When Policy Meets Statistics

The Very Real Effect That Questionable Statistical Analysis Has on Limiting Health Care Access for Bariatric Surgery

Edward H. Livingston, MD; Alan C. Elliott, MS; Linda S. Hynan, PhD; Eli Engel, MD, PhD

Hypothesis: Bariatric surgery for Medicare patients must be performed in an accredited hospital that performs at least 125 cases per year. We assessed the validity of this volume threshold and its policy implications.

Design: Using the 2001-2003 National Inpatient Survey, the effect of hospital volume on in-hospital mortality was statistically modeled and the effect of a 125–case per year threshold on access to bariatric surgery was calculated. We performed Monte Carlo modeling to investigate the effect random sampling has on the apparently high mortality rate for low-volume hospitals.

Setting: US inpatient hospitals.

Patients: Patients with hospital discharge codes indicating bariatric surgery.

Main Outcome Measure: In-house mortality.

Results: The observed in-hospital mortality distribution as a function of hospital volume was similar to the expected frequency attributable to random sampling alone. A small number of excess deaths in very low-volume facilities cause statistically significant results for volume-outcome studies. Although 74% of all bariatric surgeries are performed in high-volume centers, 73% of all hospitals currently offering these services are now classified as low volume.

Conclusions: When the results of statistical analysis are used for policy determination, the consequences for patient care may be substantial. Most studies of volume-outcome relationships rely on statistical methods that tend to amplify the effects and few fully characterize their statistical models. Despite the weak evidence for a volume-outcome relationship for bariatric surgery, a 125–case per year threshold has been set for center-of-excellence status, which eliminates most hospitals currently providing these services and disproportionately restricts access for the poor and underinsured.

Arch Surg. 2007;142(10):979-987

BARIATRIC PROCEDURES ARE BECOMING MORE FREQUENT, WITH GREATER ACCEPTANCE FOR THESE OPERATIONS. ABUNDANT INFORMATION IS AVAILABLE ESTABLISHING THE SAFETY AND EFFICACY OF THESE PROCEDURES. MOST IMPORTANTLY, ALTHOUGH THE OPERATIONS ARE PERFORMED IN A POPULATION OF PATIENTS WITH AN INHERENTLY HIGH RISK FOR COMPLICATIONS, NONSURGICAL THERAPIES ARE UNIFORMLY UNSUCCESSFUL IN CAUSING SUSTAINED WEIGHT LOSS FOR THE MORBIDLY OBSESE, JUSTIFYING THE OPERATIVE RISK. SEVERAL RECENT LARGE-SCALE REVIEWS HAVE CONCLUDED THAT BARIATRIC PROCEDURES ARE THE ONLY EFFECTIVE MEANS FOR CAUSING SUSTAINED WEIGHT LOSS WITH GOOD COMORBIDITY CONTROL FOR SERIOUSLY OBSESE INDIVIDUALS, AND FOR MOST OF THE AVAILABLE OPERATIONS, THERE IS A FAVORABLE RISK-BENEFIT BALANCE.11

For most high-risk procedures, both hospital and surgeon volume are thought to be important in influencing outcomes. Most of the procedures studied to date have been those with high mortality such as cardiac bypass and esophageal and pancreatic resections. These studies have been carried out using administrative databases derived from hospital discharge coding information. These databases are advantageous since they contain information regarding the hospitalizations for very large numbers of patients. Limitations exist in that they contain sparse data for any individual patient. Discharge coding is subject to erroneous disease or procedure classification. Nevertheless, based on an abundance of published reports, volume-outcome relationships have become widely accepted.

There have been several reports demonstrating better outcomes from high-volume bariatric surgery centers when compared with lower-volume facilities. These findings have translated into a call for minimum volume requirements for approved bariatric surgery centers by certifying bodies. Most bariatric surgery is per-

Author Affiliations: Divisions of Gastrointestinal and Endocrine Surgery (Dr Livingston) and Biostatistics (Drs Elliott and Hynan), University of Texas Southwestern School of Medicine and the Veterans Administration North Texas Health Care System, Dallas; and the Department of Biomathematics, University of California, Los Angeles (Dr Engel).
formed in what would be considered low-volume hospitals. Adoption of minimum volume standards has the effect of limiting access for this form of therapy. Despite the rise in the numbers of bariatric operations performed annually, they dwarf the total population of eligible patients. Consequently, most bariatric surgery programs are inundated, resulting in long wait times before these operations are performed, a problem that is worsened by restricting access to care because of minimum volume standards.

