The Society of Thoracic Surgeons General Thoracic Surgery Database Update on Outcomes and Quality

Christopher W. Seder, MD, Cameron D. Wright, MD, Andrew C. Chang, MD, Jane M. Han, MSW, Donna McDonald, MPH, RN, and Benjamin D. Kozower, MD, MPH

Rush University Medical Center, Chicago, Illinois; Massachusetts General Hospital, Boston, Massachusetts; University of Michigan, Ann Arbor, Michigan; The Society of Thoracic Surgeons, Chicago, Illinois; and University of Virginia, Charlottesville, Virginia

The Society of Thoracic Surgeons General Thoracic Surgery Database (STS GTSD) is a voluntary effort that provides participants with semiannual, risk-adjusted performance reports that facilitate quality improvement and comparison of institutional outcomes with national benchmarks. With improved participation in the STS GTSD, increasingly meaningful analyses are available. This year, risk models for lobectomy and esophagectomy for cancer were updated. In addition, we developed the first composite quality measure for general thoracic surgery, a composite measure for lobectomy for lung cancer. Furthermore, international collaboration with the European Society of Thoracic Surgery has facilitated our understanding of variation between American and European treatment patterns, providing a foundation for future quality improvement initiatives. This article summarizes current aggregate national outcomes in general thoracic surgery and reviews related activities in the areas of quality measurement, performance improvement, and transparency from the STS GTSD over the past 12 months.

(Ann Thorac Surg 2016;101:1646–54)

The Society of Thoracic Surgeons General Thoracic Surgery Database (STS GTSD) is the world's largest clinical thoracic surgical database. In 1989, the STS National Database was established as a means to measure, quantify, and compare performance in cardiothoracic surgery. Although data on general thoracic surgical procedures have been included since 1999, the STS GTSD was not established as an independent entity until 2002 [1].

The STS GTSD is a voluntary effort that provides participants with risk-adjusted semiannual performance reports that allow comparison of institutional outcomes against national benchmarks. This year, a composite quality measure for lobectomy for lung cancer was developed, and risk models for lung resection and esophagectomy for cancer were updated. In addition, the STS GTSD Task Force is working to have additional quality metrics receive National Quality Forum (NQF) endorsement, the gold standard for health care quality.

International collaboration with the European Society of Thoracic Surgery (ESTS) has also been a recent focus of the STS GTSD. Efforts have included harmonization of terminology and definitions between the two databases, and facilitating an intersocietal comparison of pulmonary resection practices and short-term outcomes [2, 3]. This review summarizes all national aggregate outcome, quality measurement, and improvement initiatives from the STS GTSD over the past 12 months.

Address correspondence to Dr Seder, 1725 W Harrison St, Ste 774, Chicago, IL 60612; email: christopher_w_seder@rush.edu.

Database Participation and Aggregate Outcomes

Similar to the STS adult cardiac and congenital heart surgery databases, participation in the STS GTSD has increased each year since its inception in 2002, with 261 participants contributing to the data harvest in 2014 (Fig 1) [4]. As of November 9, 2015, the STS GTSD included data from 892 surgeons (868 thoracic and 24 general surgeons) at 274 US institutions in 44 states, for a total of 438,603 operations (Fig 2). In addition, nine surgeons in the United Arab Emirates have agreed to contribute data; however, no international data have been harvested to date. Centers in two additional countries have contracts pending and will likely contribute data soon. Among other benefits of participating in the STS GTSD, performing a quality project using the STS database meets the criteria for the American Board of Thoracic Surgery maintenance of certification part IV. Table 1 summarizes the STS GTSD aggregate outcomes for lobectomy, pneumonectomy, and esophagectomy.

Harmonization of Definitions With ESTS

The STS GTSD and ESTS database were independently created for the purpose of objectively measuring processes and outcomes with the goal of improving the quality of thoracic surgical care. However, direct comparison of the databases was difficult owing to significant differences in variables included and data definitions. To facilitate efficient and accurate intersocietal collaboration, members of the STS GTSD and ESTS Database Task Forces worked to standardize terminology and harmonize definitions. The result was a comprehensive review

0003-4975/$36.00
http://dx.doi.org/10.1016/j.athoracsur.2016.02.099

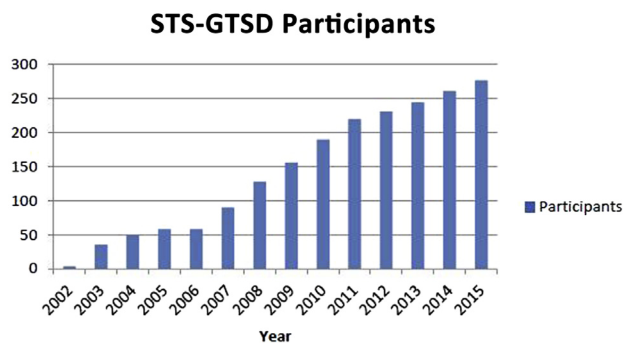

Fig 1. The Society of Thoracic Surgeons General Thoracic Surgery Database (STS-GTSD) participation, 2002 to 2015.

outlining the newly defined common terminology, history, and aims of the two databases [2]. In addition, participation, methods of data collection, important publications, and the future visions of the societal databases were reported. A list of agreed upon definitions that were not present in both databases was also included; use of these data fields in future versions of the databases will depend on individual database committee decisions.

