Title

Data-Driven Approach for Automatic Detection of Aortic Valve Opening: B point Detection from Impedance Cardiogram

Authors' names and affiliations:

Shafa-at Ali Sheikh[1,2], Nil Z. Gurel[3], Shishir Gupta[5], Ikenna V. Chukwu[5], Oleksiy Levantsevych[5], Mhmtjamil Alkhalaf[5], Majd Soudan[5], Rami Abdulbaki[5], Ammer Haffar[5], Viola Vaccarino[5], Omer T. Inan[2], Amit J. Shah[5,6,7], Gari D. Clifford[1,4], Ali Bahrami Rad[1]

[1]Department of Biomedical Informatics, Emory University, Atlanta, USA

[2]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, USA

[3]Neurocardiology Research Center of Excellence and Cardiac Arrhythmia Center, David Geffen School of Medicine at UCLA, Los Angeles, USA

[4]Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, USA

[5]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, USA

[6]Department of Medicine, Division of Cardiology, Emory University School of Medicine, Atlanta, USA

[7] Atlanta Veterans Affairs Health Care System, Atlanta, USA

Corresponding author:

Shafa-at Ali Sheikh, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, USA. Phone: +1-404-889-5212; ssheikh9@gatech.edu.

Abstract (Word limit: 250)

Pre-ejection period (PEP), an indicator of sympathetic nervous system activity, is useful in psychophysiology and cardiovascular studies. Accurate PEP measurement is challenging and relies on robust identification of the timing of aortic valve opening, marked as the B point on impedance cardiogram (ICG) signals. The ICG sensitivity to noise and its waveform's morphological variability makes automated B point detection difficult, requiring inefficient and cumbersome expert-visual annotation. In this paper, we propose a machine-learning-based automated algorithm to detect the aortic valve opening for PEP measurement, which is robust against noise and ICG morphological variations. We analyzed over 60 hours of synchronized ECG and ICG records from 189 subjects. A total of 3,657 averaged beats were formed using our recently developed ICG noise removal algorithm. Features such as the averaged ICG waveform, its first and second derivatives, as well as high-level morphological and critical hemodynamic parameters were extracted and fed into the regression algorithms to estimate the B point location. The morphological features were extracted from our proposed "variable" physiologically valid search-window related to diverse B point shapes. A subject-wise nested cross-validation procedure was performed for parameter tuning and model assessment. After examining multiple regression models, Adaboost was selected, which demonstrated superior performance and higher robustness to five state-of-the-art algorithms that were evaluated in terms of low mean absolute error of 3.5 ms, low median absolute error of 0.0 ms, high correlation with experts' estimates (Pearson coefficient = 0.9), and low standard deviation of errors of 9.2 ms. For reproducibility, an open-source toolbox is provided.

Keywords: Impedance Cardiogram; Aortic valve opening; B point detection; Pre-ejection period; Machine learning based automated algorithm;

1. Introduction

Impedance cardiography is a non-invasive technique to assess mechanical aspects of cardiac function by measuring the electrical impedance of the thorax and its time-related changes. Impedance cardiogram (ICG) is the first-order time derivative (dZ/dt) of the thoracic impedance signal (Z). Parameters and biomarkers derived using ICG have been found effective in the prediction of mortality in the general population (Silva et al., 2019), assessment of early cardiovascular disease (DeMarzo, 2020), management of hypertension (Medina et al., 2018), and evaluating cardiac autonomic functioning in stress studies (Cohen et al., 2020; Czarnek et al., 2021; Silvia et al., (2020); Peters et al., 2018; Kibler, 2018; Malcolm et al., 2016; Meijer et al., 2011; Kelsey et al., 2007). However, intra- and inter-subject morphological variability of the ICG waveform coupled with noise sensitivity are the major limitations for reliable ICG-based applications and have likely limited its use in research and clinical practice.

We addressed ICG noise sensitivity by proposing a novel automated noise removal algorithm (Sheikh, Shah, Levantsevych, et al., 2020), but automated and reliable fiducial points detection remained a challenge due to ICG's morphological variability. Among the major fiducial points, the B point, representing the aortic valve opening, is the most commonly used fiducial point since it facilitates derivation of critical parameters including pre-ejection period (PEP), left ventricular ejection time, stroke volume, and cardiac output. Unfortunately, its accurate, reliable, and automated detection is a major challenge due to its diverse shape manifestations. For instance, researchers in existing literature have identified six different B point shapes (Lababidi et al., 1970; Sherwood et al., 1990; Ono et al., 2004; Arbol et al., 2017). Figure 1 illustrates those B point morphologies extracted from our dataset.

Multiple rule-based and/or data-driven approaches were developed in the existing literature for automated B point detection, but all these efforts lacked robustness to noise and morphological variability of ICG waveform and suffered from low accuracy (for details see Appendix A). The initial B point detection methods of zero-crossing of ICG and 15% of $(dZ/dt)_{max}$ (Kubicek et al.,1966; Kubicek et al., 1970) were rendered ineffective due to the baseline shifts caused by the respiratory influences. Other rule-based algorithms, such as second-derivative rule (Debski et al., 1991; Demarzo et al., 1996); third-derivative rule

(Debski et al., 1993); inflection point rule (Ono et al., 2004), failed to encompass the B point morphological variations present in the ICG signal. Researchers also applied algorithms using the fixed search-windows (Nederend et al., 2017; Arbol et al., 2017; Nabian et al., 2018), and monotonically increasing segments on ICG cycle (Forouzanfar et al., 2018), which suffered from an incorrect B point detection in case the search-window/segment fell outside the normal PEP physiological range of 70 - 175 ms (Arbol et al., 2017). Data-driven approaches applied for automatic detection of the B point (Lozano et al., 2007; Cieslak et al., 2018) were based on a very small data set (33/37 subjects). Furthermore, the ensemble average methods used in those approaches did not perform any prior noise/artifact removal. Van et al., (2013) showed the discrepancy between the PEP estimated using Lozano's algorithm (Lozano et al., 2007) and actual PEP. In Cieslak et al., (2018), the reference B point was labeled in the RC interval using the third-derivative rule ignoring the morphological variations of the B point. Therefore, an automated B point detection algorithm that is robust to the morphological variations of ICG signal is still needed.

In this study, we aimed to develop and validate a data-driven model for automatic detection of aortic valve opening (the B point on the ICG). To do so, we examined multiple machine learning techniques such as linear regression (ordinary least squares (OLS), least absolute shrinkage and selection operator (LASSO), and Ridge), kernel ridge regression (Murphy, 2012), support vector regression (SVR) (Drucker et al., 1996), random forest (Breiman, 2001), and Adaboost (Drucker, 1997) with different sets of ICG features. We leveraged more than 60 hours of recordings from 189 subjects for training and testing of the models. The regression technique with a set of features that achieved the best results were considered our proposed model (see Section 3). The performance of this model was compared against the expert adjudication and state-of-the-art B point detection methods (see Section 4.1). If accurate, the proposed model can be used for fully automated PEP measurements that are accurate and reproducible, and greatly increase the efficiency and utility of this measurement across several research and clinical settings.

2. Materials and Methods

2.1 Dataset and the Study Protocol

We recorded synchronized ECG/ICG signals from 189 participants of two different studies approved by the Emory Institutional Review Board. We combined datasets from those two studies to increase the volume and variability for training and testing data of the machine learning models. In this paper, we will refer to the studies as study A and study B. Written informed consent was obtained from each participant before both studies. The dataset used for the design of the regression algorithm comprised 3657 ECG/ICG averaged beats derived from more than 60 hours of recordings. An averaged ECG/ICG beat was formed from a 60-second non-overlapping window of the synchronized ECG/ICG using the three-stage EA noise removal algorithm (Sheikh, Shah, Levantsevych, et al., 2020).

Study A was performed as an ancillary project on the data from the subjects in the Emory Twin Follow-up Study, which analyzed the relationship between psychological stress and heart disease (Vaccarino, 2015). Signals were collected from a group of 140 male participants (17 PTSD, and 123 non-PTSD) with ages having a mean and standard deviation (SD) of 68 and 2.5 years, respectively. Study A comprised of neutral and stress sessions. For study A, the averaged ECG/ICG beats were computed for both neutral and stress sessions. In the neutral session of approximately 4 to 7 minutes (min), the participants listened to two neutral audio scripts, followed by a short question/answer period. In the stress session of approximately 4 to 9 min, the participants listened to two trauma audio scripts (incidents from the participant's own life) followed by a short question/answer period, which was done for the purposes of eliciting acute post-traumatic stress disorder (PTSD) symptoms. The protocol details are available in our prior publication (Sheikh, Shah, Levantsevych, et al., 2020).

Study B was performed as an ancillary project on the data from the second and third day of the 3-day protocol from the project which analyzed the effects of the closed loop vagal nerve stimulations (VNS) on PTSD patients (Bremner, 2016). The project examined the potency and kinetics of the neurologic, autonomic peripheral, inflammatory, and behavioral responses to transcutaneous cervical vagal nerve stimulations (tcVNS) versus sham treatment, at baseline and in response to stressful traumatic scripts related to personal traumatic events, as well as a series of other stressors. Signals were collected from 49 (18 male

(6 PTSD), 31 female (19 PTSD)) participants with ages having a mean (SD) of 35 (12.7) years. Study B comprised baseline, preparation for the speech, speech, non-invasive tcVNS, recovery, and arithmetic sessions. The ensemble averaged ECG/ICG beats were computed for all sessions except the tcVNS stage due to the induction of the electrical stimulation artifacts in the ICG signal. After the baseline session of 8 mins, the participants prepared a response against a threat accusation in two minutes. In the speech task session, the participants delivered a speech for 2 to 3 min to defend themselves. Then active or sham tcVNS was applied for 2 min using hand-held tcVNS devices (GammaCore, ElectroCore, Basking Ridge, New Jersey). After resting for 8 min in the recovery session, all participants answered arithmetic questions for 2 to 3 min. The protocol details are available at Gurel et al., (2020).

Synchronized ECG and ICG signals were recorded by trained researchers using impedance and electrocardiography equipment (BIOPAC Systems Inc., Goleta, CA). BIOPAC module settings, and electrode configuration used for recording ECG and ICG signals for both studies can be found in our prior work (Sheikh, Shah, Levantsevych, et al., 2020; Sheikh et al., 2022) For some subjects (three in study A, and four in study B), no/noisy signal was observed due to electrode disconnect/improper connections during some tasks. This specific data was discarded from the analysis. Also, we did not remove any outlier from the analysis. The reference PEP values were computed using EA ECG/ICG beats as the RB interval (from the R peak of EA ECG beat to the B point of synchronized EA ICG beat) (Seery et al., 2016; Kelsey et al., 2010). R peaks were automatically detected via the PhysioNet Cardiovascular Signal Toolbox using "jqrs" algorithm (Behar et al., 2014; Johnson et al., 2014; Vest et al., 2018).

For Study A and Study B, three and two expert physicians with physiology training independently performed the B point annotation on the EA ICG beats. Then, the annotation discrepancies which an independent reviewer discovered were resolved via collective review by the expert physicians. The reference B points were marked in line with the methodological guidelines to encompass all morphological variations of the ICG waveform (Lababidi et al., 1970; Sherwood et al., 1990; Ono et al., 2004; Arbol et al., 2017). The "Impedance Cardiogram Manual Annotation Application (ICMAA)" from our open-source

ICG toolbox (Sheikh, Shah, Inan, and Clifford, 2020) was used to manually annotate the B point and compute PEP values.

2.2 Pre-processing and Signal Segmentation

AcqKnowledge 5.0 software was used for data storage, display, and export to ".mat" format. MATLAB R2017b (MathWorks, Inc., Natick, MA) was used for processing signals and applying the three-stage EA algorithm to generate the averaged ICG beats. The ECG signal was filtered with finite impulse response (FIR) band-pass filters, with lower and upper cutoff frequencies of 0.6 Hz and 40 Hz, to remove the low-frequency drift and high-frequency noise (Sornmo et al., 2005). R peaks were detected in the analysis window of ECG signals using an open-source toolbox (Vest et al., 2018). The median of RR intervals of the ECG beats in the analysis window was used for averaging the ECG/ICG signals. The ICG signal was filtered using a 4th-order Butterworth band-passed filter with the lower and upper cutoff frequencies of 0.5 Hz and 40 Hz. The Butterworth filter was applied in both forward and backward directions to avoid phase shift (Gustafsson, 1996). ICG beats in an analysis window with synchronized R peaks were used for generating the averaged ICG beats after noisy beats removal using the three-stage EA algorithm (Sheikh, Shah, Levantsevych, et al., 2020). A total of 1488 and 2169 synchronized averaged ECG/ICG beats were formed from the signals of study A and study B, respectively.

2.3 Machine Learning Methods

The averaged ECG/ICG beats were computed from the raw ECG/ICG signals using the three-stage EA algorithm (Sheikh, Shah, Levantsevych, et al., 2020). The feature vector consisting of ICG waveforms, its first and second-order derivatives, and high-level features was extracted from the averaged beats using MATLAB R2017b (MathWorks, Inc. Natick, MA). The feature vectors were then fed into the regression algorithm for estimating the B point location. A nested cross-validation (CV) procedure was used for hyperparameter tuning and model assessment such that an inner 5-fold CV loop with a grid-search was used for tuning the hyperparameters of the regression algorithm, and an outer 10-fold CV loop was used for

model assessment (Krstajic et al., 2014). In this work, we examined multiple candidate regression models as described in section 2.3.2. All processes were implemented using Python scikit-learn v0.24.

2.3.1 Feature extraction

The feature vector consisted of the 465 EA ICG waveform, its first and second-order derivatives, and high-level features derived from the averaged beats (see Appendix B for a detailed technical description). A total of 450 features extracted from the EA ICG waveform and its first and second-order derivatives in the time interval [t_R, t_R +150] ms, where t_R was the location of R peak (time index) on the synchronized EA ECG signal. In this work, these 450 features will be referred to as "time-series features" (in contrast to high-level features). The time-series features were extracted from the fixed interval to ensure a constant feature-vector length. Additional 15 high-level features were based on the morphology of the ICG waveform and critical hemodynamic parameters from a variable search-window.

