Procedural Volume as a Marker of Quality for CABG Surgery

Eric D. Peterson, MD, MPH
Laura P. Coombs, PhD
Elizabeth R. DeLong, PhD
Constance K. Haan, MD
T. Bruce Ferguson, MD

THE ASSOCIATION BETWEEN HOSPITALS’ CORONARY ARTERY BYPASS GRAFT (CABG) SURGERY VOLUME AND OUTCOME HAS BEEN THE SUBJECT OF MULTIPLE INVESTIGATIONS.\(^1\) Based on these studies, the Center for Medicare & Medicaid Services (CMS) and the Leapfrog Group have proposed using hospital volume as an indicator of CABG quality.\(^18\)\(^\rightarrow\)\(^20\) Prior studies of the association between volume and outcome, however, generally have been based on administrative data sources, have reflected selected patient populations, and often have not adequately accounted for patient case mix, patient clustering, and other methodological concerns. Before hospital surgery volume is accepted as a standard quality metric, further research into issues related to both analysis and policy implication is warranted.

We undertook a contemporary examination of the association between hospital CABG procedural volume and outcome using clinical data available from the Society of Thoracic Surgeons (STS) National Cardiac Database. Specifically, we considered whether hospital CABG volume was associated with operative mortality after accounting for patient case mix. Second, we examined the extent to which patient clustering within centers and site variance issues affected this association. Third, we determined how the association varied as a function of patient age and predicted surgical risk. Fourth, we determined whether the association between hospital CABG volume and outcome was influenced by individual surgeon volume. Finally, we investigated the potential health policy implications of using hospital volume as a quality indicator. This included determining the ability of hospital volume to discriminate high-mortality outliers, as well as investigating the potential number of lives saved if low-volume centers were systematically closed.

Context There have been recent calls for using hospital procedural volume as a quality indicator for coronary artery bypass graft (CABG) surgery, but further research into analysis and policy implication is needed before hospital procedural volume is accepted as a standard quality metric.

Objective To examine the contemporary association between hospital CABG procedural volume and outcome in a large national clinical database.


Main Outcome Measure Association between hospital CABG procedural volume and all-cause operative mortality (in-hospital or 30-day, whichever was longer).

Results The median (interquartile range) annual hospital-isolated CABG volume was 253 (165-417) procedures, with 82% of centers performing fewer than 500 procedures per year. The overall operative mortality was 2.66%. After adjusting for patient risk and clustering effects, rates of operative mortality decreased with increasing hospital CABG volume (0.07% for every 100 additional CABG procedures; adjusted odds ratio [OR], 0.98; 95% confidence interval [CI], 0.96-0.99; \(P\) = .004). While the association between volume and outcome was statistically significant overall, this association was not observed in patients younger than 65 years or in those at low operative risk and was confounded by surgeon volume. The ability of hospital volume to discriminate those centers with significantly better or worse mortality was limited due to the wide variability in risk-adjusted mortality among hospitals with similar volume. Closure of up to 100 of the lowest-volume centers (ie, those performing \(\leq 150\) CABG procedures/year) was estimated to avert fewer than 50 of 7110 (<1% of total) CABG-related deaths.

Conclusion In contemporary practice, hospital procedural volume is only modestly associated with CABG outcomes and therefore may not be an adequate quality metric for CABG surgery.

JAMA. 2004;291:195-201

For editorial comment see p 246.

METHODS

Data Source

The STS National Cardiac Database was established in 1989 to report surgical outcomes following cardiothoracic surgical procedures.\(^21\)\(^\rightarrow\)\(^23\) The database cur...
PROCEDURAL VOLUME AND QUALITY OF CABG SURGERY

rently captures clinical information from nearly two thirds of all US bypass procedures from more than half of all centers performing adult cardiac surgery. Sites enter patient data using uniform definitions (available online at http://www.sts.org) and certified software systems. This information is sent semi-annually to the STS Data Warehouse and Analysis Center at the Duke Clinical Research Institute. There, a series of data quality checks are performed before a site's data are aggregated into the national sample. While participation in the STS database is voluntary, data completeness is high, with overall preoperative risk factors missing in fewer than 5% of submitted cases. The accuracy of submitted data has further been confirmed in independent comparison of hospital CABG surgery volume and mortality rates reported to the STS vs those reported to the CMS.