The first project comparing the STS GTSD and ESTS database was examining the variation in pulmonary resection practices and short-term outcomes [3]. Between 2010 and 2013, 78,212 patients underwent pulmonary resection and their data were captured in the STS (n = 47,539) or ESTS databases (n = 30,673). This analysis demonstrated that a greater percentage of patients in the STS GTSD were female, had undergone previous cardiothoracic surgery, and had preoperative thoracic radiation. In addition, patients in the STS GTSD had higher American Society of Anesthesiologists (ASA) classifications and Zubrod scores, consistent with the greater observed incidence of obesity, hypertension, congestive heart failure, and coronary artery disease. More patients in the STS GTSD underwent thoracoscopic procedures (62.5% versus 21.8%, $p < 0.001$) and sublobar resections (43.3% versus 31.1%, $p < 0.001$), despite a similar proportion of surgeries performed for primary lung cancer (Table 2). Although the distribution of pathologic T stage was similar between databases, fewer patients had pathologic N2 or M1 disease at the time of surgery in the STS GTSD.

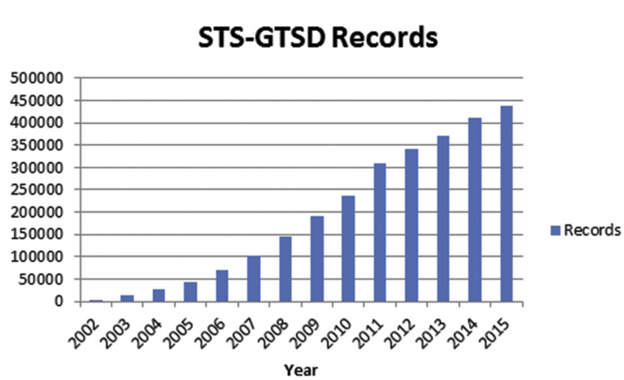

Fig 2. Number of records included in The Society of Thoracic Surgeons General Thoracic Surgery Database (STS-GTSD), 2002 to 2015.

Comparing outcomes, patients in the STS GTSD had a shorter hospital length of stay, despite being reintubated more often and having more frequent atrial arrhythmias. Thirty-day mortality was higher in the STS GTSD for patients who underwent wedge resection (1.9% versus 0.1%, $p < 0.05$). However, STS GTSD mortality was lower for lobectomy (1.4% versus 2.6%, $p < 0.05$) and pneumonectomy (4.9% versus 7.3%, $p < 0.05$) compared with the ESTS GTSD.

It must be noted that data auditing is not randomly or routinely performed in the ESTS. Only general thoracic surgical units eligible for institutional accreditation currently undergo data auditing in Europe. That adds some uncertainty regarding the quality of the ESTS data and potentially decreases the validity of any comparisons with the audited and highly accurate STS GTSD data.

Although some discrepancies due to regional influences are expected, variation in outcomes suggests the opportunity for quality improvement initiatives. These data provide the basis for further investigation with stage-specific, risk-adjusted analyses to better understand the causes of variation in specific patient populations. With the recent efforts to link the STS GTSD with longitudinal survival data from the Centers for Medicare and Medicaid Services, comparisons of intersocietal variation will become increasingly informative [5].

Lobectomy for Lung Cancer Composite Score

Composite performance measures have been developed in cardiac surgery by the STS Quality Measurement Task Force for coronary artery bypass graft surgery, aortic valve replacement, and coronary artery bypass graft surgery plus aortic valve replacement [6–8]. These measures have been made available to the public for the purpose of comparing the quality of care between programs [9]. However, until this year, no such measure has been developed in general thoracic surgery. Because lobectomy is the most common major thoracic procedure in the STS GTSD, it was chosen to create the first thoracic composite quality measure.

To develop a thoracic composite quality measure based on observed to expected mortality and major complication rates, the STS GTSD was queried for all patients who underwent lobectomy for lung cancer between July 1, 2011, and June 30, 2014 [10]. That identified 20,657 patients for analysis from 231 institutions. Participant-specific risk-adjusted rates of major morbidity and 30-day mortality were estimated in a Bayesian hierarchical model. Mortality was weighted approximately four times greater than a major complication, consistent with the previously reported adult cardiac surgery quality measures [8]. Covariates in this model were based on the most recently published lung cancer resection risk model [11], and included age, sex, year of operation, body mass index, hypertension, steroid therapy, congestive heart

Table 1. The Society of Thoracic Surgeons General Thoracic Surgery Database Aggregate Outcomes

Outcomes (Reference)	Major Morbidity (%)	30-Day Mortality (%)	LOS, Days[a] Median (IQR)
Lobectomy [19]	9.4	1.3	4 (4)
Pneumonectomy [19]	16.0	5.0	5 (4)
Esophagectomy [20]	33.1	3.4	10 (7)

[a] Data from 21st data analysis of The Society of Thoracic Surgeons General Thoracic Surgery Database.

IQR = interquartile range; LOS = length of stay.

failure, coronary artery disease, peripheral vascular disease, reoperation, preoperative chemotherapy within 6 months, cerebrovascular disease, diabetes mellitus, renal failure, dialysis, past smoker, current smoker, forced expiratory volume in 1 second percent of predicted, Zubrod score, ASA classification, and pathologic stage as defined by the American Joint Committee on Cancer, 6th edition.

Similar to previously reported STS adult cardiac composite measures [8, 12], participant composite scores were estimated and reported with 95% Bayesian credible intervals (CrIs). Participants were categorized as one-, two-, or three-star programs based on their lower-than-expected, expected, or higher-than-expected composite score, respectively. To be a one-star program or a three-star program, the 95% CrI had to fall entirely below or above the STS average score, respectively. This categorized 4.7% of participating programs (n = 8) as one-star programs and 7% (n = 12) as three-star programs. During modeling, it was found that using 98% CrIs provided higher specificity but inadequate differentiation between programs. If 80% CrIs were used, the percentage of high- and low-performing programs was substantially larger but the ability of the model to appropriately identify high-performing and low-performing programs faltered. Therefore, 95% CrIs (corresponding to 97.5% Bayesian probability) was chosen as a compromise between providing accurate classification and reasonable discrimination among programs. It is important to note that a two-star rating should not be interpreted as "average" results. A two-star rating represents expected outcomes for a participant of the STS GTSD, which have been shown to be superior to that of thoracic surgery being performed nationwide [13].