Since the B point manifests itself in various morphological shapes at different locations on an ICG waveform, we designed a variable search-window to extract the morphological features. In literature, the search-windows for detecting the B point are mostly fixed and/or defined based on the time interval related to the R peak, C point, and RC interval. This led to locating the B point out of the normal physiological range in some cases (false positive error). The normal range for PEP (measured as QB interval) is 70-175 ms (Arbol et al., 2017). With an adjustment for QR offset of 35 ms, PEP measured as RB interval will have a normal physiological range of 35-140 ms (Forouzanfar et al., 2019). However, the B point search in this fixed range can also be severely affected by ICG waveforms with shorter RC interval (< 140 ms) and/or M-shaped or notched C waves (for details, see section 4). Therefore, a variable search-window based on PEP physiological range and maximum ejection velocity of blood through aorta (associated with the amplitude of C point) as depicted in Figure 2 is proposed in this work for the extraction of morphological features with the following limits:

Lower limit: t_R+35 ms

Upper limit: minimum $\{t_R+140, t_{(0.66*(\text{C point amplitude}))}\}$ ms,

where $t_{(0.66*(\text{C point amplitude}))}$ is the time index on the ICG waveform before C point with 0.66*C point amplitude. The threshold of 0.66*C point amplitude was empirically selected to guard against the M-shaped/notched C waves, which could result in the erroneous B point detection (for details, see Section 4.1 and Figure 4).

A total of 15 high-level features were extracted by considering the morphology of the ICG waveform and critical hemodynamic parameters. The morphological features include time index and magnitude of the last inflection point, last notch, and rapid slope change on the ICG waveform in the proposed search-window (features 451-453 and 463-465, cf. Table B1). We only considered those inflection points which satisfied two conditions: (1) the first derivative of the ICG signal was positive ($(d^2Z/dt^2) > 0$), and (2) the sign of the second derivative of the ICG signal (d^3Z/dt^3) changed from negative to positive. In case of finding multiple inflection points satisfying the above-mentioned conditions, we selected the last inflection point in the search-window. A notch was detected as a negative peak in the search-window. Again, in case of finding multiple notch points, we chose the last one. The rapid slope change used to identify the "change in the gradient" (Figure 1(d)) was identified as the maximum of the second derivative of the ICG signal ($(d^3Z/dt^3)_{max}$). In case of not finding any inflection or notch point in the proposed search-window, we replaced the associated feature value with zero. The hemodynamic parameters including the C point amplitude associated with the maximum ejection velocity of blood in the aorta and initial systolic time interval (RC interval) were also calculated (features 454 and 462, cf. Table B1). Also, the differences in the amplitudes and time indices of the morphological features and hemodynamic parameters were made part of the feature vector (features 455-461, cf. Table B1).

2.3.2 Regression model and validation

Multiple candidate regression models estimated the opening of the aortic valve (B point). The models included linear regression (OLS, LASSO, and Ridge), kernel ridge regression, support vector regression

(SVR), random forest, and Adaboost regressor. All models were trained and tested via a nested cross-validation procedure.

In its OLS flavor, the linear regression algorithm tried to estimate the coefficients of a linear combination of features to minimize the sum of squares of the residuals for B point location. LASSO (L1-norm regularization) and Ridge (L2-norm regularization) were used to reduce the model complexity/flexibility and improve its generalization power. LASSO also helped to select important features by forcing the coefficients of unimportant features to be zero or small (Hastie et al., 2001). Kernel ridge regression combined ridge regression with the kernel trick leading to a nonlinear regressor (Murphy, 2012). Support vector regression differed from the kernel ridge regression in terms of its loss function (epsilon insensitive loss) (Drucker et al., 1997). In random forest regression, nearly uncorrelated decision trees were bootstrapped by using randomized subsets of data/features drawn from the original dataset (Breiman, 2001). In Adaboost regression, a sequence of weak learners was designed where each learner learned from the mistakes made by the previous learners to minimize the loss function (Freund et al., 1997).

A nested cross-validation architecture was used, with an outer 10-fold CV for model assessment, and an inner 5-fold CV for hyperparameter tuning. For each training set of the outer CV, the hyperparameters were optimized via the inner 5-fold CV and the grid-search (see Appendix C for details). The regressor models were trained and assessed in the outer loop. The data were partitioned such that to secure different subjects were used in the training and testing sets (subject-wise CV). For the featureless ICG cycles (ICG beat without any morphological shape as shown in Figure 1(f)) the third-derivative rule was applied to predict the B point location in the test set (Debski et al., 1993).

2.4 Evaluation Criteria

The comparative analysis to evaluate the performance and robustness of the candidate regression models was done in terms of the mean absolute error (MAE) with 95% confidence interval (CI), median absolute error (MedAE), and standard deviation of the error (SDE). The error was calculated as the difference

between the PEP values for expert (expert-PEP) and estimated PEP values by the regression models. The association strength between the expert-PEP and each model's estimated PEP value was also determined using Pearson correlation analysis. Then, the algorithm with the best performance metrics was proposed. In addition, we also evaluated the mean error (ME) with 95% CI and median error (MedE) to estimate the bias of regression algorithms.

In addition to the above-mentioned performance criteria, the agreement between the expert-PEP and the PEP estimated by the algorithms (proposed and popular) were assessed using Bland-Altman plots (Bland and Altman, 1986). The estimated PEP values were checked for normality (using the Lilliefors test (Lilliefors, 1969)) and compared with the expert-PEP via a set of Bonferroni-corrected paired Wilcoxon Sign Rank tests.

2.5 Feature Importance Analysis

Feature importance analysis was performed to identify the features playing a significant role in B point detection. The "mean decrease impurity" which is the total decrease in node impurity averaged over all trees (Breiman et al., 1984) was calculated and used for ranking the feature importance in the Adaboost regressor. For 10-fold outer CV, 10 feature sets were extracted via the Adaboost model with the tuned hyperparameters in the inner 5-fold CV (Table C2). Each feature set consisted of the features with an importance-value greater than the threshold (mean of all features importance values). The process was randomly repeated ten times (10 times repeated 10-fold CV), which resulted in the extraction of 100 sets of features. Feature score was computed by dividing the feature-occurrence frequency by 100. For instance, a "feature score = 1" implied that this feature had importance values greater than the threshold in all 100 feature sets obtained via 10 times repeated 10-fold CV. Feature scores were plotted against the feature indices to visualize the relative feature importance and identify the important features.

3. Results

The performances of the seven regression algorithms described earlier were evaluated using a nested cross-validation architecture with 10-fold CV in the outer loop for model assessment and 5-fold CV in the inner loop for hyperparameter tuning. Table 1 shows the detailed performance of each candidate regression algorithm. It is seen that the Adaboost regressor outperformed other regression models for B point detection in terms of the lower mean absolute error (MAE) of 3.5 ms, lower median absolute error (MedAE) of 0.0 ms, lower standard deviation of the error (SDE) of 9.2 ms, and higher Pearson correlation coefficient value ($r = 0.9$). Therefore, the proposed model for B point detection consisted of the Adaboost regressor and 465 features. Detailed results including MAE, MedAE, and hyperparameter set for each fold in the outer loop for all algorithms can be found in Appendix D.

The proposed model also demonstrated the superior performance over the state-of-the-art B point detection algorithm (for detailed analysis, see Section 4.1). It demonstrated the lowest MAE, MedAE, and SDE, as well as the highest Pearson correlation coefficient among the state-of the-art algorithms for both studies A and B combined (Table 2), and for each individual task of study A (Table 3) and study B (Table 4). In comparison with the expert-PEP, it maintained its superiority over the state-of-the-art algorithms as demonstrated by the Bonferroni corrected Wilcoxon signed-rank tests (Table 5) and Bland-Altman plot (Figure 3).

We also conducted a new set of experiments to analyze the effectiveness of the proposed variable search-window by implementing the rule-based algorithms in this window. Table 6 shows that the rule-based algorithms performed better by using the proposed variable search-window instead of Arbol's search-window for the B point detection. The algorithms displayed a significant reduction in MAE and MedAE, and improvement in Pearson correlation coefficient. We also compared fixed and variable-search windows using the select cases where it was found that use of the fixed search-windows could result in placing a B point outside the normal-PEP physiological range for rule-based algorithms (Figure 4). Performance of the proposed algorithm with expert adjudication and state-of-the-art algorithms were also demonstrated for different morphologies using select cases (Figure 5, and Appendix E).

Feature importance analysis revealed that handcrafted high-level features were most important for B point detection with consistently high feature scores (for detailed analysis, see Section 4.2). Figure 6 presents the feature scores to visualize the relative feature-importance for all features. Moreover, Table 7 presents the feature scores for important high-level features used for B point detection.

4. Discussion

In this paper, we developed an automated algorithm for the detection of aortic valve opening in ICG signal which was robust to noise and intra/inter-subject ICG morphological variability. The proposed algorithm, which was based on the Adaboost regressor and 465 features, outperformed the state-of-the-art algorithms with a significant margin (see section 4.1). It also significantly decreased the human intervention requirement for B point annotation on the ICG cycle and resulted in a reproducible, automated PEP computation that is critical for different clinical and psychophysiological studies.

4.1 Comparative Evaluation of the Proposed Algorithm with other State-of-the-Art Methods

The proposed model was compared with five state-of-the-art algorithms: three popular rule-based algorithms (zero crossing, the second derivative rule, and the third derivative rule (Arbol et al., 2017) and two data-driven algorithms (linear and quadratic regression estimates (Lozano et al., 2007). The proposed algorithm demonstrated the lowest MAE, MedAE, SDE, and highest Pearson correlation coefficient than those five B point detection algorithms (Table 2).

Performance of the proposed and the state-of the-art algorithms were also compared for different tasks of study A and study B. Table 3 depicts the performance for neutral and stress tasks of study A. Table 4 represents the performance for the baseline, preparation to speech, speech, recovery after tcVNS, and arithmetic tasks of the study B. Again, the proposed algorithm outperformed all other algorithms for the individual tasks.

From Tables 3 and 4, we observe that the performances of all algorithms (including our proposed one) are slightly different for different age groups (study A vs. B), script-driven imagery tasks (neutral vs.

stress), and emotional/cognitive tasks (speech, arithmetic, etc.). Thus, the overall performance of the proposed algorithm depends on the prevalence of these tasks in the training and testing dataset, which can be considered one of the limitations of this (and all other) data-driven algorithms. Moreover, the Bonferroni corrected Wilcoxon signed-rank tests showed that the PEP computed using our proposed algorithm was not statistically different from the expert-PEP, which was not true for other algorithms (Table 5). In addition, the Bland-Altman analysis (Figure 3) showed that our proposed algorithm had narrower limits of agreement [-18.4, 17.7] ms with the expert-PEP. Figure 3(a) also depicts that the outliers for the proposed model were mostly identified at the boundaries of the PEP physiological range. Discussion on an outlier case, marked with a red dot, can be found in Section 4.4.

The proposed model outperformed the state-of-the-art algorithms by addressing their limitations. The rule-based algorithms (i.e., zero-crossing, second and third derivative rules) focused on a particular morphology or condition in different empirically designed fixed search-windows (Arbol et al., 2017). The zero-crossing method was sensitive to the ICG signal drift and did not consider the morphological variation of the ICG signal. Individually, the second derivative rule and the third derivative rule could not detect all morphological shapes of the B point. In contrast, the B point could manifest itself in at least six different shapes at different locations on an ICG waveform (Figure 1).

It was also observed that using the fixed search-windows could result in placing a B point outside the normal-PEP physiological range for rule-based algorithms. This is so because Arbol's search-windows were based on empirical data without using the R peak which resulted in the erroneous B point detection. In addition, Arbol's search-windows also failed to accurately detect the B point on ICG waveform with notched and M-shaped C waves. Figure 4 presents two cases where these fixed search-windows failed due to the presence of notched or M-shaped C waves. In Figure 4 (a), the notched C wave resulted in the morphology of the “Plateau” near the C point which can be annotated as the B point by any automated algorithm using Arbol's search-window for the third derivative rule. In Figure 4 (b), the M-shaped C wave resulted in the morphology of a “notch” near the C point, which can be erroneously marked as the B point

by any automated algorithm being the reversal point close to the highest peak. Also, in both cases, the B point was outside the Arbol search-window for the second-derivative rule.

The data-driven Lozano's linear and quadratic estimates were dependent upon RC interval only. These approaches used a smaller dataset (33 subjects) lacking morphological variability of B point. Also, the averaged beats were generated using the conventional EA method that did not address the noise artifacts (Lozano et al., 2007).

Our proposed approach addressed these issues by (1) extracting the averaged beats using a novel noise removal three-stage EA algorithm (Sheikh, Shah, Levantsevych, et al., 2020), (2) designing time-series and high-level features (3) designing a variable PEP-based physiologically valid search-window for extraction of morphological features, and (4) labeling the reference B point to include all its morphological manifestations. Figure 5 depicts the comparison of the performance of the proposed algorithm with expert adjudication and state-of-the-art algorithms for "notch" morphology. Performance analysis for other shapes can be found in Appendix E.

4.2 Important/Relevant Features

The feature importance analysis is depicted in Figure 6 which represents the feature score vs. feature indices for the proposed regressor designed with 465 features. As depicted, the handcrafted high-level features were the most notable features for B point detection with consistently high feature scores. The morphological features designed in the proposed search-window proved to be the most relevant (Table 7). Among the critical hemodynamic parameters, the maximum ejection velocity of blood in the aorta proved its importance with a feature score of 0.71. Figure 6 also revealed that the time-series features with indices 1-100 were not significant for computing the B point location. The relevant time-series features with "feature score >0.6" are feature indices 104-110 (extracted from the ICG waveform), 249-255 and 282-298 (extracted from the first derivative of the ICG), and 365-385 (extracted from the second derivative of the ICG). As we can observe from Figure 6, the relevant time-series features are either in the B point location's

neighborhood or close to the right boundary of the search window. However, it is worth noting that the B point location in the ICG waveform presented in this figure is only an example and other ICG waveforms may have B point in slightly different locations.

4.3 Sources of Error

The proposed algorithm had some limitations. Most of the outliers/disagreements with the experts' annotations were concentrated at the lower and upper physiological limits of PEP (Bland-Altman plot - Figure 3(a)). To further study the disagreement, the cumulative distribution functions of the PEP absolute errors for the proposed algorithm and state-of-the-art methods were studied (Figure 7). The proposed method outperformed all other algorithms, but 15% of the data had a mean absolute error greater than 5 ms. On investigation, the performance limitation of the proposed algorithm over the irregularly notched ICG cycles was identified. The proposed algorithm failed for ICG cycles with highly disproportionate heights of peaks in M-shaped C waves. Figure 8 represents two cases from the same subject during the speaking task with notched C waves. In the first case, Figures 8(a) and 8(b), the proposed algorithm successfully identified the B point. However, for the second case, Figures 8(c) and 8(d), there was a difference of 67 ms between the expert-PEP and the PEP value computed via the proposed algorithm (the same point marked as a red dot in Figure 3(a)). It was also observed that most of the outliers resulting from such disproportionate C waves belong to the same participants. Another performance issue was mild or insignificant inflection on the ICG waveform. Such morphological features were invisible to the human eye and annotated as featureless beats. The B point in such cases was located using the third derivative rule. However, in the proposed (or any automated mathematical algorithm), the 'mild inflection' is detected mathematically resulting in an error. Since this aspect is related to the annotators therefore it can be considered as a limitation of this study.

