Article details: 2021-0031	
Title	**Derivation and validation of a hybrid machine learning model to predict death on the waitlist for cardiac surgery in Ontario, Canada: a population-based study**
Authors	Louise Y. Sun MD SM, Harindra C. Wijeysundera MD PhD, Douglas S. Lee MD Thierry G. Mesana MD PhD
Reviewer 1	James MacKay Brophy
Institution	
General comments (author response in bold)	Summary: Using Ontario administrative databases for coronary artery bypass grafting, valvular, and/or thoracic aorta surgeries occurring between October 1, 2008, and September 30, 2019, the authors have developed and validated a clinical model to predict death on the waitlist for surgery. General comments: The research question is definitely topical and of great interest to most healthcare professionals as they navigate these turbulent and complex times. The paper is generally well written and easy to follow. RESPONSE: We thank the reviewer for their support of our manuscript. My main concerns are: 1. The authors seem to over-estimate the strength of the evidence supporting their model as a c statistic of 0.73, while better than a coin toss of 0.5 is really more of a modest effect rather than providing definitive model support. RESPONSE: We appreciate the reviewer's concern. While we agree with the reviewer that discrimination (c-statistic) is important, we believe calibration (observed vs. expected) defines whether a model is applicable in the population for which it is intended. Our conclusion was founded on the model's high degree of calibration, where observed and expected rates were nearly identical across all surgical categories. 2. The abstract conclusions are not supported by the abstract results. In the conclusion, we are told that this new model can be combined with the CardiOttawa Length of Stay Score but no mention of this second score is provided in the methods or results. The conclusion therefore appears as a complete non- sequitur. RESPONSE/REVISION: Thank you for pointing this out. We have revised to "It has the potential to provide data-driven decision support for managing access to cardiac care and preserve system capacity during the COVID-19 crisis, the recovery period, and beyond." in the abstract. 3. "CorHealth Ontario maintains a detailed prospective registry of all patients who undergo invasive cardiac procedures" Is this a registry only of diagnostic procedures? Unclear to the non-Ontario reader exactly what is included in this "prospective" registry. Prospective with regards to what time point? RESPONSE/REVISION: Thank you for the opportunity to clarify our methods. We have revised to "CorHealth Ontario maintains a detailed prospective registry of all patients who undergo invasive diagnostic and therapeutic cardiac procedures…CorHealth Ontario data was prospectively collected from the time of

	surgical referral and undergo selected chart audits and core laboratory validation (3, 7-15).” on page 5 line 19. 4. I am unsure of the validity of their proposed “hybrid” approach. I see no reference supporting this. The authors have used combined random forest, a classification technique with logistic regression, a prediction technique without any exploration of the differences in these two measures (see fharrell.com/post/classification/) for more discussion about these contrasting approaches. RESPONSE/REVISION: We have now provided a reference to our hybrid ML approach (Reference #26 in the manuscript; PMID 24771344). We have also been able to demonstrate in a methods paper, that such a hybrid approach performed better than logistic regression with and without bootstrapping, in predicting cardiac surgery-associated acute kidney injury (manuscript under review). 5. Ultimately if this paper is to be of interest to healthcare providers, the authors need to better explain why their hybrid ML/statistical model predicts survival/death better than my model of simply predicting everyone will be alive which has a 99.8% accuracy ((112266-269)/112266). RESPONSE: We agree with the reviewer that deaths are infrequent on the waitlist. However, deaths on the waitlist reflect deficiencies of publicly funded healthcare systems, where resources are limited and would need to be allocated wisely. During the era of COVID-19, waitlists have become a reality across the globe, again highlighting the importance of evidence-based waitlist management strategies, to prioritize high-risk patients and avoid deaths while awaiting surgery. Minor comments 1. Introduction – “disrupted the care of patients with cardiovascular conditions.” Covid has not disrupted care only for CV patients RESPONSE/REVISION: We have revised to “disrupted the care of patients with cardiovascular and other health conditions” on page 4 line 4. 2. What are “salvage procedures”? RESPONSE/REVISION: Salvage procedures are rescue procedures, performed after failure of initial treatment. We have changed the word “salvage” to “rescue” to better describe this point on page 5 line 13.
Reviewer 2	Vivek Rao
Institution	Cardiac Surgery, University of Toronto, Toronto, Ont.
General comments (author response in bold)	Comments to the Author The authors have submitted a manuscript describing their derivation of a clinical model to predict waiting list deaths. The authors claim that they have "validated" this model and indeed, state so in their title. However, conventional validation uses different databases (a derivation database and a validation database).