Given the very real policy implications of the volume effect for bariatric surgery outcomes, we examined the impact that the currently recommended minimum volume thresholds have on the practice of bariatric surgery. We also performed Monte Carlo simulation to investigate the effect random sampling has on apparent volume-outcome relationships.

METHODS

SYSTEMATIC LITERATURE REVIEW

The English-language literature was searched with PubMed. A search for relevant articles was conducted by entering the terms “bariatric surgery” and “volume-outcome relationship.” For each relevant article that was found, a “related article” search was performed. Additionally, all articles relevant to this investigation were entered into the ISI Citation Index Database and all published articles citing them were reviewed.

DATABASE ACQUISITION AND CASE IDENTIFICATION

The National Inpatient Survey (NIS) was obtained from the Agency for Healthcare Research and Quality (AHRQ) for the years 2001 to 2003. These years were selected because they had very large numbers of bariatric procedures performed. The NIS is a population-representative sampling of hospital discharges obtained from 20% of all the hospitalizations in the United States in any given year. In contrast to the other large hospital discharge database (the National Hospital Discharge Survey) that samples a fraction of discharges from any given hospital, the NIS obtains information about all discharges from a select number of facilities in the United States. This has the advantage of providing the full spectrum of activity from hospitals in various regions. Although statistically corrected to be population representative, data are only collected from hospitals in 29 states and ethnicity data are often times incomplete. Thus, there are regions and populations that may not be well represented in the NIS. The major advantage of the NIS is the completeness of information from individual hospitals, facilitating volume-outcome analysis. The NIS is representative of the totality of hospital care delivered in the United States since it includes discharge information from a variety of hospital types, including small community facilities and large academic teaching hospitals.

Bariatric procedures were identified by discharges encoded with Diagnostic Related Group (DRG) 288. The DRGs are assigned at the time of a patient’s discharge based on combinations of diagnostic and procedure codes along with information from the patient’s medical record to best reflect the primary reason a patient was hospitalized. Diagnostic Related Group 288 (operating room procedures for obesity) is only used when the primary reason for hospitalization was to undergo procedures related to morbid obesity. Thus, any patient encoded with DRG 288 was admitted with the intent to perform an operation for a problem related to a patient’s morbid obesity. Procedures associated with plastic surgical procedures (ie, those with International Classification of Diseases, Ninth Revision, Clinical Modification [ICD-9-CM] procedure codes ranging from 85.XX to 86.XX) were excluded. Operations encoded as 44.XX and 45.XX along with DRG 288 were defined as bariatric surgical procedures.

Poverty was defined as living in a zip code region associated with a median income of less than $25,000/y for 2001 and 2002 or less than $35,000/y for 2003. In 2003, the AHRQ changed the income threshold for the definition of poverty. The bariatric surgery case volume for each facility was counted for each 1-year interval of the study. Each year was treated separately with no averaging between years.

MONTE CARLO SIMULATION OF THE EXPECTED SAMPLING DISTRIBUTION FOR MORTALITY IN VOLUME-OUTCOME STUDIES

We performed Monte Carlo simulation to test the effect of sampling on observed mortality rates. A 1000-member binary file set for an event rate under study was randomly sampled. Bariatric surgery mortality for the cohort studied was 0.2%. The binary file contained 2 “events” and 998 “nonevents.” The modeling program we wrote randomly generated a procedure volume from 1 to 500 cases per year, then iteratively sampled the binary file in random locations. This cycle was repeated n times, with n being the procedure volume being modeled. The sum of events was divided by the number of annual cases being modeled, yielding a simulated event rate. A file of annual case volumes and modeled event rates was generated with 1000 randomly generated volume–event rate pairs. A second file simulating individual patient events and hospital volumes was created with 65,535 cases. This file contained an event or non-event paired with the simulated hospital volume from which the case record was derived. The computer program generating these files was written in Visual Basic and executed in Microsoft Excel (Microsoft Inc, Redmond, Washington).

