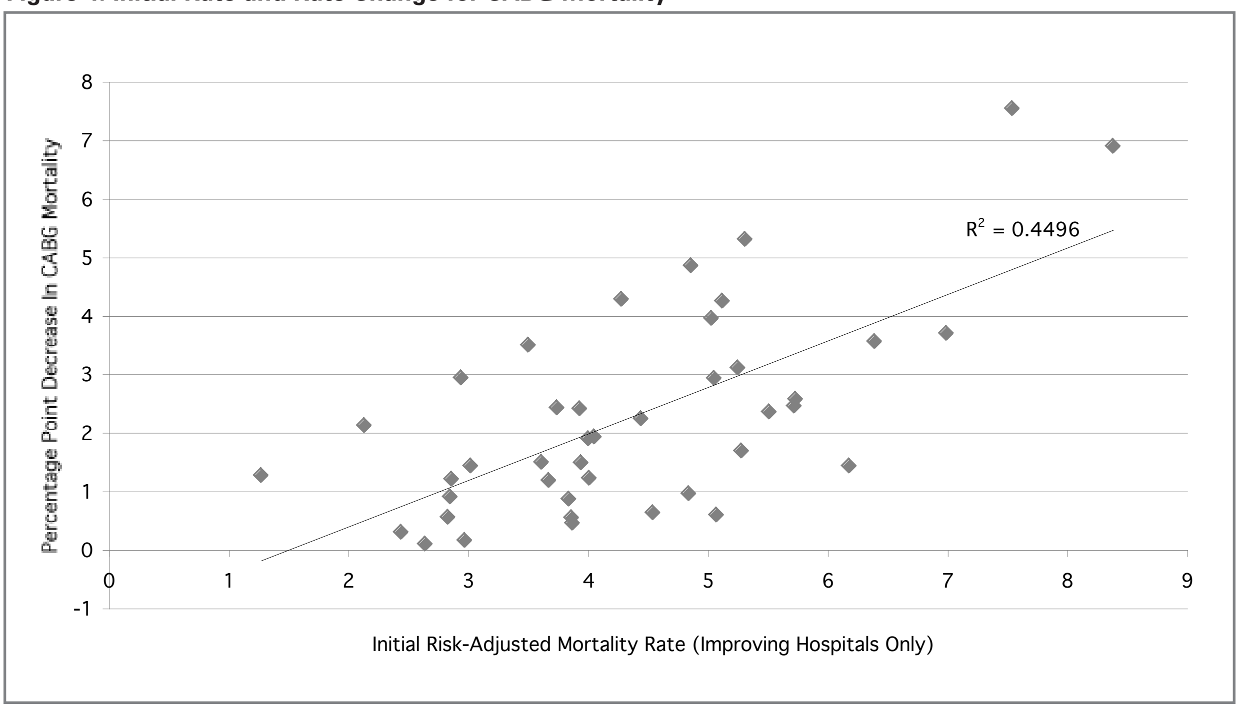
Fifty percent of the variance in CABG mortality rate change over time was explained by the initial starting rate. As discussed in the body of this report, this can be partially explained by the fact that mortality is capped at 0 percent, and therefore high-performing hospitals have much less room to improve. Another reason initial value is the strongest driver among the improvement models is that OSHPD recalibrates the risk-adjustment model annually to provide the best statistical fit for the data, irrespective of whether there is any change in clinical practices or performance. This tends to normalize performance over time, even if a hospital was improving or worsening.

Other factors that independently support high or low CABG improvement include size—large hospitals (more than 18,000 annual discharges) show less improvement—and higher gross system revenue, which yields greater improvement.

It is notable that CABG volume was not a factor in any model, nor was it present in the overall performance analysis.

The study also analyzed mortality improvement using a broader definition: the presence of improvement, rather than the amount, using logistic regression. In this broader definition, the only additional factor besides initial starting mortality rate

Figure 4. Initial Rate and Rate Change for CABG Mortality



was that lower total nurse staffing is associated with less improvement. Notably, two factors associated with increased improvement rates—large hospital size and higher gross system revenue—are not correlated with whether hospitals actually improved, even by very small amounts.