Examining only centers performing 30 or more lobectomies over the 3-year study period, the reliability of the STS lobectomy composite measure was 0.56 (95% CrI, 0.45 to 0.66), which is similar to the reliability of the STS aortic valve replacement plus coronary artery bypass graft surgery composite measure, which was 0.51 (95% CrI, 0.46 to 0.55) [8]. Reliability is used to report the suitability of a measure for profiling by describing how well performance of different providers can be distinguished. A reliability of 0 suggests that all of the variability in a measure is the result of measurement error. A reliability of 1 suggests that all variability in performance is attributable to real differences. Sample size, true performance differences between providers, and measurement error are the main drivers of reliability [14]. This volume threshold allowed a star rating to be assigned to approximately 75% of programs, with the remainder having insufficient information to provide a meaningful rating. These data have been reported to each STS GTSD participant in the semiannual institution-specific performance summary. Figure 3 illustrates participant performance in order of increasing composite score.

Thirty-day mortality and hospital length of stay have been previously shown to be lower for STS GTSD participants than national benchmarks [13]. Therefore, the STS GTSD participant results will also be reported relative to National Inpatient Sample data. By reporting the superior results of STS GTSD participants relative to the National Inpatient Sample, it is hoped that participation in the database will increase. Participants will have the option of voluntary public reporting on the STS public reporting website in 2017. It is hoped that development of the lobectomy composite measure will facilitate accurate comparison of STS GTSD participants, encourage quality assessment, and provide meaningful comparison with national benchmarks. With the planned inclusion of long-term survival data and the development of a composite measure for esophagectomy, the ability to accurately measure and differentiate STS GTSD participant performance is likely to improve.

Table 2. Distribution of Pulmonary Resections in The Society of Thoracic Surgeons General Thoracic Surgery Database and European Society of Thoracic Surgery Between 2010 and 2013

Resections	STS (n = 47,539)	ESTS (n = 30,673)
VATS, wedge resection	15,557 (32.5)	3,493 (11.3)
VATS, lobectomy	13,147 (27.5)	2,557 (8.3)
VATS, segmentectomy	923 (1.9)	605 (2)
VATS, bilobectomy	180 (0.4)	47 (0.1)
VATS, pneumonectomy	72 (0.2)	27 (0.1)
Wedge resection	3,301 (6.9)	3,836 (12.4)
Segmentectomy	954 (2)	1,657 (5.4)
Lobectomy	10,735 (22.5)	1,4175 (46)
Sleeve lobectomy	515 (1.1)	638 (2.1)
Bilobectomy	889 (1.9)	1,407 (4.6)
Pneumonectomy	1,266 (2.6)	2,249 (7.3)

Values are n (%). For all values, the p value is less than 0.05. Adapted from Seder et al [3].

ESTS = European Society of Thoracic Surgeons; STS = The Society of Thoracic Surgeons; VATS = video-assisted thoracoscopic surgery.

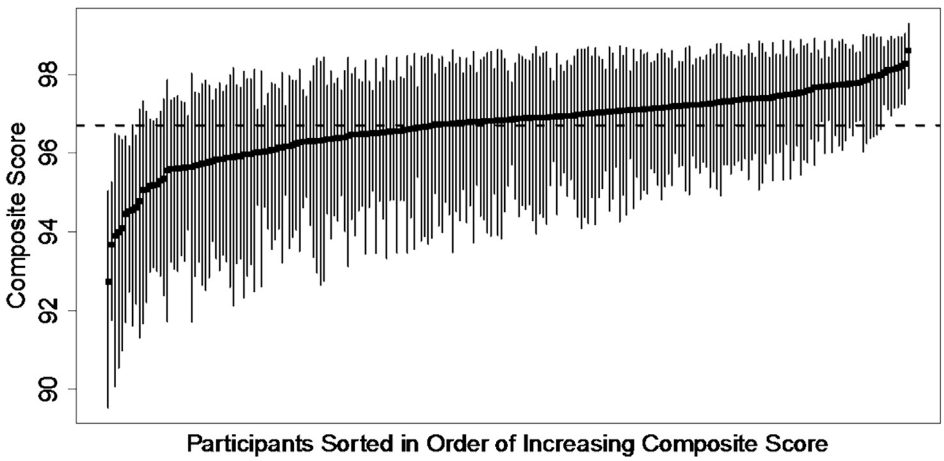

Fig 3. Distribution of participants' composite scores for lobectomy. Participating programs are sorted in order of increasing composite score. Participants on the right of the figure, with higher composite scores and credible intervals that do not cross the average, are high-performing (three-star) programs with fewer deaths and major complications. Modeling was performed using 20,657 records from 231 participants, resulting in 4.7% of programs (8 of 172) being assigned a one-star rating, 88.3% (152 of 172) a two-star rating, and 7.0% (12 of 172) a three-star rating. (Adapted from Kozower et al [10].)

Updated Lung and Esophageal Cancer Resection Risk Models

The STS has previously reported several risk-adjustment models to predict outcomes of individual patients undergoing specific cardiothoracic procedures [6–8, 11, 15, 16]. In adult cardiac surgery, that has resulted in an online calculator for common cardiac procedures that predicts risk of morbidity, mortality, and hospital length of stay, among other outcomes [17]. In 2008, the STS GTSD created the first lung resection risk model [15], which was updated in 2010 [11], with the outcome measures of major morbidity, mortality, and composite major morbidity and mortality. The same year, the STS GTSD Task Force developed the first esophageal cancer resection model to assess the risk of major morbidity, mortality, and composite major morbidity and mortality [16].

