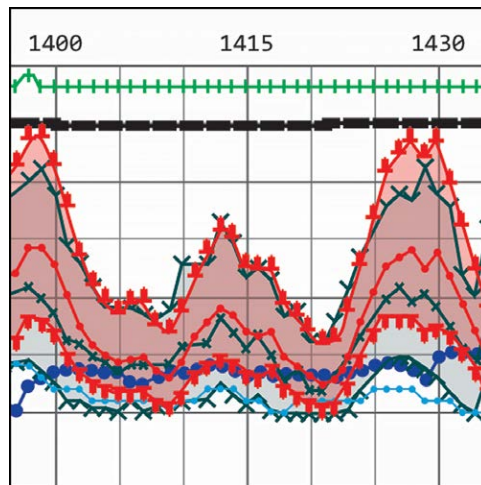


The Promise and Challenges of Predictive Analytics in Perioperative Care

Duminda N. Wijeyesundera, M.D., Ph.D., Daniel I. McIsaac, M.D., M.P.H., Martin J. London, M.D.

Intraoperative hypotension occurs frequently in clinical practice and has predictive significance. Intraoperative blood pressure below specific thresholds, including mean arterial pressure (MAP) of 65 mmHg or systolic blood pressure of 90 mmHg, is associated with elevated risks of myocardial infarction, acute kidney injury, and death.¹⁻³ Despite a growing body of research, important knowledge gaps remain. Most importantly, there remains uncertainty as to whether the association between hypotension and complications is causal. Further, we do not know which interventions can effectively reduce exposure to hypotension and also improve outcomes. Simple alarm systems appear not to help. A randomized controlled trial of 1,598 patients showed that supplemental alarms (visual alert and pager notification) for intraoperative systolic blood pressure less than 80 mmHg failed to reduce exposure to hypotension or duration of hospitalization.⁴ The lack of clinical benefit may be explained, in part, by alerts occurring only after hypotension had developed. Early warning systems for impending hypotension might plausibly allow clinicians to implement treatment strategies early and thereby reduce exposure to hypotension. One such system is commercially available: Acumen Hypotension Prediction Index software (Edwards Lifesciences, USA). Using 23 arterial waveform features measured by a pulse contour analysis monitor (FloTrac, Edwards Lifesciences), the technology alerts anesthesia providers to a high probability



“The [Hypotension Prediction] Index needs to be rigorously evaluated in high-quality validation studies that are not affected by selection bias.”

of MAP less than 65 mmHg occurring 5, 10, and 15 min in the future. In its primary industry-supported development and validation study,⁵ Hatib *et al.* screened more than 3,000 hemodynamic features to select 23 components to incorporate into the prediction index. The index was developed in a training sample of 1,334 patient records, with subsequent external validation in 204 patient records. In external validation, it showed excellent discrimination when predicting future hypotension, with an associated area under the receiver operating characteristic curve (AUC) that exceeded 0.90.⁶ Nonetheless, randomized trials comparing index-guided care to usual care have generated mixed results. Three trials, which collectively randomized 267 patients, found that index-guided care reduced exposure to hypotension,⁷⁻⁹ while another trial of 214 patients did not.¹⁰

In this issue of *ANESTHESIOLOGY*, Enevoldsen and Vistisen propose a provocative explanation for these mixed results.¹¹ Their explanation was prompted primarily by the simple observation of an atypical shape of the receiver operating characteristic curve in figure 3 of the paper by Hatib *et al.*⁵ This curve included a particular threshold value of the index, the value of which has not reported, with a sensitivity of about 55% and specificity of 100% for predicting hypotension 15 min in the future. This combination of sensitivity and specificity means that any patient meeting this threshold would always subsequently develop hypotension.

Image: J. P. Rathmell.

This editorial accompanies the article on p. 283.

Accepted for publication July 20, 2022.

Duminda N. Wijeyesundera, M.D., Ph.D.: Department of Anesthesia, St. Michael's Hospital, Toronto, Canada; Department of Anesthesiology and Pain Medicine, University of Toronto, Toronto, Canada.

Daniel I. McIsaac, M.D., M.P.H.: Department of Anesthesiology and Pain Medicine, University of Ottawa, Ottawa, Canada; Department of Anesthesiology and Pain Medicine, Ottawa Hospital, Ottawa, Canada.

