An empirically based tool for analyzing morbidity associated with operations for congenital heart disease

Marshall L. Jacobs, MD,^a Sean M. O'Brien, PhD,^b Jeffrey P. Jacobs, MD,^c Constantine Mavroudis, MD,^d Francois Lacour-Gayet, MD,^e Sara K. Pasquali, MD,^f Karl Welke, MD,^g Christian Pizarro, MD,^h Felix Tsai, MD,ⁱ and David R. Clarke, MD^j

Objective: Congenital heart surgery outcomes analysis requires reliable methods of estimating the risk of adverse outcomes. Contemporary methods focus primarily on mortality or rely on expert opinion to estimate morbidity associated with different procedures. We created an objective, empirically based index that reflects statistically estimated risk of morbidity by procedure.

Methods: Morbidity risk was estimated using data from 62,851 operations in the Society of Thoracic Surgeons Congenital Heart Surgery Database (2002-2008). Model-based estimates with 95% Bayesian credible intervals were calculated for each procedure's average risk of major complications and average postoperative length of stay. These 2 measures were combined into a composite morbidity score. A total of 140 procedures were assigned scores ranging from 0.1 to 5.0 and sorted into 5 relatively homogeneous categories.

Results: Model-estimated risk of major complications ranged from 1.0% for simple procedures to 38.2% for truncus arteriosus with interrupted aortic arch repair. Procedure-specific estimates of average postoperative length of stay ranged from 2.9 days for simple procedures to 42.6 days for a combined atrial switch and Rastelli operation. Spearman rank correlation between raw rates of major complication and average postoperative length of stay was 0.82 in procedures with n greater than 200. Rate of major complications ranged from 3.2% in category 1 to 30.0% in category 5. Aggregate average postoperative length of stay ranged from 6.3 days in category 1 to 34.0 days in category 5.

Conclusions: Complication rates and postoperative length of stay provide related but not redundant information about morbidity. The Morbidity Scores and Categories provide an objective assessment of risk associated with operations for congenital heart disease, which should facilitate comparison of outcomes across cohorts with differing case mixes. (J Thorac Cardiovasc Surg 2013;145:1046-57)

Supplemental material is available online.

Earn CME credits at
<http://cme.ctsnetjournals.org>

From the Department of Pediatric and Congenital Heart Surgery,^a Cleveland Clinic, Cleveland, Ohio; Department of Biostatistics,^b Duke University School of Medicine and Duke Clinical Research Institute, Duke University Medical Center, Durham, NC; The Congenital Heart Institute of Florida,^c All Children's Hospital and Children's Hospital of Tampa, Cardiac Surgical Associates of Florida, University of South Florida College of Medicine, St Petersburg and Tampa, Fla; Congenital Heart Institute,^d Florida Hospital for Children, Orlando, Fla; Pediatric Cardiac Surgery Department,^e Montefiore Children's Hospital, Montefiore Medical Center, New York, NY; Department of Pediatrics,^f Duke University School of Medicine and Duke Clinical Research Institute, Duke University Medical Center, Durham, NC; Cardiothoracic Surgery,^g Seattle Children's Hospital, Seattle, Wash; Nemours Cardiac Center,^h Wilmington, Del; Cardiac Surgery,ⁱ Children's Hospital of the King's Daughters, Norfolk, Va; and Department of Cardiothoracic Surgery,^j The Children's Hospital, Aurora, Colo.

Disclosures: Authors have nothing to disclose with regard to commercial support. Received for publication Dec 29, 2011; revisions received April 26, 2012; accepted for publication June 12, 2012; available ahead of print July 26, 2012.

Address for reprints: Marshall L. Jacobs, MD, 6019 Goshen Rd, Newtown Square, PA 19073 (E-mail: marshall.jacobs@comcast.net).

0022-5223/\$36.00

Copyright © 2013 by The American Association for Thoracic Surgery

<http://dx.doi.org/10.1016/j.jtcvs.2012.06.029>

Contemporary efforts to describe and compare congenital heart surgery outcomes across institutions have evolved to include (1) use of clinical registry data, rather than administrative data from the hospital bill to evaluate outcomes; (2) use of empiric rather than opinion-based models to adjust for differences in case complexity across institutions; and (3) recognition that focusing solely on in-hospital mortality overlooks 96% of patients who survive to hospital discharge and the important morbidities that they may experience.¹

In 2009, an empirically based tool for analyzing mortality associated with congenital heart surgery was introduced. The Society of Thoracic Surgeons-European Association for Cardiothoracic Surgery (STS-EACTS) Congenital Heart Surgery Mortality Score and Categories are based on analysis of 148 different types of operations performed in 77,294 patients.² Procedures are assigned to 1 of 5 categories on the basis of a similar risk of in-hospital death.

Abbreviations and Acronyms

CrI	= credible interval
ICU	= intensive care unit
PLOS	= postoperative length of stay
STSCCHSD	= Society of Thoracic Surgeons Congenital Heart Surgery Database
STS-EACTS	= Society of Thoracic Surgeons- European Association for Cardiothoracic Surgery

Category 1 has the lowest risk of death, and category 5 has the highest risk of death. In addition, each procedure receives a numeric score ranging from 0.1 to 5.0 that expresses mortality risk on a more continuous scale. The STS-EACTS Mortality Categories are intended to facilitate analysis of outcomes by grouping procedures with similar risk of in-hospital mortality.

Although congenital heart surgery outcomes analyses have traditionally focused on mortality, comprehensive assessment requires attention to other end points. Nonfatal events, such as stroke and renal failure, are major determinants of hospital cost and patients' health status after surgery. In addition, postprocedure length of hospital stay provides useful direct information about resource use and indirect proxy information about a patient's condition.^{3,4} Although such measures are captured in clinical registries, tools for analyzing these end points are lacking.

The goal of the present study was to develop a new system for classifying congenital heart surgery procedures on the basis of their potential for morbidity using empirical data from the STS Congenital Heart Surgery Database (STSCCHSD). There were 4 specific objectives:

- to develop a morbidity metric based on both the occurrence of complications that have a significant and durable impact on patient health and utilization of health care resources;
- to estimate the average amount of patient morbidity by procedure type;
- to convert these procedure-specific morbidity estimates into a scale ranging from 0.1 to 5.0 (range was chosen for consistency with the STS-EACTS Mortality Score²); and
- to group procedures with similar estimated morbidity risk into 5 relatively homogeneous categories that were designed to minimize within-category variation and to serve as a stratification variable that can be used to adjust for case mix when analyzing outcomes and comparing institutions.

The Morbidity metric was developed primarily for the purpose of grouping types of procedures to better describe case mix, as was the STS-EACTS Mortality metric. The intent was not to assess or predict outcomes for an individual patient or surgeon, for which other types of analyses may be used.

MATERIALS AND METHODS

Study Population

The STSCCHSD has been described.⁵ The Duke Clinical Research Institute serves as the data analysis center for STS databases and has an agreement and institutional review board approval to analyze the aggregate deidentified data for research purposes. For this study, operations were included if they took place between January 1, 2002, and December 31, 2008, and were 1 of the 148 types of cardiovascular procedures for which the STS-EACTS Mortality Score is defined.² Operations performed at centers with no more than 10% missing data for complications, mortality, or postoperative length of stay (PLOS) were eligible for inclusion in the analysis. From eligible centers, individual operations with missing data for complications, mortality, or PLOS were excluded. Of 63,297 potentially eligible operations, 446 individual operations were excluded on the basis of missing data regarding complications (n = 273), PLOS (n = 151), or mortality (n = 22).

Additional inclusion and exclusion criteria were identical to those used for developing the STS-EACTS Mortality Score.² Only the first operation of each hospital admission was analyzed. The final study population consisted of 62,851 operations classified into 148 procedure types at 68 centers. Results are presented for the subset of 140 procedure types having at least 10 eligible cases (62,819 total operations; 99.9%).

Classification of Multiple-Procedure Operations

Several operations in the analysis represent combinations of 2 or more procedures. These are analyzed as combined procedures because the complexity of the combination is regarded as being different from the complexity of the component procedures when performed in isolation. For each of these combined procedures, unique procedure codes were subsequently assigned in STSCCHSD version 3.0. Because all data in this analysis predate version 3.0, classification of multiple-procedure operations in this study follows guidelines set forth previously in development of the STS-EACTS Mortality Score and Categories.²

End Points

Morbidity was quantified for each procedure on the basis of the proportion of patients experiencing major complications and by the average PLOS (Table 1). Major complication was defined as the occurrence of any 1 or more of the 6 complications listed in Table 2. These complications represent definitive outcomes that can be ascertained reliably and that are likely to have significant and durable impact on patient health. The unadjusted rate of major complications is defined as the percent of operations that were associated with the occurrence of 1 or more of the major complications listed in Table 2. PLOS was defined as the number of days from the date of operation to the date of discharge and was determined for all patients, including those who died in-hospital.

