UC Davis

Reports for the Agency for Healthcare Research and Quality

Title

Outcomes of Hospital Outcome Studies. Abstract, Executive Summary and Final Report. - Report for 30 Sep 94-31 Oct 99

Permalink

https://escholarship.org/uc/item/96g7b5sv

Author

Romano, Patrick S

Publication Date

2002-04-25

eScholarship.org Powered by the California Digital Library
University of California

FINAL REPORT

TO

THE AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

FOR

GRANT #R29 HS08574

"THE OUTCOMES OF HOSPITAL OUTCOME STUDIES"

PROJECT PERIOD 9/30/94-10/31/99

PATRICK S. ROMANO, MD MPH
PRINCIPAL INVESTIGATOR
UNIVERSITY OF CALIFORNIA, DAVIS
DIVISION OF GENERAL MEDICINE AND
CENTER FOR HEALTH SERVICES RESEARCH IN PRIMARY CARE

TABLE OF CONTENTS

ABSTRACT R29 HS 08574

This study was designed to evaluate the impact of hospital report cards in New York and California. We conducted mail and telephone surveys of hospital administrators, quality improvement leaders, and managed care executives. We analyzed hospital discharge data to ascertain effects on consumer choice.

Hospital administrators and quality improvement leaders attend to report cards, but remain skeptical about their quality and value. New York's report, based on detailed clinical data, received better ratings than California's report, based on administrative data. Administrators at low-mortality hospitals rated the report better and found it more useful than those elsewhere. A few quality improvement leaders described projects resulting from public disclosure; most expressed frustration at the untimely publication schedule and the lack of actionable information about care.

There is substantial interest in hospital quality among managed care executives, but objective data receive little attention in contracting. Health plans rely more on quality measures with poor discrimination (i.e., accreditation) or subjective concepts (i.e., reputation, commitment to quality improvement, member satisfaction).

Report cards had modest, inconsistent, and transient effects on consumer choice of hospitals. These effects were larger in New York, but still lasted just two months. Volume shifts were greater among white and HMO-enrolled patients (in California) than among others.

EXECUTIVE SUMMARY R29 HS08574

Survey of hospital administrators

In the last decade, several state agencies and private coalitions have published report cards on risk adjusted hospital outcomes for specific conditions or procedures. Despite the controversy surrounding these, little is known about their value and impact. The objectives of this study were to determine whether recent hospital report cards are viewed more favorably than pioneering federal efforts, whether a report based on clinical data is viewed more favorably than one based on administrative data, and whether attitudes toward report cards are related to hospital characteristics.

We conducted a mail survey of chief executives at 374 California hospitals and 31 New York hospitals listed in report cards on acute myocardial infarction (AMI) and coronary bypass mortality. After multiple contacts, 274 hospitals in California (73.3%) and 27 in New York (87.1%) responded. California hospitals were categorized on ownership, size, occupancy, risk-adjusted mortality, teaching status, patient volume, and surgical capability. Our principal measures were the number of hospital units that received or discussed the report card, ratings of its quality, perceptions of its usefulness, and knowledge of its risk-adjustment methods.

In both states, report cards were widely disseminated within hospitals. Hospital administrators' mean quality rating, on a scale of 0-4 where 0 represents "poor" and 4 represents "excellent," was 1.4 for the California report, 1.7 for the New York report, and 0.6 for earlier Health Care Financing Administration (HCFA) reports. The New York report was rated significantly superior to the California report in its usefulness for improving quality of care, accuracy in describing hospital performance, and ease of interpretation. The two reports were rated similarly in the completeness of their risk-adjustment models, usefulness to consumers, and method of release. Hospital leaders in both states found outcomes reports to be moderately useful for improving the quality of care and (in California) the quality of ICD-9-CM coding. Thirty to fifty percent of respondents found the reports to be useful for marketing or negotiating with health plans. However, hospital leaders had limited understanding of the risk-adjustment methods used in their state's report cards. The overall mean knowledge score, on a scale of 0-4 where 4 represents correct responses to all questions, was 2.0 in California and 2.4 in New York (p=.075).

In California, fewer hospital units reviewed or discussed the report at for-profit hospitals than at nonprofit corporate hospitals, due to less frequent dissemination to medical (53% versus 82%) and nursing (59% versus 75%) staff leaders. The number of hospital units that reviewed or discussed the report, and leaders' knowledge of its risk-adjustment methods, were positively associated with both hospital size and AMI volume. Hospitals labeled as high-mortality outliers disseminated the report more widely than non-outliers. Leaders at low-mortality outlier hospitals rated the quality of the report significantly better, were more knowledgeable about its risk-adjustment methods, and found it more useful (principally for improving quality of care and marketing), than leaders at other hospitals. Leaders at for-profit hospitals found the report more useful for a variety of purposes, despite being less knowledgeable about its methods, than leaders

at nonprofit corporate hospitals.

Recent hospital report cards were rated better than pioneering federal efforts. A report based on clinical data in New York was rated better, understood better, and disseminated more often to key staff, than one based on administrative data in California. Barriers to constructive use of outcomes data persist, especially at high mortality hospitals, where administrators tend to blame the messenger instead of critically evaluating their own processes of care. State agencies and private coalitions must produce clearer and more timely report cards to overcome providers' skepticism, especially at the hospitals that might benefit the most by carefully examining their outcomes and processes of care.

Survey of hospital quality improvement leaders

Public report cards on health care organizations proliferated in the 1990s. However, it remains unclear how hospitals respond to external evaluations of their effectiveness, and whether these responses improve quality of care for present and future patients. The objectives of this study were to explore whether and how hospitals used the 1996 risk-adjusted outcome data and reports from the California Hospital Outcomes Project (CHOP). We were specifically interested in what types of quality improvement activities were undertaken in response to public disclosure.

We undertook a two-stage survey of hospital leaders in California to explore how the 1996 CHOP reports and data were used to improve organizational performance. In the first stage, described above in greater detail, we mailed a questionnaire to the chief executive officer of each hospital in the report. We asked each executive to identify an individual within his or her facility who was most familiar with the hospital outcomes report and who was involved in orchestrating the hospital's response. In the second stage, we interviewed a stratified random sample of the 129 individuals identified in this manner. These telephone interviews were semi-structured with open-ended, pretested questions, averaging 20-25 minutes.

Thirty-nine interviews were completed, representing 84% yield after replacing informants who failed to return 6 messages. Most informants (74%) were involved in quality improvement. About three-quarters found some aspect of the CHOP report to be useful, especially for "benchmarking" performance, improving ICD-9-CM coding, and educating physicians about documentation and clinical pathways. The most common criticisms were that the report was not timely and described death rates without providing practical information about the process of care. Other concerns included poor standardization of ICD-9-CM coding across hospitals, improper attribution of post-transfer deaths to the originating hospital, unfair comparison of dissimilar hospitals, and excessive complexity and technical detail.

Most interviewees reported disseminating the CHOP report to medical staff leaders and quality improvement committees. However, two-thirds of respondents indicated that no specific activities resulted. A few hospitals reestimated their risk-adjusted mortality rates after correcting ICD-9-CM codes or removing cases that were felt to bias the results. Critical pathways for AMI management were instituted or refined in three hospitals, after careful review of existing

processes of care. Three hospitals evaluated their use and timing of thrombolytic therapy in the emergency room; at least one documented improvement. At two hospitals, poor ratings led administrators to change the medical staff members responsible for treating AMI patients in the emergency room. Another hospital rated as "worse than expected" developed "a protocol for evaluation and triage of patients...with chest pain" as well as "an on-call panel (of cardiologists) for invasive procedures." Three hospitals undertook extensive activities to improve coding, such as in-service sessions to educate coders or improve physician documentation.

Most respondents felt that other available information about quality of care, such as process data from Genentech's National Registry of Myocardial Infarction and HCFA's Cooperative Cardiovascular Project, is more useful than the less timely and more general data provided by CHOP. Many specific suggestions were offered for improving future reports, such as using more recent data, simpler and shorter explanations, better graphic displays, and more information about what the "better" hospitals are doing differently.

Although the 1996 CHOP reports and data were widely disseminated within hospitals, most reported uses were ceremonial. This finding reflects two critical weaknesses of the project: non-timely data and lack of information about the process of care. Nevertheless, hospital quality managers recognize that public report cards are here to stay, and some carefully studied their outcomes data to identify areas for improvement. As outcomes reporting becomes more widespread and more timely, these activities should increase.

Survey of managed care executives

Managed care organizations (MCOs) are directly or indirectly responsible for selecting hospitals on behalf of their members. In an efficient market, these organizations should collect and analyze information about both price and quality, leading to well-informed and defensible contracting decisions. The objectives of this study were to determine whether managed care executives in California are familiar with hospital report cards, to what extent they find them useful, and how they weight such data in comparison with other factors, such as price and convenience.

We obtained a list of MCOs that had active Knox-Keene licenses, as of June 1998, from the California Department of Corporations. After identifying a contact person within each of the 47 HMOs that contracted with hospitals to provide acute inpatient medical care, we distributed questionnaires by mail and followed up with multiple telephone calls as needed. A separate list of self-insured California employers was obtained from the Department of Labor. Sixty-four unique employers, other than Kaiser Permanente, sponsored medical benefit plans with at least 1000 participants. Of the 25 employers that were still in business, were still self-insured, and agreed to answer questions, 21 used an administrator that was eligible for our MCO survey to process claims and to create a list of preferred hospitals. Therefore, we did not send separate questionnaires to self-insured employers. Health plans were categorized on the number of enrollees, the numbers of contracted hospitals and medical groups, tax status, ownership, model type, and accreditation by the National Committee on Quality Assurance.

Thirty of the 47 eligible health plans (64%) provided a usable questionnaire or interview. Three factors were almost universally rated as very important in hospital contracting decisions: accreditation by the Joint Commission for the Accreditation of Healthcare Organizations (JCAHO), geographic location, and negotiated price. Respondents from accredited plans tended to rate geographic location as less important than did respondents from non-accredited plans. Respondents from for-profit plans assigned greater importance to price than did respondents from non-profit plans. The other factors deemed very or extremely important by at least 75 percent of respondents were disciplinary actions by federal or state agencies, the hospital's reputation, and its commitment to quality improvement. The quality indicators available from outcome studies were reportedly less influential in contracting decisions, cited by 27 to 70 percent of respondents. Similarly, relatively few respondents (23 to 43 percent) viewed process-of-care measures, such as thrombolytic use for AMI patients, adherence to clinical practice guidelines, and performance on HEDIS (Health Plan Employer Data and Information Set) indicators, as very or extremely important in selecting hospitals. Text comments by respondents emphasized the importance of managing costs and maintaining member satisfaction in a competitive market.

Seventy percent of respondents said that they had reviewed at least one publicly available source of information on hospital outcomes, but most of them found outcomes reports to be only minimally to moderately useful. Reported use was not significantly related to any plan characteristic. Ten plans had conducted "in-house," comparative studies of hospital performance to help inform their contracting decisions. This behavior was much more common ($p<.001$) among accredited plans (7 of 9) than among non-accredited plans (3 of 21). The specific hospital performance measures cited by our respondents included cesarean delivery rates, readmission rates, transplant success rates, sentinel event rates, use of specific treatments (i.e., beta-blockers after heart attacks), member satisfaction, resource utilization, and average length of stay. Health plan executives assigned primary responsibility for collecting and disseminating information on hospital quality to government agencies and accrediting organizations. Almost all respondents agreed that JCAHO should take the lead in public reporting, but most also felt that HCFA and the state's OSHPD should play leading roles. Most respondents did not favor a leading role for consumer advocacy groups or for the news media. A notable minority of respondents commented that it was unlikely that any "objective" information on quality would help plans or consumers to select hospitals.

We found substantial interest in hospital quality measures among managed care executives, but little evidence that health plans weigh such measures heavily in selecting hospitals. To the extent that health plans consider hospital quality, they tend to rely on measures with poor discrimination, such as accreditation, or subjective concepts, such as reputation, commitment to quality improvement, and member satisfaction. Geographic convenience and price may be the dominant considerations in hospital contracting, especially among California's for-profit HMOs. Accredited plans apparently take a more pro-active approach than non-accredited plans in evaluating their contracted hospitals. Although there is substantial interest in information on hospital quality, and confidence that such information will improve care, health plan executives are concerned about the limitations of risk-adjusted outcomes and uncomfortable weighting these

data heavily when they select network hospitals.

Time-series analysis of the impact of report cards on hospital volumes

The objectives of this study were to determine whether hospitals recognized for good or poor performance experience volume changes in the year after publication of a report card. Secondary objectives were to test whether favorable outliers attract more patients with related conditions, or from outside their usual catchment areas; and whether disadvantaged groups are unresponsive to report cards.

We conducted a time-series analysis using ordinary least squares (OLS) and autoregressive (ARIMA) models. Study subjects included all patients admitted to nonfederal hospitals designated as outliers in three coronary bypass surgery (CABG) mortality reports in New York, two acute myocardial infarction mortality (AMI) reports in California, and one post-diskectomy complications report in California. The measures of interest included observed versus expected hospital volume for topic and related conditions and procedures, by month and quarter after a report card, with and without stratification by age, race/ethnicity, insurance, and catchment area. Potential confounders included statewide prevalence, pre-report hospital volume and market share, and unrelated volume.

California hospitals labeled as having fewer AMI deaths than expected experienced significantly increased volume in the third and fourth quarters after publication, according to the OLS model (90.4 versus 76.9 patients) but not the ARIMA model. This effect was not seen for AMI-related admissions. Hospitals labeled as having more AMI deaths than expected did not experience significantly decreased volume of either AMI or AMI-related admissions, by either OLS or ARIMA. For cervical diskectomy and diskectomy-related conditions and procedures, there were also no consistent trends in hospital volume after publication of a report card. Only for lumbar diskectomy did we find consistent evidence of increased volume at hospitals lauded for their low complication rates. Although these volume differences were numerically consistent across both models and quarters, they were statistically significant only by ARIMA and never exceeded one patient per month for the average hospital.

The stratified analyses are somewhat difficult to interpret due to the large number of comparisons. Although the ARIMA and OLS results differed, the significant increase in AMI volume during the third and fourth quarters after publication of a favorable report card was most pronounced among patients over 64 years of age, with HMO or Medicare coverage, and of white race. Among HMO patients, AMI-related volume also increased significantly during each of the first four quarters after publication of a favorable AMI report card. Uninsured AMI-related patients shifted significantly toward hospitals that were rated **poorly** in the AMI report card. Volume changes after publication of a favorable report card on lumbar diskectomy were most consistent among patients less than 65 years of age, and among African-American or white patients. We found no spillover of these modest effects to diskectomy-related admissions. However, cervical diskectomy patients with indemnity insurance shifted significantly toward hospitals with high complication rates during the second through fourth quarters after publication of an unfavorable report card, while HMO patients shifted significantly away from such hospitals toward those with low complication rates, especially in the first two quarters after public release. Stratification by age and race/ethnicity did not reveal any consistent effects on cervical

diskectomy volumes.

New York hospitals labeled as having fewer CABG deaths than expected experienced significantly increased volume in the first month after publication (74.5 versus 61.1 patients). Over the first 6 months after a favorable report was published, each of these hospitals admitted an additional 24.4 CABG patients. This effect was not seen for any of three categories of CABG-related admissions (e.g., PTCA, CHF, AMI). Hospitals labeled as having more CABG deaths than expected experienced significantly decreased volume in the second month after publication (56.7 versus 67.8 patients). This effect was much less prominent for CABG-related admissions. All report card effects disappeared within 3 months after publication. In stratified analyses, CABG volume changes were generally consistent across all age groups, except those over 74 years of age. However, most of the volume changes occurred among Medicare patients and among patients of white and other race; African-American and Hispanic patient volumes were apparently not affected by designation of a hospital as a CABG mortality outlier.

Although we identified some statistically significant effects of report cards on hospital volumes, they were generally of modest size and transient duration. In California, estimates of report card effects from autoregressive models were often smaller than estimates from linear regression. The effects observed in New York were larger, but lasted for only about two months after each public release. Consistent with our hypotheses, volume shifts were largely limited to white patients in both states, and were greater among HMO patients in California than among patients with other insurance. Only HMO patients demonstrated a clear spillover effect, with increased volume for AMI-related conditions and procedures at hospitals that were lauded for low AMI mortality. Policy-makers and "smart purchasing" advocates should not expect hospital report cards to produce dramatic volume shifts. Any shifts that occur may be limited to the sociodemographic groups that are best able to understand and act upon relatively complex information.

Content analysis of media coverage of hospital report cards in California

Recent studies suggest that newspaper reports of untoward deaths are associated with a greater loss of hospital volume than poor ratings in outcome reports, but that fear of adverse media attention is a major factor motivating hospital leaders to review these reports carefully. The objectives of this study were to determine how print media coverage of outcomes reports in California and New York has changed over time and relative to HCFA's original outcomes reports in the late 1980's, whether and how this coverage has differed between states, what newspaper and market characteristics affect coverage of outcomes reports, and what information in these reports is emphasized or overlooked in the popular press.

Through newspaper clipping services, online databases, and purchase of selected newspapers, we identified 20 newspaper articles based on the 1993 CHOP report, 39 articles based on the 1996 CHOP report, and 37 articles based on the 1997 CHOP report. Similarly, we identified 59 daily newspaper articles, 5 weekly or specialized articles, and 7 editorials and commentaries based on the 1990-1997 Cardiac Surgery Reporting System (CSRS) reports in New York. Each article was independently reviewed by two authors, with attention to: (1) its length and placement, (2)

the tone or valence of the headline, (3) the source and content of each quotation, (4) the use of graphics, (5) the number of specific hospitals mentioned, and each one's risk-adjusted outcome classification, and (6) the space devoted to responses by hospital representatives, versus rebuttals by researchers or government staff.

The mean length of newspaper articles about the CHOP reports increased from 642 words in 1993 to 514 words in 1996 and 786 words in 1997. These articles have received increasingly prominent placement, with 43% appearing on page 1, 19% elsewhere in the main news section, and 24% on the front page of the local news section, in 1997. About half of the headlines following each report had neutral valence. The percentage of articles that quoted leaders of specific hospitals increased from 60% in 1993 and 41% in 1996 to 81% in 1997, while the percentage of articles that quoted hospital association representatives fell from 45% to 18% to 14%, and the percentage of articles that quoted consumer representatives fell from 20% to 5% and 11%. About 60% of articles quoted government staff in each year, while researchers were quoted most often after the 1997 report (38%). Overall, the percentage of quoted words attributed to hospital leaders rose from 32% to 42%, while the percentage of quoted words attributed to researchers rose from 5% to 15%.

The small number of hospitals labeled as mortality outliers were disproportionately mentioned and quoted in newspaper articles. These hospitals represented fewer than 13% of all acute care facilities in 1996 and 1997, but accounted for 56% of all mentioned hospitals and nearly 59% of all quoted hospitals. Three themes were raised in these quotations: (1) the outcomes study is fundamentally flawed; (2) the data are old and do not describe current outcomes; and (3) various other justifications for poor performance, including unmeasured patient characteristics.

Our content analysis of newspaper headlines and articles in California revealed a gradual shift between 1993 and 1997 from an emphasis on limitations or criticisms of the CHOP report to an emphasis on the specific ratings of local hospitals, with more prominently placed articles that included more quotes of hospital leaders and researchers. In New York, we noted a parallel shift between 1990-1991 and 1995-1997 to headlines and articles that emphasized the perceived favorable impact of the CSRS reports on CABG mortality statewide. These findings suggest that newspaper reporters are becoming more sophisticated readers of hospital outcome studies, but remain disproportionately attentive to which hospitals are labeled as bad outliers.

FINAL REPORT R29 HS08574

INTRODUCTION AND AIMS

In the late 1980s, consumers and purchasers of health care started demanding more information about the performance of health care organizations.[1] Federal agencies, state legislatures, and private coalitions responded to this changing environment by collecting and disseminating data on risk adjusted hospital outcomes.[2] The US Health Care Financing Administration (now the Center for Medicare and Medicaid Services) published the first such report in 1987,[3] but was soon followed by the New York State Department of Health[4] and the Pennsylvania Health Care Cost Containment Council.[5] These pioneering programs subsequently expanded to report on more procedures[6] using more sophisticated methods.[7] More recently, Iowa,[8] Missouri,[9] Florida,[10] Wisconsin,[11] Virginia,[12] and California[13] joined this movement. In Cleveland, business interests and providers joined in a unique but relatively short-lived partnership to gather and report detailed measures of clinical outcomes and patient satisfaction.[14] With the advent of the *US News* ratings of "America's Best Hospitals" and the Web site http://www.healthgrades.com , even the private sector has become involved in this effort.

These report cards are intended to provide information about quality of care that consumers, employers, and health plans can use to make better decisions. A lot of time, money, and effort are being expended on generating and disseminating them. But what are we getting from this effort? Little is known about how this information is actually used in the health care marketplace.[15] In theory, hospital outcome studies could reduce morbidity and mortality either by motivating providers to improve processes of care, or by motivating consumers and health plans to select hospitals with better risk-adjusted outcomes. Our study was designed to assess both of these mechanisms for the potential effectiveness of hospital report cards.

Accordingly, the specific aims in our original proposal were:

1. To evaluate the impact of publicly released outcomes data in California and New York on hospital or surgeon volumes and long-distance referral patterns.
2. To evaluate whether the impact of publicly released outcomes data on hospital volumes and long-distance referral patterns varies by age, gender, race/ethnicity, expected source of payment, market competitiveness, diagnostic category (e.g., related or unrelated to previous outcome studies), or urgency of admission.
3. To evaluate the impact of publicly released outcomes data in California and New York on risk adjusted outcomes, both statewide and hospital-specific.
4. To evaluate the impact of publicly released outcomes data in California and New York on coding and reporting of clinical risk factors and postoperative complications.
5. To explore how hospitals in California and New York have assimilated and responded to outcomes data, and to solicit their suggestions on the optimal methods, content and format of future reports.
6. To explore how other interest groups in California and New York, such as large health plans, have assimilated and responded to outcomes data, and to solicit their suggestions on the optimal methods, content, and format of future reports.

7. To apply the results of Specific Aims 1-6 in designing and implementing hospital outcome studies that will be more responsive to the needs of consumers, providers, and purchasers of health care.

As a result of the 15% budget cut mandated by the Agency for Health Care Policy and Research (now the Agency for Healthcare Research and Quality) in Year 02, we dropped the proposed analyses using New York's Coronary Angioplasty Reporting System (CARS). Concerned about whether inter-hospital differences in postpartum maternal complication and readmission rates were actually due to socioeconomic factors and practice variation, the California Office of Statewide Health Planning and Development (OSHPD) withheld release of provider-specific obstetric outcomes data until a validation study could be completed in the late 1990s. Therefore, each of our original Specific Aims was modified to focus on OSHPD's public reports on acute myocardial infarction (AMI) mortality and post-diskectomy complications in California, and the Department of Health's public reports on coronary artery bypass surgery (CABG) mortality in New York.

