Commentary: We prefer wisdom over knowledge

Sudish C. Murthy, MD, PhD, and Eugene H. Blackstone, MD

No one disputes that esophagectomy is an extremely morbid surgical intervention despite all our advancements in patient selection, postoperative care, and “minimally invasive” approaches (our favorite oxymoron). There is just something about synchronously disrupting anatomy in contiguous body cavities, often after some sort of debilitating systemic treatment, that humbles even the fittest of patients. Yet, shockingly, some patients move through their postoperative course seamlessly and enjoy a reasonable quality of life, albeit somewhat unnatural given that much of their upper alimentary tract has been hot-wired to some degree.

How we predict who those patients are—and more importantly, who they aren’t—is the focus of the article by Bolourani and colleagues.1 The authors have leveraged a large administrative database and used ever-in-vogue machine learning to develop an algorithm to predict which esophagectomy patients are at the highest risk for 30-day hospital readmission and which might be considered reasonable surrogates for someone not doing particularly well after their intervention.

In fact, early readmission is an important predictor of major late morbidity for most interventions, and clearly for esophagectomy. The authors also demonstrate, as a smaller component of their study, that unfortunately it is a pre-terminal event for some patients. Consequently, appreciating those at greatest risk for hospital recidivism might be of particular interest if one is to intercede and interrupt some downward spiral that the patient has fallen into. That would be the idea, anyway.

Having reviewed this submission, we struggle to understand how these findings meaningfully educate us. Ultimately, we suspect that we are somewhere in that vast expanse between knowledge and wisdom and are having difficulty figuring which way to pivot. Is the message being obscured by overemphasis on the method? We lived through this in the 1970s and 1980s with logistic regression and in the 1980s and 1990s with Cox proportional hazard regression (when the proportional hazards assumption was often violated). Now, in the 2000s, is machine learning doing this to us again?

Machine learning is a modern investigative tool to help us generate new knowledge. It requires (1) a clinically important problem; (2) a thoughtful study design; (3) a comprehensive understanding of the strengths, weaknesses, and appropriateness of the data being used; and (4) most importantly, an expert application of machine learning methodology. When perfectly applied to the proper question and with appropriate data, there is little doubt that vast insight and new knowledge can be acquired; however, this is not accomplished as often as might be expected. There are some limitations to this analytic approach, and in this current study, it is possible that the authors have inadvertently found one: class imbalance. This occurs when the proportion of endpoint events (ie, hospital readmissions) is relatively small. The natural tendency of any machine learning algorithm when applied to an infrequently occurring outcome is to predict that it simply will not occur. The algorithm always defaults to the far more common outcome (ie, no return to hospital) unless some special additional methodology is used to override this

---

From the Department of Thoracic and Cardiovascular Surgery, Heart, Vascular, and Thoracic Institute, Cleveland Clinic, Cleveland, Ohio.

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 June 8, 2020; revisions received June 8, 2020; accepted for publication June 8, 2020; available ahead of print June 27, 2020.

Address for reprints: Sudish C. Murthy, MD, PhD, Department of Thoracic and Cardiovascular Surgery, Cleveland Clinic, 9500 Euclid Ave, J4-1, Cleveland, OH 44195 (E-mail: murthys1@ccf.org).


0022-5223/536.00

Copyright © 2020 Published by Elsevier Inc. on behalf of The American Association for Thoracic Surgery

https://doi.org/10.1016/j.jtcvs.2020.06.057
tendency. We wonder about both the question being asked and the application of machine learning here.

As applied, the authors ask us to use certain time-related outcomes (e.g., postoperative prolonged intubation, pneumonia, acute kidney failure, length of stay) to predict another outcome, readmission. Machine learning or not, is this really that informative?

Let’s say you knew a baseball team had committed numerous errors in a given game, its players struck out several times, and they scored very few runs. Would you be particularly surprised to learn that they had lost? And if you guessed they were going to lose before that last out was recorded, would you honestly consider yourself a modern-day Nostradamus? That is what we struggle with regarding the question being asked.

Yet there might be value buried within this proposition somewhere. Finding where, and more importantly how, is our mission.

Let’s first try to reduce the confound introduced by the time-related nature of postoperative complications required for the algorithm by simply removing them. We will limit the foreshadowing to knowing who the players are, and perhaps only those events occurring in the first inning of a 9-inning game. Anything beyond that time window gets us into the problem we are trying to circumvent. We predict that this model would be fairly poor, and not necessarily because of the obvious (remembering that it is the obvious that makes the study somewhat superfluous). Instead of giving up, we might ask what else can be added a priori? What important objective measures that might affect the outcome are we leaving outside the stadium before the game starts? In relation to recovery from esophagectomy, important physiological measurements, such as forced expiratory volume in 1 second, ejection fraction, 6-minute walk distance, and grip strength, might just be outside the turnstile. Frailty models are likely critical, supported by the authors’ identification of preoperative malnutrition as an important predictor of early readmission. Because we do not have to be clairvoyant when we use these types of data, the conclusions become far more informative.

Might there be a benefit from simply using the algorithms that the authors have developed and just imputing the data that they suggest? Possibly, but everyone needs to understand that the data spreadsheet will have to be completed at the moment of patient discharge. If there is a lag in collection, then the entire house of cards rapidly falls down.

It is possible that a highly organized data collection system might not be as difficult as it sounds, given that this approach is already being used for updating organ allocation score for heart transplantation. But it is clearly labor-intensive, and it should be appreciated that this type of data mining mandates abstractors buzzing around the patient 24/7 to capture these dynamic changes in clinical condition. It is unclear whether this is a practical solution to simply satisfy the constraints involved with meaningfully using this machine learning algorithm.

So at the end of the day, what have we got? Well, machine learning can be used to develop a prediction model for 30-day readmission after esophagectomy. Check. And perhaps this is new knowledge. Check. But knowing that a patient who struggled mightily to get through their intervention is likely to be readmitted within 30 days will almost certainly not make us much wiser, will it? Landing on the right preoperative variables to give us this same predictive capacity should be our goal, and that is the wisdom we have gained from this submission.

References