Analysis

Statistics calculated for each procedure type included the number of eligible operations, the percent of patients experiencing major complications, the 95% binomial confidence interval for the probability of major complications, and the average and interquartile range (25th and 75th percentiles) of PLOS. Model-based estimates of each procedure's average risk of major complications and average PLOS were calculated by hierarchical modeling and presented along with 95% Bayesian credible intervals (CrIs). Details of these calculations are provided in Appendix 1.

Creation of Morbidity Scores

To facilitate ranking and grouping of procedures, average risk of major complications and average PLOS were combined into a single composite morbidity measure. To account for different measurement scales, the 2 individual measures were rescaled to have the same standard deviation (Appendix 3). They were then summed together. The resulting composite morbidity measure was the basis of the proposed Morbidity Scores and

Categories. Each procedure was assigned a numeric score ranging from 0.1 to 5.0 (STS Congenital Heart Surgery Morbidity Score). The range was chosen to be the same as the existing STS-EACTS Mortality Score. Scores were assigned by shifting and rescaling the procedure-specific composite morbidity estimates to lie in the interval from 0.1 to 5.0 and then rounding to 1 decimal place.

Creation of Morbidity Categories

Procedures were sorted by increasing estimated morbidity and partitioned into 5 relatively homogeneous categories (STS Congenital Heart Surgery Morbidity Categories). This number of categories was chosen to match the number of STS-EACTS Mortality Categories.² A computer program was used to search for category cutpoints that were optimal for minimizing within-category variance and maximizing between-category variance of the composite morbidity measure. The relationship between the number of categories and the degree of within-category homogeneity was assessed (Figure E1).

Sensitivity Analyses

Sensitivity analyses were performed to assess whether the ranking of procedures depended heavily on the choice of statistical methodology. The Spearman rank correlation coefficient was used to quantify the extent to which rankings differed. Differences were also assessed graphically by plotting estimates of the same quantity calculated by 2 different statistical methods.

Assessment of Statistical Reliability

Finally, we estimated the statistical precision (reliability) of the estimated rates of major complication, average PLOS, and composite morbidity. Reliability of a set of estimates is conventionally defined as the proportion of between-unit variation that is explained by true between-unit differences (ie, signal) as opposed to random statistical fluctuations (ie, noise). A mathematically equivalent definition is the squared correlation between a measurement and the true value. In our case, reliability was defined as the squared Pearson correlation between each procedure's estimated and true amount of morbidity. Reliability could not be calculated directly (because the "true" morbidity values are unknown) but was estimated by hierarchical modeling, as described in Appendix 1.

RESULTS

Sample sizes per procedure ranged from 1 to 4868. The 140 procedures with at least 10 cases are listed in Table 1 along with their sample sizes, raw and model-based morbidity estimates, and Morbidity Scores and Categories.

Model-estimated risk of major complications ranged from 1.0% for atrial septal defect repair to 38.2% for truncus arteriosus with interrupted aortic arch repair. Procedure-specific estimates of average PLOS ranged from 2.9 days for implantable cardioverter defibrillator procedure to 34.6 days for a stage 1 Norwood procedure and 42.6 days for a combined atrial switch and Rastelli procedure for congenitally corrected transposition. The Spearman rank correlation between raw rates of major complication and average PLOS was 0.63 (Figure 1) in procedures with at least 10 cases and 0.82 in procedures with at least 200 cases. This degree of correlation suggests that complication rates and PLOS provide related, but not redundant, information about morbidity.

Procedure-specific overall morbidity was defined as $0.141 \times$ percentage rate of major complications + $0.162 \times$ average PLOS in days. The numbers 0.141 and

0.162 were calculated as the reciprocals of the standard deviations of the percentage rate of major complications and average PLOS, respectively. The STS Morbidity Score was obtained by rescaling this overall morbidity measure to lie in the interval 0.1 to 5.0. Thus, by design it ranged from 0.1 to 5.0. Procedures with the least morbidity (STS Morbidity Score = 0.1) include atrial septal defect repair and implantable cardioverter defibrillator procedures. The procedure with the greatest morbidity (STS Morbidity Score = 5.0) was repair of truncus arteriosus with interrupted aortic arch.

STS Morbidity Categories were obtained by grouping procedures into 5 unequally sized categories (1 = least morbidity, 5 = most morbidity) chosen to be maximally homogeneous with respect to overall morbidity. The number of procedures assigned to categories 1, 2, 3, 4, and 5 were 36, 43, 36, 21, and 4, respectively. The rate of major complication ranged from 3.2% in category 1 to 30.0% in category 5. The aggregate average PLOS ranged across categories from 6.3 days in category 1 to 34.0 days in category 5.

Several analyses were performed to address potential methodological concerns with the composite measures used in this analysis. First, we addressed potential issues related to "Major Complication," which is a composite designating the occurrence of any 1 or more of 6 individual complications. The observed rate of discharge mortality for patients who experienced at least 1 major complication was 23.5% (Table 2) in comparison with 2.0% among patients who experienced none of the major complications. When end points in a composite occur with differing frequencies, the more frequent end points may sometimes dominate.⁶ As shown in Table 2, the aggregate rate of major complications ranged from 0.8% for "postoperative neurologic deficit persisting at discharge" to 4.7% for "unplanned reoperation." To verify that each individual complication contributed statistical information but did not dominate the composite, we calculated the Spearman rank correlation coefficient between procedure-specific rates of each individual complication and rates of any major complication. These correlations ranged from 0.37 for heart block to 0.79 for unplanned reoperation. Thus, although unplanned reoperation explained much of the variation in the major complication end point, no single item dominated. All 6 complications contributed statistical information.

Second, we assessed the impact of modifying the list of major complications to include mortality. Although mortality was ultimately excluded, we thought it was important to know whether results would be similar or different had mortality been included. To address this, we calculated 2 versions of the major complication end point (1 including and 1 excluding mortality) and compared them. As shown in Figure 2, the 2 major complication end points were highly correlated but not perfectly related. The rank correlation coefficient between them was 0.97.

TABLE 1. Procedure names, morbidity categories and scores, and data for model development

