The effect of volume on esophageal cancer resections: What constitutes acceptable resection volumes for centers of excellence?

Robert A. Meguid, MD, MPH,^a Eric S. Weiss, MD,^a David C. Chang, PhD, MPH, MBA,^b Malcolm V. Brock, MD,^a and Steven C. Yang, MD^a

Objective: Volume–outcome relationships for esophageal cancer resection have been well described with centers of excellence defined by volume. No consensus exists for what constitutes a “high-volume” center. We aim to determine if an objective evidence-based threshold of operative volume associated with improvement in operative outcome for esophageal resections can be defined.

Methods: Retrospective analysis was performed on patients undergoing esophageal resection for cancer in the 1998 to 2005 Nationwide Inpatient Sample. A series of multivariable analyses were performed, changing the resection volume cutoff to account for the range of annual hospital resections. The goodness of fit of each model was compared by pseudo r^2 , the amount of data variance explained by each model.

Results: A total of 4080 patients underwent esophageal resection. The median annual hospital resection volume was 4 (range: 1–34). The mortality rate of “high-volume” centers ranged from 9.94% (≥ 2 resection/year) to 1.56% (≥ 30 resections/year). The best model was with an annual hospital resection volume greater than or equal to 15 (3.87% of data variance explained). The difference in goodness of fit between the best model and other models with different volume cutoffs was 0.64%, suggesting that volume explains less than 1% of variance in perioperative death.

Conclusion: Our data do not support the use of volume cutoffs for defining centers of excellence for esophageal cancer resections. Although volume has an incremental impact on mortality, volume alone is insufficient for defining centers of excellence. Volume seems to function as an imperfect surrogate for other variables, which may better define centers of excellence. Additional work is needed to identify these variables.

See related article on page 10.

Resection of the esophagus, either total or partial, is a complex surgical procedure that carries a relatively high risk of operative mortality. Because of this, a significant body of work has focused on the relationship between volume and outcome for esophageal resections. The beneficial effect of increased volume of esophagectomy on outcome has been clearly demonstrated in multiple studies.^{1–4} On the basis of the results of these and similar studies, esophageal resection has been identified as a potential procedure for volume-based regionalization, and as such resection volume has been proposed as a measurement for defining centers of

excellence. An example of this is the Leapfrog Group, which defined criteria for “evidence-based hospital referral” for esophageal resection as hospitals performing a minimum of 13 resections per year.⁵

In addition to the volume cutoff for esophageal resections set by the Leapfrog Group, various other thresholds for defining high-volume centers have been used in the literature. These annual hospital volume thresholds range from 6 to 20 esophageal resections per year.^{2,6,7} However, these cutoff points have often been imprecisely or arbitrarily defined, and there are little data to support the use of specific volume cutoffs.

The aim of this study was to determine if an objective, evidence-based threshold of operative volume associated with improved hospital-level outcomes for esophageal resection for cancer could be defined. Should this threshold be identified, it could potentially be considered a candidate in the criteria for defining high-volume hospitals for esophageal resection.

MATERIALS AND METHODS

Data Source

A retrospective analysis was performed using patient data collected from the Nationwide Inpatient Sample (NIS) file between 1998 and 2003. The NIS database comprises discharge records approximating a 20% sample of hospital discharges in the United States and is maintained by the Agency for Healthcare Research and Quality as part of the Healthcare Cost and Utilization Project.⁸ It approximates 7 million patient discharge records per year, originating from approximately 1000 different hospitals per year nationwide. Data available within the NIS include patient and hospital demographics, payer information, treating and concomitant diagnoses, inpatient

From the Division of Thoracic Surgery^a and Department of Surgery,^b Johns Hopkins School of Medicine, Baltimore, Md.

Drs Robert A. Meguid and Eric S. Weiss were supported by the Ruth L. Kirschstein National Research Service Award (T32DK007713) while undertaking this study. Dr Eric S. Weiss is an Irene Piccinini Investigator in Cardiac Surgery.

Read at the Eighty-seventh Annual Meeting of The American Association for Thoracic Surgery, Washington, DC, May 5–9, 2007.

Received for publication May 5, 2007; revisions received June 20, 2008; accepted for publication Sept 16, 2008.

Address for reprints: Robert A. Meguid, MD, MPH, Division of Thoracic Surgery, Department of Surgery, 600 N. Wolfe St, Blalock 240, The Johns Hopkins Hospital, Baltimore, MD 21287 (E-mail: rmeguid1@jhmi.edu).