Finally, the study assessed whether good-to-better hospitals (initial mortality rate below the median, and improvement over time) increased their improvement over time, and poor-to-worse (initial mortality rate above the median, and mortality increased) worsened over time. The study also used a stricter definition of "good-to-better," looking only at the hospitals that initially had below-median mortality rates (better performance) and that reduced mortality further by at least 0.9 percent. While the number of hospitals in these groups was small (nine in the looser definition and five in the stricter definition), the relationship with the initial starting rate and total nurse staffing remains significant. The result is further evidence that total nurse staffing is related to CABG mortality improvement, but not necessarily to large improvement. Although this measure reflects hospital-wide staffing levels while the specific activities of nursing in CABG programs are often quite circumscribed (i.e., CABG nursing staff are mostly exclusive to the CABG program and patients), the relationship is nonetheless present and deserves further analysis.

ICU Mortality and HAPU Improvement

No analysis was performed on improvement with ICU mortality or HAPU rates because of insufficient available data (only one point in time from the January update of the www.calhospitalcompare.org Web site).

Patient Experience Improvement

Overall patient experience in California hospitals showed virtually no overall change (median 0.601 to 0.604) during the two periods analyzed in the study. Of the CHART measures this project evaluated, patient experience had the least amount of clustering near the highest levels of performance. In other words, it is not topped out: Both relatively low and relatively high performers still have plenty of room for improvement.

With respect to patient experience, the initial level of performance demonstrates an independent relationship as a predictor of improvement, but with an explanatory value of 5 percent, it is the lowest of all the improvement models. This is at least partially due to the fact that the measure is far from being topped out.

The only other variable affecting patient experience improvement over time was the percentage of gross revenue collected. Hospitals that do a better job of collecting on their billed charges seem to improve more on this measure, while those that do not seem to improve the least. This may be a function of contracting rather than improvement, but since the charge-to-cost ratio variable is not present in the model, this is not a result of a lower relative level of charges. In other words, the prices charged, as a function of total costs, do not affect patient experience improvement.

Logistic regression offers another way to examine hospital attributes affecting the probability of improvement. This method can indicate how hospitals with the most improvement are different from the others. Since many hospitals improved patient satisfaction by a very small amount, this study evaluated the likelihood of being in the top 12 percent of hospitals (having at least a 4 percentage point improvement in patient satisfaction score). The only key drivers of higher improvement are the initial level of patient satisfaction (the higher the initial score, the less likely the hospital is to be in the high-improvement group) and being an investor-owned hospital (rather than nonprofit, county, or district hospital), which makes it significantly less likely to be in the high-improvement group.

The study also analyzed the probability of a hospital's patient experience score worsening by at least 3 percent over the period, which would put it in the bottom 9 percent of hospitals with respect to improvement. These effects turned out to mirror those for predicting the likelihood of being a high-performing hospital. For each percentage point in initial patient satisfaction score, the likelihood of being among the least-improving hospitals increases by a small amount. Being an investor-owned hospital also increases a hospital's chances of being in the worst-improving group, by a factor of more than two. Taking both logistic model results together, investor-owned hospitals are twice as likely to fall into the worst hospital group with respect to improvement, and half as likely to place among the high-performing hospitals with respect to improvement, regardless of their initial score.

CHART Core Measure Improvement

The study analyzed a subset of the CHART measures related to nationally reported core measures. Of the CHART-weighted composites for various groups of core measures, the study examined:

- Acute myocardial infarction (AMI) quality;
- Congestive heart failure quality;
- Pneumonia quality;
- Surgical Infection Prevention 1 (SIP 1); and
- Surgical Infection Prevention 3 (SIP 3).

Three composite CHART measures were not evaluated in this study: AMI timeliness, heart failure prevention, and pneumonia prevention. These were excluded because the number of patients and hospitals in the AMI timeliness measure was small, and because the prevention measures have no immediate effect on hospital outcomes.

Acute Myocardial Infarction Quality

There was a small but significant overall improvement in AMI quality scores (median 94.8 to 96.2 percent) in California hospitals over the time span analyzed in this study. Since this measure is close to topping out, with over half the hospitals above 96 percent performance, only small improvement would be expected.