Patient Population
We examined isolated CABG surgery procedures, excluding those combined with valve or other major surgical interventions, performed between January 1, 2000, and December 31, 2001, at hospitals reporting to the STS. We excluded 12 centers (2.7% of total STS centers) that reported fewer than 30 CABG procedures in a year and that had evidence for incomplete reporting. Inclusion of the 223 cases from these 12 centers had no measurable impact on the study’s measured association between CABG procedural volume and outcome.

Statistical Analysis
Our primary outcome measure was all-cause operative mortality, defined as inhospital or 30-day mortality, whichever was longer. Major morbidity was defined as any of 5 postoperative in-hospital complications: stroke, reoperation for any reason, need for mechanical ventilation for more than 24 hours following surgery, renal failure, or deep sternal wound infection.

Hospital-isolated CABG annual case volume was averaged over a 2-year period (2000 and 2001) to increase its stability. In the overall analysis, volume was considered to be a continuous variable. For display purposes, patient and hospital characteristics and unadjusted outcomes were categorized by annual hospital procedural volume (1150, 151 to 300, 301 to 450, and >450). These break points were chosen to form 4 fairly equal-sized hospital samples while maintaining similar volume differences among the groups.

The effect of hospital volume on unadjusted outcomes was tested using standard logistic regression. Expected mortality rates for patients were calculated using a logistic regression model that included 28 previously identified preoperative risk factors and year of surgery. The C-index for this model in the study population was 0.78. Risk-adjusted mortality rates for each hospital were calculated by dividing the observed mortality rate by the expected mortality rate at the hospital and multiplying by the overall (national) bypass mortality rate. Additionally, hierarchical logistic regression was used to account for patient clustering within centers (SAS Macro GLIMMIX). These analyses included clinical risk factors, procedure year, and hospital procedural volume as a fixed effect and included random intercepts for sites. To supplement this analysis, a hospital-level weighted least-squares regression of risk-adjusted mortality on volume was used to correct for nonconstant variance associated with hospital sample size, as reflected by hospital volume.

Because many previous studies have focused on Medicare populations (ie, those with patients aged ≥65 years) while others have specifically investigated the effect of hospital volume as a function of patient risk,13,15 we also tested for interactions between hospital volume and patient age (<65 years and ≥65 years) and expected risk for patients in our hierarchical analyses. We also used our hierarchical analysis to test the hospital volume × surgeon volume interaction. While the STS database does not contain the names of individual surgeons, it does contain an encrypted primary surgeon identifier that can be aggregated within and among hospital locations to approximate annual surgeon volume.

Finally, we also examined health policy implications of hospital volume. We first determined the degree to which hospital volume identified “mortality outlier” hospitals as determined by their significance in a hierarchical logistic regression model containing patient characteristics but not volume. Specifically, we used a hospital-level standard logistic regression model to test whether hospital procedural volume could discriminate mortality outliers. Second, we assessed the potential number of lives saved from closure of either the lowest hospital volume quartile in the STS database (≤150 CABG procedures per year), or alternatively using the Leapfrog-proposed referral criteria of fewer than 300 procedures per year. In these calculations, we assumed that all these patients could safely be transferred from low- to higher-volume centers without risk and that the patients would assume the expected risk predicted for the higher-volume site once there. All statistical analyses were performed using SAS release 8.2 (SAS Institute Inc, Cary, NC), with P <.05 considered statistically significant.

RESULTS
Between 2000 and 2001, 267089 isolated CABG procedures were performed at 439 STS hospitals. Average hospital procedural volumes ranged from 39 to 1734 isolated CABG procedures (median, 253; interquartile range, 165-417). Eighteen percent of STS centers performed 500 or more procedures per year.