Procedure name (STS code)	No.	Morbidity		PLOS (d)		Any major complications	
		Category	Score	Average (IQR)	Model-based estimate (95% CrI)	Raw (95% CI)	Model (95% CrI)
ASD repair, patch (30)	2469	1	0.1	4.1 (3.0-4.0)	4.1 (3.8-4.4)	0.8% (0.5-1.2)	1.0% (0.7-1.4)
ICD (AICD) implantation (1470)	358	1	0.1	3.2 (1.0-3.0)	3.2 (2.6-3.9)	1.7% (0.8-3.6)	2.1% (1.1-3.5)
ICD (AICD) procedure (1480)	156	1	0.1	2.8 (1.0-3.0)	2.9 (2.1-3.7)	2.6% (1.0-6.4)	2.9% (1.3-5.1)
ASD repair, patch + PAPVC repair (3100)	319	1	0.1	3.7 (3.0-4.0)	3.7 (3.4-3.9)	1.6% (0.7-3.6)	2.1% (1.1-3.7)
Pacemaker procedure (1460)	1413	1	0.2	3.9 (0.0-3.0)	3.9 (3.1-4.7)	2.0% (1.4-2.8)	2.1% (1.5-2.9)
ASD repair, primary closure (20)	947	1	0.2	5.0 (3.0-4.0)	5.0 (4.4-5.6)	1.3% (0.7-2.2)	1.6% (1.0-2.4)
AVC (AVSD) repair, partial (incomplete) (PAVSD) (190)	772	1	0.3	5.1 (3.0-5.0)	5.1 (4.6-5.7)	2.1% (1.3-3.3)	2.3% (1.5-3.3)
DCRV repair (570)	352	1	0.3	5.1 (3.0-5.0)	5.1 (4.6-5.6)	2.8% (1.6-5.1)	3.1% (1.7-4.9)
PAPVC repair (260)	370	1	0.4	5.4 (3.0-5.0)	5.4 (4.2-6.7)	2.4% (1.3-4.6)	2.8% (1.6-4.5)
Vascular ring repair (1360)	1006	1	0.4	6.3 (2.0-4.0)	6.3 (5.4-7.2)	1.5% (0.9-2.4)	1.8% (1.2-2.6)
Pacemaker implantation, permanent (1450)	1147	1	0.4	5.7 (1.0-4.0)	5.7 (4.8-6.6)	2.4% (1.7-3.5)	2.6% (1.8-3.6)
Aneurysm, pulmonary artery, repair (1740)	16	1	0.4	4.7 (3.0-5.0)	4.9 (2.6-7.3)	0.0% (0.0-19.4)	3.6% (1.2-7.9)
Coronary artery fistula ligation (1290)	29	1	0.4	4.9 (3.0-6.0)	5.0 (3.6-6.3)	3.4% (0.2-17.2)	3.9% (1.6-8.3)
Aortic stenosis, subvalvar, repair (780)	1573	1	0.4	4.7 (3.0-5.0)	4.7 (4.5-5.0)	4.3% (3.4-5.4)	4.2% (3.3-5.2)
Valve replacement, pulmonic (600)	1097	1	0.4	5.3 (3.0-6.0)	5.3 (5.0-5.6)	3.6% (2.7-4.9)	3.6% (2.7-4.8)
Coarctation repair, interposition graft (1250)	102	1	0.4	6.1 (4.0-6.8)	6.1 (5.0-7.3)	2.0% (0.5-6.9)	3.2% (1.4-5.9)
Valvuloplasty, pulmonic (590)	230	1	0.5	6.9 (3.0-5.8)	6.9 (5.2-8.5)	1.7% (0.7-4.4)	2.7% (1.3-4.6)
Atrial baffle procedure (non-Mustard, non-Senning) (310)	23	1	0.5	6.3 (3.0-7.5)	6.4 (4.8-8.0)	0.0% (0.0-14.3)	3.8% (1.3-8.1)
VSD repair, primary closure (100)	448	1	0.5	6.0 (3.0-5.0)	6.0 (5.2-6.8)	4.7% (3.1-7.1)	4.6% (3.0-6.6)
Valvuloplasty, aortic (660)	667	1	0.6	6.4 (3.0-6.0)	6.4 (5.6-7.2)	4.2% (2.9-6.0)	4.2% (2.9-5.8)
Conduit placement, LV to PA (620)	19	1	0.6	6.8 (4.0-9.0)	6.8 (5.0-8.7)	0.0% (0.0-16.8)	4.0% (1.4-8.8)
Aortic root replacement, valve sparing (735)	102	1	0.6	6.0 (4.0-7.0)	6.0 (5.4-6.6)	5.9% (2.7-12.2)	5.1% (2.6-8.9)
VSD repair, patch (110)	4868	1	0.6	7.7 (3.0-7.0)	7.7 (7.3-8.0)	3.3% (2.9-3.9)	3.4% (2.9-3.9)
AVC (AVSD) repair, intermediate (transitional) (180)	344	1	0.6	6.9 (3.0-6.0)	7.0 (5.8-8.1)	4.7% (2.9-7.4)	4.7% (2.9-6.9)
Coarctation repair, end to end (1210)	1214	1	0.6	9.0 (3.0-7.0)	8.9 (7.9-9.9)	2.2% (1.5-3.2)	2.5% (1.7-3.4)
TOF repair, ventriculotomy, nontransannular patch (360)	693	1	0.7	7.3 (4.0-7.0)	7.3 (6.7-7.8)	4.8% (3.4-6.6)	4.8% (3.3-6.4)
Conduit reoperation (580)	1244	1	0.7	7.2 (4.0-6.0)	7.2 (6.6-7.9)	5.4% (4.3-6.8)	5.3% (4.2-6.6)
PFO, primary closure (10)	129	1	0.7	7.4 (4.0-7.0)	7.5 (5.9-9.1)	5.4% (2.7-10.8)	5.2% (2.7-8.7)
TOF repair, no ventriculotomy (350)	768	1	0.7	8.3 (5.0-8.0)	8.3 (7.7-9.0)	4.2% (3.0-5.8)	4.3% (3.0-5.7)
Ross procedure (740)	480	1	0.7	5.9 (4.0-6.0)	5.9 (5.5-6.3)	7.9% (5.8-10.7)	7.4% (5.4-9.8)
PA, reconstruction (plasty), main (trunk) (530)	175	1	0.8	7.1 (3.0-6.0)	7.3 (5.7-8.9)	6.3% (3.5-10.9)	5.8% (3.3-9.2)
Atrial fenestration closure (85)	19	1	0.8	7.6 (4.0-10.0)	7.8 (5.2-10.5)	5.3% (0.3-24.6)	5.3% (2.0-11.2)
ASD, common atrium (single atrium), septation (50)	41	1	0.8	8.5 (3.0-5.0)	8.7 (4.0-13.5)	2.4% (0.1-12.6)	4.6% (1.7-9.5)
Valvuloplasty, mitral (830)	1365	1	0.8	8.7 (4.0-7.0)	8.7 (8.0-9.4)	5.0% (3.9-6.3)	5.0% (4.0-6.2)
Cor triatriatum repair (290)	127	1	0.8	9.6 (4.0-7.0)	9.5 (6.9-12.2)	3.1% (1.2-7.8)	4.3% (2.1-7.4)
PA debanding (1650)	75	1	0.8	8.5 (4.0-9.5)	8.5 (6.6-10.4)	5.3% (2.1-12.9)	5.4% (2.5-9.6)
AVR, bioprosthetic (690)	90	2	0.9	6.6 (3.2-6.0)	6.8 (5.4-8.3)	11.1% (6.1-19.3)	8.3% (4.4-13.8)
Aortic stenosis, supravalvar, repair (790)	268	2	0.9	7.7 (4.0-7.0)	7.8 (6.5-9.1)	7.8% (5.2-11.7)	7.3% (4.7-10.4)
Coarctation repair, end to end, extended (1220)	1773	2	0.9	11.6 (4.0-10.0)	11.5 (10.6-12.5)	2.9% (2.2-3.8)	3.1% (2.3-3.9)
Sinus of Valsalva, aneurysm repair (810)	33	2	0.9	7.0 (4.0-9.0)	7.3 (5.5-9.2)	15.2% (6.7-30.9)	8.4% (3.6-15.7)
Aortopexy (1365)	50	2	1	20.7 (2.0-8.5)	11.1 (2.6-19.3)	0.0% (0.0-7.1)	4.4% (1.4-9.3)
Valvuloplasty, tricuspid (460)	927	2	1	9.4 (4.0-8.0)	9.4 (8.2-10.7)	6.4% (5.0-8.1)	6.3% (4.9-7.9)
Aortic root replacement, bioprosthetic (715)	47	2	1	6.8 (4.0-7.0)	7.1 (5.6-8.6)	14.9% (7.4-27.7)	9.1% (4.2-16.0)
Conduit placement, other (1772)	38	2	1	9.5 (4.0-7.0)	9.9 (5.3-14.4)	5.3% (1.5-17.3)	6.0% (2.4-11.8)
Pericardiectomy (930)	63	2	1	11.8 (4.0-10.5)	11.1 (7.6-14.8)	1.6% (0.1-8.5)	4.6% (1.9-8.8)

(Continued)