Since 2010, however, the landscape of general thoracic surgery has changed substantially, with video-assisted thoracoscopic surgery assuming a prominent role [3], increasing penetrance of the STS GTSD [4], and implementation of regular data auditing [18]. External auditing has demonstrated high rates of agreement between the institutional medical record and the STS GTSD, without identification of intentional gaming or data manipulation. With these changes, this year the STS GTSD lung and esophageal cancer resection risk models were updated using substantially larger, contemporary cohorts of patients.

Lung Cancer Resection Risk Models

To revise the lung cancer resection risk model, the STS GTSD was queried for all pulmonary resections for primary lung cancers between January 1, 2012, and December 31, 2014 [19], and revealed 28,473 records from 231 US centers. A record is a patient receiving an operation. If that patient received a second operation, it is another record. During creation of the risk models, only the first record for a patient was used to prevent confounding. Data from patients who underwent wedge resection, segmentectomy, lobectomy, sleeve lobectomy, bilobectomy, or pneumonectomy were included. As in the 2010 lung resection risk model, the primary outcome measures were operative mortality and major morbidity. Major morbidity was defined as tracheostomy, reintubation, initial ventilatory support longer than 48 hours, adult respiratory distress syndrome, bronchopleural fistula, pulmonary embolus, pneumonia, unexpected return to the operating room for any cause, and myocardial infarction. Multivariate logistic regression analyses were used to create three separate risk models to predict mortality, major morbidity, and composite mortality and major morbidity.

The overall 30-day mortality rate was 1.5% (down from 2.2% between 2002 and 2008), major morbidity rate was 9.1% (up from 8.6% between 2002 and 2008), and the composite major morbidity and mortality rate was 9.5%. Sixty-one percent (up from 36.9% in the first risk model) of lung cancer resections were performed with video-assisted thoracoscopic surgery. Predictors of operative mortality included age, being male, forced expiratory volume in 1 second, body mass index, cerebrovascular disease, steroids, coronary artery disease, peripheral vascular disease, renal dysfunction, Zubrod score, ASA classification, thoracotomy operative approach, induction therapy, reoperation, tumor stage, and greater extent of pulmonary resection. For 30-day mortality, predictors included a Zubrod score of 2 or higher, ASA classification of 4 or 5, thoracotomy operative approach, stage IV disease, and bilobectomy or pneumonectomy. For the composite mortality and morbidity model, similar predictors were identified as in the morbidity model (Table 3). Diffusion capacity of the lung for carbon monoxide was excluded from analysis owing to 15% missing data. Imputation was not thought to be appropriate for diffusing capacity of lung for carbon monoxide because the missing at random assumption was not met. We are working hard to improve the completeness of this data field in the database. Standardized incidence ratios with 95% Bayesian probability intervals for composite

Table 3. Predictors of Mortality, Major Morbidity, and Composite Mortality and Major Morbidity for Lung Cancer Resection, n = 28,473 From 231 Centers

Variable	Mortality Model OR (95% CI)	*p* Value	Major Morbidity Model OR (95% CI)	*p* Value	Composite Model OR (95% CI)	*p* Value
Age, 10-year increase	1.64 (1.44–1.86)	<0.001	1.13 (1.08–1.19)	<0.001	1.14 (1.08–1.91)	<0.001
Male	1.57 (1.26–1.95)	<0.001	1.39 (1.27–1.52)	<0.001	1.40 (1.29–1.53)	<0.001
Body mass index, kg/m^2		0.004		<0.001		<0.001
≥18.5 to <25	1.00		1.00		1.00	
≥6.0 to <18.5	1.45 (0.86–2.45)		1.39 (1.12–1.71)		1.40 (1.14–1.72)	
≥25.0 to <30.0	0.95 (0.75–1.21)		0.83 (0.75–0.92)		0.84 (0.76–0.92)	
≥30.0 to <35.0	0.61 (0.44–0.85)		0.74 (0.65–0.84)		0.73 (0.65–0.83)	
≥35.0 to ≤99.9	1.16 (0.81–1.65)		0.83 (0.71–0.97)		0.84 (0.72–0.98)	
Hypertension	0.93 (0.74–1.72)	0.54	1.08 (0.99–1.19)	0.097	1.06 (0.97–1.16)	0.23
Steroids	1.75 (1.21–2.75)	0.011	1.30 (1.06–1.58)	0.014	1.34 (1.10–1.16)	0.004
Congestive heart failure	1.50 (1.01–2.23)	0.056	1.17 (0.95–1.44)	0.14	1.19 (0.97–1.46)	0.095
Coronary artery disease	1.30 (1.03–1.63)	0.028	1.17 (1.06–1.30)	0.003	1.18 (1.07–1.31)	0.001
Peripheral vascular disease	1.45 (1.10–1.90)	0.01	1.41 (1.24–1.60)	<0.001	1.41 (1.24–1.59)	<0.001
Reoperation	1.48 (1.09–2.02)	0.017	1.36 (1.17–1.57)	<0.001	1.33 (1.15–1.54)	<0.001
Cerebrovascular disease	1.40 (1.05–1.87)	0.028	1.12 (0.97–1.28)	0.13	1.15 (1.00–1.31)	0.052
Diabetes mellitus	1.07 (0.84–1.37)	0.58	1.02 (0.92–1.14)	0.75	1.02 (0.92–1.14)	0.66
% FEV_1, 10% decrease	1.06 (1.01–1.12)	0.017	1.13 (1.10–1.15)	<0.001	1.12 (1.10–1.15)	<0.001
Induction therapy	1.43 (1.04–1.96)	0.036	1.24 (1.07–1.45)	0.006	1.24 (1.07–1.44)	0.006
Renal dysfunction	1.80 (1.10–2.93)	0.029	1.05 (0.79–1.39)	0.76	1.09 (0.83–1.43)	0.55
Cigarette smoking		0.063		<0.001		<0.001
Never	1.00		1.00		1.00	
Past smoker	1.62 (1.06–2.49)		1.22 (1.04–1.43)		1.26 (1.08–1.47)	
Current smoker	1.61 (1.00–2.58)		1.68 (1.42–1.99)		1.68 (1.43–1.99)	
Zubrod score		<0.001		0.001		<0.001
0	1.00		1.00		1.00	
1	1.59 (1.25–2.00)		1.09 (1.00–1.19)		1.12 (1.02–1.22)	
2–5	2.04 (1.36–3.06)		1.41 (1.17–1.70)		1.44 (1.12–1.72)	
ASA		0.013		<0.001		<0.001
1 or 2	1.00		1.00		1.00	
3	1.40 (0.92–2.15)		1.13 (0.98–1.30)		1.14 (0.99–1.32)	
4 or 5	1.93 (1.19–3.13)		1.49 (1.25–1.78)		1.52 (1.27–1.81)	
Approach		<0.001		<0.001		<0.001
Minimally invasive	1.00		1.00		1.00	
Thoracotomy	1.92 (1.53–2.41)		1.39 (1.27–1.53)		1.43 (1.31–1.56)	
Pathologic stage		0.035		0.17		0.19
I	1.00		1.00		1.00	
II	1.15 (0.90–1.47)		1.10 (0.99–1.22)		1.08 (0.97–1.19)	
III	1.37 (1.03–1.83)		1.13 (0.99–1.28)		1.14 (1.00–1.29)	
IV	2.10 (1.17–3.76)		0.98 (0.71–1.36)		1.01 (0.74–1.38)	
Procedure		<0.001		<0.001		<0.001
Wedge	1.00		1.00		1.00	
Segmentectomy	1.01 (0.53–1.93)		1.24 (0.97–1.58)		1.28 (1.01–1.63)	
Lobectomy	1.73 (1.16–2.57)		2.00 (1.71–2.34)		1.97 (1.69–2.30)	
Sleeve	1.69 (0.71–4.00)		2.13 (1.51–3.01)		2.14 (1.53–2.99)	
Bilobectomy	3.67 (2.16–6.26)		3.08 (2.43–3.91)		2.99 (2.36–3.78)	
Pneumonectomy	4.96 (3.02–8.16)		2.83 (2.25–3.56)		2.91 (2.32–3.64)	
C statistic	0.78		0.68		0.68	

