

# A perspective on the Society of Thoracic Surgeons Composite Score for evaluating esophagectomy for esophageal cancer

Shuyin Liang, James D. Luketich, Inderpal S. Sarkaria

Department of Cardiothoracic Surgery, University of Pittsburgh School of Medicine and the University of Pittsburgh Medical Center, Pittsburgh, PA, USA

*Correspondence to:* Inderpal S. Sarkaria, 5200 Centre Avenue, Suite 715, Pittsburgh, PA 15232, USA. Email: sarkariais@upmc.edu.

*Provenance:* This is an invited Editorial commissioned by the Section Editor Mathew Thomas (Mayo Clinic, Jacksonville, Florida, USA).

*Comment on:* Society of Thoracic Surgeons General Thoracic Surgery Database Task Force. The Society of Thoracic Surgeons Composite Score for evaluating esophagectomy for esophageal cancer. *Ann Thorac Surg* 2017;103:1661-7.

Submitted Nov 29, 2017. Accepted for publication Dec 20, 2017.

doi: 10.21037/jtd.2017.12.114

View this article at: <http://dx.doi.org/10.21037/jtd.2017.12.114>

The Society of Thoracic Surgeons General Thoracic Surgery Database (STS GTSD) Task Force has proposed a composite performance measure to evaluate quality of care delivered to patients undergoing esophagectomy for esophageal cancer (1). The Task Force summarized clinical data from GTSD version 2.2 between January 01, 2012 to December 31, 2014 for the participating programs, and compared them to the National Inpatient Sample (NIS). The methodology parallels that of the previous STS composite measures for lobectomy for lung cancer (2) and aortic valve replacement (3). In contrast to composite measures derived from administrative data, such as the American College of Surgeons National Surgical Quality Improvement Program (ACS NSQIP), the STS GTSD uses prospectively collected, externally audited clinical data from voluntary participants. It captures non-fatal complications and risk-adjusted outcomes, which are important for low-volume complex operations like esophagectomy, where comparisons are adjusted for comorbidities.

The authors have proposed using the composite score to assign a program quality rating based on a 3-tiered “star” rating system. In calculating the composite score, operative mortality is defined in the GTSD as death occurring at discharge or within 30 days of operation. It is weighted four times greater than complications, such as unexpected return to operating room, anastomosis requiring medical or surgical treatment, reintubation, initial ventilator support >48 hours, pneumonia, renal failure and recurrent

laryngeal nerve paresis. The scores are presented both numerically and in the three-star system. Of the 42.7% of participating programs that are eligible to receive a star rating, 90% of them were rated as two stars, meaning their 95% Bayesian credible interval overlaps with the average score of 90.14. The five programs assigned 3 stars have 95% Bayesian credible intervals completely above 90.14, and the two programs assigned 1 star have their 95% Bayesian credible intervals completely below 90.14. Although the fact that 90% of the scores are in the two-star category, one needs to keep in mind that the denominator is the number of participants that were given a score. Out of all 167 participants, 57.3% were not given a score due to having a case volume too low for the score to be reliable. This can create confusion for the clinicians and the public to interpret why a center did not have a score: was it because they did not voluntarily participate in the STS GTSD, or because the case volume was too low? Conversely, while this does allow only the most experienced programs to set the bar for these quality indicators, it is difficult to interpret the meaning of the “star” ratings on an individual program level when the metrics have been derived only from a population of peer-programs with competitive high-level outcomes. Given this, the practical difference in quality between a “2-star” and “3-star” program may be negligible in clinical terms, but the impact and implications on public perception may be significant. Nevertheless, future efforts to expand data collection to a broader and more representative population of participants across North America has great

potential to improve the value and applicability of these scores to more surgeons and programs, and serve as a more reliable and discriminatory indicator of quality outcomes.

