The Relationship Between Hospital Surgical Case Volumes and Mortality Rates in Pediatric Cardiac Surgery: A National Sample, 1988–2005

Karl F. Welke, MD, MS, Brian S. Diggs, PhD, Tara Karamlou, MD, MS, and Ross M. Ungerleider, MD, MBA

Division of Cardiothoracic Surgery and Department of Surgery, Oregon Health and Science University, Portland, Oregon

Background. **Overall surgical volumes and raw mortality rates are frequently used to compare pediatric cardiac surgical programs, but unadjusted comparisons are potentially unreliable. We sought to quantify the relationship between hospital volume and pediatric cardiac surgical mortality.**

Methods. **Pediatric cardiac operations assigned to Risk Adjustment for Congenital Heart Surgery, version 1 (RACHS-1) categories were retrospectively identified by *International Classification of Diseases, 9th Revision, Clinical Modification* (ICD-9-CM) coding from the Nationwide Inpatient Sample, 1988–2005. Hospitals were grouped by yearly pediatric cardiac surgical volume (very small, ≤ 20; small, 21 to 100; medium, 101 to 200; large, > 200). Mortality rates were adjusted for surgical volume, case mix (RACHS-1 categories), patient age, and year of operation by logistic regression.**

Results. **We identified 55,164 operations from 307 hospitals; 188 (61%) performed 20 or fewer cases per year. The unadjusted mortality rate at very small hospitals was no different than at large hospitals (odds ratio, 1.0, 95% confidence interval [CI] 0.7 to 1.4). After adjustment for RACHS-1 category and age, large hospitals performed significantly better than all other volume groups. As a discriminator of mortality, volume performed significantly worse than a model with RACHS-1 category and age (receiver operating characteristic [ROC] curve area, 0.60 vs 0.81).**

Conclusions. **As a discriminator of mortality, volume alone was only marginally better than a coin flip (ROC curve area of 0.50). However, large-volume hospitals performed more complex operations and achieved superior results; therefore, the use of overall, unadjusted mortality rates to evaluate institution quality is misleading. Hospital comparisons and pay-for-performance initiatives must be based on robust risk-adjusted comparisons.**

(Ann Thorac Surg 2008;86:889–96)

The focus of public reporting and transparency has been hospital surgical volumes and in-hospital mortality rates. For adult cardiac surgery, the spotlight has been on coronary artery bypass grafting (CABG), an operation that is performed at relatively high volumes and for which risk-adjustment methodology has been well studied and validated. As transparency initiatives have broadened, there has been increased reporting of pediatric cardiac surgical results. In contrast to the relative homogeneity of CABG, the pediatric cardiac surgical case-mix consists of a broad variety of operations, which at any one institution are each performed at relatively low volume. In addition, risk-adjustment methodology is less mature. Despite these important differences, overall surgical volumes and raw mortality rates are frequently used to compare the performance of pediatric cardiac surgical programs. When risk adjustment is performed, it frequently encompasses patient risk factors, but not surgical case-mix. However, the broad variety of pediatric cases performed makes such unadjusted comparisons potentially unreliable.

Previous investigations of the volume–mortality relationship in pediatric cardiac surgery have been mixed in their conclusions. Although the earlier of these investigations demonstrated an inverse relationship between volume and in-hospital mortality, the most recent study suggested that this relationship might no longer exist [1–5]. These multiple studies have relied on a relatively limited group of data. All have used either administrative data from California and Massachusetts or clinical data from New York, with the data from California and New York having been studied for multiple time periods.

The purpose of our study was to determine the relationship between hospital surgical volume and mortality after pediatric cardiac surgery. For the investigation we used a national administrative database and methodology that accounts for both patient-level risk factors and surgical case-mix.

Material and Methods

This study was designed as a retrospective cohort analysis and was approved by the Oregon Health and Science University Institutional Review Board. Owing to the nature

Accepted for publication April 21, 2008.

Presented at the Forty-fourth Annual Meeting of The Society of Thoracic Surgeons, Fort Lauderdale, FL, Jan 28–30, 2008.

Address correspondence to Dr Welke, Division of Cardiothoracic Surgery L353, Oregon Health and Science University, 3181 SW Sam Jackson Park Rd, Portland, OR 97239-3098; e-mail: welkek@ohsu.edu.

