



OPEN

Non-invasive biomarkers for detecting progression toward hypovolemic cardiovascular instability in a lower body negative pressure model

Ethan K. Murphy^{1✉}, Spencer R. Bertsch¹, Samuel B. Klein^{2,3}, Navid Rashedi¹, Yifei Sun¹, Michael J. Joyner⁴, Timothy B. Curry⁴, Christopher P. Johnson⁴, Riley J. Regimbal⁴, Chad C. Wiggins⁴, Jonathon W. Senefeld⁴, John R. A. Shepherd⁴, Jonathan Thomas Elliott^{1,2,3}, Ryan J. Halter^{1,2}, Vikrant S. Vaze¹ & Norman A. Paradis^{2,3}

Occult hemorrhages after trauma can be present insidiously, and if not detected early enough can result in patient death. This study evaluated a hemorrhage model on 18 human subjects, comparing the performance of traditional vital signs to multiple off-the-shelf non-invasive biomarkers. A validated lower body negative pressure (LBNP) model was used to induce progression towards hypovolemic cardiovascular instability. Traditional vital signs included mean arterial pressure (MAP), electrocardiography (ECG), plethysmography (Pleth), and the test systems utilized electrical impedance via commercial electrical impedance tomography (EIT) and multifrequency electrical impedance spectroscopy (EIS) devices. Absolute and relative metrics were used to evaluate the performance in addition to machine learning-based modeling. Relative EIT-based metrics measured on the thorax outperformed vital sign metrics (MAP, ECG, and Pleth) achieving an area-under-the-curve (AUC) of 0.99 (CI 0.95–1.00, 100% sensitivity, 87.5% specificity) at the smallest LBNP change (0–15 mmHg). The best vital sign metric (MAP) at this LBNP change yielded an AUC of 0.6 (CI 0.38–0.79, 100% sensitivity, 25% specificity). Out-of-sample predictive performance from machine learning models were strong, especially when combining signals from multiple technologies simultaneously. EIT, alone or in machine learning-based combination, appears promising as a technology for early detection of progression toward hemodynamic instability.

Hemorrhage progressing to hemodynamic instability is a leading cause of death in trauma patients^{1,2}. Although many cases of hemorrhage are self-evident and recent improvements such as balanced resuscitation and advanced trauma life support (ATLS) protocols^{2,3} have improved outcomes, there remain some patients that present insidiously after trauma—occult subclinical hemorrhages. In these cases, there may be significant progression toward hemodynamic instability without any discernible changes in traditional clinical parameters⁴, including blood pressure (BP), heart rate (HR), and plethysmography (Pleth), until sudden hemodynamic deterioration requires rescue therapies. Delays in detection are associated with worse patient outcomes^{5–8}.

Early detection of ongoing subclinical hemorrhage before onset of frank shock is a major unmet need. In particular, a noninvasive technology that can be applied as a wearable and act as a monitor and accurate early alarm could prevent significant morbidity and mortality. Currently, none of the existing non-invasive technologies has adequate performance with respect to underlying false-positive and false-negative rates to function in such a role. We have previously hypothesized that a multiplexed combination of non-invasive technologies that incorporates anatomic and temporal patterns via machine learning may outperform existing solutions⁹ and reported that such an approach had superior diagnostic performance¹⁰ in a porcine model of occult hemorrhage¹¹. In particular, we

¹Thayer School of Engineering, Dartmouth College, Hanover, NH 03755, USA. ²Geisel School of Medicine, Dartmouth College, Hanover, NH 03755, USA. ³Dartmouth-Hitchcock Medical Center, Lebanon, NH 03756, USA. ⁴Department of Anesthesiology and Perioperative Medicine, Mayo Clinic, Rochester, MN 55902, USA. ✉email: ethan.k.murphy@dartmouth.edu

observed that direct change-from-baseline impedance measurements¹² and a multiplexed model built of absolute and impedance-change data appeared to have exceptional performances⁹.

In light of our previous results in a porcine model, we undertook evaluation of this same approach in a human model. An established method to safely mimic progression toward hemodynamic instability in humans is through use of a lower body negative pressure (LBNP) model^{4,13–15}. The application of LBNP results in “a central volume shift to the lower body, which creates central hemodynamic conditions that may mimic those obtained during actual blood loss”⁴. LBNP has been used to evaluate various diagnostic signals that may assist in earlier detection of hemorrhage^{8,13,16–19}. Absolute and relative vital sign and electrical impedance-based metrics are evaluated via direct analysis and via a machine-learning (ML) based approach.

Results

Subjects

Eighteen healthy male and female volunteers (11 males, 7 females) with an average (\pm standard deviation) age of 32.7 ± 8.2 (23–51), height of 172 ± 8.6 cm (158–193 cm), weight of 77.1 ± 10.8 kg (63–101.8 kg), and body mass index of 25.7 ± 2.6 kg/m² (21.8–30 kg/m²) were recruited to participate in this LBNP study (January to March 2020). Figure 1 provides an overview of the measurements recorded on one subject. The study was originally powered for 25 subjects, but was stopped early because of COVID. Of the 18 subjects recruited for the study, there were impedance instrumentation issues on 2 subjects. Thus, we restricted the analysis to the 16 subjects where complete data was available. LBNP levels were increased in a step-wise manner (15 mmHg per step, held stable for 5 min per step) until a clinically evident hypotensive state developed. The average LBNP reached over all subjects was 80.0 ± 11.7 mmHg (60–100 mmHg), with 2 subjects stopping at 60 mmHg and one subject continuing to 100 mmHg.

The absolute and relative data are displayed as box-and-whisker plots (25–75% with $N = 16$ samples for each) in Figs. 2 and 3. Note that Pleth_A is a metric calculated from the Pleth waveform (see “Methods” section) and EIT and EIS metrics are average, filtered impedance data corresponding to the noted sites (thx = thorax, ab = abdomen). Minute-based averages of all the data is additionally shown in Supplementary Appendix A.1. For the absolute metrics, a red square indicates significant differences between the current LBNP level and baseline ($p < 0.05$), and a green star indicates significant differences between consecutive LBNP levels ($p < 0.05$). The relative metrics in Fig. 3 show the change in metric values over the 5-min baseline period (0 mmHg) and across consecutive LBNP levels (0-to-15, 15-to-30, 30-to-45, and 45-to-60 mmHg). A description of the baseline-change calculation is provided in Supplementary Appendix A.2. Green stars in these plots indicate a significant difference between the metric changes across consecutive LBNP levels and baseline variability ($p < 0.05$). Bar graphs of area-under-the-curves (AUCs) are presented (Fig. 4A,B) for the absolute and relative metrics with overlaid 95% confidence intervals (CI), where AUCs are calculated from the receiver operating characteristic (ROC) curves.

