Hospital Volume and Failure to Rescue With High-risk Surgery

Amir A. Ghaferi, MD, MS, John D. Birkmeyer, MD, and Justin B. Dimick, MD, MPH

Introduction: Although the relationship between surgical volume and mortality is well established, the mechanisms underlying these associations remain uncertain. We sought to determine whether increased mortality at low-volume centers was due to higher complication rates or less success in rescuing patients from complications.

Methods: Using 2005 to 2007 Medicare data, we identified patients undergoing 3 high-risk cancer operations: gastrectomy, pancreatcetomy, and esophagectomy. We first ranked hospitals according to their procedural volume for these operations and divided them into 5 equal groups (quintiles) based on procedure volume cutoffs that most closely resulted in an equal distribution of patients through the quintiles. We then compared the incidence of major complications and “failure to rescue” (ie, case fatality among patients with complications) across hospital quintiles. We performed this analysis for all operations combined and for each operation individually.

Results: With all 3 operations combined, failure to rescue had a much stronger relationship to hospital volume than postoperative complications. Very low-volume (lowest quintile) hospitals had only slightly higher complications rates (42.7% vs. 38.9%; odds ratio 1.17, 95% confidence interval, 1.02–1.33), but markedly higher failure-to-rescue rates (30.3% vs. 13.1%; odds ratio 2.89, 95% confidence interval, 2.40–3.48) compared with very high-volume hospitals (highest quintile). These relationships also held true for individual operations. For example, patients undergoing pancreatectomy at very low-volume hospitals were 1.7 times more likely to have a major complication than those at very high-volume hospitals (38.3% vs. 27.7%, P < 0.05), but 3.2 times more likely to die once those complications had occurred (26.0% vs. 9.9%, P < 0.05).

Conclusions: Differences in mortality between high and low-volume hospitals are not associated with large differences in complication rates. Instead, these differences seem to be associated with the ability of a hospital to effectively rescue patients from complications. Strategies focusing on the timely recognition and management of complications once they occur may be essential to improving outcomes at low-volume hospitals.

Key Words: hospital quality, surgery, patient safety

Despite decades of research devoted to developing better measures, hospital volume remains the most widely used quality indicator in surgery.1–6 Although there is a body of literature that casts doubt on the volume-outcome relationship,7,8 the multitude of policies that have grown from this observed association has continued to give it credence. For example, the Leapfrog group, a large coalition of public and private health care purchasers, has advocated for the concentration of high-risk cancer procedures in high-volume centers.9,10 The National Cancer Policy Board has recommended selective referral to high-volume centers for pancreatectomy and esophagectomy, the 2 cancer procedures with the strongest volume-mortality relationships.11 Finally, professional organizations, such as the American College of Surgeons, use minimum volume standards for designating “centers of excellence” for certain procedures, including bariatric surgery. Although these efforts focus on concentrating patients in high volume centers, the potential adverse consequences of these policies have lead some to advocate for improving care at low-volume hospitals.12,13

However, the mechanisms underlying the superior outcomes at high-volume hospitals remain undefined. Most previous research in understanding volume-mortality relationships has focused on the occurrence of complications.14 However, based on a growing body of literature, there seems to be growing consensus that variations in mortality are due to “failure to rescue” (ie, mortality following a complication) rather than differences in postsurgical complications. It is therefore possible that the differences in mortality between high and low-volume hospitals are also related to a hospital’s ability to quickly recognize and effectively manage a major complication once it occurs.15–17

In this context, we sought to understand the extent to which failure to rescue and complication rates explain variations in mortality between low and high-volume hospitals. Using national Medicare data, we studied 3 high-risk cancer operations, which all have a well documented hospital volume-mortality relationship.
**METHODS**

**Data Source and Patient Population**

We used the Medicare Provider Analysis and Review files from 2005 to 2007. The Centers for Medicare and Medicaid Services maintains this administrative database using all claims that are submitted by hospitals for services provided to Medicare beneficiaries. Each patient record includes information on age, sex, race, admission and discharge dates, principal diagnosis codes, secondary diagnosis codes, procedure codes, and 30-day mortality.\(^ {18} \)