TABLE 1. Continued

Procedure name (STS code)	No.	Morbidity		PLOS (d)		Any major complications	
		Category	Score	Average (IQR)	Model-based estimate (95% CrI)	Raw (95% CI)	Model (95% CrI)
Bidirectional cavopulmonary anastomosis (bidirectional Glenn) (1670)	2042	2	1	9.7 (4.0-9.0)	9.7 (9.1-10.3)	6.5% (5.5-7.6)	6.4% (5.4-7.6)
RVOT procedure (510)	1426	2	1	9.8 (4.0-8.0)	9.8 (8.8-10.9)	6.6% (5.4-8.0)	6.5% (5.4-7.8)
Fontan, TCPC, lateral tunnel, nonfenestrated (980)	140	2	1	10.0 (4.8-12.0)	10.1 (8.6-11.4)	6.4% (3.4-11.8)	6.4% (3.5-10.1)
1 1/2 ventricular repair (520)	20	2	1	8.7 (4.0-7.0)	9.5 (5.2-14.0)	10.0% (2.8-30.1)	7.1% (2.6-15.1)
Anomalous systemic venous connection repair (330)	200	2	1.1	10.1 (4.0-10.0)	10.2 (8.2-12.2)	6.5% (3.8-10.8)	6.5% (3.8-9.8)
Mitral stenosis, supravalvar mitral ring repair (840)	76	2	1.1	10.0 (4.0-8.0)	10.4 (5.8-14.7)	6.6% (2.8-14.5)	6.6% (3.0-11.6)
PDA closure, surgical (1330)	1507	2	1.1	14.1 (2.0-9.0)	13.8 (12.3-15.3)	2.6% (1.9-3.5)	2.9% (2.2-3.8)
Hemi-Fontan (1700)	245	2	1.1	11.2 (6.0-12.0)	11.2 (9.7-12.8)	5.7% (3.4-9.4)	6.0% (3.7-8.8)
Aortic aneurysm repair (1380)	264	2	1.1	9.3 (4.0-8.0)	9.6 (7.7-11.5)	8.7% (5.9-12.7)	8.2% (5.4-11.5)
PAPVC, scimitar, repair (270)	64	2	1.2	10.2 (4.0-8.0)	10.6 (6.9-14.3)	7.8% (3.4-17.0)	7.3% (3.5-12.9)
Bilateral bidirectional cavopulmonary anastomosis (bilateral bidirectional Glenn) (1690)	360	2	1.2	10.9 (5.0-10.0)	11.0 (9.6-12.5)	7.2% (5.0-10.4)	7.2% (4.9-9.9)
Systemic venous stenosis repair (340)	39	2	1.2	13.7 (4.5-10.0)	12.2 (6.1-18.1)	2.6% (0.1-13.2)	5.9% (2.2-11.9)
Fontan, TCPC, lateral tunnel, fenestrated (970)	822	2	1.2	11.4 (7.0-12.0)	11.4 (10.7-12.1)	7.1% (5.5-9.0)	7.1% (5.5-8.8)
AVR, homograft (700)	18	2	1.2	7.8 (4.0-6.0)	9.0 (5.5-12.8)	22.2% (9.0-45.2)	9.9% (4.0-19.4)
Aortopulmonary window repair (210)	80	2	1.2	12.5 (5.0-14.0)	12.3 (9.5-15.1)	5.0% (2.0-12.2)	6.2% (3.0-10.7)
TOF repair, ventriculotomy, transannular patch (370)	1721	2	1.2	11.1 (5.0-11.0)	11.1 (10.3-11.9)	7.7% (6.6-9.1)	7.7% (6.6-8.9)
AVR, mechanical (680)	214	2	1.2	9.2 (5.0-8.0)	9.4 (7.9-10.8)	10.7% (7.3-15.6)	9.7% (6.4-13.9)
Glenn (unidirectional cavopulmonary anastomosis) (unidirectional Glenn) (1680)	23	2	1.2	12.3 (5.0-8.0)	12.0 (6.4-17.3)	4.3% (0.2-21.0)	6.9% (2.4-14.2)
ASD partial closure (70)	14	2	1.3	22.2 (6.2-20.0)	12.6 (2.7-22.1)	0.0% (0.0-21.5)	6.9% (1.8-17.3)
PA, reconstruction (plasty), branch, peripheral (at or beyond the hilar bifurcation) (550)	163	2	1.3	11.0 (4.0-11.5)	11.2 (8.8-13.6)	9.2% (5.7-14.6)	8.6% (5.2-12.6)
PA, reconstruction (plasty), branch, central (within the hilar bifurcation) (540)	694	2	1.3	11.7 (4.0-11.0)	11.7 (10.5-13.1)	8.4% (6.5-10.7)	8.3% (6.3-10.4)
Conduit placement, RV to PA (610)	703	2	1.3	11.9 (4.0-10.0)	11.9 (10.4-13.5)	8.1% (6.3-10.4)	8.1% (6.3-10.1)
AVC (AVSD) repair, complete AVSD (170)	2332	2	1.3	12.8 (5.0-13.0)	12.8 (12.1-13.6)	7.0% (6.1-8.1)	7.1% (6.1-8.1)
Partial left ventriculectomy (LV volume reduction surgery) (Batista) (910)	33	2	1.3	15.7 (6.0-21.0)	14.0 (8.9-18.7)	0.0% (0.0-10.4)	5.9% (2.2-12.1)
LV to aorta tunnel repair (820)	17	2	1.3	12.5 (6.0-20.0)	12.5 (8.4-16.8)	5.9% (0.3-27.0)	7.7% (2.8-15.8)
Coarctation repair, subclavian flap (1230)	179	2	1.4	13.7 (5.0-12.0)	13.5 (11.1-15.9)	6.1% (3.5-10.7)	6.7% (4.0-10.1)
Ventricular septal fenestration (150)	24	2	1.4	17.8 (4.8-9.2)	13.2 (4.5-21.5)	4.2% (0.2-20.2)	7.4% (2.4-15.8)
Glenn + PA reconstruction (3300)	458	2	1.4	11.6 (5.0-11.0)	11.7 (10.1-13.2)	9.8% (7.4-12.9)	9.6% (7.2-12.3)
VSD, multiple, repair (130)	142	2	1.5	12.5 (5.0-11.0)	12.7 (9.2-16.0)	9.2% (5.4-15.0)	8.8% (5.2-13.4)
Shunt, ligation, and takedown (1630)	58	2	1.5	11.6 (4.0-11.8)	12.0 (8.9-15.2)	12.1% (6.0-22.9)	10.0% (5.1-16.9)
Aortic root replacement, mechanical (720)	84	2	1.5	9.8 (6.0-11.0)	10.2 (8.3-12.1)	15.5% (9.3-24.7)	12.1% (7.0-18.7)
Cardiac tumor resection (1760)	226	2	1.5	15.4 (4.0-15.8)	14.9 (12.2-17.8)	6.6% (4.1-10.7)	7.2% (4.6-10.4)
Coronary artery bypass (1300)	78	2	1.5	10.5 (4.0-8.0)	11.8 (7.6-16.0)	12.8% (7.1-22.0)	10.7% (5.7-17.1)
Aneurysm, ventricular, right, repair (1720)	80	2	1.5	12.9 (4.0-10.0)	13.1 (9.4-16.8)	10.0% (5.2-18.5)	9.4% (5.0-15.3)
Coarctation repair, patch aortoplasty (1240)	172	3	1.6	14.7 (4.0-14.0)	14.4 (11.0-17.7)	8.1% (4.9-13.2)	8.3% (5.2-12.2)
Pericardial drainage procedure (920)	245	3	1.6	20.0 (4.0-18.0)	17.4 (13.3-21.4)	3.7% (1.9-6.8)	5.2% (3.0-7.9)

(Continued)

