Forecasting Acute Hypotensive Episodes in Intensive Care Patients Based on a Peripheral Arterial Blood Pressure Waveform

X Chen, D Xu, G Zhang, R Mukkamala

Michigan State University, East Lansing, MI, USA

Abstract

We entered the 10th Annual PhysioNet/Computers in Cardiology Challenge to predict which intensive care patients would experience an acute hypotensive episode (AHE) using physiologic data prior to the occurrence of the AHE. An AHE was defined through mean arterial blood pressure (ABP). We took a pragmatic approach to the Challenge. We explored six basic indices derived from ABP data near the forecast window including mean ABP and diastolic ABP. We evaluated the predictive ability of each index on the provided training dataset and employed basic classification on the testing dataset. All indices performed well on the training dataset and achieved a perfect score for Event 1 of the Challenge and scores from 32/40 to 37/40 for Event 2. However, our best official score was 36/40 for Event 2. These results stress the importance of continuous ABP monitoring in intensive care patients and indicate that sophisticated data analysis was not necessary to win the Challenge.

1. Introduction

The occurrence of acute hypotensive episodes (AHEs) in intensive care patients can significantly increase their mortality rate [1]. It is therefore important to be able to predict AHEs before they occur so that timely interventions can be administered to improve care and increase survival opportunities for these patients.

We entered the 10th Annual PhysioNet/Computers in Cardiology Challenge on predicting AHEs in the intensive care unit. The Challenge was to forecast which intensive care patients would experience an AHE within a predefined forecast window of one hour duration using physiologic data prior to the window (see Figure 1). An AHE was defined as any period of ≥ 30 minutes during which $\geq 90\%$ of the one-minute averages of the arterial blood pressure (ABP) waveform were ≤ 60 mmHg.

The Challenge is explained in detail elsewhere [1]. Briefly, physiologic signals including ECG and ABP waveforms, vital signs such as heart rate (HR) and systolic and diastolic ABP, laboratory test results, and other clinical data from a set of intensive care patients were provided. The patients were classified into two groups, H (patients with an AHE in the forecast window) and C (patients without an AHE in the forecast window). The H group was further classified into two subgroups, H1 (patients who received a pressor medication) and H2 (patients who did not receive a pressor medication). The C group was likewise classified into two subgroups, C1 (patients without an AHE outside the forecast window) and C2 (patients with at least one AHE outside the forecast window). Physiologic data from 15 patients in each of the four subgroups with correct classifications were provided as the training dataset. Physiologic data from 50 patients with partial classification information were also provided as the testing dataset. Event 1 of the Challenge was to classify five patients in the H1 subgroup amongst a total of ten patients in the H1 or C1 subgroup. Event 2 was to classify ten to 16 patients in the H group amongst 40 patients in the H or C group. Up to four entries were allowed for each event.
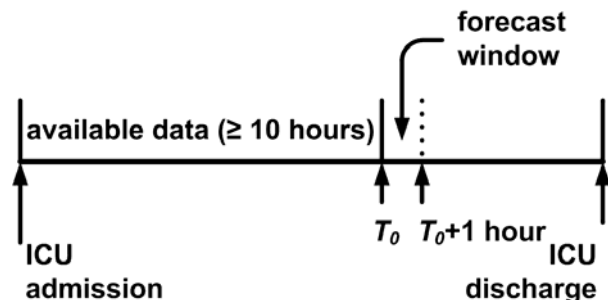

Figure 1. Segmentation of datasets. Modified from [1].

We took a pragmatic approach to the Challenge. That is, since an AHE was defined through mean ABP (MAP), we hypothesized that only basic ABP information would be needed to predict an AHE. We further hypothesized that the closer in time the ABP data are to the forecast window, the more predicative they would be of an AHE. Based on these hypotheses, we explored several different indices derived from the ABP waveform. We evaluated

these indices on the training dataset based on receiver operating characteristic (ROC) curves. We then applied straightforward classification schemes for each of the indices on the testing dataset. As required by the Challenge, our entire analysis was automated. Our results showed that all of the indices performed well on the training dataset and achieved a perfect score for Event 1 and very good scores for Event 2.

2.   Methods

2.1.   Indices

We explored six simple indices derived from the ABP vital signs or directly from the ABP waveform. All analyzed ABP vital signs were pre-processed by linear interpolation of any missing data and removal of any spurious spikes.

Index I is the 5-min average of the MAP vital signs (ABPMean) before the forecast window (see Figure 2a).

Index II is the 5-min average of the ABP waveform before the forecast window (see Figure 2b). This index was largely the same as Index I but did yield a better score for Event 2 (see Results section).

Index III is the optimal exponentially weighted average of the 10-hr ABPMean before the forecast window (see Figure 2c). The exponential weighting was optimal in the sense that it maximized the classification accuracy, as quantified via ROC curves (see below), in the training dataset. The optimal time constant of the exponential weighting turned out to be 1.2 hrs.

