BMJ Health & Care Informatics

Mortality prediction models, causal effects, and end-of-life decision making in the intensive care unit

Jason H Maley ,[1,2] Kerollos N Wanis,[3] Jessica G Young,[3,4] Leo A Celi[2,5]

To cite: Maley JH, Wanis KN, Young JG, *et al*. Mortality prediction models, causal effects, and end-of-life decision making in the intensive care unit. *BMJ Health Care Inform* 2020;**27**:e100220. doi:10.1136/bmjhci-2020-100220

Received 05 August 2020
Accepted 05 October 2020

[1]Center for Healthcare Delivery Science, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA
[2]Institute for Medical Engineering and Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
[3]Department of Epidemiology, Harvard University T H Chan School of Public Health, Boston, Massachusetts, USA
[4]Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, Massachusetts, USA
[5]Division of Pulmonary, Critical Care, and Sleep Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA

Correspondence to
Dr Jason H Maley;
jmaley@bidmc.harvard.edu

End-of-life (EOL) decision making in the intensive care unit (ICU) is challenging for both families and clinicians. This decision-making process is ideally framed around a shared understanding of a patient's values and goals, all taken in the context of their critical illness and prognosis. However, clinicians commonly face uncertainty regarding prognosis and may have difficulty offering families an accurate assessment of the likely outcomes of treatment decisions. Adding to the complexity of these scenarios, clinicians, patients and families are each susceptible to unconscious but influential cognitive biases when making decisions under stress. Given these challenges, and a rapidly growing interest in data science to inform care in the ICU, investigators have explored the use of prediction models (eg, machine learning or ML algorithms) to assist with prognostication.[1–3] Prediction models describe an outcome distribution among individuals with a particular set of characteristics, such as risk of acute kidney injury among individuals with particular laboratory values and clinical characteristics in a population. However, they do not compare how that outcome distribution would change were different treatment decisions made in that population—this requires causal effect estimation, rather than prediction modelling. Herein, we explain why prediction modelling alone is not sufficient to inform many ICU treatment decisions, including EOL decision making, and describe why causal effect estimation is necessary.

Consider the following case in which a prediction model is used, rather than causal effect estimation: a 68-year-old man is admitted to the ICU with severe pneumonia and requires mechanical ventilation. After 5 days, he requires continued full support from the ventilator and has developed delirium. His family is concerned about prolonging intensive care but worries about transitioning to comfort measures prematurely if continued intensive care could, in fact, achieve their goals. His family and clinicians would like to use the best available evidence to inform their decision. Given an increasing interest in mortality prediction models, the clinical team explores this as a tool for decision support. Their chosen algorithm is a prediction model—trained on available data containing measurements of treatments, outcomes and other characteristics of previously ventilated patients–which returns a 70% probability of death within the next 30 days.

The first question we must ask is: what is the precise interpretation of this prediction? This is an estimate of the probability of death for a population of patients 'like this patient'—patients mechanically ventilated for 5 days with similar baseline characteristics and clinical risk factors up to that moment in their ICU course. Importantly, this probability is contingent on the treatment decisions, after day 5, that were made in the population from which the training data came. For example, if all patients similar to this one in the training data transitioned after day 5 to comfort measures only, then the algorithm would predict a 100% risk of 30-day mortality. Alternatively, if most of the patients similar to this index case in the training data frequently pursued tracheostomy, the model may predict a low 30-day mortality. These two extreme conditions demonstrate that the interpretation is highly dependent on the distribution of treatment decisions among patients 'like this patient' in the training data.

Second, and more importantly, we consider the following question: how is this probability useful (or not useful) to both the family and the clinicians? We must first clearly articulate the question that clinicians and family members are truly interested in answering for this patient. They are not simply concerned about predicting if survival is possible. Instead,

they are interested in knowing: what would the outcome be if one treatment strategy was chosen compared with the outcome if a different treatment strategy was chosen. To be more concrete, they may want to know if continued mechanical ventilation and attempts at ventilator liberation for another week would result in a different 30-day survival than a more limited trial of 48 hours of ventilation. They are concerned with the balance between unnecessarily prolonging intensive care versus a missed opportunity for survival if they transition to comfort measures prematurely. More importantly, in the context of the patient's values and goals, they likely want to understand the effect of these treatment strategies on long-term quality-of-life.

CONTRASTING CAUSAL EFFECTS WITH PREDICTION MODELS

We refer to the outcome that would have been observed had, perhaps contrary to fact, a particular treatment been given, as the *counterfactual outcome* under that treatment. Using our earlier clinical scenario as an example, if our critically ill patient had decided to undergo tracheostomy on day 7, we might wonder 'what would have happened if he did not undergo tracheostomy and instead remained intubated'. The outcome under our 'what if' scenario is contrary to what actually happened, that is, the counterfactual outcome. *Causal effects* compare counterfactual outcomes for a person (or a population) under different treatment strategies, asking 'what would be the outcome if we choose treatment A compared to the outcome if we choose treatment B'. Clinicians, patients and families intuitively think in counterfactuals when weighing the risks and benefits of different decisions, including EOL decision making. Thus, causal estimates would seem to be the natural approach to support decision making. Yet prediction models, rather than causal estimates, have received rapidly growing attention in the literature while their limitations are often overlooked.