Absolute metric analysis

The absolute metrics and associated AUCs are presented in Figs. 2 and 4A. HR and Pleth_A trends appear to track with increasing LBNP, while only subtle trends appear in the remaining metrics (Fig. 2). However, there is significant overlap in the distributions between LBNP levels indicating a limited ability to differentiate states.

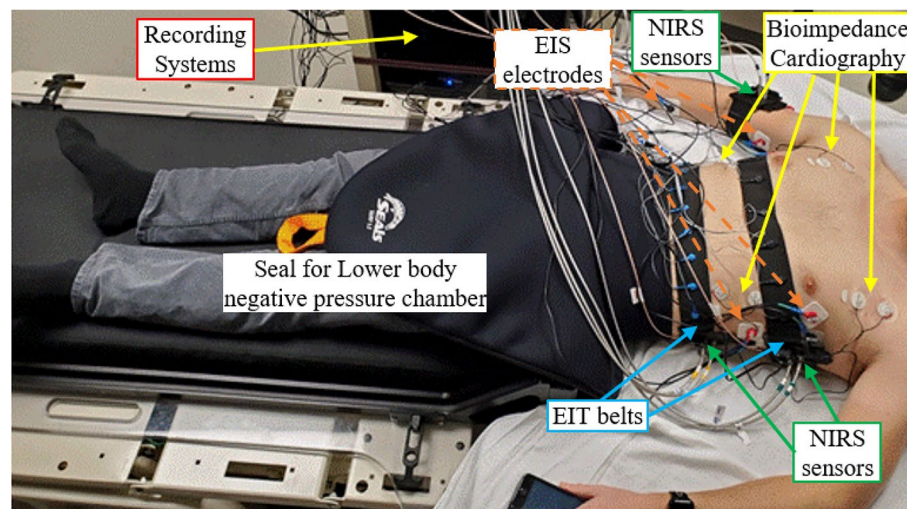


Figure 1. Photo of a subject with sensors attached prior to starting the LBNP experiment. The EIT belts (blue arrows) are positioned around the chest and abdomen, the EIS electrodes (orange) inject current across the thorax, abdomen, and arm, bioimpedance cardiography (BC) uses custom electrodes placed in prescribed locations above and below the heart, and 3 NIRS sensors recorded data in similar locations. Although not present in the photo (due to poor timing), an arterial line was placed on the patient’s left arm to invasively measure MAP.

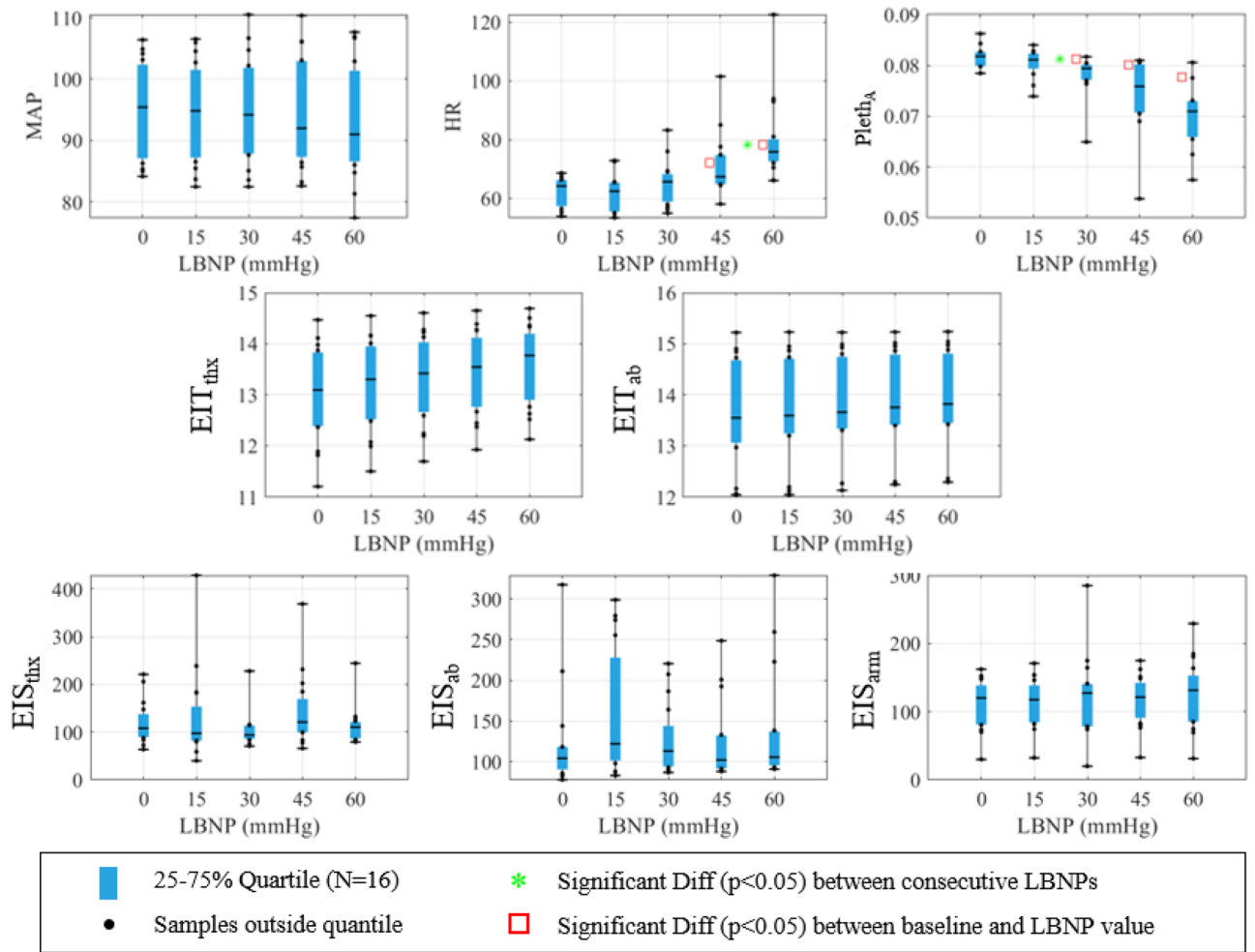


Figure 2. Absolute metric values displayed via quartile box-and-whisker plots. Black dots indicate any subject outside of the quartiles and the median value is shown as a black line segment. Red squares indicate significant differences between the given LBNP level and baseline.