Using the appropriate International Classification of Disease, 9th Revision, Clinical Modification (ICD-9-CM) codes, we identified all patients aged 65 to 99 years undergoing 1 of 3 high-risk cancer operations: gastrectomy, pancreatectomy, and esophagectomy. These operations were selected because of the significant amount of morbidity and mortality associated with each and the previously described volume-outcome relationships.\(^ {1,19–21} \)

**Hospital Volume**

Using 3 years of Medicare data, the average annual individual hospital procedural volume for each of the study procedures was tabulated. Hospitals were then ranked based on their average annual volume. Next, hospitals were divided into quintiles based on procedural volume cutoffs that most closely resulted in an equal distribution of patients through the quintiles. The average (range) annual hospital procedure volume in the very low-volume quintile was <1.3 (range, 1 to 4), <2.0 (range, 1 to 5), and <2.0 (range, 1 to 4) cases per year for esophagectomy, pancreatectomy, and gastrectomy, respectively. Conversely, the average (range) annual hospital procedure volume in the very high-volume quintile was >15 (range, 15 to 102), >27 (range, 27 to 123), and >11 (range, 11 to 110) cases per year for esophagectomy, pancreatectomy, and gastrectomy, respectively.

**Hospital Mortality, Complications, and Failure to Rescue**

Hospital mortality was defined as 30-day or in-hospital mortality. Complications were identified using specific ICD-9-CM codes previously validated by chart review in The Complications Screening Program.\(^ {22} \) Eight major postoperative complications were identified for our study: pulmonary failure (518.81, 518.4, 518.5, 518.8), pneumonia (481, 482.0 to 482.9, 483, 484, 485, 507.0), myocardial infarction (410.00 to 410.91), deep venous thrombosis/pulmonary embolism (415.1, 451.11, 451.19, 451.2, 451.81, 453.8), acute renal failure (584), hemorrhage (998.1), surgical site infection (958.3, 998.3, 998.5, 998.59, 998.51), and gastrointestinal bleeding (530.82, 531.00 to 531.21, 531.40, 531.41, 531.60, 531.61, 532.00 to 532.21, 532.40, 532.41, 532.60, 532.61, 533.00 to 533.21, 533.40, 533.41, 533.60, 533.61, 534.00 to 534.21, 534.40, 534.41, 534.60, 534.61, 535.01, 535.11, 535.21, 535.31, 535.41, 535.51, 535.61, 578.9). The coding of surgical and medical complications, including those identified in our study, was shown to be in good agreement when ICD-9-CM codes and the medical record were compared.\(^ {22,23} \) Finally, failure to rescue was defined as death in a patient with 1 or more of the defined complications. The failure-to-rescue rate for each quintile of hospitals was determined by calculating the proportion of deaths in patients who developed a postoperative complication (numerator) to the total number of patients who developed a postoperative complication (denominator).\(^ {15,24,25} \)

**Statistical Analysis**

We started by determining risk-adjusted mortality rates in each quintile of volume. The risk-adjustment model included the patient age, sex, race, urgency of operation, and comorbidities (C statistics across operations: 0.75 to 0.89). Using the methods of Elixhauser et al,\(^ {26} \) we obtained comorbidities from the secondary diagnosis codes. Using previously described methods,\(^ {24} \) we used logistic regression to predict the probability of death for each patient. Predicted mortality probabilities were then summed for patients at each hospital to estimate expected mortality rates. Next, the ratio of observed-to-expected mortality was multiplied by the overall mortality rate for each operation to obtain the risk-adjusted mortality rate for each hospital. This was repeated for each operation individually and all operations combined.

We then calculated the risk-adjusted complication and failure-to-rescue rates using the same techniques described above. The C-statistics for the complication and failure to rescue models varied by operation with a range of 0.70 to 0.79 and 0.78 to 0.88, respectively. Therefore, our models had very similar discriminatory power. Although we determined the outcomes across all quintiles of volume, we herein present the outcomes associated with the extremes of volume (ie, very high and very low-volume hospitals). In these analyses, we adjusted for the nonindependence of patients within hospitals (clustering) by generating robust standard errors. \( P \) values <0.05 were considered significant in all final analyses and all statistical analyses were performed using STATA 11.0 (College Station, TX).