The 95% confidence interval (CI) for a Poisson process with an event rate of 0.22% was determined. The expected number of events for a volume category was calculated (volume $\times 0.0022$). The lower confidence limit for this was determined from the inverse gamma function for $\alpha = 0.023$ and the value for the expected number of deaths for the volume category. The upper limit was determined from the inverse gamma function with $\alpha = 0.973$ and the expected mortality + 1.11 The CIs were converted to percentages of mortality by dividing the expected mortality by the respective volume category and multiplying by 100. This was repeated for hospital volumes ranging from 1 to 500 cases per year.

REGRESSION ANALYSIS OF THE VOLUME-OUTCOME RELATIONSHIP

Logistic regression modeling was used to assess the impact annual hospital procedure volume had on bariatric surgery mortality. In-hospital death was entered into the logistic regression equation as the dependant variable. Age and male sex were entered into the regression equations since these are well-recognized risk factors for poor outcomes from bariatric surgery. Previous studies we conducted demonstrated that the Charlson and Elixhauser comorbidity models for risk adjustment do not adequately model bariatric surgery deaths and, therefore, were not used.16 Odds ratios, P values, and the C statistic were reported for each regression.

We tested the impact of excess numbers of deaths on the results of regression analysis for volume-outcome studies. The
simulated data were examined by logistic regression to test the effect adding more than the number of expected deaths had on apparent volume-outcome relationships. Hospital volumes tested were those higher and lower than 125 cases per year since this is the minimum volume proscribed by the Surgical Review Corporation and American College of Surgeons certifying bodies. Some volume-outcome studies use $\chi^2$ analysis rather than regression. Thus, $\chi^2$ testing was performed to compare outcomes higher and lower than 125 cases per year. Since most bariatric surgery studies used 3 volume categories (<50, 50-100, and >100 cases per year), these were also tested in the regression model.

**HIERARCHICAL MODELING**

Simultaneous assessment of hospital and individual-patient factors contributing to postoperative in-hospital mortality beyond simple procedure volume effects were modeled using the SAS procedure GLIMMIX (SAS Institute Inc, Cary, North Carolina). An intercept-only model was constructed using individual hospitals as a random effect with no fixed effects to establish the variability in mortality from hospital to hospital. A second model was constructed adding the zero-centered number of cases per year (number of cases performed per year per hospital minus the mean number of operations per hospital for all hospitals) for individual hospitals as a fixed effect. The intraclass correlation was calculated by dividing the intercept covariance parameter estimates for the first model by the same estimate added to the residual covariance estimate for the model.17 Reduction of the intercept covariance parameter estimate by the annual procedure volume effect was determined by subtracting the intercept covariance parameter estimates derived from the second model from the first model and dividing by the estimate from the first model.17

**RESULTS**

The combined NIS 2001 to 2003 databases had 22 758 985 records. These were used as the basis for all subsequent analysis. There were 51 842 weight loss operations within DRG 288, with 84% of the procedures performed on female patients. These operations were carried out in 741 hospitals. Some hospitals had been reported for more than 1 year in the database. We assessed each year individually without averaging years, resulting in 921 hospital-years. The median age was 41 years. Within this cohort, there were 115 in-hospital deaths, yielding a mean in-hospital mortality rate of 0.22%. The variance was 0.22, suggesting a Poisson distribution of the mortality-volume relationship. There were progressively more gastric bypasses entered into the database in each year we examined: 11 846 in 2001, 15 950 in 2002, and 24 046 in 2003. Since the NIS database is population representative, these correspond to an annual US case volume for bariatric surgery of 58 672 in 2001, 77 335 in 2002, and 111 581 in 2003.

Logistic regression on in-hospital mortality corrected for age and male sex and accounting for overdispersion of the data yielded an odds ratio of 2.221 (95% CI, 1.428-3.450; $P <.001$; C index = 0.70) for low-volume hospitals (<125 cases per year) compared with high volume. When volume was entered into the regression as a continuous variable, the odds ratio was 1.002 (95% CI, 1.001-1.004; $P <.001$; C index = 0.71). This means that for every unit decrease in the number of cases performed annually the odds for an in-house mortality increase by 1.002. Dividing the number of cases performed per year by 50 yields an odds ratio of 1.127 (95% CI, 1.060-1.198), meaning that for every decrease in 50 cases per year for a facility the odds for an in-house mortality increase by 1.127.18

Because the volumes were increasing with time, we examined the distribution of cases by facility volume for 2003 only. Figure 1 demonstrates the distribution of case and hospital volume as a function of hospital volume categories. Approximately three-quarters of all bariatric procedures are performed in high-volume hospitals. B. The proportion of hospitals of the various volume categories. Three-quarters of all hospitals performing bariatric procedures are classified as low volume.