	RESPONSE: We agree with the reviewer that our method of validation (bootstrapping) may differ from traditional method of split sample validation within the same dataset, or validation using an external dataset. However, the low event rate (N=269 deaths) precludes us from splitting our dataset for this purpose. In addition, bootstrapping was shown by Dr. Harrell's group to be more efficient than the split sample method for validation (PMID 11470385: J Clin Epidemiol. 2001 Aug;54(8):774-81). This view is also supported by Dr. Steyerberg's book "Clinical Prediction Models" (p. 332-335). There are other concerns with the data in this study: 1. CIHI notoriously has poor accuracy for clinical data. I commend the authors for obtaining more reliable clinical data from the CorHealth registry, but several variables were still dependent on the CIHI database, and this limitation should be addressed by the authors. RESPONSE: We appreciate the reviewer's concern and would like to assure the reviewer that CIHI was only used as a secondary source of data to supplement information on comorbidities. To this end, we did not rely on diagnostic codes, which are known to be inaccurate in capturing this information. We instead used validated algorithms that have been designed for sensitive and specific data capture using integrated data sources that include CIHI and physician billing data (references were provided in the manuscript). 2. The time frame of this study ranges from 2008 to 2019 where significant changes in cardiac therapy occurred. In addition to the introduction of new surgical centers in the province, TAVI, MitraClip and other novel percutaneous therapies have been introduced. These technologies may have influenced their primary outcome: death. For example, if a patient with severe AS deteriorated on the waitlist and presented with CHF and end-organ compromise, they may no longer be a surgical candidate. Prior to 2010, this patient would likely have died. I propose that in 2020 (and likely after 2015), this patient would receive salvage TAVI and likely survive. RESPONSE/REVISION: Thank you for this astute comment. We have added this as a limitation on page 14 line 14: "Fourth, there have been advances in transcatheter techniques over the course of the study period. Further research is needed to identify how the advent of these minimally invasive procedures may have influenced the referral process and outcomes on the waitlist." 3. Further to point#2, death is an objective but insensitive outcome. If a patient waiting for CABG suffered a NSTEMI and underwent PCI and survived, they would not be captured by this model. Particularly if they were diabetic, one could argue that they received inferior therapy as a result of an adverse non-fatal event while waiting for CABG. RESPONSE: The reviewer raises an important point. We had in fact performed a follow-up study, where we developed a clinical risk model for predicting cardiac hospitalizations on the waitlist (manuscript under review). These events would be captured by the cardiac hospitalization model.

	4. The authors should have data on non-fatal adverse events while on the waiting list, particularly those that may influence perioperative risk. For example, if a patient waiting for CABG suffers a STEMI resulting in a significant decline in LV function yet still presented for CABG this would not be captured by their model. If that patient died during higher risk CABG, it is a direct result of waiting and the adverse event suffered while waiting but is not captured by their model. RESPONSE: Thank you for this astute comment. STEMI and heart failure would both be captured in our cardiac hospitalization model. Despite my comments above, I think this is an important contribution particularly in the current pandemic where decisions relating to resource allocation are being made constantly. A follow-up study where this model is used to identify patients who died waiting for cardiac surgery during the pandemic would be worthwhile and, in my opinion, "validate" this model. RESPONSE: We appreciate this suggestion and would like to thank the reviewer for their support and appreciation of our manuscript.