Estimating the potential impact of regionalizing health care delivery based on volume standards versus risk-adjusted mortality rate

LAURENT G. GLANCE1, TURNER M. OSLER2, DANA B. MUKAMEL3 AND ANDREW W. DICK4

1University of Rochester School of Medicine and Dentistry, Department of Anesthesiology, University of Rochester Medical Center, Rochester, NY, USA, 2The University of Vermont Medical College, VT, USA, 3Center for Health Policy Research, University of California, Irvine, CA, USA, and 4RAND Corporation, CA, USA

Abstract

Objective. To examine whether basing regionalization on risk-adjusted mortality would lead to better population outcomes than basing regionalization on procedure volume.

Data source. We used secondary data from the California State Inpatient Database obtained from the Healthcare Costs and Utilization Project.

Study design. A population-based retrospective cohort study of 243 thousand patients who underwent either abdominal aortic aneurysm surgery, coronary artery bypass surgery or coronary angioplasty between 1998 and 2000 in California. Four regionalization strategies were compared: (i) selective referral to high-quality hospitals; (ii) selective referral to high-volume hospitals; (iii) selective avoidance of low-quality hospitals; (iv) selective avoidance of low-volume hospitals.

Principal findings. Selective referral to high volume centers would be only moderately effective (2–20% relative reduction in mortality) and extremely disruptive (70–99% reduction in the number of hospitals treating these conditions). Selective referral to high quality centers was estimated to result in dramatic reduction in mortality (50%) but would also be highly disruptive with greater than 80% of the patients re-directed to high quality centers. Selective avoidance of low volume hospitals would not improve mortality, whereas selective avoidance of low quality hospitals was estimated to result in a small improvement in overall mortality (2–6%) while causing relatively minor disruptions in patient referral patterns.

Conclusion. Efforts to use volume standards as the basis for evidence-based hospital referrals should be re-evaluated by all stake-holders before promoting further efforts to regionalize health care delivery using volume cutoffs.

Keywords: administrative data, benchmarking, hierarchical model, outcome assessment, quality, report card, risk adjustment

According to most published studies, patients undergoing major surgery have better outcomes at high volume hospitals than at low volume hospitals [1]. In the absence of a systematic approach to measuring health care quality, selective referral of patients to high-volume hospitals has been advocated on the grounds that procedure volume can serve as a proxy for quality [2, 3]. The Leapfrog Group, ‘a large coalition of more than 150 private and public health care purchasers’ [4] covering 33 million patients, is the leading advocate of the use of market forces to promote selective referral of patients to high volume centers in order to achieve better health care outcomes [5].

Selective referral to high-volume centers is based on the premise that procedure volume is a valid quality indicator. However, the strength of the evidence supporting the existence of the volume–outcome relationship is uncertain [6, 7]. It is not clear whether volume cutoffs are an accurate surrogate for robust risk adjusted outcome measures. The association between volume and outcome tend to be true only on average [7–9] and the use of case volume as a quality proxy may not be valid. Some high-volume hospitals may have poor outcomes whereas some low-volume hospitals may have exceptional outcomes.

Address reprint requests to: Laurent G. Glance, Department of Anesthesiology, University of Rochester Medical Center, 601 Elmwood Avenue, Box 604, Rochester, NY 14642, USA. Tel: +1-585-275-4187; E-mail: laurent_glance@urmc.rochester.edu

International Journal for Quality in Health Care vol. 19 no. 4
© The Author 2007. Published by Oxford University Press on behalf of International Society for Quality in Health Care; all rights reserved
Previous studies suggest that selectively referring high-risk surgical and medical patients to high volume centers may prevent many deaths [2, 10–12]. However, the implications of a volume-based referral strategy have not been closely examined. Basing referral strategies on more direct measures of performance, such as risk-adjusted mortality, may result in better population outcomes than volume-based referral strategies. Furthermore, selective referral to high-volume centers has not been compared to alternative and potentially less disruptive referral strategies.