![Figure 1](https://example.com/figure1.png)
underwent their procedures in facilities with fewer than 125 cases per year. Most of the poor (60%) and many Medicare patients (32%) underwent bariatric surgery in facilities performing fewer than 125 cases per year. The in-hospital mortality rate for procedures performed in low-volume centers was 0.34%, meaning that 99.66% of all operations performed in low-volume centers had no mortality.

There is a disproportionate effect of volume standards in rural or suburban facilities. Six thousand of the 111,000 bariatric operations (5%) performed in 2003 were carried out in low-volume rural facilities. Of these 6000 patients, 2700 (45%) are classified as poor, with 11% insured by Medicaid and 12%, by Medicare. This contrasts with high-volume urban medical centers, in which only 15% of the population is poor, with 5% insured by Medicaid and 7%, by Medicare.

In a review of the literature, 6 articles investigated the volume-outcome relationship for bariatric surgery morbidity outcomes. One publication was not considered since it only assessed the procedure volume for Medicare-insured patients and not the total surgeon or hospital volume.20 Whatever the features are associating volume to outcome, they result from the sum of a surgeon’s experience and cannot be reflective of a single payer’s volumes, given the heterogeneity of a surgeon’s insurance case mix. For these reasons, we only considered 5 articles (Table 1).

These represent a heterogeneous group, each with its own unique definition of postoperative morbidity. They also differed in the statistical methods used for assessing the relationship and risk adjustment models. Only 1 of the articles described any sort of model testing (C index) to assess the validity of the statistical model for volume-outcome relationship testing.21 Of these articles, only 2 modeled bariatric surgery deaths.21,22 None used an objective method (eg, terciles) for determining cutoffs for definitions of high or low volume.

Figure 3 shows the results of Monte Carlo modeling for expected death rates as a function of hospital volume. The circles represent simulated expected death rates one would observe for an overall in-hospital mortality rate of 0.2%. The triangles are the actual observed death rates for bariatric surgery as acquired from the NIS database. Both sets of points substantially overlap one another. Both observed and simulated data fall within the 95% CI for Poisson-distributed data with an expected in-hospital mortality rate of 0.22%. This means that as volume decreases the uncertainty of the true mortality rate based on observed mortality rate from low-volume facilities rises. The uncertainty is especially great with very low-volume facilities.

The basis for the progressively increasing uncertainty and the general rise in the percentage of observed and modeled mortality is illustrated in Table 2. A high-volume facility performing 450 cases annually that has 1 death will have 0.22% mortality. Nine lower-volume facilities performing 50 cases per year will also have an aggregated mortality of 0.22% if there is 1 death among all of them. The single facility with the death will have a mortality rate of 2.0%. The other 9 have 0% mortality.

Because most volume-outcome studies use logistic regression to establish this relationship, we tested the sensitivity of the technique to excess deaths in low-volume facilities by modeling our simulated data. We entered the simulated data into a regression equation and used volume higher or lower than 125 cases per year as an independent variable. We also simulated high vs low volume comparing facilities performing fewer than 50 cases per year with those with more than 100 cases per year since several of the bariatric surgery volume-outcome studies used these cutoff values. Results are presented in Table 3. The simulated data had 65,536 cases with 143 deaths resulting in a death rate of 0.22%. Addition of 6 extra deaths to low-volume facilities resulted in a statistically significant difference between high-volume (>125 cases per year) and low-volume (<125 cases per year) hospitals. Addition of 7 extra deaths caused the logistic regression to be significant with this volume cutoff. When fewer than 50 cases per year were compared with more than 100 cases per year, the addition of 5 extra deaths resulted in statistically significant volume effects as determined by logistic regression.

Hierarchical modeling to determine the effect of a hospital itself on outcomes revealed a substantial effect. Modeling the hospital (ie, random) effect yielded a mean±SE intercept covariance estimate of 2.738±0.346 with a residual covariance of mean±SE 0.439±0.003.
The resultant intraclass correlation was 86.4%. This suggests a substantial amount of clustering of deaths among hospitals. Addition of the number of cases performed per year into the model reduced the intercept covariance to mean ± SE 2.123 ± 0.291. This reduced the intrahospital variance by 32.4%. A fully loaded model that adjusted for male sex, age, and Medicare status as well as the hospital volume only reduced the intrahospital variance by 32.4%.

