Preventing Intraoperative Hypotension

Artificial Intelligence *versus* Augmented Intelligence?

Mozziyar Etemadi, M.D., Ph.D., Charles W. Hogue, M.D.

Machine learning as a construct for artificial intelligence permeates many aspects of our professional lives where it offers the hope to improve health *via* multiple applications including image recognition, natural or textual language identification, "big data" analysis, and others.[1] While society may not necessarily be completely ready to have their medical care (or automobiles!) usurped by a computer, artificial intelligence has the potential to augment clinical decision-making to enhance safety and improve outcomes.[2] For example, artificial intelligence tools have been created to assist radiologists in reading lung cancer screening computed tomography images[3] and mammograms[4] and have demonstrated significantly better-than-human performance in almost all analyses.

"[Does] alerting clinicians of impending hypotension reduce its duration and severity compared with usual care?"

In this issue of Anesthesiology, Maheshwari *et al.*[5] evaluate whether clinician access to alerts from a hypotension prediction algorithm reduces hypotension compared with usual care. This hypotensive predictor is a commercially available machine learning tool previously described.[6] The proprietary algorithm uses analysis of complex features of the arterial waveform not visible with the naked eye. In a validation study, the hypotensive predictor had a sensitivity of 88% (85 to 90%) and specificity of 87% (85 to 90%) to identify a hypotension 15 min in advance (area under the receiver operating characteristic curve, 0.95).[6] That study did not include episodes of hypotension caused by surgical manipulations, and it was not tied to whether a clinical intervention was necessary.

The current pilot randomized trial thus has important relevance to this area of research in evaluating the hypotensive predictor under real-time clinical circumstances. The authors specifically ask the question of whether alerting clinicians of impending hypotension (mean arterial pressure [MAP] less than 65 mmHg) reduces its duration and severity compared with usual care. There is divergent opinion on what blood pressure constitutes "hypotension," and we have argued that its definition is an individual cutoff not accurately derived from population based data.[7,8] Nonetheless, retrospective investigations have found an association between an intraoperative MAP less than 65 mmHg and acute kidney injury, myocardial infarction, and mortality for adults undergoing noncardiac surgery.[9,10] Newer data, however, suggests that postoperative hypotension rather than intraoperative hypotension is a more important determinant of myocardial infarction.[11]

Regardless, in the current study, 214 adult patients undergoing moderate- to high-risk noncardiac surgery were randomized to an intervention group where hypotension prediction alerting was available to clinicians or a group where the alert was withheld. In the intervention arm, clinicians were given a recommended hemodynamic treatment algorithm that included any combination of administration of intravenous fluids, vasopressors, inotropic drugs, or observation. In both study arms, clinicians were advised to minimize the duration of MAP less than 65 mmHg. Eligibility included the need for direct arterial pressure measurement during surgery that was required to last for greater than 2 h. Patients were excluded for various reasons including conditions that are common in surgical patients (*e.g.*, moderate or severe valvular heart disease, atrial fibrillation, or heart failure with low ejection fraction). The main findings were that there was no difference in the primary outcome of the time-weighted average of MAP less than 65 mmHg between groups, nor were there differences for the

Image: J. P. Rathmell.

This editorial accompanies the article on p. 1214.

Accepted for publication August 26, 2020. Published online first on October 20, 2020. From the Department of Biomedical Engineering, McCormick School of Engineering, Northwestern University, Evanston, Illinois (M.E.); and the Department of Anesthesiology, Feinberg School of Medicine, Northwestern University, Chicago, Illinois (M.E., C.W.H.).

 Anesthesiology 2020; 133:1170–2. DOI: 10.1097/ALN.0000000000003561

secondary outcomes of the time-weighted average of MAP less than 60 mmHg or less than 55 mmHg. The warning for hypotension was relatively short and was often less than 2 min, thus potentially limiting the time for appropriate responses. This warning time period is certainly shorter than the 15 min previously reported in the offline validation study.[6] The authors noted the limitations of the prescriptive intervention algorithm that may have been too complex for the short interval from warning to hypotensive event.

The results of the study will be contrasted with other similar studies including a single-center randomized trial of 64 patients undergoing noncardiac surgery under general anesthesia.[12] In the latter study, the median time-weighted average of MAP less than 65 mmHg was lower in the group that received hypotensive prediction alerting compared with controls where there was no alerting. It is noteworthy that in the current study, the time-weighted average of MAP less than 65 mmHg (median [interquartile range]) in the control group (0.14 [0.03, 0.39] mmHg) appears similar to that in the intervention group in the previous study (0.10 [0.01, 0.43] mmHg) but lower than their controls (0.44 [0.23, 0.72] mmHg). This may be due to direction in the protocol in the current study to keep MAP greater than 65 mmHg even in the control group.

Of note, in the pilot randomized controlled trial, the early warning system often changed rapidly, possibly as the result of surgical manipulations. In this case, experienced clinicians would likely have observed or discussed the blood pressure with the surgeons but refrained from pharmacologic intervention. Thus, artificial intelligence might have accurately predicted hypotension, but the results did not improve clinical care. Rather, the machine learning hypotension prediction tool may have only complementary value to an astute clinician. That said, in *post hoc* analysis, they report that when considering only those hypotensive events that clinicians treated, the hypotensive predictor group had a 57% lower time-weighted average of MAP less than 65 mmHg than in the control arm (table 4 in article). Perhaps a more relevant study design in the future investigations would be to test the hypotension prediction algorithm for reducing episodes of hypotension that received treatment by deleting those where observation was recommended.