The amount of improvement depended heavily on the starting AMI quality score, which explains about 33 percent of the variation. No other variables—e.g., financial factors, DSH status, system size, region, teaching status—had a significant effect on the AMI quality composite. Similarly, when using logistic regression to analyze the probability of being among above-average-improving hospitals, the only driver is the initial score.

Congestive Heart Failure Quality

Overall heart failure quality composite scores among California hospitals improved significantly but modestly, by 4.5 percent, during the time periods evaluated (median 82.6 to 87.1 percent). This measure is moving toward topping out and, as with other such measures, the initial level of performance explains a large amount (25 percent) of the variance in change over time.

Two variables, in addition to initial level, appear to have small but statistically significant effects on change in heart failure improvement. DSH facilities tended to have declining scores over the period studied (by 1.7 points, on average), while hospitals in relatively large systems (10 or more hospitals) tended to improve more than average (by about 1.7 points). Other variables were found to have no significant net effects in the model.

In a logistic regression analysis, the study looked at the likelihood of achieving above-average improvement. In this analysis, the key drivers—in addition to initial level of performance—appear to be large size (more than 18,000 annual discharges), reflecting less likelihood to have above-average improvement, and higher pre-tax net income, which has a small but positive effect.

Pneumonia Quality

CHART hospitals overall experienced a small but significant improvement in pneumonia quality of 2.3 percent (median 89.1 to 91.4 percent) over the time period studied. This measure is also moving toward topping out in the near future.

As with other core measures, the degree of pneumonia quality improvement depends heavily on the starting score. However, the initial level is not as important for the pneumonia quality composite as for the AMI or heart failure quality scores (11 percent compared with 33 percent and 25 percent, respectively). Two other variables also appear to have statistically significant negative effects on pneumonia composite improvement: the cost-to-charge ratio and membership in a small system (fewer than three hospitals, or independent). No other variables were found to have significant net effects in the model.

When the likelihood of achieving above-average improvement was examined, the key drivers appeared to be: initial level; a higher proportion of patient days covered by managed care payers, which has a positive effect; and location in the state's central region, which has a significant negative effect. The proportion of managed care days has a fairly substantial impact on above-average improvement; this raises the interesting possibility that coordination with managed care entities, which in California often means capitated medical groups, explains some of this effect.

Surgical Infection Prevention 1

During the period studied in this project, CHART hospitals experienced a moderate overall improvement in SIP 1 performance (related to pre-surgery prophylactic antibiotic initiation) of 6 percent (median 80.6 percent to 86.6 percent). Many hospitals are topping out on this measure, but there is still room for substantial improvement among low-performing hospitals.

As with other core measures, the amount of improvement depends heavily on the initial SIP 1 score, which accounts for 35 percent of the variance in score change. Three other variables also appear to have statistically significant effects on change in the SIP 1 score:

- **Gross revenue collected.** For each percentage point increase in gross revenue collected, the SIP 1 score decreases by about a tenth of a percentage point. While small, the effect is statistically significant at the 90 percent level of confidence.
- **Ownership type.** Investor hospitals were more likely to have their scores worsen over the period—by about 4.4 percentage points, on average.
- **System size.** Hospitals in large systems tended to improve scores by about 3.4 percentage points more than other hospitals.

No other tested variables were found to have significant net effects in the model.

The study also examined the likelihood of achieving above-average improvement. In this model, the key drivers appear to be the initial level of SIP 1 (the higher the initial score, the less likely the hospital is to be in the above-average-improvement group) and being part of a large hospital system (which makes it more likely to be in the high-improving group).

Surgical Infection Prevention 3

CHART hospitals' SIP 3 performance (associated with appropriate cessation of prophylactic antibiotics) showed the largest absolute improvement of all of the CHART measures studied, at 10.3 percent (median 67.2 to 77.6 percent), over the time period evaluated in the study. As with other core measures, the degree of improvement depends heavily on the initial SIP 3 score, which accounts for 32 percent of the variance in score change. Four other factors also appear to have statistically significant effects on change with this measure:

- **Net income.** For each $1 million increase in pre-tax net income, the SIP 3 score increases by about three one-hundredths of a percent. For example, if one hospital has $100 million more in pre-tax net income than another hospital, its SIP 3 score is likely to improve, on average, by about 3 percentage points more than the other hospital. While small, the effect is statistically significant at the 90 percent level of confidence.