**COMMENT**

Using the largest database yet assembled of bariatric surgery cases, we assessed the relationship between mortality and facility volume. In contrast to prior studies, we found that although volume appears to be a significant factor explaining mortality, the effect is small and most likely results from relatively few extra deaths in low-volume facilities. Using Monte Carlo–simulated data and Poisson CIs, we have shown that the apparently higher mortality rates for lower-volume facilities result from a sampling phenomenon. As sample size decreases (low volume), the uncertainty of the actual death rate estimated from the measured death rate increases substantially. Relatively few extra deaths in low-volume facilities can result in significant volume effects when tested by χ² or logistic regression analysis. Graphic observation of observed vs simulated mortality data reveal that there is substantial overlap.

If there is a volume-outcome effect for bariatric surgery, it is small. A disproportionate number of patients cared for in low-volume facilities are poor and have Medicare or Medicaid insurance. These individuals are less likely to have resources or the ability to travel to high-volume facilities. This is especially problematic for economically disadvantaged patients in rural areas. Travel to a distant urban center may be impossible for them. There is proportionately less bariatric surgical care delivered to the disadvantaged relative to those of means. Imposition of a minimum volume standard will further limit access of care for the poor, a population that stands to gain the most from bariatric surgery.
Figure 3. Simulated and observed distributions for bariatric surgery mortality. The solid circles represent simulated data that are reflective of the sampling distribution for typical mortality information. The inverted triangles are the observed in-house mortality rates for bariatric surgery. The Poisson 95% confidence interval (CI) for the expected percentage of mortality per hospital volume category is shown with the solid lines. The CI was calculated for an expected mortality rate of 0.22%. The CI widens considerably for low-volume facilities because of the substantial degree of uncertainty in knowing a true mortality rate for a low-volume facility, secondary to sampling uncertainty. The modeled and observed mortality rates all fall within the 95% Poisson CI for sampling count data, with an expected 0.22% mortality rate.

The current minimum volume standard for the Surgical Review Corporation’s bariatric surgery centers of excellence designation or for level I American College of Surgeons–designated bariatric surgery centers is 125 cases per year. This number was not derived from findings published in the literature because none of the studies prior to adopting this standard tested this threshold. The 125 cases per year standard has the potential for limited access to care for many poor or underinsured patients who traditionally receive their care in low-volume facilities. Although nearly three-quarters of all bariatric operations are performed in high-volume hospitals, three-quarters of all hospitals performing bariatric surgery in 2003 are categorized as low volume. Survival in low-volume facilities was 99.66%, yet they now may be restricted in their ability to provide bariatric services. Our findings suggest that the volume-outcome relationship is quantitatively small, yet when translated into policy, it can seriously restrict access to care for patients with limited resources. An unmeasured downstream effect is the loss of expertise in caring for morbidly obese patients. If the lower-volume facilities cannot provide bariatric surgery care, they may become ill equipped to manage morbidly obese patients needing other types of care. With the rapid increase in the population’s obesity, hospitals and their staff must be equipped to manage large patients.

There is a substantial body of literature regarding volume-outcome relationships.25 Study designs lack consistency, making their results difficult to compare and interpret.26 Most studies have used administrative databases to capitalize on the large number of patients available for study. They are limited because administrative databases have sparse clinical information. Because of inconsistent discharge coding, they may not contain all the clinical information necessary for appropriate risk adjustment. High-volume centers tend to have a greater proportion of lower-risk patients than low-volume centers. Attempts to correct for this are made by risk adjusting volume-outcome analyses. However, the tools used for risk adjustment are rarely disease specific and, because discharge coding is not highly accurate, may not adequately characterize the absolute risk for adverse outcomes for an individual patient. Our prior work found that the frequently used Charlson score and Elixhauser methods were ineffective for risk adjustment for bariatric surgery mortality.16 Proper risk adjustment is especially problematic for bariatric surgery since one of the most important risk factors for adverse outcomes is large body size.27-29 This parameter is not available in administrative databases, highlighting the incompleteness of risk adjustment that can be achieved. Volume-outcome studies with progressively improved risk-adjustment modeling widened the CIs for volume-effect odds ratios.30 Thus, the high-volume effect weakened as risk adjustment improved.