Published by Elsevier Inc

0003-4975/08/$34.00
doi:10.1016/j.athoracsur.2008.04.077

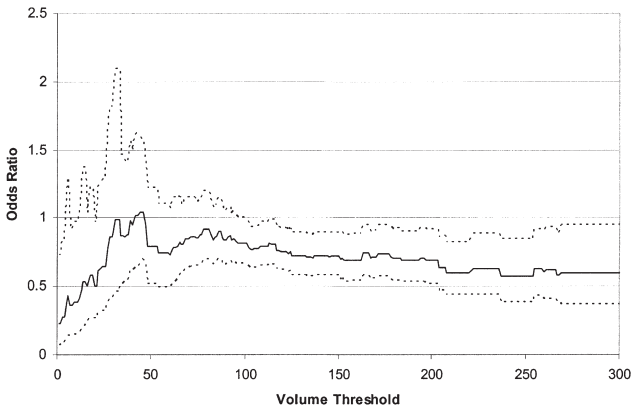

Fig 1. Identification of a hospital threshold volume associated with lower mortality. Presented are the odds ratio (solid line) and 95% confidence intervals (dashed lines) of high-volume hospitals vs low-volume hospitals from a logistic regression model of mortality. The model is also adjusted for year, Risk Adjustment in Congenital Heart Surgery categories, and age. The threshold volume dividing high- and low-volume hospitals is the value along the horizontal axis. For thresholds exceeding 100 cases per year, high-volume hospitals had a significantly lower mortality rate than low-volume hospitals.

of the study, patient consent was not required. We obtained data from the Nationwide Inpatient Sample (NIS) [6], which is the largest all-payer inpatient care database in the United States. The database is a stratified, cross-sectional sample that includes approximately 20% of all community (nonfederal) hospital discharges in the United States. NIS data are available from 1988 to 2005, over which time the number of states in the NIS has grown from 8 to 37. In 2005 the database contained discharge data on approximately 8 million hospital stays at 1054 hospitals in 37 states. The sampling frame for the 2005 NIS is a sample of hospitals that comprises approximately 90% of all hospital discharges in the United States. To ensure the representative nature of the database, the NIS is stratified by geographic region, urban vs rural location, teaching status, hospital ownership, and number of hospital beds. The NIS is managed under the Healthcare Cost and Utilization Project of the Agency for Healthcare Research and Quality. For this study we combined data from the 1988 through 2005 NIS databases.

Congenital cardiac surgical procedures performed on patients aged younger than 18 years were identified by *International Classification of Diseases, 9th Revision, Clinical Modification* (ICD-9-CM) diagnosis and procedure codes. For a patient to be included in this study, the procedure code had to match to a plausible diagnostic code. Operations were categorized by the Risk Adjustment for Congenital Heart Surgery, version 1 (RACHS-1) method [7]. This risk-stratification system groups the varied congenital cardiac surgical case-mix into six categories by similar expected short-term mortality rates. Category 1 has the lowest risk of death and category 6 the highest. The methodology has been validated and is included in the Society of Thoracic Surgeons Congenital Heart Surgery Database reports.

Mortality was defined as in-hospital mortality as indicated by the discharge disposition. Hospital volume was defined as the number of RACHS-1–categorized operations performed in a year. If a hospital appeared in the NIS in 2 or more years, each entry was treated separately. Hospital volume was not averaged for the 18-year period because all hospitals were not sampled in all years and it is likely that personnel and systems changes occurred during this time that affected volume and quality of care.

Volume was first evaluated as a continuous variable. Then, volume groups were created using the following criteria: (1) natural cut points in the data, (2) previously studied volume thresholds, and (3) maintenance of a sufficient number of hospitals in each volume group to minimize the impact of any individual hospital. The cut point of 100 cases per year was derived by dichotomizing volume and comparing odds ratios (ORs) for risk-adjusted mortality of hospitals with volumes above the cut point with those with volumes below the cut point. All volume thresholds from 1 to 300 cases per year were investigated. We found a threshold annual volume of approximately 100 cases. At this threshold and those above it, the mortality in the higher-volume cohort of hospitals was significantly lower than that in the lower-volume cohort (Fig 1). We chose the cut point of 200 cases per year from the distribution of hospital mortality rates (Fig 2). Above 200 cases per year, all hospital mortality rates were within or below the 95% confidence interval (CI) for the cohort.

Several approaches were used to define the relationships between hospital volume and mortality. First, unadjusted mortality rates across volume groups were compared using the χ^2 statistic for linear trend. Second, the

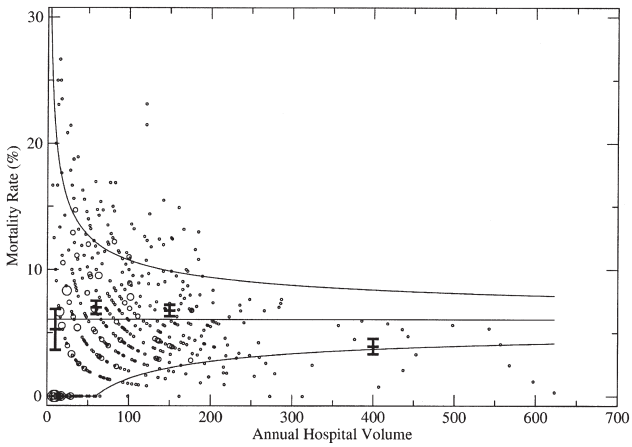

Fig 2. Individual hospital mortality rates are presented according to annual pediatric cardiac surgical volume. Not shown are 291 hospitals with fewer than 5 cases per year. The size of the circle at a given hospital annual volume and mortality rate represents the number of hospitals with that mortality rate at that volume. The horizontal line is the overall raw mortality rate (6.07%). The lines above and below this line represent the confidence intervals (CIs) of a hypothetical single institution at that volume and the average mortality rate; these give an idea of what mortality rates would be statistically different at that volume. The four marks with error bars are the collective mortality rates and 95% CIs for the four volume groups (< 20 cases/y, 20 to 100 cases/y, 101 to 200 cases/y, and > 200 cases/y).