4.4 Future Work

Future work may attempt to resolve the performance issues highlighted in Section 4.3. For incorrect detections at lower and upper PEP physiological limits, the search-window for morphological features may be designed using the actual Q point instead of a QR offset of 35 ms. The issues arising due to highly irregular and disproportionate C waves can be addressed by increasing the number of such M-shaped ICG cycles for training/testing of the regression model.

Based on the results depicted in Tables 3 and 4, it is evident that the overall results for the proposed algorithm were also influenced by the combination of study samples from different age groups undergoing different tasks (neutral, stress, speech, arithmetic, etc.). Keeping in view the age-related differences in morphology of ICG waveforms (Ermishkin et al., 2014; Tronstad et al., 2019) and having different performances for different tasks, future work may investigate their contributions to the B point detection models.

Moreover, among 189 subjects in the dataset of this study, there are 48 twin pairs, which may or may not affect the evaluation of the algorithmic performance. Measuring the impact of the twin pairs on the performance of the B point detection algorithms is out of the scope of this work and needs an independent study. However, considering the intra- and inter-subject ICG morphological variability, we conjecture that the risk of twin-related data leakage between training and testing sets is marginal. Nevertheless, since the training and testing datasets are similar for all algorithms, their respective performance and rank among other algorithms are still valid.

Apart from the performance issues, the algorithm was designed on the dataset recorded in the sitting posture only, future work may also extend to different posture positions as well as real-time algorithms for detection of the B point. The proposed algorithm can also be adapted for the long-term ambulatory monitoring and derivation of other B point-dependent ICG parameters, including systolic time ratio, left ventricular ejection time, and stroke volume.

5. Conclusion

In this work, we presented an automated data-driven regression algorithm to detect aortic valve opening (B point on an ICG signal) for robust PEP estimation. The proposed method outperformed state-of-the-art algorithms. Future work may focus on modification of the proposed search-window and new high-level features to improve performance and adapting the proposed method for the long-term ambulatory monitoring and derivation of other B point-dependent ICG parameters. The codes have been provided in the open-source toolbox for ICG analysis (Sheikh, Shah, Inan, and Clifford, 2020) to act as a benchmark for other studies.

Appendix A. Rule Based and Data Driven Approaches for B point Detection

Initially, zero-crossing of ICG (B_{zero}) and 15% of $(dZ/dt)_{max}$ ($B_{15\%}$) were identified as B point (Kubicek et al.,1966; Kubicek et al., 1970). But the respiratory influences causing baseline shifts render these methods ineffective. It was also suggested to look for the morphological feature of "notch" to mark the B point (Lababidi et al., 1970); however, there were instances where the 'notch' was not present in the ICG waveform (see Figure 1). The B point was also recommended to be detected as the initiation of the rapid upstroke of ICG as it ascends towards its maximum point $(dZ/dt)_{max}$ (C point) (Sherwood et al., 1990), which led the researchers to focus on the mathematical rules for B point detection. The second-derivative rule declared the maximum of the second derivative of impedance signal, $(d^2Z/dt^2)_{max}$, before the C point as the event of aortic valve opening (Debski et al., 1991). Later, the local minimum of the second derivative of the impedance signal $(d^2Z/dt^2)_{min,}$ along with the maximum of the third derivative of impedance signal $(d^3Z/dt^3)_{max}$ (third-derivative rule) were also suggested for B point detection (Debski et al., 1993). The B point detection algorithm based on comparative threshold analysis using the second derivative of impedance signal (d^2Z/dt^2) was also proposed (Demarzo et al., 1996). The inflection point on the ICG waveform was also identified to represent aortic valve opening (Ono et al., 2004). However, these single-rule algorithms failed to encompass the variations present in the B point morphology.

To improve the B point detection accuracy, the researchers focused on rule-based approaches in the fixed search-windows on the ICG waveform. Nederend et al., (2017) proposed a search-window of 150 ms after R peak to detect B point as maximum upstroke to the C point. The lowest point in the first difference of ICG signal in the search-window of [20, 65] % of RC interval was marked as B point by Bagal et al., (2017). Nabian et al., (2018) recommended the second-derivative rule for B point detection in 80 ms time-window before the C point. Using different search-windows, Arbol et al., (2017) compared B point detection between B_{zero} and derivative-based approaches. For the second-derivative rule, a search-window of 50 ms starting 150 ms before the C point on ICG was used. The third-derivative rule was applied in a search-window of 300 ms before the C point. In comparison with the expert visual inspection, maximum

accuracy of 78% was achieved by the third-derivative rule. The concept of fixed search-windows was modified to search for monotonically increasing segments for B point detection (Forouzanfar et al., 2018). In that algorithm, B point was detected on monotonically increasing segments between the local minimum (A point) within one-third of the beat-to-beat interval before the C point on an ICG waveform using mathematical rules. A high mean absolute error (MAE) of 12 ms was achieved by this technique. The search-window interval/segment approach can suffer from an incorrect B point detection in case the interval/segment falls outside the normal PEP physiological range of 70 - 175 ms (Arbol et al., 2017). Details with examples are discussed in Section 4.

Data-driven approaches applied for automatic detection of the B point include linear, quadratic, and Adaboost regression estimates (Lozano et al., 2007; Cieslak et al.,2018). Based on the ensemble-averaged (EA) beats from a small data set of 33 young-age subjects, linear and quadratic relationships between RB and initial systolic time interval (RC) were recommended (Lozano et al., 2007). The conventional EA method, without any prior noise rejection, was used for beat averaging. Also, Van et al., (2013) found discrepancies in its implementation under different conditions and discouraged its use as a substitute for the actual PEP obtained by the aortic valve opening detection. The Adaboost-based estimate used time-series features from EA beats of 37 subjects (Cieslak et al., 2018). The beat averaging technique (moving EA method) did not address the noise artifacts due to systematic errors. Also, the reference B point was labeled in the RC interval using the third-derivative rule ignoring the morphological variations of the B point. The accurate detection of B point on EA beats remained challenging due to these limitations (Forouzanfar et al., 2018).

Appendix B. Feature Vector for Regression Model

A total of 465 features were extracted for designing the regression model. Feature vector was formed by concatenating the ICG waveform, its first- and second-order derivatives, and high-level features. The concatenated ICG waveform, its first- and second-order derivatives will be collectively referred to as "time-series features" in this work. The high-level features consisted of morphological features and hemodynamic parameters. The morphological features were extracted from ICG waveform in the proposed search-window. The hemodynamic parameters include the initial systolic time interval (RC interval) and maximum ejection velocity of blood in the aorta (C point). These time-series features, and high-level features are listed in Table B1.

Appendix C. Training, Validation and Hyperparameter settings for Candidate Regression Models

In this appendix, the averaged beats distribution for training and validation in outer CV loop (10-fold) and hyperparameter sets to tune the model using inner 5-fold grid-search CV are presented in Tables C1 and C2, respectively.

Appendix D. Performance Analysis for Candidate Regression Models

Table D1 and Table D2 present the MAE and MedAE for validation folds in respect of all candidate regression models. Table D3 presents the best hyperparameter values for each candidate algorithm for 10 folds.

Appendix E. Performance Comparison of the B point Detection using Different Algorithms

In this appendix, performance comparison of B point detection algorithms for different morphologies is presented.

References

Árbol, J. R., Perakakis, P., Garrido, A., Mata, J. L., Fernandez-Santaella, M. C. & Vila, J. (2017). Mathematical detection of aortic valve opening (B point) in impedance cardiography: A comparison of three popular algorithms. *Psychophysiology,* 54(3), 350-357. https://doi.org/10.1111/psyp.12799

Bagal, U. R., Pandey, P. C., Naidu, S.M.M., & Hardas, S. P. (2017). Detection of opening and closing of the aortic valve using impedance cardiography and its validation by echocardiography. *Biomedical Physics & Engineering Express,* 4(1), 015012. https://doi.org/10.1088/2057-1976/aa8bf5

Behar, J., Johnson, A., Clifford, G. D., & Oster, J. (2014). A comparison of single channel fetal ECG extraction methods. *Annals of Biomedical Engineering*, 42(6), 1340-1353. https://doi.org/10.1007/s10439-014-0993-9

Bland, J. M. & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327(8476), 307-310. https://doi.org/10.1016/S0140-6736(86)90837-8

Bremner, J. D. (2016). Closed loop vagal nerve stimulation for patients with posttraumatic stress disorder. https://clinicaltrials.gov/ct2/show/NCT02992899 Accessed on: Oct 7, 2020.

Breiman, L., Friedman, J.H., Olshen, R.A., & Stone, C.J. (1984). Classification And Regression Trees (1st ed.). Routledge. https://doi.org/10.1201/9781315139470

Breiman, (2001). Random Forests. *Machine Learning*, 45(1), 5-32. https://doi.org/10.1023/A:1010933404324

Cieslak, M., Ryan, W. S., Babenko, V., Erro, H., Rathbun, Z. M., Meiring, W., ... and Grafton, S. T. (2018). Quantifying rapid changes in cardiovascular state with a moving ensemble average. *Psychophysiology*, 55(4), e13018. https://doi.org/10.1111/psyp.13018

Cohen, J. R., Thomsen, K. N., Tu, K. M., Thakur, H., McNeil, S. & Menon, S. V. (2020). Cardiac autonomic functioning and post-traumatic stress: A preliminary study in youth at-risk for PTSD. *Psychiatry Research*. 284: 112684. https://doi.org/10.1016/j.psychres.2019.112684

Czarnek, G., Richter, M., & Strojny, P. (2021). Cardiac sympathetic activity during recovery as an indicator of sympathetic activity during task performance, *Psychophysiology*, 58(2), e13724. https://doi.org/10.1111/psyp.13724

Debski, T. T., Kamarck, T. W., Jennings, J. R., Young, L. W., Eddy, M. J. and Zhang,Y. (1991). A computerized test battery for the assessment of cardiovascular reactivity. *International Journal of Biomedical Computing*, 27(3-4), 277–289. https://doi.org/10.1016/0020-7101(91)90068-P

Debski, T. T., Zhang, Y., Jennings, J. R. & Kamarck, T. W. (1993). Stability of cardiac impedance measures: aortic opening (B point) detection and scoring. *Biological Psychology*, 36(1-2), 63-74. https://doi.org/10.1016/0301-0511(93)90081-I

DeMarzo, A. P. (2020). Clinical use of impedance cardiography for hemodynamic assessment of early cardiovascular disease and management of hypertension. *High Blood Pressure & Cardiovascular Prevention*, 27(3), 203-213. https://doi.org/10.1007/s40292-020-00383-0

DeMarzo, A. P., & Lang, R.M. (1996). A new algorithm for improved detection of aortic valve opening by impedance cardiography. In *Computers in Cardiology* 1996, (pp. 373–376). IEEE. https://doi.org/10.1109/CIC.1996.542551

Drucker, H. (1997). Improving regressors using boosting techniques. In *ICML* (Vol. 97, pp. 107–115).

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1996). Support vector regression machines. *Advances in Neural Information Processing Systems*, 9, 155-161.

Ermishkin, V. V., Kolesnikov, V. A., & Lukoshkova, E. V. (2014). Age-dependent and 'pathologic'changes in ICG waveforms resulting from superposition of pre-ejection and ejection waves. *Physiological Measurement*, *35*(6), 943. https:// 10.1088/0967-3334/35/6/943

Forouzanfar, M., Baker, F. C., Colrain, I. M., Goldstone, A. & de Zambotti, M. (2019). Automatic analysis of pre-ejection period during sleep using impedance cardiogram. *Psychophysiology*, 56(7), e13355. https://doi.org/10.1111/psyp.13355

Forouzanfar, M., Baker, F. C., de Zambotti, M., McCall, C., Giovangrandi, L. & Kovacs, G. T. (2018). Toward a better noninvasive assessment of preejection period: A novel automatic algorithm for B-point detection and correction on thoracic impedance cardiogram. *Psychophysiology*, 55(8), e13072. https://doi.org/10.1111/psyp.13072

Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1), 119-139. https://doi.org/10.1006/jcss.1997.1504

Hastie, T., Friedman, J., Tibshirani, R. (2001). *The Elements of Statistical Learning* (Vol. 1, No. 10). New York: Springer series in Statistics. https://doi.org/10.1007/978-0-387-21606-5_1

Gustafsson, F. (1996). Determining the initial states in forward-backward filtering. *IEEE Transactions on Signal Processing*, *44*(4), 988-992. https://doi.org/10.1109/78.492552.

Gurel, N. Z., Huang, M., Wittbrodt, M. T., Jung, H., Ladd, S. L., Shandhi, M. M. H., ... & Inan, O. T. (2020). Quantifying acute physiological biomarkers of transcutaneous cervical vagal nerve stimulation in the context of psychological stress. *Brain Stimulation*, 13(1), 47-59. https://doi.org/10.1016/j.brs.2019.08.002

Johnson, A. E., Behar, J., Andreotti, F., Clifford, G. D., & Oster, J. (2014). R-peak estimation using multimodal lead switching. In *Computing in Cardiology* 2014 (pp. 281-284). IEEE.

Kelsey, R. M., Alpert, B. S., Dahmer, M. K., Krushkal, J. & Quasney, M. W. (2010). Beta-adrenergic receptor gene polymorphisms and cardiovascular reactivity to stress in black adolescents and young adults. *Psychophysiology,* 47(5), 863-873. https://doi.org/10.1111/j.1469-8986.2010.01006.x

Kelsey, R. M., Ornduff, S. R. and Alpert, B. S. (2007). Reliability of cardiovascular reactivity to stress: internal consistency. *Psychophysiology*, 44(2), 216-225. https://doi.org/10.1111/j.1469-8986.2007.00499.x

Kibler, J. L. (2018). An extension of the perseverative cognition hypothesis to posttraumatic stress disorder symptomatology: Cardiovascular recovery in relation to posttraumatic stress disorder severity and cognitive appraisals of stress. *Journal of Traumatic Stress* 31(1), 25-34. https://doi.org/10.1002/jts.22252

Krstajic, D., Buturovic, L. J., Leahy, D. E., & Thomas, S. (2014). Cross-validation pitfalls when selecting and assessing regression and classification models. *Journal of Cheminformatics*, 6(1), 1-15. https://doi.org/10.1186/1758-2946-6-10

Kubicek, W. G. (1966). Development and evaluation of an impedance cardiac output system. *Aerosp Med*, 37, 1208–1212.