**RESULTS**

During the years 2005 to 2007, 37,865 patients underwent 1 of the 3 high-risk cancer operations included in our study. Patients in very high and very low-volume hospitals tended to be of similar age and sex. Very low-volume hospitals tended to treat a higher proportion of blacks and slightly sicker patients (Table 1).

Figure 1 demonstrates the relationship between hospital volume and mortality, major complications, and failure to rescue for all the high-risk cancer operations combined.

The risk-adjusted mortality rates across hospitals varied from 2-fold for gastrectomy (7.5% in very high volume vs. 17.7% in very low volume) to nearly 5-fold for pancreatectomy (3.1% in very high volume vs. 13.3% in very low volume) (Table 2). However, major complication rates were similar for 2 of the 3 operations. Pancreatectomy had the largest difference between very high and very low-volume hospitals with a 1.7-fold difference in risk-adjusted major complication rates. Nonetheless, this difference is very small compared with the difference in mortality rates observed for pancreatectomy.
Failure-to-rescue rates, on the other hand, were markedly different between the very low and very high-volume hospitals for all the operations in our study (Table 2). For example, patients who developed a major complication after an esophageal resection had a 3-fold increased odds of death (ie, failure to rescue) in a very low-volume hospital compared with a very high-volume hospital (odds ratio 3.18, 95% confidence interval, 2.39–4.22).

These associations held true for individual complications as well. For example, when comparing the incidence of individual complications in patients undergoing esophagectomy in very low and very high-volume hospitals, the rates are only slightly higher in very low-volume centers. In contrast, the failure-to-rescue rates are higher in the very low-volume centers for each individual complication. However, although there was a higher rate of postoperative hemorrhage in very high-volume centers (3.8% in very high-volume hospitals vs. 2.9% in very low-volume hospitals, \( P < 0.01 \)), there was a lower failure-to-rescue rate in the very high-volume hospitals for these patients (14.9% vs. 23.1%, \( P < 0.01 \)). In other words, high-volume centers were able to rescue a patient from a postoperative bleed more effectively.

When all operations were combined (Table 3), a comparison of medical and surgical complications demonstrated a tendency toward a higher failure-to-rescue rate for medical complications compared with surgical complications. Surgical site infection was the lone surgical complication with a markedly higher failure-to-rescue rate in very low-volume centers. Furthermore, there was a paradoxically lower incidence of surgical site infections in the very low-volume hospitals.

**DISCUSSION**

This study sheds light on the potential clinical mechanisms underlying the relationship between volume and mortality with high-risk surgery. Consistent with previous studies, we found a strong relationship between hospital volume and risk-adjusted mortality for all 3 operations. We did not find a strong association between postoperative complication rates and hospital volume. However, the failure-to-rescue rates were markedly higher in the very low-volume hospitals, compared with the very high-volume hospitals. These findings not only confirm the persistent volume-mortality relationship for high-risk cancer surgery, but also provide insight into possible mechanisms underlying the association and highlight potential areas for targeted quality improvement efforts at low-volume hospitals.
Previous studies seeking to understand the volume outcome effect have yielded little insight into the clinical mechanisms underlying this association. Previous research has focused mainly on structural characteristics of hospitals that may be associated with the volume outcome relationship. For example, Hollenbeck et al. evaluated the relative effect of hospital capacity, staffing, diagnostic services, interventional services, and specialty services on the observed operative mortality rates at high and low-volume hospitals for patients undergoing complex urologic surgery. They found substantial differences in hospital structure between high and low-volume centers that explained nearly 60% of the volume mortality relationship. However, these variables are also proxies for the quality of care and do not provide insight into the mechanisms underlying the volume-mortality relationship. In a previous study examined the relationship between hospital volume and failure to rescue, Smith et al. studied patients undergoing gastrectomy in Texas. They found marked differences in failure to rescue, but little difference in complication rates between high and low-volume hospitals. This study builds on the previous work and documents the relationship for 3 high-risk cancer operations in a national cohort of Medicare patients.