Martin J. London, M.D.: University of California-San Francisco, San Francisco, California.

Copyright © 2022, the American Society of Anesthesiologists. All Rights Reserved. *Anesthesiology* 2022; 137:275-9. DOI: 10.1097/ALN.0000000000004332

Such accurate prediction is not impossible, but arguably unlikely, especially given the multiple dynamic mechanisms causing hypotension in the chaotic operative environment (e.g., blood loss, surgical manipulation, anesthesia-related vasodilation). Enevoldsen and Vistisen suggest that methodologic biases during development of the index explain this unrealistically high accuracy. They propose that selection bias led to overrepresentation of current MAP values in information used to calculate the index values, and possibly biased estimates of the ability of the index to predict future hypotension. In this editorial, we discuss the overarching principles whereby these biases affected development of the index and make recommendations for improving any future refinement and validation.

Foundationally, the Hypotension Prediction Index software is a prediction model, meaning that it uses currently available predictor information (i.e., features of the arterial waveform) to estimate the probability of an outcome occurring in the future (i.e., hypotension at 5, 10, or 15 min). When first developing a prediction model, researchers must obtain and process clinically relevant data to assemble a training (or derivation) dataset. For a prediction model to accurately predict outcomes in real-world clinical practice, the relationship between predictor (e.g., current MAP) and outcome (e.g., future hypotension) variables in the training dataset must be representative of the true relationship observed by clinicians in real-world practice. The training dataset assembled by Hatib *et al.* deviated from this assumption in two important respects.

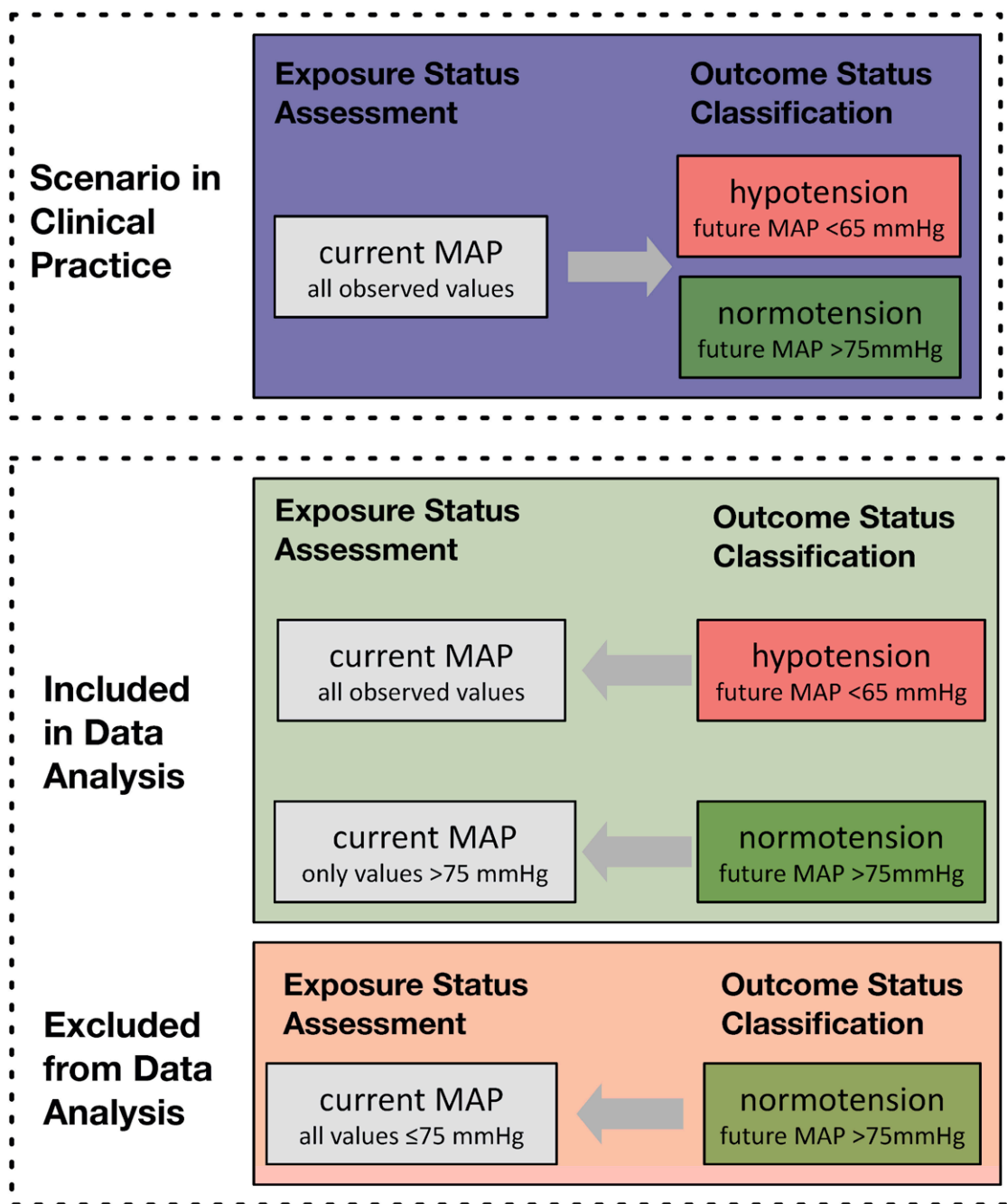
In clinical practice (fig. 1), anesthesia providers first assess a predictor variable (e.g., current MAP) and then estimate the probability of a future outcome (e.g., hypotension). The training dataset did the opposite. Patients were classified based on their outcome state (e.g., hypotension *vs.* normotension), after which temporally preceding predictor variables (e.g., current MAP) were characterized (fig. 1). This study design, in epidemiologic terms referred to as a case-control design, does not mimic the flow of information in clinical practice. Despite this difference, case-control studies can provide valid findings, if patients with (i.e., cases) and without (i.e., controls) the outcome are selected in manner that maintains the true relationship between predictor and outcome variables in the wider (e.g., clinical) population.¹²

Selection bias is the result of a dataset assembly process that distorts this true relationship. Enevoldsen and Vistisen make the critical observation of a second key problem with the original training dataset that led to significant selection bias: the operational definitions of hypotension and normotension. These definitions inadvertently restricted the allowable range of observed predictor variables based on which outcome state (hypotension *vs.* normotension) was experienced by a patient. The definition of hypotension (MAP less than 65 mmHg for more than 1 min) allowed for the full range of preceding MAP values among

patients with the hypotension outcome in the training dataset (fig. 1). *The same did not apply for the normotension outcome: normotensive episodes were defined as a continuous 30-min episode where MAP was consistently greater than 75 mmHg. All MAP values preceding an episode of normotension therefore had to exceed 75 mmHg (fig. 1).* For example, the training dataset excluded a plausible scenario in which an anesthesia provider observes a current MAP of 70 mmHg in their patient, and the MAP measured 15 min afterward was 80 mmHg. *This, and comparable scenarios in clinical practice, would have been entirely excluded from the dataset used to train the index to predict future hypotension.* Why does exclusion of these plausible scenarios matter? Stated simply, the biased dataset led the index to be taught—incorrectly—that if a patient has a current MAP less than 75 mmHg, the only foreseeable possibility at 5, 10, or 15 min in the future is that the patient will experience hypotension. Anesthesia providers will implicitly recognize that this assumption does not align with clinical reality.

These same outcome definitions were applied in the validation datasets used to test the ability of the index to predict future hypotension. In these validation datasets, the definitions artificially exaggerated differences in the range of allowable current MAP values in patients who experienced hypotension *versus* patients who experienced normotension. Consequently, the calculations used to characterize the prognostic accuracy of the index might have been substantially biased. Based on analysis of simulated data,¹¹ Enevoldsen and Vistisen show that such selection bias substantially overestimates the performance of current MAP in predicting future hypotension (AUC increased from marginally useful 0.75 to highly useful 0.93).⁶ In these same simulated data, selection bias increased the specificity of current MAP less than 75 mmHg in predicting future hypotension from about 70% to essentially perfect 100%, while sensitivity remained unchanged at about 70%.

It is important to point out that Enevoldsen and Vistisen cannot definitively prove that the prognostic performance of the prediction index is biased, as their simulation focused solely on the current MAP value, not the index (which is a proprietary algorithm). Their hypothesis is, however, indirectly supported by Jacquet-Lagrèze *et al.*, who evaluated whether linear extrapolation of two sequential MAP measurements (e.g., MAP measured 3 min apart) can predict future intraoperative hypotension (MAP less than 65 mmHg).¹³ When exposure MAP values were similarly restricted to greater than 75 mmHg in patients who experienced subsequent normotension, the performance of sequential MAP measurement in predicting future hypotension was substantially increased (e.g., AUC increased from 0.69 to 0.88). This clinical study strongly suggests that the issues raised by Enevoldsen and Vistisen cannot be ignored. The index needs to be rigorously evaluated in high-quality validation studies that are not affected by selection bias. *These validation data must include the full range of observed values for predictor variables (including current*



Downloaded from <http://asa2.silverchair.com/anesthesiology/article-pdf/137/3/275/674456/20220900-0-00007.pdf> by guest on 18 April 2024

Fig. 1. Application of predictive information by anesthesia providers in clinical practice, with comparison to process used to assemble the training dataset for the prediction index. The *top blue box* denotes how anesthesia providers consider predictive information in clinical practice. Patients are classified by their current exposure status (current mean arterial pressure [MAP]), which is used to estimate the probability of a future outcome (future hypotension). In this example, we have defined hypotension (future MAP less than 65 mmHg) and normotension (future MAP greater than 75 mmHg) based on thresholds used by Hatib *et al.*; however, anesthesia providers may use individualized and multicategorical thresholds in clinical practice.⁵ In the datasets used to train the index, patients were classified based on their outcome status as hypotension or normotension, after which the preceding prognostic indicator (current MAP) was characterized. While the dataset (*green*) included all observed current MAP values for patients classified as hypotensive, it only included current MAP values greater than 75 mmHg for patients classified as normotensive. The selective exclusion of current MAP values 75 mmHg or less in patients classified as normotensive controls led to selection bias (*orange*). The relationship between the exposure (current MAP) and outcome (hypotension) in the dataset was no longer representative of the true relationship seen in clinical practice.

MAP), regardless of whether the subsequent future outcome is hypotension versus normotension. Selection bias will be minimized by ensuring that any association between current MAP and future hypotension in these validation datasets is explained by the true relationship observed in clinical practice, not by the way the data were assembled. Furthermore, given emerging evidence that the current MAP alone may predict future hypotension,¹³ these validation studies should directly compare the prognostic performance of current MAP against the index. By comparison, Hatib *et al.* compared the index against recent change in MAP over a specified interval (e.g., 5 min).⁵ As a possible predictor of future hypotension, recent change in MAP is conceptually problematic; for example, the same absolute change in MAP from 100 to 90 mmHg is unlikely to portend impending hypotension in a manner similar to a change from 80 to 70 mmHg. Further, recent change in MAP appears to be inferior to both current MAP and linear extrapolation of MAP in predicting future hypotension.¹³ If the index is found to have lower prognostic performance than originally estimated, opportunities exist for its further refinement and improvement, which are both feasible and important. Once the selection bias in the original study is addressed, high-quality application of recommended prediction modeling methods can help identify which other complex hemodynamic parameters best augment predictive information from current MAP, thus allowing the technology to reach its true potential as an advanced predictive tool.¹⁴

What do these findings mean for anesthesia providers in the operating room? Pending further properly designed validation studies, clinicians can reasonably still use the prediction index in clinical practice with the caveat that its prognostic accuracy may be lower than initially projected. This reduced accuracy will most likely manifest as false positives, where anesthesia providers are prompted to consider mitigation interventions (e.g., IV fluid bolus, vasoactive drugs) in individuals who are unlikely to develop hypotension. The clinical importance of any such “overtreatment” remains to be determined. For predictive analytics to meaningfully improve the management of intraoperative hypotension, Enevoldsen and Vistisen have highlighted the ongoing need for high-quality epidemiologic study design, sophisticated analytical methods, careful validation by external studies, prospective assessment by randomized trials, and considered interpretation by astute clinicians.

Research Support

Dr. Wijeyesundera is supported in part by a Merit Award from the Department of Anesthesiology and Pain Medicine at the University of Toronto, Toronto, Canada, and the Endowed Chair in Translational Anesthesiology Research at St. Michael's Hospital, Toronto, Canada, and the University of Toronto. Dr. McIsaac receives salary support from the Ottawa Hospital Anesthesia Alternate Funds Association, Ottawa, Canada, and a Research Chair from the Faculty of Medicine and the University of Ottawa, Ottawa, Canada.

Competing Interests

Dr. Wijeyesundera is a member of the Scientific Advisory Board for Surgical Safety Technologies, Toronto, Canada, and has received honoraria from Edwards Lifesciences, Irvine, California, for participation in an advisory board panel. The other authors declare no competing interests.

Correspondence

Address correspondence to Dr. Wijeyesundera: d.wijeyesundera@utoronto.ca

References

1. Salmasi V, Maheshwari K, Yang D, Mascha EJ, Singh A, Sessler DI, Kurz A: Relationship between intraoperative hypotension, defined by either reduction from baseline or absolute thresholds, and acute kidney and myocardial injury after noncardiac surgery: A retrospective cohort analysis. *ANESTHESIOLOGY* 2017; 126:47–65
2. Wesselink EM, Kappen TH, Torn HM, Slooter AJC, van Klei WA: Intraoperative hypotension and the risk of postoperative adverse outcomes: A systematic review. *Br J Anaesth* 2018; 121:706–21
3. Ahuja S, Mascha EJ, Yang D, Maheshwari K, Cohen B, Khanna AK, Ruetzler K, Turan A, Sessler DI: Associations of intraoperative radial arterial systolic, diastolic, mean, and pulse pressures with myocardial and acute kidney injury after noncardiac surgery: A retrospective cohort analysis. *ANESTHESIOLOGY* 2020; 132:291–306
4. Panjasawatwong K, Sessler DI, Stapelfeldt WH, Mayers DB, Mascha EJ, Yang D, Kurz A: A randomized trial of a supplemental alarm for critically low systolic blood pressure. *Anesth Analg* 2015; 121:1500–7
5. 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
6. Staffa SJ, Zurakowski D: Statistical development and validation of clinical prediction models. *ANESTHESIOLOGY* 2021; 135:396–405
7. 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
8. Schneck E, Schulte D, Habig L, Ruhrmann S, Edinger F, Markmann M, Habicher M, Rickert M, Koch C, Sander M: Hypotension Prediction Index based protocolized haemodynamic management reduces the incidence and duration of intraoperative hypotension in

- primary total hip arthroplasty: A single centre feasibility randomised blinded prospective interventional trial. *J Clin Monit Comput* 2020; 34:1149–58
9. Tsoumpa M, Kytteri A, Matiatou S, Tzoufi M, Griva P, Pikoulis E, Riga M, Matsota P, Sidiropoulou T: The use of the Hypotension Prediction Index integrated in an algorithm of goal directed hemodynamic treatment during moderate and high-risk surgery. *J Clin Med* 2021; 10:5884
 10. 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. *ANESTHESIOLOGY* 2020; 133:1214–22
 11. Enevoldsen J, Vistisen ST: Performance of the Hypotension Prediction Index may be overestimated due to selection bias. *ANESTHESIOLOGY* 2022; 137:283–89
 12. Sutton-Tyrrell K: Assessing bias in case-control studies. Proper selection of cases and controls. *Stroke* 1991; 22:938–42
 13. Jacquet-Lagrèze M, Larue A, Guilherme E, Schweizer R, Portran P, Ruste M, Gazon M, Aubrun F, Fellahi JL: Prediction of intraoperative hypotension from the linear extrapolation of mean arterial pressure. *Eur J Anaesthesiol* 2022; 39:574–81
 14. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, Denniston AK, Faes L, Geerts B, Ibrahim M, Liu X, Mateen BA, Mathur P, McCradden MD, Morgan L, Ordish J, Rogers C, Saria S, Ting DSW, Watkinson P, Weber W, Wheatstone P, McCulloch P; DECIDE-AI Expert Group: Reporting guideline for the early stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *BMJ* 2022; 377:e070904