J Thorac Cardiovasc Surg 2009;137:23–9
0022-5223/\$36.00

Copyright © 2009 by The American Association for Thoracic Surgery
doi:10.1016/j.jtcvs.2008.09.040

Abbreviations and Acronyms

AUC = area under the curve

ICD-9 = International Classification of Diseases,
Ninth Revision

NIS = Nationwide Inpatient Sample

procedures, inpatient mortality, and length of stay.⁹ This study was approved by the Johns Hopkins Institutional Review Board, who exempted the need for patient consent.

Patient Population

Initial inclusion criteria for this study were patients from the NIS database older than 17 years of age admitted with the diagnosis of esophageal cancer as identified by the International Classification of Diseases, Ninth Revision (ICD-9) diagnosis codes (150.X).¹⁰ Inclusion criteria was further limited to patients who underwent esophageal resection as identified by ICD-9 Clinical Modification procedure codes of 42.4 and 42.40 (esophagectomy NOS), 42.41 (partial esophagectomy), 42.42 (total esophagectomy), and 43.99 (esophagogastrectomy).²

Statistical Analysis

Multivariable analysis was performed with in-hospital death as the outcome of record from the discharge summaries. Independent variables included annual hospital resection volume, teaching status of the hospital where the procedure was performed, the year the procedure was performed, patient age, gender, race, and comorbidities as measured by the Charlson Index. The NIS dataset defines teaching hospital status as hospitals that have any American Medical Association-approved residency program, belong to the Council of Teaching Hospitals, or have a ratio of no more than 4:1 beds to full-time equivalent interns and residents.¹¹

Patient comorbidities were standardized via calculation of the Deyo modification of the Charlson Index^{12,13} per the methods of Romano and colleagues.¹⁴ A standardized calculation of patient health, the Charlson Index is determined by weighted scoring of comorbidities, including cardiac, vascular, pulmonary, neurologic, endocrine, renal, hepatic, gastrointestinal, and immune diseases, as well as any documented history of cancer.

Individual annual hospital procedure volume was determined by calculating the number of esophageal resections performed using NIS-assigned unique hospital identification numbers. The annual hospital mortality rate for esophageal resections was calculated using the NIS annual hospital resection volume for esophageal resections.

Esophageal resection volume was included as a dichotomous variable to identify the volume cutoff that best models outcome. A series of sequential multiple logistic regression models with a dependent variable of in-hospital death; a set of common independent variables including patient age, gender, race, and Charlson Index of comorbidities, procedure year, and hospital teaching status; and a sequentially changing independent variable of dichotomized annual hospital resection volume were tested. This sequentially changing variable of annual hospital resection volume was dichotomized at 2 continuously up to 34, accounting for all of the esophageal resections in the NIS database in the time period studied. The resection volumes within this range are nearly continuous.

Each volume threshold dichotomizes the data and creates 2 categories for comparison: hospitals with an annual resection volume less than that cutoff and hospitals with an annual resection volume greater than or equal to that cutoff. Each volume threshold is then taken forward in the multivariable regression analysis as the independent variable.

Statistical analysis was performed using the software package STATA 10.0 (StataCorp LP, College Station, Tex). Bivariate analysis of categorical data was performed using the chi-square test. Analysis of continuous data

was performed using the Student *t* test. Multivariable analysis was performed using linear and logistic regression models. The goodness of fit, a measurement of the amount of variability in the data explained by the model, was tested for each model by calculation of McFadden's pseudo r^2 and the area under the curve (AUC). McFadden's pseudo r^2 is one such measure of goodness of fit and has been re-scaled from 0% to 100% for ease of interpretation and comparison. It represents the percent of variance in a data pattern that is explained by the set of variables in a particular model. For instance, a model explaining 7% of the variation in the data would have a pseudo r^2 of 0.07. Results are primarily reported as pseudo r^2 .¹⁵⁻¹⁷ AUC is also reported and improves as the value approaches 1.