In this study, we will determine whether selectively referring patients to high-quality hospitals with low risk adjusted mortality rates results in better population outcomes than selectively referring patients to low volume hospitals. Furthermore, since ‘selective referral’ may be very disruptive because it re-distributes patients to a small number of centers of excellence [13], we also estimated the impact of ‘selective avoidance’ [13], a strategy that diverts patients from either ‘low-volume’ or ‘low-quality’ hospitals. We used the California State Inpatient Database to estimate the potential impact of each of these referral strategies on mortality, number of patient referrals and hospital service closures for patients undergoing either coronary artery bypass surgery, coronary angioplasty or abdominal aortic aneurysm surgery.

**Methods**

**Overview**

The goal of this study was to estimate the potential impact of four regionalization strategies for patients undergoing abdominal aortic aneurysm surgery, coronary artery bypass surgery and coronary angioplasty: (i) selective referral to high-quality hospitals; (ii) selective referral to high-volume hospitals; (iii) selective avoidance of low-quality hospitals; (iv) selective avoidance of low-volume hospitals. The potential impact of regionalization was assessed by simulating in-hospital mortality rate, the number of diverted patients and the number of hospital closed for each of the four referral strategies. We used hierarchical regression to identify high-quality and low-quality hospitals. High-volume hospitals were identified using the Leapfrog volume criteria [10, 14].

**Data source**

This study is based on the 1998–2000 California State Inpatient Database. The data were obtained from the Healthcare Costs and Utilization Project, a federal-state-industry partnership sponsored by the Agency for Healthcare Research and Quality. This data set is provided by the California Office of Statewide Health Planning and Development to the Healthcare Costs and Utilization Project and is based on data abstracted by hospital coders. Discharge data reports are audited using software edit tools and trending analysis. Data reports that are outside of error tolerance levels are returned to the submitting hospital for correction [15]. Previous work has validated the accuracy of the California State Inpatient Database by comparing coded data and re-abstracted medical records [16]. The California State Inpatient Database includes as many as 30 diagnoses and 21 procedures as ICD-9-CM codes. All ICD-9-CM codes (except for E-codes) were modified using a condition present at admission field to designate whether a diagnosis was present at admission. The current study was based on three study populations specified using ICD-9-CM codes: coronary artery bypass surgery, coronary angioplasty and abdominal aortic aneurysm surgery (Table 1).

**Model development**

Separate random-intercept models were developed for each of the three study populations. The outcome variable was in-hospital death because 30-day mortality was not available. The probability of in-hospital death for each patient was described using a random-intercept model, where \( x_{ij} \) are patient risk factors (i.e. age, gender), \( \beta_0, \ldots, \beta_p \) are the estimates of the model parameters and \( \phi_j \) is the probability of in-hospital death for the \( j \)th patient in the \( i \)th hospital:

\[
\phi_j = \frac{1}{1 + e^{-\phi}}
\]

where

\[
\phi = \beta_0 + \beta_1 x_{ij} + \cdots + \beta_p x_{pj} + u_{ij}
\]

This model assumes that the hospitals represent a random sample from a population of hospitals. The effect of hospital ‘\( j \)’ on the probability of death is specified by the random-intercept \( u_{ij} \) which is a random variable with zero mean and variance \( \sigma^2 \).

Medstat Disease Staging Software (Thomson Medstat, Ann Harbor, Michigan) was used to map ICD-9-CM codes to disease stages. STATA SE (STATA Corporation, College Station, TX, USA) programming language was used to map ICD-9-CM codes to the diagnostic categories in the Elixhauser algorithm [17]. Disease staging [18] was used to code the severity of the principal diagnosis. Comorbid conditions were coded either using Disease Staging or the Elixhauser algorithm [17]. Only ICD-9-CM codes that were coded as present at admission by the present-on-admission modifier were mapped to Disease Stages and Elixhauser diagnostic categories. The set of Disease Staging categories considered as possible covariates in each model included coronary artery disease, congestive heart failure, cardiomyopathy, valvular disease (aortic stenosis, aortic insufficiency, mitral stenosis, mitral insufficiency), arrhythmias, hypertension, respiratory disease (chronic obstructive pulmonary disease, emphysema) and endocrine disease (diabetes, adrenal, thyroid). Only conditions from the Elixhauser algorithm that did not overlap with these Disease Staging categories were considered as possible risk factors (i.e. renal failure, liver disease, paralysis, other neurological disease, coagulopathy, fluid and electrolyte disorders, cancer). The set of explanatory variables considered for each model also included...
demographic variables (age and gender), transfer status and admission type (elective vs. non-elective).