Only HR and Pleth_A had significant differences; (1) with respect to baseline, starting at LBNP levels ranging from 30 mmHg (Pleth_A) to 45 mmHg (HR)—indicated by red squares, and (2) between consecutive LBNP levels from 45-to-60 for HR and 15-to-30 for Pleth_A—indicated by green stars. The best AUCs were 0.66 (EIS_{ab}), 0.83 (Pleth_A), 0.89 (Pleth_A), and 0.97 (Pleth_A) for LBNP levels of 15, 30, 45, and 60 mmHg, respectively, where CIs are only above 0.8 for HR and Pleth_A for an LBNP of 60 mmHg. Excluding HR and Pleth_A, most of the AUCs in Fig. 4A are close to 0.5 (all are under 0.7), indicating a limited ability to differentiate states. The best AUC across impedance metrics is 0.68 (EIT_{thx}) with a CI of (0.50–0.84) at an LBNP level of 60 mmHg. The EIS data shown here correspond to the impedance magnitude and frequency that yields the maximum AUC. Specifically, the best frequencies were all less than or equal to 560 Hz, except for the comparison of 30 mmHg to baseline for the thorax (26.5 kHz). One could confidently say that hypovolemia induced via the LBNP model remains subclinical to vital sign metrics for an LBNP of 15 mmHg and perhaps up to 30 mmHg, although Pleth_A shows significant differences with an AUC of 0.83 (CI 0.59–0.96).

Relative metric analysis

The relative metrics and associated AUCs are presented in Figs. 3 and 4B. Overall, the relative metrics appear very good at differentiating between increasing levels of hypovolemia (i.e., change between consecutive LBNP levels), with HR and EIT/EIS measurements recorded from the thorax and abdomen showing significant differences (green stars in Fig. 3). No metric has perfect separation from the baseline variability, but the EIT thorax data, EIT_{thx}, comes closest, achieving AUCs of 0.99, 0.98, 0.97, and 0.97 for each LBNP level studied with CIs above 0.92 for LBNP changes of 0-to-15 mmHg and 45-to-60 mmHg. The EIT_{thx} metric outperforms all other metrics (Fig. 4B) with a high degree of efficacy in identifying hypovolemia (AUC = 0.97, CI 0.95–1.00) across all levels of LBNP explored here. Some other techniques achieved AUCs > 0.9 for a subset of LBNP ranges: HR for LBNP change of 45-to-60 mmHg (0.93, CI 0.75–1.00), EIS on the thorax for LBNP changes of 0-to-15 (0.93, CI 0.79–0.99) and 45-to-60 mmHg (0.91, CI 0.71–1.00), and EIS on the abdomen for LBNP change of 45-to-60 mmHg (0.93, CI 0.78–1.00). The EIS data presented here correspond to the impedance magnitude and

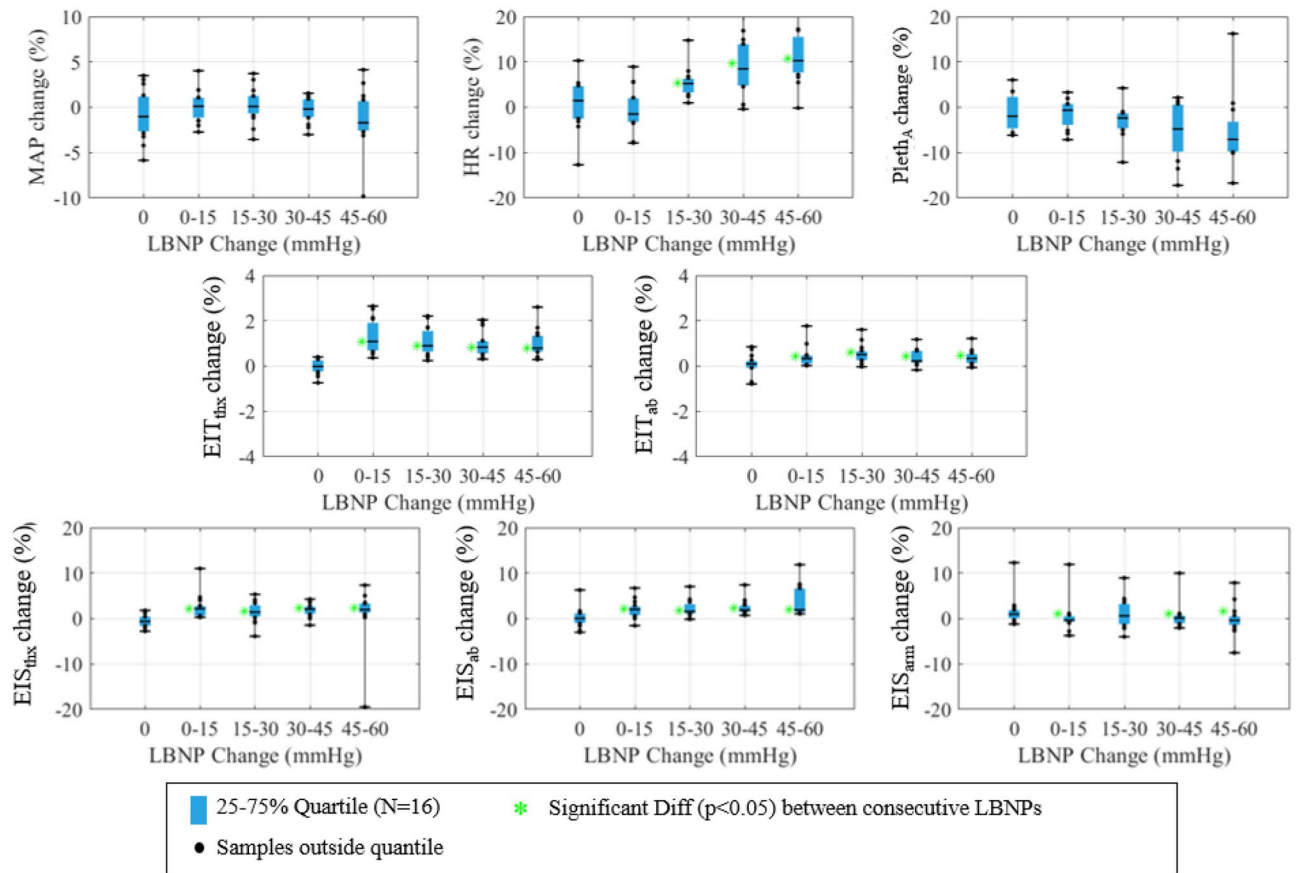


Figure 3. Relative metric value displayed via quartile box-and-whisker plots. Black dots indicate any subject outside of the quartiles and the median value is shown as a black line segment. Green stars indicate significant differences between consecutive LBNP levels.