RESULTS**Patient Population**

Analysis of the NIS dataset identified 53,168 patients with the diagnosis of esophageal cancer, of whom 4080 (7.7%) underwent esophageal resection, as defined by the previously listed ICD-9 Clinical Modification codes. Of these patients, 79.6% were male, and the median age was 64 years. These esophagectomies were performed at 1506 hospitals. The median annual hospital resection volume was 4, with the range from 1 to 34 (interquartile range 2–9). Of the patients studied, 83.9% were white, 8.8% were black, and the remainder were of unreported race. A total of 2883 patients (70.7%) underwent resection at teaching hospitals. The median Charlson Comorbidity Index for the 4080 patients studied was 3, with an interquartile range of 2 to 8 and a range of 2 to 14 (of a possible range from 0 to 33). Between 444 and 552 patients underwent esophageal resection per year. There were 387 in-hospital deaths for this patient group, resulting in an overall in-hospital mortality rate of 9.49%. See Table 1 for demographics.

Hospital Volume–Mortality Relationship

The unadjusted annual in-hospital mortality rate was calculated for each hospital. This ranged between 0% and 100%, with a median value of 0 and a mean value of 11.5% (Figure 1).

In-Hospital Mortality

A series of multiple logistic regression models were tested with a dependent variable of in-hospital death and common independent variables, including patient age, gender, race, and Charlson Index of comorbidities, procedure year, and hospital teaching status. In each model a sequentially changing variable of annual hospital resection volume threshold was inserted, dichotomizing volume into “less than” versus “greater than or equal to” that volume threshold. The mortality of patients at “high-volume” and “low-volume” hospitals defined at each threshold level and the various representations of goodness of fit (McFadden's pseudo r^2 and AUC of that particular multiple logistic regression model) are presented in Table 2.

The values represent the average mortality rate for all hospitals with esophagectomy volumes less than the volume

TABLE 1. Demographics of the study population

Characteristic	n	%
Esophageal cancer cases	53,168	
Esophageal resections	4080	7.7%
Procedures by ICD-9-CM codes:		
42.40 (esophagectomy NOS)	134	3.3%
42.41 (partial esophagectomy)	1633	40.0%
42.42 (total esophagectomy)	958	23.5%
43.99 (esophagogastrectomy)	1393	34.1%
Age median (interquartile range)	64 y	56–71 y
Male gender	3246	79.6%
Known race:	2944	
White	2471	83.9%
Black	258	8.8%
Other	215	7.3%
Median Charlson Index score (interquartile range)	3	2–8
Surgery at teaching hospital	2883	70.7%
Median annual hospital resection volume (interquartile range)	4	2–9
Year of surgery:		
1998	506	12.4%
1999	549	13.5%
2000	444	10.9%
2001	524	12.8%
2002	531	13.0%
2003	521	12.8%
2004	453	11.1%
2005	552	13.5%
In-hospital deaths	387	9.5%

ICD-9-CM, International Classification of Diseases, Ninth Revision, Clinical Modification.

threshold, as well as the average mortality rate for all hospitals with esophagectomy volumes greater than or equal to the volume threshold. An example of interpretation of the values for a volume threshold of 15 would signify that hospitals that perform greater than or equal to 15 resections per year have an average postoperative mortality rate of 5.30% compared with an average postoperative mortality rate of 10.16% at hospitals that perform less than 15 esophageal resections per year ($P < .001$). Comparison of mortality rates of hospitals with annual esophagectomy volumes for esophageal cancer above and below each volume threshold reveals that they are significantly different for all volume thresholds compared ($P < .05$), except when volume is greater than or equal to 29 esophagectomies/year ($P = .08$) and greater than or equal to 34 esophagectomies/year ($P = .19$). These results are depicted graphically in Figure 2.

The goodness of fit of different models with different volume thresholds, as measured by McFadden's pseudo r^2 , ranged from a baseline of 3.23%, where the volume variable is not included in the model, to 3.87%. The models with the best fit to data were those that defined high-volume threshold at 15 and 16, with McFadden's pseudo r^2 of 0.0387 or 3.87%. It should be noted that when the volume threshold

FIGURE 1. Graph of in-hospital mortality rates for individual hospitals by annual hospital esophageal resection volume. These data represent the annual hospital mortality rates of in-hospital death at 1506 different hospitals. The data points have been staggered to illustrate the trend because of the high frequency of overlapping values.

was set at 13, as suggested by the Leapfrog Group, the resulting model had a McFadden's pseudo r^2 of 3.80%. A graphic depiction of the changes in goodness of fit as defined by McFadden's pseudo r^2 of different volume thresholds is shown in Figure 3. The goodness of fit for the dichotomous model improved as the volume threshold was increased from 1 to 15, peaking at a threshold of 15 resections per year. Subsequently, the goodness of fit of the model deteriorated as the volume threshold was further increased.