TABLE 1. Continued

Procedure name (STS code)	No.	Category	Score	Morbidity	PLOS (d)	Any major complications	
				Average (IQR)	Model-based estimate (95% CrI)	Raw (95% CI)	Model (95% CrI)
Arrhythmia surgery ventricular, surgical ablation (1500)	40	3	1.6	13.6 (4.0-7.0)	13.7 (6.2-21.4)	10.0% (4.0-23.1)	9.6% (3.8-17.8)
Anomalous origin of coronary artery repair (1291)	287	3	1.6	13.5 (5.0-16.0)	13.5 (11.5-15.6)	10.1% (7.1-14.1)	9.9% (6.8-13.5)
Konno procedure (750)	134	3	1.6	11.5 (5.0-11.0)	12.0 (9.2-14.7)	13.4% (8.7-20.2)	11.8% (7.4-17.0)
Arrhythmia surgery atrial, surgical ablation (1490)	259	3	1.6	10.4 (5.0-11.0)	10.6 (9.2-12.0)	14.7% (10.9-19.5)	13.4% (9.6-17.7)
Fontan, TCPC, external conduit, nonfenestrated (1010)	883	3	1.6	13.5 (7.0-14.0)	13.5 (12.6-14.5)	10.4% (8.6-12.6)	10.3% (8.4-12.4)
Aortic root replacement, homograft (730)	99	3	1.7	14.3 (4.5-10.5)	14.2 (9.2-19.1)	10.1% (5.6-17.6)	9.8% (5.3-15.4)
Atrial septal fenestration (80)	18	3	1.7	12.3 (4.0-20.2)	13.3 (8.1-18.7)	16.7% (5.8-39.2)	11.2% (4.3-22.0)
Aneurysm, ventricular, left, repair (1730)	28	3	1.7	12.3 (5.8-16.0)	12.7 (9.7-15.8)	17.9% (7.9-35.6)	12.1% (5.5-22.1)
Arterial switch operation (1110)	1400	3	1.7	14.0 (8.0-15.0)	14.0 (13.2-14.7)	10.8% (9.3-12.5)	10.7% (9.2-12.4)
TOF repair, RV-PA conduit (380)	336	3	1.7	15.8 (6.0-15.0)	15.7 (13.3-18.0)	8.9% (6.3-12.5)	9.1% (6.4-12.2)
Pulmonary artery origin from ascending aorta (hemitruncus) repair (220)	45	3	1.7	15.6 (8.0-21.0)	15.2 (11.8-18.7)	8.9% (3.5-20.7)	9.6% (4.5-17.1)
Occlusion MAPCA(s) (450)	26	3	1.7	24.8 (6.0-21.8)	15.3 (6.8-24.0)	7.7% (2.1-24.1)	9.6% (3.5-19.4)
Congenitally corrected TGA repair, atrial switch and Rastelli (1060)	11	3	1.7	42.7 (10.5-23.0)	14.7 (4.4-25.1)	9.1% (0.5-37.7)	10.3% (2.7-24.3)
Fontan, atriopulmonary connection (950)	37	3	1.8	15.2 (7.0-14.0)	14.9 (10.0-19.8)	10.8% (4.3-24.7)	10.4% (4.6-18.8)
REV (1160)	10	3	1.8	21.5 (6.2-11.8)	14.8 (4.7-24.8)	10.0% (0.5-40.4)	10.6% (2.9-24.8)
Pulmonary embolectomy, acute pulmonary embolus (1802)	10	3	1.8	24.9 (3.0-30.0)	15.0 (5.1-24.9)	10.0% (0.5-40.4)	10.7% (2.8-24.9)
Ligation, pulmonary artery (1790)	12	3	1.8	13.7 (5.5-11.2)	14.4 (7.0-22.1)	16.7% (4.7-44.8)	11.6% (3.6-24.5)
Ebstein's repair (465)	117	3	1.9	13.8 (4.0-15.0)	14.1 (11.2-16.9)	13.7% (8.6-21.1)	12.5% (7.8-18.5)
Fontan, TCPC, external conduit, fenestrated (1000)	902	3	1.9	14.9 (7.0-16.0)	14.9 (14.0-15.9)	12.1% (10.1-14.4)	12.0% (10.0-14.1)
Senning (1130)	10	3	1.9	17.9 (8.2-24.8)	15.8 (8.6-22.7)	10.0% (0.5-40.4)	11.2% (3.6-24.2)
Pulmonary artery sling repair (1370)	80	3	1.9	23.8 (4.0-18.2)	17.1 (10.0-24.1)	8.8% (4.3-17.0)	9.8% (5.0-16.5)
DORV, intraventricular tunnel repair (1180)	443	3	2	14.8 (6.0-15.0)	14.9 (12.8-16.9)	13.1% (10.3-16.6)	12.8% (10.0-15.9)
Coarctation repair + VSD repair (1275)	254	3	2	18.2 (7.0-19.0)	17.8 (15.1-20.4)	9.8% (6.8-14.1)	10.2% (7.0-14.0)
Valve replacement, mitral (850)	526	3	2.1	13.9 (6.0-15.0)	14.1 (12.6-15.6)	16.0% (13.1-19.3)	15.5% (12.6-18.7)
Valve replacement, tricuspid (470)	134	3	2.1	12.3 (4.0-11.0)	13.3 (10.0-16.7)	18.7% (13.0-26.1)	16.4% (11.0-22.6)
Valve replacement, truncal valve (250)	11	3	2.1	17.5 (4.0-20.5)	16.2 (8.6-23.9)	18.2% (5.1-47.7)	13.2% (4.5-27.5)
Pulmonary venous stenosis repair (300)	286	3	2.1	18.0 (6.0-21.0)	17.7 (15.1-20.2)	11.5% (8.3-15.8)	11.7% (8.3-15.4)
VSD creation/enlargement (140)	84	3	2.1	12.8 (5.0-13.0)	13.6 (10.4-16.8)	20.2% (13.0-30.0)	16.7% (10.4-24.3)
Fontan revision or conversion (re-do Fontan) (1025)	107	3	2.2	14.0 (7.0-16.5)	14.4 (12.0-16.7)	18.7% (12.4-27.1)	16.4% (10.8-23.3)
Aortic arch repair (1280)	785	3	2.2	18.4 (6.0-20.0)	18.2 (16.5-20.1)	12.1% (10.0-14.6)	12.2% (10.0-14.6)
Pulmonary atresia VSD (including TOF, PA) repair (420)	218	3	2.2	17.6 (7.0-15.0)	17.4 (14.3-20.5)	13.8% (9.8-19.0)	13.5% (9.6-17.9)
Valvuloplasty, truncal valve (240)	14	3	2.3	16.6 (4.2-18.8)	16.6 (9.4-23.9)	21.4% (7.6-47.6)	14.8% (5.4-29.0)
TOF - AVC (AVSD) repair (390)	121	3	2.3	17.7 (7.0-19.0)	17.4 (13.8-20.7)	14.9% (9.6-22.3)	14.3% (9.2-20.2)
Ross-Konno procedure (760)	166	3	2.3	16.3 (6.0-14.8)	16.5 (13.5-19.6)	16.9% (11.9-23.3)	15.9% (11.1-21.4)
PA banding (1640)	973	4	2.5	25.1 (6.0-29.0)	24.2 (22.1-26.4)	8.2% (6.7-10.1)	8.6% (7.0-10.4)
Congenitally corrected TGA repair, atrial switch, and ASO (double switch) (1050)	27	4	2.5	17.0 (8.0-23.0)	17.2 (12.3-22.1)	22.2% (10.6-40.8)	16.8% (8.1-28.7)
Shunt, systemic to pulmonary, modified Blalock-Taussig shunt (1590)	1966	4	2.5	21.6 (7.0-23.0)	21.5 (20.1-22.8)	12.4% (11.0-13.9)	12.4% (11.0-13.9)
Shunt, systemic to pulmonary, central (from aorta or to main pulmonary artery) (1600)	482	4	2.5	24.0 (7.0-27.0)	22.8 (19.9-25.6)	10.8% (8.3-13.9)	11.3% (8.6-14.2)

(Continued)

TABLE 1. Continued

Procedure name (STS code)	No.	Category	Score	Morbidity		PLOS (d)		Any major complications	
				Average (IQR)	Model-based estimate (95% CrI)	Raw (95% CI)	Model (95% CrI)	Raw (95% CI)	Model (95% CrI)
ASO and VSD repair (1120)	670	4	2.6	17.9 (8.0-19.8)	18.0 (16.4-19.6)	17.8% (15.1-20.8)	17.5% (14.8-20.5)		
Rastelli (1150)	308	4	2.6	16.4 (6.0-18.0)	16.7 (14.2-19.1)	19.8% (15.7-24.6)	19.0% (14.9-23.4)		
TAPVC repair (280)	1327	4	2.6	19.2 (6.0-20.0)	19.1 (17.7-20.6)	16.4% (14.5-18.5)	16.4% (14.4-18.4)		
Unifocalization MAPCA(s) (440)	337	4	2.6	19.4 (6.0-21.0)	19.2 (16.4-22.0)	16.6% (13.0-21.0)	16.4% (12.8-20.4)		
Transplant, lung(s) (1410)	138	4	2.6	26.8 (13.0-28.0)	23.3 (18.8-27.7)	10.1% (6.1-16.3)	11.9% (7.5-17.1)		
Hypoplastic left heart syndrome biventricular repair (880)	38	4	2.7	33.1 (9.0-31.0)	20.7 (12.8-28.5)	15.8% (7.4-30.4)	16.2% (7.8-27.5)		
Aortic arch repair + VSD repair (1285)	353	4	2.8	21.7 (8.0-25.0)	21.4 (19.0-23.7)	15.9% (12.4-20.0)	15.9% (12.3-19.9)		
Aortic dissection repair (1390)	35	4	2.9	20.9 (5.0-28.5)	19.9 (14.1-25.5)	22.9% (12.1-39.0)	19.5% (10.5-31.2)		
ASD creation/enlargement (60)	100	4	2.9	24.0 (5.0-26.2)	21.5 (16.3-26.6)	18.0% (11.7-26.7)	18.0% (11.6-25.3)		
Transplant, heart (890)	821	4	3	23.7 (10.0-27.0)	23.4 (21.8-25.1)	16.6% (14.2-19.3)	16.6% (14.2-19.2)		
Mustard (1140)	11	4	3	19.6 (6.0-13.5)	19.5 (11.1-27.7)	36.4% (15.2-64.6)	21.5% (8.1-40.7)		
Pulmonary atresia - VSD - MAPCA (pseudotruncus) repair (430)	98	4	3.1	21.6 (8.0-22.0)	20.8 (16.2-25.3)	22.4% (15.3-31.7)	21.1% (14.1-29.1)		
TOF - absent pulmonary valve repair (400)	148	4	3.1	29.7 (6.0-27.8)	24.1 (18.5-29.7)	16.9% (11.7-23.7)	17.6% (12.2-23.6)		
Damus-Kaye-Stansel procedure (creation of aortopulmonary anastomosis without arch reconstruction) (1660)	313	4	3.3	21.3 (7.0-24.0)	21.2 (18.1-24.3)	23.3% (19.0-28.3)	22.9% (18.6-27.6)		
Truncus arteriosus repair (230)	439	4	3.4	26.1 (9.0-27.0)	25.1 (22.1-27.9)	20.0% (16.6-24.0)	20.2% (16.6-23.9)		
Interrupted aortic arch repair (1320)	476	4	3.5	28.8 (10.0-35.2)	27.4 (24.6-30.4)	18.7% (15.5-22.4)	19.0% (15.7-22.7)		
Transplant, heart and lung (900)	10	4	3.8	40.9 (17.0-53.8)	23.3 (13.7-33.0)	40.0% (16.8-68.7)	26.8% (10.0-48.2)		
Arterial switch procedure and VSD repair + aortic arch repair (1125)	119	5	4.1	23.6 (11.5-29.5)	23.5 (20.3-26.6)	32.8% (25.0-41.6)	31.0% (23.3-39.2)		
Arterial switch procedure + aortic arch repair (1123)	21	5	4.3	26.0 (10.0-31.0)	24.4 (17.9-30.7)	42.9% (24.5-63.5)	32.7% (18.0-50.3)		
Norwood procedure (870)	2273	5	4.9	34.6 (14.0-42.0)	34.1 (32.6-35.7)	29.6% (27.7-31.5)	29.7% (27.9-31.6)		
Truncus + IAA repair (3200)	34	5	5	32.6 (13.0-36.0)	27.3 (20.3-34.2)	44.1% (28.9-60.5)	38.2% (24.8-52.9)		