TABLE I displays patient and hospital characteristics as a function of hospital volume. Compared with higher-volume centers, lowest-volume hospitals (ie, those performing ≤150 procedures per year [n=98]) were more likely to operate on nonwhite patients, as well as on patients with chronic lung disease, prior stroke, recent myocardial infarction, left main artery disease, and...
emergent or salvage settings. Based on preoperative risk factors, the average expected surgical mortality risk ranged from 3.0% for hospitals performing 150 cases or fewer to 2.6% for those performing more than 450 cases. The Pacific and West South Central regions tended to have more low-volume hospitals per region, while New England, and the Mid-Atlantic, South Atlantic, and East South Central regions had more high-volume centers.

**Hospital Volume and Outcomes**

Overall, there were 7110 deaths (2.66%). Overall unadjusted mortality declined from 3.5% for hospitals in the lowest-volume group to 2.4% for hospitals in the highest-volume group ($P < .001$ for trend) (Table 2). Rates of prolonged ventilation and renal failure also declined significantly ($P < .001$ for both) with increasing hospital procedural volume. Reoperation, stroke, and rates of deep sternal wound infection were generally constant across volume groups, as was postoperative length of stay (median, 5 days for all volume groups).

After adjusting for preoperative clinical risk, year of surgery, and patient clustering within centers, hospital mortality declined significantly as a function of hospital procedural volume, ranging from 3.1% for centers performing 150 or fewer cases per year to 2.4% for those performing more than 450 cases per year. When looking at hospital volume as a continuous variable, absolute rates of mortality decreased by 0.07% for every 100 additional CABG cases per year (adjusted odds ratio [OR], 0.98; 95% confidence interval [CI], 0.96-0.99; $P = .001$).

Effect of Volume on Outcome in Patient Subgroups

The absolute effect of hospital volume was more apparent in elderly patients (Figure 2). In patients aged 65 years and older, there was a 1.0% difference in observed mortality rates between low- ($\leq 150$) and high-volume ($>450$) hospitals that was only slightly diminished after adjustment for risk ($P < .001$). In contrast, among those younger than 65 years, observed mortality declined by only 0.3% and was insignificant after adjustment for risk ($P = .53$).

**Surgeon and Hospital Volume**

Table 4 demonstrates the effect of hospital and surgeon volume on risk-adjusted operative mortality rates. While there was collinearity between sites by means of weighted least squares ($P = .001$).

Similarly, patients with intermediate and high expected preoperative risk demonstrated consistently lower mortality when treated at higher-volume centers (Table 3). In contrast, among those with expected operative risk of less than 1.5%, there was no volume effect in either observed or adjusted mortality rates. A formal test for volume $\times$ patient risk interaction in the logistic regression model was significant ($P = .002$).

**Table 1. Patient and Hospital Characteristics**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Hospital CABG Volume, Procedures per Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\leq 150$</td>
</tr>
<tr>
<td>No. of patients</td>
<td>16,929</td>
</tr>
<tr>
<td>No. of hospitals</td>
<td>98</td>
</tr>
<tr>
<td>Demographic</td>
<td></td>
</tr>
<tr>
<td>Age, median (IQR), y</td>
<td>66.0 (57-73)</td>
</tr>
<tr>
<td>Nonwhite race</td>
<td>19.6</td>
</tr>
<tr>
<td>Men</td>
<td>71.3</td>
</tr>
<tr>
<td>Clinical</td>
<td></td>
</tr>
<tr>
<td>Chronic lung disease</td>
<td>20.7</td>
</tr>
<tr>
<td>Prior stroke</td>
<td>7.6</td>
</tr>
<tr>
<td>Hypertension</td>
<td>71.8</td>
</tr>
<tr>
<td>Diabetes</td>
<td>29.7</td>
</tr>
<tr>
<td>Prior bypass</td>
<td>7.3</td>
</tr>
<tr>
<td>LVEF, median (IQR)</td>
<td>50.0 (43-50)</td>
</tr>
<tr>
<td>Left main artery disease $&gt;50%$</td>
<td>26.0</td>
</tr>
<tr>
<td>3-Vessel disease</td>
<td>76.9</td>
</tr>
<tr>
<td>NYHA class IV</td>
<td>18.8</td>
</tr>
<tr>
<td>Myocardial infarction $&lt;21$ d</td>
<td>25.7</td>
</tr>
<tr>
<td>Emergent/salvage</td>
<td>5.42</td>
</tr>
<tr>
<td>Preoperative risk, %†</td>
<td>3.0</td>
</tr>
<tr>
<td>Hospital Teaching hospital</td>
<td>21.7</td>
</tr>
<tr>
<td>Region</td>
<td></td>
</tr>
<tr>
<td>New England</td>
<td>1.1</td>
</tr>
<tr>
<td>Mid-Atlantic</td>
<td>4.4</td>
</tr>
<tr>
<td>South Atlantic</td>
<td>12.0</td>
</tr>
<tr>
<td>East North Central</td>
<td>23.9</td>
</tr>
<tr>
<td>East South Central</td>
<td>5.4</td>
</tr>
<tr>
<td>West North Central</td>
<td>10.9</td>
</tr>
<tr>
<td>West South Central</td>
<td>13.0</td>
</tr>
<tr>
<td>Mountain</td>
<td>8.7</td>
</tr>
<tr>
<td>Pacific</td>
<td>20.7</td>
</tr>
</tbody>
</table>

