Assessment of Thoracic Endografting Operative Mortality Risk Score: Development and Validation in 2,000 Patients

Arman Kilic, MD, Ibrahim S. Sultan, MD, George J. Arnaoutakis, MD, Robert S. D. Higgins, MD, MSHA, and Ahmet Kilic, MD

Department of Surgery, The Johns Hopkins Medical Institutions, Baltimore, Maryland; and Division of Cardiac Surgery, Department of Surgery, The Ohio State University Wexner Medical Center, Columbus, Ohio

Background. In this study we derive and validate a composite risk index termed the Assessment of Thoracic Endografting Operative Mortality, or ATOM, risk score.

Methods. All thoracic endovascular aortic repairs (TEVAR) in the American College of Surgeons National Surgical Quality Improvement Project (NSQIP) between 2005 and 2012 were identified. The primary outcome was operative mortality. After evaluating the association of over 60 preoperative variables and operative mortality in univariate analysis, a multivariable model was developed. Significant risk factors were assigned points equivalent to their odds ratio rounded to the nearest whole integer in the final multivariable model.

Results. Overall, 1,981 patients comprised the study population, including 1,486 (75.0%) in the derivation and 495 (25.0%) in the validation cohort. There were 173 (8.7%) operative mortalities. A 30-point risk score incorporating 10 risk factors was generated and found to be highly predictive of operative mortality in the derivation (odds ratio [OR] 1.36, $p < 0.001$) and validation cohorts (OR 1.24, $p < 0.001$). The models used to create and validate the ATOM score were robust (C indices 0.84 and 0.83, respectively). There was strong correlation between predicted mortality rates based on the derivation cohort and actual mortality rates in the validation cohort ($r = 0.75$, $p < 0.001$). Operative mortality based on low (ATOM < 5), moderate (ATOM 5 to 9), and high risk (ATOM ≥ 10) was 1.3%, 6.6%, and 24.0%, respectively ($p < 0.001$). Higher ATOM scores also correlated with higher complication rates and longer hospital stays.

Conclusions. The ATOM score is a significant predictor of operative mortality in TEVAR and can be used for preoperative risk stratification.

(Ann Thorac Surg 2015;100:860–7)

Open surgical repair for diseases involving the descending thoracic aorta was for many years the standard of care since the original description by Debakey and Cooley in 1953 [1]. Although these repair techniques offer durable results and improved long-term survival to patients, the operation itself carries significant risk of morbidity. Furthermore, even in experienced centers operative mortality rates for open repair of descending thoracic and thoracoabdominal aneurysms are 5% to 15% [2–4]. These mortality rates are undoubtedly a reflection of an increasingly older patient population with significant comorbidities who cannot tolerate the stress of a major morbidity after a substantial operation.

In this sense, less invasive approaches to treating the descending thoracic aorta seem particularly appealing. The United States Food and Drug Administration first approved the use of thoracic endovascular aortic repair (TEVAR) in 2005 [5]. Over the last decade the utilization of TEVAR in the treatment of both aneurysms and dissections involving the descending thoracic aorta has increased significantly [6, 7]. Although risk factors for operative mortality after TEVAR have been previously studied, there currently exists no preoperative risk stratification tool. In this study we derive and validate a composite risk index for operative mortality in TEVAR.

Patients and Methods

Data Acquisition

The American College of Surgeons National Surgical Quality Improvement Project (NSQIP) Participant Use Files from 2005 to 2012 were used for this study. The NSQIP provides Health Insurance Portability and Accountability Act compliant information, providing de-identified patient-level data on demographics, procedural details, and operative outcomes [8]. The number of overall cases and participating sites in NSQIP has increased with each participant use file, from 152,490

Accepted for publication Jan 13, 2015.

Address correspondence to Dr Kilic, The Ohio State University Wexner Medical Center, N-816 Doan Hall, 410 W 10th Ave, Columbus, OH 43210; e-mail: ahmet.kilic@osumc.edu.

The Appendix can be viewed in the online version of this article [http://dx.doi.org/10.1016/j.athoracsur.2015.01.040] on http://www.annalsthoracicsurgery.org.

Published by Elsevier

0003-4975/$36.00
http://dx.doi.org/10.1016/j.athoracsur.2015.01.040

cases from 121 sites in the joint 2005 to 2006 file to 543,885 cases from 374 participating sites in the 2012 file [8]. Because the data are de-identified and publically available, the Institutional Review Board granted this study exempt status.

