Looking beyond the eyeball test: A novel vitality index to predict recovery after esophagectomy

Andrew Tang, MD,[a] Usman Ahmad, MD,[a] Siva Raja, MD, PhD,[a] Jesse Rappaport, MD,[a] Daniel P. Raymond, MD,[a] Monisha Sudarshan, MD, MPH,[a] Alejandro C. Bribriesco, MD,[a] Eugene H. Blackstone, MD,[a,b] and Sudish C. Murthy, MD, PhD[a]

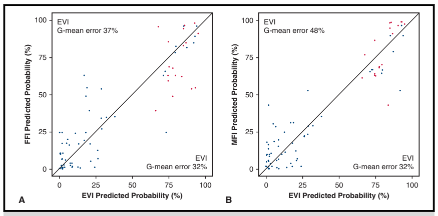
ABSTRACT

Objectives: To (1) measure 4 physiologic metrics before esophagectomy, (2) use these in an index to predict composite postoperative outcome after esophagectomy, and (3) compare predictive accuracy of this index to that of the Fried Frailty Index and Modified Frailty Index.

Methods: Grip strength (kilograms), 30-second chair sit-stands (number), 6-minute walk distance (meters), and normalized psoas muscle area (cm²/m) were measured for 77 consenting patients from January 1, 2018, to April 1, 2019. Imbalanced random forest classification estimated probability of a composite postoperative outcome, which included mortality, respiratory complications, anastomotic leak, delirium, length of stay ≥14 days, discharge to nursing facility, and readmission. G-mean error was used to compare predictive accuracy among indexes.

Results: Median grip strength was 38 kg (25th-75th percentiles, 31-44), number of sit-stands 11 (10-14), psoas muscle area to height ratio 6.9 cm²/m (6.0-8.2), and 6-minute walk distance 407 m (368-451). There was generally weak correlation between these metrics, with the highest between 30-second sit-stands and 6-minute walk distance ($r = 0.57$). Age, degree of patient-reported exhaustion, and the 4 objective metrics comprised the Esophageal Vitality Index, which had a lower G-mean error of 32% (31-33) than the Fried Frailty Index, 37% (37-38), and the Modified Frailty Index, 48% (47-48).

Conclusions: The Esophageal Vitality Index, an objective, simple assessment consisting of grip strength, 30-second chair sit-stands, 6-minute walk, and psoas muscle area to height ratio outperformed commonly used frailty indexes in predicting post-esophagectomy mortality and morbidity. The index provides a robust picture of patients' fitness for surgery beyond the qualitative "eyeball" test. (J Thorac Cardiovasc Surg 2021;161:822-32)



Esophagectomy Vitality Index (EVI) error vs Fried (FFI) & Modified Frailty (MFI) Indexes.

CENTRAL MESSAGE

Quantitative physiologic metrics of physical fitness predict complications after esophagectomy more accurately than standard qualitative frailty indexes.

PERSPECTIVE

Grip strength, 30-second chair sit-stands, psoas muscle area to height ratio, and 6-minute walk distance each provide a measure of patient fitness. Using them together—the Esophagectomy Vitality Index from this study—predicts a patient's likelihood of recovery after esophagectomy more accurately than the more commonly used Fried Frailty Index or Modified Frailty Index.

See Commentaries on pages 833 and 834.

There are few surgical interventions of greater magnitude than esophagectomy, which is accompanied by major morbidity of 33% and mortality of 3%.[1] Identifying patients at greatest risk for adverse outcomes after esophagectomy is important for caregivers and patients alike. Risk factors for adverse outcomes include "frailty," a fairly nebulous variable subject to observer bias. As a result, current assessment of vitality and fitness for esophagectomy is crude and imprecise.[2-5]

Scanning this QR code will take you to the table of contents to access supplementary information. To view the AATS Annual Meeting Webcast, see the URL next to the webcast thumbnail.

From the [a]Department of Thoracic and Cardiovascular Surgery, Heart, Vascular, and Thoracic Institute; and [b]Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio.

This study was supported in part by the Gus P. Karos Registry Fund, the Drs Sidney and Becca Fleischer Heart and Vascular Education Chair, and the Daniel and Karen Lee Endowed Chair in Thoracic Surgery. Dr Tang is a National Heart, Lung, and Blood Institute Clinical Research Scholar of the Cardiothoracic Surgical Trials Network (National Institutes of Health Grant U01 HL088955).

Clinical Trial Registration: Clinicaltrials.gov; Identifier: NCT03413449

Read at the 100th Annual Meeting of The American Association for Thoracic Surgery: A Virtual Learning Experience, May 22-23, 2020.

Received for publication June 10, 2020; revisions received Oct 2, 2020; accepted for publication Oct 18, 2020; available ahead of print Nov 13, 2020.

Address for reprints: Sudish C. Murthy, MD, PhD, Department of Thoracic and Cardiovascular Surgery, Cleveland Clinic, 9500 Euclid Ave/Desk J4, Cleveland, OH 44915 (E-mail: murthys1@ccf.org).

0022-5223/$36.00

https://doi.org/10.1016/j.jtcvs.2020.10.122

Abbreviations and Acronyms
OOB = out-of-bag
STS = Society of Thoracic Surgeons
VIMP = variable importance

Thus, to characterize patients' fitness for esophagectomy, objectives of this study were to (1) measure 4 simple physiologic metrics before esophagectomy and investigate the degree to which each provides independent information, (2) use these metrics to develop an Esophagectomy Vitality Index and test it as a predictor of 30-day composite outcome, and (3) compare this index's predictive accuracy against 2 commonly used, but qualitative, frailty indexes: the Fried Frailty Index[3] and the Modified Frailty Index.[4]

METHODS

Patients

From January 1, 2018, to April 1, 2019, 86 patients were scheduled for esophagectomy at Cleveland Clinic. With institutional review board approval (no. 17-1708), 77 were enrolled in this prospective study of physiologic metrics for estimating physical fitness for surgery (Figure E1). Patients undergoing esophagectomy for benign diseases (eg, end-stage achalasia) and those who ultimately did not undergo esophagectomy (eg, discovery of metastatic disease) were excluded. All patients were evaluated and consented by 1 clinician (A.T.) in the outpatient clinic during their preoperative visit. Median age was 65 years (25-75th percentiles, 60-72), with 66 (86%) male, and a median body mass index of 27 kg/m^2 (24-31) (Table 1). Forty-nine patients (65%) had more than a 10-pound weight loss over the previous year, and 65 (84%) and 55 (71%) had undergone neoadjuvant chemotherapy and radiation, respectively.

Physiologic Metrics

Conduct of study. We measured upper body strength (grip strength), lower body strength and balance (30-second chair sit-stands), muscle mass (psoas muscle area to height ratio), and cardiopulmonary endurance (6-minute walk distance). These metrics were assessed during the preoperative visit immediately preceding the planned operation. Typically, this visit occurred within the week before esophagectomy and after neoadjuvant treatment for patients who underwent induction therapy.

Grip strength. Using their dominant hand, patients performed 3 consecutive isometric contractions of a calibrated dynamometer for 5 seconds each time. The measurements were averaged and recorded. For reference, an average dominant-handed grip strength is 43 kg for a 60-year-old man and 25 kg for a 60-year-old woman.[6]

30-second chair sit-stands. Patients were seated on a straight-backed chair with their arms placed across their chest and feet flat on the floor. They were asked to stand up straight to a full standing position while keeping their back straight and arms across their chest. This was performed as many times as the patient could comfortably tolerate over 30 seconds. If a patient was over halfway to a full standing position when 30 seconds elapsed, this was counted as a stand. For reference, a 60-year-old man would be below average if he performed fewer than 14 sit-stands, and a 60-year-old woman would be below average if she performed fewer than 12 sit-stands.[7]

Psoas muscle area. On axial computed tomography imaging, cross-sectional areas of right and left psoas muscles[8] were measured at the inferior edge of the L3 endplate at the level of the iliac crests.[9] The average of these 2 measurements was recorded and normalized to patient height (cm^2/m). This method of normalization was data-driven for this study, chosen over the Psoas Muscle Index ([right psoas area + left psoas area]/height^2) because average psoas muscle area to height ratio was found to be a more important predictor.