The results from calculation of the AUC as a measure of goodness of fit are reported in Table 2. The range in the calculated values of AUC is similar to that of the McFadden's pseudo r^2 ; therefore, our discussion focuses on the results of the McFadden's pseudo r^2 .

CONCLUSIONS

Much attention has been focused on defining centers of excellence. To date, the role of hospital volume has been emphasized. By using the NIS dataset, we examined the relationship between surgical volume and mortality. We found this well-established inverse relationship persists in the NIS dataset. This is demonstrated in Figure 1, where a clear trend toward diminishing unadjusted annual in-hospital mortality is present.

Given the well-established inverse relationship between esophageal resection volume and in-hospital mortality, we sought to use statistical modeling to define a single value cutoff at which there is significantly reduced mortality. This would allow us to better determine hospital operative volumes required for improved outcomes for esophageal resection for cancer.

On analysis of dichotomous volume cutoff modeling, we found a statistically significant difference between mortality rates at hospitals with esophagectomy volumes above

TABLE 2. Mortality rates and goodness of fit to data of multiple high-volume threshold models†

Volume threshold	Mortality rate at centers with volume less than volume threshold (%)	Mortality rate at centers with volume equal to or greater than volume threshold (%)	P value	McFadden's pseudo r ² (shown as %)	AUC
Baseline	-	-	-	3.23	0.6354
1	-	9.94%	-	3.31	0.6364
2	13.31%	8.63%	<.001	3.43	0.6431
3	12.38%	8.03%	<.001	3.46	0.6430
4	11.54%	7.71%	<.001	3.35	0.6392
5	11.31%	7.36%	<.001	3.44	0.6414
6	11.04%	7.09%	<.001	3.43	0.6412
7	10.96%	6.54%	<.001	3.52	0.6425
8	10.86%	6.35%	<.001	3.56	0.6418
9	10.68%	5.98%	<.001	3.57	0.6435
10	10.58%	5.37%	<.001	3.73	0.6440
11	10.50%	5.22%	<.001	3.73	0.6440
12	10.35%	5.31%	<.001	3.79	0.6449
13*	10.26%	5.39%	<.001	3.80	0.6450
14	10.22%	5.50%	<.001	3.76	0.6444
15	10.16%	5.30%	<.001	3.87	0.6465
16	10.14%	5.26%	<.001	3.87	0.6461
17	10.08%	5.39%	.001	3.76	0.6441
18	10.01%	5.57%	.002	3.67	0.6425
19	9.97%	5.57%	.003	3.69	0.6428
20	9.95%	5.10%	.002	3.64	0.6417
21	9.92%	5.14%	.003	3.64	0.6417
23	9.92%	4.87%	.002	3.69	0.6427
24	9.91%	4.60%	.002	3.71	0.6437
25	9.83%	4.33%	.004	3.70	0.6449
26	9.75%	4.43%	.012	3.60	0.6430
27	9.71%	4.52%	.021	3.50	0.6408
28	9.67%	4.67%	.040	3.50	0.6408
29	9.63%	4.92%	.080	3.40	0.6387
30	9.61%	1.56%	.029	3.51	0.6390
34	9.54%	2.94%	.191	3.31	0.6359

AUC, Area under the curve. The relationship of different high-volume thresholds to mortality rates and the goodness of fit (McFadden's pseudo r² and the AUC) of each model. An example of interpretation of the values for a volume threshold of 15 would signify that, on average, hospitals that perform ≥ 15 esophageal resections per year have a mortality rate of 5.30%, compared with 10.16% at hospitals that perform < 15 resections per year. The *P* value for the chi-square statistic of these mortality rates is $< .001$, and this model has a pseudo r² of 3.87% and an AUC of 0.6465. *The definition of "high volume" as suggested by the Leapfrog Group.⁵ †Multivariable logistic regression model adjusted for patient age, gender, race, and Charlson Index of comorbidities, procedure year and hospital teaching status, and a sequentially changing independent variable of dichotomized annual hospital resection volume.