Adapted from Fernandez et al [19].

Data from January 1, 2012 through December 31, 2014.

ASA = American Society of Anesthesiologists; CI = confidence interval; FEV_1 = forced expiratory volume in 1 second; OR = odds ratio.

mortality and major morbidity demonstrated no overlap between the best and worst performing institutions, suggesting appropriate model discrimination (Fig 4).

Esophageal Cancer Resection Risk Models

Similar to pulmonary resection, there have been substantial changes in the field of esophageal resection since the first esophageal risk model was created in 2010 [16]. One important change is the increase in minimally invasive approaches to esophagectomy. With this in mind, the STS GTSD was queried for all patients treated for esophageal cancer with esophagectomy between July 1, 2011, and June 30, 2014 [20]. The revised model included 4,142 cases of esophagectomy performed at 164 centers. Multivariable risk models for major morbidity, 30-day mortality, and composite morbidity and mortality were created with the inclusion of surgical approach as a risk factor. Major morbidity was defined as the presence of one or more of the following adverse events: unexpected return to the operating room, anastomotic leak, reintubation, initial ventilatory support longer than 48 hours, pneumonia, renal failure, and recurrent nerve paresis.

The most common procedures performed during the study period were the Ivor-Lewis esophagectomy (32.5%), transhiatal esophagectomy (21.7%), and minimally invasive Ivor-Lewis esophagectomy (21.4%). Overall, minimally invasive techniques were used in 33.8% of the esophageal resections. The 30-day mortality rate was 3.1%, up from 2.7% in the previous risk adjustment model. However, it must be noted that the definition of mortality has changed from inhospital mortality to any death during the index hospitalization or within 30 days of surgery [21]. At least one major morbidity occurred in 33.1% of patients, up from 24% in the previous model. However, the definitions of major morbidity also changed, with the inclusion of renal failure, recurrent laryngeal nerve injury, and unexpected return to the operating room for any cause. Minimally invasive Ivor-Lewis esophagectomy demonstrated a significant perioperative survival advantage when compared with Ivor-Lewis esophagectomy (odds ratio 0.52, 95% confidence interval: 0.30 to 0.92, $p = 0.04$) in the mortality analysis, but not in the composite model. Predictors of composite perioperative morbidity and mortality include age greater than 65 years, congestive heart failure, Zubrod score greater than 1, past or current smoking status, body mass index greater than 35 m^2/kg, and McKeown esophagectomy (Table 4). Like the revised lobectomy model [19], standardized incidence ratios with 95% Bayesian probability intervals for composite mortality and major morbidity demonstrated no overlap between the best and worst performing institutions (Fig 5).

The 2015 revisions of the lung and esophageal cancer resection risk models demonstrate that operative morbidity and mortality rates are low among surgeons in the STS GTSD and provide updated predictors of adverse events. Identification of these predictors will allow thoracic surgeons to more accurately risk-stratify their patients based on individual characteristics and will be useful in measuring hospital performance variation.