Six-minute walk. Patients were asked to walk at a comfortable pace for 6 minutes, traveling back and forth on a flat 30-m surface. The distance traveled during this time and any oxygen desaturation events were recorded. If a patient was unable to walk for the full 6 minutes, the time until the patient stopped walking was recorded. For reference, a healthy 65-year-old man can walk 575 m and a healthy 65-year-old woman 550 m in 6 minutes.[9]

Existing Frailty Indexes

Fried Frailty Index. The Fried Frailty Index[3] categorizes frailty based on 5 qualities: exhaustion, shrinking, low physical activity level, weakness, and slow walking speed (Table E1). Level of exhaustion was determined by asking 2 questions from the Centers for Epidemiologic Studies Depression scale (Table E1).[10] If a patient had 3 or more of these qualities, the patient was labeled "frail," if 1 or 2, "intermediately frail," and if none, "not frail."

Modified Frailty Index. The Modified Frailty Index[4] is an 11-item count of comorbidities from the original 70-item Canada Study of Health and Aging Frailty Index (Table E2).[4,5] To calculate the index in this study, patients were asked about the 11 comorbidities, assigning 1 point to each self-reported comorbidity to generate a score ranging from 0 to 11. The Index was calculated a second time based on the same comorbidities abstracted by full-time registry nurses into the Society of Thoracic Surgeons (STS) General Thoracic Surgery Database. Both the patient-reported and STS-reported indexes were considered in the analysis.

End Point

The end point was a composite of 30-day mortality and STS-defined complications of delirium, respiratory and gastrointestinal systems (including anastomotic leaks), postoperative length of stay ≥14 days, discharge to a nursing facility, and readmission within 30 days of hospital discharge (Table 2). These outcomes were chosen prior to patient recruitment and study analysis.

Data Analysis

Continuous variables are summarized as median and 25th-75th percentiles and categorical variables as frequency and percentage. Random forest classification methodology using the quantile-classifier approach for class-imbalanced data was used to analyze the composite endpoint.[11] All computations used open-source randomForestSRC R-software under default settings. Missing data were pre-imputed without outcome information using R-software missForest.[12] Thereafter, 1000 trees were grown, with each tree constructed using an independent bootstrap sample consisting on average of 63% of the original patients (in-sample bootstrap data). The remaining patients were referred to as out-of-bag (OOB) observations. Each tree and its corresponding OOB observations were used to calculate an OOB (cross-validated) odds and variable importance (VIMP) measure for each of the independent variables (VIMP).[13] To identify the important candidate variables, a random forest was grown considering the 4 measured metrics: Fried Frailty Index, patient-reported and STS-abstracted Modified Frailty Index, and patient and cancer characteristics (candidate variables listed in Appendix E1).

To interpret VIMP and visualize the shape of relations between outcomes and independent variables, partial dependency graphs were produced.[14] The risk-adjusted partial dependency graph provides the forest-averaged prediction for an independent variable across 1000 trees, adjusting for the predictors in the random forest.[15] Although no formal cutoff analyses were performed, we visually identified inflection points from these partial dependency graphs.

TABLE 1. Patient and cancer characteristics and esophagectomy details

Variables	n*	No. (%) or median [25th-75th percentiles]
Demographics		
Age, y	77	65 [60-72]
Male	77	66 (86)
Body mass index, kg/m^2	77	27 [24-31]
Esophageal cancer characteristics		
Adenocarcinoma	77	72 (94)
Squamous cell carcinoma	77	5 (6.5)
Clinical T stage	77	
T1		8 (11)
T2		19 (25)
T3		50 (65)
T4		0 (0)
Grade	68	
Well differentiated		6 (8.8)
Moderately differentiated		36 (53)
Poorly differentiated		26 (38)
Comorbidities		
Previous non-esophageal cancer	77	3 (3.9)
Weight loss (>10 lb in past year)	77	49 (64)
Pulmonary function		
FEV1 (% of predicted)	76	91 [81-101]
DLCO (% of predicted)	76	81 [70-96]
Cardiac function		
Ejection fraction, %	69	63 [58-68]
Laboratory findings		
Albumin, g/dL	77	4.0 [3.8-4.2]
Creatinine, mg/dL	77	0.91 [0.76-1.1]
Hemoglobin, g/dL	77	13 [12-14]
Hematocrit, %	77	39 [35-42]
Preoperative physical fitness		
Weekly METs	77	21 [7.3-46]
CESD Exhaustion Question 1: I felt everything I did was an effort (past week)	77	
Rare (<1 d)		47 (61)
Some (1-2 d)		16 (21)
Moderately (3-4 d)		5 (6.5)
Most (5-7 d)		9 (12)
CESD Exhaustion Question 2: I could not get "going" (past week)	77	
Rare (<1 d)		61 (79)
Some (1-2 d)		7 (9.1)
Moderately (3-4 d)		5 (6.5)
Most (5-7 d)		4 (5.2)
Preoperative cancer therapy		
Chemotherapy	77	65 (84)
Radiotherapy	77	55 (70)
Days from preoperative assessment to surgery		6 [3-11]
Type of esophagectomy for cancer	77	
McKeown		42 (55)
Ivor Lewis		14 (18)
Thoracoabdominal		20 (26)

(Continued)

TABLE 1. Continued

FEV1, Forced expiratory volume in 1 second; *DLCO*, percent predicted diffusing capacity of lung for carbon monoxide; *METs*, metabolic equivalents; *CESD*, Center for Epidemiological Studies Depression scale. *Patients with data available.

We then created 3 separate imbalanced random forests, each containing the most important and universally applicable predictors from the endpoint random forest described previously with:

1. Four Measured Metrics
2. Fried Frailty Index
3. Modified Frailty Index, calculated from STS-abstracted comorbidities rather than patient self-reported ones.[5]

The geometric mean error (G-mean), 1 minus the square root of the product of sensitivity and specificity, was calculated for each random forest model. G-mean error is influenced by the number of trees grown in the random forest; however, because the number of trees in each random forest was the same for each of the 3 separate imbalanced random forests, G-mean error estimates are directly comparable. We compared these using the Wilcoxon rank-sum test to identify the most accurate predictive model.

RESULTS

Physiologic Metrics

Median grip strength was 38 kg (31-44), number of sit-stands 11 (10-14), psoas muscle area to height ratio

TABLE 2. Postesophagectomy outcomes incorporated into composite end point (n = 77)

Outcomes	No. (%)
30-d mortality	1 (1.3)
Complications*	18 (23)
Pulmonary	
Pleural effusion requiring drain	5 (6.5)
Pneumonia	6 (7.8)
Pulmonary embolism	1 (1.3)
Pneumothorax	1 (1.3)
Prolonged intubation	2 (2.6)
Gastrointestinal	
Anastomotic leak	3 (3.8)
Requiring treatment	3 (3.8)
Requiring reoperation	2 (2.6)
Conduit necrosis	0 (0)
Esophageal dilatation	0 (0)
Chylothorax	0 (0)
Clostridium difficile infection	2 (2.6)
Delirium	5 (6.5)
Prolonged length of stay (≥14 d)	14 (18)
Discharge to nursing facility	9 (12)
Readmission within 30 d of discharge†	7 (9.2)

*Not mutually exclusive. †n = 76.

6.9 cm^2/m (6.0-8.2), and 6-minute walk distance 407 m (368-451) (Table 3). Correlation between these metrics was generally weak, with the highest correlation between 30-second sit-stands and 6-minute walk distance (r = 0.57) (Figure E2). According to the Fried Frailty Index, the majority of patients (n = 56 [73%]) were classified as intermediately frail, 6 (7%) as frail, and 15 as non-frail. The median STS-abstracted Modified Frailty Index was 1 (0-2), and the median patient-reported Modified Frailty Index was 2 (1-3). On an individual patient basis, most comorbidities other than hypertension and chronic obstructive pulmonary disease (both 83% concordance) were concordant greater than 90% of the time between patient-reported and STS-abstracted data (Table E3). However, only 11 patients (14%) had patient-reported and STS-abstracted comorbidities that were identical.

Outcomes After Esophagectomy

Twenty-eight patients (36%) developed at least 1 of the composite outcomes (Table 2). Six (7.8%) developed pneumonia, 5 (6.5%) developed delirium, and 3 (3.8%) had anastomotic leaks, with 2 (2.6%) requiring reoperation One patient (1.2%) died within the 30-day perioperative period, 14 (18%) were hospitalized ≥14 days, 9 (12%) were discharged to a nursing facility, and 7 (9.2%) were readmitted within 30 days of hospital discharge.