Finally, Specific Aim 7 required modification because two key problems that we identified by surveying hospitals, physicians, and health plans could not be addressed by the relevant state agencies. Specifically, hospital administrators and quality improvement leaders were very concerned about:

The 3 year time lag between any episode of care (i.e., AMI, CABG) and the release of a public report that included that episode of care.

The lack of any linkage between outcome and process data, which would enable hospitals and physicians to identify deficiencies in care and implement process improvements.

The 1997 California Hospital Outcomes Project (CHOP) report did incorporate several recommended improvements related to formatting and presentation:

1. The report was restructured into five separate volumes. Most of the technical information about study methods was removed from the User's Guide and placed in a separate Technical Guide. Hospital response letters were also removed from the User's Guide.
2. The User's Guide was enhanced by the addition of both an overview and a graphic display directly comparing the risk-adjusted mortality rates of all hospitals, by county.
3. Highlights of OSHPD's Acute Myocardial Infarction Validation Study were added to the Consumer Guide. Relative to other facilities, hospitals with low risk-adjusted mortality were more likely to give AMI patients aspirin within 6 hours of arrival, and heparin within 24 hours of arrival, and were more likely to perform coronary angioplasty or bypass surgery within the first 24 hours.[16]
4. ICD-9-CM coding became more consistent. The number of hospitals excluded because of data problems decreased from 27 (with 2,127 patients) in 1991 to 13 (with 494 patients) in 1993.[17]

However, these changes were not substantial enough to justify another round of surveys of hospital administrators and quality improvement leaders. It was beyond OSHPD's statutory authority to collect or publish process data, and our exploratory efforts to link OSHPD's outcomes data with process data from California Medical Review, Inc. (the Peer Review

Organization for California) were unsuccessful. Due to statutory and fiscal constraints, OSHPD was unable to significantly compress the timetable for collecting, cleaning, analyzing, and publishing risk-adjusted outcome data.

We therefore chose to continue our analyses of the impact of hospital report cards in New York and California by focusing in detail on media coverage of these reports. Through our surveys of hospital excecutives, physicians, and managed care executives, we came to the conclusion that the media play a central role in disseminating hospital performance information. The responses of consumers, providers, and purchasers are likely to be influenced by how the media carry out this dissemination. Accordingly, our new specific aim was:

> To characterize how the news media, especially the print media, have covered hospital performance reports in California and New York, with particular attention to changes over time, differences between the states, and differences from the coverage of HCFA mortality reports a decade ago.

PRIOR RESEARCH

Hospital outcome studies are based on two fundamental hypotheses about health care quality and marketplace behavior:

1. *Risk-adjusted outcome studies provide reliable and valid measures of quality of care.* In other words, inter-hospital variation in outcomes due to chance or severity of illness can be estimated and removed. The remaining variation can be attributed to quality of care.
2. *Risk-adjusted outcome studies influence consumer, purchaser/payer, or provider behavior and thereby have a favorable impact on overall patient outcomes.*

If the first hypothesis is false, then hospital outcome studies are fundamentally without value. If the second hypothesis is false, then hospital outcome studies have potential value but are unlikely to be cost-effective. In either case, time and effort are being wasted and a basic reassessment of our approach is necessary.

A few studies from the early 1990s addressed the first hypothesis. As part of the Prospective Payment System Quality of Care Study,[18] RAND investigators demonstrated that patients with four medical conditions (congestive heart failure, AMI, pneumonia, and stroke) who experienced good process of care were less likely to die within 30 days of admission than those who experienced poor process of care, after using clinical variables to adjust for sickness at admission. This relationship was observed for three of the five explicit measures of process (physician cognitive, nurse cognitive, and technical diagnostic) and for an implicit measure based on physician review.[19] Hannan et al[20] reviewed the medical records of cardiac surgery patients who died in New York hospitals with low or high risk-adjusted mortality. Eighteen of the 40 deaths reviewed in high outlier hospitals had quality-of-care problems, compared to only one of 23 deaths in low outlier hospitals. Thomas et al[21] and Hartz et al[22] demonstrated significant but weak correlations between peer review organization "failure" rates and risk-adjusted mortality for certain conditions. In the Department of Veterans' Affairs National VA Surgical Risk Study, experienced site visitors were able to assign surgical services to the correct risk-adjusted outcome category 85% of the time,[23] but surgeons doing structured implicit review of medical records rated quality very similarly at low-mortality and high-mortality outlier hospitals.[24]

Despite substantial evidence supporting the validity of risk-adjusted outcomes, especially mortality, as indicators of quality of care, serious questions remain. First, several simulations have suggested that risk-adjusted outcome studies have poor sensitivity and relatively low predictive value for detecting hospitals with true quality problems.[25,26] These problems might be remedied by focusing on very high-frequency or high-mortality conditions, or by collecting more cases from each hospital. However, these potential solutions limit the utility or increase the cost of risk-adjusted outcomes reports. Even more disturbing is the fact that different methods of severity adjustment have been shown to yield quite different classifications of hospital performance.[27] Methods based solely on administrative data are particularly suspect, because they often fail to distinguish comorbidities from iatrogenic complications.

The second fundamental hypothesis was largely untested when we began our work. In two national surveys, hospital executives were extremely negative about the value of HCFA's release

of Medicare mortality data. Seventy percent of 195 responding hospitals rated the report's usefulness "to hospitals in improving quality" as poor; 85% rated its usefulness "to consumers in hospitals' area" as poor.[28] These ratings were unrelated to whether the hospital had lower than expected, expected, or higher than expected risk-adjusted mortality. In the other survey,[29] 67% of 250 responding hospitals disagreed with the assertion that "disclosure of such data...helps consumers make rational decisions about health care providers."

Indeed, the fear and skepticism of hospital administrators was understandable, given how HCFA's results were characterized in the lay press. According to a content analysis of newspaper coverage of HCFA's 1986 Medicare mortality release,[30] 41% of the articles reviewed carried negative headlines (emphasizing high-mortality hospitals) despite the fact that nearly 95% of hospitals had death rates within the expected range. Not surprisingly, 69% of the articles cited hospital spokespersons who blamed HCFA for their facilities' poor showing. Although hospitals with favorable results would be expected to support the study, their pride was apparently eclipsed by anxiety about what the next report will show. Given such negative reactions, it is hardly surprising that HCFA abandoned its effort to analyze and publish risk-adjusted outcomes data.

Consumers and purchasers of health care may be more receptive than providers to hospital outcome studies. There has long been evidence of public interest in ranking hospital performance.[31] In addition, conditional logit models of hospital choice suggest that consumers prefer hospitals with low risk-adjusted death rates, especially for AMI care, CABG surgery,[32] and high-risk obstetric care.[33] Studies of the volume-outcome relationship using simultaneous equation models also suggested that patients are "selectively referred" to hospitals with good outcomes.[34,35,36] Yet the early evidence strongly refuted any specific influence of report cards on consumer behavior. Vladek et al[37] found that HCFA's designation of 14 high-mortality and 9 low-mortality hospitals in New York City had no effect on subsequent occupancy rates. Although a few major employers reported using outcomes data to steer their employees toward particular hospitals,[38,39] this is not a widespread practice even today.

These studies suggested to us that providers may be too defensive to use outcomes data constructively, perhaps because the lay press has focused undue attention on a small number of high-mortality outliers. Purchasers may be more concerned about lowering costs than about improving quality. And consumer behavior may be so heavily influenced by tradition, convenience, and physician preference that outcomes data are superfluous. As of 1994, when we began our work, it appeared that outcome studies generate headlines for local newspapers, but are quickly forgotten and subsequently ignored.

We suspected that reporting efforts in New York and California might provide more promising prototypes for further evaluation and replication. In New York, the Department of Health developed a Cardiac Surgery Reporting System (CSRS) and started collecting, analyzing, and publishing outcomes data for CABG surgery in 1989.[4] This unique data set now includes over 40 clinical risk factors, such as cardiac ejection fraction and body mass index, that were identified by cardiac surgeons and cardiologists as potential predictors of perioperative mortality. Detailed data on risk factors and outcomes are shared with each participating hospital, in user-friendly formats. As a result of a lawsuit by a major newspaper under New York's Freedom of

Information law, surgeon-specific as well as hospital-specific outcomes data have been released to the public since 1991. For these reasons, we thought that New York's cardiac surgery studies might be more influential than HCFA's reports, and the results might be more difficult for providers to condemn or ignore. An analogous system for monitoring PTCA outcomes (the Coronary Angioplasty Reporting System, or CARS) was established in 1991, but concerns about the reporting of major complications has interfered with regular publication of these reports.[40]

Shortly before our studies began, Hannan et al[41] reported that statewide CABG volume rose 31% from 12,269 in 1989 to 16,028 in 1992 while actual mortality fell 21% from 3.5% to 2.8% (risk-adjusted mortality rate fell 41%, but more thorough reporting of risk factors may explain part of this trend). However, he found no clear correlation between risk-adjusted mortality and subsequent changes in hospital volume. Hospitals labeled as high-mortality outliers based on 1989 data gained as much volume in 1991 (up 27%) as hospitals labeled as low-mortality outliers (up 26%) or nonoutliers (up 20%). Anecdotal evidence suggested that a few hospitals responded constructively to CSRS reports. For example, one hospital that was labeled as a high-mortality outlier established a collaborative review group with representatives from cardiology, surgery, anesthesiology, nursing, and administration to study all CABG deaths.[42] This group found that the excess mortality was limited to high-risk patients with AMIs, who were less likely to receive aortic balloon pumps than similar patients elsewhere. When surgeons at this hospital began using balloon pumps more often, their death rate decreased. A public hospital in Buffalo hired a new surgeon after it received a poor rating.[43] Another hospital credited CSRS with stimulating major internal changes: two surgeons lost operating privileges, a physician's assistant program was initiated, and support personnel were replaced.[44] While the leaders of CSRS have argued that "the information supplied to hospitals prompted them to make process and personnel changes"[45] to reduce risk-adjusted mortality, others have been critical of both their methods[46] and their conclusion that outcomes reporting led to the observed drop in mortality.[47]

In California, the state legislature mandated that the Office of Statewide Health Planning and Development (OSHPD) annually generate and disseminate detailed information about risk-adjusted outcomes to all hospitals "as an aid to internal quality assurance." The legislature also mandated that less detailed reports be provided to the general public. Under contract with OSHPD, researchers at three University of California campuses developed sophisticated, condition-specific predictive models based on hospital discharge data to estimate the probability of an adverse outcome for patients with various conditions and procedures, including AMI, pneumonia, hip fracture, cervical and lumbar diskectomy, and vaginal and cesarean delivery. The first report of the California Hospital Outcomes Project (CHOP), on AMI mortality and postoperative diskectomy complications, was released to hospitals in June 1993 and to the public five months later.[13] Letters of response were received from or on behalf of 155 hospitals; most indicated that they had reviewed at least a sample of their own cases. These letters suggested that many facilities were using the data as part of their continuous quality improvement efforts. Although reports on other conditions and procedures have been released to hospitals since 1993, only the reports on AMI mortality have been publicly disseminated.

There are several important similarities and differences between New York's CSRS/CARS and California's CHOP that we thought would make our research especially instructive. Both

programs are focused at the local level, provide patient-specific outcomes data to hospitals, release relatively user-friendly data to the public, and involve substantial provider education. In this way, providers obtain the data they need to evaluate outcomes, and face strong incentives to improve substandard performance. In addition, both California and New York are populous, heavily urbanized states that have devoted significant resources to hospital outcomes studies and naturally act as trend-setters for the nation. However, there are two important differences between the states. First, New York's CSRS reports are based on detailed clinical data, whereas California's reports are based on administrative data that have less credibility with providers. Second, New York has a more regulated health care environment in which there is greater emphasis on regionalized services. By contrast, California's hospitals face greater competitive pressures because of lower occupancy rates and greater HMO market penetration.

Since we began our research, several other researchers have explored the impact of hospital outcome studies. Some of their findings prompted us to change our plans in Years 04-05. Mennemeyer, Morrisey, and Howard found that newspaper reports of untoward deaths were associated with a 9% reduction in annual Medicare discharges, although only five hospitals were affected.[48] Mukamel and Mushlin reported that the 1990 publication of CSRS risk-adjusted mortality data in the *New York Times* led to a clinically but not statistically significant decrease in the market share growth rate of hospitals with relatively high mortality (e.g., "the median hospital would have lost 8.4 of its 166 CABG procedures in 1990 for each percentage point increase in its mortality rate").[49] Our own survey of hospital quality improvement managers in California suggested that fear of adverse media attention was a major factor motivating hospital leaders to review the CHOP report carefully and disseminate it to specific units (e.g., marketing, public affairs, board of trustees). These findings led us to explore the role of the media in disseminating and interpreting risk-adjusted hospital outcomes data.

Several other studies reported since 1995 have estimated the impact of hospital report cards on hospital, physician, consumer, purchaser, and health plan behavior. In addition, other reporting schemes have emerged, and have been subjected to qualitative or quantitative assessment.[50] These studies help us place our own findings in a broader context. They are discussed in each of the next five chapters, and summarized in the concluding chapter.

OVERVIEW OF STUDIES

Five separate but related studies were conducted as part of this project. The detailed methods and results of these studies are presented in the following five chapters, and summarized in the Executive Summary. As a bloc, these studies were designed to achieve a better understanding of the outcomes of hospital outcomes studies in California and New York by evaluating the impact of these studies on consumer, provider, and health plan behavior. Using both quantitative and qualitative methods, we sought to explore how information about quality of care can be made more useful to all stakeholders in an era of managed competition and health care reform.

Survey of hospital administrators in California and New York.
This study was published as:
Romano PS, Rainwater JA, Antonius D. Grading the graders: How hospitals in California and New York perceive and interpret their report cards. *Medical Care* 1999;37:295-305.

Survey of hospital quality improvement leaders in California.
This study was published as:
Rainwater JA, Romano PS, Antonius DM. The California Hospital Outcomes Project: How useful is California's report card for quality improvement? *Joint Commission Journal on Quality Improvement* 1998;24:31-39.

Survey of managed care executives in California.
This study is currently under review for publication:
Rainwater JA, Romano PS. Are California managed care organizations using outcomes data in contracting with hospitals?
It was published in abstract form as:
Romano PS, Rainwater JA. Are California managed care organizations using hospital outcome studies and data? *Journal of General Internal Medicine*; 2000;15(April, Suppl 1):142.

Time-series analysis of the impact of report cards on hospital volumes in California and New York.
This study is currently under review for publication:
Romano PS, Zhou H. Do Well Publicized Risk-Adjusted Outcomes Reports Affect Hospital Volume?

Content analysis of media coverage of hospital report cards in California and New York.
This study is currently under review for publication:
Rainwater JA, Romano PS. A content analysis of changes in media coverage of hospital report cards in California and New York.

Please note that copyright to these manuscripts is held by the journals identified above.

These results have been formally presented at:

The 15th Annual Meeting of the Association for Health Services Research; June 21-23, 1998; Washington, DC.

The 17th Annual Meeting of the Association for Health Services Research; June 21-23, 2000; Los Angeles, CA.

The 2002 Annual Meeting of the Academy for Health Services Research and Health Policy; June 23-25, 2002; Washington, DC.

The 23rd Annual Meeting of the Society of General Internal Medicine; May 4-6, 2000; Boston, MA.

The California Office of Statewide Health Planning and Development and the California Health Policy Data Advisory Commission; April 3, 2000; Sacramento, CA.

The 2nd Scientific Forum on Quality of Care and Outcomes Research in Cardiovascular Disease and Stroke, sponsored by the American Heart Association and the American College of Cardiology; April 11, 2000; Washington, DC.

Local presentations at the University of California, Davis and the UK's National Primary Care Research and Development Centre at the University of Manchester.

In addition, the principal investigator authored, coauthored, or contributed to the following editorials, reviews, and articles with partial support from this grant:

Jollis JG, Romano PS. Pennsylvania's 'Focus on Heart Attack': Grading the scorecard. *New England Journal of Medicine* 1998;338:983-987.

Romano PS. Physician profiling: Words of caution. In: Goldfield N, Boland P, eds. *Physician Profiling and Risk Adjustment*. Gaithersburg, MD: Aspen Publishers, 1999; pp. 47-56.

Hannan EL, Stone CC, Biddle TL, DeBuono BA. Public release of cardiac surgery outcomes data in New York: What do New York state cardiologists think of it? *American Heart Journal* 1997; 134:1120-1128. (acknowledgment)

STUDY 1: Survey of hospital administrators in California and New York

Despite the controversy surrounding hospital report cards,[51] little is known about their value and impact. Three surveys of hospital leaders have been reported, but the first was based on widely criticized mortality data from the Health Care Financing Administration (HCFA)[28]and the second was limited to 17 public hospitals in California.[52] The most recent survey, involving 21 Pennsylvania and 8 New Jersey hospitals, suggested that Pennsylvania's *Consumer Guide to Coronary Artery Bypass Graft Surgery* has been used to recruit heart surgeons, monitor their performance, and stimulate improved cost-cutting and record-keeping.[53] A survey of cardiologists and cardiothoracic surgeons found considerable skepticism toward the same publication;[54] a separate survey of New York cardiologists was limited by a poor response rate (36%).[55] No previous study compared the perceived value of multiple hospital report cards, and no study was large enough to explore how hospital characteristics affect attitudes toward and understanding of outcomes reports.

We undertook a survey of hospital leaders in the 2 most populous states, which publish very different hospital report cards. The California Hospital Outcomes Project (CHOP) began in 1991, with the enactment of a law requiring the Office of Statewide Health Planning and Development (OSHPD) to produce annual reports on risk-adjusted outcomes at acute care hospitals, using ICD-9-CM coded discharge abstracts. The first report, released in 1993, classified hospital mortality rates for acute myocardial infarction (AMI) and complication rates for cervical and lumbar diskectomy into 2 categories ("better" or "not better" than expected).[13] The second report, released in May 1996 after a major validation study, classified hospital performance for AMI as "better," "worse," or "neither better nor worse" than expected[56,57] New York's CSRS began in 1989 with the creation of a special clinical data system for cardiac surgery. A Cardiac Advisory Committee was established to identify and define each data element. Hospital-specific, risk-adjusted mortality rates and 3-category ratings have been released every 12-18 months since December 1990. The first report from the Coronary Angioplasty Reporting System, focusing on risk-adjusted mortality, was released in October 1996 after extensive debates about data quality.[6]

The efforts in California and New York represent prototypes of two different approaches to reporting risk-adjusted outcomes. Florida has followed California's example in using administrative data because of its low cost and universal availability. Pennsylvania and Cleveland have followed New York's example in building new data systems with detailed clinical data to permit better risk-adjustment.[58] California's program solicits public or provider input at almost every stage, from selecting conditions for study to reviewing draft reports. New York's program includes less public input, but substantial professional input. If these methodologic differences have led to differences in how hospital report cards are perceived and interpreted, there may be important lessons for other states.

METHODS

All 398 hospitals listed in the 1996 CHOP report, and all 31 hospitals listed in the 1996 CSRS

report, were eligible for study. The name and mailing address of each California hospital's chief executive were obtained from the California Healthcare Association and updated based on the response letters submitted for publication with the 1996 CHOP report. New York hospitals were contacted by telephone to identify their chief executives and to confirm their mailing addresses. Nineteen California hospitals were not mailed surveys because they had closed, merged with other facilities, or stopped providing acute care. Five additional California hospitals were mailed surveys but were unable to respond, for one of the same reasons. The final sampling frame included 374 hospitals in California and 31 hospitals in New York.

Our 4-page printed questionnaire included 5 major components. First, respondents were asked to identify all of their hospitals' departments that received at least part of the 1996 report, or participated in discussions of it. The California questionnaire included a list of 10 options plus an open-ended "other" category; this list was abbreviated to 6 options plus "other" because of space constraints in the New York questionnaire. Second, respondents were asked to rate various aspects of the 1996 report, including its usefulness to consumers and hospitals (for improving quality), accuracy in describing hospital performance, completeness of case mix adjustment, ease of interpretation, and manner of release. To facilitate comparison with an earlier survey of hospital leaders about HCFA mortality data, we used the same questions and response options (excellent, very good, good, fair, poor).[28] The New York survey repeated these questions for both the CSRS (coronary bypass graft surgery) and coronary angioplasty reports. Because the responses were similar, only the findings related to CSRS are presented.

Third, respondents were asked to indicate their agreement or disagreement, using a 4-point Likert scale, with a series of statements about their states' report. Three (California) or 4 (New York) of these statements were correct or incorrect descriptions of the risk-adjustment method used in the 1996 report, allowing us to assess knowledge. Fourth, respondents were asked to indicate, using a 4-point Likert scale, whether the California Hospital Outcomes Project or New York's Cardiac Surgery Reporting System "is a better system for assessing quality of care" than HCFA mortality studies, administrative data systems such as All Patient Refined Diagnosis Related Groups (APR-DRGs), and clinical data systems such as APACHE and MedisGroups. Finally, respondents were invited to identify themselves and to offer suggestions for future hospital outcomes reports.

Questionnaires were mailed within 4 months of the report publication date in each state. Our cover letter emphasized that the survey was funded by the Agency for Health Care Policy and Research (now the Agency for Healthcare Research and Quality) and that individual responses would be kept confidential. Because of the first author's role as a contractor to the OSHPD, these cover letters were signed by the second author. Self-addressed stamped envelopes were provided, and nonrespondents received 2 additional mailings. With the last mailing, recipients were asked to call or write and explain why they did not wish to complete the questionnaire.

We evaluated several hospital characteristics as correlates of how hospitals used and rated the California outcomes report. Hospital ownership was categorized as for-profit, nonprofit corporate, nonprofit church, health maintenance organization (Kaiser Permanente), University of California, or public/district. Hospital size was categorized as small (<101 beds), medium (101-

209 beds), or large (>209 beds), based on the number of licensed general acute care beds on December 31, 1994. Occupancy was categorized as low (<35.4%), medium (35.4-48.0%), or high (>48.0%), based on the ratio of general acute care patient days to licensed bed days in 1994. Medicaid share was categorized as low (<10.1%), medium (10.1-26.4%), or high (>26.4%), based on the proportion of patients whose primary payer was Medicaid in 1994. We constructed these tertiles to include equal numbers of hospitals with AMI patients and valid data from the OSHPD's Annual Report of Hospitals.[59] AMI volume in 1991-1993 was categorized as low (<87 patients), medium (87-206 patients), or high (>206 patients).