- **DSH status.** Perhaps surprisingly, DSH facilities were more likely to have their scores improve over the period—on average by about 3.4 percentage points more than non-DSH facilities.
- **System size.** Hospitals in large systems tended to improve their scores more, on average by about 3.7 percentage points more than other hospitals.
- **Staffed beds.** Staffed bed occupancy had a positive impact on SIP 3 improvement: With each unit increase in the staffed bed occupancy rate, the SIP 3 composite score increased by 4.7 points, on average.

No other variables were found to have significant net effects in the model.

Appendix E: Statistical Results of Major Hospital Attributes

The following table displays statistically significant (of at least $p \leq 0.10$) performance measures for 14 major hospital attributes (which aggregate all 30 subcategories) analyzed in this study as independent variables.

INDEPENDENT VARIABLE	STATISTICALLY SIGNIFICANT PERFORMANCE MEASURE										
	CABG	CABG CHG	ICU	HAPU	PAT SAT CHG	PAT SAT	AMI CHG	HF CHG	PN QUAL CHG	SIP 1 CHG	SIP 3 CHG
Region			Central – worse performance‡	Northern California outside of the Bay Area – better performance*							
Hospital Size	Large (18,000 or more annual discharges) – worse performance*	Large (18,000 or more annual discharges) – worse improvement*		Small (fewer than 6,500 annual discharges) – better performance‡		Medium (4,000 to 16,500 annual discharges) – worse performance†					
Ownership						Nonprofit – better performance*				Investor-owned – worse improvement*	
System Membership			CHW – worse performance*								
System Size								Large (10+ hospitals) – better improvement*	Small/none (1–3 hospitals) – worse improvement*	Large (10+ hospitals) – better improvement†	
DSH Status	DSH hospital – worse performance		DSH hospital – worse performance			DSH hospital – worse performance*		DSH hospital – worse improvement			DSH hospital – better improvement*
Income			Higher percentage of gross revenue collected – better performance*		Higher percentage of gross revenue collected – better performance*					Higher percentage of gross revenue collected – worse performance*	Higher percentage of gross revenue collected – better performance*
Profit						Higher total margin – better performance‡			Higher cost to charge ratio – worse improvement*		

*$p \leq 0.05$

†$p \leq 0.005$

‡$p \leq 0.0005$

INDEPENDENT VARIABLE	STATISTICALLY SIGNIFICANT PERFORMANCE MEASURE										
	CABG	CABG CHG	ICU	HAPU	PAT SAT CHG	PAT SAT	AMI CHG	HF CHG	PN QUAL CHG	SIP 1 CHG	SIP 3 CHG
Staffing	Higher nurse staffing levels – better performance*										
Occupancy Rates						Higher licensed bed occupancy rate – worse performance‡					Higher staffed bed occupancy rate – better improvement*
Proportion of Managed Care-Days				Higher proportion third-party managed care days – better performance*							
IPS			Non-submission of intensivists data to Leapfrog – worse performance*								
System Revenue	Higher gross system revenue – better performance	Higher gross system revenue – better improvement									
Initial Level		Higher mortality rate in 2003 – better improvement‡			Lower initial level of performance – better improvement†		Lower initial level of performance – better improvement‡				

*p ≤ 0.05
†p ≤ 0.005
‡p ≤ 0.0005

Note: The following attributes were not found in any of the models: rural location; teaching hospital designation; value (net PPE, bad debt); length of stay; presence of hospitalists; CABG volume; and computerized physician order entry.

Appendix F: Statistical Results from Regression Models

The specific statistical regression models for each dependent variable are listed below. The explanatory power of the model is indicated by the R^2 value, with higher results indicating a better statistical fit with the data. Unstandardized coefficients describe the absolute impact of each variable in the model, using the unit of measurement (number of discharges, billions of dollars, etc.) specific to each variable. The standardized coefficients describe the relative impact of each variable and can be compared to one another. For example, in the CABG model, having 18,000 or more discharges (Beta=0.279) has a relatively larger impact on CABG mortality than does being a DSH hospital (Beta=0.181). Only variables with statistically significant effects are shown; if a variable is not shown, it means that it does not have a significant effect on the given measure once the other variables in the model are included.