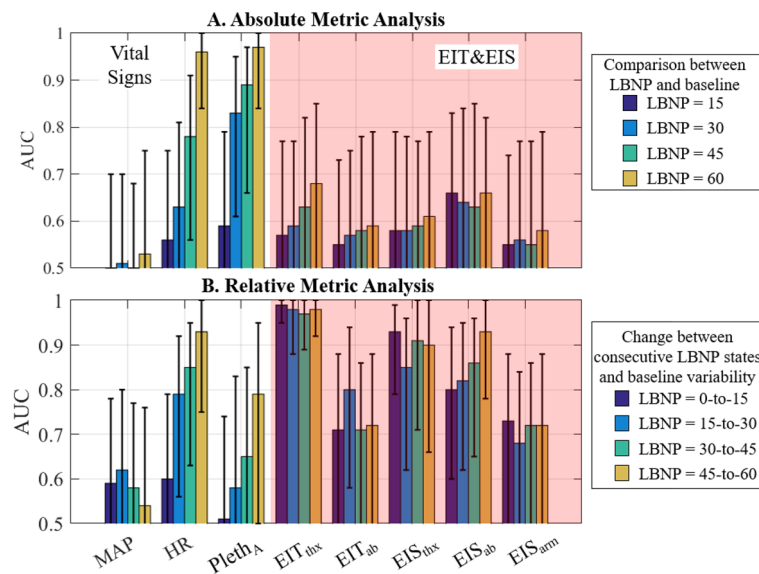


Figure 4. AUCs (bars) and 95% confidence intervals (whiskers) of the (A) absolute and (B) relative analysis for the vital signs and EIT/EIS metrics at each LBNP level (A) or each LBNP change (B). All analyses compare metric values (at the specific LBNP level) or change in metric values (over the specified LBNP change) to baseline or baseline variability, respectively—i.e. these are bleed versus no-bleed analyses performed at different LBNP levels. No comparisons are performed between the different LBNP levels, e.g. comparisons between metric changes from LBNPs of 0-to-15 to those from LBNPs of 15-to-30 are not considered.

frequency that yields the maximum AUC. In general, these optimum frequencies fell between 8 and 27 kHz. Although the impedance changes are very consistent, they are small. For example, the median percent increases in EIT_{thx} are 1.09%, 0.89%, 0.83%, and 0.79% with minimum increases of 0.36%, 0.24%, 0.31%, and 0.29% for LBNP changes from 0-to-15, 15-to-30, 30-to-45, and 45-to-60 mmHg, respectively. ROC curves (Fig. 5A–D) from MAP, HR, the best EIT metric (EIT_{thx}), and best EIS metric (EIS_{ab}) illustrate the superior performance of EIT_{thx} compared to other metrics. When requiring 100% sensitivity at the smallest LBNP change (0-to-15 mmHg), EIT_{thx} has a much higher specificity of 87.5% compared to the best vital sign specificity of 25% (MAP).

ML analysis: individual technologies

The ML approach relied on a binary time-series classification (TSC) method using a random forest model, 15-fold cross-validation, and its overall performance was assessed using out-of-sample diagnostic performance. The out-of-sample AUC and F1 performance results for the binary TSC models using 7-min slope windows filtered for each LBNP change are shown in Fig. 6. F1 scores were used in addition to AUC to accurately measure the predictive performance of the ML models; unlike the balanced data set used in the above absolute and relative metric analyses, the ML data set often contained imbalanced data with more hypovolemic samples than normovolemic

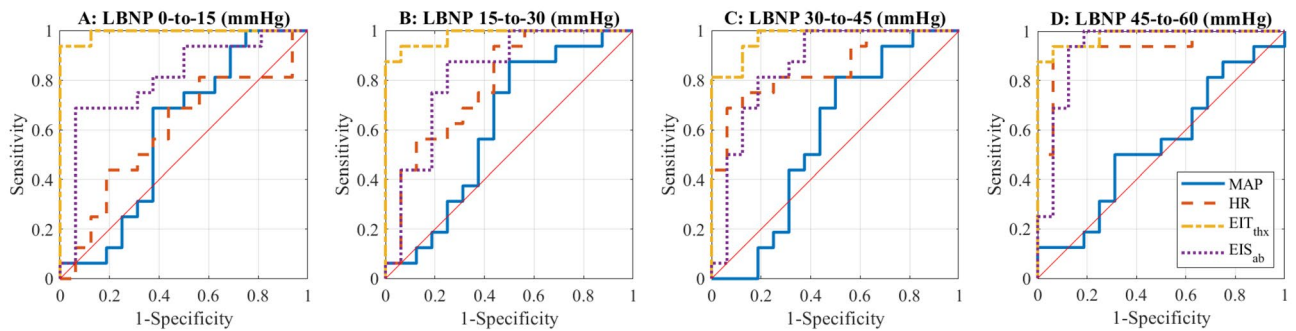


Figure 5. (A–D) ROC curves corresponding to MAP, HR, and the best metrics from EIT (thorax) and EIS (abdomen) for the relative analysis of change between consecutive LBNP levels and baseline change.

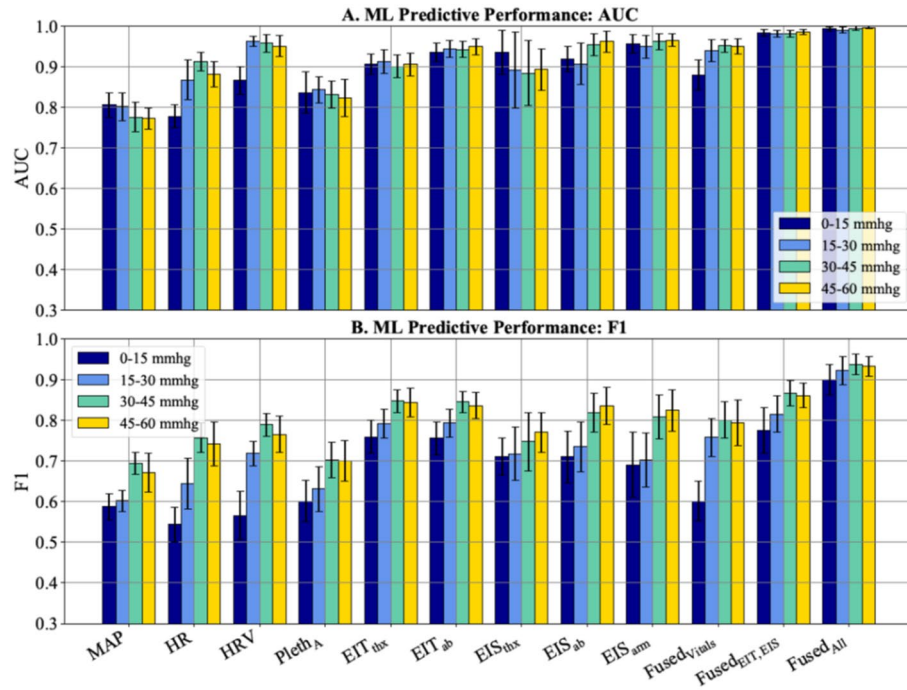


Figure 6. Out-of-sample random forest performance results (bars) and 95% confidence intervals for each 15-fold cross validation (whiskers) for models trained on 7-min slopes from vital signs, impedance, and fused metrics, shown for each LBNP change. Fused results represent performance of majority voting ensembles that combine vital sign metrics (Fused_{Vitals}), impedance metrics (Fused_{EIT,EIS}), and all metrics (Fused_{All}). All analyses compare slopes over the specified LBNP change to baseline—i.e. they perform bleed versus no-bleed analyses. No comparisons are performed between different LBNP levels, e.g. comparing slopes from LBNPs of 0-to-15 to those from LBNPs of 15-to-30 are not considered.