Abbreviations: CABG, coronary artery bypass graft; IQR, interquartile range; LVEF, left ventricular ejection fraction; NYHA, New York Heart Association.

All values are percentages unless otherwise indicated.

†Based on Society of Thoracic Surgeons Mortality model.25
these factors, both surgeon volume and hospital volume were significant predictors of mortality. Combined, the highest mortality rates (3.3%) were observed when patients were treated by low-volume surgeons at low-volume hospitals and best results (2.4%) were obtained by high-volume surgeons at high-volume hospitals.

**Policy Implications**

The use of hospital procedural volume was of limited value in discriminating those hospitals with significantly better or worse risk-adjusted mortality outcomes (“outlier centers”) (C-index, 0.60 and 0.67, respectively). In fact, as procedural volume dropped, a hospital’s likelihood of being singled out as a center with significantly better or worse outcomes declined due to increasing variance associated with the measurement of end point at low-volume sites.

An upper boundary for the potential numbers of lives saved by closure of low-volume CABG centers was also estimated. If 25% of STS hospitals with the lowest annual CABG volume (<150 procedures) were closed and their patients undergoing CABG surgery were safely transferred to another, higher-volume center, then approximately 45 CABG procedural deaths could be averted annually among all 439 STS centers. If the criteria for selective referral were based on Leapfrog’s referral criteria of 500 cases per year,18 82% of STS centers would close and the high-end estimate of avertable CABG deaths would be 212 per year.

**Validity of STS Data**

In order to determine the validity of STS data for patients aged 65 years and older, we performed a systematic comparison with data supplied to the CMS. There were 415 centers that released data to the STS and the CMS; more cases (96330 vs 78788) were reported to the STS, likely due to nonreporting of health maintenance organization and managed care cases to the CMS. The reported mortality rates were 4.50% and 4.48%, respectively. Reported mortality rates among centers that performed 150 or fewer procedures per year were 5.06% and 4.91% according to the STS and the CMS, respectively, while corresponding rates for centers that performed more than 150 cases per year were 4.39% and 4.35%. Thus, we
doubt that our results were biased by underreporting or selective reporting to the STS.

We also systematically compared mortality rates among participants and nonparticipants in the STS. Mortality rates in participating sites were lower (4.5% vs 5.2%), but the differential mortality between lower-volume sites (≤150 CABG procedures/year) and higher-volume sites (≥151 CABG procedures/year) was larger in the STS sites than in nonparticipating sites (0.6% vs 0.1%, respectively). Thus, the association between volume and outcome may have been overestimated by focusing on STS sites.

**COMMENT**

The recent interest in using hospital procedure volume as a quality indicator for CABG surgery is understandable. This structural characteristic is readily available via administrative claims data, requires no complex adjustment techniques, is easily interpretable by the lay public, and is consistent with the common belief that “practice makes perfect.”20 However, our study indicates that hospital volume has only a modest association with risk-adjusted mortality and has important limitations as a quality metric for CABG surgery.