Table 4. CABG Mortality

	Unstandardized Coefficients		Standardized Coefficients	
Variable	B	Std. Error	Beta	Sig.
(Constant)	5.749	1.138		0.000
Large Hospital (more than 18,000 annual discharges)	1.241	0.438	0.279	0.006
System Gross Revenue* (in billions of dollars)	−0.067	0.039	−0.159	0.094
DSH Hospital Status	0.981	0.522	0.181	0.064
Nurse Staffing Level	−0.337	0.139	−0.234	0.017

Notes: N=100 and R^2=0.188

*System Gross Revenue is calculated from 2005 tax year forms for all nonprofit systems. For non-system hospitals and for-profit systems, gross system revenue is the sum of gross revenue across relevant hospitals listed in the full OSHPD 2005 annual financial data.

Sources: CHART CABG Data (2005) and OSHPD Annual Financial Data (2005).

Table 5. ICU Mortality

	Unstandardized Coefficients		Standardized Coefficients	
Variable	B	Std. Error	Beta	Sig.
(Constant)	14.79	0.75		0.00
Central Region	1.98	0.68	0.21	0.00
CHW	1.19	0.49	0.18	0.02
DSH Hospital Status	0.96	0.50	0.14	0.06
Did Not Submit Leapfrog Intensivist Physician Standard	1.80	0.69	0.19	0.01
Percentage of Gross Revenue Collected	−0.06	0.03	−0.16	0.03

Notes: N=166 and R^2=0.184

Source: CHART ICU Data (Q1 – Q3 2007) and OSHPD Quarterly Financial Data (Q4 2006 – Q3 2007).

Because the system membership performance varied as noted in the table above, the study tested all systems (not just CHW) for significant effects but had to remove the variable Percentage of Gross Revenue Collected since OSHPD does not collect that information for Kaiser Permanente hospitals. Tables 6, 7 and 8 present the results of that analysis.

Table 6. ICU Model Results Including Kaiser Permanente Hospitals

VARIABLE	UNSTANDARDIZED COEFFICIENTS B	UNSTANDARDIZED COEFFICIENTS STD. ERROR	STANDARDIZED COEFFICIENTS BETA	STANDARDIZED COEFFICIENTS SIG.
(Constant)	13.191	0.257		0.000
Central	1.865	0.699	0.186	0.008
DSH Hospital Status	0.956	0.532	0.127	0.074
Did Not Submit Leapfrog Intensivist Physician Standard	1.660	0.725	0.160	0.023
Kaiser Permanente	1.210	0.549	0.156	0.029
CHW	1.207	0.524	0.162	0.022

Notes: N=195 and R^2=0.126

Sources: CHART ICU Data (Q1 – Q3 2007) and OSHPD Quarterly Financial Data (Q4 2006 – Q3 2007).

Table 7. HAPU

VARIABLE	UNSTANDARDIZED COEFFICIENTS B	UNSTANDARDIZED COEFFICIENTS STD. ERROR	STANDARDIZED COEFFICIENTS BETA	STANDARDIZED COEFFICIENTS SIG.
(Constant)	0.048	0.003		0.000
Northern California (excluding Bay Area)	–0.009	0.004	–0.171	0.021
Small Hospital (fewer than 6,500 annual discharges)	–0.016	0.004	–0.305	0.000
Proportion of Third-Party Managed Care Days	–0.028	0.009	–0.208	0.004
Sharp HealthCare	–0.027	0.010	–0.173	0.011

Notes: N=191 and R^2=0.174

Sources: CHART HAPU Data (Q1 – Q3 2007) and OSHPD Quarterly Financial Data (Q4 2006 – Q3 2007).