Kubicek, W., Patterson, R., & Witsoe, D. (1970). Impedance cardiography as a noninvasive method of monitoring cardiac function and other parameters of the cardiovascular system. *Annals of the New York Academy of Sciences*, 170(2), 724–732. https://doi.org/10.1111/j.1749-6632.1970.tb17735.x

Lababidi, Z., Ehmke, D., Durnin, R. E., Leaverton, P. E., & Lauer, R. M. (1970). The first derivative thoracic impedance cardiogram. *Circulation,* 41(4), 651–658. https://doi.org/10.1161/01.CIR.41.4.651

Lilliefors, H. W. (1969). On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. *Journal of the American Statistical Association*, 64(325), 387-389.

Lozano, D. L., Norman, G., Knox, D., Wood, B. L., Miller, B. D., Emery, C. F., & Berntson, G. G. (2007). Where to B in dz/dt. *Psychophysiology*, 44(1), 113–119. https://doi.org/10.1111/j.1469-8986.2006.00468.x

Malcolm, L., Kibler, J. L., Ma, M., Tursich, M., Augustin, D., Greenbarg, R., & Gold, S. N. (2016). Psychophysiological reactivity and PTSD symptom severity among young women. *International Journal of Psychology and Neuroscience*, 2(3): 17.

Medina-Lezama, J., Narvaez-Guerra, O., Herrera-Enriquez, K., Morey-Vargas, O. L., Bolaños-Salazar, J. F., Abugattas, J. P., ... & Chirinos, J. A. (2018). Hemodynamic patterns identified by impedance cardiography predict mortality in the general population: the PREVENCION study. *Journal of the American Heart Association*, 7(18), e009259. https://doi.org/10.1161/JAHA.118.009259

Meijer, J. H., Elbertse, E., Boesveldt, S., Berendse, H. W. & Verdaasdonk, R. M. (2011). Using the initial systolic time interval to assess cardiac autonomic nervous function in Parkinson's disease. *Journal of Electrical Bioimpedance*, 2(1), 98-101. https://doi.org/10.5617/jeb.216

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT Press.

Nabian, M., Yin, Y., Wormwood, J., Quigley, K. S., Barrett, L. F., & Ostadabbas, S. (2018). An open-source feature extraction tool for the analysis of peripheral physiological data. *IEEE Journal of Translational Engineering in Health and Medicine*, 6, 1–11. https://doi.org/10.1109/JTEHM.2018.2878000

Nederend, I., Ten Harkel, A. D., Blom, N. A., Berntson, G. G., & de Geus, E. J. (2017). Impedance cardiography in healthy children and children with congenital heart disease: Improving stroke volume assessment. *International Journal of Psychophysiology*, 120, 136-147. https://doi.org/10.1016/j.ijpsycho.2017.07.015

Newlin, D. B., & Levenson, R. W. (1979). Pre-ejection period: Measuring beta adrenergic inuences upon the heart. *Psychophysiology* 16(6), 546-552. https://doi.org/10.1111/j.1469-8986.1979.tb01519.x

Ono, T., Miyamura, M., Yasuda, Y., Ito, T., Saito, T., Ishiguro, T., … & Yambe, T. (2004). Beat-to-beat evaluation of systolic time intervals during bicycle exercise using impedance cardiography. *The Tohoku Journal of Experimental Medicine*, 203(1), 17–29. https://doi.org/10.1620/tjem.203.17

Peters, J. R., Eisenlohr-Moul, T. A., Walsh, E. C. & Derefinko, K. J. (2018). Exploring the pathophysiology of emotion-based impulsivity: The roles of the sympathetic nervous system and hostile reactivity, *Psychiatry Research*, 267, 368-375. https://doi.org/10.1177/1077559514538115

Seery, M. D., Kondrak, C. L., Streamer, L., Saltsman, T., & Lamarche, V. M. (2016). Preejection period can be calculated using R peak instead of Q. *Psychophysiology* 53(8), 1232-1240. https://doi.org/10.1111/psyp.12657

Sheikh, S. A., Shah, A. J., Levantsevych, O., Soudan, M., l Alkhalaf, M., Rad, A. B., … & Clifford, G. D. (2020). An open-source automated algorithm for removal of noisy beats for accurate impedance cardiogram analysis, *Physiological Measurement*, 41(7), 075002. https://doi.org/10.1088/1361-6579/ab9b71

Sheikh, S. A., Shah, A. J., Inan, O. T., and Clifford, G. D. (2020). Open-Source Toolbox for ICG Analysis. v1.0 (https://github.com/cliffordlab/ICG_OSToolbox/ B point Detection Accessed: 20 Nov 2021)

Sheikh, S. A. A., Gurel, N. Z., Gupta, S., Chukwu, I. V., Levantsevych, O., Alkhalaf, M., ... & Shah, A. J. (2022). Validation of a new impedance cardiography analysis algorithm for clinical classification of stress states. *Psychophysiology*, e14013. https://doi.org/10.1111/psyp.14013

Sherwood, A., Allen, M. T., Fahrenberg, J., Kelsey, R. M., Lovallo, W. R., & Van Doornen, L. J. (1990). Methodological guidelines for impedance cardiography. *Psychophysiology* 27(1), 1-23. https://doi.org/10.1111/j.1469-8986.1990.tb02171.x

Silva Lopes, B., Craveiro, N., Firmino-Machado, J., Ribeiro, P. & Castelo-Branco, M. (2019). Hemodynamic differences among hypertensive patients with and without heart failure using impedance

cardiography. *Therapeutic Advances in Cardiovascular Disease*, 13: 1753944719876517. https://doi.org/10.1177/1753944719876517

Silvia, P. J., McHone, A. N., Mironovov_a, Z., Eddington, K. M., Harper, K. L., Sperry, S. H., & Kwapil, T. R. (2020). RZ interval as an impedance cardiography indicator of effort-related cardiac sympathetic activity. *Applied Psychophysiology and Biofeedback*, 46(1), 83-90. https://doi.org/10.1007/s10484-020-09493-w

Sornmo, L., & Laguna, P. (2005). Bioelectrical signal processing in cardiac and neurological applications. Vol. 8, Academic Press. https://doi.org/10.1016/B978-012437552-9/50003-9

Tronstad, C., Høgetveit, J. O., Elvebakk, O., & Kalvøy, H. (2019). Age-related differences in the morphology of the impedance cardiography signal. *Journal of Electrical Bioimpedance*, *10*(1), 139-145. https://doi.org/10.2478/joeb-2019-0020

Vaccarino, V. (2015). PTSD and Ischemic Heart Disease Progression: A Longitudinal Twin Study. https://grantome.com/grant/NIH/R01-HL125246-01A1 Accessed on Jul 13, 2021

Van Lien, R., Schutte, N. M., Meijer, J. H., & de Geus, E. J. (2013). Estimated preejection period (PEP) based on the detection of the R-wave and dz/dt-min peaks does not adequately reflect the actual PEP across a wide range of laboratory and ambulatory conditions. *International Journal of Psychophysiology*, 87(1), 60-69. https://doi.org/10.1016/j.ijpsycho.2012.11.001

Vest, A. N., Da Poian, G., Li, Q., Liu, C., Nemati, S., Shah, A. J., & Clifford, G. D. (2018). An open-source benchmarked toolbox for cardiovascular waveform and interval analysis. *Physiological Measurement*, 39(10), 105004. https://doi.org/10.1088/1361-6579/aae021

Acknowledgement

Shafa-at Ali Sheikh is funded by Fulbright Scholarship Program. The authors wish to acknowledge the National Institutes of Health (Grant # NIH K23HL127251, R01HL136205, R01HL125246, R03HL146879, and R01AG026255), the National Science Foundation Award 1636933, and Emory University for their financial support of this research. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Institutes of Health, the National Science Foundation, Georgia Institute of Technology and Emory University.

Table 1: Performance metrics for the candidate data-driven models with designed feature vector. Linear regression (OLS, LASSO, Ridge), kernel ridge regression, support vector regression, random forest, and Adaboost models were applied to detect the B point for PEP computation. Performance metrics were computed in comparison to the expert-PEP. Adaboost outperformed all other models as depicted by lower MAE, MedAE, SDE, and a higher Pearson correlation coefficient. We evaluated the mean error (ME) with 95% CI and median error (MedE) to estimate the bias of regression algorithm.

Data-driven models	Performance metrics over validation folds							
	MAE (ms)	95% CI of absolute errors (ms)	MedAE (ms)	SDE (ms)	Pearson correlation coefficient	ME (ms)	95% CI of errors (ms)	MedE (ms)
Linear regression (OLS)	8.9	[8.5, 9.2]	5.8	13.7	0.78	-0.36	[-0.81, 0.08]	0.0
Linear regression (LASSO)	8.3	[7.5, 9.1]	6.1	11.5	0.85	-0.61	[-1.78, 0.55]	0.0
Linear regression (Ridge)	8.5	[7.6, 9.3]	6.6	11.6	0.84	-0.87	[-2.04, 0.31]	0.0
Kernel ridge regression	11.3	[10.2, 12.4]	9.1	15.5	0.71	-0.61	[-2.18, 0.96]	0.0
Support vector regression	6.9	[7.7, 6.0]	3.6	10.4	0.88	-2.27	[-3.32, -1.22]	0.0
Random forest	4.7	[4.4, 5.0]	0.78	10.4	0.88	0.06	[-0.28, 0.40]	0.05
Adaboost regressor	3.5	[3.3, 3.8]	0.0	9.2	0.90	-0.38	[-0.68, 0.09]	0.0

Table 2: Performance metrics for the proposed and state-of-the-art algorithms. The popular rule-based algorithms (Zero crossing, second derivative rule, and third derivative rule) were applied in the Arbol search-window for PEP measurement. Data-driven Lozano's linear and quadratic estimates were applied using RC interval. Performance metrics were computed in comparison to the expert-PEP. The proposed algorithm performed better than all other algorithms with lower MAE, MedAE, SDE, and higher Pearson correlation coefficient. We evaluated the mean error (ME) with 95% CI and median error (MedE) to estimate the bias of algorithms.

Algorithms	Performance metrics							
	MAE (ms)	95% CI of absolute errors (ms)	MedAE (ms)	SDE (ms)	Pearson correlation coefficient	ME (ms)	95% CI of errors (ms)	MedE (ms)
Proposed	3.5	[3.3, 3.8]	0.0	9.2	0.90	-0.38	[-0.68, 0.09]	0.0
Zero crossing	33.1	[32.3, 33.9]	28.0	27.0	0.37	31.5	[30.6, 32.4]	28
Second derivative rule	49.8	[49.0, 50.5]	47.0	26.6	0.41	47.9	[47.0, 48.7]	47
Third derivative rule	25.1	[24.1, 26.1]	15.0	39.6	0.46	0.19	[-1.09, 1.48]	-9
Lozano linear estimate	13.5	[13.1, 14.0]	8.5	18.1	0.54	-5.4	[-6.0, -4.8]	-0.7
Lozano quadratic estimate	15.0	[14.7, 15.4]	12.4	18.7	0.55	0.79	[0.18, 1.40]	5.4

Table 3: Performance metrics for the proposed and state-of-the-art algorithms for all tasks of the study A. The popular rule-based algorithms (Zero crossing, second derivative rule, and third derivative rule) were applied in the Arbol search-window for PEP measurement. Data-driven Lozano's linear and quadratic estimates were applied using RC interval. Performance metrics were computed in comparison to the expert-PEP. The proposed algorithm performed better than all other algorithms with lower MAE, MedAE, SDE, and higher Pearson correlation coefficient. We evaluated the mean error (ME) with 95% CI and median error (MedE) to estimate the bias of algorithms. The neutral and stress tasks resulted in relatively similar accuracy for all algorithms.

Algorithms	Performance metrics							
	MAE (ms)	95% CI of absolute errors (ms)	MedAE (ms)	SDE (ms)	Pearson correlation coefficient	ME (ms)	95% CI of errors (ms)	MedE (ms)
Study A: Neutral task								
Proposed	4.7	[4.0, 5.4]	1.0	9.5	0.87	-0.9	[-1.7, -0.1]	0.0
Zero crossing	35.4	[33.5, 37.2]	34.0	25.0	0.17	30.7	[28.4, 33.0]	33.0
Second derivative rule	38.0	[36.6, 39.4]	36.0	18.9	0.50	34.9	[33.1, 36.7]	36.0
Third derivative rule	32.4	[29.7, 35.0]	19.0	35.5	0.31	6.7	[3.1, 10.2]	-10
Lozano linear estimate	18.5	[17.4, 19.7]	13.2	15.4	0.53	-16.0	[-17.3, -14.7]	-11.9
Lozano quadratic estimate	16.9	[16.0, 17.9]	13.2	12.9	0.49	-9.6	[-11.0, -8.2]	-5.9
Study A: Stress task								
Proposed	4.0	[3.4, 4.6]	0.7	8.5	0.89	-0.8	[-1.5, -0.1]	0
Zero crossing	35.2	[33.4, 37.0]	33.0	25.3	0.13	30.6	[28.4, 32.8]	32

Second derivative rule	39.5	[38.1, 40.9]	37.0	20.1	0.49	37.0	[35.3, 38.7]	37
Third derivative rule	34.8	[32.2, 37.5]	20	37.7	0.29	11.8	[8.3, 15.3]	-1
Lozano linear estimate	18.5	[17.4, 19.5]	13.3	15.2	0.54	-16.2	[-17.4, -15.0]	-12.5
Lozano quadratic estimate	16.6	[15.7, 17.5]	12.0	13.3	0.50	-10.6	[-11.9, -9.3]	-6.5

Table 4: Performance metrics for the proposed and state-of-the-art algorithms for all tasks of the study B. The popular rule-based algorithms (Zero crossing, second derivative rule, and third derivative rule) were applied in the Arbol search-window for PEP measurement. Data-driven Lozano's linear and quadratic estimates were applied using RC interval. Performance metrics were computed in comparison to the expert-PEP. The proposed algorithm performed better than all other algorithms with lower MAE, MedAE, SDE, and higher Pearson correlation coefficient. We evaluated the mean error (ME) with 95% CI and median error (MedE) to estimate the bias of algorithms. All tasks yielded relatively similar accuracy for all algorithms.

