characteristics data, may be used to personalize treatment decisions and subsequent management in advanced cardiovascular disease.⁸ As we await developments in explainable AI, the use of ML will continue to rise exponentially in the biomedical literature; therefore, it is prudent that clinicians have a cursory understanding of the advantages and disadvantages of AI methods to understand the benefits and limitations of prediction modeling using these novel techniques.

References

- Handelman GS, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H. eDoctor: machine learning and the future of medicine. *J Intern Med*. 2018;284:603-19.
- Rajkumar A, Dean J, Kohane I. Machine learning in medicine. *N Engl J Med*. 2019;380:1347-58.
- Benedetto U, Dimagli A, Sinha S, Cocomello L, Gibbison B, Caputo M, et al. Machine learning improves mortality risk prediction after cardiac surgery: systematic review and meta-analysis. *J Thorac Cardiovasc Surg*. 2022;163:2075-87.e9.
- Tam DY, Fang J, Tran A, Tu JV, Ko DT, Deb S, et al. A clinical risk scoring tool to predict readmission after cardiac surgery: an Ontario administrative and clinical population database study. *Can J Cardiol*. 2018;34:1655-64.
- van der Ploeg T, Austin PC, Steyerberg EW. Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints. *BMC Med Res Methodol*. 2014;14:137.
- Chen P-HC, Liu Y, Peng L. How to develop machine learning models for healthcare. *Nat Mater*. 2019;18:410-4.
- Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol*. 2019;110:12-22.
- Krittanawong C, Johnson KW, Rosenson RS, Wang Z, Aydar M, Baber U, et al. Deep learning for cardiovascular medicine: a practical primer. *Eur Heart J*. 2019;40:2058-73.

See Article page 2075.

Commentary: Machine learning in cardiothoracic surgery: From evidence-based to intelligence-based practice

Arman Arghami, MD, and Joseph A. Dearani, MD

One of the marvels of modern medical practice is the shift from clinical decision making based-on “intuition, unsystematic clinical experience, and pathophysiologic rationale”¹ to scientific, clinically relevant evidence-based decision making. For the past decades, common statistical methods (regressions) have allowed researchers to extract meaningful outcome predictions based on set of known variables. In more recent years, the exponential growth in information and advancement in “big data” storage and sharing and improvements in computational power have matured this field for application of machine learning

Arman Arghami, MD (left), and Joseph A. Dearani, MD (right)

CENTRAL MESSAGE

Machine learning helps improve prediction models of mortality after cardiac surgery and will transform today's evidence-based-practice to tomorrow's practice assisted with artificial intelligence.

(ML). Benedetto and colleagues² report a systemic review of predictive modeling

with ML (also referred to as artificial intelligence). They reviewed published reports from 1966 to 2019 of ML modeling to predict mortality in cardiac surgical patients and compared them with the traditional logistic regression

From the Department of Cardiovascular Surgery, Mayo Clinic, Rochester, Minn.

Disclosures: The 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.

Received for publication Aug 28, 2020; revisions received Aug 28, 2020; accepted for publication Aug 31, 2020; available ahead of print Sept 3, 2020.

Address for reprints: Joseph A. Dearani, MD, Department of Cardiovascular Surgery, Mayo Clinic, 200 First St SW, Rochester, MN 55905 (E-mail: jdearani@mayo.edu).

J Thorac Cardiovasc Surg 2022;163:2094-5

0022-5223/\$36.00

Copyright © 2020 by The American Association for Thoracic Surgery

<https://doi.org/10.1016/j.jtcvs.2020.08.103>

(LR) models. Despite their limitations, which include a wide range of different cardiac surgical procedures targeted in each included study, or the fact that 30% of their reviewed articles suffered from poor methodology and reporting, from the 15 articles reviewed they were able to demonstrate that, compared with LR, ML models achieved better discrimination ability in predicting operative mortality after cardiac surgery (c-statistic: ML 0.88; 95% credible interval, 0.83-0.93 vs LR 0.81; 95% credible interval, 0.77-0.85; $P = .03$).

Risk predictions are becoming even more important in this era of expanding multimodal therapy and percutaneous approaches to allow better patient selection for each approach, identifying unexpected risk factors for perioperative mortality and morbidity, and perhaps even reimbursement practices. Most all current predictive models (Society of Thoracic Surgeons score, European System for Cardiac Operative Risk Evaluation) rely on traditional LR methods for creating a model. The potential advantage of ML modeling is its ability to capture nonlinearity and the interactions among features without the need for the modeler to manually specify all interactions, as needed with LR. Unlike LR, which requires structured datasets with a limited number of variables, ML performs best by using high dimensional data from large validated datasets, the electronic medical record, or even unstructured data like images and clinical notes (with deep learning). This could prove to be far better at quantification of mortality and morbidity risk profiles for procedures in health care.

ML has potential pitfalls. Therefore, researchers and clinicians should verify the validity of the ML methods just like any other diagnostic or prognostic tool. Effective ML modeling requires large numbers of validated, secure data variables for training, eg, the benefit of big datasets such as the Society of Thoracic Surgeons National Database.

Readers of studies reporting the results of ML systems should assess the most crucial ML modeling element, ie, whether the model has been validated on an independent dataset not used for training or “tuning” the model. Naive use of ML modeling can easily result in “overfitting,” whereby random variations in the data that are not associated with outcome are used to predict the outcome. This would be evident on significant decline in the model’s performance between the tuning dataset to the validation dataset. Finally, the results of ML should be judged against expert assessments since results that sound too good to be true probably are. The supersonic train of ML has already left the station, so clinicians and researchers should familiarize themselves with basic methods of ML and methods to read and critique ML publications. Underestimating the power of ML is just as big a mistake as overestimating it—avoid both.³

Cardiothoracic surgery is arguably the most technically demanding of all surgical specialties, with a risk of mortality associated with every operation. The public and the profession aims for and expects a “perfect” outcome in most all circumstances. ML and deep learning will inevitably change the landscape and power of data acquisition and analytics in the health care industry and facilitate continuous improvement in outcomes above and beyond our current methods.

References

1. Guyatt GH. Evidence-based medicine. *ACP J Club*. 1991;114:A-16.
2. Benedetto U, Dimagli A, Sinha S, Cocomello L, Gibbison B, Caputo M, et al. Machine learning improves mortality risk prediction after cardiac surgery: systematic review and meta-analysis. *J Thorac Cardiovasc Surg*. 2022;163:2075-87.e9.
3. Liu Y, Chen PC, Krause J, Peng L. How to read articles that use machine learning: users’ guides to the medical literature. *JAMA*. 2019;322:1806-16.