Table 8. Patient Experience

VARIABLE	UNSTANDARDIZED COEFFICIENTS B	UNSTANDARDIZED COEFFICIENTS STD. ERROR	STANDARDIZED COEFFICIENTS BETA	STANDARDIZED COEFFICIENTS SIG.
(Constant)	0.657	0.028		0.000
Medium-Size Hospital (4,000 to 16,500 annual discharges)	-0.040	0.011	-0.221	0.001
DSH Hospital Status	-0.042	0.014	-0.183	0.005
Percentage of Gross Revenue Collected	0.002	0.001	0.148	0.032
Total Margin	0.191	0.052	0.242	0.000
Licensed Bed Occupancy Rate	-0.084	0.024	-0.226	0.001
St. Joseph Health System	0.071	0.031	0.140	0.024
Scripps Health	0.067	0.034	0.121	0.050
Sharp HealthCare	0.103	0.034	0.187	0.003
Sutter Health	0.044	0.016	0.174	0.007

Notes: N=184 and R^2=0.362

Sources: CHART Patient Satisfaction Data (Q3 2006 – Q2 2007) and OSHPD Quarterly Financial Data (Q3 2006 – Q2 2007).

CABG Mortality Improvement

The first analysis looked at potential factors that affect the absolute amount of improvement over time. Table 9 displays the results.

Table 9. Variables in the Equation: Percentage Point Improvement in CABG Mortality, 2003–2005 (OLS Regression)

	UNSTANDARDIZED COEFFICIENTS		STANDARDIZED COEFFICIENTS	
VARIABLE	B	STD. ERROR	BETA	SIG.
(Constant)	–3.374	0.422		0.000
Initial Mortality Rate (2003)	1.064	0.112	0.708	0.000
Large Hospital (more than 18,000 annual discharges)	–0.895	0.416	–0.157	0.034
Gross System Revenue (in billions of dollars)	0.069	0.039	0.132	0.078

Notes: N=91 and R^2=0.548

Another way to describe improvement is whether any improvement occurred during the time period studied. Using logistic regression, the result is the model shown in Table 10.

Table 10. Variables in the Equation: Any Improvement, 2003–2005 (Logistic Regression)

VARIABLE	B	S.E.	WALD	DF	SIG.	EXP(B)
(Constant)	–7.799	2.197	12.604	1	0.000	0.000
Initial Mortality Rate (2003)	1.026	0.217	22.461	1	0.000	2.791
Total Nursing Staff	0.520	0.230	5.125	1	0.024	1.683

Notes: N=91 and Nagelkerke R^2=0.541

The project examined good-to-better hospitals (initial mortality rate below median, with improvement) and poor-to-worse (initial mortality rate above median, and worse mortality over time), which yielded the models shown in Tables 11 and 12.

Table 11. Variables in the Equation: Good-to-Better Hospitals,* 2003–2005 (Logistic Regression)

VARIABLE	B	S.E.	WALD	DF	SIG.	EXP(B)
(Constant)	–6.159	1.935	10.131	1	0.001	0.002
Total Nursing Staff	0.454	0.210	4.664	1	0.031	1.574

Notes: N=91 and Nagelkerke R^2=0.105

*Below median-level mortality in 2003 and improved by 2005.

Table 12. Variables in the Equation: Poor-to-Worse Hospitals,* 2003–2005 (Logistic Regression)

VARIABLE	B	S.E.	WALD	DF	SIG.	EXP(B)
(Constant)	–3.821	0.928	16.949	1	0.000	0.022
Initial Mortality Rate (2003)	0.356	0.185	3.691	1	0.055	1.427
Large Hospital (more than 18,000 annual discharges)	1.273	0.673	3.572	1	0.059	3.571

Notes: N=91 and Nagelkerke R^2=0.162

*Above median-level mortality in 2003 and no improvement by 2005.

Patient Experience Improvement

Table 13. Variables in the Equation: Percentage Point Improvement in Patient Satisfaction, July 2007–January 2008 (OLS Regression)

VARIABLE	UNSTANDARDIZED COEFFICIENTS		STANDARDIZED COEFFICIENTS	
	B	STD. ERROR	BETA	SIG.
(Constant)	0.044	0.017		0.013
Initial Patient Satisfaction	–0.090	0.028	–0.255	0.001
Percentage of Gross Revenue Collected	0.001	0.000	0.167	0.034

Notes: N=165 and R^2=0.071

The study also analyzed the likelihood of being a top 12 percent hospital (improving by at least 4 percentage points) using logistic regression.

