Commentary: Can we crack the black box of machine learning for aortic aneurysms?

Edgar Aranda-Michel, PhD,a and Ibrahim Sultan, MDa,b

Machine learning (ML) is rapidly becoming a powerful tool in the armamentarium of modern cardiac surgeons. The primary benefit of ML is the lack of assumptions. This is apparent in no a priori defining of variable functional forms (eg, linear vs exponential) and the use of unsupervised learning, where grouping of observations is blind to the outcome (eg, complication vs routine postoperative course). Ostberg and colleagues1 leverage this aspect of ML to train various ML models to predict complications in this patient population. The authors demonstrate 2 major findings: outcome prediction was improved across all ML models, and aortic aneurysm size was the most important feature across all end points. The integration of additional variables, in an unassuming capacity, certainly enhances predictive capability.

The first point that needs to be discussed is the use of a composite end point that included death not directly attributable to an aortic etiology. Composite end points always present with significant bias. There are 2 aspects to this issue. The first and readily apparent aspect is that this may bias or reduce significance of the results—training on a patient population with an aortic aneurysm to an end point that includes nonaortic deaths is not ideal. It is important to note that the authors state this limitation clearly. The other side deals with model overfitting. Although there are several techniques to address overfitting, the caveat is the need for a frequent event rate.2 Low event rates are more difficult for model training and generating robust metrics of the model’s performance. Creating a composite end point allows for an increased event rate and better model performance. However, this is not to say that the results are without merit or devoid of progression.

Although the increase in predictive performance is impressive, there is a larger, more alluring message in this work. Across all ML models, the prediction was improved, underscoring the necessity to account for variables in unassuming capacities. Mainly, not assuming that variables are independent of each other and linearly related to the outcome of interest. ML models, by eliminating these assumptions, can probe various relationships between the variables (eg, covariate or exponential), uncovering crucial dynamics.3 To illustrate the importance of nonlinearity, we turn to renal function and creatinine. The relationship between the 2 is a \( \frac{1}{x} \) dynamic, with increases in creatinine from 1.0 to 1.5 mg/dL indicating a larger decrease in renal function than from 2.0 to 2.5 mg/dL—a distinctly nonlinear relation. However, we are not yet here with ML in cardiac surgery.

This work demonstrates the importance of nonlinearity and indicates which models have high importance. It must be stressed that importance, as referenced here, means...
how does the inclusion of a variable in the model improve its prediction. It does not indicate whether or not that variable increases or decreases the chance of the outcome. And here we finally are, faced with obscurity. We have these black box models that have substantially improved risk stratification—no small feat—and with direct clinical utility, but lacking interpretability. However, the next evolution of this work is to explore the functional forms of the important variables to uncover exactly how they relate to outcomes. By doing so, certain variables could come to light as targets for intervention. Perhaps there is a specific combination of hypertension, weight, and aortic aneurysm size that confers the highest acute aortic dissection risk and focused preventative efforts could be offered.

We have learned over time that aortic aneurysm size is a poor predictor for acute aortic syndrome. ML models may help identify other baseline clinical and radiologic characteristics that may help with prediction of acute aortic syndromes. These techniques may not only be helpful in identifying risk for acute aortic syndromes but perhaps also in identifying morbidity and mortality associated with acute aortic syndromes after surgical intervention.

ML, although improving prediction, hints at variables that are important in nonlinear and nonisolated capacities. We can leverage this to explicitly explore these potential relations—an important undertaking in patient populations where intervention versus surveillance is an ongoing discussion and a moving target. The authors should be commended for the immense undertaking of this work and demonstrating the possible utility of ML techniques. With the appropriate approach, perhaps we will be able to uncover what is hidden in the black box of ML.

References