Predictors of Esophagectomy Outcomes

The 4 measured frailty metrics, in addition to age (Figure E3), smoking pack-years, patient-reported history of transient ischemic attack or cerebrovascular accident, percent predicted forced expiratory volume in 1 second, neoadjuvant chemotherapy, Fried Frailty Index, and American Society of Anesthesiologists score were the most important predictors of the composite outcome (Figure 1, *A*). Fried Frailty Index and patient-reported Modified Frailty Index were more important than that calculated from STS-abstracted data (Figure 1, *B*). Patients with an average grip strength lower than 35 kg and those performing fewer than 10 sit-stands in 30 seconds had a higher predicted risk of developing the composite outcome (Figure 2, *A* and *B*), as did patients with a psoas muscle area to height ratio <7 cm^2/m (Figure 2, *C*). Patients who were able to walk a longer distance had a linearly lower predicted risk of developing the composite end point without a plateauing effect (Figure 2, *D*).

Esophagectomy Vitality Index Versus Fried Frailty Index and Modified Frailty Index

Grip strength, 30-second sit-stands, psoas muscle area, and 6-minute walk distance, in combination with age and self-reported level of exhaustion according to the Centers for Epidemiologic Studies Depression scale questionnaire, comprised the Esophagectomy Vitality Index.

TABLE 3. Physiologic metrics measured and values for Fried Frailty Index and Modified Frailty Index (n = 77)

Variables	No. (%) or median [25th-75th percentiles]	
Physiologic metrics		
Grip, kg		
#1	39 [31-44]	
#2	38 [30-44]	
#3	36 [30-42]	
Average	38 [31-44]	
30-s chair sit-stands, n	11 [10-14]	
Psoas muscle area, cm^2/m	6.9 [6.0-8.2]	
6-min walk distance, m	407 [368-451]	
Fried Frailty Index components (Table E1)		
Shrinking	49 (64)	
Exhaustion	19 (25)	
Low physical activity	16 (21)	
Weakness	12 (16)	
Slow walking speed	5 (6.5)	
Frailty level		
None	15 (19)	
Intermediate	56 (73)	
Frail	6 (7.8)	
	STS	**Patient reported**
Modified Frailty Index comorbidities (Table E2)		
Diabetes	22 (29)	21 (27)
Congestive heart failure	0 (0)	4 (5.2)
Hypertension	48 (62)	45 (58)
Transient ischemic attack/ cerebrovascular accident	6 (7.8)	5 (6.5)
Dependent functional status	2 (2.6)	2 (2.6)
Myocardial infarction	6 (7.8)	7 (9.1)
Peripheral arterial disease	6 (7.8)	4 (5.2)
Cerebrovascular accident with deficits	2 (2.6)	2 (2.6)
Chronic obstructive pulmonary disease	4 (5.2)	9 (12)
Coronary artery disease	11 (14)	10 (13)
Dementia	2 (2.6)	2 (2.6)
Total score	1 [0-2]	2 [1-3]

STS, Society of Thoracic Surgeons.

Predictions based on the random forest for this index had lower prediction error, 32% (31%-33%) than the Fried Frailty Index, 37% (37%-38%; P < .001) (Figure 3, *A*). Similarly, these predictions had lower error than the STS-abstracted Modified Frailty Index (G-mean error 48% [47%-48%]; P < .001) (Figure 3, *B*). Compared with the Fried Frailty Index and STS-abstracted Modified Frailty Index, our Esophageal Vitality Index assigned greater predicted risk of the composite outcome for patients who ultimately developed the outcome (Figure 3, red circles). Similarly, it assigned a lower predicted risk of the composite outcome for patients who did not ultimately develop the outcome (Figure 3, blue circles).

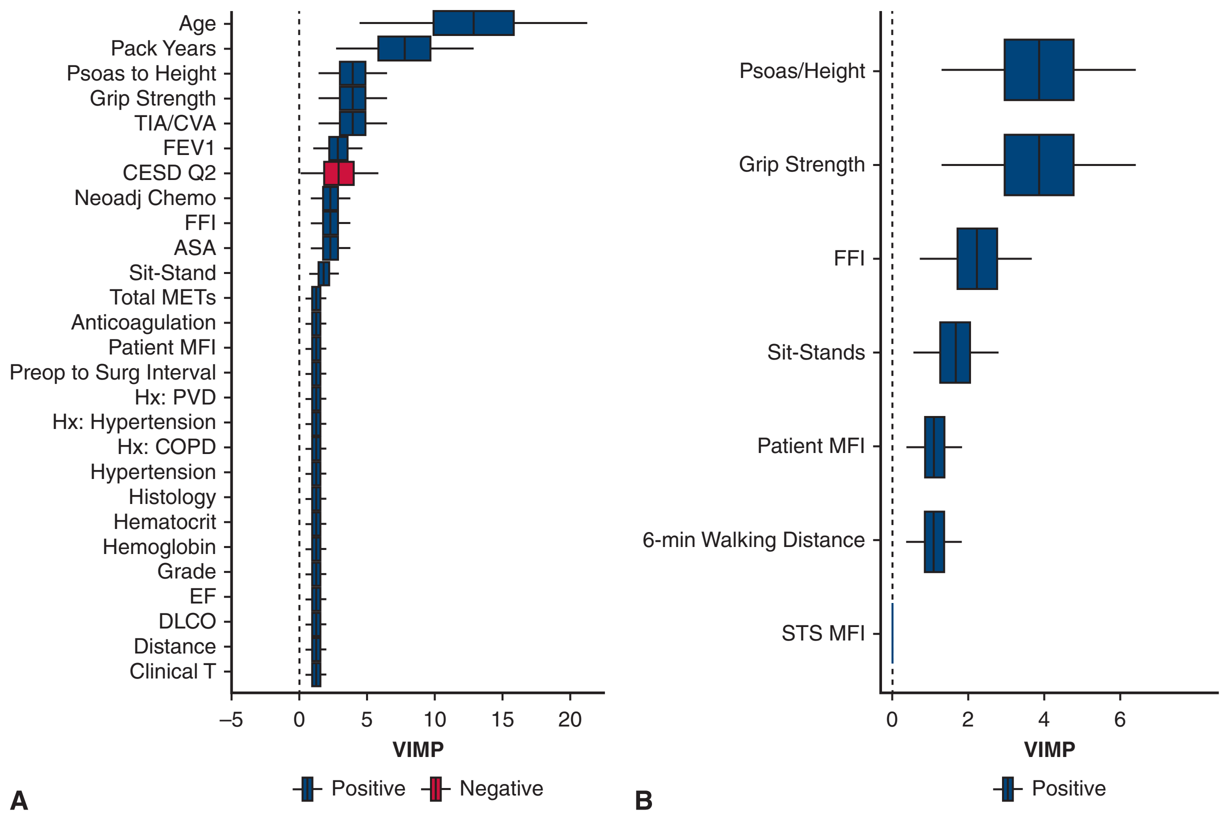


FIGURE 1. Standardized VIMP for predicting complications. A, Top candidate variables. Note that the prefix Hx represents comorbidities abstracted into our general thoracic surgery database, and the duplicated hypertension variable was patient-reported. B, Esophagectomy Vitality Index variables consisting of grip strength, 30-second chair sit-stands, 6-minute walk distance, and psoas muscle area to height ratio, with FFI, patient-reported Modified Frailty Index, and STS-Modified Frailty Index. *Blue boxes* represent important variables with lower 95% confidence interval not extending below 0. *Red box* represents VIMP value with lower 95% confidence interval extending below 0. *Boxes* encompass median (line) and 25th and 75th percentile confidence limits, and *whiskers* 95% confidence limits. *Black vertical line* at 0.0 VIMP represents the point at which a variable does not contribute predictive ability. *TIA/CVA*, Transient ischemic attack/cerebrovascular accident; *FEV1*, forced expiratory volume in 1 second; *CESD Q2*, Center for Epidemiological Studies Depression scale, question 2; *Neoadj Chemo*, neoadjuvant chemotherapy; *FFI*, Fried Frailty Index; *ASA*, American Society of Anesthesiologists; *METS*, metabolic equivalents; *MFI*, Modified Frailty Index; *Preop*, preoperative; *PVD*, peripheral vascular disease; *COPD*, chronic obstructive pulmonary disease; *EF*, ejection fraction; *DLCO*, diffusing capacity of lung for carbon monoxide; *VIMP*, variable importance; *STS*, Society of Thoracic Surgeons General Thoracic Surgery Database comorbidity definition.