Table 14. Variables in the Equation: Likelihood of Top 12 Percent Improvement re Patient Satisfaction, July 2007–January 2008 (Logistic Regression)

VARIABLE	B	S.E.	WALD	DF	SIG.	EXP(B)
(Constant)	2.967	1.091	7.391	1	0.007	19.428
Initial Patient Satisfaction	–0.042	0.018	5.666	1	0.017	0.959
Investor-Owned Hospital	–0.873	0.452	3.733	1	0.053	0.418

Notes: N=192 and Nagelkerke R^2=0.060

Finally, the study looked at the likelihood of poor performance over time (being among the bottom 9 percent of hospitals—decreasing by at least 3 percentage points).

Table 15. Variables in the Equation: Likelihood of Poor Performance re Patient Satisfaction Improvement, July 2007–January 2008 (Logistic Regression)

VARIABLE	B	S.E.	WALD	DF	SIG.	EXP(B)
(Constant)	–2.868	1.098	6.821	1	0.009	0.057
Initial Patient Satisfaction	0.038	0.018	4.666	1	0.031	1.039
Investor-Owned Hospital	0.818	0.448	3.331	1	0.068	2.265

Notes: N=165 and Nagelkerke R^2=0.316

Core Measure Improvement

AMI Quality Composite

Table 16. Variables in the Equation: Percentage Improvement in AMI Quality Composite, 2006–07
(OLS Regression)

VARIABLE	UNSTANDARDIZED COEFFICIENTS		STANDARDIZED COEFFICIENTS	
	B	STD. ERROR	BETA	SIG.
(Constant)	30.778	3.426		0.000
AMI QC, Time 1	−0.311	0.036	−0.573	0.000

Notes: N=184 and R^2=0.329

Another way to assess this relationship is to examine whether a hospital had above-average improvement rather than the percentage point improvement.

Table 17. Variables in the Equation: Above-Average AMI Quality Improvement, 2006–07 (Logistic Regression)

VARIABLE	B	S.E.	WALD	DF	SIG.	EXP(B)
(Constant)	27.047	5.351	25.551	1	0.000	5.577E11
Initial AMI Quality Score	−0.287	0.056	26.212	1	0.000	0.750

Notes: N=153 and Nagelkerke R^2=0.321

Heart Failure Quality Composite

Table 18. Variables in the Equation: Percentage Improvement in Heart Failure Composite, 2006–07
(OLS Regression)

VARIABLE	UNSTANDARDIZED COEFFICIENTS		STANDARDIZED COEFFICIENTS	
	B	STD. ERROR	BETA	SIG.
(Constant)	26.440	2.537		0.000
Initial Heart Failure Value	−0.272	0.031	−0.550	0.000
DSH Hospital Status	−1.658	0.899	−0.114	0.067
Large System Hospital (membership in system with 10+ hospitals)	1.663	0.693	0.150	0.017

Notes: N=201 and R^2=0.290

Table 19. Variables in the Equation: Above-Average Improvement in Heart Failure Composite, 2006–07
(Logistic Regression)

VARIABLE	B	S.E.	WALD	DF	SIG.	EXP(B)
(Constant)	10.580	2.047	26.726	1	0.000	39358.722
Initial Heart Failure Value	−0.125	0.024	27.819	1	0.000	0.882
Large Hospital (more than 18,000 annual discharges)	−1.546	0.521	8.815	1	0.003	0.213
Pre-Tax Net Income (in millions of dollars)	0.012	0.005	5.426	1	0.020	1.012

Notes: N=170 and Nagelkerke R^2=0.369

Pneumonia Quality Composite

Table 20. Variables in the Equation: Percentage Improvement in Pneumonia Composite, 2006–07
(OLS Regression)

VARIABLE	UNSTANDARDIZED COEFFICIENTS B	UNSTANDARDIZED COEFFICIENTS STD. ERROR	STANDARDIZED COEFFICIENTS BETA	STANDARDIZED COEFFICIENTS SIG.
(Constant)	26.073	4.344		0.000
Initial Pneumonia Quality Value	–0.252	0.045	–0.442	0.000
Cost-to-Charge Ratio	–5.331	2.380	–0.171	0.026
Small System Hospital (membership in system with 1–3 hospitals)	–0.805	0.397	–0.150	0.044