For each of the three models, initial variable selection was performed with stepwise selection using standard logistic regression. Predictors thus selected were reviewed for clinical plausibility. The method of fractional polynomials developed by Royston [19] was used to obtain the optimal transformation for continuous variables (age). These risk factors were then used to construct random-intercept models (base model).

Identification of hospital quality outliers

Hospital quality outliers for each of the three patient populations were identified using the shrinkage estimators obtained from the random-intercept models [22]. The shrinkage estimator of the random-intercept term for each hospital was exponentiated to obtain the adjusted odds ratio for each hospital. The shrinkage estimator is based on a weighting procedure that ‘shrinks’ the hospital intercept term, in the random-intercept model, towards the overall population mortality rate [22]. Hospital intercept terms that are based on small hospital volumes are considered less reliable and are ‘shrunk’ more than estimates based on large hospital volumes. Hospital intercept terms that are further away from zero (corresponding to estimated hospital mortality rates that deviate more from the average hospital mortality rate) are ‘shrunk’ more than estimates that are closer to the mean [23]. Hospitals with an adjusted odds ratio significantly greater than 1 were labeled as low-performance outliers, whereas providers with an adjusted odds ratio significantly less than 1 were classified as high-performance outliers.

Identification of high-volume and low-volume hospitals. High-volume hospitals were identified using the Leapfrog volume criteria: abdominal aortic aneurysm surgery > 50 cases/year; coronary artery bypass surgery > 450 cases/year, and coronary angioplasty > 400 cases/year [10]. In order to identify low-volume hospitals, we first divided hospitals into volume quartiles (based on annual volume) so that each quartile contained approximately one-fourth of the hospitals. We then defined low volume hospitals as those hospitals in the lowest volume quartile.

Table 1  ICD-9-CM codes for study populations

<table>
<thead>
<tr>
<th>Disease staging</th>
<th>Number of patients</th>
<th>Hospitals</th>
<th>ICD-9-CM codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdominal aortic aneurysm surgery</td>
<td>8 855</td>
<td>301</td>
<td>38.34, 38.44, 38.64</td>
</tr>
<tr>
<td>Coronary angioplasty</td>
<td>149 375</td>
<td>162</td>
<td>36.01, 36.02, 36.05, 36.06</td>
</tr>
<tr>
<td>Coronary artery bypass grafting</td>
<td>84 656</td>
<td>123</td>
<td>36.10–36.39</td>
</tr>
</tbody>
</table>

Estimating the impact of regionalization

In order to estimate the number of deaths if all coronary artery bypass surgery patients at low volume hospitals had been treated at high volume hospitals, we added a binary variable to the base model which indicated whether a patient was treated at a high volume center or a low volume center. We then used this model to calculate the predicted probability of death for patients selectively referred to high volume centers by changing the value of the indicator variable for the patients in the low volume group to reflect the counter factual assumption that they underwent coronary artery bypass surgery at the high volume centers. Because we assumed that patients were diverted to the ‘average’ quality hospital, the hospital random intercept terms were set equal to zero. We also assumed that there would be no decrement in the quality and outcomes at the referral hospitals (i.e. high volume hospitals) and that all patients (elective and non-elective) could be safely diverted to these hospitals. The estimated mortality rate for the patients referred to high volume hospital and the observed mortality rate for those patients who did not change hospitals were then used to calculate the overall mortality rate under this regionalization scenario.

We estimated the impact of each of our regionalization strategies for each of the three study populations in a similar fashion: (i) selective referral to high volume centers based on the Leapfrog volume cutoffs; (ii) selective referral to high-quality centers; (iii) selective avoidance of low-volume centers (hospital volume in bottom 25th percentile); (iv) selective avoidance of low-quality centers.

Data management, logistic regression models and regression diagnostics were performed using STATA, and hierarchical models were constructed in SAS version 8.2 (SAS Corp, Cary, NC, USA) using PROC GLIMMIX.