In contrast to causal estimates, mortality prediction models are mapping inputs (or 'features') to a chosen outcome, such as mortality. They might help estimate if a patient is at a higher risk of death, but they offer little help in making the best decision in that scenario. However, as we noted previously, these estimates depend on the distribution of treatments that were given to patients like ours in the training data. As such, if historic treatment distributions differ from those in current practice, then the prediction will be inaccurate.

The appeal of predictive approaches, from a data science perspective, is that they can be readily applied with existing healthcare data and established machine learning algorithms. However, these models ignore assumptions about causal structure— the relationship between the variables that can only be informed by expert knowledge. For example, users of a prediction model may claim that a high fraction of inspired oxygen is associated with higher ICU mortality but that prediction could not justify a claim that an intervention to reduce the fraction of inspired oxygen administered would reduce mortality without first reasoning about how these variables are connected to one another.[4] Specifically, they must defend assumptions about how the treatment and outcome are related, causally or by associational pathways. Because prediction models omit this step, they cannot provide an estimate of expected outcomes when divergent treatment decisions are chosen.

APPLYING CAUSAL INFERENCE TO ICU DATA

Having established the need for causal effects of different ICU decisions, rather than predictions of mortality, we will describe how they can be estimated. An intuitive and effective approach to designing observational analyses for estimation of causal effects is to specify a hypothetical pragmatic randomised trial (ie, a 'target trial'), one that would answer the question of interest but may be impossible or impractical to conduct in practice.[5] This hypothetical trial helps us be explicit about the important aspects of our analysis, including the causal question it aims to answer and avoid biases introduced by the study design (eg, immortal time bias).[6] Specifically, we need to define the eligibility criteria, the treatment strategies of interest, the follow-up period (including a clear definition of 'time zero', the start of follow-up, eg, mechanical ventilation day 5 in the above example), the outcomes of interest and the statistical analysis plan. This also requires expert knowledge of the clinical context.

We describe one such trial, for example, in table 1. This trial is ethically and logistically infeasible; therefore, an analysis of observational data, designed with identical features as the trial, is the next best approach. In particular, to emulate the trial described in table 1, we would, after obtaining appropriate observational data: (1) restrict our data to individuals meeting the eligibility criteria, (2) classify those who immediately discontinue mechanical ventilation as adherent to strategy one and those continuing mechanical ventilation on day 6 as adherent to strategy two and (3) compare estimates of the risk of 30-day mortality among those adherent to strategy one versus strategy two, adjusted for measured prebaseline prognostic factors (ie, *measured confounders*). Adjustment is required because treatment is not randomly assigned in observational data (in other words, treatment is related to the outcome via associational pathways). If all the relevant confounders are measured and adjusted for, then the same effect estimates will be obtained from the observational data analysis as from the trial had it been conducted (except for random variation).

While this may be a useful example for the technical aspects of this process, the trial described in table 1 is not the one of interest to decision makers. For example, treatment strategies such as 'continue mechanical ventilation for another week, unless liberation from the ventilator is achieved' versus 'continue mechanical ventilation for another 48 hours, unless liberation from the ventilator is achieved' address questions around time-limited trials

Table 1 A hypothetical randomised trial

Trial component	Description
Eligibility criteria	Individuals aged 65 years or older admitted to a critical care unit with severe pneumonia requiring intubation who have received 5 days of mechanical ventilation and cannot yet be liberated from the ventilator.
Treatment strategies	1. Immediate transition to comfort care measures. 2. Immediate continuation of mechanical ventilation.*
Assignment procedures	Unblinded random assignment to one of the treatment strategies.
Follow-up period	Beginning at baseline, the time of randomisation, individuals are followed until death or the end of 30 days.†
Outcome	All-cause mortality by the end of 30-day follow-up.
Causal contrast of interest	Intention to treat effect.