Algorithms	Performance metrics							
	MAE (ms)	95% CI of absolute errors (ms)	MedAE (ms)	SDE (ms)	Pearson correlation coefficient	ME (ms)	95% CI of errors (ms)	MedE (ms)
Study B: Baseline task								
Proposed	3.1	[2.5, 3.7]	0	8.1	0.88	0.15	[-0.5, 0.8]	0
Zero crossing	31.1	[29.3, 32.9]	22	24.5	0.31	28.1	[26.0, 30.1]	21
Second derivative rule	57.7	[56.1, 59.3]	58	21.5	0.42	56.3	[54.5, 58.2]	58
Third derivative rule	20.1	[18.2, 22.0]	12	26.0	0.39	-7.0	[-9.3, -4.6]	-10
Lozano linear estimate	10.5	[9.7, 11.2]	7.4	10.1	0.62	3.0	[2.0, 4.1]	5.3
Lozano quadratic estimate	15.0	[14.3, 15.7]	14.4	9.6	0.67	10.1	[9.0, 11.1]	11.8
Study B: Preparation for speech task								
Proposed	3.8	[2.5, 5.1]	0	9.0	0.84	-0.1	[-1.5, 1.3]	0
Zero crossing	31.8	[28.6, 35.1]	25	22.5	0.36	29.0	[25.2, 32.8]	25

Second derivative rule	57.6	[54.3, 60.8]	57	22.3	0.47	56.4	[52.7, 60.0]	57
Third derivative rule	19.3	[15.8, 22.7]	11.5	23.6	0.44	-3.0	[-7.4, 1.4]	-7
Lozano linear estimate	10.1	[8.6, 11.5]	7.2	9.9	0.61	0.6	[-1.4, 2.7]	3.1
Lozano quadratic estimate	12.7	[11.4, 14.1]	10.9	9.3	0.64	6.3	[4.2, 8.4]	8.3
Study B: Speech task								
Proposed	2.9	[1.8, 4.1]	0	9.2	0.84	0.2	[-1.0, 1.4]	0
Zero crossing	31.8	[29.1, 34.5]	27	21.6	0.29	30.2	[27.2, 33.1]	27
Second derivative rule	57.2	[54.6, 59.8]	55	21.0	0.54	56.2	[53.3, 59.1]	55
Third derivative rule	20.9	[17.6, 24.2]	13	26.6	0.46	-1.2	[-5.4, 3.0]	-9
Lozano linear estimate	8.7	[7.5, 9.9]	6.1	9.7	0.67	0.6	[-1.1, 2.2]	3
Lozano quadratic estimate	11.5	[10.3, 12.6]	10.1	9.1	0.68	6.0	[4.4, 7.7]	7.5
Study B: Recovery after tcVNS task								
Proposed	2.8	[2.2, 3.3]	0	7.5	0.89	0.05	[-0.5, 0.6]	0
Zero crossing	31.1	[29.2,33.0]	21	25.7	0.25	28.4	[26.3, 30.5]	20
Second derivative rule	56.5	[54.9, 58.2]	54	22.1	0.38	55.1	[53.2, 57.0]	54
Third derivative rule	18.6	[17.0, 20.2]	11	22.0	0.35	-8.3	[-10.5, -6.3]	-9
Lozano linear estimate	10.2	[9,4, 10.9]	7.1	10.0	0.59	2.6	[1.6, 3.7]	4.9

Lozano quadratic estimate	14.2	[13.5, 14.8]	12.9	9.1	0.66	9.2	[8.2, 10.3]	10.7
Study B: Arithmetic task								
Proposed	2.7	[1.8, 3.6]	0	8.0	0.89	-0.9	[-1.8, 0.06]	0
Zero crossing	33.2	[30.2, 36.2]	28	26.7	0.25	30.7	[27.3, 34.0]	28
Second derivative rule	57.9	[55.4, 60.5]	58	22.5	0.48	56.9	[54.0, 59.7]	58
Third derivative rule	17.5	[15.2, 19.9]	12	21.1	0.53	-4.8	[-7.9, -1.8]	-8
Lozano linear estimate	10.2	[9.0, 11.4]	7.1	10.8	0.60	-0.5	[-2.1, 1.2]	3.2
Lozano quadratic estimate	12.9	[11.8, 13.9]	11.8	8.8	0.66	5.0	[3.3, 6.7]	8.4

Table 5: Mean, median, and SD of PEP values computed using the expert visual inspection and proposed/popular algorithms. Bonferroni corrected Wilcoxon signed-rank test p values for all algorithms with an adjusted alpha value of 0.0083 (0.05/6) are also shown in the last column. The PEP values computed using the proposed Adaboost regressor were not statistically different from the expert-PEP. All other algorithms generated PEP values that are statistically different from the expert-PEP value.

Algorithms	Mean (ms)	Median (ms)	SD (ms)	Wilcoxon Sign Rank (p value)
Expert visual inspection	79.9	84.0	21.5	-
Proposed	80.3	84.0	19.3	0.3
Zero crossing	54.7	48.0	34.2	$\leq$0.0001
Second-derivative rule	32.0	32	26.8	0.0
Third-derivative rule	79.7	94	44.7	$\leq$0.0001
Lozano linear estimate ($\underline{\Delta t_{RB}} = 0.55*(t_C - t_R) + 4.45$)	85.3	85.3	11.6	$\leq$0.0001
Lozano quadratic estimate ($\underline{\Delta t_{RB}} = 1.233*(t_C - t_R) - 0.0032*(t_C - t_R)^2 - 31.59$)	79.1	80.5	6.6	$\leq$0.0001

Table 6: Performance metrics for the three popular rule-based B point detection algorithms using Arbol and proposed search-windows. The algorithms were applied on EA ICG beats from 189 subjects and compared with the expert-PEP. All popular rule-based B point detection algorithms performed better with the proposed physiologically valid search-window with a lower MAE/MedAE, and a higher Pearson correlation coefficient.

Algorithm	MAE (ms)		MedAE (ms)		Pearson correlation coefficient	
	Search-window		Search-window		Search-window	
	Arbol	Proposed	Arbol	Proposed	Arbol	Proposed
Zero-crossing	33.1	24.4	28.0	19.0	0.37	0.41
Second derivative rule	49.8	30.7	47.0	29.0	0.41	0.60
Third derivative rule	25.1	15.7	15.0	12.0	0.46	0.65

Table 7: Feature scores for important high-level features used for B point detection. A feature score of 1.0 was obtained mostly by the morphological features designed in the proposed search-window. The C point amplitude on ICG associated with maximum ejection velocity of blood in aorta also proved its importance by obtaining a feature score of 0.71.

High-level feature description	Feature Score
ICG amplitude in search-window for last inflection point	1.0
ICG amplitude difference: C point minus last inflection point	1.0
Time interval between initial systolic time interval (t_{RC}) and last inflection point in search-window: t_{RC} - t_{LI}	1.0
Time index in search-window for last notch (t_{LN})	1.0
Time index in search-window for last inflection (t_{LI})	1.0
Time index in search-window for maximum of the second-order derivative of ICG (d^3Z/dt^3)	1.0
ICG amplitude in search-window for maximum of second-order derivative of ICG	0.95
Time interval between initial systolic time interval (t_{RC}) and maximum of the second-order derivative of ICG (d^3Z/dt^3) (t_{3Dmax}) in search-window: t_{RC} - t_{3Dmax}	0.76
C point amplitude on ICG associated with the maximum ejection velocity of blood in aorta	0.71

Table B1: Feature vector formed using time-series features and high-level features along with indices for designing regression model. The time-series features consist of the EA ICG signal, its first and second-order derivative in the interval [t_R, t_R+150] ms. High-level features are based on the morphology of the ICG waveform (notch, inflection point, maximum of the second-order derivative of ICG cycle) and critical hemodynamic parameters (initial systolic time interval (RC), and maximum ejection velocity of the blood in the aorta (C point)). The morphological features were extracted from the proposed search-window.

Features	**Feature Index**
Time-Series features	
ICG waveform (dZ/dt) in the interval [t_R, t_R+150] ms	1-151
First-order derivative of ICG waveform (d^2Z/dt^2) in the interval [t_R, t_R+150] ms	152-301
Second-order derivative of ICG waveform (d^3Z/dt^3) in the interval [t_R, t_R+150] ms	302-450
High-Level Features	
ICG amplitude in the search-window	
Last notch	451
Last inflection point	452
Maximum of second-order derivative of ICG (d^3Z/dt^3)	453
C point amplitude on ICG (associated with the maximum ejection velocity of blood in aorta)	454
ICG amplitude difference	
C point minus last notch	455
C point minus last inflection point	456
C point minus maximum of second-order derivative of ICG (d^3Z/dt^3)	457
Time intervals on ICG Cycle	
Time interval between last notch (LN) and last inflection point (LI) in search-window: $t_{LN} - t_{LI}$	458

Time interval between initial systolic time interval (t_{RC}) and	
Last notch in search-window: t_{RC} - t_{LN}	459
Last inflection in search-window: t_{RC} - t_{LI}	460
Maximum of the second-order derivative of ICG (d^3Z/dt^3) (t_{3Dmax}) in search-window: t_{RC} - t_{3Dmax}	461
Initial systolic time interval (RC interval)	462
Time index in search-window	
Last notch (t_{LN})	463
Last inflection (t_{LI})	464
Maximum of the second-order derivative of ICG (d^3Z/dt^3)	465

Table C1: Number of averaged beats in training and validation folds used to train and test the regression models in the 10-fold outer CV. Data were partitioned subject-wise.

Fold number	Number of averaged beats in 10-fold CV	
	Training fold	Validation fold
1	3314	343
2	3282	375
3	3312	345
4	3266	391
5	3349	308
6	3267	390
7	3312	345
8	3226	431
9	3305	352
10	3280	377

Table C2: Hyperparameter sets for tuning candidate data-driven models via 5-fold nested grid-search CV. The variables ‘lin’, ‘poly’, and ‘rbf’ stands for linear, polynomial, and radial basis function, respectively.

Regressor Model	Hyperparameter Set
Linear regression (LASSO)	alpha = [0.001, 0.01, 0.1, 0.5, 1, 2, 10]
Linear regression (Ridge)	alpha = [0.001, 0.01, 0.1, 0.5, 1, 2, 10]
Kernel ridge regression	alpha = [0.001, 0.01, 0.1, 0.5, 1, 2, 10] kernel = ['lin', 'poly', 'rbf']
Support vector regression	C = [0.001, 0.1, 0.5, 1, 2, 5, 10] kernel = ['lin', 'poly', 'rbf']
Random forest	Number of estimators = 500 max features = [0.1, 0.3, 0.5, 0.7] max depth = [none, 5, 7, 10, 12, 15, 17, 20, 22]
Adaboost	Number of estimators = 500 loss = [‘exponential’] learning rate = [0.01, 0.05, 0.1,0.5,1] base estimator =Decision tree regressor max features = [0.1, 0.3, 0.5, 0.7] max depth = [none, 5, 7, 10, 12, 15, 17, 20, 22]

Table D1: MAE for the candidate data-driven models in the validation folds. Adaboost regressor produced the best results over all folds. For the Adaboost, the minimum MAE (2.0 ms) was obtained in fold-3, whereas the maximum MAE (6.3 ms) was achieved in fold-2.

Regression model	Mean Absolute Error (MAE) for validation folds (ms)										MAE over all folds
	1	2	3	4	5	6	7	8	9	10	
Linear regression (OLS)	7.3	12.7	6.8	8.6	8.4	10.0	8.1	8.8	8.5	9.3	8.9
Linear regression (LASSO)	6.0	12.5	5.7	7.5	7.6	9.5	6.7	7.7	7.1	8.3	8.3
Linear regression (Ridge)	6.0	12.4	5.9	7.3	7.6	9.5	6.7	7.8	6.9	8.4	8.5
Kernel ridge regression	10.6	13.0	5.6	9.1	7.9	10.5	13.2	9.5	6.8	11.3	11.3
Support vector regression	6.2	9.5	5.4	6.5	4.4	8.0	6.2	8.7	5.2	6.9	6.9
Random forest	4.3	8.1	3.0	3.6	2.7	6.1	4.3	6.2	3.3	4.7	4.7
Adaboost	2.8	6.3	2.0	2.5	2.2	5.4	3.0	4.5	2.3	3.7	3.5

Table D2: MedAE for the candidate regression models in the validation folds. Adaboost regressor produced the best results over all folds.

Regression model	Median Absolute Error (MedAE) for validation folds (ms)									
	1	2	3	4	5	6	7	8	9	10
Linear regression (OLS)	4.5	6.0	5.1	5.2	6.2	6.2	5.0	6.2	6.4	7.1
Linear regression (LASSO)	3.2	5.6	3.7	3.3	5.7	5.6	4.3	5.1	5.8	6.1
Linear regression (Ridge)	3.3	5.7	4.0	3.8	5.6	6.0	4.3	5.2	5.7	6.6
Kernel ridge regression	6.1	7.1	3.3	4.7	4.1	6.2	6.4	5.9	5.4	9.1
Support vector regression	3.0	3.9	2.3	3.1	1.6	3.8	3.7	5.2	3.3	3.6
Random forest	1.5	1.1	0.4	0.3	0.5	0.6	1.9	1.1	0.6	1.5
Adaboost	0.1	0.7	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0

Table D3: Best hyperparameter values for the candidate regression algorithms found via 5-fold nested grid search CV using the train sets of the 10-folds. The variables 'lin', 'poly', and 'rbf' stands for linear, polynomial, and radial basis function, respectively.