the volume threshold in comparison with mortality rates of hospitals with esophagectomy volumes below the volume threshold, irrespective of annual hospital resection volume cutoff (Figure 2). For example, defining high volume at 13 or more, as suggested by the Leapfrog Group,⁵ the resulting high-volume hospitals have a mortality rate of 5.39% in comparison with 10.26% at low-volume hospitals ($P < .001$). However, even defining the high-volume threshold at a volume of 2 resections per year produces significant differences in mortality rates between hospitals with esophagectomy volumes above and below that threshold. Our study confirms previous findings by Christian and colleagues,¹⁸ who also showed that the Leapfrog standards may not have been optimal for other surgical procedures; for example, they empirically found different thresholds for coronary artery bypass graft, abdominal aortic aneu-

rysm, and esophagectomies compared with the Leapfrog standards; moreover, in contrast with Leapfrog, they found no good empiric threshold for carotid endarterectomies.

This finding reveals the true conundrum of volume modeling: No matter what the volume cutoff is set at, the mortality rates above and below it are almost always significantly different. Therefore, to determine the best model for high-volume centers, we examined goodness of fit of the model to the data instead of differences in mortality.

When multiple logistic regression of in-hospital death after esophageal resection includes the variables of patient age, gender, race, and Charlson Index of comorbidities and calendar year, but not resection volume, the resulting model explains 3.23% of the variance in the data. Adding hospital volume as a dichotomous variable, ranging from 2 to 34 resections per year, improves the explanatory power

FIGURE 2. Graph showing the different mortality rates above and below each volume threshold. Squares indicate the mortality rates of hospitals with annual esophagectomy volumes for esophageal cancer less than the volume threshold. Triangles indicate the mortality rates of hospitals with annual esophagectomy volumes for esophageal cancer greater or equal to the volume threshold. The difference between each pair of mortality rates at a given annual hospital volume is statistically significant for all volume thresholds except more than 29 and 34. Associated data are listed in Table 2.

of the model, with pseudo r^2 ranging between 3.35% and 3.87%. By using these criteria, the best model is one that defines a “high-volume” cutoff as 15 or more esophageal resections per hospital per year, because this has the highest McFadden’s pseudo r^2 value and accounts for the most variability in the data. It is interesting to note that the inclusion of volume into the multivariable model only accounts for a maximum of 0.64% of the variability in the data. Therefore, varying the volume threshold did not substantially change the explanatory power of the different dichotomous volume models for defining high-volume centers for esophageal resection. This is noteworthy given the attention that resection volume for esophageal surgery, among other procedures, has been afforded in the literature.

Although there is an overall trend of increased operative volume associated with decreased postoperative mortality, a curious finding is present in Table 2: The mortality rate at centers with annual resection volumes equal to or greater than the volume threshold tested do not necessarily have continuously diminishing values. The mortality rates given in Table 2 are calculated by averaging the mortality rates of every hospital that performs esophagectomies above or below the volume threshold. As can be seen, increased volume does not strictly correlate with decreased postoperative, in-hospital mortality. Therefore, factors other than annual hospital volume must certainly contribute to mortality rate.

The NIS database was chosen over other available databases because of the extensive nature of its records and the ability to provide a large sample size with which to compare outcomes across the United States. As in analyses of all administrative databases, the current analysis has several limitations. They include the retrospective database design

FIGURE 3. Graph of goodness of fit versus annual hospital resection volume. Each point represents one resection volume. McFadden’s pseudo r^2 is shown as percent. Dashed line represents “baseline” McFadden’s pseudo r^2 from the model without resection volume (3.23%).

and the associated constraints at the level of the data used for analysis, the inability to account for surgeon experience, the difficulty in examining other postoperative outcomes such as cause of death, and the inability to measure 30-day mortality, as opposed to in-hospital death. In examining the NIS database, we are unable to check the accuracy of the diagnostic and procedure coding. Although the validity of the coding may be verified, the appropriateness of the coding used for diagnosis and procedures may not. However, we assume that this type of error would be equally distributed across all groups of interest. The overall in-hospital mortality rate of 9.49% is consistent with reported mortality rates of other large series using 30-day mortality,¹⁹ adding validity to the data reported in the NIS database, and our use of in-hospital mortality as an outcome. In addition, it has been argued that for complex operations, in-hospital mortality may be a better measure of postoperative mortality than 30-day mortality because of improved capabilities of intensive care management to rescue critically ill patients.¹¹