PLOS, Postoperative length of stay; STS, Society of Thoracic Surgeons; IQR, interquartile range; CrI, credible interval; CI, confidence interval; ASD, atrial septal defect; ICD, implantable cardioverter defibrillator; AICD, automatic implantable cardioverter defibrillator; PAPVC, partial anomalous pulmonary venous connection; AVSD, atrioventricular septal defect; PAVSD, partial atrioventricular septal defect; DCRV, double-chambered right ventricle; VSD, ventricular septal defect; LV, left ventricle; PA, pulmonary artery; AVC, atrioventricular canal; TOF, tetralogy of Fallot; PFO, patent foramen ovale; AVR, aortic valve replacement; RVOT, right ventricular outflow tract; TCPC, total cavopulmonary connection; PDA, patent ductus arteriosus; RV, right ventricle; AVC, atrioventricular canal; MAPCA, major aortopulmonary collateral artery; TGA, transposition of the great arteries; REV, réparation à l'étage ventriculaire; DORV, double-outlet right ventricle; ASO, arterial switch operation; TAPVC, total anomalous pulmonary venous connection; IAA, interrupted aortic arch.

Third, although morbidity was calculated as an equally weighted combination of complication rate + average PLOS, strong consideration was given to an alternative composite consisting of the rate of major complications and the average time on ventilator. Rank correlation between these 2 composite morbidity measures was 0.93, suggesting that the 2 methods tend to give similar, but not completely identical, results. The version using PLOS was preferred in part because PLOS was collected with high (>99.9%) completeness, whereas ventilation time was more than 15% missing. Moreover, during the time period of this study, the STS definition of time on ventilator only included the time until the first extubation and did not include the additional time on ventilator for patients who were subsequently reintubated.

Fourth, we assessed the reliability (ie, statistical precision; see “Materials and Methods”) of the various measures that were used for ranking procedures in this study. For

major complications, average PLOS, and the composite morbidity measure, the estimated reliability values were 0.80 (95% CrI, 0.71-0.87), 0.88 (95% CrI, 0.82-0.92), and 0.90 (95% CrI, 0.85-0.94), respectively. Thus, reliability was greatest for composite morbidity, which was the basis of the proposed Morbidity Score and Categories. The estimated reliability of composite morbidity increased to 0.95 (95% CrI, 0.92-0.97) when considering only procedures with at least 30 cases (N = 115 procedures) and to 0.99 (95% CrI, 0.98-0.99) when considering only procedures with at least 200 cases (N = 67 procedures).

Finally, we assessed the degree of association between the proposed Morbidity Score and the existing STS-EACTS Mortality Score. A weak association would suggest poor content validity because, conceptually, we know that morbidity and mortality are closely related. On the other hand, a perfect association would suggest that the morbidity score is redundant with mortality and thus is unneeded.

TABLE 2. Major complications

Complication description (STS code*)	No. (%) of events†	Mortality N (%)	Rank correlation with “major complication”
Postoperative acute renal failure requiring temporary or permanent dialysis (220 or 230)	705 (1.1%)	396 (56.2%)	0.59
Postoperative neurologic deficit <i>persisting at discharge</i> (320)	500 (0.8%)	152 (30.4%)	0.45
Postoperative AV block requiring permanent pacemaker (60)	593 (0.9%)	28 (4.7%)	0.37
Postoperative mechanical circulatory support (IABP, VAD, ECMO, or CPS) (40)	1110 (1.8%)	617 (55.6%)	0.54
Phrenic nerve injury/paralyzed diaphragm (300)	578 (0.9%)	35 (6.1%)	0.40
Unplanned reoperation (20 or 240)	2942 (4.7%)	636 (21.6%)	0.79
Major complication (defined as any 1 or more of the above)	5059 (8.0%)	1187 (23.5%)	N/A

STS, Society of Thoracic Surgeons; AV, atrioventricular; IABP, intra-aortic balloon pump; VAD, ventricular assist device; ECMO, extracorporeal membrane oxygenation; CPS, cardiopulmonary support. *Complication codes in the STSCHSD Data Collection Form, Version 2.50.¹⁸ †Denominator is 62,851 operations.

To address these issues, the proposed Morbidity Score and the STS-EACTS Mortality Score were plotted and compared. As shown in Figure 3, they are closely related (rank correlation = 0.79), but far from being redundant.

Descriptive characteristics of the 5 Morbidity categories are shown in Table 3. The association between Morbidity Categories and Mortality Categories is summarized in Table 4. The Morbidity and Mortality Categories were identical for 74 procedures, differed by 1 or fewer positions for 135 procedures, and differed by 2 or fewer positions for 139 procedures. One procedure (pulmonary artery debanding) was in category 4 for mortality but category 1 for morbidity.

DISCUSSION

Measuring morbidity is a challenging but important element of outcomes reporting and quality assessment.^{7,8} Morbidity is a major determinant of health status after

surgery and of hospital cost.^{3,4,8} The importance of developing a morbidity metric was articulated in 2004 by Phillippe Kolh,⁹ who described quantitation of morbidity in cardiac surgery as follows: “Being more frequent than mortality, it could carry more information and be measured in terms of postoperative complications and length of hospital stay.... Furthermore, because of the heterogeneity of morbidity events, future scoring systems should probably generate separate predictions for mortality and major morbidity events.” In this report, we introduce an empirically derived tool that estimates the relative risk of morbidity associated with congenital heart surgery procedures on the basis of elements of both complications and PLOS.

Formal risk modeling using logistic regression is practical for common “adult cardiac procedures,” such as coronary artery bypass grafting and valve replacement. No operation for congenital heart disease is performed in

FIGURE 1. Unadjusted average PLOS (days) and unadjusted rate of major complications (percentage) are measured on the horizontal and vertical axes, respectively. Squares represent the 140 procedure types with n greater than 10. PLOS, Postoperative length of stay.

FIGURE 2. The proportions (%) of Any Major Complication and of Any Major Complication or Mortality are measured on the horizontal and vertical axes, respectively. The squares represent the 140 procedure types with n greater than 10.

FIGURE 3. The relationship between the STS-EACTS Mortality score² and the STS Morbidity Score. Squares represent the 140 procedure types with *n* greater than 10.

numbers comparable to coronary artery bypass grafting. The diverse spectrum of distinct procedures is reflected by 140 procedure types in this study. Bayesian modeling is a particularly appropriate tool in this setting where denominators may be small. Thus, the product of this analysis is not a series of procedure-specific risk models, but rather a metric of procedure-based estimates of morbidity that can be used to describe case mix.

Composite Development

At the outset, we appreciated the importance of including both a complications element and a resource utilization element in a morbidity metric. We felt obligated, however, to demonstrate that use of either alone would be inadequate, that is, a model that assumed a direct 1-to-1 relationship between major complications and PLOS would not fit the data as well, and therefore be an incomplete and less informative morbidity metric.

Resource utilization variables used in previous analyses were considered. Analysis based on inclusion of ventilation time is described earlier in this article. Previous analyses from individual institutions have included length of intensive care unit (ICU) stay.^{10,11} It is less useful at

TABLE 3. Summary of morbidity categories

	1	2	3	4	5
No. of procedures	36	43	36	21	4
Aggregate average PLOS (d)	6.3	11.3	15.2	22.3	34.0
Rate of major complications	3.2%	6.5%	11.9%	15.2%	30.0%

PLOS, Postoperative length of stay.

TABLE 4. Association between morbidity categories and mortality categories

Morbidity categories	Mortality categories				
	1	2	3	4	5
1	21	13	1	1	0
2	5	26	10	2	0
3	0	12	10	13	1
4	0	0	3	15	3
5	0	0	0	2	2

a multi-institutional level because of lack of a uniform definition of ICU, and because some institutions keep postoperative patients in an ICU environment until discharge. Cost is another measure of resource utilization that may be associated with morbidity. Cost data are not included in STS registries; furthermore, true cost data can be difficult to estimate.