Most volume-outcome studies find relatively weak effects. Odds ratios for these effects are relatively small and C indexes demonstrate models perform only moderately well. Most articles using multivariate logistic regression to assess outcomes use these techniques inappropriately or fail to report important aspects of the model’s characteristics.31 Since these models can be misleading, it is important to fully characterize them, something rarely reported in the volume-outcome literature. When model characteristics are shown, they tend to show relative poor performance by the predictive models. This is caused by insufficient clinical information available in administrative databases, precluding adequate characterizing of patients and their comorbidities. Volume-outcome relationships generally disappear when clinical databases are used that contain high-fidelity disease-specific clinical information, allowing for high-quality risk adjustment. Several large-scale studies have found that the volume-outcome effect disappears when

Table 2. Illustration of the Sampling Effect on Apparent Mortality

<table>
<thead>
<tr>
<th>No. of Cases per Year</th>
<th>No. of Deaths</th>
<th>Mortality, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>450^a</td>
<td>1</td>
<td>0.22</td>
</tr>
<tr>
<td>450^b</td>
<td>1</td>
<td>0.22</td>
</tr>
<tr>
<td>50^b</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>50^b</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>50^b</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>50^b</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>50^b</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>50^b</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>450^c</td>
<td>1</td>
<td>0.22</td>
</tr>
</tbody>
</table>

^a If there is 1 death in a year for a facility where 450 cases per year are performed, the mortality will be 0.22%.
^b If one examines 9 hospitals that year, each performing 50 cases annually, and there is a death at 1 of them, that hospital’s annual mortality will be 2.00% compared with 0.00% for the others.
^c In aggregate, the mortality for the 9 hospitals is 0.22%.
Abbreviations: CI, confidence interval; NA, not applicable; OR, odds ratio.

a Randomly generated data to model a 0.2% mortality rate. There were 65,536 data points corresponding to individual patients. The algorithm resulted in 143 deaths for a 0.22% mortality rate. The effect of adding extra deaths (more than random) to the very lowest hospital volume categories is examined in this Table.

b For data stratified into hospital volume greater or less than 125 cases annually.

c Probability that there is no difference in mortality between volume of 125 cases per year or higher and fewer than 125 cases per year.

d Odds ratio for mortality in low-volume vs high-volume hospital (≥ 125 cases per year).

e Odds ratio compared with volumes of more than 100 cases per year.

Table 3. Effect of Additional Deaths on Regression and χ² Models of Volume-Outcome Relationships