Other outcomes, such as complications associated with surgery or perioperative care and postdischarge outcomes, including deaths occurring outside of the surgical hospitalization, are not ascertainable from this database. Complications occurring after surgery cannot be differentiated from comorbidities existing preoperatively. This prevents us from examining and comparing postoperative complications. In addition, because these patients have undergone esophagectomy for cancer, it would be meaningful to measure disease-free and overall survival. When calculating the Charlson Index we assume that preexisting conditions and those same conditions arising after surgery have the same impact on patient outcomes. Proxies of non-death hospital outcome, including the need for postoperative procedural intervention and length of hospital stay, have been used by others studying different databases.^{20,21}

There has been much recent postulation as to factors that influence postoperative outcomes at the hospital level. These focus on processes of care, which may be associated with improved outcome after surgery. Billingsley and colleagues²¹ have correlated improved outcomes after surgery for colon cancer with the presence of solid organ transplantation teams, as a proxy for patient care indices associated with improved postoperative outcomes. Other processes of care studied and correlated to improved outcomes include dedicated surgical intensive care units managed by dedicated intensive care specialists,^{3,22} patient safety initiatives,²³ and the use of multidisciplinary teams and standardized clinical care pathways at high-volume centers, for example.²⁴ We believe it is likely that these hospital-level processes of care are more readily available at high-volume centers, and as such, high-volume status may serve as a proxy for them in large administrative databases such as the NIS.

We show that, although there is a trend toward an inverse relationship between volume and mortality, volume is not sufficient for defining centers of excellence. Volume seems to function as an imperfect surrogate for other variables, which may better define centers of excellence, such as quality of dedicated intensive care, postoperative monitoring, clinical care pathways, and other processes of care.^{20,21,25,26} Additional work is needed to identify those variables associated with improved outcome after esophageal resection.

In addition, using a comparison of mortality rates and goodness of fit of different volume thresholds, we were unable to identify a clear, optimal volume threshold for improved outcomes after esophagectomy for esophageal cancer. We conclude that the use of volume thresholds alone for determining centers of excellence does not appropriately represent the variance in the data or necessarily guide appropriate decision making and should be avoided.

The authors thank Nita Ahuja, MD, for critical review of the methods used in this study.

References

- Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. *N Engl J Med*. 1979;301:1364-9.
- Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. *JAMA*. 1998;280:1747-51.
- Dimick JB, Cattaneo SM, Lipsett PA, Pronovost PJ, Heitmiller RF. Hospital volume is related to clinical and economic outcomes of esophageal resection in Maryland. *Ann Thorac Surg*. 2001;72:334-41.
- Kuo EY, Chang Y, Wright CD. Impact of hospital volume on clinical and economic outcomes for esophagectomy. *Ann Thorac Surg*. 2001;72:1118-24.
- Leapfrog Group "Evidence-based hospital referral" for esophageal cancer. 2003. Available at: leapfroggroup.org. Accessed November 21, 2006.
- Chang AC, Birkmeyer JD. The volume-performance relationship in esophagectomy. *Thorac Surg Clin*. 2006;16:87-94.
- Milstein A, Galvin RS, Delbanco SF, Salber P, Buck CR Jr. Improving the safety of health care: the leapfrog initiative. *Eff Clin Pract*. 2000;3:313-6.
- Healthcare Cost and Utilization Project (HCUP): Nationwide Inpatient Sample. Rockville, MD: Agency for Healthcare Research and Quality; 2005.
- Ahuja N, Chang D, Gearhart SL. Disparities in colon cancer presentation and in-hospital mortality in Maryland: a ten-year review. *Ann Surg Oncol*. 2007;14:411-6.
- International Classification of Diseases, Ninth Revision, Ninth Edition. Washington, DC: Public Health Service, US Department of Health and Human Services; 1988.
- Casson AG, van Lanschot JJ. Improving outcomes after esophagectomy: the impact of operative volume. *J Surg Oncol*. 2005;92:262-6.
- Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40:373-83.
- Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol*. 1992;45:613-9.
- Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. *J Clin Epidemiol*. 1993;46:1075-9; discussion 1081-90.
- Allison PD. Comparing logit and probit coefficients across groups. *Sociol Methods Res*. 1999;28:186-208.
- Long JS, Freese J. *Regression Models for Categorical Dependent Variables Using Stata*. 2nd ed. College Station, TX: Stata Press; 2006.
- Long JS. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage Publications; 1997.
- Christian CK, Gustafson ML, Betensky RA, Daley J, Zinner MJ. The Leapfrog volume criteria may fall short in identifying high-quality surgical centers. *Ann Surg*. 2003;238:447-57.
- Bailey SH, Bull DA, Harpole DH, Rentz JJ, Neumayer LA, Pappas TN, et al. Outcomes after esophagectomy: a ten-year prospective cohort. *Ann Thorac Surg*. 2003;75:217-22.
- Birkmeyer JD, Dimick JB. Potential benefits of the new Leapfrog standards: effect of process and outcomes measures. *Surgery*. 2004;135:569-75.
- Billingsley KG, Morris AM, Dominitz JA, Matthews B, Dobie S, Barlow W, et al. Surgeon and hospital characteristics as predictors of major adverse outcomes following colon cancer surgery: understanding the volume-outcome relationship. *Arch Surg*. 2007;142:23-32.
- Angus DC, Shorr AF, White A, Dremsizov TT, Schmitz RJ, Kelley MA. Critical care delivery in the United States: distribution of services and compliance with Leapfrog recommendations. *Crit Care Med*. 2006;34:1016-24.
- Makary MA, Sexton JB, Freischlag JA, Millman EA, Pryor D, Holzmüller C, et al. Patient safety in surgery. *Ann Surg*. 2006;243:628-35.
- Zehr KJ, Dawson PB, Yang SC, Heitmiller RF. Standardized clinical care pathways for major thoracic cases reduce hospital costs. *Ann Thorac Surg*. 1998;66:914-9.
- Birkmeyer JD, Sun Y, Goldfaden A, Birkmeyer NJ, Stukel TA. Volume and process of care in high-risk cancer surgery. *Cancer*. 2006;106:2476-81.
- Hannan EL, Popp AJ, Feustel P, Halm E, Bernardini G, Waldman J, et al. Association of surgical specialty and processes of care with patient outcomes for carotid endarterectomy. *Stroke*. 2001;32:2890-7.