Table 1. Descriptive Characteristics of Nationwide Inpatient Sample Hospitals Performing Pediatric Cardiac Surgery (1988–2005)

Variable	Hospital Case Volume (cases/y)				
	All Hospitals	Very Small (≤ 20)	Small (21–100)	Medium (101–200)	Large (> 200)
Unique hospitals, No.	307	188	51	41	27
Hospital-year combinations, No.	828	364	250	164	50
Hospital location/teaching status, No (%)[a]					
Rural	27 (3.3)	25 (6.9)	1 (0.4)	1 (0.6)	
Urban nonteaching	211 (25.5)	132 (36.3)	43 (17.2)	28 (17.1)	8 (16.0)
Urban teaching	590 (71.3)	207 (56.9)	206 (82.4)	135 (82.3)	42 (84.0)
Control/ownership of hospital, No (%)[b,c]					
Government/private (collapsed category)	301 (36.4)	101 (27.7)	104 (41.6)	63 (38.4)	33 (66.0)
Government, nonfederal (public)	117 (14.1)	55 (15.1)	43 (17.2)	18 (11.0)	1 (2.0)
Private, not-for-profit (voluntary)	386 (46.6)	192 (52.7)	96 (38.4)	82 (50.0)	16 (32.0)
Private, investor-owned (proprietary)	20 (2.4)	12 (3.3)	7 (2.8)	1 (0.6)	
Private (collapsed category)	4 (0.5)	4 (1.1)			

[a] $p < 0.0001$ for linear trend across volume groups. [b] $p = 0.31$ for linear trend across volume groups. [c] States report hospital control/ownership with varying detail. When less detail is available, hospital control/ownership is reported in collapsed categories.

discrimination of volume alone as a predictor of mortality was assessed by the C statistic [8, 9] as determined from logistic regression analyses. The C statistic is numerically equivalent to the area under a receiver operating characteristic (ROC) curve. To maximize potential discrimination, volume categorizations were modeled using indicator variables for individual volume categories in lieu of a single ordinal variable.

Multivariable analyses examined the increase in model discrimination (ie, C statistic values) resulting from the addition of RACHS-1 category and patient age to the models containing hospital volume. Models accounted for patients clustered within hospitals [10] and included robust methods for calculating the 95% CIs for regression coefficients [11]. Two sets of models were developed, one for the years 1988 to 2005 and the other for the years 2001 to 2005. The first range of years was chosen to include all available data and achieve the largest possible sample size and most stable relationships. The second range of years was chosen to determine the volume–mortality relationship for the most contemporary time period. All models were adjusted for year of surgery. To check for the possible effect of miscoding on the results, all models were also run excluding hospitals that performed fewer than 5 cases in a year. When this was done, there were no significant changes in the results. Our analyses used sampling weights provided within the NIS database to derive national estimates. Analyses were performed using SAS 9.1 software (SAS Institute, Cary, NC).

Results

We identified 55,164 operations, representing a national estimate of 289,749 operations, in the NIS from 1988 to 2005 that could be placed in RACHS-1 categories. These operations took place at 307 hospitals, giving a total of 828 hospital-year combinations. Most hospitals were urban teaching hospitals, with the proportion being lowest in the group of hospitals that performed fewer than 20 operations per year (Table 1).

Patient characteristics and surgical case-mix varied significantly by hospital volume category (Table 2). The percentage of neonates undergoing operations increased with hospital case volume, with 12.2% of operations at very small hospitals being performed on neonates compared with 19.9% at large hospitals. At very small hospitals, 74.4% of operations were performed on patients older than 1 year. Case-mix varied significantly by hospital volume. The proportion of RACHS-1 category 1 cases decreased with increasing hospital volume, and the proportion of complex cases (RACHS-1 categories 4, 5, and 6) increased with volume.

The overall unadjusted mortality rate for the cohort declined steadily from 9.5% in the earliest 3-year period (1988 to 1990) to 4.0% in the most recent period (2003 to 2005). The overall unadjusted mortality rate at very small-volume hospitals was lower than that at small- and medium-volume hospitals (5.3% vs 7.0% and 6.8%) and slightly higher than that at large-volume hospitals (4.0%). When mortality rates were adjusted only for volume and year of operation, very small hospitals performed no differently than large volume hospitals (OR, 0.99; $p = 0.94$; Table 3). In contrast, small- and medium-volume hospitals underperformed large-volume hospitals. The C statistic for this model was low (0.60), indicating that volume alone was a poor predictor of mortality. The addition of RACHS-1 category and patient age to the model improved the discrimination of the model substantially (C statistic, 0.81). In this model, mortality at very small-volume hospitals was significantly higher than mortality at large-volume hospitals (OR, 1.88, $p = 0.0009$). Small- and medium-volume hospitals continued to underperform large-volume hospitals with slightly higher ORs than in the model containing only volume.

Table 2. Descriptive Characteristics of Patients Undergoing Pediatric Cardiac Operations at Nationwide Inpatient Sample Hospitals (1988–2005)[a]

	Hospital Case Volume (cases/y)				
Variable	All Hospitals	Very Small (≤ 20)	Small (21–100)	Medium (101–200)	Large (> 200)
Patients, No.	289,749	6606	75,418	132,618	75,107
Age group, No. (%)[b]					
Missing	16,649 (5.7)	67 (1.0)	3348 (4.4)	6385 (4.8)	6848 (9.1)
≤30 days	52,898 (18.3)	804 (12.2)	13,073 (17.3)	24,090 (18.2)	14,932 (19.9)
31 days–1 year	75,829 (26.2)	821 (12.4)	19,477 (25.8)	35,713 (26.9)	19,818 (26.4)
>1 year	144,373 (49.8)	4914 (74.4)	39,520 (52.4)	66,430 (50.1)	33,510 (44.6)
RACHS-1 category, No. (%)[b]					
1	48,065 (16.6)	2570 (38.9)	13,890 (18.4)	21,823 (16.5)	9783 (13.0)
2	110,751 (38.2)	1867 (28.3)	30,663 (40.7)	51,255 (38.6)	26,965 (35.9)
3	100,497 (34.7)	1897 (28.7)	24,652 (32.7)	46,366 (35.0)	27,582 (36.7)
4	24,040 (8.3)	267 (4.0)	5195 (6.9)	10,852 (8.2)	7725 (10.3)
5 and 6	6397 (2.2)	5 (0.1)	1018 (1.3)	2322 (1.8)	3052 (4.1)

[a] Weighted national case estimates. [b] $p < 0.0001$ for linear trend across volume groups.

RACHS-1 = Risk Adjustment in Congenital Heart Surgery, version 1.

We then focused on the most recent 5 years of data (2001 to 2005) to have the most up-to-date picture of the volume–mortality relationship. Mortality rates adjusted only for hospital volume and year of operation were no different across volume groups ($p = 0.45$). However, adjustment for RACHS-1 category and age revealed that high-volume hospitals continued to outperform all other volume categories. The discrepancy was most marked for very low-volume hospitals, where the OR increased to 2.54 (95% CI, 1.29 to 5.02; Table 4).

Comment

Our investigation of the relationship between hospital pediatric cardiac surgical volume and mortality in a national sample during an 18-year period found that in aggregate, hospitals that performed more than 200 operations per year outperformed all smaller-volume groups. This relationship persisted in the most recent 5-year period (2001 to 2005). Volume alone was an unreliable discriminator of mortality. These findings clearly illustrate the need for data adjusted for patient

Table 3. Logistic Regression Models for the Relationship Between Pediatric Cardiac Surgical Case Volume and In-Hospital Mortality (1988–2005)

	Model 1[a]		Model 2[b]	
Variable	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value
Annual hospital volume		<0.0001		<0.0001
≤ 20 cases/y	0.99 (0.70–1.39)	0.94	1.88 (1.30–2.73)	0.0009
21–100 cases/y	1.47 (1.25–1.73)	<0.0001	1.85 (1.56–2.20)	<0.0001
101–200 cases/y	1.29 (1.10–1.52)	0.0023	1.48 (1.24–1.77)	<0.0001
>200 cases/y	1.0 (Ref)		1.0 (Ref)	
RACHS-1 category				<0.0001
1			1.0 (Ref)	
2			4.27 (3.19–5.71)	<0.0001
3			10.47 (7.999–13.69)	<0.0001
4			15.59 (11.30–21.50)	<0.0001
5 and 6			38.62 (27.72–53.80)	<0.0001
Age group				<0.0001
≤30 days			4.14 (3.708–4.610)	<0.0001
31 days–1 year			1.97 (1.725–2.251)	<0.0001
>1 year			1.0 (Ref)	

[a] Model 1 contains 55,164 operations with 3145 deaths and has a C statistic of 0.60; model is adjusted for year. [b] Model 2 contains 51,703 operations with 2912 deaths and has a C statistic of 0.81; model is adjusted for year.

CI = Confidence interval; OR = odds ratio; RACHS-1 = Risk Adjustment in Congenital Heart Surgery, version 1.