The goal of developing a composite score is to improve quality and transparency at a national level. There are additional unique challenges in developing a composite score for esophagectomy. Resectable esophageal cancer remains a relatively rare disease entity, and most programs have low case volumes, as demonstrated in Figure 1 of the article. As mentioned, out of 167 participating programs, less than half of them (42.7%) reported an average of 5 esophagectomies per year during the study period, and only these programs were included in the analysis and eligible to receive a proposed “star” quality rating. This finding by itself is troubling. We know that high volume centers for complex procedures benefit not only from surgeon’s experience, but also the multidisciplinary team’s experience: medical oncology, radiation oncology, anesthesia, radiology, pathology, nursing, and physiotherapy all actively participate in the patient’s pre-operative, intra-operative and postoperative recovery. Past literature has repeatedly shown that high volume regionalized centers have improved outcomes in many complex cancer operations (4,5). For esophagectomy, data from a Canadian database of 6,985 patients showed that high volume centers (>20 cases per year) have decrease in odds of in-hospital mortality by 64% and decreased LOC by 38% (6). With an increase to 10 cases per year, in hospital mortality further decreased by 15% and LOS by 10%. Fuchs *et al.* studied the Nationwide Inpatient Sample database of 23,751 patients, and identified that hospital volume was the only variable that impacted outcomes, with high volume centers (>20 cases/year) having less than half of the perioperative mortality rate as the low volume centers (4.01% high volume *vs.* 11.4% low volume) (7). Thus the fact that 42.7% of the STS GTSD voluntarily reported programs have an average of 5 esophagectomies per year and fewer is particularly concerning, and the real number of low volume centers is likely much higher, and not represented in the quality rating metric.

Focusing on the operation itself, the GTSD collects data on the location of the cancer and the associated operative approach. Most of the cases from the STS GTSD are cancers located in the lower third of the esophagus (60.6%) or esophagogastric junction (30.9%). More than half of the procedures were performed open (66.2%), as opposed to a minimally invasive esophagectomy (MIE). An Ivor Lewis approach accounted for 53.9%, followed

by transhiatal (24.5%). Certainly the characteristic of the approach can affect outcomes. For example, the MIE technique, including robotic approaches, has been shown to be feasible and safe (8-11). Although long term oncologic outcomes appear to be similar, studies have shown that MIE is associated with less intra-operative blood loss, improved lymph node yields, and fewer postoperative complications (12-14). Unfortunately, given that 57.3% of the participating programs do not meet the average 5 esophagectomies per year, the sample size is likely too small to derive composite scores with adequate reliability to differentiate the variety of approaches, and that also more accurately reflect the nature of esophageal resections across a broader population of hospitals.

Unlike cardiac procedures, esophagectomy has traditionally been performed by surgeons with vast heterogeneity in their training background and experience. Across the country, esophagectomy can be performed by general surgeons, foregut/minimally invasive trained general surgeons, surgical oncologists and general thoracic surgeons. At some institutions, the general surgeons and the thoracic surgeons cooperate to perform the esophagectomy together as a team. It is difficult to classify and compare the outcomes of esophagectomy at a general surgeon predominant program, thoracic surgeon predominant program, and a cooperative team program. Dimick *et al.* used the national Medicare database from 1998–1999 to show that mortality rates were 37% higher for surgeons without thoracic specialty training (20.7% without *vs.* 10.7% with) (15). However, this difference was not as great as the difference in mortality between high volume *vs.* low volume centers (11.4% high *vs.* 24.3% low). Ten years later, Smith *et al.* reported that the majority of esophagectomies are performed by thoracic surgeons (59%) who prefer the Ivor Lewis technique, though thoracic surgery training did not affect outcomes such as hospital stay, complications and mortality (16). In 2016, Khoushal *et al.* analyzed the NSQIP database 2006–2013, and found that out of the 5,142 esophagectomies identified, 70.3% were performed by general surgeons and 29.7% by cardiothoracic surgeons (CTS) (17). Despite the fact that the CTS patients had significantly higher comorbidities and cancer rates, there was a lower incidence of infection and shorter hospital stay, and equivalent mortality rates. In the article of interest, Table 7 compares the STS and NIS outcomes for esophagectomy, and shows that the STS GTSD participants as a whole have comparable discharge mortality rates, but less variability, shorter length of stable than NIS

participants. Since many STS participants are also in the NIS database, one may argue that if taken out the STS participants from the NIS database, the NIS performance would be worse.

One may counter-argue that STS participants have better outcomes than NIS because STS participants are voluntary, whereas NIS is an “all-comers” dataset based on administrative data. Nevertheless, given the trend of surgical education and increased surgical sub-specialization, current data suggests a significant benefit from having a high-volume, experienced general thoracic surgeon performing esophagectomy.