National Quality Forum-Endorsed Measures

The NQF currently administers the national library of endorsed performance measures. The NQF is contracted by the US Department of Health and Human Services to establish “a portfolio of quality and efficiency measures that will allow the federal government to more clearly see how and whether healthcare spending is achieving the best results for patients and taxpayers” [22]. By 2017, nearly 10% of all Medicare payment is expected to be performance-based with appropriate measures designated by the NQF. With this in mind, STS leadership

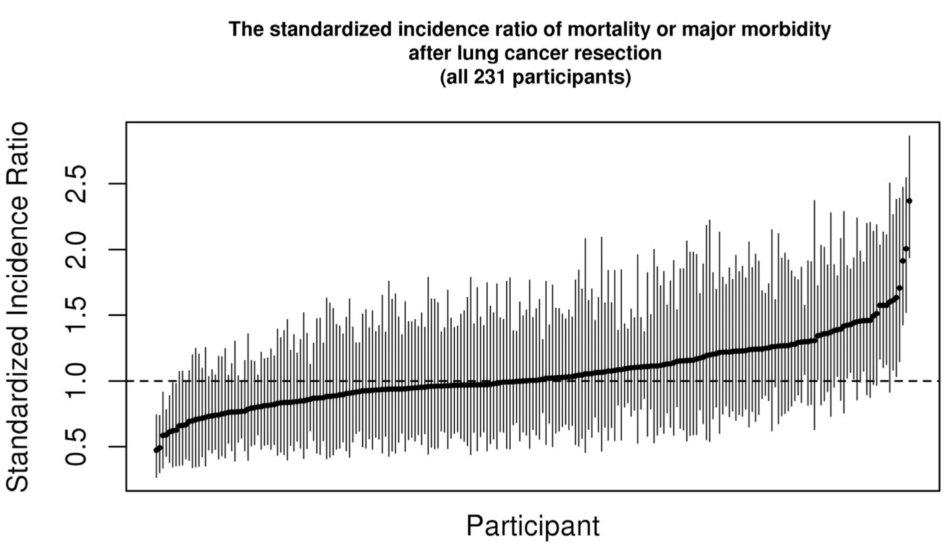

Fig 4. Hospital performance variability for lung resection for cancer. The standardized incidence ratios with 95% Bayesian probability intervals are shown for the composite measure of mortality and major morbidity after lung cancer resection among The Society of Thoracic Surgeons database participating centers. Modeling was performed using 28,473 records from 231 participants, identifying 3.9% of programs (9 of 231) of programs as high performers (Bayesian intervals do not cross 1) and 6.9% (16 of 231) as low performers. (Adapted from Fernandez et al [19].)

Table 4. Predictors of Mortality, Major Morbidity, and Composite Mortality and Major Morbidity for Esophageal Cancer Resection, n = 4,321 From 164 Centers

	Morbidity Model		Mortality Model		Composite Model	
Variable	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value	OR (95% CI)	*p* Value
Age, per 10-y increase <65 years	1.01 (0.89–1.15)	0.85	1.36 (0.90–2.05)	0.14	1.02 (0.90–1.16)	0.74
Age, per 10-y increase >65 years	1.32 (1.12–1.56)	0.001	1.73 (1.17–2.56)	0.006	1.32 (1.12–1.56)	<0.001
Female	0.93 (0.77–1.12)	0.43	0.67 (0.39–1.16)	0.15	0.93 (0.77–1.12)	0.44
Black	1.19 (0.83–1.71)	0.35	0.95 (0.36–2.50)	0.92	1.23 (0.86–1.77)	0.25
Congestive heart failure	1.93 (1.26–2.96)	0.002	1.99 (0.88–4.52)	0.10	1.99 (1.30–3.04)	0.002
Coronary artery disease	1.09 (0.91–1.31)	0.35	1.17 (0.76–1.84)	0.47	1.08 (0.90–1.30)	0.40
Peripheral vascular disease	1.28 (0.93–1.75)	0.13	0.68 (0.28–1.67)	0.40	1.25 (0.92–1.71)	0.16
Zubrod score, versus 0		<0.001		<0.001		<0.001
1	0.92 (0.77–1.10)		1.03 (0.64–1.66)		0.93 (0.78–1.10)	
>1	1.83 (1.31–2.57)		3.21 (1.66–6024)		1.84 (1.32–2.58)	
ASA risk class, versus I or II		0.09		0.55		0.06
III	1.19 (0.97–1.46)		1.29 (0.69–2.43)		1.22 (0.9–1.49)	
IV or V	1.41 (1.03–1.93)		1.58 (0.70–3.78)		1.45 (1.05–1.98)	
Diabetes mellitus	1.03 (0.87–1.21)	0.77	1.22 (0.80–1.86)	0.35	1.05 (0.89–1.24)	0.56
Hypertension	1.13 (0.97–1.32)	0.11	1.09 (0.72–1.66)	0.67	1.13 (0.97–1.31)	0.12
Corticosteroid use	1.45 (0.86–2.45)	0.16	3.33 (1.34–8.28)	0.01	1.53 (0.91–2.57)	0.11
Renal dysfunction	1.46 (0.82–2.60)	0.19	1.29 (0.38–4.45)	0.69	1.43 (0.80–2.54)	0.23
Induction therapy	0.86 (0.72–1.03)	0.10	1.12 (0.70–1.79)	0.64	0.87 (0.72–1.03)	0.11
Cigarette smoking, versus never		<0.001		0.79		<0.001
Past	1.24 (1.05–1.46)		0.86 (0.56–1.33)		1.25 (1.05–1.47)	
Current	1.67 (1.33–2.10)		0.94 (0.50–1.74)		1.65 (1.31–2.07)	
Body mass index, versus normal		<0.001		0.52		<0.001
<18.5	1.17 (0.77–1.78)		2.16 (0.90–5.20)		1.25 (0.82–1.89)	
≥25 and <30	0.9 (0.77–1.08)		1.05 (0.67–1.64)		0.91 (0.77–1.08)	
≥30 and <35	1.01 (0.8–1.24)		0.96 (0.55–1.68)		1.03 (0.84–1.26)	
≥35	1.53 (1.20–1.96)		1.04 (0.51–2.10)		1.53 (1.19–1.95)	
Clinical stage, versus I		0.71		0.84		0.76
II	0.94 (0.76–1.17)		0.87 (0.48–1.57)		0.94 (0.75–1.17)	
III	0.89 (0.70–1.14)		1.01 (0.53–1.92)		0.89 (0.70–1.13)	
IV	0.75 (0.42–1.37)		1.26 (0.32–4.86)		0.81 (0.45–1.45)	
Procedure, versus Ivor-Lewis		<0.001		0.04		<0.001
Transhiatal	1.13 (0.90–1.41)		0.61 (0.35–1.06)		1.13 (0.90–1.40)	
McKeown	1.75 (1.34–2.29)		1.53 (0.86–2.69)		1.76 (1.35–2.30)	
Thoracoabdominal	1.53 (0.92–2.54)		0.68 (0.15–3.06)		1.49 (0.90–2.48)	
MIE Ivor-Lewis type	0.88 (0.71–1.10)		0.52 (0.30–0.92)		0.88 (0.71–1.10)	
MIE transhiatal type	1.12 (0.72–1.74)		0.91 (0.30–2.68)		1.15 (0.74–1.79)	
MIE McKeown type	1.34 (1.00–1.79)		0.65 (0.3–1.39)		1.33 (0.99–1.77)	
C statistic	0.62		0.70		0.62	