Table 4. Logistic Regression Models for the Relationship Between Pediatric Cardiac Surgical Case Volume and In-Hospital Mortality (2001–2005)

	Model 3[a]		Model 4[b]	
Variable	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value
Annual hospital volume		0.45		0.013
≤20 cases/y	0.89 (0.48–1.65)	0.71	2.54 (1.29–5.02)	0.007
21–100 cases/y	1.13 (0.89–1.45)	0.31	1.44 (1.08–1.93)	0.014
101–200 cases/y	1.20 (0.94–1.53)	0.15	1.43 (1.05–1.96)	0.025
>200 cases/y	1.0 (Ref)		1.0 (Ref)	
RACHS-1 category				<0.0001
1			1.0 (Ref)	
2			4.24 (2.17–8.29)	<0.0001
3			10.42 (5.27–20.62)	<0.0001
4			16.44 (8.33–32.45)	<0.0001
5 and 6			36.06 (17.82–72.95)	<0.0001
Age group				<0.0001
≤30 days			5.28 (4.09–6.82)	<0.0001
31 days–1 year			1.94 (1.49–2.53)	<0.0001
>1 year			1.0 (Ref)	

[a] Model 3 contains 18,593 operations with 795 deaths and has a C statistic of 0.57; model is adjusted for year. [b] Model 4 contains 16,202 operations with 632 deaths and has a C statistic of 0.81; model is adjusted for year.

CI = Confidence interval; OR = odds ratio; RACHS-1 = Risk Adjustment in Congenital Heart Surgery, version 1.

and surgical case-mix for identifying better-performing, higher-quality hospitals.

Previous studies of the volume–mortality relationship in pediatric cardiac surgery have relied on state-level data. Three previous articles have used California administrative data from different time periods, 1988 (with the addition of Massachusetts administrative data from 1989), 1995–1997, and 1998–2003 [1, 4, 5]. Two other articles used data from the New York Cardiac Surgery Reporting System in overlapping time periods: 1992 to 1995 and 1990 to 1995 [2, 3]. All but the most recent of these investigations found an inverse relationship between hospital volume and mortality.

Although the numbers of patients in each of these studies were adequate, these studies were limited by the number of hospitals available for analysis (range, 16 to 52). More important, the number of hospitals in the largest volume categories ranged from 1 to 9. In contrast, our study contained 307 unique hospitals (27 in the largest volume group of > 200 cases/year) in the 1988 to 2005 cohort and 140 unique hospitals (17 in the largest volume group of > 200 cases/year) in the 2001 to 2005 cohort. A low number of hospitals in a volume group increases the relative impact of each individual hospital's mortality rate on the group mortality rate. As a result, an apparent volume–mortality relationship may instead be due to the influence of one high performer and not reflective of a "true" high performing group [5].

We did find a relationship between hospital case volume and mortality, but this finding must be interpreted with caution. First, the relationship was nonlinear. Although the ORs decreased as volume increased, there was no significant difference in risk-adjusted in-hospital mortality between the three volume categories of fewer than 200 cases per year. Second, although the threshold of 200 cases per year seems a reasonable division point based on the distribution of individual hospital mortality rates (Fig 2), it is still somewhat arbitrary. Our strategy of aggregating hospitals into volume groups gave us sufficient statistical power to analyze important relationships, but it disguised individual hospitals. One should not conclude that all hospitals performing fewer than 200 cases per year have high mortality rates. On average this was true; however, there were individual hospitals below this threshold that had low mortality rates and those that had volumes too low for any mortality rate difference to be observed.

Our analysis illustrates the inappropriateness of using volume alone as a marker of quality. The broad range of individual hospital mortality rates within volume groups resulted in volume alone being a poor discriminator of mortality. Notably, volume did not significantly improve the discrimination, as measured by the C-statistic, of a model that already included RACHS-1 category and patient age. This means that the effect of volume should be interpreted carefully, particularly from the perspective of the patient. On average, a patient's own risk characteristics and level of disease burden account for the vast majority of their mortality risk, and the impact of hospital volume on the mortality risk of an individual patient may be small. For example, the absolute difference in risk-adjusted mortality of hospitals with more than 200 cases per year and those at 200 or fewer cases per year was 1.7%, indicating that roughly 58 patients would

have to undergo operations in high-volume hospitals to save one life.

We chose to use the RACHS-1 method of risk adjustment for this investigation. RACHS-1 was developed to compare the mortality for groups of patients undergoing congenital cardiac operations. Certain procedures that occur less commonly are not included in the RACHS-1 method; however, low-frequency procedures are less important when comparing overall performance. In addition, the methodology does not allow for specific conditions that may be risk factors for specific operations. However, such potential risk factors are likely to be evenly distributed in large populations. Despite these limitations, the RACHS-1 method is a widely used and validated risk-adjustment methodology for congenital heart surgery [12–14].

Our analysis has several limitations. First, although the NIS is the largest all-payer inpatient care database in the United States, with a sampling frame that comprises 90% of all hospital discharges, it is a sample and not a compete database of all hospital discharges. As a result, although it is designed to be representative of national practice, there is the possibility for error. However, the NIS is the largest collection of real data that can currently be used for addressing the present question.

Second, the NIS is an administrative database. Administrative databases were designed for claims data collection and billing, not heath care research, and can be limited by erroneous coding of congenital diagnoses [15, 16]. As a result, we may not have adequately accounted for differences in case-mix; however, there is no evidence to suggest that miscoding is related to hospital volume. Given the large sample size, random miscoding would be unlikely to significantly influence our findings. Any potential impact would likely bias the results towards the null. Thus, it is possible that volume would have had a larger effect on mortality rates that were adjusted using a more robust set of clinical variables than those that are found in administrative data. We reduced error from miscoding of data in this study by only including patients where the procedure code was matched with a plausible diagnostic code.

Notwithstanding these limitations, our study was conducted using a large, national data set with adequate power to generate current, stable mortality rates. Administrative data are widely available to payers, governmental agencies, and groups that produce Internet reports and other reports that reach patients, their families, and providers. At present, administrative data are being used to evaluate heath care quality. Such data provide a national picture of the practice of medicine. By including information from both high- and low-performing and high- and low-volume hospitals, the data can be used to evaluate the practice of hospitals that are less likely to participate in voluntary, clinical databases.

The present study found that volume alone was a poor discriminator of mortality and only marginally better than a coin flip (ROC curve area, 0.50). However, in aggregate, large-volume hospitals performed more complex operations and achieved superior results; therefore, the use of overall, unadjusted mortality rates to evaluate institution quality is misleading. Hospital comparisons and pay-for-performance initiatives must be based on robust risk-adjusted comparisons. Many factors contribute to the mortality risk of a pediatric patient undergoing a cardiac operation. Volume is not a measure of quality, but rather an easily obtained structural attribute associated with quality. It is likely a surrogate for process measures and characteristics of systems of care that lead to better outcomes, but are not currently captured in administrative or most clinical databases. As a result, institution-specific risk-adjusted outcomes are likely to be more informative than a volume threshold. Future efforts should focus on identifying the fundamental elements that are apt to explain the variation in outcomes currently attributed to volume.

References

1. Jenkins KJ, Newburger JW, Locke JE, et al. In-hospital mortality for surgical repair of congenital heart defects: preliminary observations of variation by hospital caseload. Pediatrics 1995;95:323–30.
2. Hannan EL, Racz M, Kavey RE, Quaegbeur JM, Williams R. Pediatric cardiac surgery: the effect of hospital and surgeon volume on in-hospital mortality. Pediatrics 1998;101:963–9.
3. Sollano JA, Gelijns AC, Moskowitz AJ, et al. Volume-outcome relationships in cardiovascular operations: New York State, 1990–1995. J Thorac Cardiovasc Surg 1999;117:419–28.
4. Chang RK, Klitzner TS. Can regionalization decrease the number of deaths for children who undergo cardiac surgery? A theoretical analysis. Pediatrics 2002;109:173–81.
5. Bazzani LG, Marcin JP. Case volume and mortality in pediatric cardiac surgery patients in California, 1998–2003. Circulation. 2007;115:2652–9.
6. HCUP Nationwide Inpatient Sample (NIS). Healthcare Cost and Utilization Project (HCUP). 1988–2005. Agency for Healthcare Research and Quality, Rockville, MD. www.hcup-us.ahrq.gov/nisoverview.jsp.
7. Jenkins KJ, Gauvreau K, Newburger JW, Spray TL, Moller JH, Iezzoni LI. Consensus-based method for risk adjustment for surgery for congenital heart disease. J Thorac Cardiovasc Surg 2002;123:110–8.
8. Ash AS, Shwartz M. Evaluating the performance of risk-adjustment methods: dichotomous outcomes. In: Iezzoni LI, ed. Risk adjustment for measuring healthcare outcomes.2nd ed. Chicago, IL: Health Administration Press; 1997:427–70.
9. Agresti A. Categorical data analysis. New York, NY: John Wiley & Sons; 1990:306–46.
10. White H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 1980;48:817–30.
11. Hannan EL, Racz MJ, Jollis JG, Peterson ED. Using Medicare claims data to assess provider quality for CABG surgery. Does it work well enough? Health Serv Res 1997;31:659–78.
12. Welke KF, Shen I, Ungerleider RM. Current assessment of mortality rates in congenital cardiac surgery. Ann Thorac Surg 2006;82:164–71.
13. Jenkins KJ, Gauvreau K. Center-specific differences in mortality: oreliminary analyses using the Risk Adjustment in Congenital Heart Surgery (RACHS-1) method. J Thorac Cardiovasc Surg 2002;124:97–104.