DISCUSSION

Principal Findings

Grip strength, number of 30-second chair sit-stands, psoas muscle area to height ratio, and 6-minute walk distance are simple physiologic metrics, each providing independent information reflecting unique aspects of physical vitality and fitness for esophagectomy. When considered individually, these metrics were among the most important predictors of poor outcome after esophagectomy for cancer, and when considered together in what we term the Esophagectomy Vitality Index, they appeared to be more informative in predicting an unfavorable outcome following esophagectomy than the Fried Frailty Index or Modified Frailty Index (Figure 4).

Clinical Significance

Risk estimation for any medical intervention depends on the specific outcome being investigated. Mortality, the most important outcome, has become so infrequent for many "high-risk" interventions (including esophagectomy) that risk models are rapidly becoming underpowered and inaccurate. However, as mortality recedes, new adverse outcomes gain importance. For esophagectomy, these include major specific morbidity (eg, anastomotic leak), length of stay, 30-day readmission, and destination at time of hospital discharge. Although not of equal importance, it has become a pragmatic necessity to formulate composites of these outcomes to provide sufficient power to assess risk, despite loss of specific interpretability.[16,17]

Frailty Indexes

Two standard indexes, the Fried Frailty Index[3] and the Modified Frailty Index,[2,4,5] represent opposite ends of the spectrum of assessing frailty. The Fried Frailty Index was prospectively developed through the Cardiovascular Health Study,[3] broadly categorizing a general population

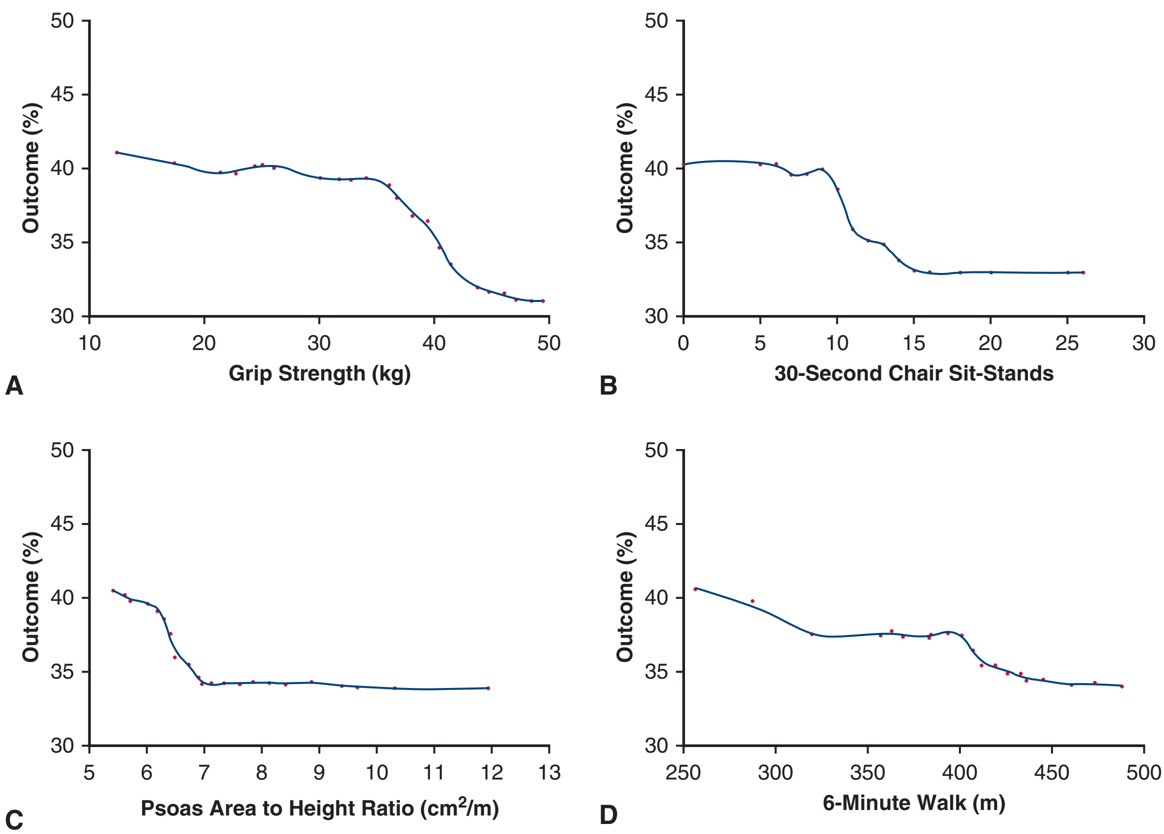


FIGURE 2. Risk-adjusted partial dependency plots based on the Esophagectomy Vitality Index estimating the risk of a patient developing the composite outcome after esophagectomy based on (A) grip strength, (B) 30-second chair sit-stands, (C) psoas muscle area to height ratio, and (D) 6-minute walk distance. *Red circles* indicate risk-adjusted estimates, and *blue lines* indicate smooth fit of risk-adjusted estimates.

of patients age 65 years or older as frail, intermediately frail, or not frail. Its primary objective was to predict falls, hospitalization, and all-cause mortality within 5 years. It more closely captures what clinicians think of as the "eyeball test." In contrast, the Modified Frailty Index was developed as a comorbidity index through the American College of Surgeons National Surgical Quality Program.[2] As a simple tally of patient comorbidities—a measure of disease complexity—it does not measure any physical ability and may be underestimated in large databases due to missing documentation.[18] Neither index was designed specifically for general thoracic surgery patients. In addition, it appears, based on our findings, that what the patient reports and what is abstracted into quality databases often are not perfectly concordant. This is related in part to incomplete medical record documentation, but, particularly in the case of STS variables, more directly to strict definitions of comorbidities that are necessary to mitigate inappropriate up-coding.

Quantifying Physical Status and Fitness for Esophagectomy

Accurate estimation of patient risk is important in general thoracic surgery, where diseases are often debilitating, and operations such as esophagectomy are of considerable magnitude. Patients undergoing esophagectomy for cancer frequently are malnourished and weakened from prolonged dysphagia and induction therapy regimens, which are associated with decline of physiologic function. A 70-year-old man with well-compensated congestive heart failure and early-stage esophageal cancer who walks 10,000 steps a day may be a lower-risk operative candidate than a 60-year-old previously healthy man with locally advanced cancer who has lost 30 pounds over the past 6 months and subsequently lacks the energy to even walk to the front door. Historically, the eyeball test was used to characterize such patients' risk for planned intervention. Subjective variables such as posture, ambulation into the office, and attention to the discussion often served as the basis for surgical candidacy. This is almost certainly not reproducible or accurate when screening patients for operative interventions of formidable magnitude.

Our Esophagectomy Vitality Index uses 4 independent objective metrics to assess upper and lower body strength and balance, muscle mass, and cardiopulmonary endurance. These are representative of patients' real-time physiologic status and do not require costly resources to measure. All study patients performed these metrics within 10 minutes.