Study Population

All TEVARs in the NSQIP database between 2005 and 2012 were identified using the Common Procedural Terminology codes 33870 and 33881. Patients with an International Classification of Diseases, Ninth Edition (ICD-9) code for descending thoracic aortic aneurysm (441.1, 441.2, 441.6, or 441.7) or descending thoracic aortic dissection (441.01 or 441.03) were included. Ruptured cases were identified by ICD-9 codes 441.1 or 441.6. Trauma-related cases were identified by the ICD-9 code 901.0. The extent of the aneurysm or dissection was classified using ICD-9 codes 441.01, 441.2, and 441.2 for isolated descending thoracic and 441.03, 441.6, and 441.7 for thoracoabdominal extension.

Risk Score Generation

The study population was randomly divided in a 3:1 fashion into a derivation and validation cohort. The primary outcome was operative mortality defined by NSQIP as occurring within 30 days of the operation. In the derivation cohort, univariate logistic regression analysis was conducted to evaluate the association between each variable and operative mortality. All variables that were evaluated are shown in the Appendix.

Variables were entered into a multivariable logistic regression model for operative mortality if they were associated with operative mortality in univariate analysis ($p < 0.20$) and had less than 10% missing data. The multivariable model was evaluated using the area under the receiver operating characteristic curve, or C index, Akaike information criterion, and Hosmer-Lemeshow goodness-of-fit. Continuous variables were categorized using clinical meaningful thresholds or Lowess smoothing plots to identify linear breakpoints. Points were assigned to significant risk factors ($p < 0.05$) based on their odds ratios rounded to the nearest whole integer. The composite risk score termed the Assessment of Thoracic Endografting Operative Mortality, or ATOM, was generated by adding the individual risk points, and a probability curve for operative mortality was generated using this composite score.

Validation of the Risk Score

The multivariable model used to generate the ATOM score was applied to the validation cohort and the same statistical methods were used to evaluate the robustness of the model in this patient subset. The C index of the composite score was calculated as well. Weighted linear regression analysis was used to evaluate the association between predicted operative mortality rates in the derivation cohort based on ATOM score and actual mortality rates in the validation cohort for the same ATOM score, with weights assigned to the number of patients with each score. Categories of risk (low, moderate, high) based on score were also generated to make the application of the ATOM score more useful clinically, and operative mortality rates were compared between these categories. Secondary outcomes, including rates of pneumonia, reintubation, pulmonary embolism, prolonged ventilation greater than 48 hours, new-onset dialysis, urinary tract infection, wound infection, stroke, cardiac arrest, myocardial infarction, bleeding requiring transfusion, and sepsis were also compared between categories, as was the length of hospitalization. All continuous data are presented as mean ± standard deviation and all categoric data as number (percentage). Data analyses were performed with version 11 STATA statistical software (StataCorp LP, College Station, TX).

Results

Baseline Data

There were 1,981 patients who underwent TEVAR during the study period and were included in the analysis. The mean age was 69.2 ± 13.1 years, with 875 (44%) females. Demographics and comorbidities were comparable between the derivation and validation cohorts with the exception of transfusion in the last 3 days prior to TEVAR (Table 1). With respect to preoperative laboratory parameters, the validation cohort had statistically higher sodium, albumin, and hematocrit levels although the absolute differences were small and clinically insignificant (Table 1). Most patients presented with aneurysms without dissection (n = 1,468; 74%) and with disease limited to the descending thoracic aorta (n = 1,567; 79%). A total of 372 (19%) TEVARs were considered emergent. These disease-related, operative, and other characteristics were comparable between the derivation and validation cohorts (Table 1).

Development of Risk Score

Operative mortality occurred in 173 (9%) patients. In the derivation cohort 29 variables met entry criteria for the multivariable model based on univariate analysis, and of these 10 were found to be significant risk factors in multivariable analysis: age 70 or greater years, body mass index less than 30 kg/m^2, chronic obstructive pulmonary disease, requiring total functional assistance, blood urea nitrogen greater than 25 mg/dL, white blood cell count greater than 12,000 cells/μL, emergency operation, left subclavian artery coverage, thoracoabdominal extension, and mesenteric debranching (Table 2). The multivariable model performed well, with a C index of 0.84, Akaike information criterion of 604, and a nonsignificant Hosmer-Lemeshow goodness-of-fit p value of 0.89, suggesting the data were a good fit. The ATOM score was out of 30 possible points (Table 2).

Predicted Operative Mortality Based on ATOM Score in the Derivation Cohort

The mean ATOM score in the derivation cohort was 6.5 ± 3.5, with a median of 6 and range of 0 to 21. There was a wide range of predicted operative mortality, from 0.8%