samples, making F1 score a more apt performance metric²⁰. Individual and fused results are shown for models trained on 7-min slopes because that window length yielded the best out-of-sample results (see Supplementary Appendix A.3 for performance of other time windows). The mean vital sign F1 scores for an LBNP change of 0-to-15 mmHg are 0.59, 0.54, 0.57, and 0.60 for MAP, HR, HRV, and Pleth_A, respectively. The EIT and EIS technologies performed generally better than vital signs, yielding mean F1 scores of 0.76, 0.76, 0.71, 0.71, and 0.69 for EIT_{thx}, EIT_{ab}, EIS_{thx}, EIS_{ab}, and EIS_{arm}, respectively.

ML analysis: multi-technology

In addition to measuring the out-of-sample predictive performance of each individual technology, each model's predictions were combined and used in a majority voting scheme to measure the predictive performance of combined technologies. Combining traditional vital sign technologies (Fused_{Vitals}) yielded F1 scores of 0.60, 0.76, 0.80 and 0.79 for the 0-to-15 mmHg, 15-to-30 mmHg, 30-to-45 mmHg, and 45-to-60 mmHg LBNP changes, respectively. Combining impedance-based technologies (Fused_{EIT,EIS}) yielded F1 scores of 0.78, 0.82, 0.87 and 0.86 for each LBNP change respectively. Combining all non-invasive technologies (Fused_{All}) yielded mean F1 scores of 0.90, 0.92, 0.94, and 0.93 for each LBNP change respectively, far outperforming all individual technologies and other multi-technology models (see last 3 sets of results in Fig. 6).

Discussion

We have previously described the high performance of non-invasive impedance-based measurements in the early identification of occult hemorrhage in a porcine model¹² and the ability of machine learning to improve this performance through the application of multivariable algorithms⁹. We have now extended this observation into an established human model of LBNP progressive circulatory instability from central hypovolemia. The present investigation evaluated the performance of multiple technologies to detect early subclinical changes in LBNP in an absolute and relative (change) sense compared to standard vital sign metrics.

Impedance-based non-invasive sensing performed best when used as relative metrics. The performance of relative EIT_{thx} is worth noting. It had an AUC of 0.99 (CI 0.95–1.00) at the smallest LBNP change (0-to-15 mmHg) and when requiring 100% sensitivity it yielded an extremely high specificity of 87.5% especially when compared to the 25% specificity achieved by the best vital sign metric (MAP) at this LBNP change. LBNP 0-to-15 mmHg is almost always clinically occult and existing detection solutions perform poorly at this LBNP level with high rates of false negatives. Given these encouraging results, it is important to highlight the clinical value of analyzing *relative metrics*. Based on Ref.¹⁴, LBNP levels of 15, 30, 45, and 60 mmHg respectively correspond to blood losses of 93, 193, 313, and 451 ml, and consequently, for LBNP levels *changes* of 0-to-15, 15-to-30, 30-to-45, and 45-to-60 mmHg there should be a corresponding increased blood volume loss of 93, 100, 120, and 138 ml. Thus, the relative metric-based analysis essentially evaluates if one can detect an increase in blood loss relative to no blood loss for these particular volumes. The data here suggests that a simple threshold applied to the change in EIT_{thx} is sufficient to detect these small occult-level blood volume changes (note that EIT_{thx} change data are all above zero in Fig. 3), i.e. no apparent baseline (healthy state) is needed in the relative metric-based analysis. At the same time, Fig. 3 shows that the change in the EIT_{thx} metric cannot determine the severity of the bleed, i.e. all changes evaluated (0-to-15 up to 45-to-60) appear essentially the same. Although not considered here, it would be interesting, in a larger study, to investigate if any combination of absolute or relative metrics using ML techniques could help determine the severity of the bleeds in terms of cumulative volume or rate.

Similar to previous investigations¹², we found that traditional vital signs measured with standard monitors such as ECG and plethysmography performed poorly in early detection of LBNP progression. EIT outperformed these metrics, with an observed AUC of 0.99 (CI 0.95–1.00) for EIT_{thx}, while the best clinical metric at the same LBNP level exhibited an AUC of only 0.60 (CI 0.38–0.79) (HR).

As in our earlier porcine study, impedance performs poorly as an absolute metric (e.g., see Table 3 in Ref.¹²). This can largely be attributed to the variation in body size and habitus between study participants. Given similar tissue impedance characteristics of all subjects (e.g., assuming muscle or adipose impedance is similar between subjects), a large subject will have a larger impedance compared to a small subject due to variations in electrode spacing and underlying constituent tissues. Unfortunately, simple calibration factors are generally not adequate, as the impedance is dependent on not only the subject's size (cross-sectional area), but also tissue composition (e.g., percent adipose tissue)²¹, hydration status (electrolyte content)²², and recent activity levels²³. ROC analysis of EIT metrics including various height, weight, body mass index (BMI), and body surface area (BSA) factors were evaluated; no factor explored yielded notable improvements in the absolute analysis.

With respect to out-of-sample predictive performance, a clear contrast emerges between the strong out-of-sample performance achieved using data generated by the impedance-based technologies as compared to the other technologies. Only impedance-based technologies were able to achieve AUCs > 90% across all four consecutive LBNP levels (see Fig. 6A). Mean AUC and F1 scores for impedance-based technologies dominated scores from all other individual technologies for the 0-to-15 mmHg LBNP change. HRV models tended to prioritize specificity over precision (i.e., do not flag subjects as hypovolemic unless they are very certain), resulting in strong AUCs, but weaker F1 scores.

Pleth-based results from Convertino and others, e.g.^{8,18,19}, appear promising but may have limitations when applied broadly. For example, investigators have established that in LBNP studies their Pleth-based metric can be separated into subjects that have high and low tolerance to hypovolemia^{15,17}. This may impair detection of blood loss in "high tolerance" individuals. As we hypothesize that impedance is a measure of the sensitivity to the blood loss and not to the body's *response* to the blood loss, our impedance-based technology might not be susceptible to false-negative findings in subgroups of subjects.

The near perfect performance of EIT for the initial 0–15 mmHg step is particularly encouraging. This early period in the transfer of circulating blood volume is generally not detectable clinically and or with alternative existing technologies (Fig. 5A). A clinical alarm that activated just after initial hemorrhage would give clinicians the largest possible window for evaluation and treatment before onset of frank deterioration.