*Note that these treatment strategies are not *sustained* because they only direct initial treatment. In other words, individuals assigned to 'immediate continuation of mechanical ventilation' may later be transitioned to comfort measures.
†For simplicity, we assume no loss to follow-up. This assumption is reasonable for ventilated patients for an endpoint at 30 days from randomisation.

better than those proposed in table 1. These strategies are *sustained*[7] because they specify a treatment over time, rather than simply at baseline, and they are *dynamic*[8] because the treatment assigned under each strategy depends on a patient's time-evolving characteristics, such as respiratory status and liberation from the ventilator. Almost all real-world treatment strategies in the ICU are sustained and dynamic, yet clinical researchers infrequently apply the methods necessary to account for this.[9]

In addition to considering different treatment strategies than those in table 1, clinicians and families are also interested in outcomes other than 30-day survival. For instance, they may be more interested in quality of life at 6 months. Because the quality of life at 6 months is not defined among individuals who die before the end of 6 months of follow-up, defining a meaningful causal effect of interest requires careful handling of competing events.[10]

For many clinically relevant questions, causal inference researchers have developed the methodological tools required for computing the function of the observed data that identify the causal effect; that is, the effect that could be directly estimated from a perfect execution of the target trial. This occurs under assumptions about causal structure, informed by clinical expertise.[9 11] In particular, the function depends on all of the covariates within the causal structure of the clinical scenario that are needed for confounding adjustment. Estimating this typically high-dimensional function of the data does, in fact, involve obtaining a form of predictions as interim steps. For example, inverse probability weighting, which under particular assumptions can yield causal effect estimates, requires as an interim step that the probability of treatment conditional on the confounders be estimated. This estimate needs to be an accurate mapping between the treatment and the confounders. These predictions are used to construct the weights for the final causal effect estimation. Methods that address sustained and dynamic treatment strategies may incorporate multiple predictions into an inverse probability weighted approach to account for the time-varying nature of real-world care. Therefore, while we have argued that prediction modelling is not, in itself, ideal for ICU decision making, prediction algorithms are necessary interim steps for obtaining causal effect estimates (which are the basis of such decision making). Moreover, just as the use of modern machine learning algorithms (eg, neural networks, random forests and gradient boosting) may perform better than traditional models (eg, logistic regression) when the end goal is a prediction, these modern algorithms may ultimately provide better causal effect estimates than traditional models when used during interim steps.[12–14]

CONCLUSION

Amidst our enthusiasm to apply machine learning to ICU healthcare data, we should remember to start with the end in mind—questions that matter to patients and families. The process requires clinical expertise to identify specific treatment strategies and outcomes of interest. It also entails close collaboration between clinicians, data scientists, causal inference experts, patients and families. Rather than prediction (which might help us identify a problem), we should estimate causal effects (which help us understand the impact of actions we may take when faced with that problem) by applying the tools developed by causal inference researchers over the past two and a half decades. While machine learning plays an important role in this process, it is relevant only after the careful mapping of a causal structure and consideration of design elements of a target trial. In doing this, data analysis may begin to complement the existing sound principles of EOL communication in the ICU and answer many other important questions faced by clinicians.

Twitter Jason H Maley @jhmaley

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iD
Jason H Maley http://orcid.org/0000-0002-3632-6585

REFERENCES

1 Barnato AE, Angus DC. Value and role of intensive care unit outcome prediction models in end-of-life decision making. *Crit Care Clin* 2004;20:345–62.
2 Johnson AEW, Mark RG. Real-time mortality prediction in the intensive care unit. *AMIA Annu Symp Proc* 2017;2017:994–1003.
3 Hadique S, Culp S, Sangani RG, *et al*. Derivation and validation of a prognostic model to predict 6-month mortality in an intensive care unit population. *Ann Am Thorac Soc* 2017;14:1556–61.
4 Grasselli G, Greco M, Zanella A, *et al*. Risk factors associated with mortality among patients with COVID-19 in intensive care units in Lombardy, Italy. *JAMA Intern Med* 2020. doi:10.1001/jamainternmed.2020.3539. [Epub ahead of print: 15 Jul 2020].
5 Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*.
6 Hernán MA, Sauer BC, Hernández-Díaz S, *et al*. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol* 2016;79:70–5.
7 Danaei G, García Rodríguez LA, Cantero OF, *et al*. Electronic medical records can be used to emulate target trials of sustained treatment strategies. *J Clin Epidemiol* 2018;96:12–22.
8 Hernán MA, Lanoy E, Costagliola D, *et al*. Comparison of dynamic treatment regimes via inverse probability weighting. *Basic Clin Pharmacol Toxicol* 2006;98:237–42.
9 Robins J. *A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect*, 1986: 1393–512.
10 Young JG, Stensrud MJ, Tchetgen Tchetgen EJ, *et al*. A causal framework for classical statistical estimands in failure-time settings with competing events. *Stat Med* 2020;39:1199–236.
11 Hernán M, Robins J. *Causal inference: what if*. Chapman & Hall/CRC, 2020.
12 Liu L, Mukherjee R, Robins J. On assumption-free tests and confidence intervals for causal effects estimated by machine learning. Available: https://arxiv.org/abs/1904.04276
13 Van der Laan MJ. *Targeted learning in data science : causal inference for complex longitudinal studies*. Springer Berlin Heidelberg, 2017.
14 Chernozhukov V, Chetverikov D, Demirer M, *et al*. Double/debiased machine learning for treatment and structural parameters. *Econom J* 2018;21:C1–68.