Results

Approximately two hundred and forty thousand patients were hospitalized for one of the three procedures or conditions included in this analysis. All estimated models had good to excellent discrimination with C-statistics ranging between 0.81 and 0.88. The Hosmer-Lemeshow statistics for the estimated models ranged between 11 and 53. Calibration curves demonstrated acceptable calibration for all models. Results on model performance are available on request.

The association between the case volume for individual hospitals and hospital quality ranking (low quality, medium...
quality and high quality) is shown in the Fig. 1. Using the hospital as the unit of analysis, hospital volume does not appear to represent an accurate proxy for hospital quality in any of the study populations. Some hospitals with high case volumes deliver low quality care, whereas some of the low-volume hospitals deliver high quality care.

**Potential impact of regionalization**

The potential impact of selectively referring patients to high volume centers based on the Leapfrog volume standards in California was estimated (Table 2a). Adoption of the Leapfrog volume standard for coronary artery bypass would result in a decrease in the mortality rate from 3.02 to 2.42% ($P \leq 0.05$), but would require that 70% of coronary artery bypass patients be transferred to high volume centers and would lead to the closure of 91% of the cardiac surgery units. Adoption of the Leapfrog volume standard for abdominal aortic aneurysm surgery would result in a decrease in the mortality rate from 12.6 to 11.6% ($P \leq 0.05$), but would require that 94% of the patients be transferred to high volume centers and would lead to the elimination of abdominal aortic aneurysm surgery at 99% of the hospitals currently performing abdominal aortic aneurysm surgery. Implementation of the Leapfrog criteria for coronary angioplasty would not lead to a significant reduction in mortality.

Selective referral to high quality hospitals would lead to the greatest reduction in mortality rate: approximately ≥ 50% reduction for abdominal aortic aneurysm surgery, coronary artery bypass surgery and coronary angioplasty (Table 2a). However, this would require that ~80–90% or more of the patients be diverted to centers of excellence, and the elimination of these clinical services at ~92–99% of the hospitals.

Selective avoidance of low volume hospitals would not result in significant reductions in mortality for any of the three procedures (Table 2b). Selective avoidance of low-quality hospitals would lead to small, but statistically significant, reductions in mortality rate for coronary artery bypass surgery and coronary angioplasty (Table 2b).

**Discussion**

We find that selective referral to high volume centers would be only moderately effective (0–20% relative reduction in mortality) and extremely disruptive (70–99% reduction in the number of hospitals treating these conditions). Selective referral to high quality centers was estimated to result in dramatic improvements in mortality but would not be a realistic policy option given the need to re-direct greater than 80% of the patients to a small number of centers of excellence. Selective avoidance of low volume hospitals would not result in lower mortality, whereas selective avoidance of low quality hospitals was estimated to result in a small improvement in overall mortality (2–6%) while causing relatively minor disruptions in patient referral patterns.

Based on the extensive evidence showing that, on average, higher volumes are associated with better outcomes, many analysts and policy makers have supported the use of volume cutoffs as the basis for regionalizing health care delivery [3, 24]. Birkmeyer, after exploring the volume outcome association for 2.5 million Medicare admissions, concluded that in the absence of hospital quality information, patients ‘can significantly reduce their risk of operative death by selecting high-volume hospitals [11].’ Despite the potentially significant disruptions involved with such an approach, relatively few studies have been undertaken to assess the potential impact of regionalizing health care in this manner [2, 10, 25, 26]. The largest and most comprehensive study to
<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Leapfrog</th>
<th></th>
<th>High quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mortality (%)</td>
<td>Patients</td>
<td>Hospitals</td>
<td>Mortality (%)</td>
</tr>
<tr>
<td>Selective referral</td>
<td>AAA 12.55</td>
<td>8855 301</td>
<td>11.51 8.3</td>
<td>93.7 99.0</td>
</tr>
<tr>
<td></td>
<td>CABG 3.02</td>
<td>84 656 123</td>
<td>2.42 20</td>
<td>69.6 91.1</td>
</tr>
<tr>
<td></td>
<td>PCI 1.56</td>
<td>149 375 162</td>
<td>1.53 1.9</td>
<td>29.6 67.3</td>
</tr>
<tr>
<td>Selective avoidance</td>
<td>AAA 12.55</td>
<td>8855 301</td>
<td>12.36 1.5</td>
<td>3.2 28.2</td>
</tr>
<tr>
<td></td>
<td>CABG 3.02</td>
<td>84 656 123</td>
<td>2.94 2.6</td>
<td>7.1 26</td>
</tr>
<tr>
<td></td>
<td>PCI 1.56</td>
<td>149 375 162</td>
<td>1.56 0</td>
<td>0.20 25.3</td>
</tr>
</tbody>
</table>