A limitation of the study was the small sample size. A larger study could have allowed for additional sub-analysis investigating dependencies on factors such as age, sex, height, and BMI. In terms of non-impedance related metrics, other studies have performed these types of analyses, e.g. see²⁴ for a review. If the performance described here is replicated in larger, future studies, development of a clinical early detection alarm system seems possible as these results indicate that low rates of false negative and false positive may be possible even in heterogenous clinical populations. A future wearable system may be reduced to a belt of electrodes and a small wireless device.

There are a number of limitations to the results reported here. While LBNP may be among the best available mimics for early hemorrhage it has all the limitations of a model system, and performance in actual clinical setting may be inferior. LBNP subjects in this study are all healthy volunteers. Patients will have significant phenotypic heterogeneity, may be taking medications that alter physiologic responses, and will likely be receiving intravenous fluids at variable rates²⁵. A shortcoming of the ML-based approach is the reliance on calculating slopes of metrics from an overall LBNP-step-based experiment (5 min at each level). The slope calculations allowed for a large-enough database for ML training, but do not quite match the stepped, physical scenario of the LBNP model. However, the slope-based approach is how we envision the technology being implemented in practice. One additional limitation associated with the length of time-windows used for electrical impedance-based metric calculations should be noted. While the performance observed is exceptional using long time-windows (5-min here and 1-min in Ref.¹²), for shorter duration windows, one might expect respiratory and cardiac-related signals to reduce the performance as these signals will no longer be removed through the averaging operation. Window length should continue to be considered as the technology is further developed.

Conclusions

Impedance-based noninvasive technologies appear to have promising diagnostic performance in the early detection of hypovolemia in a LBNP model. This is especially true when evaluated as a change with respect to baseline, and when incorporated in a multivariate machine learning algorithm.

Methods

A previously described step protocol for LBNP¹⁴ was used to model hypovolemic cardiovascular instability in a cohort of healthy subjects. Prior to the study day, all subjects provided written informed consent after all procedures and risks of the study were fully explained, and the study was approved by Mayo Clinic's Institutional Review Board. All human research was performed in accordance with relevant guidelines/regulations. Subjects were continuously monitored during the experiment via observation and vital signs to ensure safety.

Experimental design

Subjects were instrumented with ECG and plethysmography for non-invasive vital sign monitoring, and a brachial arterial line was invasively introduced to capture mean arterial pressure (MAP). Non-invasive *test* devices included electrical impedance tomography (EIT) (SenTec AG, Landquart, Switzerland), electrical impedance spectroscopy (EIS) (Sciospec Scientific Instruments GmbH, Germany), Near infrared spectroscopy (NIRS) (custom-device⁹), and bioimpedance cardiography (BC) (Starling, Baxter, USA) as shown in Fig. 1. Only vital sign metrics and EIT/EIS data are presented in this report. High subject-to-subject variability with the particular NIRS configuration and its coupling to the subject skin limited its value, and BC did not provide additional value relative to past reports of its performance in LBNP studies (see Supplementary Appendix A.4)¹³. The impedance spectroscopy device (EIS) recorded electrical impedance spectra from 3 sites (chest, abdomen, and arm), and the impedance tomography device (EIT) collected multiple impedance measurements from two 16-electrode belts positioned around the subjects' chest and abdomen. The belt was placed just below the nipple line (see Fig. 1), so belt placement was essentially the same across sexes. The pressure waveforms/values, ECG voltages, Pleth voltages, and LBNP levels were captured using LabChart 7.0 (ADI, Sydney, Australia).

Prior to LBNP progression each subject rested in the supine position for approximately 5 min. The LBNP levels proceeded as follows: baseline, 15 mmHg, 30 mmHg, 45 mmHg, 60 mmHg, 70 mmHg, 80 mmHg, 90 mmHg, and 100 mmHg. The experiment aimed to hold the negative pressure stable for 5 min at each level. There was a necessary transition period, which occurred *within* each 5-min interval. Consequently, the average stable period of each level was 4.2 min, i.e. ~48 s transition periods. At each 5-min time-point the LBNP moved to the next level. Each subject underwent an LBNP progression from a baseline hemodynamic state until a clinically evident hypotensive state developed. The target hypotensive state was achieved when any of the following criteria was met: (a) systolic BP of ≤ 70 mmHg or (b) the subject asked to stop because of symptoms typically including, but not limited to, lightheadedness or nausea.

Data

The collected data was averaged during the stable periods of each LBNP level. Only the data associated with LBNP levels of 0, 15, 30, 45, and 60 mmHg were considered—to focus on clinically occult states. The different LBNP levels are referred to as *hypovolemia states* as they are meant to mimic a volume contraction secondary to processes such as hemorrhage.

Baseline data was collected for each subject for five minutes before the LBNP process was initiated (LBNP = 0 mmHg), and the protocol took between 28 and 54 min. The beat-to-beat HR was calculated from R-peaks extracted from the ECG data. Because only a single channel of Pleth data was available for some subjects,

a standard blood oxygen calculation was not possible. As a surrogate, the trough-to-trough area-under-the-curve was computed for each Pleth time-series (similar to that used in Ref.¹²) and denoted as $Pleth_A$.

The EIS data recorded the complex impedance (real and imaginary or magnitude and phase) at 100 logarithmically-spaced frequencies ranging from 100 Hz to 1 MHz at a rate of ~1 spectra per minute at three anatomical locations. Results are presented for the impedance magnitude and frequency yielding the maximum AUC at each site considered. We denote the impedance spectroscopy results as EIS_{thx} , EIS_{ab} , and EIS_{arm} for the thorax, abdomen, and arm sites, respectively. There is an abundance of EIT data, which could be used in a standard way to image cross-sections of the thorax and abdomen. However, here we average a subset of impedance data (specified in Supplementary Appendix A.5) at each time and perform checks to ensure that the data is valid (i.e., electrode–skin contact is good), similar to that described in Ref.¹². This approach yields a single belt location-specific mean impedance value \pm (standard deviation) for each LBNP level. We denote the impedance tomography results as EIT_{thx} and EIT_{ab} for the thorax and abdomen sites, respectively.

Analysis

The data is visualized in terms of 25–75 quartile box-and-whisker plots for each metric and LBNP level. Each box-and-whisker plot is produced using one metric value from each subject. Two main types of analyses were performed: absolute and relative. The *absolute analysis* explored the potential of using the raw values of each metric at a specific LBNP level to distinguish between normovolemia (LBNP = 0 mmHg) and hypovolemia (LBNP > 0 mmHg) states. The *relative analysis* investigated how significantly different a metric change between two LBNP levels (e.g., a metric difference between LBNP levels of 0 and 15 mmHg) was with respect to baseline variability (metric change over the 5-min of baseline). As not all metrics values yielded normal distributions, Wilcoxon rank sum tests were used with significance defined for p-values < 0.05. Additionally, receiver operating characteristic (ROC) analysis was performed on both absolute and relative metrics and the area-under-the ROC curves (AUCs) are reported²⁶; confidence intervals of the AUCs are also computed using 500 bootstrap samples.