To date, there have been multiple analyses of volume and outcome for CABG surgery. Most14-17 but not all16,17 of these concluded that hospital procedural volume was associated with lower rates of hospital mortality. The largest analysis of the association between volume and outcome was conducted by Birkmeyer et al.14 Using national Medicare claims data from 1994 to 1999, they found a 1.3% absolute difference in unadjusted rates of mortality between the lowest to highest quartiles of hospital volume.14 While their study was graphically inclusive, it was limited to patients aged 65 years or older, could not adequately adjust data for potential differences in case mix, and did not control for patient clustering within sites or for site variance issues.

One of the most complete evaluations of CABG volume-outcome relationships performed using clinical data was completed by Hannan et al.9 Using data from 30 centers performing CABG surgery in New York from 1997 to 1999, the study by Hannan et al also noted significant differences in risk-adjusted mortality among low- and high-volume hospitals. The study, however, was limited to a single state where a strict certificate-of-need program limits the number of low-volume centers (ie, <3% of all patients receive CABG surgery at centers performing <300 procedures per year).

Our contemporary analysis using clinical data from the STS National Cardiac Database found an association between procedural volume and unadjusted mortality similar to that found in past analyses. Our study, however, expanded on these prior analyses using contemporary analytic techniques to properly account for clinical factors, differences in site variability, and clustering within sites. We found that, compared with high-volume hospitals, low-volume hospitals tended to operate on patients with higher risk and under more emergent conditions (Table 1). Reasons for these differences may include adverse selection,7,10 variance in clinical coding among hospitals, or a differential threshold for surgery due to altered center experience and/or institutional financial pressure.

We also found that the association between hospital volume and mortality was not constant among all patients. In particular, the volume-outcome effect was nonsignificant in patients younger than 65 years and in those with low preoperative risk (Figure 2, Table 3). These results imply that prior analyses performed exclusively in Medicare patients would

---

**Table 3.** Association Between Hospital CABG Volume and Mortality, by Expected Risk

<table>
<thead>
<tr>
<th>Expected Risk, %</th>
<th>Unadjusted Mortality Rate, % (95% CI)</th>
<th>Adjusted Mortality Rate, % (95% CI)†</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤150</td>
<td>0.9 (0.7-1.1)</td>
<td>0.7 (0.6-0.9)</td>
</tr>
<tr>
<td>151-300</td>
<td>0.8 (0.7-0.9)</td>
<td>0.7 (0.6-0.8)</td>
</tr>
<tr>
<td>301-450</td>
<td>0.7 (0.5-0.9)</td>
<td>0.5 (0.4-0.6)</td>
</tr>
<tr>
<td>≥450</td>
<td>0.5 (0.4-0.6)</td>
<td>0.3 (0.2-0.5)</td>
</tr>
</tbody>
</table>

**Table 4.** Risk-Adjusted Mortality, by Hospital and Surgeon Volume†

<table>
<thead>
<tr>
<th>Surgeon CABG Volume, Procedures per Year</th>
<th>Hospital CABG Volume, Procedures per Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤150</td>
<td>151-300</td>
</tr>
<tr>
<td>151-300</td>
<td>301-450</td>
</tr>
<tr>
<td>≥450</td>
<td>Overall</td>
</tr>
</tbody>
</table>

Abbreviations: CABG, coronary artery bypass graft. *Mortality rates are presented as percentages. †Accurate assessment not possible due to insufficient overlap of high-volume surgeons in hospitals with very low procedural volume.
inflated the volume-outcome effect. Additionally, it is presumed that while young, low-risk patients would be the group most likely to use public information from CABG quality metrics, these are also the patients in whom hospital volume had no measurable association with outcome.

We further demonstrated that associations between hospital volume and outcome were confounded by the concomitant effects of surgeon volume (Table 4). These results expand those previously reported from a single state9 to a national cohort. Even if a patient elected to receive surgery at a high-volume center, their risk-adjusted mortality rates could vary from 2.4% to as much as 2.9%, depending on whether or not their actual procedure was performed by a low- vs a high-volume surgeon.