A hospital’s proficiency in minimizing failure to rescue could be related to a variety of factors surrounding the ability of its staff to recognize postoperative complications and take actions necessary to mitigate further complications and death. Such factors include those specific to individual caregivers, to aspects of the clinical microsystem in which postoperative patients receive their care [ie, hospital wards and the intensive care unit (ICU)], and to the broader hospital environment or macrosystem characteristics. Previous research in this area is limited almost exclusively to factors in the last category. Several studies have described relationships between the relative availability of technology, nurse staffing levels, and teaching status, and failure to rescue. However, it is the ability of these resources to integrate and collaborate effectively and efficiently that influences postsurgical care. Other possible attributes that may describe high-volume hospitals and contribute to a postsurgical care environment conducive to effectively rescuing patients include surgeon experience and judgment, presence of experienced house staff or hospitalists at night, or the structure and design of ICUs. Currently, there is little research exploring how these factors are related to failure to rescue.

Our results should be considered in the setting of several limitations. First, by using the Medicare population as our cohort, we limited the analysis to patients aged 65 years and older. Although this may threaten the generalizability of our results, this elderly population accounts for a large proportion of patients undergoing the operations in our study. Second, appropriate risk-adjustment remains a limitation in any study using administrative data. To limit this bias, we used an established method of risk-adjustment set forth by Elixhauser et al. There were statistically significant differences in the racial composition and comorbidity burden between our quintiles of volume (Table 1). Although we attempted to adjust for these differences in our risk adjustment models using the methods described above,

### TABLE 2. Comparison of Overall Mortality, Major Complication Rates, and Failure-to-rescue Rates Between Very High and Very Low-volume Hospitals

<table>
<thead>
<tr>
<th>Operation</th>
<th>Overall Mortality</th>
<th>Major Complications</th>
<th>Failure to Rescue</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR* (95% CI)</td>
<td>OR* (95% CI)</td>
<td>OR* (95% CI)</td>
</tr>
<tr>
<td>Gastrectomy</td>
<td>2.67 (2.24–3.18)</td>
<td>1.05 (0.92–1.19)</td>
<td>2.26 (1.90–2.70)</td>
</tr>
<tr>
<td>Esophagectomy</td>
<td>3.70 (2.74–4.98)</td>
<td>1.35 (1.11–1.65)</td>
<td>3.18 (2.39–4.22)</td>
</tr>
<tr>
<td>Pancreatectomy</td>
<td>4.85 (3.53–6.68)</td>
<td>1.72 (1.39–2.13)</td>
<td>3.21 (2.18–4.72)</td>
</tr>
</tbody>
</table>

*Odds ratios were adjusted for age, sex, race, urgency of operation, and comorbidities. CI indicates confidence interval; OR, odds ratio.

### TABLE 3. Complication Incidence and Failure-to-rescue Rates Between Very Low and Very High-volume Hospitals for all Operations Combined Stratified by Medical and Surgical Complications

<table>
<thead>
<tr>
<th>Complication Incidence</th>
<th>Very Low Volume</th>
<th>Very High Volume</th>
<th>OR (95% CI)</th>
<th>Failure to Rescue</th>
<th>Very Low Volume</th>
<th>Very High Volume</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All complications</td>
<td>42.7%</td>
<td>38.9%</td>
<td>1.17 (1.02–1.33)</td>
<td>30.3%</td>
<td>13.1%</td>
<td>2.89 (2.40–3.48)</td>
<td></td>
</tr>
<tr>
<td>Medical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulmonary failure</td>
<td>16.7%</td>
<td>8.2%</td>
<td>2.25 (1.89–2.68)</td>
<td>47.0%</td>
<td>22.1%</td>
<td>3.13 (2.42–4.05)</td>
<td></td>
</tr>
<tr>
<td>Pneumonia</td>
<td>6.8%</td>
<td>5.1%</td>
<td>1.36 (1.11–1.67)</td>
<td>37.4%</td>
<td>13.7%</td>
<td>2.77 (2.38–5.98)</td>
<td></td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>2.7%</td>
<td>1.5%</td>
<td>1.80 (1.29–2.51)</td>
<td>40.1%</td>
<td>16.7%</td>
<td>3.35 (1.61–6.95)</td>
<td></td>
</tr>
<tr>
<td>DVT/PE</td>
<td>2.4%</td>
<td>2.9%</td>
<td>0.85 (0.66–1.10)</td>
<td>31.6%</td>
<td>10.8%</td>
<td>3.84 (2.16–6.85)</td>
<td></td>
</tr>
<tr>
<td>Acute renal failure</td>
<td>10.6%</td>
<td>4.9%</td>
<td>2.29 (1.86–2.81)</td>
<td>46.8%</td>
<td>31.1%</td>
<td>1.95 (1.42–2.68)</td>
<td></td>
</tr>
<tr>
<td>Surgical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemorrhage</td>
<td>2.9%</td>
<td>3.8%</td>
<td>0.75 (0.61–0.93)</td>
<td>23.1%</td>
<td>14.9%</td>
<td>1.71 (1.05–2.79)</td>
<td></td>
</tr>
<tr>
<td>Surgical site infection</td>
<td>6.7%</td>
<td>8.9%</td>
<td>0.74 (0.62–0.88)</td>
<td>28.9%</td>
<td>8.6%</td>
<td>4.33 (2.96–6.32)</td>
<td></td>
</tr>
<tr>
<td>GI bleed</td>
<td>2.1%</td>
<td>2.3%</td>
<td>1.06 (0.84–1.32)</td>
<td>25.8%</td>
<td>16.1%</td>
<td>1.81 (1.08–3.05)</td>
<td></td>
</tr>
</tbody>
</table>