Hospitals with unusually low or high risk-adjusted mortality, using the official $p<.01$ threshold, were identified from the 1996 CHOP report. An alternative threshold of $p<.05$ was tested, but was abandoned because marginal outliers (i.e., $.01<p<.05$) were never publicly identified and responded similarly as nonoutliers to all key questions. Hospitals sponsoring an accredited residency program in either internal medicine or family practice, and reporting a "major" or "graduate" medical school affiliation, were classified as teaching hospitals.[60] Hospitals with cardiology training programs were found not to differ from other teaching hospitals with regard to the variables of interest. Hospitals licensed to perform cardiovascular surgery were also identified.[59]

Three composite scores were created using multiple survey questions related to the same concept. An overall quality rating was computed by averaging ordinal scores on the 6 questions adapted from Berwick and Wald.[28] A usefulness score was computed based on agreement or disagreement with 4 statements regarding possible uses of the state's outcomes data: improving the quality of care, improving the quality of medical records coding, negotiating with health plans, and marketing or public relations. These 2 scores demonstrated excellent reliability, with Cronbach's alpha of 0.75 and 0.86 respectively. A knowledge score was computed based on agreement or disagreement with factual statements about whether the state's risk-adjustment method accounts for demographic factors, chronic medical conditions, acute physiologic problems, and intraoperative factors (in New York). To compare data across states, the California knowledge score was rescaled by multiplying the number of correct answers by 1.33.

All data were entered using EpiInfo; most analyses were performed using SPSS/PC+. A few surveys with missing or uninterpretable responses were excluded on an item-specific basis. We first identified the hospital characteristics associated with nonresponse in California. We then performed univariate analyses of each survey question, followed by bivariate analyses with each of the hospital characteristics defined above. For "yes/no" survey questions, the reported p values are based on continuity-adjusted chi-square tests for dichotomous predictors, Pearson chi-square tests for other categorical predictors, and linear association chi-square tests for ordinal predictors. For scaled survey questions, the reported p values are based on analysis of variance, except that the Kruskal-Wallis or Wilcoxon rank sum test was used when the Levene test for homogeneity of variance was significant ($p<.10$). When a multi-sample test was statistically significant, all possible 2-way comparisons were evaluated using the Wilcoxon test. Because of the exploratory nature of this study, all p values less than .05 are noted.

RESULTS

Of the 374 eligible hospitals in California, 249 (66.6%) provided usable responses. An additional 25 hospitals indicated that they could not provide usable responses, either because they had never received the report or because they were not familiar with its contents. Therefore, the total response rate was 73.3%. For the 3 hospitals that provided 2 responses, only the response of the chief executive officer was analyzed. Of the 31 eligible hospitals in New York, 27 (87.1%) provided usable responses. Table 1 shows the characteristics of responding and nonresponding hospitals in California. The response rate was lower at for-profit hospitals than at nonprofit corporate, nonprofit church, and public/district hospitals. The response rate was positively associated with AMI volume.

Hospital leaders generally agreed that the conditions, procedures, and outcomes studied in California and New York are important. Specifically, at least 74% of California respondents acknowledged the importance of studying mortality after AMI, in-hospital complications after diskectomy, and maternal readmissions after delivery (the 3 conditions and procedures evaluated to date). Strong agreement with these selections was reported by 28%, 11%, and 22%, respectively. At least 81% of New York respondents acknowledged the importance of studying mortality after coronary bypass surgery, mortality and complications after coronary angioplasty, and mortality after other types of cardiac surgery (the procedures for which data are collected).

Table 2 shows that in both states, the hospital outcomes report was disseminated widely within hospitals. A mean of 5.0 units in California hospitals and 5.2 units in New York hospitals either received a copy of the report or participated in discussions of its contents. These means are not directly comparable because the New York questionnaire listed fewer response options. Over 90% of hospitals in both states shared or discussed the outcomes report with high-level administrators and quality improvement staff. Medical staff leaders, boards of directors, and public relations or marketing staff were significantly more likely to receive or discuss the report in New York than in California.
Table 3 shows how hospital leaders rated the 1996 state outcomes reports along 6 dimensions. Data from Berwick and Wald's survey regarding HCFA mortality data are displayed, but cannot be compared statistically.[28] The New York report received significantly better ratings than the California report in its usefulness for improving quality of care, accuracy in describing hospital performance, and ease of interpretation. The 2 reports were not rated significantly different in the completeness of their risk-adjustment models, usefulness to consumers, and method of release. The overall mean rating, on a scale of 0-4 where 0 represents "poor" and 4 represents "excellent," was highest (1.8) for New York's CSRS report, intermediate (1.4) for California's CHOP report, and lowest (0.6) for earlier HCFA reports (NY versus CA, $p=.0074$).

When hospital leaders were asked to compare their own state's outcomes reporting system with other products, 73% in California and 93% in New York agreed that the state's system was better than HCFA mortality reports. About 50% in California and 81% in New York agreed that the state's system was better than systems using administrative data on broad patient cohorts, such as APR-DRGs. Only 24% in California and 50% in New York agreed that the state's system was better than systems using detailed clinical data on broad patient cohorts, such as APACHE and MedisGroups. All of these differences between states were statistically significant ($p<.05$).

Table 4 shows that hospital leaders in California and New York found their state's outcomes reports to be moderately useful for improving the quality of care and (in California) the quality of ICD-9-CM coding. About one-third of California respondents, and over half of New York respondents, found the reports to be useful for negotiating with health plans or marketing their programs. New York respondents rated their report as significantly more useful in each respect, and overall, than did California respondents. California respondents who rated the quality of the CHOP report as fair or poor did not differ from other respondents in dissemination, but described the report as considerably less useful in each respect, and overall (not shown).

Hospital leaders had limited understanding of the risk-adjustment methods used in their state's report cards. Only 25% of California respondents and 48% of New York respondents were aware that the reports adjusted for chronic diseases (e.g., diabetes); this difference was statistically significant ($p=.027$). About 38% of California respondents and 88% of New York respondents were aware that the reports adjusted for demographic factors (e.g., age); this difference was highly significant ($p<.001$). Only 60% and 54% of New York respondents were aware that the CSRS report adjusted for physiologic (e.g., shock) and intraoperative factors (e.g., number of arteries bypassed), respectively. About 86% of California respondents realized that the CHOP report did not adjust for physiologic factors. The overall mean knowledge score, on a scale of 0-4 where 4 represents correct responses to all questions, was 2.0 in California and 2.4 in New York ($p=.075$).

Using only data from California, we examined factors associated with hospital leaders' dissemination of the CHOP report, ratings of its quality, perceptions of its usefulness, and understanding of its risk-adjustment methods (Table 5). Fewer hospital units reviewed or discussed the report at for-profit hospitals than at nonprofit corporate hospitals. This difference resulted from less frequent dissemination to medical (53% versus 82%) and nursing (59% versus 75%) staff leaders, and board members (43% versus 59%), at for-profit hospitals. The number of hospital units that reviewed or discussed the report was positively associated with hospital size and AMI volume. These findings resulted from less frequent dissemination to marketing staff and medical staff leaders at both small and low-volume hospitals, and less frequent dissemination to board members at low-volume hospitals. Hospitals labeled as high-mortality outliers disseminated the report to more units than those not labeled as outliers. This difference resulted from more frequent dissemination to nursing staff leaders (93% versus 68%), board members (93% versus 47%), legal counsel (27% versus 10%), and marketing (87% versus 33%) and medical records staff (80% versus 63%) at high-mortality hospitals. Finally, more hospital units reviewed or discussed the report at facilities with low percentages of Medicaid patients than at those with high percentages. This difference resulted from more frequent dissemination to quality improvement (96% versus 86%) and medical records staff (76% versus 54%) at hospitals with low Medicaid shares.

Leaders at hospitals labeled as low-mortality outliers rated the quality of the CHOP report significantly better than leaders at either nonoutlier or high-mortality outlier hospitals. This difference was significant ($p<.05$) for each item in Table 3, but especially so ($p<.005$) for "accuracy as a descriptor of performance," "completeness of case-mix adjustment," and "way report was released." Respondents' overall ratings of quality were unrelated to the ownership,

size, occupancy, teaching status, AMI volume, surgical capability, and payer mix of their facilities.

Leaders at for-profit hospitals found the CHOP report more useful than leaders at nonprofit corporate and University of California hospitals. This difference was significant ($p<.05$) for all items in Table 4 except improving coding. Leaders at hospitals labeled as low-mortality outliers found the report more useful than leaders at nonoutlier and high-mortality outlier hospitals. This difference was significant only for improving quality of care and marketing. Respondents' overall ratings of usefulness were unrelated to the size, occupancy, teaching status, AMI volume, surgical capability, and payer mix of their facilities.

Knowledge about the CHOP risk-adjustment methods was positively related to leadership of a nonprofit corporate or church hospital (versus a for-profit hospital), large hospital, high or medium-volume hospital, low-mortality outlier hospital, or hospital with an intermediate Medicaid share (Table 5). Respondents' apparent knowledge was unrelated to the occupancy, teaching status, and surgical capability of their facilities.

DISCUSSION

In a systematic survey of hospital leaders in California and New York, we found that the hospital report cards produced by state agencies are viewed as focusing on important conditions, procedures, and outcomes. These reports are widely disseminated within hospitals, and hospital leaders rate them better than the reports produced by the HCFA in the late 1980s. However, their overall ratings were still only fair-to-good. New York's CSRS report, which has a longer track record and is based on detailed clinical data, was rated better than California's CHOP report, which is based on administrative data. This difference was also apparent when respondents were asked to compare their state's outcomes reporting system with commercially available systems such as APR-DRGs, APACHE, and MedisGroups. Hospital leaders demonstrated limited understanding of risk-adjustment; only 14% of California respondents and 38% of New York respondents correctly answered all questions about the methodology.

These findings are generally consistent with prior studies. Specifically, we confirmed Berwick and Wald's finding that hospital leaders are skeptical about risk-adjusted outcomes data.[28] This is even true, to a lesser degree, in a state where risk-adjusted mortality rates are estimated using a detailed clinical data set that was established under the guidance of cardiologists and cardiac surgeons with extensive peer review. This skepticism was most apparent when hospital leaders were asked to evaluate the usefulness of the state's outcomes report to consumers and, in California, its usefulness in improving quality of care. Our results differed from Berwick and Wald's in that high-mortality outliers rated the CHOP report significantly worse than other hospitals. This difference may reflect the wider distribution of quality ratings in our study. One other survey was not directly comparable but revealed substantial concern that the HCFA's mortality reports "unfairly damaged the reputations of some hospitals" and were not helpful to consumers.[29]

Our results are also consistent with Luce et al's finding that hospital leaders view publicly

reported risk-adjusted outcomes data as having limited usefulness.[52] Although many hospital units reviewed or discussed the report, they were unenthusiastic about its quality and value. Only 8% of hospital leaders in California and 22% in New York rated the report "very good" or "excellent" in facilitating quality improvement. A follow-up telephone survey of 39 quality improvement managers at California hospitals confirmed that the two key concerns are the excessive delay before outcomes data are released and the lack of specific information about modifiable processes of care.[61]

We used California data to evaluate the association between hospital characteristics and their leaders' internal dissemination of the CHOP report, ratings of its quality, perceptions of its usefulness, and understanding of its risk-adjustment methods. The number of hospital units that reviewed or discussed the report was positively associated with hospital size, AMI volume, and being labeled as a high-mortality outlier. Respondents at large hospitals appeared more knowledgeable than those at small hospitals; respondents at high-volume and medium-volume hospitals appeared more knowledgeable than those at low-volume hospitals. These findings may reflect the ability of large and high-volume hospitals to develop specialized units to focus on cardiac care or quality improvement. Personnel in these specialized units have the time and resources to study risk-adjusted outcomes data.

Leaders at hospitals labeled as low-mortality outliers rated the overall quality of the CHOP report significantly better, found it more useful, and better understood its risk-adjustment methods, than leaders at either nonoutlier or high-mortality outlier hospitals. Leaders at for-profit hospitals found the report more useful (especially for marketing and negotiating contracts), despite having less understanding of its methods and disseminating it to fewer units, than leaders at nonprofit corporate hospitals. These findings may reflect greater organizational efficiency, competitiveness, and external focus at for-profit hospitals.

The response rate to our survey was excellent in both California (73.3%) and New York (87.1%). Although hospitals with high AMI volumes were more likely to respond than those with low volumes, our response rate in California was at least 54% in every identifiable stratum. We suspect that hospital leaders with especially negative or positive views of their state's report were more likely to respond than those with neutral or poorly formed views. If this hypothesis is correct, our survey responses might be skewed toward the extremes, but the effect on mean score values is uncertain.

The major limitation of our findings is that they represent case studies which may not generalize to other states and communities. The attitudes and opinions of hospital leaders are likely to reflect local factors, such as when a report was first published, who led the effort, how they interacted with providers, how much publicity the report received, and what market conditions existed at the time. Therefore, the better ratings of New York's report may not be attributable to its use of detailed clinical data. Those ratings may instead reflect New York's longer track record, shorter lag time between submission of data and publication, greater oversight by a Cardiac Advisory Committee, and limited population of hospitals. All of the hospitals licensed to perform CABG surgery in New York are large, high-volume facilities. Other agencies or coalitions using administrative data may achieve better ratings through attention to deficiencies

in methodology and communication.

A final limitation is that we only surveyed one of the three target audiences for hospital report cards: providers, consumers, and purchasers/payers. Relatively little is known about the role of consumers, but focus group participants in Oregon viewed undesirable event indicators (e.g., mortality rates) as more difficult to understand and less important than patient satisfaction ratings and desirable event indicators (e.g., cancer screening rates).[62] However, over two thirds of participants weighted undesirable event indicators more heavily when they were asked to select a hypothetical health plan.[63] One consumer-oriented report card appears to have stimulated many hospitals to offer expanded services and reduce cesarean rates.[9] Future studies should address how consumers and purchasers/payers value and use hospital outcomes information.[64]

The era of scorecard cardiovascular medicine has arrived.[65] Hospital leaders in New York and California are adapting to this new era, and appear more favorably disposed toward recent hospital report cards than their predecessors were toward the pioneering federal reports in the late 1980s. Although these report cards are intended to improve care by stimulating local quality improvement efforts, considerable lack of knowledge and skepticism persist among hospital leaders. The use of detailed clinical data may reduce this skepticism and enhance dissemination of outcomes data to key staff. By involving hospitals in the data collection process, clinical data systems may promote more education and discussion within the provider community. Small and low-volume hospitals face particular challenges in using outcomes data, given the inherent difficulty estimating outcome rates for small numbers of patients. For-profit hospitals find report cards useful for externally directed activities, such as negotiation and marketing, but should involve their medical and nursing staff more in reviewing and responding to outcomes data.

High-mortality outlier hospitals pose a special challenge. Hospital leaders tend to blame the messenger when their facilities are rated poorly, and argue that the risk-adjustment methods are inadequate. The 1996 CHOP report featured a detailed description of the analytic approach, including a list of all predictor variables and results from a major validation study based on reabstracting medical records. Obviously, this information was not communicated in a clear and compelling manner. Supplementary educational efforts through trade publications, professional meetings, and opinion leaders may be helpful. In the long run, state agencies and private coalitions must produce clearer and more timely report cards to overcome providers' skepticism, especially at the hospitals that might benefit the most by carefully examining their outcomes and processes of care.

Table 1. Number of respondents and response rate by hospital characteristic, California[a]

Characteristic	Respondents Number	(%)	Nonrespondents Number	(%)
Ownership[b]				
For-profit	51	(54)[c]	43	(46)[c]
Nonprofit corporate	93	(70)	40	(30)
Nonprofit church	36	(75)	12	(25)
Kaiser Permanente	17	(71)	7	(29)
University of California	4	(67)	2	(33)
Public/district	48	(74)	17	(26)
Size				
Small (<101 beds)	78	(63)	46	(37)
Medium (101-209 beds)	83	(67)	41	(33)
Large (>209 beds)	88	(72)	34	(28)
Occupancy[d]				
Low (<35.4%)	76	(62)	46	(38)
Medium (35.4-48.0%)	83	(67)	41	(33)
High (>48.0%)	90	(73)	33	(27)
Mortality outlier status				
Low (p<0.01)	14	(82)	3	(18)
Nonoutlier	220	(65)	116	(35)
High (p<0.01)	15	(71)	6	(29)
Housestaff training programs				
Yes	18	(72)	7	(28)
No	231	(66)	118	(34)
AMI volume 1991-1993[e]				
Low (<87 patients)	67	(58)	49	(42)
Medium (87-206 patients)	85	(68)	40	(32)
High (>206 patients)	97	(73)	36	(27)
Cardiovascular surgery				
Yes	71	(66)	36	(34)
No	178	(67)	89	(33)

[a] Twenty-five unusable responses were assigned to the nonrespondent category.

[b] Difference across ownership categories is marginally significant by Pearson chi-square, p=.065.

[c] Differ significantly from nonprofit corporate (p=.023), nonprofit church (p=.027), and public/district hospitals (p=.019) by continuity-adjusted chi square.

[d] Difference across occupancy categories is marginally significant by linear association chi-square, p=.070.

[e] Difference across AMI volume categories is statistically significant by linear association chi-square, p=.012.

Table 2. Hospital units or departments that received at least part of the 1996 state outcomes report, or discussed it with other units or departments

	California (n=249)				New York (n=27)			
	Received		Received or Discussed		Received		Received or Discussed	
Unit or department	no.	(%)	no.	(%)	no.	(%)	no.	(%)
President, CEO, or administrator	180	72.3[a]	226	90.8	26	100[a]	27	100
Medical records or coding staff	110	44.2	158	63.5				
Quality improvement staff	186	74.7	228	91.6	20	74	26	96
Medical director or medical staff	121	48.6[c]	193	77.5[a]	27	100[c]	27	100[a]
Nursing director or nursing staff	110	44.2	172	69.1	10	37	20	74
Board of directors	48	19.3	125	50.2[b]	8	30	22	81[b]
Public relations or marketing staff	55	22.1[c]	98	39.4[c]	116	59[c]	22	81[c]
Hospital legal counsel	13	5.2	28	11.1				
Contracting health plans	7	2.8	19	7.6				
Outside consultant	1	0.4	3	1.2				
Total number of units (mean)[d]	3.3		5.0		3.8		5.2	

[a] Difference between CA and NY is significant by continuity-adjusted chi square, $p<.05$.

[b] Difference between CA and NY is significant by continuity-adjusted chi square, $p<.005$.

[c] Difference between CA and NY is significant by continuity-adjusted chi square, $p<.0001$.

[d] These means are not directly comparable because the New York questionnaire listed fewer options (as shown).

Table 3. Hospital ratings of risk-adjusted mortality data

	Excellent, %			Very Good, %			Good, %			Fair, %			Poor, %			Mean score*		
Item	CA	NY	HCFA	CA	NY	HCFA	CA	NY	HCFA	CA	NY	HCFA	CA	NY	HCFA	CA	NY	HCFA
Usefulness in improving hospital quality[a]	2	7	0	6	15	1	26	48	4	34	19	25	32	11	70	1.1	1.9	0.4
Accuracy as a descriptor of hospital performance[b]	3	7	1	9	22	6	34	41	13	36	22	26	18	7	54	1.4	2.0	0.7
Completeness of case-mix adjustment model	2	4	1	16	22	2	34	22	8	37	37	35	10	15	55	1.6	1.6	0.6
Ease of interpretation[c]	4	15	1	14	48	4	37	19	20	35	7	37	10	11	38	1.7	2.5	0.9
Usefulness to consumers in market area	0	0	0	6	19	1	20	15	2	32	44	12	41	22	85	0.9	1.3	0.2
Way report was released to hospital and public	2	0	1	9	30	5	40	26	24	37	26	37	12	19	33	1.5	1.7	1.0
Overall score (mean)[d]																1.4	1.8	0.6

[a] Difference between CA and NY is significant by linear association chi square, p=.0002.

[b] Difference between CA and NY is significant by linear association chi square, p=.0067.

[c] Difference between CA and NY is significant by linear association chi square, p<.0001.

[d] Mean scores range between a minimum of 0 (if all respondents answered "poor") and a maximum of 4 (if all respondents answered "excellent"). Values of 1, 2, and 3 correspond to "fair," "good," and "very good," respectively. The difference between CA and NY is significant by linear association chi square, p=.0074.

Table 4. Perceived usefulness of risk-adjusted mortality reports in California and New York

	California (n=249)				New York (n=27)			
	Agree		Disagree		Agree		Disagree	
Activity	no.	(%)	no.	(%)	no.	(%)	no.	(%)
Improving quality of care[a,b]	168	68.0	79	32.0	24	89	3	11
Improving quality of ICD-9-CM coding	139	57.4	103	42.6				
Negotiating with health plans[c]	74	30.8	166	69.2	14	52	13	48
Marketing and public relations[d]	85	35.3	156	64.7	18	67	9	33
Overall score (mean)[e]	1.9				2.8			

[a] We inferred that individuals who rated the CHOP or CSRS report "poor" in its usefulness "in helping to improve your hospital's quality" would disagree with a hypothetical statement that the report was useful for this purpose.

[b] Difference between CA and NY is significant by continuity-adjusted chi square, p=.043.

[c] Difference between CA and NY is significant by continuity-adjusted chi square, p=.047.

[d] Difference between CA and NY is significant by continuity-adjusted chi square, p=.0030.

[e] Mean scores range between a minimum of 0 (if all respondents disagreed with all statements) and a maximum of 4 (if all respondents agreed with all statements). The difference between CA and NY is significant by linear association chi square, p=.0026.