Regression algorithm	Hyperparameters for 10 folds found via 5-fold nested CV										
	Hyperparameter	1	2	3	4	5	6	7	8	9	10
LASSO	Alpha	0.01	0.1	0.1	1	0.001	0.1	0.1	0.01	0.1	0.1
Ridge	Alpha	10	10	10	2	10	10	10	10	10	10
Kernel ridge regression	Alpha	10	0.1	10	0.01	10	10	0.1	10	10	0.01
	Kernel	poly	poly	poly	poly	poly	poly	poly	poly	poly	rbf
SVR	C	10	10	10	0.1	10	10	10	10	10	10
	Kernel	rbf	rbf	rbf	lin	rbf	rbf	rbf	rbf	rbf	rbf
Random forest	max depth	17	17	20	17	17	15	15	17	17	17
	max features	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
Adaboost	max depth	12	10	20	15	15	20	15	12	22	15
	max features	0.7	0.3	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7
	learning rate	0.1	0.05	1	0.1	0.5	1	1	1	0.1	0.01

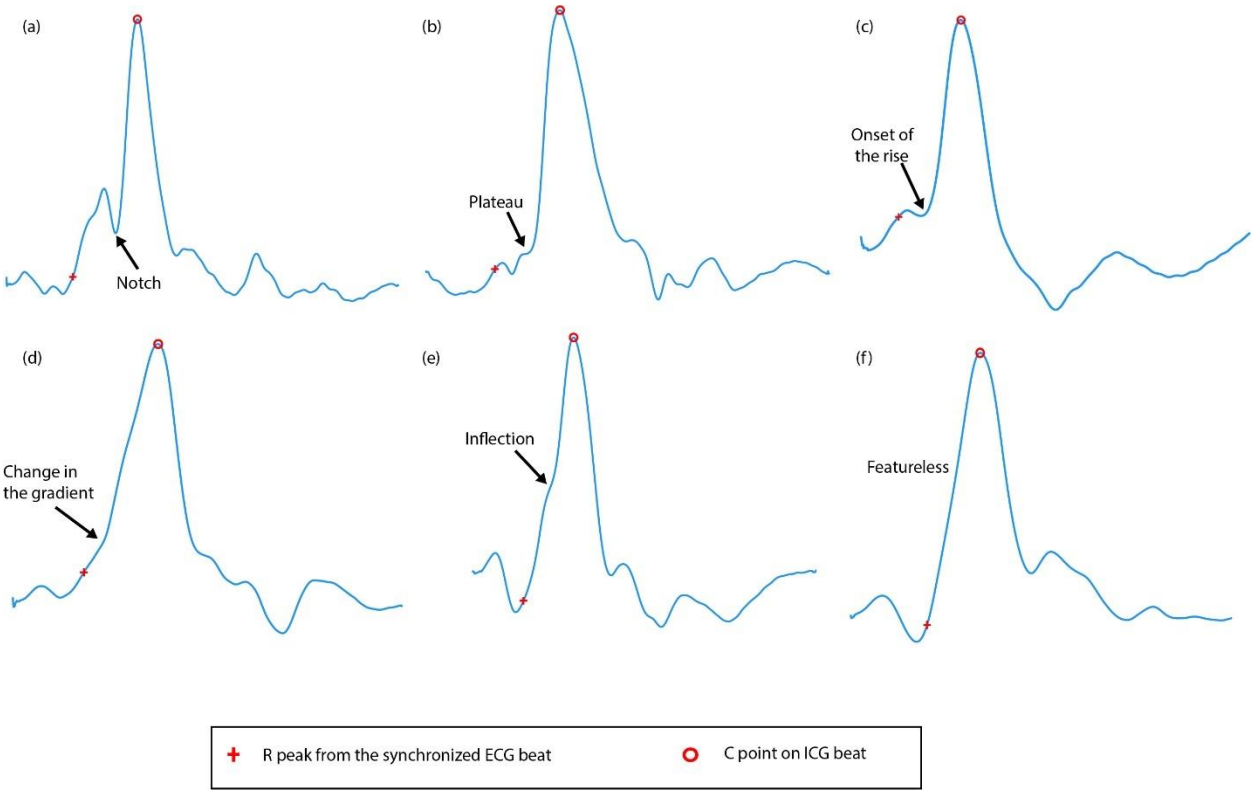

Figure 1: Morphological variations of the B point on different ICG cycles. ICG signals suffer from intra- and inter-subject morphological variabilities. The B point manifests itself in different morphological shapes which include (a) notch, (b) plateau, (c) onset of the rise, (d) change in the gradient, (e) inflection, and (f) featureless. In the case of featureless, no distinct feature is present to identify the B point on the ICG cycle. R peak from the synchronized ECG beat and C point on ICG signal are also marked for reference on the ICG cycle. Image (c) Emory University, CC-BY-SA.

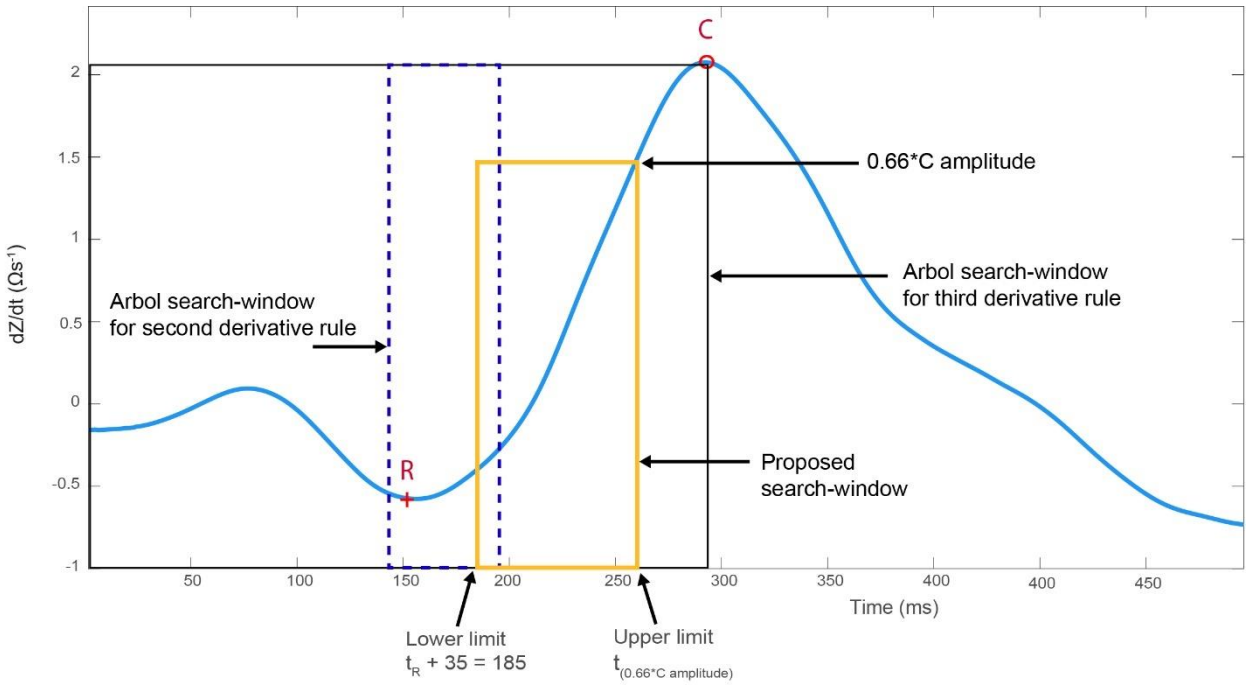

Figure 2: Proposed search-window for the extraction of morphological features. Proposed search-window based on physiologically valid PEP range, initial systolic time interval (RC interval) and maximum ejection velocity of blood in aorta (C point amplitude) is shown in solid amber color with lower and upper limits. Lower limit is R peak time index plus 35 ms, and upper limit is time index of 0.66*C amplitude. Arbol search-window of 300 ms duration prior to C point for third derivative rule is depicted in solid black color. For second derivative rule, Arbol search-window of 50 ms starting 150 ms before the C point is shown in dashed blue line (Arbol et al.,2017). Image (c) Emory University, CC-BY-SA.

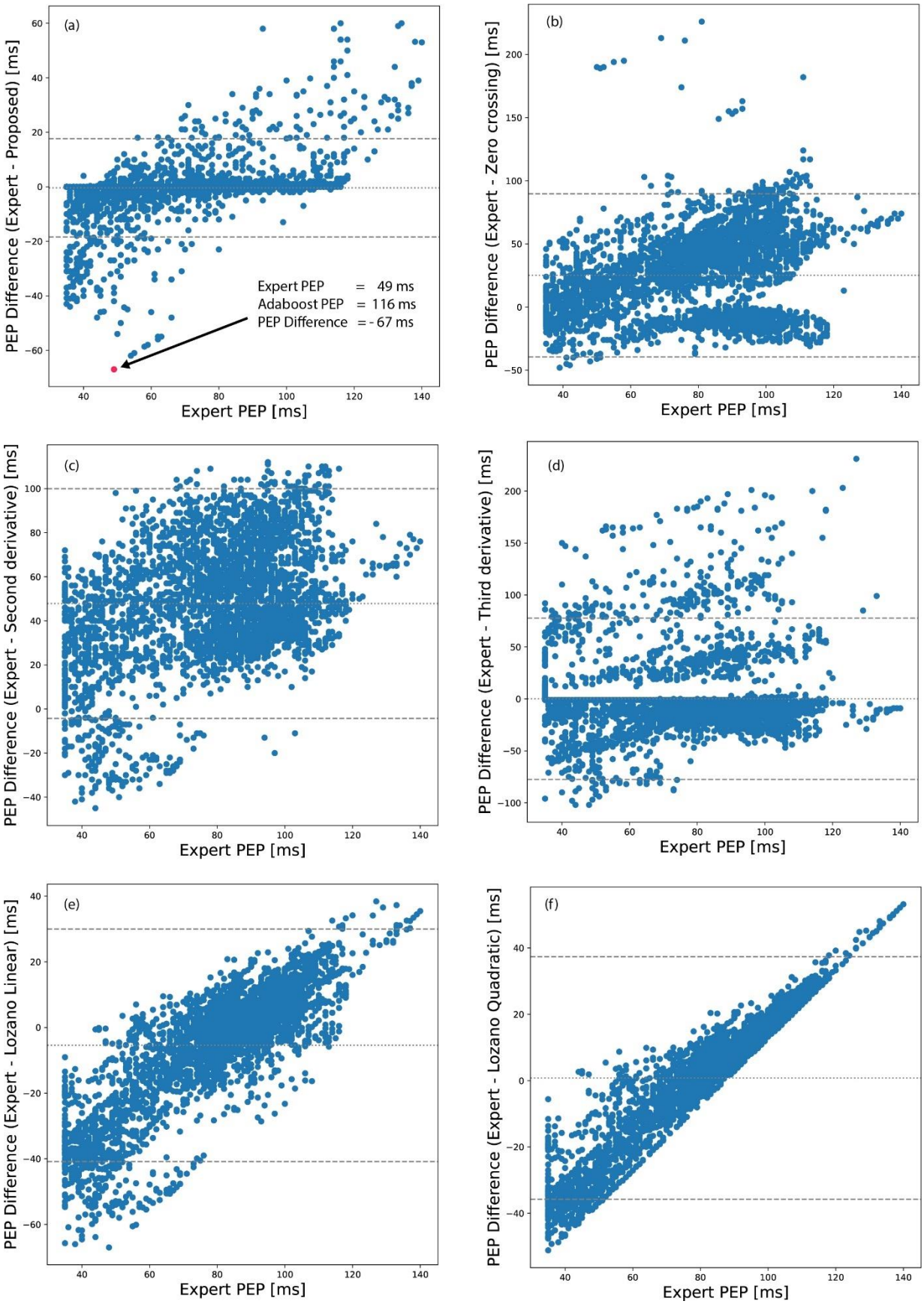
(a)
PEP Difference (Expert - Proposed) [ms]
Expert PEP [ms]
Expert PEP = 49 ms
Adaboost PEP = 116 ms
PEP Difference = - 67 ms
(b)
PEP Difference (Expert - Zero crossing) [ms]
Expert PEP [ms]
(c)
PEP Difference (Expert - Second derivative) [ms]
Expert PEP [ms]
(d)
PEP Difference (Expert - Third derivative) [ms]
Expert PEP [ms]
(e)
PEP Difference (Expert - Lozano Linear) [ms]
Expert PEP [ms]
(f)
PEP Difference (Expert - Lozano Quadratic) [ms]
Expert PEP [ms]

Figure 3: Bland–Altman plots for PEP difference between the expert-PEP and proposed/popular algorithms against the expert-PEP over 3657 EA ICG beats. Figure 3 (a) depicts the plot for proposed Adaboost regressor. A small portion of these beats were found to be outliers. On investigation, it was found that mostly the outlier beats belong to the same subjects. Discussion on an outlier case (highlighted as red dot) can be found in Section 4 and Figure 8(c)(d). Figure 3 (b) – Figure 3 (f) represents the Bland-Altman plots for zero crossing, second derivative rule, third derivative rule, Lozano linear estimate, and Lozano quadratic estimate. The horizontal dotted line shows the bias (mean of the difference). The horizontal dashed lines show the limits of agreement (bias ± 1.96∗SD) in all cases. Image (c) Emory University, CC-BY-SA.

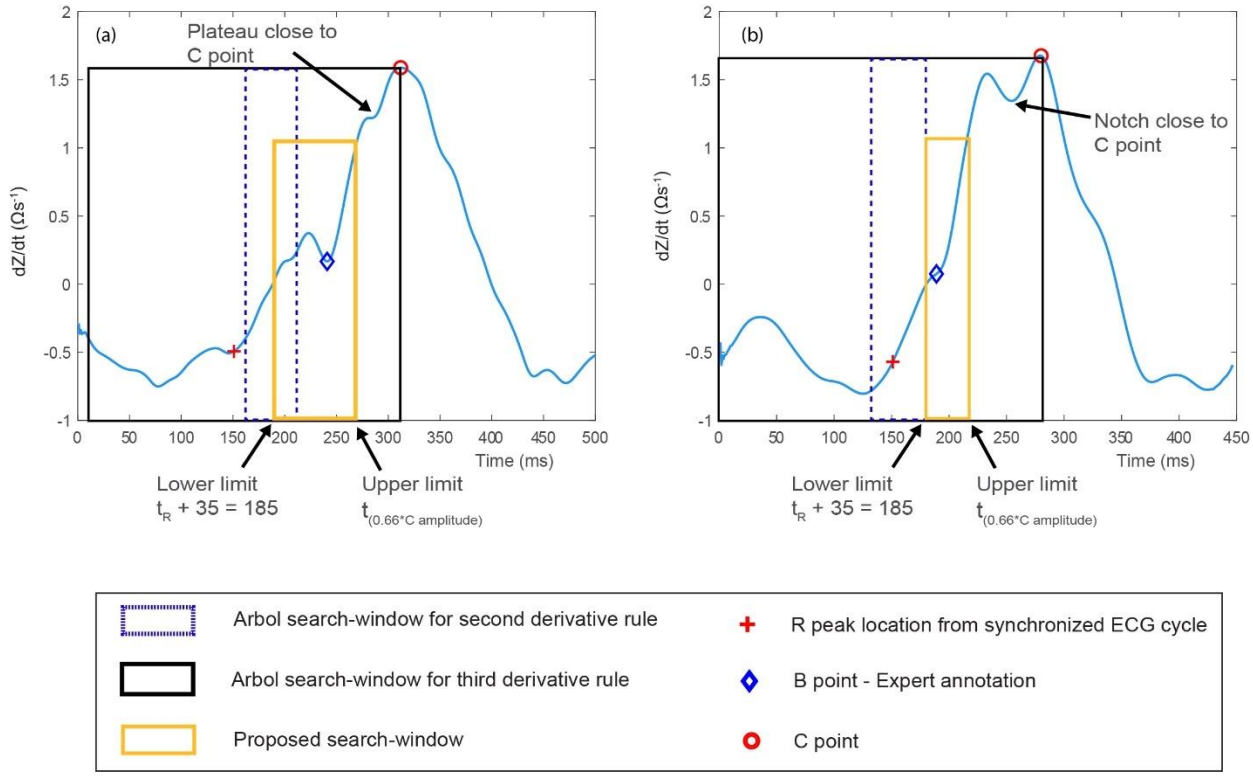

Figure 4: Search-window comparison. Figure 4 (a) and Figure 4 (b) represents the ICG waveform with notched and M-shaped C waves respectively. Arbol search-windows for second and third derivative rules are shown in dotted blue and solid black color on both plots. The proposed variable search-window, in solid amber color, is shown along with the lower and upper limits. Figure 4 (a): The B point manifests itself in the "notch" shape and falls outside the Arbol search-window for the second derivative rule. The notched C wave presents the morphology of "Plateau" near the C point, which can be mistakenly marked as B point in the Arbol search-window for the third derivative rule. The proposed window successfully captured the interval with B point in it. Figure 4 (b): The B point manifests itself in the "inflection" form and falls outside the Arbol search-window for the second derivative rule. The notched C wave presents the morphology of "notch" near the C point, which can be mistakenly marked as B point in the Arbol search-window for the third derivative rule. The proposed window successfully captured the interval with B point in it. Image (c) Emory University, CC-BY-SA.