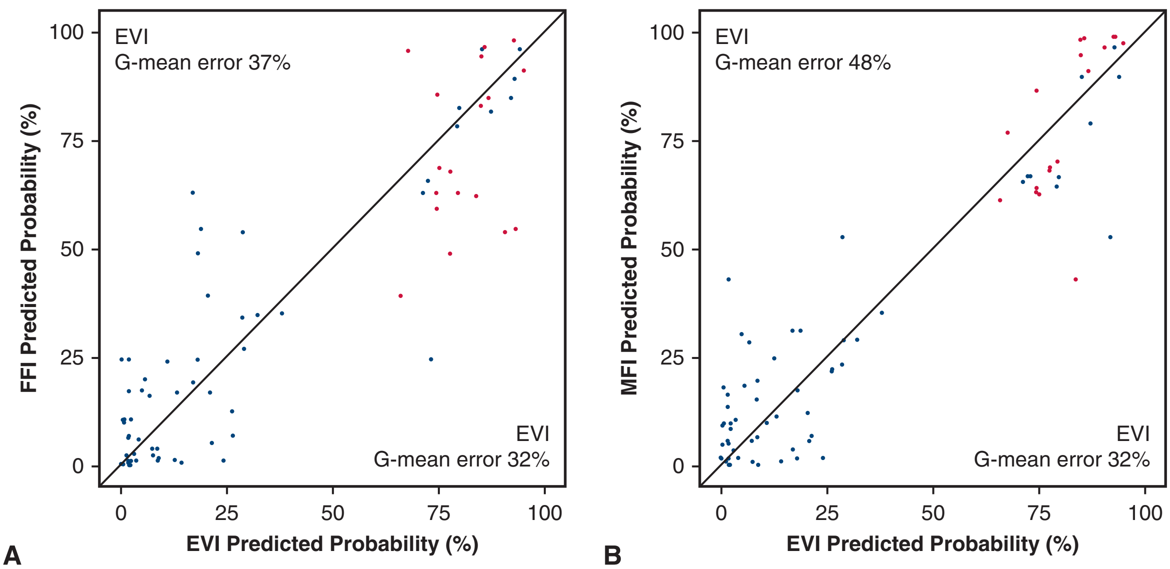


FIGURE 3. Scattergram of predicted probability of developing the composite outcome and G-mean error based on EVI random forest versus FFI random forest (A) and versus MFI from the STS-based comorbidities random forest (B). *Circles* represent each patient with their associated predicted probability for each model and observed outcome. *Solid lines* are line of identity. *Red circles* indicate observed outcome. *Blue circles* indicate observed no outcome. *FFI*, Fried Frailty Index; *EVI*, Esophagectomy Vitality Index; *MFI*, Modified Frailty Index.

Prehabilitation Potential

Frailty is a marker of physiologic fitness and therefore may be modifiable. Although studies have shown that using targeted nutritional and exercise regimens can improve various metrics, their effects on postoperative outcomes are not as well known.[19,20] We have visually identified inflection points in the 4 measured metrics that can serve as tangible target goals for patients to achieve in a prehabilitation program.

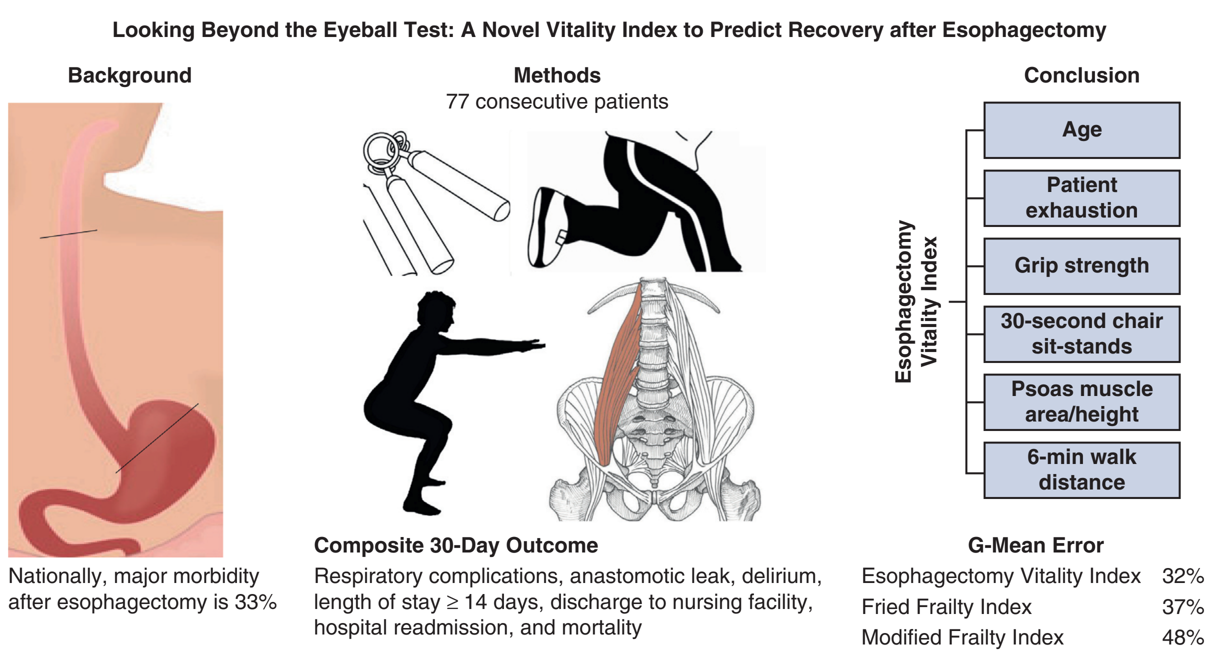


FIGURE 4. Few surgical interventions are of greater magnitude than esophagectomy, which is accompanied by major morbidity of 33% and mortality of 3%. Risk factors for adverse outcomes include "frailty." Thus, to characterize patients' fitness for esophagectomy, this study measured 4 physiologic metrics before esophagectomy: grip strength, number of 30-second chair sit-stands, 6-minute walk distance, and psoas muscle area to height ratio. These, plus age, outperformed commonly used frailty indexes in predicting postesophagectomy mortality and morbidity. The resulting Esophageal Vitality Index provides a robust picture of patients' fitness for surgery beyond the qualitative "eyeball" test.

Limitations

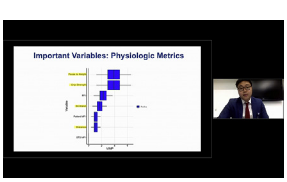
This is a single-institution prospective study representing only a 15-month experience, but is a contemporary series. Patients studied were a select group who already had demonstrated adequate cardiopulmonary function, as assessed by cardiac stress and pulmonary function tests. Our analyses were limited by number of outcome events, and thus we constructed a composite outcome based on frequency and perceived importance of individual postoperative adverse events. Because of this, we could not identify predictors of individual complications. Further, we equally weighted each outcome in the composite for simplicity, without any attempt to hierarchically assign individual weights. If the simple and objective physical status and fitness metrics we measured were added to national quality databases, they may be able to provide more accurate predictions for individual outcomes. It is difficult to know if these metrics are the most predictive frailty ones to use. Others may include nutritional indexes such as recent percentage weight loss, neurocognitive function, or adequacy of social support. Future refined frailty indexes may need to incorporate these as well.

CONCLUSIONS

A simple assessment of frailty, which we call the Esophagectomy Vitality Index, consisting of grip strength, number of 30-second chair sit-stands, 6-minute walk distance, and psoas muscle area to height ratio, combined with standard documented clinical variables, outperformed commonly used frailty models in predicting morbidity after esophagectomy for cancer. These objective measures provide a robust picture of a patient's physical status and fitness for surgery and appear to reflect and look beyond the "eyeball" test.

Webcast

You can watch a Webcast of this AATS meeting presentation by going to: https://aats.blob.core.windows.net/media/20AM/Presentations/Quantifying%20the%20Eye-Ball%20Test_%20A%20Nov.mp4.



Conflict of Interest Statement

Dr Raja is a consultant for Smiths-Medical. All other authors reported no conflicts of interest.

The *Journal* policy requires editors and reviewers to disclose conflicts of interest and to decline handling or reviewing manuscripts for which they may have a conflict of interest. The editors and reviewers of this article have no conflicts of interest.