<table>
<thead>
<tr>
<th>No. of Added Deaths</th>
<th>Deaths, %, Volume &lt; 125 Cases per year</th>
<th>Deaths, %, Volume &gt; 125 Cases per year</th>
<th>Absolute Difference</th>
<th>χ²</th>
<th>OR (95% CI), Volume &lt; 125 Cases per year</th>
<th>P Value</th>
<th>C Index</th>
<th>OR (95% CI), Volume = 50-100 Cases per year</th>
<th>P Value</th>
<th>C Index</th>
<th>OR (95% CI), Volume 50-100 Cases per year</th>
<th>P Value</th>
<th>C Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.22</td>
<td>0.22</td>
<td>0</td>
<td>0.964</td>
<td>1.017 (0.518-1.999)</td>
<td>0.96 &lt; 0.001</td>
<td>NA</td>
<td>0.97</td>
<td>1.523 (0.672-3.454)</td>
<td>0.31</td>
<td>0.505</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.25</td>
<td>0.22</td>
<td>0.03</td>
<td>0.708</td>
<td>1.131 (0.594-2.151)</td>
<td>0.71 NA</td>
<td>0.713 (0.100-5.158)</td>
<td>0.74</td>
<td>1.523 (0.672-3.454)</td>
<td>0.74</td>
<td>0.501</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.27</td>
<td>0.22</td>
<td>0.05</td>
<td>0.461</td>
<td>1.244 (0.672-2.303)</td>
<td>0.49 NA</td>
<td>1.429 (0.353-5.784)</td>
<td>0.62</td>
<td>1.523 (0.672-3.454)</td>
<td>0.31</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.3</td>
<td>0.22</td>
<td>0.08</td>
<td>0.39</td>
<td>1.357 (0.751-2.452)</td>
<td>0.31 NA</td>
<td>2.150 (0.684-6.761)</td>
<td>0.11</td>
<td>1.523 (0.672-3.454)</td>
<td>0.31</td>
<td>0.505</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.32</td>
<td>0.22</td>
<td>0.11</td>
<td>0.189</td>
<td>1.471 (0.832-2.602)</td>
<td>0.18 NA</td>
<td>2.867 (1.058-7.772)</td>
<td>0.38</td>
<td>1.523 (0.672-3.454)</td>
<td>0.31</td>
<td>0.509</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.34</td>
<td>0.22</td>
<td>0.12</td>
<td>0.098</td>
<td>1.584 (0.913-2.750)</td>
<td>0.10 NA</td>
<td>3.589 (1.465-8.790)</td>
<td>0.005</td>
<td>1.523 (0.672-3.454)</td>
<td>0.31</td>
<td>0.512</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.37</td>
<td>0.22</td>
<td>0.15</td>
<td>0.0496</td>
<td>1.689 (0.995-2.898)</td>
<td>0.052 NA</td>
<td>4.315 (1.898-9.807)</td>
<td>0.001</td>
<td>1.523 (0.672-3.454)</td>
<td>0.31</td>
<td>0.515</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.39</td>
<td>0.22</td>
<td>0.17</td>
<td>0.0229</td>
<td>1.182 (1.078-3.046)</td>
<td>0.03 NA</td>
<td>5.041 (2.349-10.816)</td>
<td>&lt;0.001</td>
<td>1.523 (0.672-3.454)</td>
<td>0.31</td>
<td>0.518</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.42</td>
<td>0.22</td>
<td>0.2</td>
<td>0.0097</td>
<td>1.926 (1.162-3.193)</td>
<td>0.01 NA</td>
<td>5.769 (2.816-11.819)</td>
<td>&lt;0.001</td>
<td>1.523 (0.672-3.454)</td>
<td>0.31</td>
<td>0.522</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.44</td>
<td>0.22</td>
<td>0.22</td>
<td>0.0038</td>
<td>2.040 (1.246-3.340)</td>
<td>0.005 NA</td>
<td>6.501 (3.297-12.828)</td>
<td>&lt;0.001</td>
<td>1.523 (0.672-3.454)</td>
<td>0.31</td>
<td>0.525</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.47</td>
<td>0.22</td>
<td>0.25</td>
<td>0.0014</td>
<td>2.153 (1.330-3.485)</td>
<td>0.002 0.531</td>
<td>7.237 (3.790-13.816)</td>
<td>&lt;0.001</td>
<td>1.523 (0.672-3.454)</td>
<td>0.31</td>
<td>0.528</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

high-quality risk adjustment is applied or when highly refined statistical methods are applied to the data analysis.25,32-46 We found that when volume is represented as a continuous variable for observed mortality following bariatric surgery the odds ratios remained statistically significant but were 1.00, demonstrating that the magnitude of the odds ratio depends on how the variable is defined. Since a continuous variable explains more of the observed variance than a categorical one, the odds ratio derived from a continuous variable is a more reliable indicator of its effect.41

Prior studies of bariatric surgery volume-outcome relationships modeled procedure complications between outcomes. They range from 0.5 (no discrimination) to 1.0 (perfect classification of outcomes by the model), with a C statistic greater than 0.75 considered to be indicative of reasonable model performance.24 In that study, the C statistic ranged from 0.60 to 0.70. The C statistic provides an index of how well a model can discriminate between outcomes. They range from 0.5 (no discrimination) to 1.0 (perfect classification of outcomes by the model), with a C statistic greater than 0.75 considered to be indicative of reasonable model performance.24

All prior bariatric surgery volume-outcome studies also found that the number of bariatric surgery procedures was increasing with time. Some found a secular decrease in comorbidity scores.20 Acceptance of bariatric surgery by the medical and insurance communities has been recent. In the past, only very ill patients were referred for these operations. With time, individual bariatric surgery centers have become busier with an enlarging population of lower-risk patients. Prior studies have consistently found that in higher-volume centers patients had lower comorbidity scores than low-volume centers. Since the statistical models from which odds ratios and other statistical measures were derived have not been fully characterized, it is possible that the lower risk associated with these changes was not completely explained by the models. Not fully accounting for surgical risk would lead to falsely concluding that high volume is associated with better outcomes.

