HHS Public Access
Author manuscript
J Am Coll Surg. Author manuscript; available in PMC 2023 August 03.

Published in final edited form as:
J Am Coll Surg. 2023 August 01; 237(2): 385–386. doi:10.1097/XCS.0000000000000676.

Improving the American College of Surgeons NSQIP Surgical Risk Calculator with Machine Learning

Kevin A Chen, MD,
Shawn M Gomez, EngScD,
Muneera R Kapadia, MD, MME, FACS
Chapel Hill, North Carolina

Dear Dr. Eberlein,

We read with interest the recent publication by Liu et al. which validates the work of multiple previous researchers showing superior accuracy for a machine learning algorithm, gradient boosting (XGB), compared with traditional logistic regression (1-4) in predicting post-operative complications. We agree that the improved accuracy afforded by XGB algorithm supports its use as the American College of Surgeons (ACS) Surgical Risk Calculator.

However, we suggest three important points for consideration. First, the authors maintain that a single model for all operations is better than procedure-specific models for simplicity, ease of use, and updating. However, this approach masks some deficiencies in predicting procedure-specific outcomes. For example, while the model may show high discriminatory accuracy between procedures (i.e. risk of surgical site infection (SSI) between hernia repair and colectomy), it shows poor accuracy within specific procedures, which is far more useful to individual patients and surgeons. In the initial publication describing the ACS Surgical Risk Calculator, the area under the receiver operating characteristic curve was 0.82 for predicting SSI overall but dropped to 0.65 for colectomy specifically (5). This discrepancy has been validated in external cohorts (6). We hope that procedure-specific model development is considered for future iterations of the ACS Surgical Risk Calculator.

Second, the authors note that one downside of XGB is that it is less interpretable compared with logistic regression. We agree but note that important advancements have been made in machine learning interpretability and that specific approaches, especially Shapley additive explanations, can be very helpful in interpreting XGB. In fact, both the European Union General Data Protection Regulation (GDPR) and the United States Food and Drug Association (FDA) have emphasized the importance of machine learning interpretability for decision-support tools in healthcare.

Third, and most importantly, the authors should publish their code. This allows external machine learning researchers to learn from, improve, and validate the ACS Surgical Risk Calculator. Making the code public would help the ACS overcome the limitations of having a relatively small research team in developing procedure-specific models and allow input from machine learning interpretability specialists. Open-source development is a broadly successful approach, even used by XGB itself (https://github.com/dmlc/xgboost). Because of these factors, publishing code is recommended by the International Committee of Medical Journal Editors (ICJME), which the Journal of the American College of Surgeons endorses.

In summary, we applaud Liu et al. in adding credence to the mounting evidence showing improved predictive accuracy for machine learning approaches. We agree that XGB is a suitable replacement for the current ACS Surgical Risk Calculator. However, we believe that open-source development of procedure-specific models, with interpretability analyses, will be of greater benefit to the surgeons and patients who use the ACS Surgical Risk Calculator's predictions.

References:

1. Liu Y, Ko CY, Hall BL, et al. ACS NSQIP Risk Calculator Accuracy Using a Machine Learning Algorithm Compared to Regression. J Am Coll Surg 2023;E-published.
2. Bonde A, Varadarajan KM, Bonde N, et al. Assessing the utility of deep neural networks in predicting postoperative surgical complications: a retrospective study. Lancet Digit Heal 2021;3:e471–e485.
3. Bihorac A, Ozrazgat-Baslanti T, Ebadi A, et al. MySurgeryRisk: Development and Validation of a Machine-learning Risk Algorithm for Major Complications and Death After Surgery. Ann Surg 2019;269:652–662. [PubMed: 29489489]
4. Bertsimas D, Dunn J, Velmahos GC, et al. Surgical Risk Is Not Linear: Derivation and Validation of a Novel, User-friendly, and Machine-learning-based Predictive {OpTimal} Trees in Emergency Surgery Risk ({POTTER}) Calculator. Ann Surg 2018;268:574–583. [PubMed: 30124479]
5. KY B, Y L, JL P, 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.
6. Hyde LZ, Valizadeh N, Al-Mazrou AM, et al. ACS-NSQIP risk calculator predicts cohort but not individual risk of complication following colorectal resection. Am J Surg 2019;218:131–135. [PubMed: 30522696]