ML and analysis

A binary time series classification (TSC) method was employed to assess the out-of-sample diagnostic performance of (1) each individual technology and (2) the combination of all technologies for identifying hypovolemic states. A random forest model was used to classify subjects as either hypovolemic or normovolemic using 15-fold cross validation in which data from 13 randomly chosen subjects were used for training and the remaining 3 for testing. The random forest models; did not set a max tree depth (as specifying the tree depth can result in underfitting²⁷), 500 estimators (trees) were used because that number represents a typical default value²⁸, and splitting criteria was based on the Gini index, which are appropriate for classification problems²⁹. Similar to the absolute analysis, the ML approach was designed to determine whether or not subjects were in a normovolemic (LBNP = 0 mmHg) or hypovolemic state (LBNP > 0 mmHg). Data produced from each technology was initially modeled individually; each model's predictions were then used in a majority voting scheme to evaluate any performance improvements associated with the multi-modal signals captured. A majority voting ensemble was employed as this approach often exceeds performance of a single machine learning model in clinical settings^{30–32}. Combinations considered for ML ensembles included vital metrics (Fused_{vitals}), impedance metrics (Fused_{EIT,EIS}), and all metrics (Fused_{All}). Instead of using the entire training set for each ML model in the ensembles, we use each technology's data individually, and then perform the majority vote using predictions made by each technology's random forest model. This simplified majority voting method is easier to interpret because it separates performance achieved by each technology, and yet achieves strong combined performance.

In order to apply supervised learning models to the time series data for each technology, deterministic features were generated by calculating the slope across rolling time windows of increasing length. Slopes generated during the normovolemic and hypovolemic states were treated as class 0 and class 1, respectively. Due to the limited normovolemic data produced during the study, the normovolemia dataset was created by concatenating the slopes across the 2-, 3-, and 4-min rolling time windows. This is a valid approach to data augmentation given that the subject's baseline values were observed to be stable across all technologies. Each random forest model was trained and tested using 111 normovolemic samples and between 257 and 284 hypovolemic samples depending on the window length used to calculate hypovolemic slopes, representing a range of binary class imbalances of 43.2% to 39.1%. The number of features used for each technology along with a list of feature names are provided in Supplementary Appendix A.6 (their corresponding importance are noted in Supplementary Appendix A.7). F1 scores were used in addition to AUC to accurately measure the predictive performance of the ML models; unlike the balanced data set used in the absolute and relative metric analyses, the ML data set often contained imbalanced data with more hypovolemic samples than normovolemic samples, making F1 score a more apt performance metric²⁰. Additional features were generated for some technologies; for example, raw, millisecond-level HR data was used to generate 20 individual minute-level heart rate variability (HRV) features from the time, frequency, and nonlinear domains using the PyHRV library (Ref.³³ and/or see Supplementary Appendix A.8 for the list of features). Many modern TSC applications can be effectively approached through the use of deep learning³⁴, but the limited size of the dataset meant that relying on deep learning models for TSC in this instance was not feasible.

Data availability

The data that support the findings of this study have been made into an open-source data set and are available at <https://doi.org/10.5281/zenodo.10119427>.