References

1. Seder CW, Raymond D, Wright CD, Gaissert HA, Chang AC, Becker S, et al. The Society of Thoracic Surgeons general thoracic surgery database 2018 update on outcomes and quality. *Ann Thorac Surg*. 2018;105:1304-7.
2. Velanovich V, Antoine H, Swartz A, Peters D, Rubinfeld I. Accumulating deficits model of frailty and postoperative mortality and morbidity: its application to a national database. *J Surg Res*. 2013;183:104-10.
3. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol A Biol Sci Med Sci*. 2001;56:M146-56.
4. Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I, et al. A global clinical measure of fitness and frailty in elderly people. *CMAJ*. 2005;173:489-95.
5. Subramaniam S, Aalberg JJ, Soriano RP, Divino CM. New 5-factor modified frailty index using American College of Surgeons NSQIP data. *J Am Coll Surg*. 2018;226:173-81.
6. Hanten WP, Chen WY, Austin AA, Brooks RE, Carter HC, Law CA, et al. Maximum grip strength in normal subjects from 20 to 64 years of age. *J Hand Ther*. 1999;12:193-200.
7. Centers for Disease Control and Prevention. Assessment 30-Second Chair Stand. Available at: https://www.cdc.gov/steadi/pdf/STEADI-Assessment-30Sec-508.pdf. Accessed December 1, 2017.
8. Derstine BA, Holcombe SA, Goulson RL, Ross BE, Wang NC, Sullivan JA, et al. Quantifying sarcopenia reference values using lumbar and thoracic muscle areas in a healthy population. *J Nutr Health Aging*. 2017;21:180-5.
9. Casanova C, Celli BR, Barria P, Casas A, Cote C, de Torres JP, et al. The 6-min walk distance in healthy subjects: reference standards from seven countries. *Eur Respir J*. 2011;37:150-6.
10. Radloff L. A self-report depression scale for research in the general population. *Appl Psychol Meas*. 1977;1:385-401.
11. O'Brien R, Ishwaran H. A random forests quantile classifier for class imbalanced data. *Pattern Recognit*. 2019;90:232-49.
12. Tang F, Ishwaran H. Random forest missing data algorithms. *Stat Anal Data Min*. 2017;10:363-77.
13. Ishwaran H. Variable importance in binary regression trees and forests. *Electron J Stat*. 2007;1:519-37.
14. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat*. 2001;29:1189-232.
15. Cafri G, Baily BA. Understanding variable effects from black box prediction: quantifying effects in tree ensembles using partial dependence. *J Data Sci*. 2016;14:67-96.
16. Hara H, van Klaveren D, Kogame N, Chichareon P, Modolo R, Tomaniak M, et al. The "A, B, C" of multiple statistical methods for composite endpoints. *EuroIntervention*. April 28, 2020 [Epub ahead of print].
17. Mascha EJ, Imrey PB. Factors affecting power of tests for multiple binary outcomes. *Stat Med*. 2010;29:2890-904.
18. Gani F, Canner JK, Pawlik TM. Use of the modified frailty index in the American College of Surgeons national surgical improvement program database: highlighting the problem of missing data. *JAMA Surg*. 2017;152:205-7.
19. Dunne MJ, Abah U, Scarci M. Frailty assessment in thoracic surgery. *Interact Cardiovasc Thorac Surg*. 2014;18:667-70.
20. Bolger JC, Loughney L, Tully R, Cunningham M, Keogh S, McCaffrey N, et al. Perioperative prehabilitation and rehabilitation in esophagogastric malignancies: a systematic review. *Dis Esophagus*. 2019;32:1-11.

Key Words: esophagectomy, frailty, esophageal cancer, complications, Esophageal Vitality Index

Discussion

Presenter: Dr Andrew Tang

Dr Shanda H. Blackmon *(Rochester, Minn)*. Thank you, Dr Luketich, I'd also like to take a moment and appreciate Dr Starnes, who's the president of the American Association for Thoracic Surgery (AATS), and congratulate him on celebrating the 100th Annual Meeting of the AATS and for planning such a timely and important session for this General Thoracic Breakout. As I look back on last year's meeting that took place in Toronto, I discovered an entire session at that meeting was dedicated to similar topics on frailty and preoperative rehabilitation as they relate to cardiac surgery. It's timely that we now focus our attention on similar topics as they relate to the practice of general thoracic surgery.

I believe thoracic surgeons should take an inventory of the preoperative condition of our patients and their ability to tolerate the surgeries we offer them. In the era of ERAS, prehab, mobile applications, video visits, and telehealth, forward-thinking general thoracic surgeons will no doubt start to formalize the way we optimize our patients to minimize what is now termed, according to last year's AATS meeting, a core functional survival. No longer will we merely look at survivorship and the surgery to be acceptable; we will now look at the quality of life after surgery.

As patients advocate for themselves, they now look for programs that offer them something beyond a simple surgery and rolling the dice. Assessing patients for baseline vulnerability, including their cognitive function, frailty, preoperative nutrition, age, and psychosocial stressors, I believe should now be the standard of care. Further assessing a patient for frailty, including things like wasting, vulnerability, cognitive impairment, activities of daily living, slowing, social situations, malnutrition, depression, weakness, and other comorbidities as well as mental health, can now be performed by offering patients increasingly popular apps, like the one that was discussed at the AATS meeting, the iOS Android frailty tool, created by Dr Jonathan Afilalo, and this is available free.

It identifies patients who would have a predicted drop below a functional level that's recoverable or reasonable and will help them make the best choice about having surgery or delaying to optimize. Once these patients have been identified, prehab, which is a conglomeration of preoperative nutritional optimization, coaching, strengthening, and supporting patients so that they are optimized before surgery, can no doubt improve outcomes. I'm pleased to see that AATS is placing emphasis on such matters.

Reviewing Dr Tang and colleagues' manuscript on quantifying the eyeball test, a novel vitality index to predict recovery after esophagectomy, I applaud the senior author, Dr Sid Murthy, for leading an important investigation, and I'd like to thank them for sending me both versions of the manuscript ahead of time for review. I have only a few questions.

My first question is: Your conclusion states that your esophageal vitality index outperformed other frailty models to predict morbidity following esophagectomy for cancer. Using the Fried frailty index, which is based on 5 qualities, and the Modified Frailty Index, which is 11 comorbidities from the Society of Thoracic Surgeons database for comparison, my question for you is: Why did you choose those and not generalizable apps, like the frailty tool that I mentioned, or the frailty app, or the clinical frailty scale, or something like the frailty wheel? These are more widely used and are readily available and more generalizable beyond patients that are just having esophagectomy. The comparisons you elected to make were based on indices that have to be performed by an institution or a provider, and cannot be performed by the patient readily. How and why did you choose these 2 indices for comparison?

Dr Andrew Tang *(Cleveland, Ohio)*. Thank you, Dr Blackmon, for your questions and for reading over my manuscripts. When we were designing the study originally, we looked at the multitude of publications on frailty across all specialties, and it seems that the common denominator was the Fried Frailty Index, which was one of the first ones to be developed in the early 1990s as part of the cardiovascular health study through Johns Hopkins. A lot of the other studies base the definition of frailty on how Dr Fried at Hopkins had defined it. So, that's the reason we chose that index.

Specifically, for the Modified Frailty Index, the reason we looked at that was because it was a relatively simple one that has been widely published across disciplines including colorectal surgery, general thoracic surgery, urology, and orthopedic surgery, and we felt that we wanted to provide a spectrum. The Fried Frailty Index was the original, and the Modified Frailty Index was a relatively simple count of comorbidities. We wanted to show that our index can be used as a generalizable way of measuring physiologic status to predict a multitude of outcomes for a multitude of procedures, because we're not boxing patients into 3 separate buckets such as "not frail," "intermediately frail," or just "frail." What we're trying to say is that using these 4 basic metrics, if your goal is to look at, for example, leak after esophagectomy, you can say that a patient who was able to walk 300 m or more is less likely to develop a leak, or a patient who had a grip strength of more than 40 kg was less likely to develop delirium.

Unfortunately, with 77 patients, we didn't have enough events to model those events individually, which is why we think it is worthwhile to build this into a larger database so that moving forward we can tell our patients, "You walked 500 meters today, so based off the multitude of patients who just had their esophagectomy done in Pittsburgh, your risk of a leak is much lower than how patients did there."

Dr Blackmon. My second question is a plead to change the title of your novel index. You, in your manuscript called it the "esophageal vitality index." Some people might think that they're talking about the vitality of the esophagus. Instead, shouldn't you call it the esophagectomy patient vitality index, recognizing that this is an assessment of the patient and not the esophagus?

Dr Tang. Yes, I think that is a fair point. We definitely do not want to give people the sense that we only care about their esophagus. We want them to understand that this is a big operation and to make sure that they're strong enough to undergo it. So yes, we will definitely consider changing that to make it a bit more user-friendly.

Dr Blackmon. My third and final question is: Why in your assessment did you not include patients who were assessed for fitness of esophagectomy and then not selected for esophagectomy? I think that would be the most value. Looking at people that were assessed by a surgeon, and either because of a comorbidity evaluation, or some other conglomeration. If a young surgeon with no experience says "I think that patient might be high risk," versus an elderly surgeon who's had a whole lifetime of experience assessing patients and looking at them and decided that they weren't, and looking at your assessment model to determine if that correlated with other people deciding that perhaps this patient is not fit. I was just curious if you plan on looking at that, or if you did look at it and just didn't include it in your manuscript.

Dr Tang. Thank you, Dr Blackmon. You hit the nail on the head. For this specific study, we wanted to identify which physiologic metrics were most important and how they're important first, and we included the patients who underwent esophagectomy because we wanted to show that those patients who are weaker than this, or walked fewer steps than this, were more likely to develop a complication. This was more to give us the idea that, okay, these are the things that are easily reproducible and well validated, and they work.

Moving forward, we will start to evaluate all consult patients with this measure to set up a baseline that will serve 2 purposes. Number one, it will help us better stratify who's at greater risk for surgery and who's at lower risk for surgery, but it will also give us a very easy-to-understand bar to set for the patient. "Today you were able to walk 250 m. By the time you're finished with your induction therapy, and by the time you've regained your strength and your nutrition, I want to see you be able to walk x amount, because we know that this will decrease your risk of complication." You hit the nail head because this is meant for all clinicians, not just the ones who have a multitude of years of experience, where all it takes is a handshake, or all it takes is watching the patient get up and say, "I think they might be high risk," or "I think they might be able to pass." So, the purpose of this is to make it easy for anyone to use.

Dr Blackmon. Did you find that you were able to look at which patients were selected for salvage esophagectomy versus pre-emptive esophagectomy 4 to 6 weeks after treatment?

Dr Tang. We did not look at that specifically. We actually are looking back now through our institutional data at our salvage esophagectomy.

Dr Blackmon. That would be a nice group to look at.

Dr Tang. Certainly.

Dr Blackmon. Congratulations.

Dr Sudish Murthy *(Cleveland, Ohio)*. Thanks very much for the nice commentary. I think those are very valid and important points you bring up, no question. I think one of the key things you have brought up on numerous occasions, whether it's here or elsewhere, is informing the patient. Some of these issues that we are trying to quantitate, we really aren't just quantitating for ourselves, but also for the patients. Very much as you point out, to give them a better sense of what they might expect, what they should anticipate, etc. to help us make a more informed decision for both of us. I think that is critical, and I think it's a very important point, and we appreciate you sharing those thoughts with us.

Dr Shaf Keshavjee *(Toronto, Ontario, Canada)*. Congratulations on a very nice study. I think it's a very important area because there are frailty experts out there who have built careers in studying frailty, and they have created all kinds of frailty indices. I think many of these indices have received unwarranted validity, if you will; in that just because they're a quantitative number they garner more respect—if it's quantitative it must be right. That becomes a self-fulfilling prophecy, which I think isn't helping the clinician. As Shanda alluded to, experience helps and is critically important, Griff Pearson used to call it the "foot of the bed test," similar to you calling it the "eyeball test." We see it all the time in major surgery such as esophagectomy or even lung transplant too. On paper, the patient looks like a disaster and you go look at them in the bed and say, "I can pull them through a lung transplant," and we do. Attempting to accurately quantify that is very important, and it is really important in all of surgery because the referring doctors are pulling out their calculators, calculating a "frailty index" and then saying, "Well, I think this patient is too frail for surgery, maybe we will just send him for radiation," and those decisions are being made. I don't care if a patient's frail or old, I care if they're fit for surgery or not, and whether I can get them to a meaningful life on the other side of surgery. I think trying to quantify that surgical judgment is very, very important.

Dr Tang. Yes, thank you very much. That was the purpose of this, so hopefully that came across clearly. Thank you very much.

Dr James D. Luketich *(Pittsburgh, Pa)*. Regarding the choice of outcomes, did you think about attempting to look at other metrics like a quality of life outcome, etc.? Some kind of return to premorbid status, such as, "When can I go back to work?" "When can I really get back to my normal life?" Did you begin to address these factors?

Dr Tang. We haven't yet, but we still have these patients who we follow regularly. Moving forward, I'd like to know, not just immediately after operation, "what's your quality of life," but even a year or 2 years out. Each variable may predict a different thing. What we found was whenever we looked at this, and we also performed this for patients who underwent a lobectomy or pneumonectomy, it just depends on what we're specifically looking at. If I'm looking at length of stay, the walk distance might be more important. If I'm looking at just activities of daily living, it might have been that the sit-stands are more important. So, moving forward, we're going to go back and talk to these patients and say, "Okay, now that you've recovered, what's your quality of life a year out, 6 months out from your operation?" We're going to see how well our frailty index correlates with that.

Dr Luketich. Along the lines of Sid's comments, I think there's a bit of a misconception that esophagectomy may lead to never being able to eat normally, or constant dumping, or other issues related to quality of life. So, I think it will be important to look at that, but very nice presentation.

APPENDIX E1. VARIABLES CONSIDERED IN ANALYSES

Patient Characteristics

Male, age (years), race, Hispanic, body mass index (kg/m^2), albumin (g/dL), creatinine (mg/dL), hemoglobin (g/dL), hematocrit (%), weight loss (kilograms), weight loss over last 3 months (kilograms), Center for Epidemiological Studies Depression scale question 1, Center for Epidemiological Studies Depression scale question 2, previous cancer, renal disease, previous cardiothoracic surgery, American Society of Anesthesiologists physical status, history of smoking, anticoagulation use, steroid use, narcotic use, alcohol use, psychiatric history.

Cardiopulmonary Function

Forced expiratory volume in 1 second (% of predicted), lung diffusion test (% of predicted), ejection fraction (%).

Cancer Details

Clinical tumor size, tumor grade, neoadjuvant chemotherapy, neoadjuvant radiation, histology, interval from preoperative visit to surgery (days).

Esophagectomy Vitality Index

Grip strength (kilograms), 30-second chair sit-stands (number), psoas muscle area to height ratio (cm^2/m), 6-minute walk distance (meters).

Fried Frailty Index

Modified Frailty Index (patient reported). Diabetes mellitus, congestive heart failure, hypertension, transient ischemic attack or cerebrovascular accident, dependent functional status, myocardial infarction, peripheral arterial disease, cerebrovascular accident with neurologic deficits, chronic obstructive pulmonary disease, coronary artery disease, dementia, total score.

Modified Frailty Index (Society of Thoracic Surgeons–abstracted). Diabetes mellitus, congestive heart failure, hypertension, transient ischemic attack or cerebrovascular accident, dependent functional status, myocardial infarction, peripheral arterial disease, cerebrovascular accident with neurologic deficits, chronic obstructive pulmonary disease, coronary artery disease, dementia, total score.

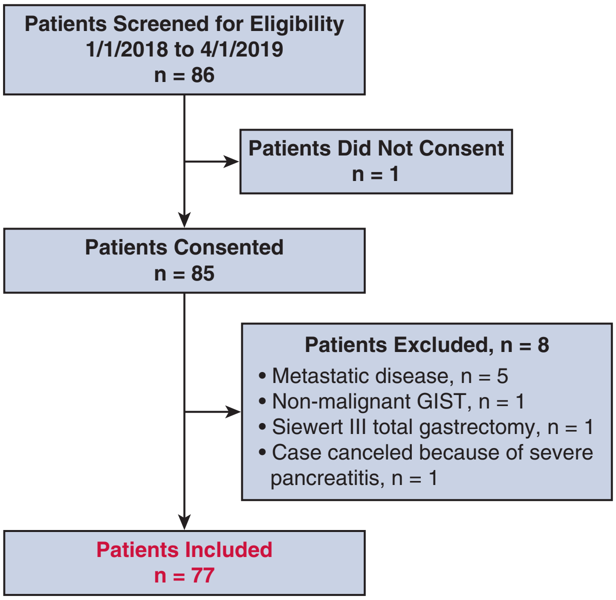


FIGURE E1. STROBE diagram of study group. *GIST*, Gastrointestinal stromal tumor.