Individual elements of the complication end point were considered on the basis of their potential impact on patients' health status, including durable, long-lasting effects. We acknowledge that validated data describing relationships between some individual complications and late health status are not readily available. Some complications that are not included, such as Postoperative Cardiac Arrest, have been evaluated in other studies and shown to potentially be associated with mortality.¹² Despite the fact that all complication codes in the STSCHSD have corresponding definitions since 2006, the coding of complications such as Postoperative Cardiac Arrest may still be subject to a degree of interpretation, and thus potentially variable ascertainment, in contrast to complications included in our list, such as postoperative mechanical circulatory support. Still other complications, such as sternal dehiscence and mediastinitis, are not included in this metric because they result in unplanned reoperations, and thus are accounted for in the composite. A variety of other complications that are not counted in the major complications end point are likely to be reflected in increased PLOS. Although some have argued that mortality is the "ultimate morbidity," the decision to exclude in-hospital mortality from the morbidity metric was deliberate, based on several principles. First, the Morbidity Scores and Categories are designed to be used in conjunction with the STS-EACTS Mortality Scores and Categories. Second, analyses of potential associations between morbidity and mortality require the use of separate metrics for each.^{9,13}

Our analysis confirms that morbidity and mortality indices provide related but not redundant information. Outcomes assessment should include measures of both, as suggested by Kohl.⁹ The decision to include patients who died before discharge may seem obvious, but it was debated. An alternative strategy that eliminates from analysis those who died before discharge would have resulted in an incomplete and potentially misleading

picture of morbidity, and would have compromised any possible efforts to explore relationships between morbidity and mortality. The concept of “Failure to Rescue,” that is, probability of death following the occurrence of a complication or adverse event, is emerging as a potentially important tool for measuring performance and directing quality improvement initiatives.^{14,15} Quantitative estimation of morbidity only among hospital survivors would overlook the importance of this concept.

Widely used systems for stratifying risk or complexity of congenital heart surgery procedures have focused entirely on in-hospital mortality¹⁶ or included a morbidity element that was based on expert opinion rather than objective data.¹⁷ The Morbidity Score and Categories presented in the current study are complementary to the empirically based STS-EACTS Mortality Score and Categories² and were derived using data in the largest congenital heart surgery registry.

Study Limitations

Despite the advantages of an empirically based tool for analyzing morbidity, this study has important limitations. The analysis focuses on estimation of morbidity at the procedure level. We did not address methods of incorporating these procedural variables into statistical models for performing inter-institutional outcomes comparisons. Nor does our methodology address the appropriateness and timing of individual procedures in relation to overall disease management or include consideration of patient-specific risk factors. Second, despite a large database size, it is possible that patients and data in the STSCHSD are not entirely representative of other populations. Third, several individual procedures had small sample sizes, and the true morbidity associated with these procedures may have been estimated with error. We attempted to minimize this error by using a composite measure that combines statistical information from several related end points into a single end point. Fourth, both the occurrence and the impact of morbidity extend beyond the duration of the surgical hospital admission. The nature of the STSCHSD precludes inclusion of complications recognized or therapeutic interventions occurring after discharge from the “surgical admission.”¹⁸ A “long-term database” will ultimately be needed to achieve a more comprehensive estimation of morbidity associated with surgery for congenital heart disease. This study represents an important first step and acknowledges the need for a quantitative morbidity metric and the mandate that it be empirically derived.

CONCLUSIONS

The STS Morbidity Score and Categories is a tool for analyzing morbidity associated with operations for congenital heart disease and for grouping procedures with similar

empirically estimated risk of morbidity. Together with the STS-EACTS Mortality Score and Categories, this tool enhances our ability to accurately characterize case mix. It should add a new dimension and precision to outcome assessments and may provide important information to guide quality-improvement initiatives.

References

- Jacobs JP, O'Brien SM, Pasquali SK, Jacobs ML, Lacour-Gayet FG, Tchervenkov CI, et al. Variation in outcomes for benchmark operations: an analysis of The Society of Thoracic Surgeons' Congenital Heart Surgery Database. *Ann Thorac Surg*. 2011;92:2184-92.
- O'Brien SM, Clarke DR, Jacobs JP, Jacobs ML, Lacour-Gayet FG, Pizarro C, et al. An empirically based tool for analyzing mortality associated with congenital heart surgery. *J Thorac Cardiovasc Surg*. 2009;138:1139-53.
- Mahle WT, Wernovsky G. Long-term developmental outcome of children with complex congenital heart disease. *Clin Perinatol*. 2001;28:235-47.
- Pasquali SK, Sun JL, d'Almada P, Jaquiss RD, Lodge AJ, Miller N, et al. Center variation in hospital costs for patients undergoing congenital heart surgery. *Circ Cardiovasc Qual Outcomes*. 2011;4:306-12.
- Jacobs ML, Jacobs JP, Franklin RCG, Mavroudis C, Lacour-Gayet F, Tchervenkov CI, et al. Databases for assessing the outcomes of the treatment of patients with congenital and paediatric cardiac disease—the perspective of cardiac surgery. *Cardiol Young*. 2008;18(Suppl 2):101-15.
- Ferreira-González I, Permanyer-Miralda G, Busse JW, Bryant DM, Montori VM, Alonso-Coello P, et al. Methodologic discussions for using and interpreting composite endpoints are limited, but still identify major concerns. *J Clin Epidemiol*. 2007;60:651-7.
- Jacobs JP, Jacobs ML, Mavroudis C, Maruszewski B, Tchervenkov CI, Lacour-Gayet FG, et al. What is operative morbidity? Defining complications in a surgical registry database: a report from the STS Congenital Database Task Force and the Joint EACTS-STS Congenital Database Committee. *Ann Thorac Surg*. 2007;84:1416-21.
- Benavidez OJ, Connor JA, Gauvreau K, Jenkins KJ. The contribution of complications to high resource utilization during congenital heart surgery admissions. *Congenit Heart Dis*. 2007;2:319-26.
- Kolh P. Importance of risk stratification models in cardiac surgery. *Eur Heart J*. 2006;27:768-9.
- Pagowska-Klimek I, Pychynska-Pokorska M, Krajewski W, Moll JJ. Predictors of long intensive care unit stay following cardiac surgery in children. *Eur J Cardiothorac Surg*. 2011;40:179-84.
- Bojan M, Gerelli S, Gioanni S, Pouard P, Vouhé P. Evaluation of a new tool for morbidity assessment in congenital cardiac surgery. *Ann Thorac Surg*. 2011;92:2200-4.
- Benavidez OJ, Gauvreau K, Del Nido P, Bacha E, Jenkins KJ. Complications and risk factors for mortality during congenital heart surgery admissions. *Ann Thorac Surg*. 2007;84:147-55.
- Ghaferi AA, Birkmeyer JD, Dimick JB. Complications, failure to rescue, and mortality with major inpatient surgery in Medicare patients. *Ann Surg*. 2009;250:1029-34.
- Silber JH, Romano PS, Rosen AK, Wang Y, Even-Shoshan O, Volpp KG. Failure-to-rescue: comparing definitions to measure quality of care. *Med Care*. 2007;45:918-25.
- Pasquali SK, Li JS, Burstein DS, Sheng S, O'Brien SM, Jacobs ML, et al. The Association of Center Volume with Mortality and Complications in Pediatric Heart Surgery. *Pediatrics*. 2012;129:e370-6.
- 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.
- Lacour-Gayet F, Clarke D, Jacobs J. The Aristotle score: a complexity-adjusted method to evaluate surgical results. *Eur J Cardiothorac Surg*. 2004;25:911-24.
- The Society of Thoracic Surgeons Congenital Cardiac Surgery Database Data Collection Form. Available at: http://www.sts.org/sites/default/files/documents/pdf/DataCollectionForm250_07102006_Nonannotated.pdf. Accessed December 14, 2011.
- Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput*. 2000;10:325-37.

APPENDIX 1. STATISTICAL APPENDIX

Statistical Model

A bivariate hierarchical model with normally distributed random effects was used to estimate the distribution of procedure-specific probabilities of major complication and average PLOS. For the i -th patient undergoing the j -th procedure, let y_{ji} denote the occurrence of major complication (0 = no, 1 = yes) and x_{ji} denote the patient's PLOS. The model was as follows:

$$y_{ji} \stackrel{iid}{\sim} \text{Bernoulli}(\pi_j); \quad x_{ji} \stackrel{iid}{\sim} \text{Normal}(\mu_j, \sigma_j^2);$$

$$\pi_j = 1 / (1 + \exp[-\eta_j])$$

$$(\eta_j, \mu_j) \sim \text{BivariateNormal}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

where π_j denotes the unknown theoretic probability of major complication for the j -th procedure; μ_j and σ_j^2 denote the unknown mean and variance of PLOS for the j -th procedure; and $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ denote unknown parameters of the assumed bivariate normal random effects distribution.

Estimation

Model parameters were estimated in a Bayesian statistical framework by specifying a prior probability distribution for unknown parameters $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, and $\sigma_1^2, \dots, \sigma_{148}^2$. Because our prior knowledge was limited, we specified a vague proper prior distribution that consisted of independent normal distributions for the elements of $\boldsymbol{\mu}$, independent inverse Gamma distributions for the σ_j^2 s, and an inverse Wishart distribution for $\boldsymbol{\Sigma}$. Posterior means and CrIs were calculated using Markov Chain Monte Carlo (MCMC) simulations as implemented in WinBUGS version 1.4 software (Medical Research Council Biostatistics Unit, Cambridge, UK, and the Imperial College School of Medicine at St Mary's, London, UK).¹⁹ Posterior summaries were based on 4000 sets of simulated parameter values generated after a long burn-in period to ensure convergence.