AAA, abdominal aortic aneurysm surgery; CABG, coronary bypass surgery; PCI, percutaneous coronary intervention.

*P* ≤ 0.05 compared to baseline.
evaluate the impact of regionalization [2], also conducted using California data, estimated that selective referral to high volume centers could result in 602 fewer deaths each year. This landmark study by Dudley et al. motivated us to further explore the potential impact of selective referral. However, we chose to use both volume standards and risk-adjusted mortality as the basis for identifying high-quality hospitals. Furthermore, unlike the study by Dudley [2], we did not use the results of external studies to estimate the odds ratio of death at high volume hospitals compared to low volume hospitals. Instead, we constructed risk adjustment models using the California data set and used those same models to estimate the potential impact of regionalization in California. In addition, we explored potentially less disruptive regionalization strategies based on selective avoidance of low volume hospitals and selective avoidance of low quality hospitals—as alternatives to a policy based on selective referral to high volume hospitals.

Of all the regionalization scenarios, only selective referral to high quality hospitals achieves major reductions in mortality rate. However, this scenario could never be implemented due to the small number of high quality centers. Moreover, the large increase in the caseload of high quality hospitals is likely to adversely affect the quality of those institutions. Similar reductions in mortality may be achievable if the overall quality of all hospitals could be incrementally improved to approach the quality of high-performance hospitals. Although this latter approach may at first appear unrealistic, there is evidence that coupling quality improvement to clinical benchmarking improves outcomes. The effectiveness of this approach has been demonstrated for coronary artery bypass surgery by the Northern New England group [27] and for non-cardiac surgery in the Veterans Administration hospital system [28]: linking quality monitoring to quality improvement was associated with a 24% reductions in hospital mortality rates for coronary artery bypass surgery in the Northern New England group hospitals and a 27% reduction in mortality rates for non-cardiac operations in the Veterans Administration hospitals. Thus, raising the overall quality of care is less disruptive and may have as great, or greater, an impact on outcome compared to regionalization.

With the exception of coronary artery bypass surgery, neither selective referral to high-volume centers, nor selective avoidance of low-volume centers achieved very significant reductions in mortality rate. This finding is a result of the relatively weak association between hospital quality and hospital volume. Most high-volume hospitals are not high-quality, and most low-volume hospitals are not low-quality.

Even if regionalization based on volume standards were shown to be effective and practical, it is questionable whether the use of volume as a surrogate marker for quality would continue to be effective over the long term. If volume was in fact an accurate proxy for quality, then selective referral to high volume centers would ‘freeze’ the status of high volume centers as ‘high quality’ centers by insuring that these high volume centers continue to receive high volumes of patients. However, there is no guarantee that these centers would continue to deliver high quality care because hospital volume would now become largely structural in nature and would be determined by fixed patient referral patterns. In addition, other factors affecting hospital referral patterns are independent of quality. For example, selective contracting in a managed care environment may lead to higher volumes in lower-priced hospitals, irrespective of quality considerations [9].

One of the key findings of this study is that selective avoidance of low-volume centers does not appear to result in decreased mortality. Although a policy of selective avoidance is minimally disruptive, is based on an easily verifiable structural measure (volume), and has face validity [29], our analyses suggest that such a policy would not lead to better population outcomes. It is unlikely that this finding is due to insufficient statistical power since the point estimates for the mortality rates in the baseline groups and in the groups based on selective avoidance of low volume centers were nearly identical.