Table 5. Hospital leaders' internal dissemination of materials from the California Hospital Outcomes Project, ratings of its quality, perceptions of its usefulness, and knowledge about its methods, by hospital characteristic

Characteristic	**Units using** Mean no. (0-10)	**Quality** Mean rating (0-4)	**Usefulness** Mean score (0-4)	**Knowledge** Mean score (0-4)
Ownership				
For-profit	4.4[a]	1.5	2.4[b]	1.6[c]
Nonprofit corporate	5.4	1.3	1.6	2.2
Nonprofit church	5.1	1.5	1.9	2.2
Kaiser Permanente	5.0	1.6	2.1	2.0
University of California	5.3	1.1	0.7	1.3
Public/district	4.9	1.5	1.9	1.8[d]
Size				
Small (<101 beds)	4.4[e]	1.4	1.9	1.8
Medium (101-209 beds)	5.1	1.4	2.0	1.9
Large (>209 beds)	5.5	1.4	1.7	2.2[f]
Occupancy				
Low (<35.4%)	4.9	1.4	2.0	2.0
Medium (35.4-48.0%)	5.0	1.4	1.8	2.0
High (>48.0%)	5.1	1.3	1.8	2.0
Mortality outlier status				
Low (p<0.01)	5.9	2.2	2.8[i]	2.9[i]
Nonoutlier	4.8	1.4[h]	1.8	2.0
High (p<0.01)	6.8[g]	0.9	1.5	1.7
Housestaff training				
Yes	4.8	1.5	1.9	2.0
No	5.0	1.4	1.9	2.0
AMI volume 1991-1993				
Low (<87 pts)	4.3[j]	1.5	2.1	1.6[j]
Medium (87-206 pts)	5.2	1.5	1.9	2.1
High (>206 pts)	5.3	1.3	1.7	2.2
Cardiovascular surgery				
Yes	5.2	1.4	1.8	2.1
No	4.9	1.4	1.9	1.9
Medicaid share				

Low (<10.1%)	5.4[k]	1.4	1.9	2.0
Medium (10.1-26.4%)	5.1	1.3	1.7	2.2[l]
High (>26.4%)	4.7	1.5	2.0	1.9

[a] Differs significantly from nonprofit corporate hospitals by Wilcoxon rank sum test, p=.025.

[b] Differs significantly from nonprofit corporate (p=.0030) and University of California (p=.047) hospitals by Wilcoxon rank sum test.

[c] Differs significantly from nonprofit corporate (p=.0041) and nonprofit church (p=.0086) hospitals, by Wilcoxon rank sum test.

[d] Differs significantly from nonprofit corporate hospitals by Wilcoxon rank sum test, p=.044.

[e] Differs significantly from medium (p=.016) and large (p=.0004) hospitals, by Wilcoxon rank sum test.

[f] Differs significantly from small hospitals by Wilcoxon rank sum test, p=.010.

[g] Differs significantly from nonoutlier hospitals by Wilcoxon rank sum test, p<.0001.

[h] Differs significantly from low-mortality (p=.0002) and high-mortality (p=.0078) outlier hospitals, by Wilcoxon rank sum test.

[i] Differ significantly from nonoutlier (p=.018 usefulness; p=.0038 knowledge) and high-mortality outlier (p=.0086 usefulness; p=.0036 knowledge) hospitals, by Wilcoxon rank sum test.

[j] Differs significantly from medium-volume (p=.0050 units using; p=.0096 knowledge) and high-volume (p=.0012 units using; p=.0017 knowledge) hospitals, by Wilcoxon rank sum test.

[k] Differs significantly from high-Medicaid hospitals by Wilcoxon rank sum test, p=.038.

[l] Differs significantly from high-Medicaid hospitals by Wilcoxon rank sum test, p=.041.

STUDY 2: Survey of hospital quality improvement leaders in California

How do hospitals respond to external evaluations of their effectiveness, such as public report cards?[66, 67] Are these responses likely to improve quality of care, and outcomes, for present and future patients? At one extreme, hospitals may ignore or dismiss their evaluations. A hospital responding in this manner would make no effort to disseminate the report or its findings, and would deride the report as invalid or irrelevant. At the other extreme, a critical report may be viewed as very salient because of its potential impact on consumer choice, managed care contracts, and market share. A hospital responding in this manner would distribute the report to units concerned with its core activity of patient care. As a result, the process of care would be evaluated and improved. Gormley and Weimer identify three other "functional responses": reallocating inputs to change the mix of outputs, focusing managerial attention on key issues, and enhancing the organization's mission by changing norms and beliefs.[66]

Between these two extremes is another type of organizational response, described as ceremonial.[68] An organization responding in this manner would alter observable activities to create the impression that established processes are working. For example, a hospital may create a quality improvement committee to review treatment protocols and even produce formal guidelines. But if the actual care of patients does not change, the response is ceremonial. This type of response is exemplified by distributing the report to units concerned with the hospital's relationship to the external environment, such as marketing staff, legal counsel, and the board of directors. The activities likely to result are described by Gormley and Weimer as "dysfunctional responses" because they create the illusion of improvement: cherry-picking low-risk cases, manipulating the data, and blaming the messenger.[66]

Hospitals may adopt different responses depending on the extent to which a report card is seen as a threat to their legitimacy or their ability to compete in the marketplace. If the report card is viewed as threatening the continued supply of patients, a hospital might alter its treatment protocols to improve outcomes. If the report card is deemed to have little potential effect on patient volume, but to represent a threat to the hospital's legitimacy, a ceremonial response might be elicited instead. The hospital's capacity to respond mediates its response. For example, larger hospitals or hospital systems have more resources with which to respond, such as well-developed quality improvement committees and research units.

We undertook a two-stage survey of hospital leaders in California to explore the impact of reports from the California Hospital Outcomes Project on efforts to improve quality of care and patient outcomes. Our hypothesis was that public dissemination of information about outcomes would motivate providers to investigate ways to improve their outcomes. What specific activities resulted from dissemination of the report? Which aspects of the report were felt to be particularly useful to hospitals? Which aspects were felt to make the report less useful?

BACKGROUND

When the California legislature enacted Assembly Bill (AB) 524 in November 1991, an

ambitious program to analyze and disseminate risk-adjusted hospital outcomes data was established.[69] The California Hospital Outcomes Project (CHOP) uses ICD-9-CM coded hospital discharge abstracts, routinely collected by California's Office of Statewide Health Planning and Development (OSHPD), to develop disease-specific, customized risk-adjustment models. The first report, released to the public in December 1993, classified inpatient mortality rates for acute myocardial infarction (AMI) and complication rates for cervical and lumbar diskectomy into two categories: "better" or "not better" than expected.[13] The second report, released to the public in May 1996, classified AMI mortality rates into three categories: "better," "worse," and "not significantly different" than expected.[56] The third report, which also focused on AMI mortality, was released to hospitals in July 1997 and to the public in December 1997.[16]

Each CHOP report has several volumes. The *User's Guide* includes an overview of the study objectives and the risk-adjustment methods, with a table showing each hospital's name, county, and rating based on each risk-adjustment model. Beginning in 1997, this volume also presents graphs showing each hospital's risk-adjusted mortality rate and 98% confidence interval, sorted by county. The *Technical Guide* includes a background literature review, defines all outcome variables and risk factors, describes the data linkage methods, and displays all of the risk-adjustment models and their performance characteristics. The *Detailed Statistical Tables* show more specific numeric data from each hospital, including the number of cases, observed and expected death rates, risk-adjusted death rates with confidence limits, and exact *p* values.

As required by AB 524, drafts of all three components of the report are sent to hospitals for comment before they are released to the public. Each hospital also receives a diskette containing all available information on its individual patients, and a *Hospital Guide* that explains how to use these spreadsheets and how to interpret the results. During the comment period for both the first and second reports, the regional hospital associations organized educational workshops for hospital staff throughout the state. At these workshops, OSHPD staff and consulting researchers helped participants to understand their outcome statistics and to use the diskette data to examine coding accuracy and provider performance. Hospitals were invited to respond to the report within 60 days of receipt. These response letters were appended to the *User's Guide*, along with a summary of their major themes, before public release.

The present study focuses on the second report of the California Hospital Outcomes Project, which presented risk-adjusted 30-day inpatient death rates for AMI patients treated in 1990-1992. This report was unique in that it also summarized the results of a comprehensive validation study that involved detailed reabstraction of both clinical and ICD-9-CM data on nearly 1,000 randomly selected AMI patients from 30 randomly selected hospitals.[56] This validation study confirmed that the classification of hospitals in the public report was only modestly affected by variation in ICD-9-CM coding practices and uncodable clinical risk factors. The validation study also found that several specific therapies for AMI, such as early use of aspirin and coronary revascularization, were performed more often at hospitals with low risk-adjusted mortality than at those with high risk-adjusted mortality.

SURVEY OF HOSPITAL ADMINISTRATORS

In the first phase of the study, a four-page questionnaire was mailed to the chief executive officer of each hospital included in the 1996 report. In this questionnaire, respondents were asked to identify which units or departments received or participated in discussions of the 1996 report. Respondents were also asked to rate the quality and usefulness of the report (alone and in comparison with other outcome monitoring systems), on a variety of dimensions.[70] Of the 374 hospitals included in the 1996 AMI report that were still providing acute care, 249 (66.6%) provided usable responses. Respondents from 25 additional hospitals (6.7%) indicated that they could not complete the survey, either because they had never received the report or because they were unfamiliar with its contents. Ninety-two percent of respondents indicated the report had been shared with quality improvement staff. Somewhat smaller proportions of hospitals involved their medical staff (78%), nursing staff (69%), and medical records or coding units (64%) in reviewing or discussing the report. Less commonly, the report was sent to or discussed with the board of directors (50%), public relations or marketing staff (39%), legal counsel (11%), contracting health plans (8%), and outside consultants (1.2%). On average, the report was shared with 5 hospital units.

This survey established that the CHOP report was disseminated widely both within and outside California hospitals. The second phase of our study was designed to explore whether the report had been used to improve quality of care or other aspects of organizational performance, such as abstraction of medical records. Of the 249 hospitals that returned the written questionnaire, 129 identified a contact person who was willing "to provide more detailed information" about the hospital's impressions and responses. We undertook a telephone survey of a stratified random sample of these 'key informants'.

TELEPHONE SURVEY OF KEY INFORMANTS

All 129 hospitals that identified a key informant in the written questionnaire were eligible for a telephone interview. To generate a balanced study sample with adequate representation of key types of hospitals, we stratified hospitals by three factors (Table 6):

1. *Risk-adjusted mortality*. Five categories were created: significantly better than expected AMI mortality at $p<.01$ (hospitals assigned a "star" in the 1996 report), marginally better than expected mortality at $.01<p<.05$, neither better nor worse than expected mortality, marginally worse than expected mortality at $.01<p<.05$, and significantly worse than expected mortality at $p<.01$ (hospitals assigned a "black dot" in the 1996 report).
2. *AMI volume*. Two categories were created; high volume hospitals were defined as those with more than the median number (175) of AMI patients included in the public report.
3. *Report Dissemination*. This variable represents the extent to which the report was disseminated within the hospital, based on responses to our written survey. A hospital with "limited dissemination" involved two or fewer units in reviewing or discussing the CHOP report, whereas a hospital with "extensive dissemination" involved three or more units. We reasoned that nearly every administrator would send the report to one other unit, such as quality improvement. More extensive dissemination suggests that the information was considered salient to the hospital.

Six of the twenty strata created in this manner contained no hospitals. Specifically, only two

hospitals with either better or worse than expected outcomes reported limited dissemination. Because we were especially interested in how hospitals labeled as outliers used the report, we sampled 100% of the hospitals with better or worse than expected outcomes. We randomly sampled three hospitals from each of the four strata containing non-outlier hospitals, creating a target sample size of 45.

We attempted to reach the identified informant at each of these 45 hospitals at least six times over a period of six weeks. If these six attempts were unsuccessful, a replacement hospital was sampled randomly from the same stratum. Replacement was ultimately impossible for one stratum (neither better nor worse than expected outcomes, low volume, limited dissemination) in which we attempted to reach all six hospitals but could only complete two interviews.

A trained interviewer with extensive previous experience introduced herself and established that the interviewee was familiar with the CHOP report. The interviews were semi-structured with open-ended questions, averaging 20-25 minutes in duration. Our interview questions were pretested on six hospitals with neither better nor worse than expected outcomes, high volume, and extensive dissemination. Although the questions were reordered and slightly modified after this pretest, the answers obtained were consistent with the final interviews described below. Informants were asked about:

1. Whether and how they used the patient-specific data provided on diskette (and if not, why not);
2. What specific aspects of the 1996 CHOP report they found to be useful or not useful;
3. What specific activities or discussion, if any, occurred within the hospital as a result of the 1996 CHOP report;
4. Whether and why the CHOP report was shared anyone outside the hospital staff (and if not, why not);
5. The usefulness of the CHOP report compared with other quality of care systems;
6. Suggestions for making the CHOP report more useful, informative, or understandable;
7. Why the hospital did or did not submit a response letter for publication with the final CHOP report in 1996; and
8. Whether risk-adjusted outcomes data should be reported to the public, and if so, what entity or group should be responsible (and if not, why not).

Follow-up questions were asked as needed to obtain more complete descriptions of relevant projects, activities, and opinions. Respondents were assured of the confidentiality of their responses.

Although our sample was not intended to be representative of a larger population, it allows us to describe how one report card was used in different types of hospitals. If an interview question elicited a “yes” or “no” response from all interviewees, the results were tallied. Open-ended questions were coded into a subset of the most frequently occurring responses, after we developed a list of all responses. For some items, we have elected to quote respondents verbatim or list important themes.

RESULTS

Table 6 shows that 39 telephone interviews were completed, representing an overall yield of 84% (38 of 45) after replacement. As shown in Table 7, about three-quarters of the interviewees held positions involving quality improvement (74.4%). Other job classifications included hospital administrator (but not CEO, 10.3%), medical director (7.7%) and director of medical records (5.1%).

Although about 50% of respondents reported opening the diskette with patient-specific data, most did not thoroughly review the spreadsheet. About one third of those using the data pulled specific medical records for review. A few hospitals reestimated their risk-adjusted mortality rates after correcting ICD-9-CM codes or removing cases that were felt to bias the results, such as those with do-not-resuscitate orders. The proportion of hospitals that reported using the diskette was similar between outlier and non-outlier hospitals. Among the respondents who did not use the diskette, many were uncertain whether they had ever received it, or reported difficulty opening it or using it on an appropriate computer (i.e., IBM-compatible PC).

Only about one quarter of respondents recalled attending one of the half-day courses on the use of risk-adjusted outcome data that were sponsored by the regional hospital associations in 1993 and 1995. About half of these attendees reported that they did not find the course to be very useful. Among those who had not attended, most said they either did not know about the courses or did not see the need to go.

About three-quarters of respondents found some aspect of the CHOP report to be useful. The most frequent comment was that the report was at least somewhat useful for "benchmarking" performance. These respondents viewed the CHOP report as one means of comparing their hospitals' risk-adjusted outcomes with those at comparable institutions statewide and competing hospitals in the local market. Several hospitals also mentioned that the report was useful not only for improving ICD-9-CM coding, but also for educating physicians about the importance of coding. According to one interviewee, "it stimulated attention and discussion among medical staff. It showed physicians that their documentation affects coding, which affects our (risk-adjusted) outcomes".

When we asked respondents what they found least useful about the report, the most common answer was that the report was not timely and the data did not reflect current practices. In addition, respondents complained that the report described outcome rates without providing practical information about the process of care. As one hospital quality manager put it, "the 'key drivers' for good outcomes should be determined from a performance improvement standpoint and there should be sharing of known processes to improve outcomes". Many other concerns were expressed, including poorly standardized ICD-9-CM coding, excessive complexity and technical detail, attribution of deaths after transfer to the originating hospital, and inclusion of superfluous information which made the report time-consuming to review. Another complaint was that the report compared very dissimilar hospitals, such as those of different size and those providing highly specialized care, such as cancer care, with non-specialty hospitals.

Despite these perceived drawbacks, most interviewees said that the CHOP report was disseminated to other individuals or groups within the hospital, such as medical staff leaders

(e.g., cardiologists) and quality improvement committees. Several said that the report was summarized at committee meetings or in a hospital staff newsletter. Nearly all interviewees at hospitals with better or worse than expected mortality shared the report with public relations or marketing units. Some hospitals with better than expected mortality also shared the report or its contents with managed care organizations or the general public (through brochures or newspaper advertisements). By contrast, only one nonoutlier hospital shared the CHOP report with units outside the hospital. The most common reasons given for not doing so were that the hospital's outcomes were unremarkable or "no one asked." "Payors aren't interested anyway," according to one interviewee. Several other respondents noted that their hospitals are small, geographically isolated, or in environments where advertising and marketing are not warranted.

Interviewees were asked to describe specific activities, if any, that resulted from receipt of the CHOP report. Two-thirds of respondents indicated that no specific actions were taken and no changes were made in the care of AMI patients. In one hospital, an ongoing review of AMI cases was instituted, which contributed to the development of a new critical pathway for AMI management. Three hospitals reported evaluating their overall use and timeliness of thrombolytic therapy in the emergency room (ER); at least one documented improvement. In another hospital, this evaluation led to refinement of existing critical pathways for AMI care. At two hospitals, poor ratings in the CHOP report led administrators to examine and then change the medical staff members assigned to treating AMI patients in the ER. At one hospital that received a “worse than expected” rating, the respondent said “we developed a protocol for evaluation and triage of patients in the ER with chest pain to receive tissue plasminogen activator (TPA) within 30 minutes. Cardiology developed an on-call panel for invasive procedures. We developed a pathway for treatment of ‘uncomplicated’ AMI.” Another hospital reviewed its TPA policy and added thrombolytics to its standard patient consent form.

Three hospitals undertook fairly extensive activities to improve coding. One hospital reported that after all AMI cases had been pulled and reviewed, coding errors were summarized and an in-service was held to educate hospital coders. Another hospital hired a coding consultant, implemented coding changes, and advised physicians to improve their documentation.

Most respondents felt that other information they receive about quality of care is more useful than that provided by the California Hospital Outcomes Project. For example, respondents appreciated and preferred the more detailed process data they receive from Genentech’s National Registry of Myocardial Infarction and the Health Care Financing Administration’s Cooperative Cardiovascular Project. The availability of rapid feedback on processes of care was felt to be a major advantage of these systems. Some quality managers viewed the CHOP report as confirming information they had already received from other systems or from internal monitoring programs (e.g., Voluntary Hospitals of America). Some noted that private vendors, such as HCIA (now Solucient), offered more favorable and timely evaluations of quality.

About half of our informants made specific suggestions about how future reports could be improved. The majority indicated a need for more timely data. Many suggested a simpler, more user-friendly presentation of results, with better graphics. Several respondents indicated the report should be shorter and provide an executive summary or an explanation in lay terms.

Improving the consistency of coding was seen as an important way to improve the data upon which the report depends. Many respondents wanted to know what the "better" hospitals were doing differently; several argued that the report should identify process-of-care factors that are correlated with "better than expected" outcome status. One respondent suggested that the report list the "top ten hospitals" for AMI and describe their treatment protocols.

When asked whether information about the quality of hospital care should be made available to the public, nearly every respondent said "yes, but…". Many expressed concern that risk-adjusted outcomes data are not easily explained to consumers, the current methodology is complex, and its presentation is overly detailed. According to one quality manager, "the public should be able to select the hospital with the best care, but this information is too difficult for them to understand." And there was concern that report cards should be based on measures that are widely accepted and validated. As one respondent said, "before public release, there needs to be consensus about what's being measured. We need better design of measures in the first place so we can tell patients something tangible." There were also concerns that mortality data do not reflect the overall quality of hospital care, are 'emotionally charged", and either should not be the focus of public disclosure or should be reported along with other measures such as length of stay, readmissions, and cost. Two-thirds of our respondents believed that a government agency should be responsible for public reporting of hospital outcomes data.

DISCUSSION

Several clear messages emerge from our detailed conversations with quality managers and other hospital leaders about the CHOP report. First, the timeliness of outcomes information is a key concern. Efforts to monitor and improve care depend on promptly identifying ineffective processes of care and poorly performing providers. The CHOP report is viewed as having little practical application due to the age of the data. Therefore, it is not surprising that most (but not all) reported uses of CHOP data, such as dissemination to marketing staff, summarization in newsletters and brochures, and evaluation or improvement of ICD-9-CM coding, are largely ceremonial responses.

Second, hospital quality managers desire information about processes of care, not just outcomes. Two-thirds of our interviewees indicated that no specific quality improvement activities resulted from receipt of the CHOP report. According to our informants, this lack of activity confirms that CHOP data do not meet the quality improvement needs of most hospitals. Quality managers suggested that the project should scrutinize hospitals with better than expected outcomes, determine "best practices" for AMI, and then identify hospitals using such practices. Although this idea would require a new effort by the state, hospitals seem receptive to improving patient care in this manner.

The California Hospital Outcomes Project, similar to projects in New York, Pennsylvania, Florida, and elsewhere, has focused on publicly releasing outcomes data, hoping that these data will be used constructively. Our interviews revealed that despite the age of the data and the lack of process-of-care information, some quality managers made serious attempts to study their AMI outcomes data and identify areas for improvement. Several hospitals were stimulated to develop

or refine clinical pathways, improve thrombolytic use, or reassign medical staff. These efforts are consistent with previously reported case studies from California [52,57] and New York.[42,44] A comprehensive educational component might help even more hospitals to make use of AMI outcomes data. One respondent even suggested that the state provide a profile of "best practices for using the CHOP report."

This study has several limitations. Most importantly, we surveyed just one of several potential audiences for risk-adjusted outcomes data: hospital quality managers. One would expect different evaluations of the project from different audiences. Other potential users of hospital report cards include hospital administrators; physicians, nurses, and other health professionals; employers; managed care organizations; media; government or regulatory agencies (e.g., the Joint Commission on the Accreditation of Healthcare Organizations); and the general public.

Relatively few studies have examined how these other audiences use hospital outcomes data. Hospital administrators were extremely negative toward Medicare hospital mortality reports in the late 1980s.[28] Our own survey of hospital leaders in New York and California found somewhat more favorable views toward state-sponsored report cards; New York's system was felt to be more useful than California's.[70] In both New York[55] and Pennsylvania,[54] cardiologists were skeptical about the value of risk-adjusted outcomes reports on coronary artery bypass surgery, and reported that the data have had little impact on their referral practices. A recent survey of managed care organizations found that only 9% considered outcomes data important to their success in the marketplace.[71] The 1986 Medicare hospital mortality report received considerable media attention, but newspaper articles focused heavily on outlier hospitals and provided little guidance to consumers.[30] Finally, there is little evidence that hospital report cards affect consumer choice.[48,49] Indeed, one study suggests that consumers find hospital death rates less important as quality-of-care indicators than desirable-event rates, complication rates, patient satisfaction ratings, and disciplinary records.[63]

Second, our sample was derived from the subset of hospitals that responded to a prior written survey and identified an individual willing to discuss the CHOP report in depth. This method allowed us to contact key hospital staff members who were already familiar with the report and had been directly involved in coordinating the hospital's response. However, hospitals whose quality managers have particularly favorable or unfavorable views of the report, or are particularly willing to express their views, are probably overrepresented in our sample. Although the response rate to our original survey was 73% (274 of 374), 9% of these responses were not usable and only 52% of the remainder (129 of 249) identified a key informant. Fortunately, these 129 hospitals represent a broad cross-section of the hospital population, with adequate representation of low mortality, high mortality, low volume and high volume hospitals.