Index IV is the predicted ABPMean at the midpoint of the forecast window via linear regression of the 1-hr ABPMean before the forecast window (see Figure 2d).

Index V is the 5-min average of the diastolic ABP vital signs (ABPDias) before the forecast window (see Figure 2e). Diastolic ABP is known to be related to MAP but could offer complementary prognostic value.

Index VI is a combined index of the 5-min averages of the ABP waveform (Index II) and ABPDias (Index V) before the forecast window (see Figure 2f). This index specifically represents a voting strategy in which an AHE is concluded, if predicted by both Index II and Index V (see below).
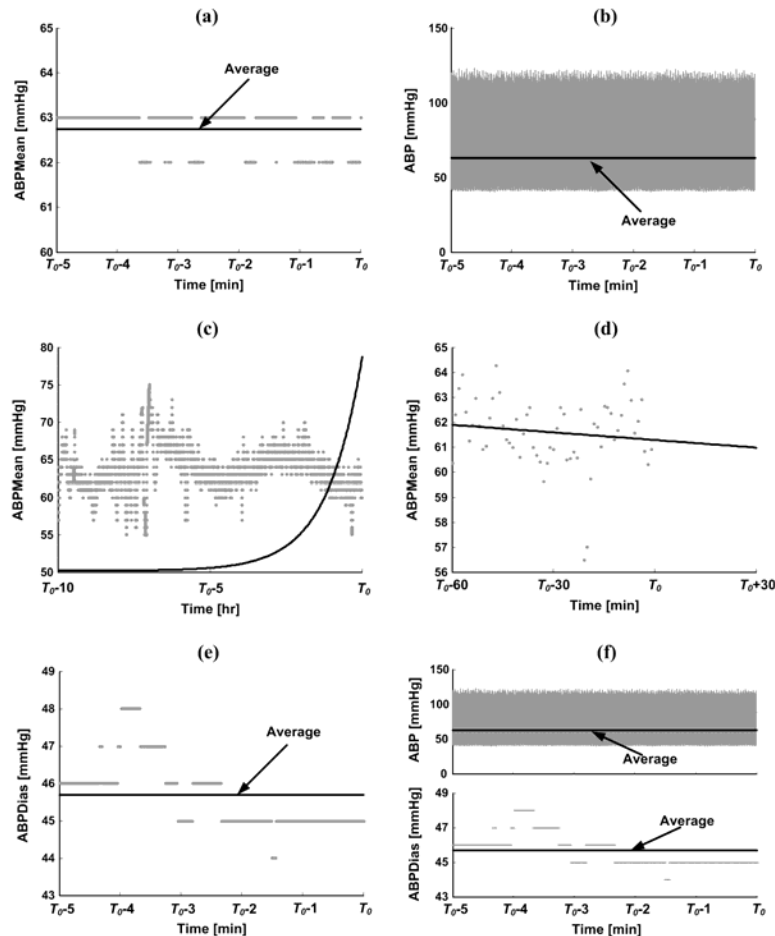

Figure 2: Indices derived from the ABP waveform.

2.2. Evaluation on training dataset

We evaluated the efficacy of each index on the training dataset. We specifically tested the ability of each index to discriminate between patients in the H1 and C1 subgroups and patients in the H and C groups so as to emulate Event 1 and Event 2 of the Challenge. We quantified the discriminatory ability of each index in terms of the ROC area under the curve (AUC). The ROC AUC indicates the probability that two samples, one drawn from each class, will be accurately classified [2].

2.3. Classification on testing dataset

We used straightforward classification schemes for each index on the testing dataset. For Event 1, we classified the five patients with the lowest index levels in the H1 subgroup (see Figure 3a). For Event 2, we classified ten to 16 patients with the lowest index levels in the H group, wherein the exact number of patients in this group was determined by maximizing the difference between the highest index level in the H group and the lowest index level in the C group (see Figure 3b).
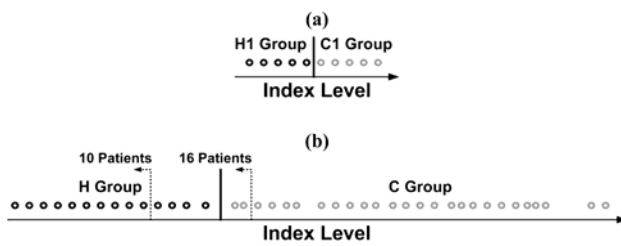

Figure 3. Classification on testing dataset.

3. Results

Table 1 summarizes the results of the six ABP derived indices on the training and testing datasets. As can be seen, all of the indices performed well on the training dataset and achieved a perfect score for Event 1 and good, but somewhat different, scores for Event 2.