Our study further demonstrates the limitations of using hospital volume as an indicator of the quality of CABG surgery. Hospital volume had generally poor predictive accuracy as a means of identifying hospitals with significantly better or worse CABG mortality rates (C-index, <0.7). Similarly, using volume as a sole referral criterion for selecting a provider would unfairly defer cases from nearly half of very-low-volume centers with outcomes equal or better than overall STS mortality results (Figure 1).

Regardless of its value at the level of the individual center, some still have argued for using hospital-volume limits for their “aggressive societal goal.” A study from California estimated that up to 27% of deaths at low-volume centers may be averted with selective referral.20 In contrast, our study estimated that the total number of lives saved by eliminating up to 100 low-volume STS hospitals (25% of total sites) would be less than 1% from deaths (n=45). If Leapfrog-proposed referral criteria of 500 or more cases per year were applied, up to four fifths of STS sites would be closed while fewer than 3% of national STS CABG mortality events would potentially be averted. This study does not fully preclude the use of procedural volume in certain roles. In particular, the small but consistent association between hospital procedural volume and outcomes should be considered when prospectively determining the need for new programs.30 Additionally, given that quality assessment based on comparison of risk-adjusted outcomes is limited at low-volume sites due to wide 95% CIs surrounding their outcome estimates, it seems reasonable to monitor low-volume centers using trend analyses of multiyear adjusted outcomes to ensure quality of care.

Several limitations should be noted. While this study represents the largest clinical evaluation of hospital volume to date, the STS National Cardiac Database currently collects data from roughly half of all US surgical centers. Participating centers tend to be larger and have slightly better outcomes than nonparticipating centers. However, unadjusted results in this study were similar to those from prior studies. Our evaluation of surgeon volume was limited to a blinded proxy and may underestimate this impact. Third, while our study indicated that low-volume centers may have been more willing to operate on higher-risk patients, it could not determine whether this lower threshold for case selection was indicative of better or worse patient care.

Further study of the appropriateness of case selection by hospital volume would be indicated. Finally, our study was not designed to isolate specific clinical mechanisms for the association between volume and outcome. Differences in surgical teams, quality of postoperative care, surgical techniques, and other unmeasured factors all may contribute to the effect observed in this study.

In this national study we found that hospital procedural volume was only modestly associated with risk-adjusted CABG mortality rates; however, there were many low-volume hospitals with low mortality rates and some high-volume centers with rates higher than expected. This study suggests that hospital CABG surgery volume is best considered as a surrogate for quality in a setting where other more direct process and outcome assessments are not available.20 Instead it seems more reasonable to support the continued growth of national clinical databases, which are capable not only of tracking risk-adjusted surgical care patterns and outcomes, but also of improving them.23,31

Author Contributions: Dr Peterson, as principal investigator of this study, had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analyses. Study concept and design: Peterson, DeLong, Haan, Ferguson. Acquisition of data: Peterson, Ferguson. Analysis and interpretation of data: Peterson, Coombs, DeLong, Haan. Drafting of the manuscript: Peterson, Coombs. Critical revision of the manuscript for important intellectual content: Peterson, Coombs, DeLong, Haan, Ferguson. Statistical expertise: Peterson, Coombs, DeLong. Obtained funding: Peterson, Ferguson. Administrative, technical, or material support: Ferguson. Study supervision: Peterson.

Role of the Sponsor: This study was sponsored by the Society of Thoracic Surgeons (STS). Specifically, the Duke Clinical Research Institute (DCRI) has a contract with the STS to be their National Cardiac Data Warehouse and Analysis Center (Dr Peterson, principal investigator of this subcontract). In this role, the DCRI independently harvests data from each participating STS center, creates a national analysis database, and performs statistical analyses. The proposal for this study was submitted by Dr Peterson to the STS national database publications committee (a 6-member board representing both the STS and the DCRI). After approval, the manuscript was reviewed by the coauthors and a final version approved by the publications committee.

REFERENCES

**Life is not an exact science, it is an art.**
—Samuel Butler (1833-1902)