Third, response bias is always a concern in surveys that inquire about potentially sensitive beliefs and activities. Our informants may have exaggerated their actual use of the CHOP report to avoid seeming unresponsive to information about quality of care. Alternatively, they may have underreported their actual use of the report to emphasize its lack of value. We tried to minimize these potential biases by assuring informants of the confidentiality of their responses, and by emphasizing that our research was Federally funded. Because the first and second authors were

involved in producing the 1996 CHOP report, as contractors to California's OSHPD, all interviews were conducted by the third author.

The responses to our telephone interviews reflect an underlying tension between the quality improvement needs of health care organizations and the mandate for California's OSHPD to provide information to the public on outcomes of care. Information that can readily be used by hospitals and physicians to improve their processes of care may not satisfy the desire of health care consumers, purchasers, payers, and government agencies to track risk-adjusted outcomes. Most of the hospital quality managers we interviewed recognize that report cards are here to stay, and believe that both hospitals and consumers are entitled to receive valid, comparative risk-adjusted outcomes information in a timely manner. Given the obvious limitations of the CHOP report, including non-timely data, substantial inter-hospital variation in the reporting of ICD-9-CM diagnoses, and important missing predictors (e.g., do-not-resuscitate orders), it is remarkable that even a few hospitals undertook meaningful quality improvement activities based on CHOP data. As outcomes reporting becomes more widespread and more timely, these activities should increase.

Table 6. Hospital characteristics, interviews targeted, and interviews completed

Hospital characteristic			Interview Results		
Outcome[a]	**Degree of Report Dissemination[b]**	**AMI Volume[c]**	**Potential Interviews**	**Targeted Interviews**	**Completed Interviews (%)**
Better (p < .01)	Extensive	Low	2	2	2 (100)
Better (p < .01)	Extensive	High	5	5	4 (80)
Better (p < .05)	Extensive	Low	2	2	2 (100)
Better (p < .05)	Extensive	High	7	7	5 (71)
Neither	Extensive	Low	44	3	3 (100)
Neither	Limited	Low	4	3	2 (67)
Neither	Extensive	High	39	3	4 (133)[d]
Neither	Limited	High	7	3	3 (100)
Worse (p <.05)	Extensive	Low	3	3	3 (100)
Worse (p <.05)	Limited	Low	1	1	0[e]
Worse (p <.05)	Extensive	High	3	3	1 (33)
Worse (p <.01)	Extensive	Low	3	3	3 (100)
Worse (p <.01)	Extensive	High	6	6	6 (100)
Worse (p <.01)	Limited	High	1	1	1 (100)
Total (%)			127	45	39 (86.7)

[a] Categories refer to acute myocardial infarction (AMI) risk-adjusted mortality as reported in the 1996 California Hospital Outcomes Project (CHOP) report.

[b] Limited dissemination was defined as 1 or 2 hospital units; extensive dissemination was defined as more than 2 hospital units.

[c] Low AMI volume was defined as <175 AMI cases in the CHOP report; high AMI volume was defined as ≥ 175 AMI cases.

[d] One additional interview took place after the target number of interviews had been completed.

[e] Hospital closed before interview could be conducted.

Table 7. Hospital outcome status and interviewee position/title for completed interviews

Hospital outcome	**Interviewee position/title**				**Completed interviews**
	Administrator	**Medical Director**	**Quality Manager/ Director**	**Medical Records Manager/Director**	
Better[a]	1	1	11	0	13
Neither	1	0	10	1	12
Worse[b]	2	2	9	1	14
Total (%)	4 (10)	3 (8)	30 (77)	2 (5)	39

[a] Includes both significantly better ($p < 0.01$) and marginally better ($0.01 < p < 0.05$) than expected mortality.

[b] Includes both significantly worse ($p < 0.01$) and marginally worse ($0.01 < p < 0.05$) than expected mortality.

STUDY 3: Survey of managed care executives in California

Managed care organizations, such as health maintenance organizations (HMOs), are directly or indirectly responsible for selecting hospitals on behalf of their members. In an efficient market, these organizations should collect and analyze information about both price and quality, leading to well-informed and defensible contracting decisions. However, there is reason for concern that publicly available quality information may have little influence on the market for hospital services. Three studies of hospital volume before and after publication of Federal or State report cards on risk-adjusted mortality found either no effect or only modest and transient effects.[37,48,49] In addition, surveys of employers have generally shown moderate attention to health plan quality measures and virtually no attention to hospital quality measures.[72,73,74]

This study was part of a comprehensive effort to explore the impact of publicly reported information about hospital quality in California. We were especially interested in whether managed care executives in this mature market are familiar with hospital report cards, to what extent they find them useful (and if not, why not), and how they weight such data in comparison with other factors, such as price and convenience. By improving our understanding of how HMOs currently select providers, we hope to identify opportunities for enhancing the importance of quality.

METHODS

We obtained a list of 49 HMOs with active Knox-Keene licenses, as of June 1998, from the California Department of Corporations. Two of these HMOs contracted only for specialized services, such as transplantation, and were therefore excluded from this study. By telephone, we attempted to identify an individual in each organization who was responsible for collecting or sharing information about hospital quality for potential use in contracting decisions. About two thirds of these individuals had the title of Director, Manager, or Vice President of quality management, improvement, or assurance; most of the remainder had a title such as "Medical Director" or "Director, Health Services." We mailed questionnaires to all 47 contacts, and followed up with multiple telephone calls to those who did not return the questionnaire or who agreed to a verbal interview.

We obtained a list of employers that were self-insured for acute inpatient care, based on filings of form 5500, "Annual Return/Report of Employee Benefit Plan," with the California Department of Labor in 1994 (the most recent available data). Ninety employee medical benefit plans had at least 1000 active or retired participants. Of these 90 plans, 11 were sponsored by Kaiser Permanente, which was already eligible to participate in the survey as an HMO. Fifteen employers sponsored multiple plans, leaving 64 self-insured employers to contact by telephone.

We confirmed that each employer met our definition of "self insured" by maintaining a special fund to pay health care claims directly to provider organizations. Our screening calls focused on whether anyone in the organization was involved in selecting hospitals for plan participants. After at least four attempts, we established contact with knowledgeable individuals in 41 companies. Of these 41, 4 refused to answer our questions, 3 were no longer in business, and 9

were no longer self-insured. Of the 25 eligible employers, 22 (88%) used a third party administrator to process claims and to create a list of preferred providers, including hospitals. In all but one case, the employer named a third party administrator that was eligible for our HMO survey. Therefore, we did not send separate questionnaires to self-insured employers.

We asked health plan executives to review a list of factors that might have affected the health plan's decision to contract with specific hospitals. Each factor was rated on a 1-5 scale, where 1 was "not at all important" and 5 was "extremely important." We then asked respondents which sources of quality information they had used in the past year to help select hospitals, including the California Hospital Outcomes Project reports on risk adjusted heart attack mortality, the *Consumers' Guide to Hospitals* summary of risk-adjusted Medicare death rates and structural indicators, the Pacific Business Group on Health's "HealthScope" data on risk-adjusted cesarean rates, neonatal readmission rates, and coronary bypass surgery volume, *US News and World Report*'s rankings of "America's Best Hospitals," and their own in-house analyses. Finally, we asked respondents to indicate who should "take a leading role" in collecting and distributing information about hospital quality, and to rate the utility and potential impact of hospital outcome studies.

We obtained data on health plan characteristics, including the number of enrollees, the numbers of contracted hospitals and medical groups, tax status, and model type, from a directory published by the California Association of HMOs.[75] For HMOs that were not members of this association or did not have data included in its directory, we obtained data on enrollment, the number of contracted hospitals and medical groups, tax status, and model type from the California Medical Association's "Knox-Keene Health Plan Expenditures Summary," a compilation of data derived from reports to the Department of Corporations and statements provided by publicly traded plans to the U.S. Securities and Exchange Commission for FY 1997/98.[76] Enrollment data were also confirmed using data reported directly by the Department of Corporations.[77] Unless otherwise noted, HMOs were classified as smaller or larger than the statewide median of 57,000 enrollees. Tax status was classified as for-profit or non-profit. Model type was dichotomized as staff-model or "other." Ownership was classified as private or public, where public HMOs represent county-managed Medicaid plans. Plans that had full or provisional accreditation from the National Committee on Quality Assurance (NCQA), regardless of duration, were compared to unaccredited plans.[78]

RESULTS

Respondents and nonrespondents

Thirty of the 47 (63.8%) eligible HMOs provided usable responses: 19 in writing and 11 by telephone. There were no statistically significant differences between respondents and nonrespondents, but responding organizations tended to be larger and to have contracts with more hospitals (Table 8).

Among these health plans, NCQA accreditation was strongly associated with the number of plan enrollees, the number of contracted hospitals, and ownership. Accredited plans had an average of 1.47 million enrollees, compared with an average of 62,000 enrollees for unaccredited plans

($p<.05$). Accredited plans also contracted with more hospitals than did nonaccredited plans (544 versus 46, $p<.001$). All of the accredited plans were privately owned. These findings made it impossible to disentangle the independent associations between accreditation, size, and ownership, and health plan behavior.

Factors affecting contracting decisions

Three factors were rated as most important in hospital contracting decisions (Table 9): accreditation by the Joint Commission for the Accreditation of Healthcare Organizations (JCAHO), geographic location, and negotiated price. Every respondent said that accreditation and hospital location (ease of access for members) were either very or extremely important factors in their selection of hospitals. However, respondents from accredited plans tended to rate geographic location as less important than did respondents from non-accredited plans. Nearly all (97%) respondents also felt that negotiated prices were either very or extremely important. Respondents from for-profit plans assigned greater importance to price than did respondents from non-profit plans. The other factors deemed very or extremely important by at least 75% of respondents were disciplinary actions against the hospital by federal or state agencies, the hospital's reputation, and the hospital's apparent commitment to quality improvement.

The quality indicators available from hospital outcome studies were reportedly less influential in contracting decisions. For example, mortality rates after heart attack, cardiac surgery, or other conditions or procedures were considered very or extremely important by only 57% of respondents. Readmission rates, overuse (i.e., cesarean delivery, hysterectomy) rates, organ transplant survival, and preventable complication rates were rated as very or extremely important by 27% to 70% of respondents. We created an internally consistent scale by combining these five hospital outcome measures (Cronbach's alpha = 0.918). This scale was not significantly related to any health plan characteristic, except that respondents from staff model HMOs viewed these quality measures as more important in contracting decisions than did respondents from other types of plans.

Three other factors were considered very or extremely important by a majority of respondents: member satisfaction, the frequency of malpractice judgments against the hospital, and the duration of the plan's relationship with the hospital. Fewer than half of our respondents felt that administrative factors, such as the accuracy of claims submitted by the hospital, were very or extremely important. Similarly, relatively few respondents (23% to 43%) viewed process-of-care measures, such as thrombolytic use for heart attack patients, adherence to clinical practice guidelines, and performance on HEDIS (Health Plan Employer Data and Information Set) indicators, as very or extremely important in selecting hospitals.

Through an open-ended inquiry, we further explored how health plans weigh quality information against other considerations in selecting hospitals. Several comments emphasized the importance of managing costs and maintaining member satisfaction in a competitive market. Two respondents acknowledged that "we have to compete on price (or) we are out of the market"; hence "cost concerns override other factors." Another reported that "we have to take a lot of things into consideration" but "it is member satisfaction that makes a difference." Different departments within the same health plan may weigh these factors differently: "quality

managers have their perspective, which may be different from the contracting department, and less based on price." Therefore, deselection efforts tend to focus on "eliminating the very inefficient and low quality providers."

Sources of useful data on hospital outcomes

Seventy percent of respondents said that they had reviewed at least one publicly available source of information on hospital outcomes when selecting hospitals. Among those who had used publicly available hospital outcomes data, most reported them to be only minimally to moderately useful (Table 10). PBGH's HealthScope report on hospitals had been reviewed by 63% of respondents, but was rated as very good or excellent in usefulness by only 32% of those who had used it. Other publicly available sources were similarly familiar to our respondents, but were rated as even less useful. Reported use of these sources was not significantly related to any plan characteristic.

A health plan may select hospitals based upon measures of hospital performance that it develops and monitors itself, or with the help of external consultants. Ten of our thirty respondents (33%) reported that their plans had conducted "in-house," comparative studies of hospital performance to inform their contracting decisions. This behavior was much more common ($p<.001$) among accredited plans (7 of 9) than among non-accredited plans (3 of 21), and marginally more common ($p=.052$) among very large plans with more than 150,000 enrollees than among other plans. No other plan characteristics were associated with doing independent studies. The specific hospital performance measures cited by our respondents included cesarean delivery rates, readmission rates, transplant success rates, sentinel event rates, use of specific treatments (i.e., beta-blockers after heart attacks), member satisfaction, resource utilization, and average length of stay.

A major theme throughout these interviews was that "hospital quality is very difficult to measure and compare" in the absence of "empirical data on costs and outcomes which is (*sic*) validated by an external source." Several respondents complained about other limitations of the currently available data on hospital performance. According to one respondent, "hospital outcome studies are ineffective without follow-up in other settings, like outpatient care." Another respondent agreed that "outcome data need to be longitudinal," but also argued that public reports should match hospitals "on severity and service type" because "simple geographic differences, such as the availability of physicians…cause differences in outcomes." In this environment, one respondent said, "we really have to rely on our own experiences and our own monitoring."

Views toward the future

Health plan executives assigned primary responsibility for collecting and disseminating information on hospital quality to government agencies and accrediting organizations (Table 11). Consistent with the emphasis placed on hospital accreditation in contracting decisions, almost all respondents agreed that JCAHO should take the lead in public reporting. At least 80% of respondents also felt that the federal Health Care Financing Administration (HCFA) and the state's OSHPD should play leading roles. Most agreed that peer review organizations should also be involved. About 73% of respondents said that health plans should collect and analyze

their own information on hospital quality, and 69% looked to employers to take a leading role. Narrower majorities felt that provider-purchaser coalitions and provider associations should play a major role. Most respondents did not favor a leading role for consumer advocacy groups or for the news media.

Respondents from large plans were more likely than those from small plans to agree that provider-purchaser coalitions and provider associations should take a leading role. However, plan characteristics did not otherwise affect ratings of who should play a leading role in collecting and distributing hospital outcomes information.

A notable minority of respondents commented that it was unlikely that any "objective" information on quality would help plans or consumers to select hospitals; for example, "patients are influenced by the location of the hospital and the hospital's role in the community, not by government studies." However, a substantial majority agreed that hospital outcome studies would lead to improved quality of care (87%) and to less unnecessary and inappropriate care (70%) in the "next few years." Only 53% of our respondents endorsed the notion that hospital outcome studies would lead to lower costs. Responses to these questions were unrelated to plan characteristics.

DISCUSSION AND POLICY IMPLICATIONS

In this survey of managed care executives, we found substantial interest in hospital quality measures, but little evidence that health plans weigh such measures heavily in selecting hospitals. To the extent that health plans consider hospital quality, they tend to rely on measures with poor discrimination, such as accreditation, or subjective concepts, such as "reputation," "commitment to quality improvement," and "member satisfaction." Plans find it more efficient to "flag" problematic hospitals, based on JCAHO reviews or rare disciplinary actions by federal or state agencies, than to use other available measures of quality. Consequently, geographic convenience and price may be the dominant considerations in hospital contracting, especially among California's for-profit HMOs.

Most plans have made little effort to identify top-performing hospitals. However, ten of our thirty respondents reported that their plans had conducted "in-house," comparative studies of hospital performance. This behavior was strongly associated with NCQA accreditation, although the direction and mechanism of this association are unclear. Accredited plans apparently take a more pro-active approach than non-accredited plans in evaluating their contracted hospitals.

Two previous studies have reported on how managed care executives contract for hospital services. A detailed case study of three market areas showed that only in the largest, most mature market did health plans use objective information about quality of care to select tertiary care hospitals.[79] However, the authors of that study reported that "priorities were difficult to ascertain because most plans do not have an explicit, ordered list of criteria." In New York, 60% of respondents to a survey of HMOs stated that "quality is the most important consideration" in contracting with hospitals for coronary surgery; however, only 66% of these executives (64% of all respondents) had reviewed public reports on risk-adjusted mortality.[80] Consistent with our Table 10, only 20% of respondents viewed these reports as a "major source" of information.

The major strength of this study is that it was based on in-depth written or telephone interviews with key decision-makers in nearly two-thirds of all licensed HMOs in a populous, diverse state with high managed care penetration. We addressed the full spectrum of hospital quality indicators available to these decision-makers, not just a single published report card. Finally, our results are consistent with other surveys of managed care executives. When consultants at Foster Higgins asked managed care executives to rank "seven factors in order of importance to their marketplace success over the next three to five years," price was ranked first or second by 69%, patient satisfaction by 50%, provider access by 31%, and published outcomes data by 9%. Hence, the behavior of HMO executives reflects their perception of the environment in which they compete.[71]

However, our findings are based entirely on self-reported knowledge, attitudes, and behaviors, which may be biased. First, our respondents may have intentionally or unintentionally exaggerated the importance of quality information to please the researchers and to burnish the industry's image. In fact, many of our respondents were required to obtain approval from their supervisors before they could participate, and some were required to obtain approval for their specific responses. We have no information on how responses might have been altered through this process. Second, we may have identified respondents in a manner that made them unrepresentative of the health plans for which they worked. Among the employees involved in hospital contracting within each plan, our respondents may have been the most favorably disposed toward hospital quality information. We tried to overcome these concerns by recruiting high-level executives whenever possible, and by emphasizing that our study was Federally funded and that we would protect the confidentiality of all responses. These efforts were at least partially successful, in that most respondents candidly acknowledged the primacy of cost and convenience in selecting hospitals.

The other major limitation is that our results may not generalize to all managed care organizations in California or nationwide. For example, non-responding health plans might be even less interested in quality information, and less likely to use it, than our respondents were. Although there were several nonsignificant but notable differences between nonrespondents and respondents, our sample included a broad representation of HMOs of different sizes and types. Health plans in less mature markets may be less interested in hospital quality information than those in California, if they are less likely to design exclusive networks that "deselect" certain providers. Conversely, they may be more interested in such information, if they are more likely to pay hospitals directly, on a per diem or discounted fee-for-service basis, rather than through shared risk pools with globally capitated medical groups.

Report cards on hospital quality can only achieve their full potential if they are used by consumers or organizations that act on consumers' behalf to select hospitals. Previous studies have shown that consumers have little awareness of hospital outcome studies, and that such report cards have little impact on consumer decision-making.[81,82] Our study extends these findings to health plans, showing that key decision-makers in these organizations are aware of hospital report cards, but find them less useful than relatively subjective or non-discriminating measures of quality. Although there is substantial interest in information on hospital quality, and confidence that such information will improve care, health plan executives are concerned about

the limitations of risk-adjusted outcomes and uncomfortable weighting these data heavily when they select network hospitals. Although some of these concerns may be well-founded, it will be important for policy-makers and producers of hospital report cards to address them in the coming years.

Table 8. Characteristics of responding and nonresponding health plans

Characteristic	**Respondents**[a] (N=30)	**Nonrespondents**[a] (N=17)
Median number of enrollees	77,499	48,681
Median number of contracted hospitals	49	11
Median number of contracted medical groups	19	24
Tax status		
For profit	17 (57%)	7 (41%)
Nonprofit	12 (40%)	7 (41%)
Unknown	1 (3%)	3 (18%)
Ownership		
Public (county)	5 (17%)	6 (35%)
Private	25 (83%)	11 (65%)
Model type		
Staff model (includes mixtures)	7 (23%)	2 (12%)
Other (IPA, network, medical group)	19 (63%)	9 (53%)
Unknown	4 (13%)	6 (35%)
NCQA accreditation	12 (30%)	2 (12%)

[a] Differences between responding and nonresponding organizations are not statistically significant for any of the listed characteristics.

Table 9. Managed care executives' ratings of factors affecting hospital contracting decisions

Factor affecting decision	Importance (%)					Mean Score
	Not at all	Slightly	Somewhat	Very	Extremely	
JCAHO accreditation (full, provisional)	0	0	0	20	80	4.80
Hospital location[a]	0	0	0	40	60	4.60
Price (negotiated)[b]	3	0	0	53	43	4.33
Federal/state disciplinary actions	0	7	3	40	50	4.33
Reputation of the hospital	0	0	17	53	30	4.13
Commitment to quality improvement	0	13	7	43	37	4.03
Member satisfaction with hospital[c]	0	7	27	30	37	3.97
Long-standing relationship with hospital	3	13	17	40	27	3.73
Frequency of malpractice judgments	3	10	20	47	20	3.70
Readmission rates [c]	3	10	23	43	20	3.67
Organ transplant success rates	13	3	13	47	23	3.63
Average length of stay	3	10	30	40	17	3.57
Accuracy of claims submitted[d]	3	10	40	23	23	3.53
Mortality after heart attack, cardiac surgery, other conditions	3	20	20	37	20	3.50
Performance on HEDIS 3.0 indicators	7	23	27	30	13	3.20
Incidence of overused procedures	7	20	30	33	10	3.20
Development and adherence to practice guidelines[c]	10	20	33	27	10	3.07
Preventable complication rates	7	20	47	17	10	3.03
Information on process of care	10	20	47	13	10	2.93
Membership in multi-hospital system[b]	13	20	50	7	10	2.80
Recommendation of a consultant	27	43	13	10	7	2.27
Index of 5 Hospital Outcomes Indicators[e]	3	13	33	40	10	3.39

[a] Respondents from accredited plans rated this factor as marginally less important than did respondents from non-accredited plans (p=.053).

[b] Respondents from for-profit plans rated these factors as more important than did respondents from non-profit plans (p<.05).

[c] Respondents from staff model HMOs rated these factors as more important than did respondents from other model types (p<.05).

[d] Respondents from large health plans rated this factor as more important than did respondents from small health plans (p<.05).

[e] Index of 5 items (readmission rate, average length of stay, mortality rate, incidence of overused procedures, preventable complication rate). Respondents from staff model HMOs rated this index more highly than did respondents from other model types (p<.05).

Table 10. Managed care executives' ratings of sources of useful data on hospital quality

Source of Hospital Quality Data	Usefulness (%)					
	Excellent	Very Good	Good	Fair	Poor	Never used
US News and World Report "America's Best Hospitals"	6	17	22	39	17	40
Pacific Business Group on Health HealthScope report or website	0	32	42	16	10	37
Office of Statewide Health Planning and Development California Hospital Outcomes Project	3	17	33	33	11	40
Consumers' Checkbook "Consumers' Guide to Hospitals"[a]	6	12	41	29	12	43

[a] Respondents from nonprofit health plans were marginally more likely than those from for-profit plans to report having ever used this source (p=.053).

