STS-PROM score (Class I) is preferred over the EuroSCORE II (Class IIb) to estimate in-hospital mortality.⁶

Mori and colleagues show that one size certainly may not fit all, and that a critical examination and understanding of the risk prediction models that we use daily is necessary to convey the most accurate message to our patients. Although risk calculators help supplement clinical decision making, the heart team approach is central to optimizing patient selection and ensuring that patients undergo the most appropriate treatment.⁷

References

1. Bilimoria KY, Liu Y, Paruch JL, Zhou L, Kmieciak TE, Ko CY, et al. Development and evaluation of the universal ACS NSQIP surgical risk calculator: a decision aid and informed consent tool for patients and surgeons. *J Am Coll Surg*. 2013;217:833-42. e1-e3.
2. Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 1—coronary artery bypass grafting surgery. *Ann Thorac Surg*. 2009;88(1 Suppl):S2-22.
3. O'Brien SM, Shahian DM, Filardo G, Ferraris VA, Haan CK, Rich JB, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 2—isolated valve surgery. *Ann Thorac Surg*. 2009;88(1 Suppl):S23-42.
4. Ad N, Holmes SD, Patel J, Pritchard G, Shuman DJ, Halpin L. Comparison of EuroSCORE II, Original EuroSCORE, and The Society of Thoracic Surgeons Risk Score in cardiac surgery patients. *Ann Thorac Surg*. 2016;102:573-9.
5. Mori MS, Shahian DM, Huang C, Li S, Normand ST, Geirsson A, et al. Surgeons: buyer beware – does 'universal' risk prediction model apply to patients universally? *J Thorac Cardiovasc Surg*. 2020;160:176-9.e2.
6. Neumann F-J, Sousa Uva M, Ahlsson A, Alfonso F, Banning AP, Benedetto U, et al. 2018 ESC/EACTS Guidelines on myocardial revascularization. *Eur Heart J*. 2019;40:87-165.
7. Patterson T, McConkey HZR, Ahmed-Jushuf F, Moschonas K, Nguyen H, Karamasis GV, et al. Long-term outcomes following heart team revascularization recommendations in complex coronary artery disease. *J Am Heart Assoc*. 2019;8:e011279.

See Article page 176.

Commentary: The risk of risk models

Paul Kurlansky, MD

Surgical risk prediction models have had and will continue to have a major impact on surgical decision-making and quality improvement efforts.² However, all risk models are subject to limitation. Perhaps the most important but least discussed limitation of surgical risk models is the fact that they are built on patients who actually underwent surgery, rather than patients who were potentially evaluated for surgery. In other words, every case in the database represents a patient whom a surgeon judged appropriate for surgery. Those who might share the same exact risk profile according to the variables identified in the model but were turned down based on frailty, social situation, or other factors not captured, are not included. The

Scatterplots demonstrating correlation of different cardiac surgical risk models.¹

CENTRAL MESSAGE

As useful as they have become, all surgical risk models are subject to limitations. The more precise the methodology, the more likely the model will be to provide useful information.

model-derived predicted risk for a prospective patient may not be well correlated with the actual surgical risk for that specific patient. Therefore, risk models tend to be more useful for evaluating surgery performed than for selecting appropriate surgical candidates.

Within this arena of surgical risk prediction, Mori and colleagues³ raise an interesting question: does the added power that might be gained by increasing sample size of a combined “universal” risk prediction model misrepresent

From the Division of Cardiac Surgery, Department of Surgery, Columbia University, New York, NY.

Disclosures: Author has nothing to disclose with regard to commercial support.

Received for publication Dec 15, 2019; revisions received Dec 15, 2019; accepted for publication Dec 16, 2019; available ahead of print Jan 7, 2020.

Address for reprints: Paul Kurlansky, MD, Division of Cardiac Surgery, Department of Surgery, Columbia University, Black Building 210, 650 West 168th St, New York, NY 10032 (E-mail: Pk2245@cumc.columbia.edu).

J Thorac Cardiovasc Surg 2020;160:181-2

0022-5223/\$36.00

Copyright © 2020 by The American Association for Thoracic Surgery

<https://doi.org/10.1016/j.jtcvs.2019.12.063>

the performance of its individual component parts? Although their hypothetical demonstration strongly suggests the need to provide not only the overall model performance, but the procedure-specific performance of such generalized models, several caveats are worthy of note. First, it is actually quite impressive for those of us who advocate the benefit of clinical rather than administrative data sources in the construction of these models to note the very favorable performance (c-statistic and Brier score—discrimination and calibration) of the models generated from the National Inpatient Survey administrative data. Second, as noted by the authors, the discrimination of their hypothetical model may have been somewhat inflated by the heterogeneity of including such dramatically different patient populations as those chosen (although such heterogeneity might not be all that different from what is found in the National Surgical Quality Improvement Program database). Unfortunately, what the authors do not reveal is what, if any, procedure-specific interaction terms were entered into their models that may have more adequately accounted on the analytical side for the heterogeneity of procedures included. Lastly, as might be expected, the model performance was least favorable for the least common procedure—esophagectomy. However, there is no comparison with how this “universal” model performed compared with

an esophagectomy-specific model. Had appropriate interactions been applied to the universal model, would there really have been a difference in performance?

In short, one of the advantages of the “universal” approach—the ability to provide information regarding less common procedures—may also be its greatest limitation. Whether pooling data to leverage potentially common risk factors is preferable to attempting an independent model cannot be answered by the analysis by Mori colleagues. Perhaps this is an area in which emerging machine-learning techniques, with their ability to identify unanticipated associations, may be of increasing assistance in the future.

References

1. Ad N, Holmes SD, Patel J, Pritchard G, Shuman DJ, Halpin L. Comparison of EuroSCORE II, Original EuroSCORE, and The Society of Thoracic Surgeons Risk score in cardiac surgery patients. *Ann Thorac Surg.* 2016;102:573-9.
2. Blackstone EH, Swain J, McCardle K, Adams DH, Governance Committee, American Association for Thoracic Surgery Quality Assessment Program. A comprehensive American Association for Thoracic Surgery quality program for the 21st century. *J Thorac Cardiovasc Surg.* 2019;158:1120-6.
3. Mori MS, Shahian DM, Huang C, Li S, Normand ST, Geirsson A, et al. Surgeons: buyer beware – Does “universal” risk prediction model apply to patients universally? *J Thorac Cardiovasc Surg.* 2020;160:176-9.e2.