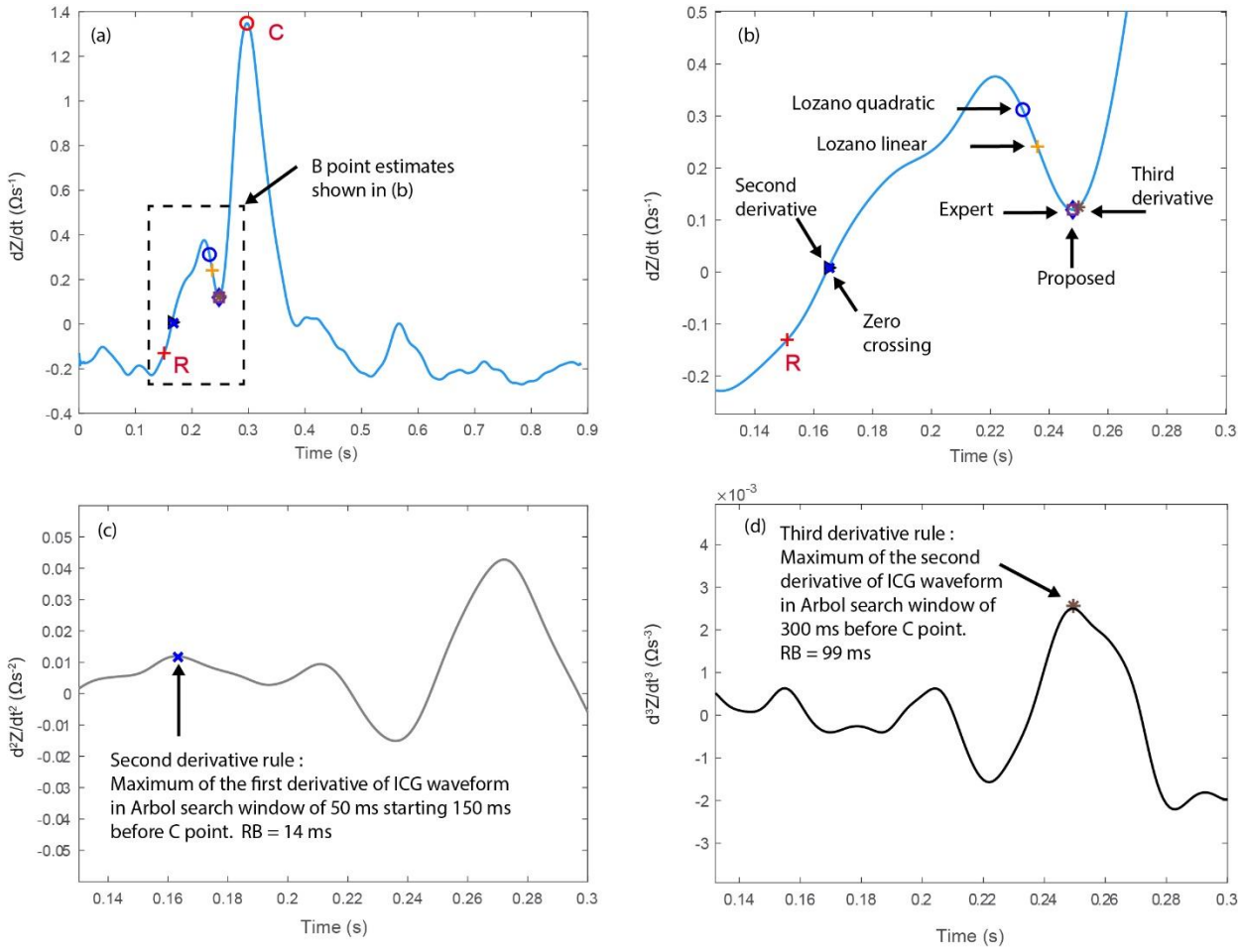

Figure 5: Performance comparison of B point detection using the proposed and state-of-the-art algorithms for 'notch' morphology. Figure 5 (a) presents the averaged ICG cycle with R peak, C point and B point locations estimates obtained using different algorithms. The RC interval for this ICG beat is 164 ms. Figure 5 (b) represents the zoom-in of B point estimates. The proposed algorithm was successful in identifying the "notch" morphology of the B point with an accurate estimation of PEP = 97 ms as per expert annotation. Zero-crossing (14 ms) and second derivative rule (14 ms) estimated the B point out of physiologically valid PEP range. Lozano quadratic (PEP = 80 ms) and Lozano linear (85 ms) underestimated the B point location. Whereas the third derivative rule (99 ms) slightly overestimated the B point location. Figure 5 (c), and Figure 5 (d) illustrate the application of second derivative and third derivative rules on the first and second derivative of ICG waveform in Arbol search-windows, respectively. Image (c) Emory University, CC-BY-SA.

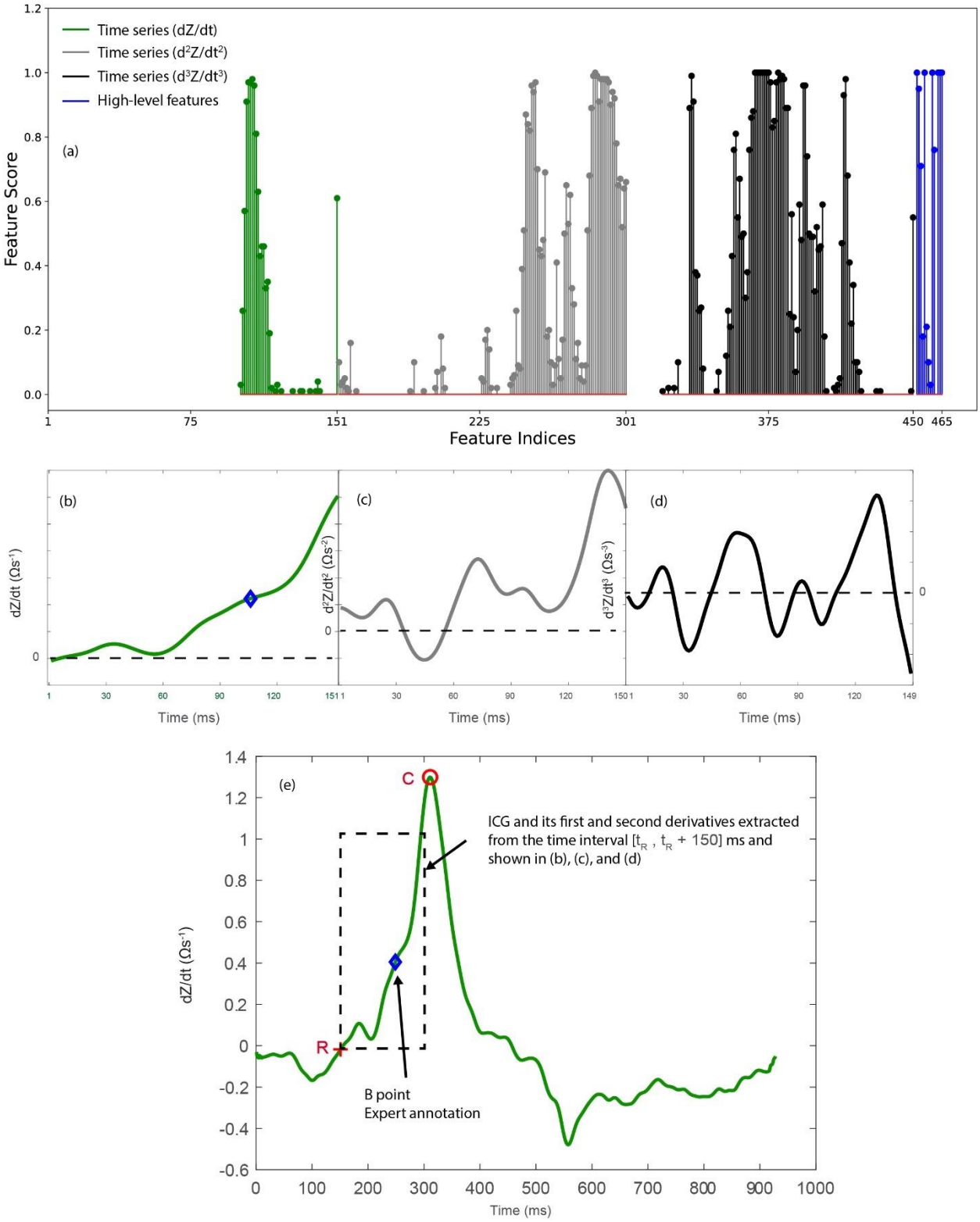

Figure 6: Feature importance analysis for proposed algorithm. A total of 100 sets of important features with feature-importance threshold greater than overall mean were extracted using 10 times repeated 10-fold CV

for the proposed algorithm. Feature occurrence frequency was normalized by dividing the feature occurrences by 100 to obtain the feature score for each feature index. as depicted in Figure 6(a). Feature indices on the x-axis consist of time-series features (index 1 - 450) and high-level features (index 451 - 465). Figure 6(e) represents an example of the ICG cycle from which the interval $\{t_R\ t_R + 150\}$ was extracted. Zoomed-in time-series features consisting of ICG waveform, its first and second-order derivatives in the time interval $\{t_R\ t_R + 150\}$ are depicted in Figure 6(b), 6(c), and 6(d) respectively. The black dashed lines in Figures 6(b), 6(c), and 6(d) represent the reference zero crossing for each plot. The ICG time-series features (with indices 1-100) were not significant for computing the B point location. The important features appeared in the form of small groups for time-series features. For instance, important features with feature score >0.6 appeared in small pockets of groups for ICG waveform (feature indices: 104 to 110), first derivative of ICG (feature indices: 249~255 and 282 ~ 298), and second derivative of ICG, features (feature indices =365 ~ 385). A total of 9 out of 15 high-level features with feature score >0.7 were also identified to play a pivotal role in predicting the B point. Image (c) Emory University, CC-BY-SA.

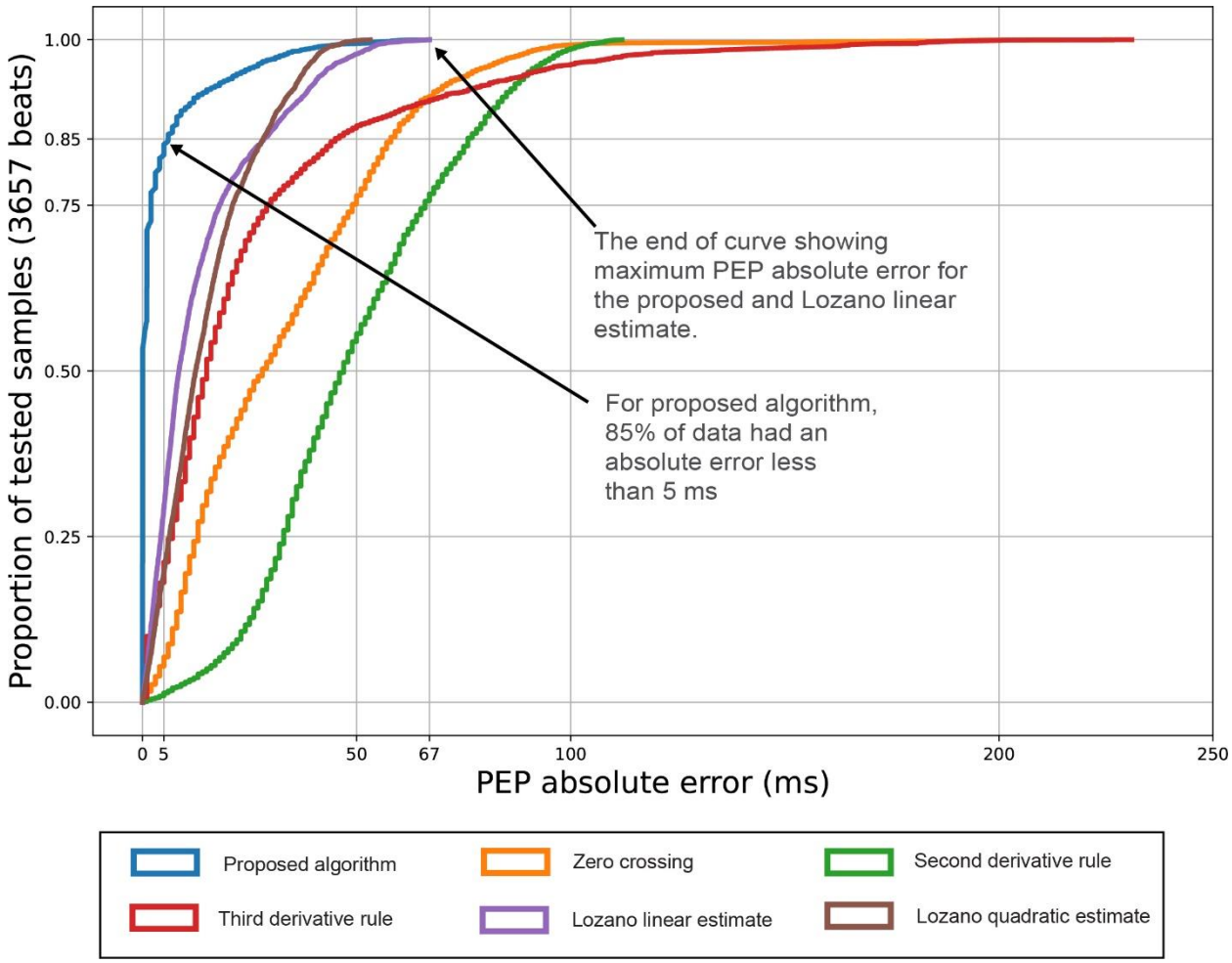

Figure 7: Cumulative distribution functions of PEP absolute errors for the proposed, and state-of-the-art algorithms. The height of each of the curves at any chosen PEP absolute error represents the proportion of the 3657 ICG beats that had x ms of error or less. For instance, 85% of the tested dataset had less than 5 ms of PEP absolute error for the proposed algorithm. The curves are ending at the maximum PEP absolute error for each algorithm. For example, the curves for the proposed and Lozano linear algorithm had maximum PEP absolute error of 67 ms. Image (c) Emory University, CC-BY-SA.