The primary limitation of this study is that the use of administrative data may bias risk adjustment. Problems associated with the use of administrative data sets are well known and include errors in the abstraction process, under-coding of comorbidities, variation in the quality of charting by physicians, lack of precise definitions for ICD-9-CM codes and ‘overcoding’ of patient diagnoses to maximize reimbursements. On the other hand, clinical data sets are not widely available, whereas in most states and for all Medicare enrollees, administrative data are available for every hospitalization and could serve as an ‘information infrastructure’ [30] for performance measurement. One of the most important sources of bias in administrative data is the difficulty in distinguishing complications from pre-existing conditions. However, all ICD-9-CM codes within the California State Inpatient data set are date stamped so that conditions present at the time of admission can be distinguished from complications. Recent analyses provide evidence to support the validity of the date stamp [31, 32].

Our study has several other potentially important limitations. First, we did not consider referral strategies which focused on individual surgeons, as opposed to hospitals. Birkmeyer [33] has shown that surgeon volume accounts for a significant proportion of the difference in mortality between low and high volume hospitals. However, we could not extend our analysis to include surgeons because the California data set lacks surgeon identifiers. Second, we did not exclude emergency cases from this analysis and, instead, assumed that regionalization would apply equally to elective and non-elective cases. We believe that it is unrealistic to simulate a clinical scenario where only the highest risk emergency patients remain in the lowest quality hospitals.

Third, we did not exclude rural hospitals from the analyses. It may not be practical to regionalize health care for rural populations because of the impact of regionalization on the ability of patients residing in rural areas to access medical care. Nor did we limit patient referrals to hospitals within a pre-specified distance radius from patients’ residence (the California State Inpatient Database does not contain patient zip codes). Therefore, our estimates provide the upper bounds on the number of lives that might be saved with regionalization and may be biased upwards (towards showing
a greater benefit for regionalization). Fourth, we did not take into account the unintended consequences of regionalization. For example, it is unlikely that a hospital’s caseload could double or triple without it experiencing some worsening in outcomes. As noted previously, our calculations represent the upper bounds on the potential benefit of regionalization. Moreover, even without taking into account the potential down side of regionalization, most of the regionalization scenarios—other than selective referral to high quality centers—resulted in only modest or no improvements in population mortality rates.

Fifth, in-hospital mortality was used instead of 30-day mortality because 30-day mortality was not available in our data set. In-hospital mortality may be influenced by hospital discharge policies and could be a source of bias. Additionally, this study relied on mortality as the sole measure of quality. Mortality is the outcome measure used in most studies evaluating the volume-outcome relation [1, 2] and the potential benefit of regionalization [2, 10] because of the importance and reliability of mortality as an outcome, and the inherent difficulty of reliably defining and measuring complication outcomes.

Sixth, we did not model the potential benefit of the targeted regionalization of high-risk patients. Targeted regionalization may be more realistic because it is less disruptive. However, in the case of coronary artery bypass surgery patients, one study suggests that high risk patients benefit more than low risk patients from undergoing surgery at high-volume centers [34], while two studies have shown the opposite [35, 36]. Because it is unclear whether being treated at high-volume or high-quality hospitals benefits high-risk patients to a greater extent than low-risk patients, it is not possible to estimate the potential benefits of targeted regionalization.

Finally, this study is based on California alone and our findings may not be easily generalizable to other states. Health care is highly decentralized in California compared to other regions in North America [25]. On the other hand, it is precisely in this type of setting where regionalization would be expected to yield the greatest benefit—as opposed to states like New York State where care is already centralized to a much greater extent.

Conclusion

We find that selective referral to high volume centers may, at best, yield only modest improvement in outcomes but would result in massive disruption in hospital services in California, and possibly in other decentralized health care environments. Selective avoidance of low volume hospitals would not lead to any improvement in outcomes. Selective avoidance of low quality hospitals might yield some improvement in mortality rates without causing large-scale disruptions in referral patterns. Further studies are needed to validate our findings. The use of procedure volume as the basis for evidence-based hospital referrals should be re-evaluated by all stake-holders before undertaking further efforts to regionalize health care delivery using volume-based referral strategies.

Acknowledgements

This project was supported by a grant from the Agency for Healthcare and Quality Research (RO1 HS 13617).

References


Accepted for publication 23 April 2007