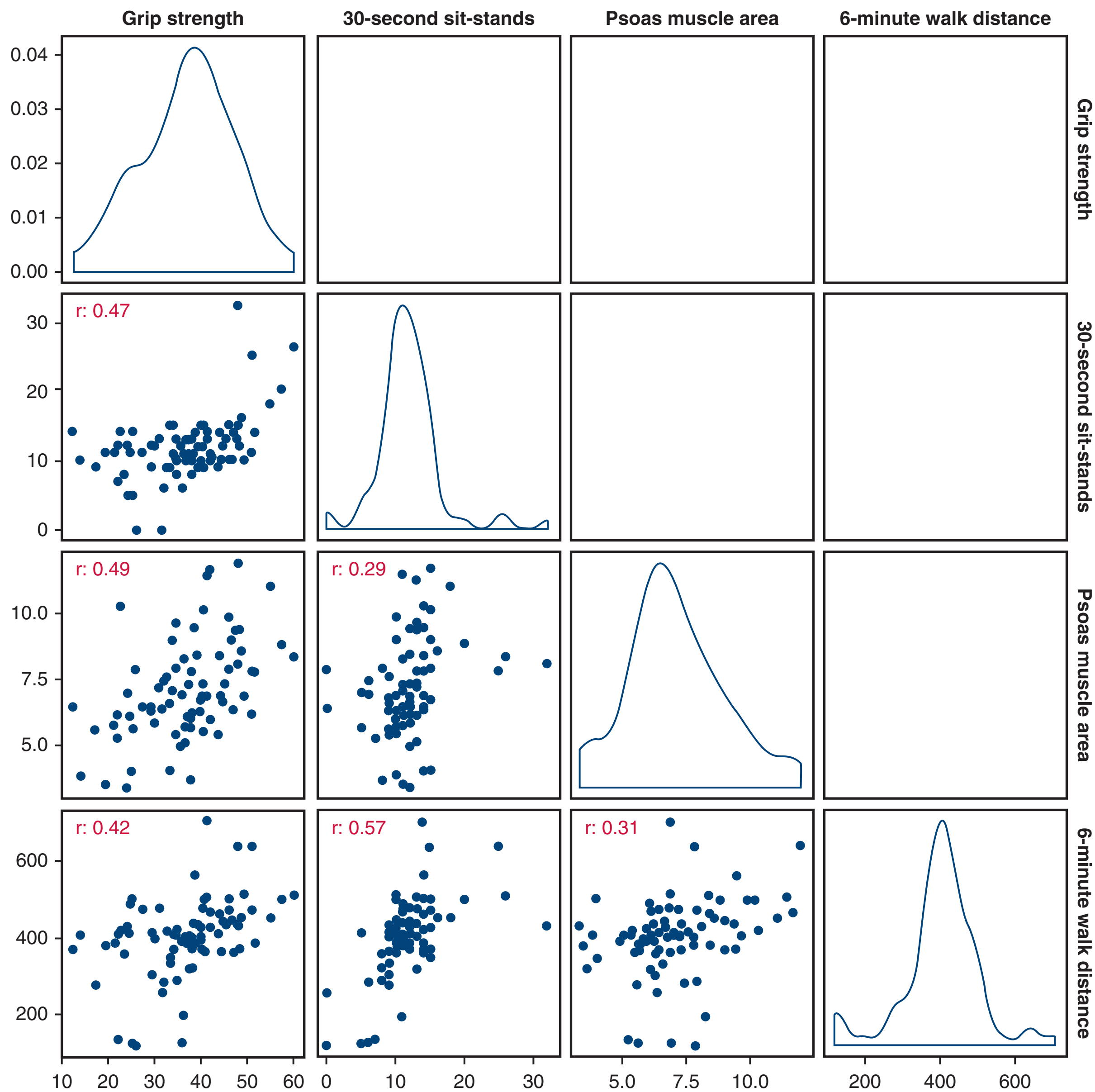


FIGURE E2. Distribution of measured frailty metrics and correlation between measured frailty metrics: grip strength, 30-second chair sit-stands, psoas muscle area/height ratio, and 6-minute walk distance. Scatterplots (*bottom left half*) and distribution semiparametric plots (*along the diagonal*) show the distribution of values for each frailty metric. Pearsons' correlation coefficient (r) between the frailty metrics are displayed in the *top left corner* of each scatterplot.

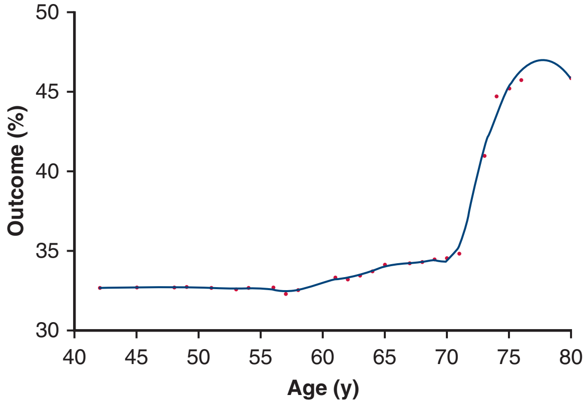


FIGURE E3. Risk-adjusted partial dependency plot based on the Esophagectomy Vitality Index estimating the risk of a patient developing the composite outcome after esophagectomy versus age at esophagectomy. *Red circles* indicate risk-adjusted estimates, and *blue line* indicates smooth fit of risk-adjusted estimates.

TABLE E1. Fried Frailty Index

Criteria	Frailty status			
Shrinking	Frailty cut point: Baseline: Self-reported unintentional weight loss ≥10 lb in previous year Follow-up: Unintentional weight loss ≥5% of previous year's body weight OR Body mass index <18.5 kg/m^2			
Physical endurance/exhaustion	Geriatric Depression Scale: 1. How often would you say I felt everything I did was an effort in the past week? 2. How often would you say I could not get going in the past week? Response options: <1 d, 1-2 d, 3-4 d, most of the time Frailty cut point: Either 3-4 d or most of the time			
Low physical activity	Frequency of mildly energetic, moderately energetic, and very energetic physical activity. Response options: ≥3 times per week, 1-2 times per week, 1-3 times per month, hardly ever/never Frailty cut point: Hardly ever/never for very energetic			
Weakness	Hand grip strength in kg: GRIP-D hand held dynamometer, dominant hand, average of 3 measures. Frailty cut point: Grip strength: lowest 20% (by sex and BMI)			
	Men (BMI)	**kg**	**Women (BMI)**	**kg**
	≤24	≤29	≤23	≤17
	24.1-26	≤30	23.1-26	≤17.3
	26.1-28	≤30	26.1-29	≤18
	>28	≤32	>29	≤21
Slow walking speed	Walking time in seconds (usual pace) over 15 feet Frailty cut point: Slowest 20%, stratified by sex and median standing height.			
	Men (height)	**Time**	**Women (height)**	**Time**
	≤173 cm	≥7 s	≤159 cm	≤7 s
	>173 cm	≥6 s	>159 cm	≥6 s
	OR Time to complete "timed up and go test" (TUG) Frailty cut point: TUG time ≥19 s			

Frail = ≥3 criteria present; intermediate or pre-frail = 1 or 2 criteria present; not frail = 0 criteria present. *BMI*, Body mass index. Adapted from Fried and colleagues.[3]

TABLE E2. Modified Frailty Index

Comorbidities (1 point each)
Diabetes mellitus
Hypertension requiring medication
Coronary artery disease
Myocardial infarction
Congestive heart failure
Chronic obstructive pulmonary disease or pneumonia
Peripheral vascular disease or rest pain
Transient ischemic attack or cerebrovascular attack
Cerebrovascular attack with neurologic deficit
Dementia
Functional status 2 (not independent)
Total: 0-11 points

TABLE E3. Concordance of Modified Frailty Index between patient-reported and STS database–recorded comorbidities

Comorbidities	No. (% of 77)
Diabetes mellitus	74 (96)
Hypertension	64 (83)
Coronary artery disease	74 (96)
Myocardial infarction	74 (96)
Congestive heart failure	73 (95)
Chronic obstructive pulmonary disease	64 (83)
Peripheral arterial disease	71 (92)
Cerebrovascular accident	76 (98)
Cerebrovascular accident with deficits	76 (98)
Dementia	75 (97)
Dependent functional status	73 (95)
Patients whose self-reported and STS-abstracted comorbidities were 100% identical	11 (14)

Concordance of Modified Frailty Index between patient-reported comorbidities and comorbidities abstracted by registry nurses according to the STS database definitions from the electronic medical record. Concordance means that either "yes" or "no" from query of the patient at preoperative visit is identical to that abstracted by registry nurses. *STS*, Society of Thoracic Surgeons.