Received: 10 November 2023; Accepted: 8 April 2024

Published online: 15 April 2024

References

- Johnson, M. C. *et al.* Compensatory reserve index: Performance of a novel monitoring technology to identify the bleeding trauma patient. *Shock* **49**(3), 295–300. <https://doi.org/10.1097/SHK.0000000000000959> (2018).
- Callcut, R. A. *et al.* The why and how our trauma patients die: A prospective Multicenter Western Trauma Association study. *J. Trauma Acute Care Surg.* **86**(5), 864–870. <https://doi.org/10.1097/TA.0000000000002205> (2019).
- Van Breugel, J. M. M. *et al.* Global changes in mortality rates in polytrauma patients admitted to the ICU—A systematic review. *World J. Emerg. Surg.* **15**(1), 1–14. <https://doi.org/10.1186/s13017-020-00330-3> (2020).
- Johnson, B. D. *et al.* Reductions in central venous pressure by lower body negative pressure or blood loss elicit similar hemodynamic responses. *J. Appl. Physiol.* **117**(2), 131–141. <https://doi.org/10.1152/jappphysiol.00070.2014> (2014).
- Bellamy, R. F. Causes of death in conventional warfare: Implications for combat casualty care research. *Military Med.* **149**, 55–62 (1984).
- Eastridge, B. J. *et al.* Death on the battlefield (2001–2011): Implications for the future of combat casualty care. *J. Trauma Acute Care Surg.* **73**, S431–S437. <https://doi.org/10.1097/TA.0b013e3182755dcc> (2012).
- Shackelford, S. A. *et al.* Early identification of uncontrolled hemorrhage after trauma: Current status and future direction. *J. Trauma Acute Care Surg.* **77**, S222–S227. <https://doi.org/10.1097/TA.000000000000198> (2014).
- Convertino, V. A., Grudic, G., Mulligan, J. & Moulton, S. Estimation of individual-specific progression to impending cardiovascular instability using arterial waveforms. *J. Appl. Physiol.* **115**(8), 1196–1202. <https://doi.org/10.1152/jappphysiol.00668.2013> (2013).
- Elliott, J. T. *et al.* A portable polyanatomic optico-impedance system for early warning of systemic disease: Evaluation in occult hemorrhage. In *Optical Diagnostics and Sensing XXI: Toward Point-of-Care Diagnostics* 14. <https://doi.org/10.1117/12.2583327> (2021).
- Klein, S. *et al.* 1292: A multivariate machine learning algorithm for occult hemorrhage in a porcine model. *Crit. Care Med.* **49**(1), 652–652. <https://doi.org/10.1097/01.ccm.0000731056.53582.4c> (2021).
- Minichiello, J. M. *et al.* A Porcine Model of Clinically Occult Hemorrhage. <https://mhsrcs.amedd.army.mil/sites/mhsrcs2020/submissions/Abstracts/Print.aspx?AID=1621> (2020).
- Murphy, E. K. *et al.* Detection of subclinical hemorrhage using electrical impedance: A porcine study. *Physiol. Meas.* **43**(5), 055001. <https://doi.org/10.1088/1361-6579/ac6cc6> (2022).
- Convertino, V. A., Ludwig, D. A. & Cooke, W. H. Stroke volume and sympathetic responses to lower-body negative pressure reveal new insight into circulatory shock in humans. *Auton. Neurosci. Basic Clin.* **111**(2), 127–134. <https://doi.org/10.1016/j.autneu.2004.02.007> (2004).
- Hinojosa-Laborde, C. *et al.* Validation of lower body negative pressure as an experimental model of hemorrhage. *J. Appl. Physiol.* **116**(4), 406–415. <https://doi.org/10.1152/jappphysiol.00640.2013> (2014).
- Schiller, A. M., Howard, J. T. & Convertino, V. A. The physiology of blood loss and shock: New insights from a human laboratory model of hemorrhage. *Exp. Biol. Med.* **242**(8), 874–883. <https://doi.org/10.1177/1535370217694099> (2017).
- Schiller, A. M., Howard, J. T., Lye, K. R., Magby, C. G. & Convertino, V. A. Comparisons of traditional metabolic markers and compensatory reserve as early predictors of tolerance to central hypovolemia in humans. *Shock* **50**(1), 71–77. <https://doi.org/10.1097/SHK.0000000000001034> (2018).
- Schlottman, T. E. *et al.* Evidence for misleading decision support in characterizing differences in tolerance to reduced central blood volume using measurements of tissue oxygenation. *Transfusion* **60**(S3), S62–S69. <https://doi.org/10.1111/trf.15648> (2020).
- Convertino, V. A. *et al.* Use of advanced machine-learning techniques for noninvasive monitoring of hemorrhage. *J. Trauma Inj. Infect. Crit. Care* **71**, 25–32. <https://doi.org/10.1097/TA.0b013e3182211601> (2011).
- Moulton, S. L., Mulligan, J., Grudic, G. Z. & Convertino, V. A. Running on empty? The compensatory reserve index. *J. Trauma Acute Care Surg.* **75**(6), 1053–1059. <https://doi.org/10.1097/TA.0b013e3182aa811a> (2013).
- DeVries, Z. *et al.* Using a national surgical database to predict complications following posterior lumbar surgery and comparing the area under the curve and F1-score for the assessment of prognostic capability. *Spine J.* **21**(7), 1135–1142. <https://doi.org/10.1016/j.spinee.2021.02.007> (2021).
- Coppini, L. Z., Waitzberg, D. L. & Campos, A. C. L. Limitations and validation of bioelectrical impedance analysis in morbidly obese patients. *Curr. Opin. Clin. Nutr. Metab. Care* **8**(3), 329–332. <https://doi.org/10.1097/01.mco.0000165013.54696.64> (2005).
- O'Brien, C., Young, A. J. & Sawka, M. N. Bioelectrical impedance to estimate changes in hydration status. *Int. J. Sports Med.* **23**(5), 361–366. <https://doi.org/10.1055/s-2002-33145> (2002).
- Dehghan, M. & Merchant, A. T. Is bioelectrical impedance accurate for use in large epidemiological studies? *Nutr. J.* **7**(1), 1–7. <https://doi.org/10.1186/1475-2891-7-26> (2008).
- Goswami, N., Blaber, A. P., Hinghofer-Szalkay, H. & Convertino, V. A. Lower body negative pressure: Physiological effects, applications, and implementation. *Physiol. Rev.* **99**(1), 807–851. <https://doi.org/10.1152/physrev.00006.2018> (2019).
- Harris, T., Coats, T. J. & Elwan, M. H. Fluid therapy in the emergency department: An expert practice review. *Emerg. Med. J.* **35**(8), 511–515. <https://doi.org/10.1136/emered-2017-207245> (2018).
- Bradley, A. P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit.* **30**(7), 1145–1159. [https://doi.org/10.1016/S0031-3203\(96\)00142-2](https://doi.org/10.1016/S0031-3203(96)00142-2) (1997).
- Zhou, S. & Mentch, L. Trees, forests, chickens, and eggs: When and why to prune trees in a random forest. *Stat. Anal. Data Min.* **16**(1), 45–64. <https://doi.org/10.1002/sam.11594> (2021).
- Probst, P., Wright, M. & Boulesteix, A. *Hyperparameters and Tuning Strategies for Random Forest*, Vol. 2010, 1–19. <https://doi.org/10.1002/widm.1301> (2018).
- Nembrini, S., König, I. R. & Wright, M. N. The revival of the Gini importance? *Bioinformatics* **34**(21), 3711–3718. <https://doi.org/10.1093/bioinformatics/bty373> (2018).
- Atallah, R. & Al-Mousa, A. Heart disease detection using machine learning majority voting ensemble method. In *2019 2nd Int. Conf. New Trends Comput. Sci. ICTCS 2019—Proc.* 1–6. <https://doi.org/10.1109/ICTCS.2019.8923053> (2019).
- Osamor, V. C. & Okezie, A. F. Enhancing the weighted voting ensemble algorithm for tuberculosis predictive diagnosis. *Sci. Rep.* **11**(1), 1–11. <https://doi.org/10.1038/s41598-021-94347-6> (2021).
- Ben Jabra, M., Koubaa, A., Benjdira, B., Ammar, A. & Hamam, H. Covid-19 diagnosis in chest X-rays using deep learning and majority voting. *Appl. Sci.* **11**(6), 884. <https://doi.org/10.3390/app11062884> (2021).
- Gomes, P., Margaritoff, P. & da Silva, H. P. pyHRV: Development and evaluation of an open-source python toolbox for heart rate variability (HRV). In *Electr. Electron. Comput. Eng.* 822–828. https://www.researchgate.net/publication/333611305_pyHRV_Development_and_Evaluation_of_an_Open-Source_Python_Toolbox_for_Heart_Rate_Variability_HRV (2019).
- Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L. & Muller, P. A. Deep learning for time series classification: A review. *Data Min. Knowl. Discov.* **33**(4), 917–963. <https://doi.org/10.1007/s10618-019-00619-1> (2019).

Acknowledgements

This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs endorsed through the Department of Defense through the Defense Medical Research and Development Program, Joint Program Committee 6, Combat Casualty Care Research Program under Award No. W81XWH1820076. Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the Department of Defense.

Author contributions

EKM led the writing of the manuscript with significant contributions from SRB and NAP. The Mayo team (MJJ, TBC, CPJ, RJR, CCW, JWS, and JRAS) led by MJJ aided in (1) modifying their lower body negative pressure model for this study in collaboration with SBK and the PIs of the study (JTE, RJH, VSV, and NAP), and in (2) all parts of the LBNP study on 16 subjects (recruitment and data collection). The direct data analysis was led by EKM whereas the ML effort was led by SRB. Additional data analysis was performed by NR, YS, CPJ, and JTE. VSV provided expert advice on ML efforts. Significant editing was additionally performed by RJH and VSV, and other important editing information was provided by SBK and NR. All authors reviewed the manuscript.

Competing interests

NP has issued patents in this subject, NP/RH/JE/VV/EM have a patent pending in related technology, NP/VV/RH/JE are founders of a company developing technologies for diagnosis of acute system disease. All other authors declare that they do not have any competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-59139-8>.

Correspondence and requests for materials should be addressed to E.K.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024