Composite Morbidity

The overall composite morbidity of the j -th procedure was defined as follows:

$$\theta_j = \frac{\pi_j}{s_\pi} + \frac{\mu_j}{s_\mu}$$

where

$$s_\pi = \sqrt{\frac{1}{148} \sum_{j=1}^{148} \left(\pi_j - \frac{1}{148} \sum_{h=1}^{148} \pi_h \right)^2} \quad \text{and}$$

$$s_\mu = \sqrt{\frac{1}{148} \sum_{j=1}^{148} \left(\mu_j - \frac{1}{148} \sum_{h=1}^{148} \mu_h \right)^2}.$$

The parameter θ_j was estimated as $\hat{\theta}_j = \sum_{l=1}^{4000} \theta_j^{(l)} / 4000$,

where $\theta_j^{(l)}$ denotes the simulated value of θ_j at the l -th iteration of the MCMC procedure. A 95% Bayesian CrI for θ_j was obtained by calculating the 100th lowest and 100th highest values of $\theta_{(l)}$ across the 4000 simulated values.

Creation of Categories

To create internally homogeneous categories, procedures were first sorted in order of increasing estimated morbidity and then grouped by choosing cutpoints that were optimal according to a least squares variance criterion, as described later. We first sorted procedures so that $\hat{\theta}_1 < \hat{\theta}_2 < \dots < \hat{\theta}_{148}$. Let K denote the number of categories and let $\mathbf{c}_K = c_1 < c_2 < \dots < c_{K-1}$ denote a set of category cutpoints that partition the categories into K groups. The symbol $\mathbf{c}_K$ denotes a number between 1 and 148 and represents the index of the highest-morbidity procedure in the k -th category. Also, define $c_0 = 0$ and $c_K = 149$. For any particular choice of K and $\mathbf{c}_K$, within-category homogeneity was measured by the weighted sum-of-squares criterion:

$$\text{WSS}(\mathbf{c}_K; \boldsymbol{\theta}) = \sum_{k=1}^K \sum_{j=c_{k-1}+1}^{c_k} n_j (\theta_j - \bar{\theta}_k)^2$$

where n_j is the number of patients in the denominator for the

j -th procedure and $\bar{\theta}_k = \sum_{j=c_{k-1}+1}^{c_k} n_j \theta_j / \sum_{j=c_{k-1}+1}^{c_k} n_j$ is the average

morbidity of all procedures in the k -th category weighted by their respective sample sizes. If the θ_j were known instead of unknown, then the "optimal" cutpoints could (in theory) be determined by enumerating all possible choices for the c_k and choosing the one that minimizes the WSS. Because the θ_j are unknown, we instead chose cutpoints that minimize the estimated value of WSS ($\mathbf{c}_K$). Specifically, we chose cutpoints that minimize the posterior mean

$$\widehat{\text{WSS}}(\mathbf{c}_K) = \frac{1}{4000} \sum_{l=1}^{4000} \text{WSS}(\mathbf{c}_K; \boldsymbol{\theta}^{(l)})$$

where $\boldsymbol{\theta}^{(l)}$ is the value of $\boldsymbol{\theta} = (\theta_1, \dots, \theta_{148})$ on the l -th iteration of the MCMC procedure. An unpublished dynamic programming algorithm was used to determine the set of cutpoints that made this quantity a minimum. The WSS criterion gets smaller as K , the number of categories, increases. The value $K = 5$ was selected for consistency with the published STS-EACTS mortality categories.

Estimation of Reliability

Reliability is conventionally defined as the proportion of variation in a measure that is due to true between-unit

differences (ie, signal) as opposed to random statistical fluctuations (ie, noise). Equivalently, it is the squared correlation between a measurement and the true value. Accordingly, reliability was defined as the square of the Pearson correlation coefficient between the set of procedure-specific estimates $\hat{\theta}_1, \dots, \hat{\theta}_{148}$ and the corresponding unknown true values $\theta_1, \dots, \theta_{148}$, that is:

$$\rho^2 = \frac{\left[\sum_{j=1}^{148} \left(\hat{\theta}_j - \frac{1}{148} \sum_{h=1}^{148} \hat{\theta}_h \right) \left(\theta_j - \frac{1}{148} \sum_{h=1}^{148} \theta_h \right) \right]^2}{\left[\sum_{j=1}^{148} \left(\hat{\theta}_j - \frac{1}{148} \sum_{h=1}^{148} \hat{\theta}_h \right)^2 \right] \left[\sum_{j=1}^{148} \left(\theta_j - \frac{1}{148} \sum_{h=1}^{148} \theta_h \right)^2 \right]}$$

The quantity ρ^2 was estimated by its posterior mean, namely

$$\hat{\rho}^2 = \frac{1}{4000} \sum_{l=1}^{4000} \rho_{(l)}^2$$

where

$$\rho_{(l)}^2 = \frac{\left[\sum_{j=1}^{148} \left(\hat{\theta}_j - \frac{1}{148} \sum_{h=1}^{148} \hat{\theta}_h \right) \left(\theta_j^{(l)} - \frac{1}{148} \sum_{h=1}^{148} \theta_h^{(l)} \right) \right]^2}{\left[\sum_{j=1}^{148} \left(\hat{\theta}_j - \frac{1}{148} \sum_{h=1}^{148} \hat{\theta}_h \right)^2 \right] \left[\sum_{j=1}^{148} \left(\theta_j^{(l)} - \frac{1}{148} \sum_{h=1}^{148} \theta_h^{(l)} \right)^2 \right]}$$

with $\theta_j^{(l)}$ denoting the value of θ_j at the l -th MCMC iteration and $\hat{\theta}_j = \sum_{l=1}^{4000} \theta_j^{(l)} / 4000$ denoting the posterior mean of θ_j . A 95% CrI for ρ^2 was obtained by calculating the 100th smallest and 100th largest values of $\rho_{(l)}^2$ across the 4000 MCMC iterations. Analogous calculations were used to estimate reliability for subsets of procedures with at least 10, 30, or 200 cases. An identical approach was also used for estimating the reliability of procedure-specific probability parameters $\pi_1, \dots, \pi_{148}$ and mean parameters $\mu_1, \dots, \mu_{148}$.

APPENDIX 2. CLASSIFICATION OF MULTIPLE-PROCEDURE OPERATIONS

Several procedures listed in Table 1 are actually combinations of 2 or more procedures. These combinations were previously identified during development of the STS-EACTS Mortality Score and Categories.² They occur

frequently in the STS and EACTS databases, and the complexity of the combination is regarded as being different from the complexity of the component procedures when performed in isolation. For all other operations involving combinations of procedures, the operation was classified according to the most technically complex procedure, as determined by the difficulty component of the 2007 update of the Aristotle Basic Complexity score. The Aristotle Basic Complexity score contains some ties and is not defined for 3 of the procedures listed in Table 1. To deal with undefined or tied Aristotle scores, 6 of the study authors independently ranked the difficulty of each procedure. Undefined or tied Aristotle scores were adjudicated by assigning the operation to the procedure with the highest average ranking determined by the 6 graders. The difficulty rankings were published together with the STS-EACTS Mortality Score and Categories.² The identical methodology and rankings were used to classify multiple-procedure operations during development of the STS Morbidity Score and categories.

APPENDIX 3. RESCALING OF INDIVIDUAL OUTCOME MEASURES (AVERAGE RISK OF MAJOR COMPLICATIONS AND AVERAGE POSTOPERATIVE LENGTH OF STAY)

Procedure-specific complication rates are percentages measured on a scale from 0 to 100. Procedure-specific average PLOS is measured in days ranging from 0 to infinity. These are different measurement scales. We rescaled these so that the new scales would have approximately the same standard deviation. This guarantees that approximately half of the variation of the composite measure will be attributable to complications and half to PLOS. If we did not rescale them, then the amount of variation contributed by each item would be dependent on the scale we used for measuring it. For example, we would get different results depending on whether complication rates were expressed as percentages or proportions, or whether PLOS was measured in days, weeks, or months. Rescaling makes it possible to have a composite measure in which dominance of a single element is avoided.

For complications, first, we calculated each procedure's complication rate. Next, we calculated the standard deviation of the set of procedure-specific complication rates. Finally, to obtain a rescaled complication rate, we divided each of the original 140 complication rates by their common standard deviation. The same process was used for standardizing average PLOS.

FIGURE E1. The relationship between the number of categories and the degree of within-category homogeneity. Within-category homogeneity is defined as 1 minus the ratio of within-category variance to total variance of procedure-specific morbidity estimates.