While the prediction tool offers guidance, the crucial limitation of the algorithm is that it is completely agnostic to the clinical situation, and it directs clinicians to interpret the recommended treatment on the monitor, then decide if the episode can be ignored or requires intervention. The clinical reality is that vasoactive drugs are often immediately given while more definitive treatments (*e.g.*, fluids, blood products, *etc.*) are prepared, then initiated. Nonetheless, what if instead of an algorithm that strictly predicts MAP less than 65 mmHg regardless of the clinical setting, a family of algorithms was trained to predict only the need for each intervention itself delivered to the clinician in an easily interpreted format, such as a "you will need to administer phenylephrine" predictor, or a "you will need to give an intravenous fluid bolus" predictor, and so on? This way, the nuanced decision-making of the clinician is captured in the training data, and therefore the clinical practice is incorporated in the algorithms. This approach fundamentally improves the ability of clinical decision-making without bypassing the clinician's skills. To be sure, the development of such a family of algorithms requires significant, interdisciplinary collaboration between machine learning engineers, medical device manufacturers, and clinicians, all with tremendous amounts of data that are labeled iteratively and collaboratively. Our belief is that the time has come to augment our clinical intelligence and not necessarily only with artificial intelligence. For those important split-second decisions we make as anesthesiologists, we can train an artificial intelligence tool to "give us a heads-up." The study evaluating a machine learning hypotension prediction tool clearly demonstrates the pitfalls of predictive analytics in clinical medicine that will need to be refined to enhance clinical relevance.

Research Support

Supported in part by grant No. 1R01HL130619-01A1 from the National Institutes of Health (Bethesda, Maryland; to Dr. Etemadi) and grant No. RO1HL092259 from the National Institutes of Health (to Dr. Hogue).

Competing Interests

Dr. Hogue has received speaking and consultant fees from Edwards Lifescience (Irvine, California). He has consulted with Medtronic, Inc. (Minneapolis, Minnesota).

Correspondence

Address correspondence to Dr. Hogue: charles.hogue@nm.org

References

1. Connor CW: Artificial intelligence and machine learning in anesthesiology. Anesthesiology 2019; 131:1346–59
2. Rusin CG, Acosta SI, Shekerdemian LS, Vu EL, Bavare AC, Myers RB, Patterson LW, Brady KM, Penny DJ: Prediction of imminent, severe deterioration of children with parallel circulations using real-time processing of physiologic data. J Thorac Cardiovasc Surg 2016; 152:171–7
3. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, Tse D, Etemadi M, Ye W, Corrado G, Naidich DP, Shetty S: End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nat Med 2019; 25:954–61

4. McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, Back T, Chesus M, Corrado GC, Darzi A, Etemadi M, Garcia-Vicente F, Gilbert FJ, Halling-Brown M, Hassabis D, Jansen S, Karthikesalingam A, Kelly CJ, King D, Ledsam JR, Melnick D, Mostofi H, Peng L, Reicher JJ, Romera-Paredes B, Sidebottom R, Suleyman M, Tse D, Young KC, De Fauw J, Shetty S: International evaluation of an AI system for breast cancer screening. Nature 2020; 577:89–94
5. Maheshwari K, Shimada T, Yang D, Khanna S, Cywinski JB, Irefin SA, Ayad S, Turan A, Ruetzler K, Qiu Y, Saha P, Mascha EJ, Sessler DI: Hypotension prediction index for prevention of hypotension during moderate- to high-risk noncardiac surgery: A pilot randomized trial. Anesthesiology 2020; 133:1214–22
6. Hatib F, Jian Z, Buddi S, Lee C, Settels J, Sibert K, Rinehart J, Cannesson M: Machine-learning algorithm to predict hypotension based on high-fidelity arterial pressure waveform analysis. Anesthesiology 2018; 129:663–74
7. Bijker JB, van Klei WA, Kappen TH, van Wolfswinkel L, Moons KG, Kalkman CJ: Incidence of intraoperative hypotension as a function of the chosen definition: Literature definitions applied to a retrospective cohort using automated data collection. Anesthesiology 2007; 107:213–20
8. Brady K, Hogue CW: Intraoperative hypotension and patient outcome: Does "one size fit all?" Anesthesiology 2013; 119:495–7
9. Sun LY, Wijeysundera DN, Tait GA, Beattie WS: Association of intraoperative hypotension with acute kidney injury after elective noncardiac surgery. Anesthesiology 2015; 123:515–23
10. Walsh M, Devereaux PJ, Garg AX, Kurz A, Turan A, Rodseth RN, Cywinski J, Thabane L, Sessler DI: Relationship between intraoperative mean arterial pressure and clinical outcomes after noncardiac surgery: Toward an empirical definition of hypotension. Anesthesiology 2013; 119:507–15
11. Liem VGB, Hoeks SE, Mol KHJM, Potters JW, Grüne F, Stolker RJ, van Lier F: Postoperative hypotension after noncardiac surgery and the association with myocardial injury. Anesthesiology 2020; 133:510–22
12. Wijnberge M, Geerts BF, Hol L, Lemmers N, Mulder MP, Berge P, Schenk J, Terwindt LE, Hollmann MW, Vlaar AP, Veelo DP: Effect of a machine learning-derived early warning system for intraoperative hypotension vs standard care on depth and duration of intraoperative hypotension during elective noncardiac surgery: The HYPE randomized clinical trial. JAMA 2020; 323:1052–60