Adapted from Raymond et al [20].

Data from July 1, 2011 through June 30, 2014.

ASA = American Society of Anesthesiologists; CI = confidence interval; MIE = minimally invasive esophagectomy; OR = odds ratio.

has worked to obtain endorsement for numerous cardiothoracic performance measures. Currently, STS has 34 NQF-endorsed measures, including 24 in adult cardiac surgery, 6 in general thoracic, and 4 in congenital heart surgery [4]. The thoracic measures that have been endorsed include the following: participation in a systematic national database for general thoracic surgery; recording of clinical stage before lung and esophageal cancer resection; recording of performance status before lung or esophageal cancer resection; risk-adjusted morbidity; length of stay more than 14 days after elective lobectomy for lung cancer; risk-adjusted morbidity and mortality for esophagectomy for cancer; and risk-adjusted morbidity and mortality for lung resection for lung cancer. Between January 2012 and December 2014, clinical staging was performed and recorded in 95.9% of cases lung cancer resections captured by the STS GTSD. The STS GTSD Task Force is in the process of submitting additional metrics for endorsement by the NQF.

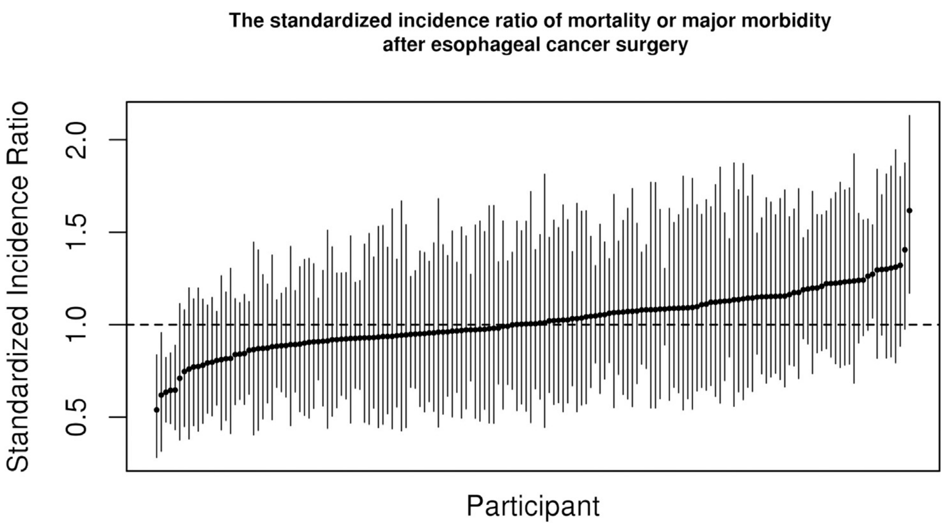

Fig 5. Hospital performance variability for esophageal resection for cancer. The standardized incidence ratios with 95% Bayesian probability intervals are shown for the composite measure of mortality and major morbidity after lung cancer resection among The Society of Thoracic Surgeons database participating centers. Modeling was performed using 4,321 records from 164 participants, identifying 3.0% of programs (5 of 164) as high performers (Bayesian intervals do not cross 1) and 1.2% (2 of 164) as low performers. (Adapted from Raymond et al [20]).

Initiatives for 2016

The GTSD has begun work on several additional projects for 2016. First, we are developing the second composite quality measure for esophagectomy for esophageal cancer. Next, we are working hard to incorporate long-term survival in the database. The GTSD incorporated data fields into version 2.3 to capture 5-year survival for all lung and esophageal cancer resections. In addition, Dr Felix Fernandez is leading STS on a project funded by the Agency for HealthCare Research and Quality to link STS lung cancer resection records with Medicare data. This project will provide 5-year survival on non-Health Management Organization Medicare beneficiaries, a rich database for comparative effectiveness research, and will help us understand resource utilization after lung cancer resection. Finally, the GTSD and Public Reporting Task Forces are working together to provide the opportunity for members to participate in the STS public reporting initiative. Because the STS-GTSD outperforms national benchmarks, we are working on the optimal display to report our outcomes.

Certain unique challenges face the STS GTSD, prime among which is a low data capture rate. A large number of general and cardiac surgeons intermittently perform general thoracic procedures but do not contribute these data to the STS GTSD. That has resulted in an overall data capture rate that is lower than the STS adult cardiac or congenital databases. We have seen with our lobectomy model that a program needs to perform a given procedure often enough to have sufficient reliability for provider profiling. A star rating could not be assigned to nearly 25% of programs owing to low lobectomy volumes, and it is expected that even more programs will not be able to be reliably assigned an esophagectomy star rating.

Conclusion

With improved participation in the STS GTSD, increasingly meaningful analyses are possible. With data from nearly a half million cases available, risk models for lung resection and esophagectomy for cancer have been revised and the first general thoracic composite quality measure has been developed. In addition, international collaboration has opened new doors to understanding variation between American and European treatment patterns, building a foundation for future quality-improvement initiatives. The STS GTSD is dedicated to providing the highest-level data and outcomes measures that are capable of obtaining NQF endorsement, with the goal of improving overall thoracic surgical care.

References

1. Wright CD, Edwards FH. The Society of Thoracic Surgeons General Thoracic Surgery Database. Ann Thorac Surg 2007;83:893–4.
2. Fernandez FG, Falcoz PE, Kozower BD, Salati M, Wright CD, Brunelli A. The Society of Thoracic Surgeons and the European Society of Thoracic Surgeons General Thoracic Surgery Databases: joint standardization of variable definitions and terminology. Ann Thorac Surg 2015;99:368–76.
3. Seder CW, Salati M, Kozower BD, et al. Variation in pulmonary resection practices between The Society of Thoracic Surgeons and the European Society of Thoracic Surgeons General Thoracic Surgery Databases. Ann Thorac Surg March 26, 2016 [epub ahead of print; doi: http://dx.doi.org/10.1016/j.athoracsur.2015.12.073].
4. Jacobs JP, Shahian DM, Prager RL, et al. Introduction to the STS national database series: outcomes analysis, quality improvement, and patient safety. Ann Thorac Surg 2015;100:1992–2000.
5. Jacobs JP, Edwards FH, Shahian DM, et al. Successful linking of The Society of Thoracic Surgeons Adult Cardiac Surgery Database to Centers for Medicare and Medicaid Services Medicare data. Ann Thorac Surg 2010;90:1150–7.
6. Shahian DM, O'Brien SM, Filardo G, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 1—coronary artery bypass grafting surgery. Ann Thorac Surg 2009;88(Suppl):2–22.
7. Shahian DM, He X, Jacobs JP, et al. The Society of Thoracic Surgeons isolated aortic valve replacement (AVR) composite score: a report of the STS Quality Measurement Task Force. Ann Thorac Surg 2012;94:2166–71.

8. Shahian DM, He X, Jacobs JP, et al. The STS AVR+CABG composite score: a report of the STS Quality Measurement Task Force. Ann Thorac Surg 2014;97:1604–9.
9. The Society of Thoracic Surgeons. Public reporting. Available at http://www.sts.org/quality-research-patient-safety/sts-public-reporting-online. Accessed November 14, 2015.
10. Kozower BD, O'Brien SM, Kosinski AS, et al. The Society of Thoracic Surgeons composite score for rating program performance for lobectomy for lung cancer. Ann Thorac Surg 2016;101:1379–87.
11. Kozower BD, Sheng S, O'Brien SM, et al. STS database risk models: predictors of mortality and major morbidity for lung cancer resection. Ann Thorac Surg 2010;90:875–83.
12. Rankin JS, He X, O'Brien SM, et al. The Society of Thoracic Surgeons risk model for operative mortality after multiple valve surgery. Ann Thorac Surg 2013;95:1484–90.
13. LaPar DJ, Bhamidipati CM, Lau CL, Jones DR, Kozower BD. The Society of Thoracic Surgeons General Thoracic Surgery Database: establishing generalizability to national lung cancer resection outcomes. Ann Thorac Surg 2012;94:216–21.
14. Adams J. The reliability of provider profiling: a tutorial. 2009. Available at http://www.rand.org/pubs/technical_reports/tr653.html. Accessed February 18, 2016.
15. Wright CD, Gaissert HA, Grab JD, et al. Predictors of prolonged length of stay after lobectomy for lung cancer: a Society of Thoracic Surgeons General Thoracic Surgery Database risk-adjustment model. Ann Thorac Surg 2008;85:1857–65.
16. Wright CD, Kucharczuk JC, O'Brien SM, et al. Predictors of major morbidity and mortality after esophagectomy for esophageal cancer: a Society of Thoracic Surgeons General Thoracic Surgery Database risk-adjustment model. J Thorac Cardiovasc Surg 2008;137:587–95.
17. The Society of Thoracic Surgeons. Available at http://riskcalc.sts.org/stswebriskcalc/#/. Accessed November 14, 2015.
18. Magee MJ, Wright CD, McDonald D, Fernandez FG, Kozower BD. External validation of The Society of Thoracic Surgeons General Thoracic Surgery Database. Ann Thorac Surg 2013;96:1734–9.
19. Fernandez FG, Kosinski AS, Burfeind W, et al. The Society of Thoracic Surgeons lung cancer resection risk model: higher quality data and superior outcomes. Ann Thorac Surg 2016 [in press].
20. Raymond DP, Seder CW, Wright CD, et al. Predictors of major morbidity or mortality after resection for esophageal cancer: a Society of Thoracic Surgeons General Thoracic Surgery Database risk-adjustment model. Ann Thorac Surg 2016 [under review].
21. Overman DM, Jacobs JP, Prager RL, et al. Report from The Society of Thoracic Surgeons National Database Workforce: clarifying the definition of operative mortality. World J Pediatr Congenit Heart Surg 2013;4:10–2.
22. National Quality Forum. Available at http://www.qualityforum.org/home. Accessed November 14, 2015.