Index I and Index II represented our first two entries in the Challenge. These two indices were naturally similar and, in fact, ranked the patients in the same order for Event 1 and Event 2. However, for Event 2, Index I classified 13 patients in the H group, whereas Index II classified 16 patients in this group. These three additional patients ascribed to the H group by Index II turned out to be correct and improved our score to 36/40.

Index III represented our third entry in the Challenge. This index performed the best on the training dataset, as it was the only index that was optimised on this dataset. However, to our dismay, this optimal index yielded a score of only 32/40 for Event 2.

Index IV produced the same entries for Event 1 and Event 2 as Index II. This index was therefore not submitted as an entry.

Index V performed the best on the training dataset amongst the non-optimised indices. Further, it achieved a score of 37/40 for Event 2. We did not submit this entry, because we ascertained through simple logic that it could not yield a score > 37/40 for Event 2, which was the winning score.

Index VI did provide a chance to win Event 2 and therefore represented our fourth and final allowed entry. However, this index only achieved a score of 36/40 for Event 2.

4. Discussion

Previous studies have shown that MAP is a good predictor of hypotension (e.g., [3]). Further, for the purposes of the Challenge, an AHE was defined based on MAP. It is therefore not too surprising that indices derived from MAP were accurate, but not perfect, in predicting AHEs in both the training and testing datasets. Indices derived from diastolic ABP showed similar prognostic capabilities perhaps due to its tight relationship with MAP. Based on the overall results of the Challenge from all entrants [1], it is unclear whether perfect prediction was achievable.

From our first and second entries, we determined that the patients numbered 214, 217, and 224 in Event 2 belonged to the H group. However, the level of the investigated indices for these three patients were usually the highest amongst the 16 lowest index levels (see Figure 4), which made it difficult to detect false positive patients and further improve our score for Event 2.

Table 1. Results of the six ABP derived indices on the training and testing datasets.

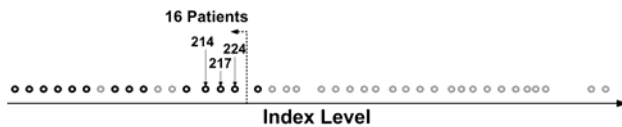|	Training Dataset ROC AUC		Test Dataset Scores		Entry
	H1 vs C1	H vs C	Event 1	Event 2	
Index I	0.81	0.76	10/10	33/40	first entry
Index II	0.85	0.75	10/10	36/40	second entry
Index III	0.93	0.82	10/10	32/40	third entry
Index IV	0.76	0.72	10/10	36/40	= second entry
Index V	0.89	0.79	10/10	37/40	not submitted
Index VI	0.82	0.75	10/10	36/40	fourth entry

Figure 4. Difficulty in improving the score for Event 2.

We did also investigate other indices derived from ABP (systolic ABP, pulse pressure, cardiac output [CO] by ABP waveform analysis, and the ratio of the number of ABPMean $\leq$ 60 mmHg over the total number of ABPMean), ECG (HR variability [HRV] spectral powers and premature beats), and both ABP and ECG (ratio of MAP to HR and baroreflex sensitivity). Indeed, some of these indices (e.g., HRV and CO) have been shown to be good predictors of hypotension [3,4]. However, these indices showed less or even little prognostic value on the training dataset. Nevertheless, some of the indices here may offer value in other clinical scenarios such as predicting hypotension caused by certain mechanisms (e.g., hemorrhage) or as defined by a reduction in ABP.

In conclusion, both MAP and diastolic ABP were excellent predictors of AHEs in the context of the Challenge. Our best scores from the submitted entries were 10/10 for Event 1 and 36/40 for Event 2. Our results emphasize the importance of continuous ABP monitoring in intensive care patients and demonstrate that sophisticated data analysis was not necessary to win the Challenge.

Acknowledgements

References

[1] Moody GB, Lehman LH. Predicting acute hypotensive episodes: the 10th annual PhysioNet/Computers in Cardiology Challenge. Computers in Cardiology 2009;36.
[2] http://www.anaesthetist.com/mnm/stats/roc/Findex.htm
[3] Reich DL, Hossain S, Krol M, Baez B, Patel P, Bernstein A, Bodian CA. Predictors of hypotension after induction of general anesthesia. Anesth Analg 2005;101:622-28.
[4] Hanss R, Bein B, Ledowski T, Lehmkuhl M, Ohnesorge H, Scherkl W, Steinfath M, Scholz J, Tonner PH. Heart rate variability predicts severe hypotension after spinal anesthesia for elective cesarean delivery. Anesthesiology 2005;102:1086-93.
[5] Barcroft H, Edholm OG, McMichael J, Sharpey-Schafer EP. Posthaemorrhagic fainting. Lancet 1944;i:489-91.

Address for correspondence

Xiaoxiao Chen
2120 Engineering, East Lansing, MI 48824, USA
chenxia7@msu.edu