While several significant concerns regarding the current proposed quality rating system have been outlined, there is significant future potential from a well-organized STS GTSD. Currently, the STS GTSD score focuses on short term quality measures. However, it is the long term oncologic outcome that is used to justify the invasiveness and toxicity of treatment. In order to differentiate high performance program from low performance program, all these factors need to be considered: short term safety outcomes, post-treatment quality of life and long term oncologic outcomes. After ensuring superior post-operative recovery, the next step in esophagectomy outcome score development would incorporate measures for better quality of life and long-term oncologic survival. As the STS GTSD expands and matures over time, with broader participation from all programs, it is highly likely these composite scores will ultimately be able to provide meaningful measures of quality to better gauge and guide the state of esophageal resection for cancer in North America.

## Acknowledgements

None.

## Footnote

*Conflicts of Interest:* Dr. Inderpal S. Sarkaria: speaker and education for Intuitive Surgical, Inc. The other authors have no conflicts of interest to declare.

## References

1. Task-Force S. The Society of Thoracic Surgeons Composite Score for Evaluating Esophagectomy for Esophageal Cancer. *Ann Thorac Surg* 2017;103:1661-7.
2. 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;101:1379-86; discussion 1386-7.
3. 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.
4. Ho V, Town RJ, Heslin MJ. Regionalization versus competition in complex cancer surgery. *Health Econ Policy Law* 2007;2:51-71.
5. Finley CJ, Bendzsak A, Tomlinson G, et al. The effect of regionalization on outcome in pulmonary lobectomy: a Canadian national study. *J Thorac Cardiovasc Surg* 2010;140:757-63.
6. Finley CJ, Jacks L, Keshavjee S, et al. The effect of regionalization on outcome in esophagectomy: a Canadian national study. *Ann Thorac Surg* 2011;92:485-90; discussion 490.
7. Fuchs HF, Harnsberger CR, Broderick RC, et al. Mortality after esophagectomy is heavily impacted by center volume: retrospective analysis of the Nationwide Inpatient Sample. *Surg Endosc* 2017;31:2491-7.
8. Luketich JD, Pennathur A, Franchetti Y, et al. Minimally invasive esophagectomy: results of a prospective phase II multicenter trial-the eastern cooperative oncology group (E2202) study. *Ann Surg* 2015;261:702-7.
9. Luketich JD, Pennathur A, Awais O, et al. Outcomes after minimally invasive esophagectomy: review of over 1000 patients. *Ann Surg* 2012;256:95-103.
10. Okusanya OT, Sarkaria IS, Hess NR, et al. Robotic assisted minimally invasive esophagectomy (RAMIE): the University of Pittsburgh Medical Center initial experience. *Ann Cardiothorac Surg* 2017;6:179-85.
11. Sarkaria IS, Rizk NP, Grosser R, et al. Attaining Proficiency in Robotic-Assisted Minimally Invasive Esophagectomy While Maximizing Safety During Procedure Development. *Innovations (Phila)* 2016;11:268-73.
12. Ahmadi N, Crnic A, Seely AJ, et al. Impact of surgical approach on perioperative and long-term outcomes following esophagectomy for esophageal cancer. *Surg Endosc* 2017. [Epub ahead of print].
13. Kauppi J, Rasanen J, Sihvo E, et al. Open versus minimally invasive esophagectomy: clinical outcomes for locally advanced esophageal adenocarcinoma. *Surg Endosc* 2015;29:2614-9.
14. Biere SS, van Berge Henegouwen MI, Maas KW, et

- al. Minimally invasive versus open oesophagectomy for patients with oesophageal cancer: a multicentre, open-label, randomised controlled trial. *Lancet* 2012;379:1887-92.
15. Dimick JB, Goodney PP, Orringer MB, et al. Specialty training and mortality after esophageal cancer resection. *Ann Thorac Surg* 2005;80:282-6.
16. Smith BR, Hinojosa MW, Reavis KM, et al. Outcomes of esophagectomy according to surgeon's training: general vs. thoracic. *J Gastrointest Surg* 2008;12:1907-11.
17. Khoushhal Z, Canner J, Schneider E, et al. Influence of Specialty Training and Trainee Involvement on Perioperative Outcomes of Esophagectomy. *Ann Thorac Surg* 2016;102:1829-36.

**Cite this article as:** Liang S, Luketich JD, Sarkaria IS. A perspective on the Society of Thoracic Surgeons Composite Score for evaluating esophagectomy for esophageal cancer. *J Thorac Dis* 2018;10(1):94-97. doi: 10.21037/jtd.2017.12.114