Discussion

Dr A. Pennathur (Pittsburgh, Penn). Dr Meguid, that was an excellent presentation and I congratulate you on your efforts to define a cutoff volume. Dr Meguid and colleagues from the Johns Hopkins Hospital have analyzed the outcomes of more than 3000 patients from more than 1000 hospitals derived from the NIS. They analyzed the variable of esophagectomy volume with the primary outcome being in-hospital mortality. Their main objective was to evaluate how valid hospital volume cutoffs are for defining centers of excellence. The number of esophagectomies performed in these hospitals ranged from a minimum of 1 to a maximum of 29. The median was somewhat lower, just 4 esophagectomies per year. The overall mortality rate was 9.9%. After using esophagectomy volume as a dichotomous variable initially, they went on to perform a goodness-of-fit model, concluding that they were unable to establish a cutoff value. They then concluded that volume cutoffs cannot be used to determine centers of excellence and further work is needed to investigate this.

Several large single-institution series have been published with remarkably low mortality rates, including your institution; ours,

where in a series of 222 consecutive esophagectomies had a mortality rate of 1.4%; and excellent results from other larger series, such as those from Drs Orringer, Altorki, and Swanson, to name a few.

However, this relationship between esophagectomy volume and outcome is complex, with several factors playing a role. These factors include the surgeon volume, specialty training of the surgeon, comorbidities in the patient, and provision for critical care services. Adding to these factors is the case mix seen at a particular hospital, which may contribute. For example, the referral pattern of a private hospital may be much different than that of an inner-city hospital.

So my first question is, given the complexities of the volume–outcome relationship, do you think it is possible that you can reduce this to a single number across all hospitals in the United States without taking into consideration other important factors, such as surgeon volume, expertise, and patient population?

Dr Meguid. The impetus for our study was the curiosity to see how the seemingly arbitrary cutoff of 13 fared against other volume cutoffs. We expected to see a dramatic difference in mortality rates, for instance, but were, quite frankly, shocked by the apparent lack of the difference between choosing 13 and any other volume cutoff.

As you point out, the relationship between individual and hospital operative volume and the processes of care and outcome is very complex. Unfortunately, it's difficult to study the effect of different processes of care because of lack of information available in these multi-institutional databases. Using a patient-focused database, such as the Society of Thoracic Surgeons' database, would be ideal; however, at the present time data are lacking for such analysis, and unfortunately, use of all of our own single-institution databases lends to bias.