Table 11. Managed care executives' opinions about which groups should take a leading role in collecting and distributing data on hospital quality

	Agreement (%)			
Organization	**Strongly Agree**	**Agree**	**Disagree**	**Strongly Disagree**
Employers or employer organizations	3	67	20	10
State agencies	23	60	10	7
Federal agencies	37	53	3	7
Peer review organizations	23	57	17	3
Regional provider-purchaser coalitions[a]	10	50	30	10
Hospital or provider organizations[a]	7	48	35	10
News media	0	17	53	30
Accrediting organizations	55	38	7	0
Consumer advocacy groups	10	37	37	17
Health plans should collect and analyze their own information on hospital quality	27	47	17	10

[a] Respondents from small health plans were less likely to agree than respondents from large health plans ($p<.05$).

STUDY 4: Time-series analysis of the impact of report cards on hospital volumes in California and New York

Studies of hospital choice suggest that consumers, in the absence of public information, prefer hospitals with low risk-adjusted death rates, especially for acute myocardial infarction (AMI), coronary artery bypass graft (CABG) surgery,[32] and high-risk obstetric care.[33] Studies of the volume outcome relationship also suggest that patients are "selectively referred" to hospitals with good outcomes.[34,35,36] However, there is little evidence that the publication of risk-adjusted outcomes data affects referral practices. The Health Care Financing Administration's (HCFA) designation of 14 high-mortality and 9 low-mortality hospitals in New York City had no effect on subsequent occupancy rates.[37] A twofold increase in the standardized mortality ratio based on HCFA's 1986-1993 data releases was associated with 46 fewer Medicare discharges, on average, in each subsequent year, but this result varied with different model specifications.[48] The authors interpreted this finding as statistically significant but too small to be meaningful.

Report cards published by agencies in California and New York may have more impact on hospital volumes, because they are based on better data, focus on specific conditions or procedures, incorporate clinical expertise, and address regional concerns. The California Hospital Outcomes Project (CHOP) began in 1991, with the enactment of a law requiring the Office of Statewide Health Planning and Development (OSHPD) to produce periodic reports on risk adjusted outcomes at acute care hospitals, using ICD-9-CM coded discharge abstracts.[69] New York's Cardiac Surgery Reporting System (CSRS) began in 1989 with the creation of a special clinical data system for cardiac surgery. Hospital-specific, risk-adjusted mortality rates and 3-category ratings have been released every 12-24 months since December 1990.

The efforts in California and New York represent prototypes of two contrasting approaches to reporting risk-adjusted outcomes. If one approach appears to have had a greater impact on hospital volumes than the other, there may be important lessons for other states and countries. Hannan et al found no relationship between unpublished risk-adjusted mortality rates from CSRS (in 1989) and subsequent market share.[41] Mukamel and Mushlin found that the 1990-1992 CSRS reports had minimal effect on hospital volume, but did affect surgeon volume as hypothesized.[49] However, their analysis failed to adjust for several possible determinants of provider volume, and failed to consider the temporal course of a report's effect. No previous studies have examined the impact of the CHOP reports.

METHODS

Conceptual framework and hypotheses

Surveys of American consumers have shown that physician recommendations have the strongest influence on hospital choice.[83,84,85] Previous personal experiences and the recommendations of friends or family members are also important factors; indeed, 62-72% of consumers would

choose a hospital they “have used for many years without any problems” over a hospital that “is rated much higher in quality by the experts.” Accordingly, quality information may influence consumer choice in three ways. First, health plans may selectively contract with high quality providers for specialized services, such as coronary revascularization and back surgery.[79,80] Second, physicians may selectively refer patients to these providers,[55] especially if they do not already have well established referral relationships. Third, consumers (or family members acting on their behalf) may postpone or cancel elective surgery that is scheduled to be performed at a poorly rated hospital, or may request referral to a highly rated hospital.[81]

We therefore hypothesized that hospitals with lower than expected mortality or complication rates experience significant volume increases, and hospitals with higher than expected mortality or complication rates experience significant volume decreases, in the year after publication of a report card. These hypotheses were based on a recently validated assumption that consumers have little, if any, implicit knowledge about hospital quality before a report is published.[48] We analyzed volume data by month and quarter to determine whether any observed change was immediate or delayed. We did not analyze risk-adjusted mortality as a continuous predictor, as others have done,[48,49] because both the public reports and the attendant media coverage have focused on designated outliers.

We further hypothesized that hospitals with lower than expected mortality or complication rates attract more patients from long distances, or from outside their usual catchment areas, after a report is published. We label this a "bypass effect," as patients bypass hospitals that they would otherwise use to find favorable outliers. A "spillover effect" would be expected to lead to an increased volume of clinically related conditions or procedures. We expect this increase to be smaller than that for the condition or procedure analyzed in the report, except for AMI, because the volume of elective cardiologic procedures may be more sensitive to favorable or adverse publicity than the volume of AMI itself. Both the "bypass effect" and the "spillover effect" should lead to parallel volume decreases at hospitals with higher than expected mortality or complication rates.

Finally, we hypothesized that certain sociodemographic groups are more likely to hear about the release of a hospital report card, and are better able or more likely to use this information to select a hospital, than other groups. Specifically, changes in hospital volume may be greater among patients who are young, white, or privately insured than among those who are elderly, African-American, or publicly insured or uninsured. Older patients are hypothesized to have more inertia resulting from their established relationships with specific providers. Patients from minority groups, or without health insurance, are hypothesized to have more limited sets of hospitals from which to choose.

Data

Our California analysis was based on the California Patient Discharge Data Set, which includes computerized discharge abstracts from every licensed, non-federal hospital in the state. The variables collected include a hospital facility number; the patient's date of birth, social security number (SSN), zip code, race/ethnicity, sex, and disposition; the dates of admission and

discharge; the principal source of payment and total charges; the source and type of admission; the principal diagnosis and up to 24 other diagnoses; the principal procedure and up to 20 other procedures; and up to 4 external cause of injury codes. All diagnoses and procedures are coded using ICD-9-CM.

Our New York analysis was based on the Statewide Planning and Research Cooperative System (SPARCS), which also includes computerized discharge abstracts from every licensed, non-federal hospital in the state. New York's abstract is similar to California's, except that it includes a maximum of 8 other diagnoses (14 as of 4/1/94) along with a set of dummy variables indicating whether each secondary diagnosis was present at admission.

Subjects and hospitals

Using these administrative data sets, we identified all admissions to acute care nonfederal hospitals for target conditions and procedures, certain related conditions and procedures, and unrelated reasons (Table 12). The related conditions and procedures associated with each target condition or procedure were analyzed separately, but then aggregated in California because of similar effects and marginal power. We excluded children less than 18 years of age at admission and patients admitted for psychiatric conditions (ICD-9-CM principal diagnosis 290.xx-319), injury or poisoning (ICD-9-CM 800.xx-995.xx), or rehabilitation (ICD-9-CM V57.x). All of these patients are subject to specialized referral arrangements, which were beyond the scope of this study. We also excluded patients transferred from other acute care hospitals, because these transfers generally reflect the capabilities of different facilities, or insurance arrangements, rather than consumer choice. The OSHPD definition of transfers was fully inclusive, but the SPARCS definition did not capture unscheduled transfers for which the "admit source" was "physician" and scheduled transfers for which the "admit source" was "emergency," "outpatient," or "physician."

All analyses were limited to hospitals that were included in the relevant report card. In California, the CHOP report published in December 1993 evaluated AMI mortality at 394 hospitals, complications after lumbar diskectomy at 344 hospitals, and complications after cervical diskectomy at 277 hospitals. The CHOP report published in May 1996 evaluated AMI mortality at 398 hospitals. Kaiser hospitals and state developmental and correctional hospitals were included in the CHOP report, but were excluded from the present study because their patients did not have the freedom to choose a different facility. In New York, the CSRS report published in December 1992 evaluated 30 hospitals, whereas the reports published in December 1993 and June 1995 evaluated 31 hospitals.

The study period for the 1993 CHOP report began 24 months before publication (January 1, 1992) and ended 12 months after publication (December 31, 1994). The study period for the 1996 CHOP report began 24 months before publication (June 1, 1994) and ended 7 months after publication (December 31, 1996). The study period for each CSRS report began 36 months before publication and ended 12 months after publication (except 6 months after publication of the 1995 report).

Dependent variables

Our primary dependent variable was the total number of patients with a topic condition or procedure, or related condition or procedure, who were admitted to a specific hospital in a specific calendar month. Using hospital-months as the unit of analysis enabled us to track temporal changes in hospital volume. Our secondary dependent variables included the number of patients in each clinical category, stratified by age at admission, race/ethnicity, and insurance status. Age was categorized as younger than 55, 55-64, 65-74, or greater than 74 years of age. Race/ethnicity was categorized as white, African-American, Hispanic, or other; Hispanics could only be explicitly excluded from each non-Hispanic category in New York. Insurance status was categorized based on the expected principal source of payment as Medicare, Medicaid, private (including Blue Cross, Blue Shield, or commercial), HMO (non-Medicare, non-Medicaid), uninsured (including self-pay or indigent care), or other.

Finally, we identified patients who traveled from outside a hospital's catchment area. Catchment areas were defined using all of a hospital's topic, related, and unrelated discharges during the entire study period. For each patient at each hospital, we computed the air distance between the geographic centroid of his or her 5-digit zip code of residence and his or her hospital, using the Pythagorean theorem. The latitude and longitude of each zip code centroid was obtained from GDT, Inc.; the exact coordinates of each hospital were determined from US Geological Survey topographical maps and confirmed by GDT, Inc. We then defined each hospital's catchment area as the set of zip codes (rank-ordered by numerical importance) that contributed 60% of that facility's discharges, plus additional zip codes for which that hospital was the majority provider of inpatient, acute care (excluding pediatric, psychiatric, trauma, and rehabilitation care, as defined above) before publication of the first official report.[86] The data used to define catchment areas were from 1992-1993 in California, and from 1991-1992 in New York. An alternative definition of catchment area was the radius surrounding a hospital within which 60% of that hospital's patients originated. We did not use a fixed distance cutoff of 15 miles,[87] because the proportion of long distance patients using this definition varied widely, from less than 20% to over 95%.

Independent variables

Several independent variables were used to predict each hospital's volume after publication of a report card. First, we used statewide hospital volume for the same condition or procedure to represent its overall prevalence in a given month. Second, we used a set of dummy variables to represent random hospital effects, or the mean monthly volume of each hospital before publication of a report card. Using a complete set of interactions between statewide volume and hospital, we allowed the effect of statewide volume to differ across hospitals, reflecting the fact that some hospitals were more successful than others at maintaining market share, even before publication of a report card. Finally, we used unrelated volume in the same hospital-month as an indicator of changes in the size or geographic distribution of the local population, the accessibility of a hospital, and the size and scope of its referral network. Using a complete set of interactions between unrelated volume and hospital, we allowed the effect of unrelated volume to vary across hospitals, reflecting the fact that the product lines of interest (e.g., CABG, PTCA,

AMI) are inherently more susceptible to local demographic changes at some hospitals than at others. Invariant hospital characteristics, such as teaching status and size, were omitted because they would not explain temporal changes within the same hospitals.

Because a favorable report card could affect unrelated volume, as well as the volume of topic and related conditions and procedures, we tested alternative models without unrelated volume and its interactions. These models were often less powerful than our preferred models, but generated very similar results.

Following Mennemeyer et al,[48] we also tested mean hospital charges as a predictor of volume in California. Using a complete set of interactions between mean charges and hospital, we estimated the elasticity of demand at each hospital before publication of a report card. However, these variables were omitted from our final models because: (1) their statistical significance was marginal, (2) the direction of the relationship between mean charges and hospital volume was inconsistent; and (3) mean charges may be endogenous, in that favorable report cards may inspire hospitals to raise prices due to increased demand. Although this endogeneity could theoretically be removed using an instrumental variable, we were unable to identify any appropriate instruments for hospital charges.

Statistical methods

Our initial models were estimated using ordinary least squares (OLS), including only the hospital-months that preceded publication of a report card. Using the parameter estimates from this time period, and subsequent values of the independent variables, we predicted what each outlier hospital's volume should have been in each of the 12 months following publication of a report card. These predicted volumes were aggregated for all hospitals assigned to the same risk-adjusted performance category (e.g., higher than expected AMI mortality) in that report card. To smooth out month-to-month variation and improve statistical power, we then averaged predicted hospital volumes in California by calendar quarter. We estimated 95% confidence intervals surrounding these averaged predictions, for each quarter after a report was published. If the observed volume for the hospitals in a performance category fell outside this 95% (CI) confidence interval, we labeled the report's effect as statistically significant during that quarter. Monthly data from New York did not require averaging by quarter, because all facilities were large, high-volume centers.

The Durbin-Watson D statistic suggested significant first order autocorrelation of residual volumes in several OLS models based on California data. We therefore estimated time-series models using autoregressive integrated moving average (ARIMA) methods, with $p=1$, $q=0$, and the same main effects. Due to software limitations, we were unable to include two-way interactions. Second-order autoregressive models were also tested, but generated very similar results. Because the residual autocorrelations were nonsignificant ($p>0.10$) for all but one model, we did not estimate moving average models. The ARIMA and OLS results differed somewhat, so we report both sets of numbers to demonstrate the sensitivity of our results to different statistical methods. We report only OLS results from New York, because autocorrelation was minimal in that state. SAS was used for all analyses.

RESULTS

Numbers of outlier hospitals

The 1992, 1993, and 1995 CSRS reports labeled one, one, and two hospitals as having fewer CABG deaths than expected ($p<.025$), respectively. The same set of reports labeled three, one, and three hospitals as having more CABG deaths than expected ($p<.025$), respectively. Each hospital's study period was constructed around the publication date of the first report in which it was identified as a performance outlier.

The 1993 CHOP report labeled 10 hospitals as having fewer AMI deaths than expected ($p<.01$) and 12 hospitals as having more AMI deaths than expected ($p<.01$), using either of two different risk adjustment models. For cervical diskectomy complications, there were 6 favorable ($p<.005$) and 7 unfavorable ($p<.005$) outliers. For lumbar diskectomy complications, there were 32 favorable ($p<.005$) and 27 unfavorable ($p<.005$) outliers.

Changes in California hospital volume

Hospitals labeled as having fewer AMI deaths than expected experienced significantly increased volume in the third and fourth quarters after publication according to the OLS model, but not according to the ARIMA model (Table 13). The average such hospital was predicted by OLS to admit 76.9 (95% CI, 67.9-85.9) AMI patients during this 6-month period, but actually admitted 90.4 AMI patients, for an 18% increase. This effect was not seen for AMI-related admissions; indeed, the OLS model (but not the ARIMA model) suggested fewer AMI-related admissions during the first quarter after publication. Hospitals labeled as having more AMI deaths than expected did not experience significantly decreased volume of either AMI or AMI-related admissions, by either OLS or ARIMA models.

For cervical diskectomy and diskectomy-related conditions and procedures, there were also no consistent trends in hospital volume after publication of a report card. Hospitals labeled as having fewer complications than expected after cervical diskectomy had no significant deviations from expected volume, by either OLS or ARIMA. Hospitals labeled as having excess complications after cervical diskectomy operated upon slightly fewer patients during the first quarter after publication, and slightly more patients during the third quarter, but only using OLS methods. Although some volume changes for diskectomy-related admissions were statistically significant by OLS, most were in the opposite direction from what we had hypothesized. Only for lumbar diskectomy did we find consistent evidence of increased volume at hospitals lauded for their low complication rates. Although these volume differences were numerically consistent across both models and quarters, they were statistically significant only by ARIMA and never exceeded one patient per month for the average hospital.

The stratified analyses are somewhat difficult to interpret due to the large number of comparisons. Although the ARIMA and OLS results differed, the significant increase in AMI volume during the third and fourth quarters after publication of a favorable report card was most

pronounced among patients over 64 years of age, with HMO or Medicare coverage, and of white race (Table 14). Among HMO patients, there were also significant increases in AMI-related volume (data not shown) during each of the first four quarters after publication of a favorable AMI report card (1.40, 2.81, 1.99, and 2.15 additional cases per month in quarters 1 to 4 by ARIMA; 1.66, 2.33, 1.24, and 1.45 additional cases per month in quarters 1 to 4 by OLS). Such increases were not observed in any other insurance category; indeed, there was a significant shift of uninsured AMI-related patients (0.53, 0.65, 0.67, and 0.69 additional cases per month by ARIMA; 1.45, 1.44, 0.56, and 0.43 additional cases per month by OLS) toward hospitals that were rated poorly in the AMI report card. We found inconsistent evidence of increased AMI-related volume among white patients during the second and third quarters after publication of a favorable AMI report card (0.00, 3.57, 3.36, and -0.58 additional cases per month by ARIMA; -2.44, -0.26, -1.06, -0.45 additional cases per month by OLS).

Stratified analyses of lumbar diskectomy volume revealed few clear patterns. However, volume changes after publication of a favorable report card were most consistent among patients less than 65 years of age, and among African-American or white patients (Table 15). We found no spillover of these modest effects to diskectomy-related admissions (data not shown). However, cervical diskectomy patients with indemnity insurance shifted significantly toward hospitals with high complication rates during the second through fourth quarters after publication of an unfavorable report card (0.39, 1.37, 1.79, 1.16 additional cases per month by ARIMA; 0.10, 0.39, 0.95, and 0.66 additional cases per month by OLS), while HMO patients shifted significantly away from such hospitals (-0.85, -0.86, -0.63, -0.57 additional cases per month by ARIMA; -0.92, -0.76, -1.01, -1.20 additional cases per month by OLS) toward hospitals that were lauded for low complication rates (0.90, 0.45, -0.12, -0.09 additional cases per month by ARIMA; 1.59, 1.02, 0.98, 0.93 additional cases per month by OLS), especially in the first two quarters after public release. Stratification by age and race/ethnicity did not reveal any consistent effects on cervical diskectomy volumes.

Changes in New York hospital volume

New York hospitals labeled as having fewer CABG deaths than expected experienced significantly increased volume in the first month after publication. The average such hospital was predicted to admit 61.1 CABG patients during this month, but actually admitted 74.5 (a 22% increase). Over the first 6 months after a favorable report was published, each of these hospitals admitted an additional 24.4 CABG patients. This effect was not seen for any of three categories of CABG-related admissions (e.g., PTCA, CHF, AMI). Hospitals labeled as having more CABG deaths than expected experienced significantly decreased volume in the second month after publication. The average such hospital was predicted to admit 67.8 CABG patients in the first two months after an unfavorable report was published, but actually admitted 56.7 (a 16% decrease). This effect was much less prominent for CABG-related admissions. All of the report card effects disappeared within 3 months after publication.

In stratified analyses, the CABG volume changes that followed publication of a report card were generally consistent across all age groups, except those over 74 years of age. However, most of the volume changes occurred among Medicare patients (41.3 observed versus 32.8 predicted at

low-mortality outliers during the first month after release; 23.5 observed versus 31.8 predicted at high-mortality outliers during the first 2 months after release). Nearly all of the CABG volume changes occurred among patients of white and other race; African-American and Hispanic patient volumes were apparently not affected by designation of a hospital as a CABG mortality outlier.

DISCUSSION

In this study, we attempted to estimate changes in patient volumes that followed publication of hospital report cards in California and New York. Although we identified some statistically significant effects, they were generally of modest size and transient duration. In California, where time-series data showed significant autocorrelation, our estimates of report card effects from autoregressive models were often smaller and less statistically significant than our estimates from linear regression models. The effects observed in New York were larger, but lasted for only about two months after each public release. We did, however, find some support for our hypothesis of heterogeneous effects across population strata. Specifically, the observed volume shifts were largely limited to white patients in both states, and were greater among HMO patients in California than among patients with indemnity insurance, Medicaid, or no insurance. Only HMO patients demonstrated a clear spillover effect, with increased volume for AMI-related conditions and procedures at hospitals that were lauded for low AMI mortality.

These findings generally confirm those of previous studies,[37,48,49] in that the publication of risk-adjusted hospital report cards appears to have relatively small and transient effects on consumer behavior, as measured by hospital volumes. This study extends previous work by showing that report cards about an emergency diagnosis (AMI) have had little effect on hospital volumes for elective cardiac procedures in California, and vice versa in New York. In addition, this study is the first to demonstrate that volume changes may be limited to relatively advantaged populations, including white patients and those with Medicare or HMO coverage. This finding supports recently expressed concerns that outcomes reporting may not benefit, or may even hurt, vulnerable populations because of their inability to understand or act upon comparative performance information.[88]

There are several possible explanations for these findings. First, consumers may not have known about the ratings of local hospitals when they had to make their choices. Although considerable publicity surrounded each release, this publicity generally lasted only a few days, and may not have reached the population at risk.[81] Indeed, only 20% of respondents to a telephone survey in California reported ever having seen or heard comparative information about hospitals.[85] Anecdotal press accounts of single, unexpected deaths may be more salient to consumers than abstract performance data.[48] Second, even if consumers did remember the last hospital report card, they may have received conflicting advice from more trusted sources, such as their physicians,[54] friends, or family members. Third, even in the absence of such advice, consumers may have misinterpreted mortality data[89] or dismissed the report cards as incomplete, biased, out-of-date, or inapplicable to their own circumstances. For example, the 42-63% of consumers who believe that there are “small” or “no” differences in quality across hospitals[83,84,85] would probably attribute any variation in outcomes to variation in patient risk. Fourth, even if they believed the report card results, consumers may not have been able to act upon these beliefs because of

established relationships with local physicians and hospitals, preferred provider panels, geographic constraints, and other limitations on access to the highest-rated hospitals.[82] More generally, consumers may prefer to entrust decisions regarding hospital care to family members, emergency medical personnel, or other health professionals. Finally, older studies suggest that patients may be selectively referred to better hospitals, even in the absence of public reporting.[29] If so, then public reporting may have little impact because the market has already identified and rewarded the hospitals that provide better quality of care.

The major strength of this study is that we used the two largest statewide databases to assess the impact of two of the most prominent hospital performance reporting initiatives in the US. This design allowed us to compare the impact of report cards based on different data sources and different methodologies, to explore variation across population strata, and to test the robustness of our findings with a variety of modeling approaches. In addition, this study fits within a larger body of work (summarized elsewhere in this report) that describes how publicly reported hospital performance data are used and interpreted by hospital administrators, hospital quality improvement leaders, health plan executives, and the news media. Together, these studies help us to understand how all of the target groups use and respond to hospital report cards.

