

Structural causal models' application in the development of clinical trials

Marvi Bikak MD

Machine learning has revolutionized all aspects of our lives. From Amazon accounts that recommend items to buy based on past purchases to Netflix recommended movies, computers are generating assumptions about our lives which often are very accurate. Several areas in medicine have embraced machine learning to help refine our diagnostic abilities. A result of one such application was the development of a software that uses image analysis to screen for diabetic retinopathy.¹

While there are numerous machine learning algorithms and methods available, for clinicians it is important to understand that not all will have the same degree of dependence on a computer to derive results. For example, a clinician prescribing a statin to a patient with an LDL >200 mg/dL is an entirely independent human decision, whereas the calculation of the MELD risk score that uses regression analysis requires some degree of dependence on a computer to analyze the results. Conversely, convolutional neural networks, such as image analysis software, rely heavily on a computer's ability to analyze thousands of images to generate the diagnostic prediction model.² Similarly, Structural Causal Models are a type of machine learning algorithm that involves a robust selection of variables to understand their links to outcomes of interest. Structural causal models have been used and applied in various epidemiological research projects and are becoming widely known to help answer clinical questions using observational data.³

A systematic approach is needed to create a structural causal model to understand a causal relationship between a variable and outcome of interest.⁴

The first step is to create a causal model based on prior knowledge about the question, as, for example, how low tidal volume ventilation decreases mortality. This model must also include all the variables that are linked with the outcome directly or indirectly. The model is then tailored according to the observed data that are available for analysis. Applying this model to the observed data generates relationships that help clinicians not only compare to the original model but also identifies other potential variables that may represent important relationships. To analyze the validity of these generated relationships, statistical analysis is applied. The more robust the accountability for confounders, the stronger the causal model is.

In critical care medicine one of the most poorly understood diseases is the acute respiratory distress syndrome (ARDS) given the variations in the disease progression from patient to patient. Numerous randomized controlled trials have been conducted to understand the best ventilation strategies, yet there are numerous unanswered questions. For example, what PEEP setting has better outcomes? According to ALVEOLI,⁵ LOVS⁶ and EXPRESS⁷ trials, we know there is no difference in mortality with high PEEP vs low PEEP strategies but are unable to confidently choose the perfect PEEP setting in all ARDS patients simply because it may not apply to all patients. The challenges that are common to all RCTs is that in a highly selected population, the generalizability of the results is mainly poor given the heterogeneity among patients.⁸ We can utilize the application of causal models to understand our critical care patient population better and identify not only the best treatment strategies for them but also discover new phenotypes like PaO₂:FiO₂ that impact outcomes and response to treatments.

We created a structural causal model based on 3 landmark trials⁹ in ARDS, ARMA,¹⁰ ALVEOLI,⁵ and ACURASYS.¹¹ Variables and outcomes in trials were used to create the models. These models were then

Corresponding author: Marvi Bikak
Contact Information: marvi.bikak@gmail.com
DOI: 10.12746/swrccc.v7i31.587

applied to a large database (MIMIC-III). Data were represented by direct acyclic graphs which mathematically analyzed the causal relationships between the variables and the outcomes. Using this method, we were able to produce results like the trials. However, the most remarkable feature about using structural causal models in such clinical scenarios is how quickly results become available as opposed to trials that take years to complete.

Structural causal models can have vast applications. They can be used to design observational studies that emulate RCTs, rigorously controlling confounders and determining important relationships to test a hypothesis.¹² Will structural causal models replace RCTs as the gold standard method to generate evidence? This seems tough to imagine but is not impossible if clinicians start applying them regularly to answer simple clinical questions and ultimately learn the best ways to eliminate confounders.

Keywords: causal inference, structural causal models, observational studies, acute respiratory distress syndrome

Article citation: Bikak M. Structural causal models' application in development of clinical trials. *The Southwest Respiratory and Critical Care Chronicles* 2019;7(31):1–2

From: Department of Internal Medicine, Rush College of Medicine, Chicago, Illinois

Submitted: 9/2/2019

Accepted: 10/17/2019

Conflicts of interest: none

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

REFERENCES

1. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 2016; 316(22):2402–2410.
2. Beam AL, Kohane IS. Big data and machine learning in health care. *JAMA* 2018;319(13):1317–1318.
3. van der Laan MJ, Rose S. Targeted Learning: Causal Inference for Observational and Experimental Data. Springer Science & Business Media; 2011.
4. Petersen ML, van der Laan MJ. Causal models and learning from data: integrating causal modeling and statistical estimation. *Epidemiology* 2014;25(3):418–426.
5. Brower RG, Lanken PN, MacIntyre N, et al. Higher versus lower positive end-expiratory pressures in patients with the acute respiratory distress syndrome. *N Engl J Med* 2004 Jul 22; 351(4):327–36.
6. Meade MO, Cook DJ, Guyatt GH, et al. Ventilation strategy using low tidal volumes, recruitment maneuvers, and high positive end-expiratory pressure for acute lung injury and acute respiratory distress syndrome: a randomized controlled trial. *JAMA* 2008;299(6):637–645.
7. Mercat A, Richard JCM, Vielle B, et al. Positive end-expiratory pressure setting in adults with acute lung injury and acute respiratory distress syndrome: a randomized controlled trial. *JAMA* 2008;299(6):646–655.
8. Bos LD, Martin-Loeches I, Schultz MJ. ARDS: challenges in patient care and frontiers in research. *Eur Respir Rev* 2018;27(147):
9. Bikak M, Adibuzzaman M, Jung Y, et al. Regenerating evidence from landmark trials in ARDS using structural causal models on electronic health record. *Am J Resp Crit Care Med* 2018;197:A4290:2.
10. Acute Respiratory Distress Syndrome Network. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *N Engl J Med*. 2000 May 4;342(18):1301–8.
11. Papazian L, Forel JM, Gacouin A, et al. Neuromuscular blockers in early acute respiratory distress syndrome *N Engl J Med* 2010;363:1107–1116.
12. Lederer DJ, Bell SC, Branson RD, et al. Control of confounding and reporting of results in causal inference studies. Guidance for Authors from Editors of Respiratory, Sleep, and Critical Care Journals. *Ann Am Thorac Soc* 2018;16(1):22–28.
1. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic