ANESTHESIOLOGY
Trusted Evidence: Discovery to Practice®

Journal Pre-proof

The Hypotension Prediction Index is equally effective in predicting intraoperative hypotension during non-cardiac surgery compared to a mean arterial pressure threshold: a prospective observational study

Marijn P. Mulder, M.Sc., Mirjam Harmannij-Markusse, M.Sc., Libera Fresiello, Ph.D., Dirk W. Donker, M.D., Ph.D., Jan-Willem Potters, M.D., Ph.D.

DOI: https://doi.org/10.1097/ALN.0000000000004990

To appear in: Anesthesiology

Submitted for publication: September 26, 2023

Accepted for publication: March 22, 2024

Please cite this article as: Mulder MP, Harmannij-Markusse M, Fresiello L, Donker DW, Potters J-W: The Hypotension Prediction Index is equally effective in predicting intraoperative hypotension during non-cardiac surgery compared to a mean arterial pressure threshold: a prospective observational study. Anesthesiology. 2024; https://doi.org/10.1097/ALN.0000000000004990

Anesthesiology Publish Ahead of Print

DOI: 10.1097/ALN.0000000000004990

The Hypotension Prediction Index is equally effective in predicting intraoperative hypotension during non-cardiac surgery compared to a mean arterial pressure threshold: a prospective observational study

Marijn P. Mulder, M.Sc.Cardiovascular and Respiratory Physiology, TechMed Centre, University of Twente, Enschede, The Netherlands, technical physician, PhD student, m.p.mulder-1@utwente.nl

Mirjam Harmannij-Markusse, M.Sc., Technical Medicine, University of Twente, Enschede, The Netherlands, technical physician, mirjam.markusse@live.nl mailto:

Libera Fresiello, Ph.D., Cardiovascular and Respiratory Physiology, TechMed Centre, University of Twente, Enschede, The Netherlands, biomedical engineer, assistant-professor, l.fresiello@utwente.nl

Dirk W. Donker, M.D., Ph.D., Cardiovascular and Respiratory Physiology, TechMed Centre, University of Twente, Enschede, The Netherlands & Intensive Care Center, University Medical Center Utrecht, Utrecht, The Netherlands, cardiologist-intensivist, professor, d.w.donker@utwente.nl

Jan-Willem Potters, M.D., Ph.D., Department of Anesthesiology, Medisch Spectrum Twente, Enschede, The Netherlands, anesthesiologist, janwillem.potters@mst.nl

Corresponding Author: Marijn P. Mulder, Cardiovascular and Respiratory Physiology, TechMed Center, Faculty of Science and Technology, University of Twente, Technohal 3184, P.O. box 217, 7500 AE Enschede, The Netherlands, email: m.p.mulder-1@utwente.nl, phone: +31 53 489 5625

Clinical trial number and registry URl, if applicable: Not applicable

Prior Presentations: Not applicable

Acknowledgments: Not applicable

Word and Element counts: Abstract: 295 words, Introduction: 376 words, Discussion: 1556 words, Figures: 4, Tables: 4, Supplemental Digital Content: 1

Abbreviated Title (Running Head): Hypotension Prediction Index versus MAP thresholds

Summary Statement: Not applicable

Funding Statement: Support was provided solely from institutional and/or departmental sources.

Conflicts of Interest: MPM, DWD and LB provide research consultancy to Maquet Critical Care AB, DWD and LB have a research corporation with Sonion BV, Hoofddorp, The Netherlands and DWD provides research consultancy to HBOX Therapies GmbH, Aachen, Germany. None of these authors receive personal fees, all other authors declare no competing interests.

Abstract (297/300 words)

Background: The Hypotension Prediction Index is designed to timely predict intraoperative hypotension and is based on arterial waveform analysis using machine learning. It has recently been suggested that this algorithm is highly correlated with the mean arterial pressure (MAP) itself. Therefore, the aim of this study was to compare the Index with MAP based prediction methods and it is hypothesized that their ability to predict hypotension is comparable.

Methods: In this observational study, the Hypotension Prediction Index was used in addition to routine intraoperative monitoring during moderate- to high-risk elective non-cardiac surgery. The agreement in time between the default Hypotension Prediction Index alarm (>85) and different concurrent MAP thresholds was evaluated. Additionally, the predictive performance of the Index and different MAP based methods were assessed within five, ten and fifteen minutes before hypotension occurred.

Results: A total of 100 patients were included. A MAP threshold of 73 mmHg agreed 97% of the time with the default Index alarm, while a MAP threshold of 72 mmHg had the most comparable predictive performance. The areas under the receiver operating characteristic curve of the Hypotension Prediction Index (0.89 (0.88-0.89)) and concurrent MAP (0.88 (0.88-0.89)) were almost identical for predicting hypotension within five minutes, outperforming both linearly extrapolated MAP (0.85 (0.84-0.85)) and delta MAP (0.66 (0.65-0.67)). The positive predictive value was 31.9 (31.3–32.6)% for the default Index alarm and 32.9 (32.2–33.6)% for a MAP threshold of 72 mmHg.

Conclusion: In clinical practice, the Hypotension Prediction Index alarms are highly similar to those derived from MAP, which implies that the machine learning algorithm could be substituted by an alarm based on a MAP threshold set at 72 or 73 mmHg. Further research on intraoperative hypotension prediction should therefore include comparison with MAP based alarms and related effects on patient outcome.

Introduction (376/500 words)

Intraoperative hypotension is common in surgical patients and associated with an increased risk for postoperative mortality, acute kidney injury and myocardial injury[1,2]. A method to predict hypotension could help anesthesiologists to treat hypotension in a proactive instead of a reactive way, and thus ultimately prevent it. The Hypotension Prediction Index has recently been introduced as an innovative method, consisting of a validated machine learning model based on the characteristics of the arterial waveform[3]. It is specifically developed to predict hypotension, defined as a mean arterial pressure (MAP) below 65 mmHg for at least one minute, up to fifteen minutes in advance[3].

Recently, a modelling study hypothesized that the performance of the Hypotension Prediction Index may be overestimated due to the data selection process in the development phase of the machine learning algorithm, which can introduce systematic bias[4]. This elaborate theoretical underpinning further substantiates the notion that MAP may indirectly be overrepresented in the Index. This is also supported by our recent observational pilot study revealing a high cross-correlation and no time difference between the Hypotension Prediction Index and MAP signals[5]. It is conceivable that with MAP available, the added predictive ability of the Hypotension Prediction Index algorithm is relatively limited, as it largely represents an inverse reflection of the concurrent MAP.

The predictive performance of the Hypotension Prediction Index was compared to several standard clinical hemodynamic variables, including MAP[6]. For the Index, an area under the curve (AUC) of the receiver operating characteristic of 0.926 was reported for the prediction of hypotension five minutes in advance. The concurrent MAP was found to be the best predictor among the standard hemodynamic variables tested (AUC of 0.807)[6]. Yet, in a recent erratum to this publication it was clarified that different data analyses were applied for the Hypotension

Prediction Index and other tested variables and therefore it could not be concluded that the Index outperforms concurrent MAP[7]. This erratum was not supported by numeric data or new results, so a formal direct comparison between the Hypotension Prediction Index and MAP is still lacking.

Therefore, we set out to compare the Hypotension Prediction Index with MAP based predictions in a large group of moderate- to high-risk non-cardiac surgery patients and hypothesized that respective alarms are equally effective in predicting hypotension.

Materials and methods

This single center prospective, observational cohort study was conducted within the department of Anesthesiology of the Medisch Spectrum Twente, Enschede, the Netherlands, a tertiary referral hospital. This manuscript is prepared in accordance with the STROBE and STARD reporting guidelines[8,9]. Ethical approval for this study was waived by the local ethical committee (#K22-42). All consecutive adult patients aged eighteen years or older and undergoing moderate- to high-risk, non-cardiac, elective surgery were included. All patients were scheduled for and received general anesthesia with the need for invasive blood pressure monitoring by means of an arterial line. The inclusion period ranged from March 2022 until July 2023 while aiming for a convenience sample size of 100 patients. There was no follow-up of this cohort. A sub analysis of the first 33 patients from this cohort was published in a research letter before[5].

In addition to standard care in line with current practice guidelines, patients were monitored with the HemoSphere advanced monitor including the Acumen Hypotension Prediction Index software (version 2.1, Edwards Lifesciences, Inc, Irvine, USA). The HemoSphere advanced monitor, including all available additional parameters and secondary screen, was not blinded.

The attending anesthesiologists could manage blood pressure according to their preferences; no specific treatment protocol was used. Demographic data, medical history and details on the surgery performed were obtained from the electronic health records of the patients enrolled. The measured and averaged MAP (in mmHg) and computed Hypotension Prediction Index (unitless, ranging from 0 to 100) were downloaded from the HemoSphere monitor, with 20-seconds interval samples.

Data analysis was performed in Python (version 3.9, Python Software Foundation, Beaverton, Orlando, USA). The Python script, featuring a documented, step-by-step methodology along with example data is available at https://github.com/crph-utwente/HPIvalidation. All data from the start of incision till the end of surgery was used for the analysis; no data was removed. Missing data and artifacts, where a single data point dropped ≥30 mmHg[3] below its surrounding values, were replaced by a linear interpolation. The analysis included the incidence and severity of hypotension, defined as a MAP <65 mmHg for at least one minute[3]. All other 20 second datapoints were defined as non-hypotensive. The 'area under the threshold' was calculated by multiplying the depth of hypotension (in mmHg) below a MAP of 65 mmHg by the duration of time spent in hypotension (in minutes). This calculation yields an area value that takes both depth and duration of hypotensive events into account. Additionally, the area under the threshold was normalized by dividing it by the total duration of surgery, resulting in the 'time-weighted average' in hypotension.

Statistical analysis

All statistical analysis have been performed in RStudio (version 2023.09.1, RStudio, PBC, Boston, MA, USA). Descriptive statistics are presented as mean ± SD for normally distributed data, or median [25^{th}, 75^{th} percentiles] for not normally distributed data. Normality of distribution was assessed visually with histograms and Q-Q plots, categorical data are presented as n (%).

A cross-correlation analysis between the time series of the Hypotension Prediction Index and the MAP was conducted per patient. This is a technique commonly used in the fields of signal processing and statistical time series analysis, to quantify the similarity between the two signals as a function of the time shift relative to each other[10,11]. Both the Index and MAP signals were normalized for this analysis by subtracting their means and dividing by their SDs. Since this method incorporates the temporal relation between data points, it is also used to estimate the time delay between two signals. The cross-correlation at zero time shift is essentially Pearson's correlation. Since a non-linear, but monotonic, relation was found between the Hypotension Prediction Index and MAP samples (see Figure 2) and the residuals generated during the correlation estimation process were not normally distributed, performing merely Pearson's correlation may formally not be completely sufficient. Therefore, a Spearman's rank correlation was also calculated to better capture the relationship between the Hypotension Prediction Index values and the concurrent MAP values at zero time shift. The latter two correlation measures were calculated on a per patient basis and for the entire pooled dataset.

Additionally, the timewise relationship between the alarms derived from the Hypotension Prediction Index and different MAP thresholds was evaluated. The Hypotension Prediction Index alarm was defined at the default value >85, while different MAP thresholds were defined at concurrent values ranging from 65 to 80 mmHg retrospectively. The proportionate agreement between different MAP thresholds and the Index alarm, defined as percentage of time the thresholds were surpassed or not and the alarms were simultaneously on or off, was calculated. Also, the Cohen's kappa coefficient was analyzed, which takes into account the possibility of this agreement occurring by chance.

The performance of predicting hypotension within five, ten and fifteen minutes in advance was evaluated for the Hypotension Prediction Index and concurrent MAP. The prediction windows were defined as follows: 1-5 minutes, 1-10 minutes and 1-15 minutes. The first one minute was excluded from the prediction windows, because this short prediction time is not deemed clinically relevant when aiming to prevent intraoperative hypotension by starting an intervention. The performance was assessed by means of a receiver operating characteristic curve and the area under this curve (AUC) including 95% CI using bootstrap method with 2000 repetitions. These were constructed using a generalized estimating equations model which allowed to correct for repeated measures within a single patient[12]. Comparably, also precision recall curves were constructed, where sensitivity (recall) is plotted against positive predictive value (precision) for each threshold value. The latter curves are more commonly used to assess the performance of tests in an imbalanced dataset, since they hold more clinical relevance. For different Hypotension Prediction Index and MAP thresholds sensitivity, specificity and positive and negative predictive values were assessed. Additionally, the threshold values at a minimal difference between sensitivity and both specificity and positive predictive value were defined.

For the five minutes prediction window also the performance of delta MAP (ΔMAP) and linearly extrapolated MAP was assessed[13]. ΔMAP was defined as the difference between the concurrent MAP and the MAP five minutes before. The linearly extrapolated MAP served as prediction of the MAP five minutes ahead and was determined on a trendline between concurrent MAP and MAP five minutes before, assuming linearity (Figure 1).

A sensitivity analysis was performed to evaluate the potential effect of different data selection methods, (non)hypotension definitions and prediction windows.

To achieve a more clinically relevant positive predictive value estimate a forward analysis was conducted[14]. In this forward analysis alarms were defined as a Hypotension Prediction Index >85 or MAP <72 mmHg for at least a minute. From the onset of the alarm, i.e. the first sample, the presence of hypotension was searched in the upcoming 1-5, 1-10 and 1-15 minutes windows. The positive predictive value was defined as the percentage of alarms truly followed by hypotension. Additionally, this forward analysis was repeated, but then excluding alarms that were followed by a sudden increase in MAP, to correct for possible anti-hypotensive treatments that might conceal the potential predictive performance by preventing hypotension. Sudden increases in MAP were defined as MAP increase of >5 mmHg in 20 seconds or >8 mmHg in 2 minutes, when MAP was initially <75 mmHg[6].

Results

A total of 100 patients were included in the study, no patients were excluded from the analysis. Table 1 shows the baseline characteristics of the study population: 49/100 (49%) American Society of Anesthesiologists (ASA) Physical Status II, 42/100 (42%) ASA III and 9/100 (9%) ASA IV patients, with hypertension as the most common chronic disease (49/100, 49%), followed by diabetes mellitus (24/100, 24%). The mean age was 68 ± 11 years and about half of the patients (51/100, 51%) underwent gastrointestinal surgery. The total amount of datapoints analyzed was 68,583, with 2,061 (3.0%) defined as hypotensive, grouped into 231 hypotensive events. Table 2 shows the hypotension metrics for this study population. Hypotension was present in 62/100 (62%) patients, with an area under the threshold of 5 [0, 35] mmHg·min and time weighted average of 0.03 [0.0, 0.2] mmHg.

The maximal cross-correlation between the Hypotension Prediction Index signal and MAP signal was -0.91 ± 0.05 with 0.0 ± 0.0 minutes time shift. This indicates a strong inverse relation between the two signals and no time delay. For the pooled dataset of the Index and concurrent

MAP data points the Pearson's correlation was -0.88. A sigmoidal monotonic relation between Hypotension Prediction Index and concurrent MAP values is visualized in the scatter plot in Figure 2. The Spearman's rank correlation coefficient between the Index and concurrent MAP values per patient was -0.99 ± 0.01, with a -0.97 correlation for the pooled dataset.

In Figure 2, also the default Index alarm and the most corresponding MAP thresholds (72 and 73 mmHg) resulting from the analyses detailed below are indicated. At least one Hypotension Prediction Index alarm was detected in 89/100 (89%) patients, with a concurrent MAP of 71 ± 4 mmHg at the alarm onset (histogram displayed in Supplemental Digital Content 1 Figure S1, https://links.lww.com/ALN/D530). The proportionate agreement between the default Index alarm and the different MAP thresholds (65-80 mmHg) ranges between 0.79 and 0.97. The specific results of the agreement and Cohen's kappa analysis can be found in Table 3 and more details in Supplemental Digital Content 1 Table S1 (https://links.lww.com/ALN/D530). The highest proportionate agreement (0.97) and Cohen's kappa coefficient (0.85) were found with a MAP threshold of 73 mmHg, indicating this value has the highest agreement with the default Hypotension Prediction Index alarm. In Figure 3 the agreement between these alarms is displayed in time frames of two representative hypotensive patients.

In Figure 4a the receiver operating characteristic curves of the Hypotension Prediction Index, concurrent MAP, ΔMAP and linearly extrapolated MAP to predict hypotension up to five minutes are shown. Hypotension Prediction Index and concurrent MAP receiver operating characteristic curves are almost identical, with an AUC of respectively of 0.89 (0.88-0.89) and 0.88 (0.88-0.89), yielding a higher AUC than the other two methods: AUC of ΔMAP is 0.66 (0.65-0.67) and AUC of linearly extrapolated MAP is 0.85 (0.84-0.85). Figure 4b shows the precision recall curves for the four prediction methods within five minutes. Again, Hypotension Prediction Index and concurrent MAP are similar with overlapping CI of the AUC, namely 0.56

(0.55-0.57) and 0.55 (0.54-0.56). Linearly extrapolated MAP has a substantially lower precision recall AUC of 0.32 (0.32-0.33). The coinciding performance curves for the Index and concurrent MAP are also present at the other prediction windows. In Supplemental Digital Content 1 Figure S2-3 (https://links.lww.com/ALN/D530) receiver operating characteristic and precision recall curves for the ten and fifteen minutes prediction windows can be found.

The performance metrics of a selection of the Hypotension Prediction Index alarms with corresponding MAP thresholds and the other MAP based methods are reported in Table 4. When predicting hypotension within five minutes, the default HPI alarm has a sensitivity of 72.4 (71.1-73.8)%, specificity of 89.9 (89.7–90.1)% and positive predictive value of 31.9 (31.3–32.9)%. The MAP threshold set at 72 mmHg was the most comparable to the default Hypotension Prediction Index alarm, gaining a 72.6 (71.3–73.9)% sensitivity, 90.3 (90.1–90.5)% specificity and 32.9 (32.2–33.6)% positive predictive value. When balancing sensitivity and specificity the thresholds are 63 for the Hypotension Prediction Index, 75 mmHg for concurrent MAP, -3 mmHg for ΔMAP and 73 mmHg for linearly extrapolated MAP. Those thresholds are 97 for the Hypotension Prediction Index, 67 mmHg for concurrent MAP, -16 mmHg for ΔMAP and 57 mmHg for linearly extrapolated MAP when aiming for a minimal difference between sensitivity and positive predictive value. A more complete overview of the performance metrics for other Hypotension Prediction Index and MAP thresholds at different prediction windows can be found in Supplemental Digital Content 1 Table S2-4 (https://links.lww.com/ALN/D530).

The sensitivity analysis revealed that absolute performance of both HPI and MAP based prediction methods are affected by data processing, while they still perform equally, both being affected in the same way. The performance curves for different data selection and prediction times are displayed in Supplemental Digital Content 1 Figure S4-5 (https://links.lww.com/ALN/D530). At times, we observed an oscillatory behavior in the precision recall curves at low, not clinically relevant, sensitivity, which is caused by few data points with extremely low Hypotension Prediction Index or MAP values.

The results of the forward analysis are shown in Table 5. The time to hypotension within 15 minutes from start of a Hypotension Prediction Index alarm was 2.3 [1.0, 7.3] minutes. The longer the prediction window, the higher the positive predictive value, all the values increase with correction for anti-hypotensive interventions. The positive predictive values for the Hypotension Prediction Index range between 24.1-56.8%, for a MAP alarm of 72 mmHg they are systematically slightly higher, ranging between 27.8-66.6%.

Discussion (1556/1500 words)

This observational clinical investigation in 100 surgical patients demonstrates a high correlation between the Hypotension Prediction Index and MAP values in real-world intraoperative practice. Specifically, the MAP thresholds of 72 and 73 mmHg were found to correspond closely to the default Hypotension Prediction Index alarm. When comparing the predictive performance of those methods using AUCs, those for Hypotension Prediction Index and concurrent MAP are virtually identical, while outperforming ΔMAP and linearly extrapolated MAP.

<u>Comparison of Hypotension Prediction Index and MAP</u>

The cross-correlation analysis revealed that the MAP signal is highly similar to the Index signal, being largely its mirrored version, as suggested before[5,15–17]. When accounting for the non-linear relation between the two values an even higher correlation was found (Spearman's $\rho = -0.99$). This may well be due to a presumed selection bias in the training of the machine learning model[4] and the fact that MAP is related to the arterial pressure waveform features from which the Index is derived[3]. Since no time delay between the two signals was found at maximum correlation, the Hypotension Prediction Index does not seem to be ahead of concurrent MAP in time. The cross-correlation analysis performed on a subgroup published earlier is confirmed by the results found in this larger study population[5]. These results indicate that the Hypotension Prediction Index value, although based on multiple waveform features, does essentially not comprise more or different predictive information than the MAP itself.

Hypotension Prediction Index was originally not compared to concurrent MAP, but only to ΔMAP, which showed AUCs of 0.50-0.62 for predicting hypotension[3]. This low predictive power of ΔMAP is confirmed in the present study; ΔMAP being a parameter which does not take the actual absolute blood pressure into account. Instead, a linear extrapolation of the MAP, as suggested recently, includes both the change and absolute value of MAP[13] and shows a substantially higher performance. Therefore, linearly extrapolated MAP seems a better comparator than ΔMAP when evaluating machine learning models or other innovative methods to predict hypotension, but it still performs substantially less than concurrent MAP.

In our present study, we directly compared Hypotension Prediction Index and concurrent MAP using identical data and methodology and found that the receiver operating characteristic curves are coinciding. This indicates that not only the overall performance of the two prediction methods is the same, as reflected by the AUCs, but that for every specific alarm threshold the same sensitivity and specificity are achieved. There is no added predictive value of the Index over concurrent MAP for any chosen threshold. This is in agreement with the receiver operating characteristic curves in the supplementary material of a recent study on the Hypotension Prediction Index performance in liver transplant surgery[18]. In addition, the precision recall curves of the Index and concurrent MAP are also almost identical. The sensitivity analyses revealed that despite the spectrum of studied definitions, data selection and prediction windows, the Index and MAP perform similar. This clinical study highlights the importance of a proper and accurate comparison between the Hypotension Prediction Index and other clinical measures to prevent ambiguous assumptions about its added value.

Strengths and limitations

During this clinical study no specific protocol was used for treating hypotension, allowing to assess the net effect of the presence of the Hypotension Prediction Index. This resulted in real-world data of a large group of patients showing the potential of the Hypotension Prediction Index in the intended clinical context of use and without strict boundaries imposed by a rigid research protocol. We realize that our study represents an observational, non-blinded analysis, which might have led the attending anesthesiologists to act upon Hypotension Prediction Index alarms. The resulting performance measures, especially positive predictive value, may therefore be underestimated since hypotension might have been prevented by timely intervention. We aimed to retrospectively correct for this in the forward analysis. This limits general conclusions about the absolute performance of the Index and comparison with blinded studies. However, MAP is never blinded in a routine clinical setting, for a proper comparison of both performances it is therefore beneficial that the Index is also visible to the clinician.

Compared to the original Hypotension Prediction Index paper, the AUCs of the Index's performance found in this study for five, ten and fifteen minutes prediction are somewhat lower than reported before, where all AUCs were above 0.95[3]. In this study all data points were included in order to prevent a selection bias, as recently underscored in this context[4], and consequently reducing the performance of the Index. In addition, we used prediction windows instead of a single prediction time point in the future. The current approach may thus result in a more accurate reflection of the Index performance in everyday clinical practice.

Although our study is confined to a single center analysis, its adequate representation of real-world clinical data is underlined by a comparison to the recently published European multicenter prospective observational registry: EU-HYPROTECT[20]. This registry contains data of 749 surgeries in which the Hypotension Prediction Index was also used in a non-blinded fashion and without a treatment protocol, showing comparable hypotension metrics: 59% incidence with 1 [0, 3] event per patient, 2 [0, 9] minutes duration and 0.03 [0.00, 0.20] mmHg time weighted average[20] (see also Supplemental Digital Content 1 Figure S1, https://links.lww.com/ALN/D530).

Clinical implications

Probably, the most clinically relevant measure is the positive predictive value resulting from the forward analysis. This indicates the likelihood of hypotension occurring upon the onset of an alarm, which would be the trigger to consider a therapeutic intervention. Our results reveal that only around a quarter to half of the Hypotension Prediction Index alarms are actually followed by hypotension. Despite performing a retrospective correction of possible preventive interventions, the positive predictive value still remains rather limited. The MAP alarm of 72 mmHg even seems to perform slightly better than the default Hypotension Prediction Index alarm. The positive predictive values for the Index in this study are lower than in previous validation studies using a prediction window of 0-20 minutes[14,19].

Overall, the results of this study imply that the Index's performance to predict hypotension can equally be achieved by adjusting MAP thresholds to higher set points than usually applied in conventional clinical routine. Notably, the default Hypotension Prediction Index alarm is fixed at 85, while theoretically a better balance between sensitivity and specificity could be achieved by setting a much lower alarm value (46-63, see Table 4). This might not be adequate in a clinical setting, potentially inducing alarm fatigue and overtreatment from false positive alarms. Nevertheless, the Hypotension Predicting Index alarm could be set lower to improve sensitivity, for example at 75 as it is currently proposed for the planned HYPE-2 trail[21], or alternatively concurrent MAP thresholds could be set higher (75-79 mmHg, see Table 4), as supported by our data. Depending on the clinical need also a higher Index threshold (90-97), or lower MAP (67-71 mmHg), could be chosen to achieve a better precision.

Additionally, an important limitation of the Hypotension Prediction Index is that it can only be used to predict hypotension defined as a MAP <65 mmHg, while MAP thresholds can easily be adjusted to timely predict any desired blood pressure. Therefore, intraoperative alarms based on MAP thresholds seem more suitable than the Index, when aiming for patient-specific blood pressure targets in relation to the patient's baseline blood pressure, co-morbidities and type of surgery in order to balance the risks of organ ischemia and bleeding[22]. Besides, the predictive value of the MAP thresholds found in this study are limited to our specific population, optimal thresholds could be different for patient groups with different age, gender and ASA score[23].

In general, innovative methodologies like the Hypotension Prediction Index are deemed beneficial to prevent intraoperative hypotension, as reflected by a recent systematic review[24]. The data synthesis of several randomized controlled clinical trials[25–29] showed a decrease in the occurrence, duration and severity of hypotension during non-cardiac surgery while using the Index, yet, confined to limited and low- to very-low-quality evidence[24]. It still needs to be verified whether a MAP based alarm would yield the same hypotension reduction in clinical practice as the Hypotension Prediction Index in every single patient at every single moment. The current study was not designed to investigate hypotension reduction, treatment decisions or the use of any other parameters available the HemoSphere platform. Though, hypotension was found to be less prevalent than in both the control and intervention groups of the randomized clinical trials as reflected by different hypotension metrics[24]. However, this might also be based on a psychological effect that arises from the presence of the Hypotension Prediction Index technique, its novelty and potentially induced increased awareness, leading to more proactive hemodynamic management. The same phenomenon could be present in those randomized controlled trails.

Moreover, since most studies used a treatment protocol in addition to the Hypotension Prediction Index, it remains unresolved which of the two would lead to improved outcomes.

Conclusion

Hypotension Prediction Index and concurrent MAP values are highly correlated, their respective alarms and thresholds show strong agreement and interchangeable performance in predicting intraoperative hypotension. A MAP threshold set at 72 or 73 mmHg seems a straightforward clinical alternative to the Hypotension Prediction Index. It is important to compare new hypotension prediction methods against parameters derived from routine hemodynamic monitoring, notably concurrent MAP, to study their added value including patient outcome.

Supplemental Digital Content

Additional results: Figures and Tables, https://links.lww.com/ALN/D530

References (29/50 references)

1. Bijker JB, Klei WA Van, Kappen TH, Wolfswinkel L Van, Moons KGM, 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
2. Wesselink EM, Kappen TH, Torn HM, Slooter AJC, Klei WA van: Intraoperative hypotension and the risk of postoperative adverse outcomes: a systematic review. Br J Anaesth 2018; 121:706–21
3. 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
4. Enevoldsen J, Vistisen ST: Performance of the Hypotension Prediction Index May Be Overestimated Due to Selection Bias. Anesthesiology 2022; 137:283–9
5. Mulder MP, Harmannij-Markusse M, Donker DW, Fresiello L, Potters J-W: Is Continuous Intraoperative Monitoring of Mean Arterial Pressure as Good as the Hypotension Prediction Index Algorithm?: Research Letter. Anesthesiology 2023; 138:657–8
6. Davies SJ, Vistisen ST, Jian Z, Hatib F, Scheeren TWL: Ability of an Arterial Waveform Analysis-Derived Hypotension Prediction Index to Predict Future Hypotensive Events in Surgical Patients. Anesth Analg 2020; 130:352–9
7. Ability of an Arterial Waveform Analysis–Derived Hypotension Prediction Index to Predict Future Hypotensive Events in Surgical Patients: Erratum. Anesth Analg 2023; 137:e33–e33

8. Elm E von, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP: The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008; 61:344–9

9. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, Lijmer JG, Moher D, Rennie D, Vet HCW De, Kressel HY, Rifai N, Golub RM, Altman DG, Hooft L, Korevaar DA, Cohen JF: STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015; 351

10. Stoica Petre, Moses RL: Spectral analysis of signals. Upper Saddle River, NJ, Prentice Hall, 2005

11. Chatfield C: The Analysis of Time Series : An Introduction, Sixth Edition. Chapman and Hall/CRC, 2003 doi:10.4324/9780203491683

12. Hardin J.W. HJM: Generalized Estimating Equations. New York, Chapman and Hall, 2003

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. Wijnberge M, Ster BJP Van Der, Geerts BF, Beer F De, Beurskens C, Emal D, Hollmann MW, Vlaar APJ, Veelo DP: Clinical performance of a machine-learning algorithm to predict intra-operative hypotension with noninvasive arterial pressure waveforms: A cohort study. Eur J Anaesthesiol 2021; 38:609–15

15. Michard F, Futier E: Predicting intraoperative hypotension: from hope to hype and back to reality. Br J Anaesth 2023; 131:199–201

16. Michard F, Biais M, Futier E, Romagnoli S: Mirror, mirror on the wall, who is going to become hypotensive? Eur J Anaesthesiol 2023; 40:72–4

17. Enevoldsen J, Hovgaard HL, Vistisen ST: Selection Bias in the Hypotension Prediction Index: Reply. Anesthesiology 2023; 138:450–2

18. Yang SM, Cho HY, Lee HC, Kim HS: Performance of the Hypotension Prediction Index in living donor liver transplant recipients. Minerva Anestesiol 2023; 89:387–95

19. Ven WH van der, Terwindt LE, Risvanoglu N, Ie ELK, Wijnberge M, Veelo DP, Geerts BF, Vlaar APJ, Ster BJP van der: Performance of a machine-learning algorithm to predict hypotension in mechanically ventilated patients with COVID-19 admitted to the intensive care unit: a cohort study. J Clin Monit Comput 2021 doi:10.1007/S10877-021-00778-X

20. Kouz K, Monge García MI, Cerutti E, Lisanti I, Draisci G, Frassanito L, Sander M, Ali Akbari A, Frey UH, Grundmann CD, Davies SJ, Donati A, Ripolles-Melchor J, García-López D, Vojnar B, Gayat É, Noll E, Bramlage P, Saugel B: Intraoperative hypotension when using hypotension prediction index software during major noncardiac surgery: a European multicentre prospective observational registry (EU HYPROTECT). BJA Open 2023; 6:100140

21. Rellum SR, Schuurmans J, Schenk J, Ster BJP Van Der, Ven WH Van Der, Geerts BF, Hollmann MW, Cherpanath TGV, Lagrand WK, Wynandts P, Paulus F, Driessen AHG, Terwindt LE, Eberl S, Hermanns H, Veelo DP, Vlaar APJ: Effect of the machine learning-derived Hypotension Prediction Index (HPI) combined with diagnostic guidance versus standard care on depth and duration of

intraoperative and postoperative hypotension in elective cardiac surgery patients: HYPE-2 - study protocol of a randomised clinical trial. BMJ Open 2023; 13

22. Meng L, Yu W, Wang T, Zhang L, Heerdt PM, Gelb AW: Blood Pressure Targets in Perioperative Care. Hypertension 2018; 72:806–17

23. Schnetz MP, Danks DJ, Mahajan A: Preoperative Identification of Patient-Dependent Blood Pressure Targets Associated With Low Risk of Intraoperative Hypotension During Noncardiac Surgery. Anesth Analg 2023; 136:194

24. Li W, Hu Z, Yuan Y, Liu J, Li K: Effect of hypotension prediction index in the prevention of intraoperative hypotension during noncardiac surgery: A systematic review. J Clin Anesth 2022; 83:110981

25. 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

26. 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 - Journal of the American Medical Association 2020; 323:1052–60

27. 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

28. Tsoumpa M, Kyttari 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

29. Murabito P, Astuto M, Sanfilippo F, Via L La, Vasile F, Basile F, Cappellani A, Longhitano L, Distefano A, Li Volti G: Proactive Management of Intraoperative Hypotension Reduces Biomarkers of Organ Injury and Oxidative Stress during Elective Non-Cardiac Surgery: A Pilot Randomized Controlled Trial. J Clin Med 2022; 11

Figure legends

Figure 1

Calculation of MAP based predictors of hypotension, including concurrent MAP, delta MAP and linearly extrapolated MAP, illustrated as based on a representative time frame of a patient (*MAP = mean arterial pressure, Δ = delta*)

Figure 2

Scatter plot of Hypotension Prediction Index and concurrent MAP values of all 100 patients; red points indicate presence of hypotension, blue points its absence within 1-5 minutes. Values for the default Index and proposed MAP thresholds of 72 and 73 mmHg (dashed lines) (*MAP = mean arterial pressure*)

Figure 3

Agreement of the Hypotension Prediction Index default alarm (>85, blue) and the proposed alarm based on a MAP threshold (<73 mmHg, red) during two representative time frames of patients experiencing hypotension (threshold dashed) (*MAP = mean arterial pressure*)

Figure 4

Performance curves for hypotension prediction with Hypotension Prediction Index (blue), concurrent MAP (red dashed), linearly extrapolated MAP (green) and ΔMAP (orange) within a time window of 1-5 minutes (*AUC = area under the curve, MAP = mean arterial pressure, Δ = delta)*

4a. Receiver operating characteristic curve

4b. Precision recall curve

Table 1: Baseline characteristics of the study population

Variable	**n = 100**
Male gender	48 (48%)
Age (years)	68 ± 11
Body mass index (kg/m^2)	25.8 ± 4.3
ASA classification	
1	0 (0%)
2	49 (49%)
3	42 (42%)
4	9 (9%)
Medical history	
Hypertension	59 (59%)
Heart failure	8 (8%)
Vascular disease	30 (30%)
Diabetes mellitus	24 (24%)
COPD	9 (9%)
Kidney failure	8 (8%)
MAP in outpatient clinic (mmHg)	98 ± 15
Type of surgery	
Gastrointestinal	51 (51%)
Urological	21 (21%)
Vascular	13 (13%)
Gynecological	12 (12%)
Other	3 (3%)
Surgical approach	
Open	63 (63%)
Laparoscopic/robot	30 (30%)
Endovascular	4 (4%)
Conversion	1 (1%)
Other	2 (2%)

ASA = American Society of Anesthesiologists (Schaumburg, Illinois) Physical Status, COPD = Chronic obstructive pulmonary disease, MAP = mean arterial pressure

Table 2: Hypotension metrics

Variable	n = 100
Monitoring time (min)	228.6 ± 107.6
Incidence (%)	62
Total duration (min)	1.8 [0.0, 7.0]
Proportion of surgery time (%)	1.0 [0.0, 3.4]
Number of events	1 [0, 3]
Length of events (min)	2.0 [1.3, 3.0]
Area under the threshold (mmHg*min)	5 [0, 35]
Time weighted average (mmHg)	0.03 [0.0, 0.2]

Table 3: Agreement between default Hypotension Prediction Index (>85) and concurrent MAP based alarms

MAP threshold [mmHg]	Proportionate agreement	Cohen's kappa coefficient
< 80	0.79	0.45
< 75	0.94	0.78
< 73	0.97	0.85
< 70	0.94	0.71
< 65	0.90	0.37

MAP = mean arterial pressure

Table 4: Performance of different Hypotension Prediction Index and MAP based thresholds in predicting hypotension

	Prediction window [min]	Receiver operating characteristic AUC (95% CI)	Precision recall AUC (95% CI)	Threshold	Sensitivity (95% CI) [%]	Specificity (95% CI) [%]	Positive predictive value (95% CI) [%]	Negative predictive value (95% CI) [%]
Hypotension Prediction Index	1-15	0.78 (0.77-0.78)	0.48 (0.47-0.49)	90*	46.8 (45.7-47.8)	92.7 (92.5-92.9)	46.5 (45.6-47.4)	92.8 (92.7-93.0)
				85	49.9 (48.9-51.0)	90.9 (90.7-91.2)	42.6 (41.8-43.4)	93.1 (93.0-93.2)
				46**	70.1 (69.1-71.1)	70.5 (70.2-70.9)	24.2 (23.9-24.6)	94.6 (94.4-94.8)
	1-10	0.81 (0.81-0.82)	0.50 (0.49-0.51)	95*	49.6 (48.4-50.8)	95.0 (94.8-95.1)	49.9 (48.8-50.9)	94.9 (94.8-95.0)
				85	58.0 (56.7-59.2)	90.5 (90.3-90.8)	38.2 (37.5-39.0)	95.5 (95.4-95.6)
				51**	73.3 (72.3-74.3)	73.3 (73.0-73.7)	21.7 (21.4-22.1)	96.5 (96.3-96.6)
	1-5	0.89 (0.88-0.89)	0.56 (0.55-0.57)	97*	51.8 (50.3-53.3)	97.8 (97.7-97.9)	60.5 (59.1-62.0)	96.9 (96.8-97.0)
				85	72.4 (71.1-73.8)	89.9 (89.7-90.1)	31.9 (31.3-32.6)	98.0 (97.9-98.1)
				63**	80.1 (78.9-81.4)	80.2 (79.9-80.5)	20.9 (20.5-21.2)	98.4 (98.3-98.5)
MAP [mmHg]	1-15	0.77 (0.77-0.78)	0.47 (0.46-0.48)	71*	45.8 (44.7-46.9)	93.1 (92.9-93.3)	47.2 (46.2-48.2)	92.7 (92.6-92.9)

				72	49.1 (48.1-50.2)	91.3 (91.1-91.5)	43.1 (42.3-43.9)	93.0 (92.9-93.2)
				79**	70.5 (69.5-71.5)	69.9 (69.5-70.3)	24.0 (23.6-24.3)	94.6 (94.4-94.8)
	1-10	0.81 (0.80-0.82)	0.49 (0.48-0.49)	70*	50.7 (49.5-51.9)	94.3 (94.1-94.5)	47.4 (46.4-48.4)	95.0 (94.9-95.1)
				72	57.5 (56.3-58.7)	90.9 (90.7-91.2)	39.0 (38.2-39.8)	95.5 (95.4-95.6)
				78**	74.2 (73.1-75.3)	72.9 (72.5-73.2)	21.6 (21.3-22.0)	96.5 (96.4-96.7)
	1-5	0.88 (0.88-0.89)	0.55 (0.54-0.56)	67*	55.1 (53.5-56.6)	97.3 (97.2-97.4)	57.1 (55.8-58.4)	97.1 (97.0-97.2)
				72	72.6 (71.3-73.9)	90.3 (90.1-90.5)	32.9 (32.2-33.6)	98.1 (98.0-98.2)
				75**	79.8 (78.5-81.1)	82.3 (82.0-82.6)	22.8 (22.4-23.2)	98.4 (98.3-98.5)
Linearly extrapolated MAP [mmHg]	1-5	0.85 (0.84-0.85)	0.32 (0.32-0.33)	57*	37.4 (35.9-38.9)	96.2 (96.0-96.3)	38.6 (37.3-39.9)	96.0 (95.9-96.1)
				73**	78.6 (77.4-79.9)	78.3 (78.0-78.6)	18.9 (18.5-19.2)	98.3 (98.2-98.4)
ΔMAP [mmHg/s]	1-5	0.66 (0.65-0.67)	0.13 (0.13-0.14)	-16*	17.6 (16.4-18.8)	95.1 (94.9-95.3)	18.7 (17.6-19.9)	94.7 (94.6-94.8)
				-3**	59.5 (58.0-61.0)	63.8 (63.4-64.1)	9.6 (9.3-9.8)	96.1 (95.9-96.2)

*) Value defined by minimizing difference between sensitivity and positive predictive value

***) value defined by minimizing difference between sensitivity and specificity*

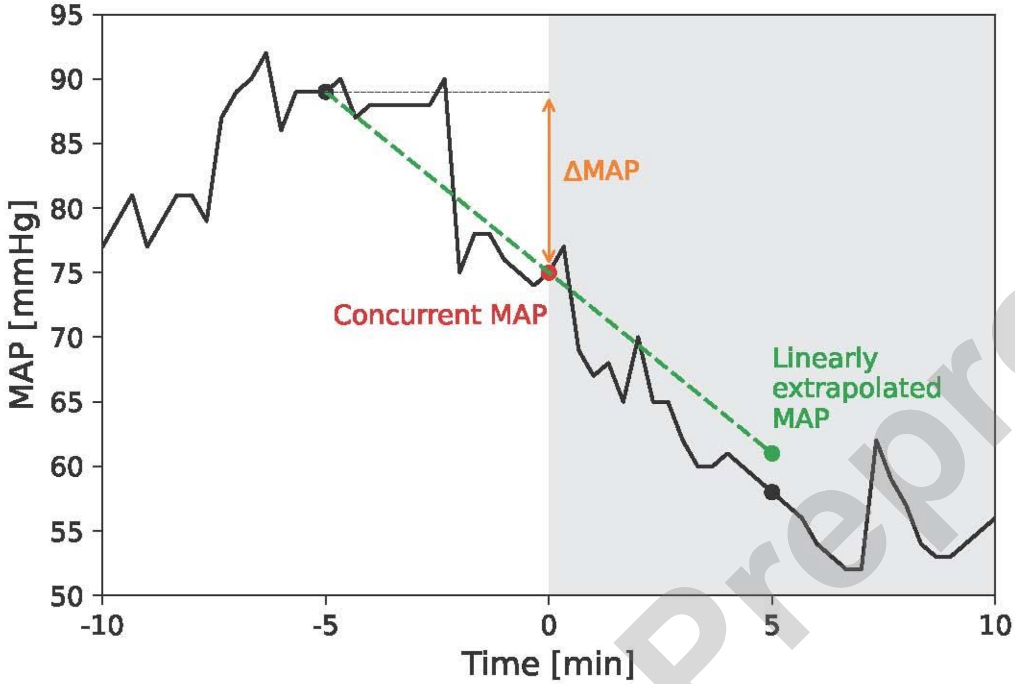
AUC = area under the curve, CI = confidence interval, MAP = mean arterial pressure, Δ = delta

Table 5: Positive predictive values from forward analysis for default Hypotension Prediction Index (>85) and MAP (<72) alarm

	All alarms		Exclude alarms with interventions	
Prediction window	PPV Hypotension Prediction Index (95% CI) [%]	PPV MAP (95% CI) [%]	PPV Hypotension Prediction Index (95% CI) [%]	PPV MAP (95% CI) [%]
1-5 min	24.1 (20.7-27.8)	27.8 (24.1-31.8)	34.6 (29.9-39.5)	43.3 (38.0-48.6)
1-10 min	30.8 (27.0-34.7)	34.9 (30.9-39.0)	47.7 (42.5-52.9)	59.4 (53.8-64.8)
1-15 min	36.4 (32.4-40.5)	39.6 (35.5-43.8)	56.8 (51.6-62.0)	66.6 (61.2-71.6)

MAP = mean arterial pressure, PPV = positive predictive value

Figure 1



MAP [mmHg]
95
90
85
80
75
70
65
60
55
50
-10
-5
0
5
10
Time [min]
ΔMAP
Concurrent MAP
Linearly extrapolated MAP

Figure 2

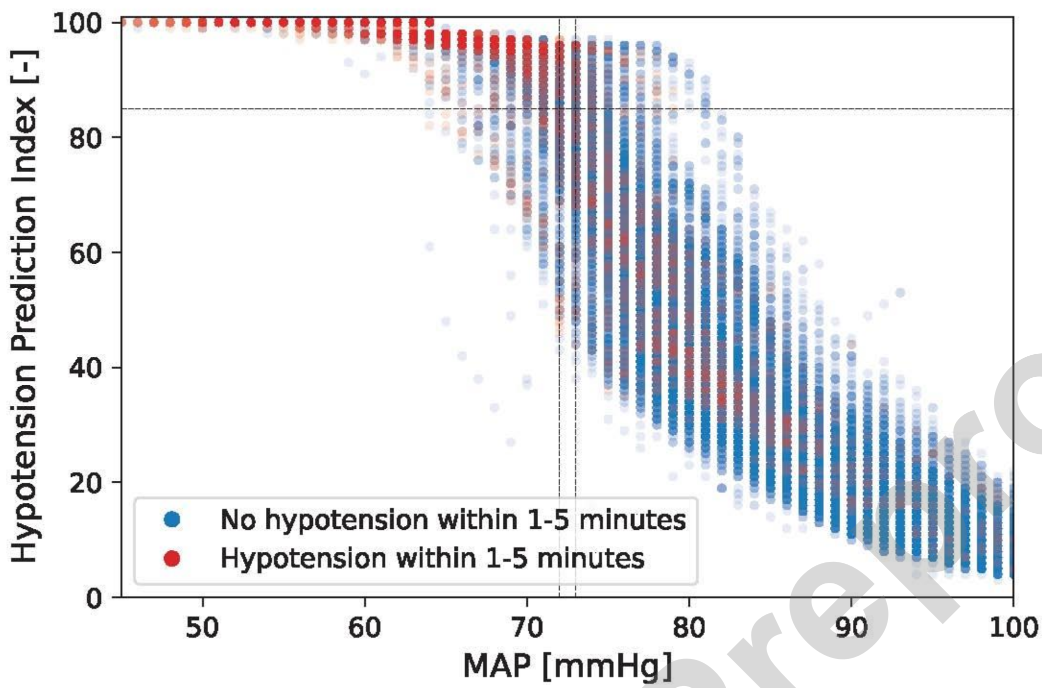

Hypotension Prediction Index [-]
100
80
60
40
20
0
No hypotension within 1-5 minutes
Hypotension within 1-5 minutes
50
60
70
80
90
100
MAP [mmHg]

Figure 3

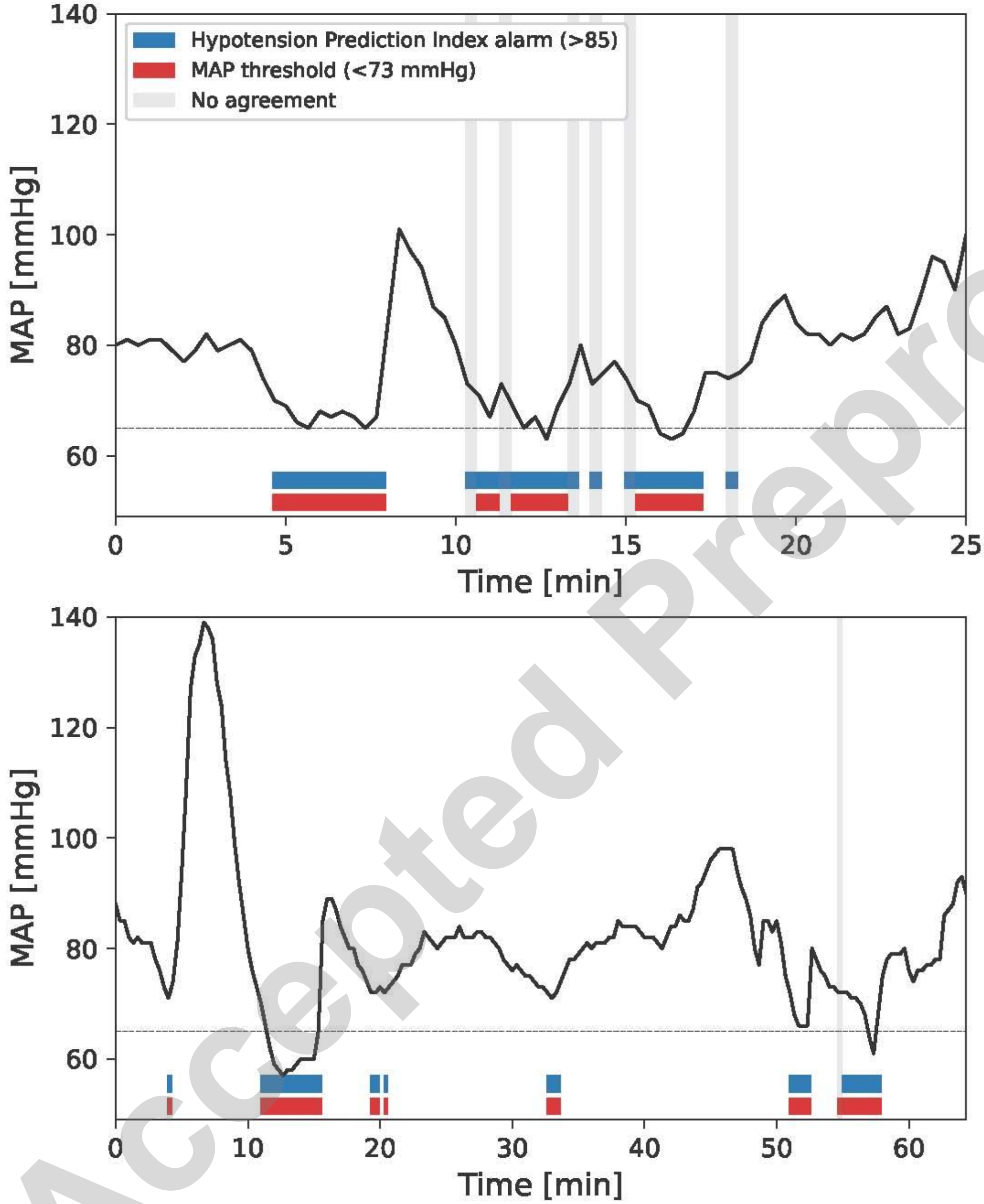
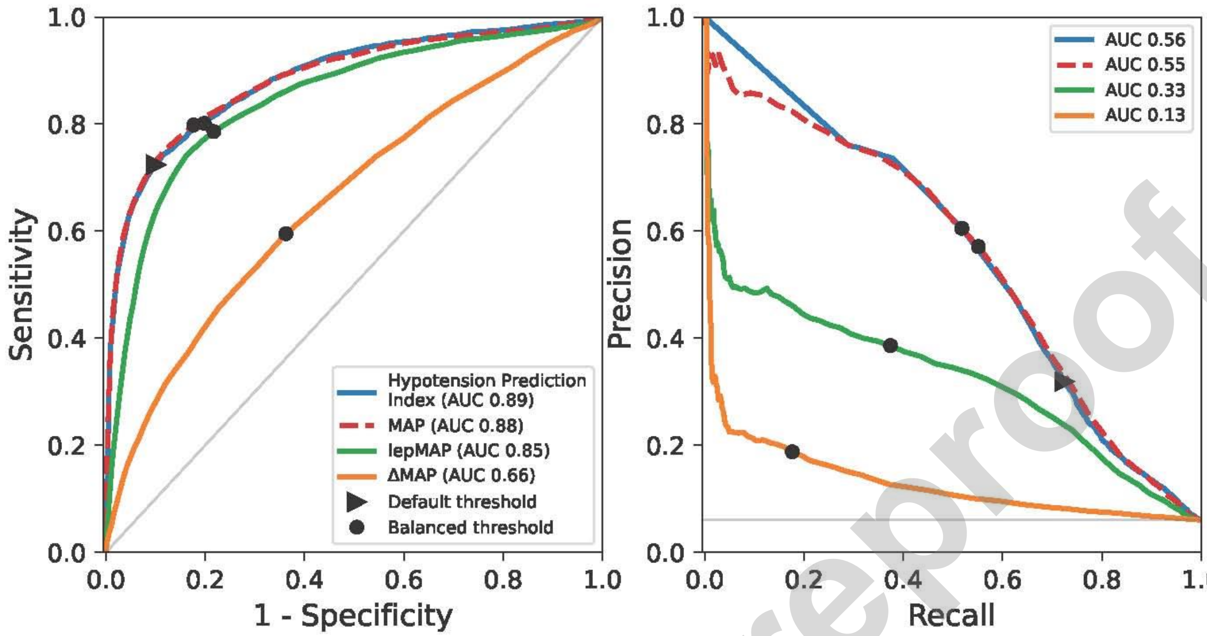

Hypotension Prediction Index alarm (>85)
MAP threshold (<73 mmHg)
No agreement
140
120
100
80
60
MAP [mmHg]
0
5
10
15
20
25
Time [min]
140
120
100
80
60
MAP [mmHg]
0
10
20
30
40
50
60
Time [min]

Figure 4



1.0
0.8
0.6
0.4
0.2
0.0
Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
1 - Specificity
Hypotension Prediction Index (AUC 0.89)
MAP (AUC 0.88)
lepMAP (AUC 0.85)
ΔMAP (AUC 0.66)
Default threshold
Balanced threshold
Precision
Recall
AUC 0.56
AUC 0.55
AUC 0.33
AUC 0.13