Our study has several weaknesses. First, the relatively small number of hospitals identified as outliers limited our ability to find statistically significant effects, especially in analyses stratified by age, race/ethnicity, insurance, or area characteristics. The 95% confidence intervals for predicted volume were reasonably narrow in our overall analyses, but widened substantially in stratified analyses. In addition, we were better able to detect consistent effects for high volume conditions and procedures, such as CABG in New York and AMI in California, than for lower volume procedures, such as cervical diskectomy. Second, we were only authorized to use data for up to 12 months after the publication of each report card. This interval seems appropriate for report cards that are intended for annual release. However, it precluded looking for very delayed or cumulative effects, especially for hospitals that repeatedly received either excellent or poor ratings. Finally, our predictors of monthly hospital volume may be inadequate surrogates for the underlying factors of interest:(1) the prevalence of a condition or procedure in the population at risk for admission to a given hospital, and (2) the relative attractiveness of each hospital to which a patient could go. Both of these factors could change over time, but our ability to capture these changes was limited by the variables available to us.

The major policy implication of our findings is that policy-makers and "smart purchasing" advocates should not expect hospital report cards to produce dramatic shifts in volume. Any shifts that occur may be limited to the sociodemographic groups that are best able to understand and act upon relatively complex information. On the other hand, there is evidence that at least some consumers (or the managed care organizations that purchase health care on their behalf) attend to hospital report cards. These consumers or their surrogate decision-makers may be numerous enough to drive certain local markets toward higher quality. The major unanswered question from this and prior studies is whether public reporting makes providers more responsive to performance data, and therefore leads to greater subsequent improvements in quality of care, than nonpublic dissemination of report cards or league tables. Fear of public embarrassment and professional competitiveness may be important motivations for providers to avoid a low

performance rating,[90] even if such a rating would be unlikely to cause a significant loss of business.

Table 12. Definitions of target conditions and procedures, related conditions and procedures, and unrelated admissions in New York and California.*

State	Condition or Procedure	Defining Variable	Definition
CA	AMI (target)	Principal diagnosis	410.x1, 410.x2 (or 426.0, 427.1, 427.41, 427.42, 427.5, 429.5- 429.7x, 429.81, 518.4, 780.2, 785.51 with a secondary diagnosis of AMI)
CA	CABG (AMI-related)	Any procedure	36.1x (without principal diagnosis of AMI)
CA	Percutaneous coronary angioplasty (AMI-related)	Any procedure	36.01-36.02, 36.05 (without principal diagnosis of AMI or any procedure of CABG)
CA	Congestive heart failure (AMI-related)	Principal diagnosis	428.x, 402.x1, 404.x1, 404.x3, 398.91 (without principal diagnosis of AMI or any procedure of CABG or PTCA)
CA	Cervical diskectomy (target)	Any procedure	80.51 (with an associated condition or procedure code indicating cervical level)
CA	Lumbar diskectomy (target)	Any procedure	80.51 (with an associated condition or procedure code indicating lumbar level)
CA	Back or neck procedures (diskectomy-related)	DRG	214, 215 (without any procedure of diskectomy)
CA	Medical back problems (diskectomy-related)	DRG	243
CA	Knee arthroplasty (diskectomy-related)	Any procedure	81.54, 81.55 (without any procedure of hip arthroplasty or diskectomy)
CA	Hip arthroplasty (diskectomy-related)	Any procedure	81.51-81.53 (without any procedure of diskectomy)

NY	CABG (target)	Any procedure	36.1x
NY	AMI (CABG-related)	Principal diagnosis	410.x1, 410.x2 (without any procedure of CABG)
NY	Percutaneous coronary angioplasty (CABG-related)	Any procedure	36.01-36.02, 36.05 (without any procedure of CABG or principal diagnosis of AMI)
NY	Congestive heart failure (CABG-related)	Principal diagnosis	428.x, 402.x1, 404.x1, 404.x3, 398.91 (without any procedure of CABG or PTCA, or principal diagnosis of AMI)

* Unrelated admissions were defined by exclusion as all other acute care admissions among persons at least 18 years of age, excluding psychiatric conditions (ICD-9-CM principal diagnosis 290.xx-319), injury or poisoning (ICD-9-CM 800.xx-995.xx), and rehabilitation (ICD-9-CM V57.x).

Table 13. Mean difference between actual and predicted monthly patient volume for the average outlier hospital, over four consecutive quarters (in California) or months (in New York) after publication of a risk-adjusted outcome study, by model (ordinary least squares versus autoregressive integrated moving average)*

State	Condition or procedure	Outlier group ‡	Actual minus predicted monthly patient volume							
			Quarter 1		Quarter 2		Quarter 3		Quarter 4	
			OLS	ARIMA	OLS	ARIMA	OLS	ARIMA	OLS	ARIMA
CA	AMI (target)	Better (D=0.93)	1.00	1.93	-0.74	-1.14	**1.66†**	-0.56	**2.84†**	1.12
		Worse (D=1.15)	1.30	0.67	0.25	1.04	-0.55	0.03	0.15	0.65
CA	AMI-related	Better (D=2.26)	**-3.83†**	-1.10	-0.60	4.19	-0.23	3.75	0.44	-0.07
		Worse (D=1.78)	**2.36†**	1.04	0.29	0.36	0.52	-0.38	0.98	-0.97
CA	Cervical diskectomy (target)	Better (D=2.71)	0.04	0.22	0.13	-0.30	-0.55	-1.61	-0.66	-0.59
		Worse (D=1.47)	**-1.14†**	-0.97	-0.57	0.34	**1.38†**	1.07	0.53	0.86
CA	Lumbar diskectomy (target)	Better (D=1.90)	0.60	**0.58†**	0.61	0.30	0.68	0.52	0.61	**0.78†**
		Worse (D=1.21)	-0.18	-0.13	-0.22	-0.13	-0.33	-0.31	-0.52	-0.52
CA	Diskectomy-related	Better (D=1.98)	0.47	0.36	-0.43	-0.87	**-1.13†**	-1.15	-1.05	0.36
		Worse (D=1.42)	-0.75	**-1.37†**	**1.08†**	0.18	**1.13†**	0.03	**1.19†**	0.24
NY	CABG (target)	Better (D=1.92)	**13.45†**		5.55		6.73		2.96	
		Worse (D=1.91)	-4.04		**-7.11†**		-2.66		-0.93	
NY	CABG-related (AMI)	Better (D=1.96)	-4.93		-1.44		-1.95		0.55	
		Worse (D=1.38)	**-4.53†**		-1.24		-1.61		**-6.00†**	
NY	CABG-related (PTCA)	Better (D=2.14)	3.75		1.12		0.60		-1.15	

		Worse (D=1.34)	-2.62	-1.43	0.36	-2.07
NY	CABG-related (CHF)	Better (D=1.74)	-2.81	-3.97	-0.52	-1.72
		Worse (D=2.14)	-0.98	-1.97	-1.73	-0.07

* Positive numbers indicate that hospitals in that category had more admissions than predicted; negative numbers indicate that hospitals in that category had fewer admissions than predicted. To estimate the total difference in patient volume for the average hospital in each quarter, the numbers shown should be multiplied by three.

† $p<0.05$

‡ The Durbin-Watson statistics in this column are based on the OLS models. All Durbin-Watson statistics for the first-order ARIMA models were between 1.71 and 2.14, and none were statistically significantly different from 2.

Table 14. Mean difference between actual and predicted monthly patient volume for the average California hospital lauded for its low risk-adjusted AMI mortality, over four consecutive quarters after publication of the report card, stratified according to relevant patient characteristics and model (ordinary least squares versus autoregressive integrated moving average)*

Patient characteristic	Stratum	Actual minus predicted monthly patient volume							
		Quarter 1		Quarter 2		Quarter 3		Quarter 4	
		OLS	ARIMA	OLS	ARIMA	OLS	ARIMA	OLS	ARIMA
Age	<55 years	0.07	0.18	0.31	0.10	0.04	-0.39	0.41	0.19
	55-64 years	0.01	-0.20	-0.08	**-1.24†**	-0.07	**-1.14†**	**0.65†**	-0.47
	65-74 years	0.75	**0.94†**	-0.20	-0.16	0.81	**1.19†**	0.71	0.59
	>74 years	0.21	0.27	-0.25	0.06	1.13	0.28	**2.27†**	**1.76†**
Insurance	Commercial indemnity	**0.69†**	0.51	**0.64†**	0.02	0.40	-0.51	**0.82†**	0.01
	HMO/PPO	**0.75†**	0.60	-0.15	-0.47	**1.62†**	**0.90†**	**2.34†**	**0.88†**
	Medicaid	‡	0.04	‡	0.07	‡	-0.15	‡	0.33
	Medicare	-0.12	0.63	-0.49	-0.01	1.09	0.37	**1.18†**	1.18
	None/self-pay	0.09	0.07	0.10	0.14	0.25	0.07	**0.46†**	0.05
Hospital catchment area	Outside#	0.09	0.28	-0.19	0.08	-0.15	-0.35	**0.86†**	**1.41†**
	Inside#	0.69	1.69	-0.57	-1.24	**1.63†**	-0.28	**2.00†**	-0.27
Race/ethnicity	African-American	-0.14	-0.14	-0.24	0.07	-0.13	0.29	0.25	0.00
	Hispanic	-0.13	-0.18	0.00	-0.18	0.14	0.21	**0.48†**	0.12
	White	1.05	1.48	-0.34	-0.70	**1.94†**	0.07	**2.50†**	1.45

* Positive numbers indicate that hospitals in that category had more admissions than predicted; negative numbers indicate that hospitals in that category had fewer admissions than predicted. To estimate the total difference in patient volume for the average hospital in each quarter, the numbers shown should be multiplied by three.

† $p<0.05$

‡ These OLS models could not be estimated, because of data limitations.

Hospital catchment areas were defined using the first method described in the Methods; the alternative method generated similar but generally smaller effects.

Table 15. Mean difference between actual and predicted monthly patient volume for the average California hospital lauded for its low risk-adjusted postoperative complication rate after lumbar diskectomy, over four consecutive quarters after publication of the report card, stratified according to relevant patient characteristics and model (ordinary least squares versus autoregressive integrated moving average)*

Patient characteristic	**Stratum**	**Actual minus predicted monthly patient volume**							
		Quarter 1		**Quarter 2**		**Quarter 3**		**Quarter 4**	
		OLS	ARIMA	OLS	ARIMA	OLS	ARIMA	OLS	ARIMA
Age	<55 years	0.45	0.08	0.39	0.05	**0.86†**	0.48	**0.78†**	**0.75†**
	55-64 years	0.24	0.23	0.19	0.07	**0.24†**	0.00	0.23	-0.20
	65-74 years	0.08	0.00	0.20	-0.05	-0.01	-0.13	**0.28†**	0.17
	>74 years	-0.05	0.00	0.05	0.04	-0.04	-0.09	**0.33†**	-0.02
Insurance	Commercial indemnity	-0.00	-0.19	0.31	-0.01	0.39	0.00	0.30	-0.05
	HMO/PPO	0.42	0.16	-0.19	-0.40	-0.32	-0.42	-0.40	-0.07
	Medicaid	‡	0.01	‡	**-0.17†**	‡	0.06	‡	-0.01
	Medicare	0.10	0.10	0.14	0.02	0.07	-0.06	**0.44†**	0.26
	None/self-pay	0.20	0.03	0.18	0.00	**0.14†**	0.02	**0.38†**	0.00
Hospital catchment area	Outside#	0.43	0.38	0.30	-0.01	0.39	0.21	0.41	0.09
	Inside#	0.21	0.22	0.38	0.34	0.35	0.34	**0.48†**	**0.71†**
Race/ethnicity	African-American	0.13	0.08	**0.16†**	**0.14†**	**0.16†**	0.02	**0.58†**	**0.20†**
	Hispanic	-0.01	0.03	-0.02	-0.02	0.08	-0.01	**0.39†**	0.08

White	**0.74†**	0.48	0.68	0.14	**0.83†**	0.41	0.56	0.53

* Positive numbers indicate that hospitals in that category had more admissions than predicted; negative numbers indicate that hospitals in that category had fewer admissions than predicted. To estimate the total difference in patient volume for the average hospital in each quarter, the numbers shown should be multiplied by three.

† $p<0.05$

‡ These OLS models could not be estimated, because of data limitations.

Hospital catchment areas were defined using the first method described in the Methods; the alternative method generated similar but generally smaller effects.

STUDY 5: Content analysis of media coverage of hospital report cards in California and New York

As information on the performance of health care organizations becomes a higher priority for government agencies, purchasers, and consumers, the manner in which this information is conveyed takes on increased importance. A leading mechanism by which consumers learn about hospital performance is through the mass media, which play a critical gatekeeping role in communicating and filtering complex technical information. As Rudd and Glanz noted,[30] "journalistic decisions on what information to publish, how to slant it, and who to consult for interpretation or reaction, set the tone and boundaries of what reaches consumers." Hospital report cards have received attention from both the broadcast and print media since they were first released by the Health Care Financing Administration (now the Center for Medicare and Medicaid Services) in the mid-1980's, but press coverage of these report cards has not been systematically evaluated for a decade.[30]

The California Hospital Outcomes Project (CHOP) uses ICD-9-CM coded hospital discharge abstracts, routinely collected by California's Office of Statewide Health Planning and Development (OSHPD), to develop disease-specific, customized risk-adjustment models. The first report, released to the public in December 1993, classified inpatient mortality rates for acute myocardial infarction (AMI) and complication rates for cervical and lumbar diskectomy into two categories: "better" or "not better" than expected.[13] The second report, released to the public in May 1996, classified AMI mortality rates into three categories: "better," "worse," and "not significantly different" than expected.[56] The third report, which also focused on AMI mortality, was released to hospitals in July 1997 and to the public in December 1997.[16] For each cycle, OSHPD issued a press release and a media packet with advance copies of each report (embargoed for several days to allow ample time for questions and interviews), and held a press conference at which OSHPD staff and consulting researchers described the study and answered questions. Reporters were encouraged to interview local physicians and hospital administrators for perspectives on both low and high-performing hospitals.

In this paper, we analyze the amount and type of media coverage of California Hospital Outcomes Project (CHOP) reports in 1993, 1996, and 1997. Our goals are to describe the extent and content of newspaper coverage of CHOP reports, to explore how this coverage has changed over time, and to contrast it with coverage of earlier Federal reports on hospital quality. What are consumers being told? What findings or data are being emphasized? Whose views receive the most attention? How balanced has the coverage been? Has this coverage become more prominent, more in-depth, or more balanced over time? (A parallel analysis of media coverage in New York is nearly complete, but is not included in this report)

METHODS

Sample

We searched newspapers, newspaper clipping services and online databases to collect reports in the print media about the California Hospital Outcomes Project and its public reports. Most

articles were found using Lexus-Nexus, a reference service that contains full-text articles from major media outlets nationwide. Major newspapers, wire service articles, major magazines, newsletters, and smaller newspapers were included in our search, while television transcripts were excluded. Additional articles were acquired from the public relations unit of the Office of Statewide Health Planning and Development (OSHPD), which subscribes to a newspaper clipping service. We also obtained copies of all major daily newspapers in California for several days after the embargo date for each public report, and searched manually for articles that were not in the Lexus-Nexus database. An article was selected for review if it discussed or described the public report during one of the three publicity periods (December 1993, May 1996, December 1997). Articles that described the project in general or its history, without reference to specfic performance data, were not eligible. Using these criteria, a total of 96 articles were selected for analysis. Table 1 shows the number and type of articles included in the final sample.

Coding

All newspaper articles were identified by title, publication, and date. In addition, the following information was recorded for each article: (1) total number of words; (2) page number and section in which the article began; (3) headline valence toward CHOP in general and toward specific hospital(s); (4) name and number of hospitals mentioned in the text; (5) public performance evaluation of each hospital mentioned (better than expected, worse than expected, or neither better nor worse); (6) number of words in quotation marks and the source of each quote; (7) themes expressed in quotations.

A trained research assistant and both authors independently reviewed each article. All articles were first coded by a research assistant and a randomly selected subsample of 10 articles in each of the three years of reporting were separately coded by the first author. Agreement rates between coders were generally high for length of article, placement, headline valence, hospital mentions, quote sources, and content themes (95% observed agreement, kappa= 0.89). All disagreements in the subsample were resolved prior to analyses.

Measurements

Extent of news coverage. The length and placement of articles are measures of the perceived significance of a news story, and a newspaper's level of interest in it. We measured article length in words, not lines or column inches, to facilitate comparisons across hard copy and electronic formats. Longer articles were presumed to reflect greater perceived significance and interest. We measured article placement by the page and section of the newspaper on which the article began. Placement of the front page of the main news section was presumed to reflect the greatest perceived significance, followed by the front page of the local news section, the interior of the main news section, and finally elsewhere in the newspaper.

Headline valence. Headline valence refers to the bias or tone of the headline. We coded each headline as positive, negative, or neutral with respect to the study in general, and also with respect to local hospitals. Positive headlines emphasized positive aspects of the report or the favorable performance of specific hospitals. For example, the headlines "3 Area Hospital Lauded" and "French Places High in Cardiac Care" were coded as positive valence. Negative

headlines noted negative aspects of the report or the unfavorable performance of specific hospitals. For example, the headlines "Local Doctors Dispute Data on Heart Attacks" and "Heart Attack Report Dings 2 Orange County Centers" were coded as negative valence. Neutral headlines emphasized neither positive nor negative aspects of the public release, such as "State Rates Hospitals on Heart Attack Care" or "Most Hospitals Earn 'B' Saving Cardiac Victims."

Hospital mentions. We recorded the name and outlier status of each hospital mentioned in each article, not counting any associated graphs or tables. The hospital's outlier status was obtained from the relevant public report. The number of mentions of hospitals with better or worse than expected outcomes, compared to the number of mentions of hospitals with neutral outcomes, is a measure of whether newspapers view their primary role as providing news or information. Overrepresentation of outlier hospitals was presumed to reflect a focus on providing news rather than information.

Quote sources. We counted the number of quoted words, defined as words within quotation marks that could be attributed to an individual speaker, by source:

1. Hospital spokesperson or representative, including the CEO, public relations staff, medical leaders, or quality improvement leaders. Quotes in this category had to be attributed to a representative of a specific hospital mentioned in the article.
2. Government representative, such as staff from OSHPD or other state or Federal agencies. Direct quotes from the OSHPD report were not counted.
3. Hospital, provider, or insurance industry representatives, such as the California Healthcare Association or the Healthcare Association of Southern (or Northern) California (representing hospitals), the California Medical Association or local/county medical associations (representing physicians), the Society of Thoracic Surgeons, and health plan executives.
4. Consumer, health advocacy, or purchaser groups, such as Consumers First, the American Heart Association, and the Pacific Business Group on Health.
5. Researchers or college professors, including the contract researchers at the University of California who analyzed the hospital outcomes data on behalf of OSHPD.

The relative frequency of quotes from various sources is a measure of how a story is viewed by the newspaper and how it will likely be interpreted by readers. The quoted sources were presumed to reflect who reporters believed to be the major and peripheral players in a story.

Content themes in hospital quotations. Reporters often interview individuals in organizations that are affected by or mentioned in hospital report cards, and offer them an opportunity to expain or comment upon their performance. The direct quotes of all hospital representatives were coded as belonging to none, or one or more, of three overlapping content areas:

1. Flawed study or methods. Examples include "the report is misleading because it simplifies too many variables" or "to give the patient just one number is misleading".
2. Data are too old to be useful. For example, "it doesn't reflect what is going on now."
3. Explanations for poor performance. These interviewees apologized for or tried to remove blame for worse-than-expected results. Examples include "we are seeing sicker patients" and "once they have told us they want 'do not resuscitate' orders, then there is nothing we can do to prevent those deaths." This theme was often connected with the first (for example, "a lot of the patients that we were seeing were of advanced age… and that was not factored as part

of the results"), but comments coded to this theme emphasized particular problems that affected the interviewee's hospital rather than general defects in the report.

Statistical analyses. We tested for temporal trends in categorical characteristics of newspaper articles using the chi square for trend. We tested for temporal trends in continuous variables using analysis of variance. Although some articles appear to have been based on wire service stories or articles published in more prominent newspapers, we were unable to assess this phenomenon accurately enough to adjust for the non-independence of some observations.

RESULTS

Extent of News Coverage. The number of retrieved newspaper articles about CHOP reports increased from 20 in 1993 to 39 in 1996 and 37 in 1997 (Table 16). The mean length of these articles increased modestly from 642 words in 1993 and 514 words in 1996 to 786 words in 1997. Between 1993 and 1997, articles about CHOP reports received increasingly prominent placement. About 67% of the articles describing the 1996 or 1997 release appeared on the front page of the main or local news section, compared with 45% following the 1993 release. In 1997, 43% of the articles appeared on the front page of the main news section, 19% elsewhere in the main news section, and 24% on the front page of the local news section.

Headline Valence. Following all three releases, at least 65% of headlines either focused on the performance of specific hospitals or were neutral toward the CHOP study in general (Table 17). Negative headlines toward the study in general declined nonsignificantly from 25% in 1993 and 33% in 1996 to 14% in 1997. We also examined whether headlines emphasized favorable (i.e., better than expected mortality), unfavorable (i.e., worse than expected mortality), or neutral hospital performance (i.e., neither better nor worse than expected mortality). The percentage of headlines emphasizing unfavorable hospital performance increased significantly from 25% in 1993 and 26% in 1996 to 51% in 1997.

Hospital Mentions. Most articles mentioned the name of at least one hospital. The percentage of articles mentioning one or more specific hospitals increased marginally from 79%-80% after the first two reports to 95% after the 1997 report. The mean number of hospitals mentioned per article was 5.9 in 1993, 4.3 in 1996 and 5.6 in 1997. Table 18 displays the outlier status of mentioned hospitals as a percent of the total number of mentioned hospitals. Following each public release, 56% to 58% of all mentioned hospitals had either better or worse than expected performance. Note that "worse than expected" performers were not identified in the 1993 report. The distribution of mentioned hospitals across these performance categories differed markedly from the distribution of all hospitals (Table 20). In 1996 and 1997, there was a tendency for articles to mention "worse than expected" hospitals more often than "better than expected" hospitals. About 32% of all mentioned hospitals had "worse than expected" ratings and 24% had "better than expected" ratings, whereas the percentage of outlier hospitals was about 6.5% at each end in 1996 and 1997.