14. Al-Radi OO, Harrell FE Jr, Caldarone CA, et al. Case complexity scores in congenital heart surgery: A comparative validation study of the Aristotle Basic Complexity score and the Risk Adjusted Congenital Heart Surgery (RACHS-1) system. J Thorac Cardiovasc Surg 2007;133:865–75.
15. Cronk CE, Malloy ME, Pelech AN, et al. Completeness of state administrative databases for surveillance of congenital heart disease. Birth Defects Res A Clin Mol Tetratol 2003;67:597–603.
16. Frohnert BK, Lussky RC, Alms MA. Validity of hospital discharge data for identifying infants with cardiac defects. J Perinatol 2005;25:727–42.

DISCUSSION

DR J. WILLIAM GAYNOR (Philadelphia, PA). I would like to congratulate Karl and his coworkers on a very interesting study and an excellent presentation. As they noted, there is increased demand for reporting outcomes of pediatric cardiac surgical programs by parents, payers, and government agencies, and these data can be used for benchmarking, quality assurance, and, increasingly, pay for performance. There has been concern as to whether or not there is a true relationship between institutional volume and mortality. This relationship has been controversial and the data are often conflicting.

Dr Welke and his colleagues used a large data set derived from a national sample of administrative discharge data to examine the relationship between institutional volume and hospital mortality. Not surprisingly, they found that institutional volume alone was a very poor predictor of outcome. They then examined the impact of patient factors (age at surgery) and case-mix on this relationship. They identified an intriguing relationship between volume and case complexity: large institutions cared for younger, more complex patients but yet had superior results. The case complexity and the age were the most important predictors of outcome, and the addition of volume to the model added little to the predictive accuracy of the statistical model. Nevertheless, volume cannot be ignored, because the large centers operated on the sickest patients with the best results.

This study highlights the pitfalls and the risk of simple comparisons of outcomes data which are frequently used in the media. It demonstrates the need for risk or complexity adjustment. We could have a whole discussion over which method to use for that. But we need to have some method to adjust for risk or complexity when we are comparing outcomes.

The relationship between volume and case-mix is intriguing. What factors lead to a high surgical volume? Do the centers have high volume because they are doing complex cases with good results and attract patients, or do they have good results because they are larger and they have instituted structural changes and processes that improve their outcomes? Which comes first, the chicken or the egg?

I have three questions for Dr Welke. Could you identify any of the characteristics in the hospitals that might affect the outcomes, such as did a stand-alone children's hospital have different outcomes from a hospital that cares for both children or adults, or were they more likely to be represented in the high-volume programs? The confirmation of a relationship between volume and outcome would obviously have significant impact for patients, physicians, and payers, which could lead to changes in referral patterns and even closure of programs. Therefore, it is imperative to be sure that such a relationship exists.

There are significant problems with administrative data sets. They have the advantages of large numbers, but the details may be very inaccurate. At our institution, I compared the coding from the STS [Society of Thoracic Surgeons] database, which is done by the attending surgeon, with the discharge coding for 20 patients, and in this very small sample there were significant errors in coding in over 50% of the patients on the administrative data set. So what are the next steps that are needed to verify whether or not this relationship between volume and outcome really exists?

And finally, how can we identify the factors that are associated with the improved outcomes at the large volume centers and apply these to the smaller centers?

Once again I would like to thank the Society for the privilege of discussing the paper and congratulate Dr Welke.

DR WELKE: Thank you, Dr Gaynor, for your insightful comments and a thorough review of our work. I will address your questions in order.

First, it is often difficult to find associations between hospital-level characteristics and outcomes. There are a number of hospital-level characteristics available in administrative data and in the Nationwide Inpatient Sample (NIS) specifically. None of those that we examined were associated with mortality. Designation as a stand-alone children's hospital is not noted in the NIS. It is available in the Kids Inpatient Database (KID); however, we chose not to use the KID for this study for two reasons. First, the KID is only available for three nonconsecutive years, which would have limited our hospital sample size. Second, the sampling methodology of the KID is such that only 80% of cases at each sampled hospital are included. In contrast, the NIS captures all admissions to a sampled hospital, making it better suited for determining hospital volumes. Children's hospital designation is recorded in the American Hospital Association (AHA) database and from there can be linked to the NIS. We are linking AHA data with the NIS so that we can examine the question. I suspect that designation as a children's hospital may be associated with volume, but I do not have the data to support that.