Only 1 study presented information regarding the adequacy of the regression model to describe bariatric surgery volume-outcome relationships.42 In that study, the C statistic ranged from 0.60 to 0.70. The C statistic provides an index of how well a model can discriminate between outcomes. They range from 0.5 (no discrimination) to 1.0 (perfect classification of outcomes by the model), with a C statistic greater than 0.75 considered to be indicative of reasonable model performance.24 We have shown that the addition of a few extra deaths can result in significant odds ratios when logistic regression is used. These extra deaths or morbidities are more likely to be observed in low-volume facilities because of the sampling uncertainties that have been described by others25 and highlighted with our simulation studies. Conceivably, better characterization of the patients undergoing bariatric surgery would yield better models. These improved models might erase any volume effect. When risk
adjustment improves, more sophisticated statistical methods are used, or when clinical databases with high-quality risk adjustment are assessed, high volume ceases to be predictive of mortality.25,32-40

The attributes and limitations of volume-outcome studies have been comprehensively reviewed.25 Many questions remain about the true effect procedure volume has on outcomes. There is consensus that volume per se does not influence outcomes. More likely, accumulated experience and the establishment of multidisciplinary teams result in better outcomes. Theoretically, processes and other factors associated with high volume could be identified and translated to low-volume hospitals. We have found the volume-outcome relationship for bariatric surgery is weak. Most deaths occurred within the CIs expected for sampling uncertainty, which increase as volumes diminish. Imposing policies limiting surgery to high-volume facilities will reduce access to care for those who stand to benefit the most from these operations: the uninsured and impoverished. The current volume standard of 125 cases per year eliminates three-quarters of all hospitals currently performing bariatric operations. The benefits of this sacrifice are equivocal. Greater experience managing obese patients at facilities performing bariatric surgery results in better care for obese patients with nonbariatric problems. Regionalization results in the global reduction of experience in managing obese patients at a time when obesity is rapidly increasing in the population.

Hierarchical modeling enables the simultaneous assessment of individual and group effects. When applied to the NIS data set, substantial covariance between hospitals was observed, suggesting that mortality tends to cluster at certain facilities or in hospitals that share certain characteristics. Facility procedure volume resulted in a modest decrease in the clustering effect as did adjustment with the known bariatric surgery mortality risk factors age, male sex, and Medicare status. These findings suggest that hospital characteristics other than these account for bariatric surgery mortality. These factors could be the surgeons’ skills at those hospitals, the type of patients accepted for surgery, or other as of yet unidentified processes of care that contribute to mortality. Another important observation derived from this modeling was the substantial clustering observed. Regression methods require independence of data from one observation to another. Clustering violates the independence assumption that serves as the basis for most regression techniques, leading to inaccurate parameter estimates for the statistical models that have been used for volume-outcome studies.

Since high volume is only a proxy for other processes that determine outcomes, it is imperative that these be identified. Denying care at low-volume hospitals is not acceptable if those facilities can mimic the processes responsible for good outcomes at higher-volume facilities. Since the vast majority of low-volume hospitals had no in-hospital mortality, they should be able to provide bariatric care. Rather than restrict access to care by arbitrary volume criteria, centers should be evaluated by their risk-adjusted outcomes.41 Implementation of mature programs such as the National Surgical Quality Im-
provement Program44-45 achieves this and should replace volume criteria for the certification of bariatric surgery centers.

Accepted for Publication: June 2, 2007.

Correspondence: Edward H. Livingston, MD, Division of Gastrointestinal and Endocrine Surgery, UT Southwestern Medical Center, 5323 Harry Hines Blvd, Room E7–126, Dallas, TX 75390-9136 (edward.livingston@utsouthwestern.edu).

Author Contributions: Study concept and design: Livingston and Engel. Acquisition of data: Livingston. Analysis and interpretation of data: Livingston, Elliott, and Hynan. Drafting of the manuscript: Livingston. Critical revision of the manuscript for important intellectual content: Livingston, Elliott, Hynan, and Engel. Statistical analysis: Livingston, Elliott, Hynan, and Engel. Obtained funding: Livingston. Administrative, technical, and material support: Livingston.

Financial Disclosure: None reported.

REFERENCES