Quote Sources. As shown in Table 19, the majority of direct quotes were from either hospital or government sources following all three public releases. The percentage of articles that quoted leaders of specific hospitals increased significantly from 60% in 1993 and 41% in 1996 to 81%

in 1997. The percentage of articles that quoted industry representatives fell significantly from 45% to 18% to 14%, while the percentage of articles that quoted consumer representatives fell nonsignificantly from 20% to 5% and 11%. About 60% of articles after each report quoted government staff.

The overall percentages of directly quoted words were approximately balanced between hospital and government representatives after each of the three releases. However, hospital representatives were quoted somewhat more intensively over time, as they accounted for 32% of all quoted words in 1993, 35% in 1996, and 42% in 1997. Provider industry representatives were quoted less intensively over time, as they accounted for 21% of all quoted words in 1993, 12% in 1996, and 3% in 1997. Quotes from researchers comprised 5% of quoted words in 1993, but 15% of quoted words in 1997.

Representatives of hospitals that were identified in the report as "worse than expected" outliers were quoted disproportionately in newspaper articles; this phenomenon has not changed over time. As shown in Table 20, 17 hospitals (4%) were classified as "better than expected" and 22 (6%) as "worse than expected" in 1996. Representatives of the 17 "better" hospitals accounted for 9% of direct quotes from hospital representatives, whereas representatives of the 22 "worse" hospitals accounted for 43%. Similarly, after the 1997 release, 35 "better" hospitals represented 8% of all hospitals listed in the report and accounted for 14% of direct quotes from hospital representatives. The 31 "worse" hospitals (7%) identified in the report accounted for 47% of all direct quotes from hospital representatives. In summary, hospitals labeled as mortality outliers were disproportionately represented in newspaper articles. These hospitals represented fewer than 16% of all acute care facilities in 1996 and 1997, but accounted for 56% of all mentioned hospitals and 59% of all quoted hospital representatives.

Content Themes in Hospital Quotations. Table 20 displays the frequency of different content themes by year. In 1993, the most common content theme was "flawed study" (40%). In the face of worse than expected outcomes, hospital representatives were apt to criticize some aspect of the CHOP study itself. The blame was often laid on a methodological or statistical flaw, typically described in very general terms. Some examples of this theme are:

"The report is misleading because it simplifies too many variables. Hospitals code records differently, but the state treats them all the same. We don't give this a lot of credence."

"There's so much that doesn't meet the eye. To give a patient just one number is misleading."

"We feel that it is seriously flawed and not worthy of consideration."

"We know there are things counted in there (as problems) that should not be. We think we're doing better than projected."

"The study was flawed in that the sample was small, it did not include patients who signed 'do not resuscitate' cards (sic) and it did not take age into consideration."

"It is not a bona-fide study. They took numbers and interpreted data from them. The study didn't even get the hospital's name right..."

In later years, hospitals were significantly more likely to focus their criticisms on the age of the data (i.e., 43% in 1997 versus 5% in 1993 and 13% in 1996), thereby avoiding broad rejection of the study methods. Typical comments included:

"It doesn't reflect what's going on now. In terms of quality of care, there is no comparison."

"We feel the report doesn't reflect our heart attack mortality accurately. Before 1993, clot-dissolving drugs weren't the norm in any ER. But since 1994, they're the first thing used."

Hospitals also became more likely over time to offer specific explanations or excuses for their failure to show "better than expected" outcomes (i.e., 43% in 1997 versus 15% in 1993 and 21% in 1996). Quotes of this type often attributed poor apparent performance to the illness level of patients, even though the report's results were risk adjusted. For example,
"We are seeing sicker patients".
"A lot of the patents that we were seeing here were of advanced age…"

Several hospital representatives correctly noted that patients with do-not-resuscitate orders were included in the analysis and that patients were sometimes admitted from nursing homes with such orders in place. For example, "once they have told us they want 'do not resuscitate' orders, then there is nothing we can do to prevent those deaths. The fact that this facility serves an older population has much to do with our rating."

DISCUSSION

Overall, hospital outcomes reports in California have received increasing attention from newspapers over time, with slightly longer and more prominently placed articles. Reporters became more likely to interview local hospital representatives, and less likely to rely on press releases or "canned statements" from industry representatives. Both the headlines and the content of these articles, across all three years, overrepresented hospitals with better or worse than expected ratings and underrepresented those within the expected range. The disproportionate number of hospital mentions of, and direct quotes from representatives of, "worse than expected" outliers indicates that newspapers have tried to create "news stories" out of CHOP reports instead of just conveying potentially useful information. We noted a gradual shift between 1993 and 1997 in the responses of hospital representatives, from broad criticism of the report and its methods to more specific justification of poor performance and focused criticism of the lengthy publication delay.

Our findings are generally consistent with those of Rudd and Glanz, who analyzed newspaper articles published after the release of Health Care Financing Administration's 1987 Medicare mortality report.[30] They reported more hospital mentions per article than we did (i.e., 13.8 versus 4.3-5.9 in the present study), and more focus on hospitals with lower-than-expected mortality (i.e., 56% of mentioned hospitals versus only 3% of all rated hospitals). As in the present study, quotes from hospital and medical sources dominated the articles, and the number of such quotes was highly correlated with the length of the article. A major difference was that newspaper articles after the Medicare mortality report contained more vigorous criticism of the report, with 69% of articles quoting a hospital representative who blamed the study, and 29% quoting a hospital representative who denied "that the hospital was in any way to blame." Another study did not describe the content of newspaper articles about the Medicare mortality reports, but found that such articles had no apparent effect on hospital volume.[48]

This study is unique in examining the print media response to a later generation of hospital report cards, which are being produced by state agencies and private coalitions around the country. It is

also unique in examining temporal trends in print media coverage. However, two limitations should be noted. First, we may have missed some articles in newspapers that are not indexed in Lexus-Nexus, not reviewed by newspaper clipping services, and not available for purchase in California's major cities. Such articles may have appeared in weekly or small town newspapers, and would therefore be expected to have less statewide impact than the articles that we were able to retrieve. Second, we focused only on newspaper coverage of hospital report cards, because Lexus-Nexus contains more limited information about radio and television stories. Our anecdotal impression is that such coverage is very spotty, as very few broadcast reporters attended OSHPD's press conferences and the key government and academic representatives did fewer than three broadcast interviews after each release.

These findings suggest that newspaper reporters are becoming more sophisticated and appreciative readers of hospital outcome studies, but remain disproportionately attentive to which hospitals are labeled as bad outliers. They try to present the results fairly, providing ample opportunity for government and academic representatives to defend their work, and for hospital representatives to explain their reported performance. They invest the time and effort needed to interview key local stakeholders, although they tend to accept and report those comments uncritically. Key challenges for the near future are: (1) to increase the timeliness of hospital outcome studies, thereby diffusing what has emerged as the single most important public criticism in California; (2) to encourage reporters to view their role as providing information (i.e., publishing data on all local hospitals) and not simply reporting "news" about outliers; (3) to lengthen the period of media attention, so that hospital report cards are not forgotten within days after their relese; and (4) to increase broadcast reporting of the availability of hospital quality information.

Table 16. Number, type, length and placement of articles included in sample.

Article characteristic	Report release date			
	1993	1996	1997	All Years
Total number of articles	20	39	37	96
Number (percent) of articles appearing in daily newspapers	15 (75%)	29 (74%)	33 (89%)	77 (80%)
Number (percent) of articles appearing in main news section	10 (50%)	21 (54%)	23 (62%)	54 (56%)
Number (percent) of articles appearing on page one of main news section[a]	3 (15%)	13 (33%)	16 (43%)	32 (33%)
Number (percent) of articles appearing on page one of local news section	6 (30%)	13 (33%)	9 (24%)	28 (29%)
Mean number of words per article (standard deviation)	642 (323)	514 (211)	786 (375)	646
Total number of words (all articles)	12,849	20,055	29,074	

1. Sisk JE, Dougherty DM, Ehrenhaft PM, Ruby G, Mitchner BA. Assessing information for consumers on the quality of medical care. *Inquiry* 1990; 27:263-272.

2. Iezzoni LI, Shwartz M, Restuccia J. The role of severity information in health policy debates: A survey of state and regional concerns. *Inquiry* 1991; 28:117-128.

3. *Medicare Hospital Mortality Information, 1986*. Washington, DC: US Department of Health and Human Services; 1987. Pub. No. 01-002.

4. *Coronary Artery Bypass Surgery in New York State, 1992-1994.* Albany, NY: New York State Department of Health, 1996.

5. *Hospital Effectiveness Report*. Harrisburg, PA: Pennsylvania Health Care Cost Containment Council; 1993.

6. *Angioplasty in New York State, 1994.* Albany, NY: New York State Department of Health, 1996.

7. *Focus on Heart Attack in Pennsylvania, Research Methods and Results*. Harrisburg, PA: The Pennsylvania Health Care Cost Containment Council, 1996.

8. *Iowa Hospital Resource & Outcome Report, January 1992 through December 1992*. Des Moines, IA: Iowa Health Data Commission; 1993.

9. Longo DR, Land G, Schramm W, Fraas J, Hoskins B, Howell V. Consumer reports in health care: Do they make a difference in patient care? *JAMA* 1997;278:1579-1584.

10. Taulbee P. Florida agency readies hospital outcomes "report card." *Rep Med Guidelines Outcomes Res* 1994; 5(1):2,5-6.

11. *Hospital Morbidity and Mortality in Wisconsin, General Medical-Surgical Hospitals, 1990*. Madison, WI: Office of Health Care Information, Department of Health and Social Services; 1991.

12. *Obstetrical Services: A Consumer Guide.* Richmond, VA: Virginia Health Information; 1998.

13. *Annual Report of the California Hospital Outcomes Project, December 1993, Volume One: Study Overview and Results Summary*. Sacramento, CA: Office of Statewide Health Planning and Development, 1993.

14. *Cleveland-area Hospital Quality Outcome Measurements and Patient Satisfaction Report*. Cleveland, OH: Quality Information Management Corporation; 1993.

15. Marshall MN, Shekelle PG, Leatherman S, Brook RH. The public release of performance data. What do we expect to gain? A review of the evidence. *JAMA* 2000;283:1866-1874.

16. Zach AP, Romano PS, Luft HS, Rainwater JA. *Report on Heart Attack 1991-1993, Volume 1: User's Guide*. Sacramento, CA: California Office of Statewide Health Planning and Development, 1997.

17. Romano PS, Luft HS, Rainwater JA, Zach AP. *Report on Heart Attack 1991-1993, Volume 2: Technical Guide*. Sacramento, CA:

California Office of Statewide Health Planning and Development, 1997.

18. Kahn KL, Rogers WH, Rubenstein LV, Sherwood MJ, Reinisch EJ, Keeler EB, et al. Measuring quality of care with explicit process criteria before and after implementation of the DRG-based prospective payment system. *JAMA* 1990; 264:1969-1973.

19. Rubenstein LV, Kahn KL, Reinisch EJ, Sherwood MJ, Rogers WH, Kamberg C, et al. Changes in quality of care for five diseases measured by implicit review, 1981 to 1986. *JAMA* 1990; 264:1974-1979.

20. Hannan EL, Kilburn H Jr, O'Donnell JF, Lukacik G, Shields EP. Adult open heart surgery in New York State. An analysis of risk factors and hospital mortality rates. *JAMA* 1990; 264:2768-2774.

21. Thomas JW, Holloway JJ, Guire KE. Validating risk-adjusted mortality as an indicator for quality of care. *Inquiry* 1993; 30:6-22.

22. Hartz AJ, Gottlieb M, Kuhn EM, Rimm AA. The relationship between adjusted hospital mortality and the results of peer review. *Health Serv Res* 1993; 27:765-777.

23. Daley J, Forbes MG, Young GJ, et al. Validating risk-adjusted surgical outcomes: Site visit assessment of process and structure. *J Am Coll Surg* 1997;185:341-351.

24. Gibbs J, Clark K, Khuri S, Henderson W, Hur K, Daley J. Validating risk-adjusted surgical outcomes: Chart review of process of care. *Intl J Qual Health Care* 2001;13:187-196.

25. Hofer TP, Hayward RA. Identifying poor-quality hospitals: Can hospital mortality rates detect quality problems for medical diagnoses? *Med Care* 1996;34:737-753.

26. Zalkind DL, Eastaugh SR. Mortality rates as an indicator of hospital quality. *Hosp Health Serv Admin* 1997;42:3-15.

27. Iezzoni LI. The risks of risk adjustment. *JAMA* 1997;278:1600-1607.

28. Berwick DM, Wald DL. Hospital leaders' opinions of the HCFA mortality data. *JAMA* 1990; 263:247-249.

29. Johnsson J. CEOs: Mortality data no help to consumers. *Hospitals* 1990; 64(19):61.

30. Rudd J, Glanz K. A survey of newspaper coverage of HCFA hospital mortality data. *Public Health Rep* 1991; 106:517-523.

31. Podolsky D, Beddingfield KT. America's best hospitals. *US News World Rep*. July 12, 1993; 66-74.

32. Luft HS, Garnick DW, Mark DH, Peltzman DJ, Phibbs CS, Lichtenberg E, McPhee SJ. Does quality influence choice of hospital? *JAMA* 1990; 263:2899-2906.

33. Phibbs CS, Mark DH, Luft HS, Peltzman-Rennie DJ, Garnick DW, Lichtenberg E, McPhee S. Choice of hospital for delivery: A comparison of high-risk and low-risk women. *Health Serv Res* 1993; 28:201-222.

34. Luft HS. The relation between surgical volume and mortality: An exploration of causal factors and alternative models. *Med Care* 1980; 18:940-959.

35. Hughes RG, Garnick DW, Luft HS, McPhee SJ, Hunt SS. Hospital volume and patient outcomes. The case of hip fracture patients. *Med Care* 1988; 26:1057-1067.

36. Luft HS, Hunt SS, Maerki SC. The volume-outcome relationship: Practice-makes-perfect or selective referral patterns? *Health Serv Res* 1987; 22:147-182.

37. Vladek BC, Goodwin EJ, Myers LP, Sinisi M. Consumers and hospital use: The HCFA "death list." *Health Affairs* 1988; 7(1):122-125.

38. Zablocki E. Businesses share outcomes with employees. *Rep Med Guidelines Outcomes Res* 1992; 3(9):1,6-7.

39. Sardinha C. Hershey picks hospitals based on quality data from PaHC4. *Rep Med Guidelines Outcomes Res* 1992; 3(18):11.

40. Hannan EL, Arani DT, Johnson LW, Kemp HG Jr, Lukacik G. Percutaneous transluminal coronary angioplasty in New York State. Risk factors and outcomes. *JAMA* 1992; 268:3092-3097.

41. Hannan EL, Kilburn H Jr, Racz M, Shields E, Chassin MR. Improving the outcomes of coronary artery bypass surgery in New York State. *JAMA* 1994;271:761-766.

42. Dziuban SW Jr., McIlduff JB, Miller SJ, Dal Col RH. How a New York cardiac surgery program uses outcomes data. *Ann Thorac Surg* 1994;58:1871-1876.

43. Darby M. New York's latest cardiac data show trend in improvement. *Rep Med Guidelines Outcomes Res* 1994; 5(1):9-10.

44. Oberman L. Valuable input? Risk-adjusted data credited for better outcomes. *Amer Med News*. December 28, 1992; 1,24.

45. Hannan EL, Kumar D, Racz M, Siu AL, Chassin MR. New York State's Cardiac Surgery Reporting System: Four years later. *Ann Thorac Surg* 1994;58:1852-1857.

46. Green J, Wintfeld N. Report cards on cardiac surgeons: Assessing New York State's approach. *N Engl J Med* 1995;332:1229-1232.

47. Ghali WA, Ash AS, Hall RE, Moskowitz MA. Statewide quality improvement initiatives and mortality after cardiac surgery. *JAMA* 1997;277:379-382.

48. Mennemeyer ST, Morrisey MA, Howard LZ: Death and reputation: How consumers acted upon HCFA mortality information. *Inquiry* 34:117-128, 1997.

49. Mukamel DB, Mushlin AI. Quality of care information makes a difference. An analysis of market share and price changes after publication of the New York State cardiac surgery mortality reports. *Medical Care* 1998;36:945-954.

50. The Severyn Group, Inc. *Voices of Experience: Case Studies in Measurement and Public Reporting of Health Care Quality*. Oakland, CA: California HealthCare Foundation; 2001.

51. General Accounting Office. *"Report cards" are useful but significant issues need to be addressed*. Washington, DC: Government Printing Office, 1994. (Pub. no. GAO/HEHS 94-219).

52. Luce JM, Thiel GD, Holland MR, Swig L, Currin SA, Luft HS. Use of risk-adjusted outcome data for quality improvement by public hospitals. *West J Med* 1996;164:410-414.

53. Bentley JM, Nash DB. How Pennsylvania hospitals have responded to publicly released reports on coronary artery bypass graft surgery. *Joint Comm J Qual Improvement* 1998;24:40-49.

54. Schneider EC, Epstein AM. Influence of cardiac-surgery performance reports on referral practices and access to care: A survey of cardiovascular specialists. *N Engl J Med* 1996; 335:251-256.

55. Hannan EL, Stone CC, Biddle TL, DeBuono BA. Public release of cardiac surgery outcomes data in New York: What do New York State cardiologists think of it? *Am Heart J* 1997;134:55-61.

56. *Report of the California Hospital Outcomes Project. Acute Myocardial Infarction. Volume Two: Technical Appendix*. Sacramento, CA: Office of Statewide Health Planning and Development, 1996.

57. Romano PS, Zach A, Luft HS, Rainwater JA, Remy LL, Campa D. The California Hospital Outcomes Project: Using administrative data to compare hospital performance. *Joint Comm J Qual Improvement* 1995;21:668-682.

58. Iezzoni LI, Greenberg LG: Risk adjustment and current health policy debates. In Iezzoni LI (ed): *Risk Adjustment for Measuring Health Care Outcomes. Second Edition*. Ann Arbor, MI: Health Administration Press, 1997.

59. *Annual Report of Hospitals, 1994*. Sacramento, CA: Office of Statewide Health Planning and Development, 1995.

60. *Directory of Graduate Medical Education Programs 1994-1995*. Chicago, IL: American Medical Association; 1994.

61. Rainwater JA, Romano PS, Antonius DM. The California Hospital Outcomes Project: How useful is California's report card for quality improvement? *Joint Comm J Qual Improvement* 1998;24:31-39.

62. Hibbard JH, Jewett JJ. Will quality report cards help consumers? *Health Affairs* 1997; 16:218-228.

63. Hibbard JH, Jewett JJ. What type of information do consumers want in a health care report card? *Med Care Res Rev* 1996;53:28-47.

64. Cleary PD, Edgman-Levitan S. Health care quality: Incorporating consumer perspectives. *JAMA* 1997;278:1608-1612.

65. Topol EJ, Califf RM. Scorecard cardiovascular medicine: Its impact and future directions. *Ann Intern Med* 1994;120:65-70.

66. Gormley WT, Weimer DL: *Organizational Report Cards*. Cambridge, MA: Harvard University Press; 1999.

67. Scott WR. The organization of medical care services: Toward an integrated theoretical model. *Med Care Rev* 271-303, 1993.

68. Meyer JW, Rowan B: Institutional organizations: Formal structure as myth and ceremony. *Am J Sociol* 83 (September):440-463,

1977.

69. Johns L. Measuring quality in California. *Health Aff* 11:266-270, 1992.

70. Romano PS, Rainwater JA, Antonius DM. Grading the graders: How hospitals in California and New York perceive and interpret their report cards. *Med Care*; 1999;37:295-305.

71. *Survey on Outcomes Management.* Princeton, NJ: Foster Higgins, 1994.

72. Why organizations select their HMOs. *Drug Benefit Trends* 1998;10(1):8,31.

73. Gabel JR, Hunt KA, Hurst K. *When Employers Choose Health Plans: Do NCQA Accreditation and HEDIS Data Count?* KPMG Peat Marwick; 1998.

74. Hibbard JH, Jewett JJ, Legnini MW, Tusler M. Choosing a health plan: Do large employers use the data? *Health Affairs* 1997;16:172-180.

75. *California Health Maintenance Organizations 1997 Profile.* Sacramento, CA: California Association of HMOs; 1998.

76. 1998 Managed Care Profiles. *California Medicine* 1998(May/June):MC3-MC15. Also accessed at <www.cmanet.org/knoxfiles> on 29 June 1998.

77. *Summary of 1997 Enrollee Requests for Assistance*, accessed at <www.corp.ca.gov/pub/97rfa.htm> on 5 May 1999.

78. *Accreditation Summary Report*, accessed at <www.ncqa.org> on 3 April 1998.

79. Schulman KA, Rubenstein LE, Seils DM, Harris M, Hadley J, Escarce JJ. Quality assessment in contracting for tertiary care services by HMOs: A case study of three markets. *Joint Comm J Qual Improv* 1997;23:117-127.

80. Mukamel DB, Mushlin AI, Weimer D, et al. Do quality report cards play a role in HMOs' contracting decisions? Evidence from New York State. *Health Serv Res* 2000;35:319-332.

81. Schneider EC, Epstein AM. Use of public performance reports: A survey of patients undergoing cardiac surgery. *JAMA* 1998;279:1638-1642.

82. Schneider EC, Lieberman T. Publicly disclosed information about the quality of health care: response of the US public. *Qual Health Care* 2001;10:96-103.

83. Robinson S, Brodie M. Understanding the quality challenge for health consumers: The Kaiser/AHCPR survey. *Joint Comm J Qual Improv* 1997;23:239-244

84. *National Survey on Americans as Health Care Consumers: An Update on the Role of Quality Information.* Menlo Park, CA: The Kaiser Family Foundation/Agency for Healthcare Research and Quality, 2000.

85. Berry SH, Brown JA, Spranca MD, Monroe AF, Simon LP. Taking the pulse of Californians: A survey of consumer attitudes on health care quality. The RAND/California Health Care Foundation Report, Volume I. Presented at the U.S. Agency for Healthcare Research and Quality; November 27, 2000

86. Goody B. Defining rural hospital markets. *Health Serv Res* 1993; 28:183-200.

87. Welch HG, Larson EB, Welch WP. Could distance be a proxy for severity-of-illness? A comparison of hospital costs in distant and local patients. *Health Serv Res* 1993; 28:441-458.

88. Davies HTO, Washington AE, Bindman AB. Health care report cards: Implications for vulnerable patient groups and the organizations who provide care for them. *J Health Politics Policy Law*; in press.

89. Jewett JJ, Hibard JH. Comprehension and quality care indicators: Differences among privately insured, publicly insured, and uninsured. *Health Care Financing Rev* 1996;18:75-94.

90. Davies HTO. Public release of performance data and quality improvement: internal responses to external data by US health care providers. *Qual Health Care* 2001;10:104-110.