Dr Pennathur. The second question is, did you first conduct an analysis of all the data in this particular cohort of patients to establish a functional relationship between mortality and volume in this cohort before dichotomizing these patients?

Dr Meguid. Yes, sir. Before dichotomizing the data at the different cutoff points, we did examine the unadjusted relationship between volume and in-hospital death and found this to be an inversely linear relationship. When we adjusted for age, patient gender, patient race, and patient comorbidities, we also saw that that persisted.

Dr Pennathur. The next thing is, using mortality as an outcome variable, how do you do a risk-adjusted mortality rate? For example, from your article, when the volume was greater than 10, the mortality rate was 5.3%. However, when the volume was greater than or equal to 29, the mortality was actually higher; it was 8.6%. Is this because sicker patients are going to high-volume hospitals? Perhaps a more useful reporting might be a risk-adjusted mortality rate.

Dr Meguid. Yes, I fully agree with you. In dichotomizing a continuous variable, one is combining all of the values below and above that cutoff into 2 values. As a result, the lowest mortality rate was observed at a volume cutoff of 10, but when the volume cutoff was raised, a lower mortality rate was not observed. Again, that's one of the prime complications of using a dichotomous model for continuous variables.

Dr Pennathur. Along the same lines, are you going to attempt to take other variables into consideration, such as nutritional status of patients, which has been shown to be important; socioeconomic status; specialty training of the surgeon, which has been shown to be important; and elective versus emergent procedures, all of which are going to have an impact on mortality?

Dr Meguid. Yes, I fully agree with you regarding the importance of these factors. Unfortunately, because of the administrative nature of the NIS database, one is unable to account for many germane factors, such as cancer staging, preoperative nutritional status, and neoadjuvant or postoperative chemo- and radiotherapy. One can control for gender, some patient demographics, which aren't necessarily specific to operative mortality, the Charlson Comorbidity Index, and some hospital demographics. This is, again, a big limitation with these administrative datasets.

Dr M. Jaklitsch (*Boston, Massachusetts*). I'm just deeply concerned that your highest output hospital was 29 cases per year, and we all know of excellent academic centers of excellence that do more than 29 per year. So either they are not in the database, the most sterling outcomes are not in this database, or they are in this database but with incomplete data. I don't know how you can draw conclusions from this if either 1 of those 2 cases is true.

Dr Meguid. Dr Jaklitsch, that's an excellent point. In fact, when we looked at a similar example using pancreatic resections, we found a similarly small range, and that motivated us at Johns Hopkins to look into why we don't find institutions with 100 resections per year. Subsequently, in this nationally representative sample, we find that a lot of the hospitals are lower-volume hospitals and we don't see a lot of the larger academic centers included every year in this database. So that is, again, a limitation specific to the NIS database.

Dr Jaklitsch. Can I ask, how many cases a year do you do at Hopkins?

Dr Meguid. I'm not sure, I believe it's in the 50s. Dr Yang?

Dr Yang. It's about 75.

Dr Jaklitsch. So Hopkins' data, for instance, is not in this?

Dr Meguid. It's not in the NIS, no, sir.

Dr T. Grodzki (*Szczecin, Poland*). Did you make a differentiation between hospital volume and surgeon volume? Because it's not the same. And did you analyze the mortality reasons? Were they due to technical failure or the imperfections in postoperative care?

Dr Meguid. Those are very good points. We chose to use hospital volume, although we could have examined physician volume. We chose to use hospital volume because that is what has been used in other models, including the Leapfrog Group, and we wanted to analyze these cutoffs in particular. Unfortunately, we can't tease out what postoperative complications occur in these patients because of the limitations of this dataset.

Dr D. Wood (*Seattle, Wash*). I have the same concern that Dr Jaklitsch expressed, but also a concern about how you have represented the conclusion. It would seem that volume is an important surrogate for mortality outcomes, yet you have found that there is not a good cutoff for volume. That does not mean that volume isn't important, which it sounds like in the conclusion. Rather, volume is very important; we just cannot create a cutoff to define an "adequate" volume. So I think that it is very important to refine the message, because policymakers, like Leapfrog, need a clear message that volume is important in terms of quality of outcomes, unless you think that this research disputes that premise.

Dr Meguid. Dr Wood, you have made an excellent point. I don't want to misrepresent our findings. Our findings are that a specific cutoff is an inappropriate way to determine centers of excellence. However, in this example, increased volume is correlated with decreased complications and mortality, and that should not be overlooked.