Your comments about administrative data are absolutely correct. Administrative data are not perfect and coding errors do exist. Because of these issues, it is important that we restrict our analyses of administrative data to large populations rather than specific hospitals or surgeons. If coding errors are random, which is more likely with a large cohort, their impact is likely to bias the results toward the null. So, if associations exist, they are more likely to be real. The Society of Thoracic Surgeons (STS) is currently working with the Agency for Healthcare Research and Quality to compare the International Classification of Diseases, Ninth Revision, Clinical Modification codes used in administrative data with the codes used in the STS Congenital Heart Surgery Database in order to quantify and then rectify the coding problems. One way to check the validity of a finding is to test the same hypothesis in a different cohort. We have looked at the volume mortality relationship in clinical data by using the STS Congenital Heart Surgery Database and have found complementary results. We are going to present those results later this year.

Lastly, volume is a surrogate for quality, not a measure of quality in and of itself. We need to identify the structural measures and process measures that account for the apparent contribution of volume to outcome. This may be not come from analyzing large administrative or even clinical databases. Instead it may require multicenter quality improvement initiatives

such as the Northern New England Cardiovascular Disease Study Group. Site visits to well-performing hospitals, regardless of size, may elucidate common factors that can be implemented at lower-performing hospitals. Some of these factors are likely to be the process measures in the list developed by this Society for presentation to the National Quality Forum. However, it will take some time to find out which of these factors are most influential. Thank you very much.

The Thoracic Surgery Foundation for Research and Education

Grants, Fellowships, and Career Development Awards

The Thoracic Surgery Foundation for Research and Education (TSFRE) was founded in 1991 to bring together the research and education support efforts of the four major societies in cardiothoracic surgery in the United States: The American Association for Thoracic Surgery, The Society of Thoracic Surgeons, The Southern Thoracic Surgical Association, and the Western Thoracic Surgical Association. Because of its close and continuing relationship with organized cardiothoracic surgery, TSFRE attracts the highest quality research award applicants and truly outstanding reviewers.

Any surgeon who meets the eligibility requirements is invited to submit an application. Research grants will be judged separately from research fellowship applications. In general, top-scoring applications in each category will receive priority with respect to funding.

Multiple fellowship applications under the sponsorship of an individual mentor or multiple grant applications from a single institution will be accepted and reviewed, as long as there is no significant scientific overlap. Only under extraordinary circumstances will the TSFRE fund simultaneous awards to a single institution.

This year, TSFRE is proud to offer the following awards to the most promising cardiothoracic surgeon-scientists:

The Nina Starr Braunwald Career Development Award
Provides a biennial award of $110,000 for 2 years to support the research career development of a woman cardiac surgeon who holds a full-time faculty appointment and who is within 10 years of completion of thoracic surgery residency. Deadline: October 15

TSFRE Research Grants
Provides operational support of original research efforts by cardiothoracic surgeons who have completed their formal training, and who are seeking initial support and recognition for their research program. Awards of up to $30,000 a year for up to 2 years are made each year to support the work of an early-career cardiothoracic surgeon (within 5 years of first faculty appointment). Deadline: October 15

TSFRE Research Fellowships
Provides support of up to $35,000 a year for up to 2 years for surgical residents who have not yet completed cardiothoracic surgical training. Deadline: October 15

TSFRE Career Development Awards
Provide salary support of up to $50,000 a year for up to 2 years for applicants who have completed their residency training and who wish to pursue investigative careers in cardiothoracic surgery. Deadline: October 15

TSFRE/NHLBI Jointly Sponsored Mentored Clinical Scientist Development Award—K08 or K23
Provides support to outstanding clinician research scientists who are committed to a career in cardiothoracic surgery research and have the potential to develop into independent investigators. The award is $150,000 a year ($75,000 from TSFRE and $75,000 from NHLBI) plus $25,000 indirect support from the NHLBI and supports a 3-, 4-, or 5-year period of didactic training and supervised research experience. Deadline: May 31

TSFRE/NCI Jointly Sponsored Mentored Clinical Scientist Development Award—K08 or K23
Provides support to outstanding clinically trained professionals who are committed to a career in laboratory or field-based research and have the potential to develop into independent investigators. The award is $150,000 a year ($75,000 from TSFRE and $75,000 from NCI) plus $30,000 indirect support from the NCI and supports a 5-year period of supervised research that integrates didactic studies with laboratory or clinically based research. Deadline: February 1 and October 1

The American Association for Thoracic Surgery Awards
Provides $75,000 in support for 1 year through the *Evarts A. Graham Memorial Traveling Fellowship* to a non-North American young cardiothoracic surgeon future international leader for further development in the United States. The AATS also provides $75,000 a year for 2 years of support for young North American cardiothoracic surgeons committed to pursuing an academic career in cardiothoracic surgery through the AATS Research Scholarship. Additionally, AATS provides $5,000 travel grants to broaden the educational experience of North American residents in their final year of residency through the Resident Traveling Fellowship. Deadline: July 1

Applications will be available online only and can be found at *www.tsfre.org.* For more information, please address inquiries to:

Chair, Research Committee
Thoracic Surgery Foundation for Research and Education
900 Cummings Center, Suite 221-U
Beverly, MA 01915
Telephone: (978) 927-8330