Notes: N=201 and R^2=0.157

Table 21. Variables in the Equation: Above-Average Improvement in Pneumonia Quality Composite, 2006–07
(Logistic Regression)

VARIABLE	B	S.E.	WALD	DF	SIG.	EXP(B)
(Constant)	15.047	3.947	14.535	1	0.000	3425900.876
Initial Pneumonia Quality Value	–0.177	0.045	15.602	1	0.000	0.838
Proportion of Managed Care Days	2.493	0.987	6.385	1	0.012	12.100
Central Region	–1.146	0.588	3.799	1	0.051	0.318

Notes: N=171 and Nagelkerke R^2=0.164

Surgical Improvement Prevention 1 (SIP 1)

Table 22. Variables in the Equation: Percentage Improvement in SIP 1 Composite, 2006–07
(OLS Regression)

VARIABLE	UNSTANDARDIZED COEFFICIENTS B	UNSTANDARDIZED COEFFICIENTS STD. ERROR	STANDARDIZED COEFFICIENTS BETA	STANDARDIZED COEFFICIENTS SIG.
(Constant)	38.613	4.118		0.000
Initial SIP 1 Value	–0.374	0.040	–0.577	0.000
Percentage of Gross Revenue Collected	–0.126	0.074	–0.111	0.091
Investor-Owned Hospital	–4.398	1.605	–0.182	0.007
Large System Hospital (membership in system with 10+ hospitals)	3.390	1.048	0.203	0.001

Notes: N=201 and R^2=0.441

Table 23. Variables in the Equation: Above-Average Improvement in SIP 1 Composite, 2006–07
(Logistic Regression)

VARIABLE	B	S.E.	WALD	DF	SIG.	EXP(B)
(Constant)	10.441	1.871	31.129	1	0.000	34231.565
SIP 1, Time 1	–0.135	0.022	36.338	1	0.000	0.873
Large System Hospital (membership in system with 10+ hospitals)	0.720	0.353	4.151	1	0.042	2.054

Notes: N=197 and Nagelkerke R^2=0.401

Surgical Improvement Prevention 3 (SIP 3)

Table 24. Variables in the Equation: Percentage Improvement in SIP 3 Composite, 2006–07
(OLS Regression)

	UNSTANDARDIZED COEFFICIENTS		STANDARDIZED COEFFICIENTS	
VARIABLE	B	STD. ERROR	BETA	SIG.
(Constant)	23.151	3.441		0.000
Initial SIP 3 Value	−0.321	0.035	−0.579	0.000
Pre-Tax Net Income (in millions of dollars)	0.031	0.014	0.141	0.027
DSH Hospital Status	3.420	1.504	0.143	0.024
Large System Hospital (membership in system with 10+ hospitals)	3.693	1.172	0.197	0.002
Staffed Bed Occupancy	4.665	2.009	0.144	0.021

Notes: N=165 and R^2=0.401

Table 25. Variables in the Equation: Above-Average Improvement in SIP 3 Composite, 2006–07
(Logistic Regression)

VARIABLE	B	S.E.	WALD	DF	SIG.	EXP(B)
(Constant)	3.362	0.732	21.083	1	0.000	28.837
Initial SIP 3 value	−0.061	0.011	30.816	1	0.000	0.941
Large System Hospital (membership in system with 10+ hospitals)	0.645	0.338	3.652	1	0.056	1.907

Notes: N=19 and Nagelkerke R^2=0.195

Table 26. Variables in the Equation: Decrease in Performance in SIP 3 Composite, 2006–07
(Logistic Regression)

VARIABLE	B	S.E.	WALD	DF	SIG.	EXP(B)
(Constant)	−8.772	2.258	15.097	1	0.000	0.000
Initial SIP 3 Value	0.089	0.028	10.057	1	0.002	1.094
Pre-Tax Net Income (in millions of dollars)	−0.032	0.015	4.720	1	0.030	0.969

Notes: N=165 and Nagelkerke R^2=0.316

California HealthCare Foundation

1438 Webster Street, Suite 400
Oakland, CA 94612
tel: 510.238.1040
fax: 510.238.1388
www.chcf.org