CI indicates confidence interval; OR, odds ratio.
standard risk adjustment may not account for unobserved differences in patient factors. Although this could potentially account for a fraction of the differences in complications and mortality, there is little concern that these differences would impact the relative importance of complications and failure to rescue in explaining differences in mortality between hospital volume quintiles. Third, the ability to identify postoperative complications using administrative data can prove difficult, especially without “present on admission” codes. In light of this limitation, we chose a subset of complications that have high sensitivity and specificity based on the work of Iezzoni et al.22 Fourth, given our exposure variable of hospital volume, there are inherent limitations in the analysis of low-volume hospitals. The reliability of morbidity and mortality data reported in such low volumes has been well described.33–35 Therefore, our grouping of hospitals into quintiles aims to alleviate this limitation by studying hospitals in aggregate. Thus, our analysis may be difficult to perform in specific hospitals as a measure of individual hospital quality. In addition, given this grouping of hospitals, we did not adjust for specific hospital characteristics that may be associated with surgical outcomes, such as nurse staffing ratios or teaching status. Finally, our definition of failure to rescue15,24,25 may differ from others in the published literature.36–38 Our definition includes mortality after several major medical and surgical complications and is most closely related to the Agency for Healthcare Research and Quality definition.39,40 These complications were selected because of the high sensitivity and specificity of detection in administrative data as described above. In addition, they are nearly identical to the clinical complications set forth in our previously published work, which used prospectively collected clinical data.15

Our findings may have important implications for efforts aimed at improving surgical quality. Although most efforts surrounding hospital volume have focused on the selective referral of patients to high-volume centers, I overlooked opportunity is the improvement of quality at low-volume hospitals. Many critics have described the potential unintended consequences of selective referral policies, with many citing the obstacles to implementation and increased disparities in the quality of care.41–44 Some have cited difficulties with implementation in a fractured United States health care system, and the potential increase in the cost of health care delivery, and finally resistance from patients with variable preferences for care.3,12,45,46 Accordingly, our findings support a policy aimed at improving the care delivered at low-volume hospitals as an alternative to the concentration of care at high-volume hospitals. Failure to rescue is a clinical phenomenon that is influenced by the availability of resources and the processes of care within which those resources are applied. In fact, there are several modifiable elements that could improve failure-to-rescue rates at low-volume hospitals, such as nurse-to-bed ratios, presence of certified intensivists in ICUs, or hospital technology.47 Nonetheless, the existing literature provides very little actionable insight as to what hospitals can do to improve their proficiency in managing patients with postoperative complications. Another perspective may be necessary. Namely, the delivery of high-quality postoperative care can be viewed as an intersection of specific types of hospital resources, and the behaviors and attitudes among caregivers in inpatient wards and ICUs, where postoperative patients receive their care. Further research understanding the safety cultures and mindful practices of these clinical microsystems that distinguish high and low failure-to-rescue hospitals will be the next step in improving care at low-volume hospitals. Ultimately, with failure to rescue, we have provided a potential mechanism through which volume influences outcomes. As such, perhaps future policy efforts should place an increased emphasis on improving a hospital’s ability to successfully rescue a patient from major postoperative complications.

REFERENCES