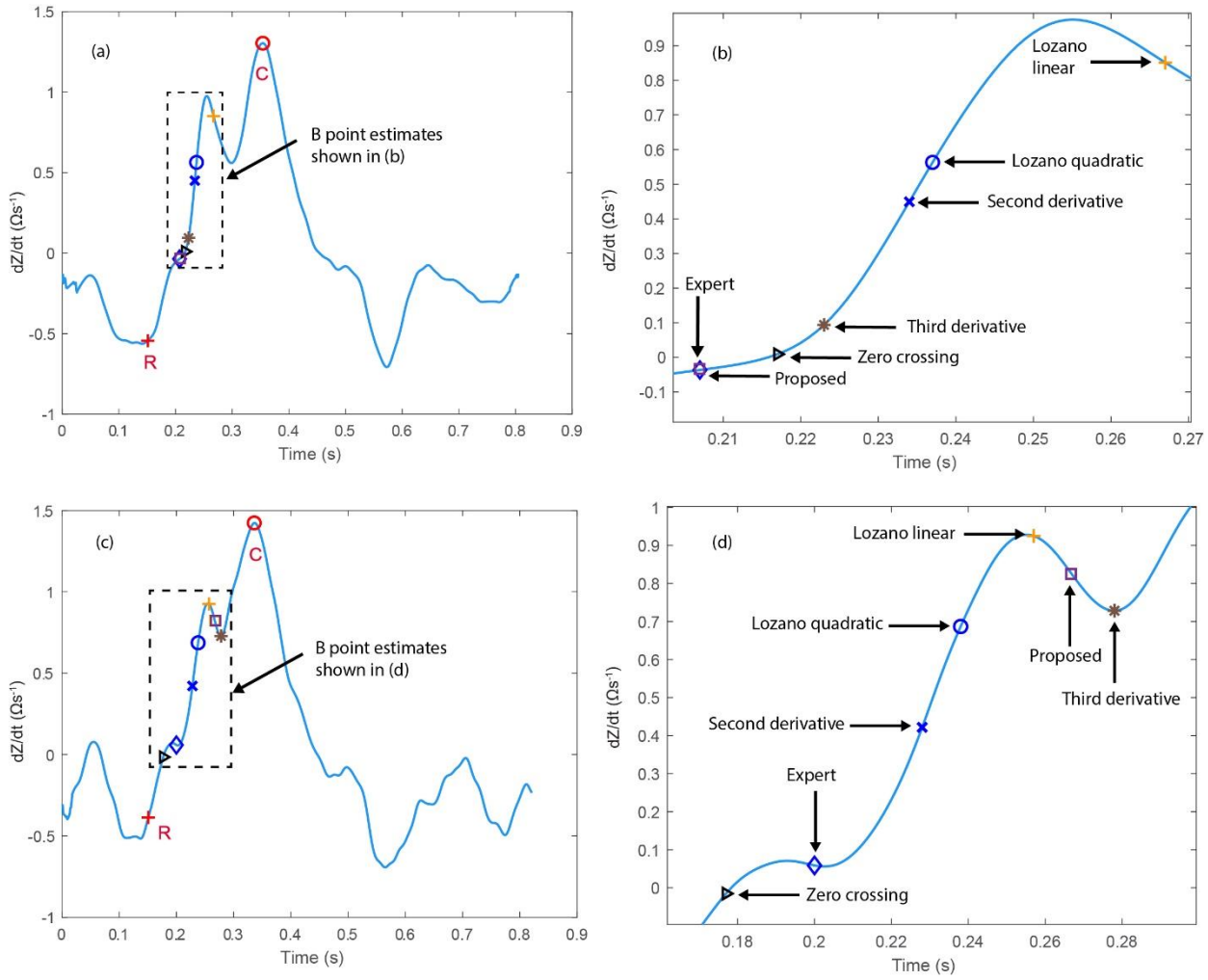

Figure 8: Performance comparison of B point detection using different algorithms for ICG waveforms with notched/M-shaped C wave. Figure 8 (a) and 8 (c) presents the ICG cycles from the same participant (study B in the same speaking task on two different days). Both ICG cycles were annotated with R peak, C point, and B point estimates using different algorithms. For Figure 8 (a), the RC interval was 203 ms. Figure 8 (b) represents the zoom-in of the B point estimates. Our proposed method was successful in identifying the "inflection" morphology of the B point with an accurate estimation of "PEP = 56 ms" as per expert annotation. The B point location were over-estimated via all other algorithms: zero crossing (66 ms), third derivative rule (72 ms), Lozano quadratic estimate (87 ms), Lozano linear estimate (116 ms), and second derivative rule (83 ms). For Figure 8(c), the RC interval was 184 ms. None of the algorithms were successful in identifying the "onset of the rise" morphology of the B point. Expert-PEP was 49 ms. Zero-crossing

underestimated the B point with an estimate of 25 ms. All other algorithms overestimated the B point: second derivative rule (77 ms), Lozano quadratic estimate (87 ms), Lozano linear estimate (106 ms), our proposed method (116 ms), and third derivative rule (127 ms). Image (c) Emory University, CC-BY-SA.

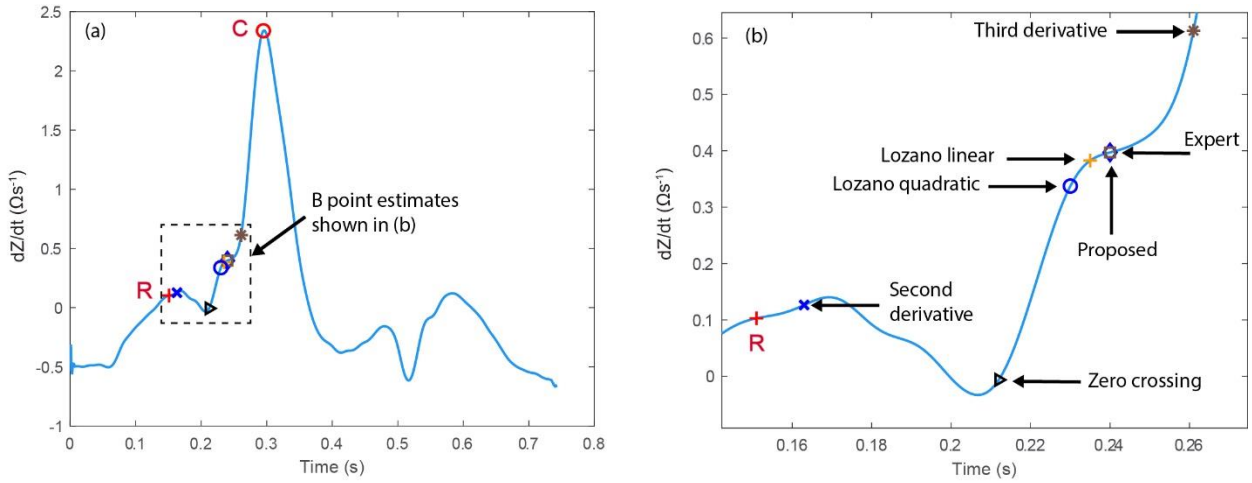

Figure E1: Performance comparison of B point detection using different algorithms for "Plateau" morphology. Figure E1 (a) presents the ICG cycle with R peak, C point and B point estimates using different algorithms. The RC interval for this example is 144 ms. Figure E1 (b) represents the zoom-in of B point estimates. The proposed algorithm was successful in identifying the "Plateau" morphology of the B point and resulted in accurate estimation of PEP = 89 ms as per expert annotation. B point estimated via the second derivative rule (12 ms) was physiologically invalid. Whereas zero crossing (62 ms), Lozano quadratic estimate (80 ms), and Lozano linear estimate (84 ms) under-estimated and third derivative rule (110 ms) overestimated the B point. Image (c) Emory University, CC-BY-SA.

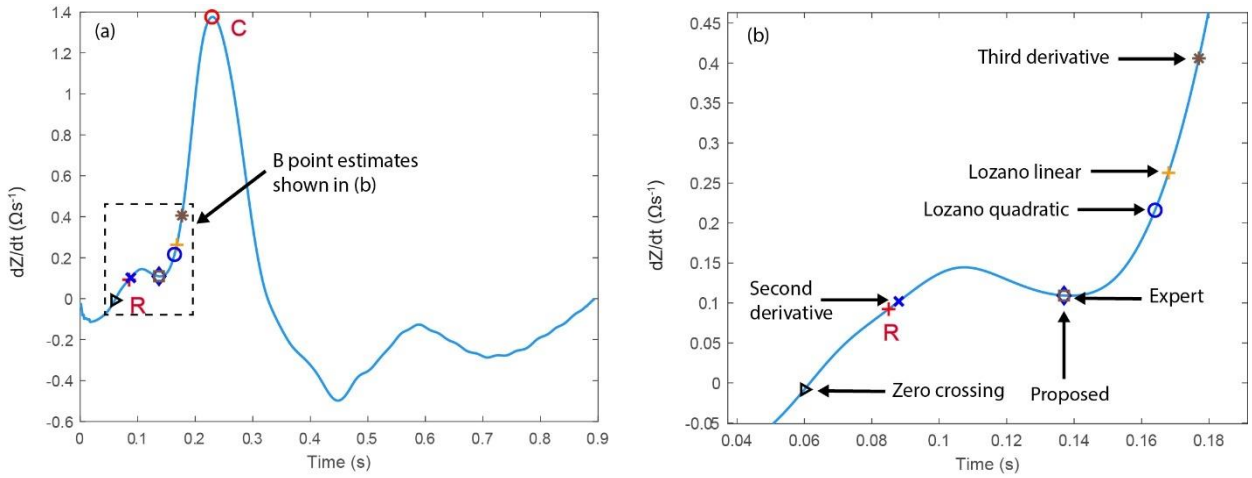

Figure E2: Performance comparison of B point detection using different algorithms for "Onset of the rise" morphology. Figure E2 (a) presents the ICG cycle with R peak, C point and B point estimates using different algorithms. The RC interval for this example is 143 ms. Figure E2 (b) represents the zoom-in of B point estimates. The proposed regressor was successful in identifying the "Onset of the rise" morphology of the B point and resulted in an accurate estimation of PEP = 52 ms as per expert annotation. B points detected by zero crossing and second derivative rule were physiologically invalid. Zero crossing detected B point prior to R peak, while B point detected by the second derivative rule was close to R peak (RB =3 ms). All other algorithms over-estimated the B point: Lozano quadratic estimate (79 ms), and Lozano linear estimate (83 ms), and third derivative rule (91 ms). Image (c) Emory University, CC-BY-SA.

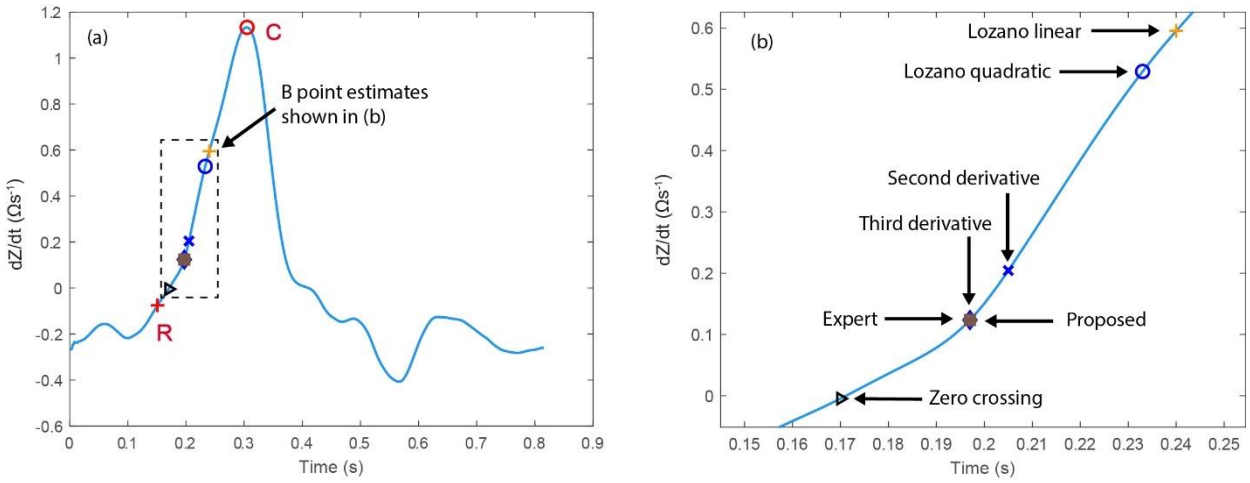

Figure E3: Performance comparison of B point detection using different algorithms for "Change in the gradient" morphology. Figure E3 (a) presents the ICG cycle with R peak, C point, and B point estimates using different algorithms. The RC interval for this example is 154 ms. Figure E3 (b) represents the zoom-in of the B point estimates. Adaboost regressor, and third derivative rule were successful in identifying the "Change in the gradient" morphology of the B point and resulted in accurate estimation of PEP = 46 ms as per expert annotation. Zero crossing underestimated the B point with PEP = 19 ms. Other algorithms overestimated the B point: second derivative rule (54 ms), Lozano quadratic estimate (82 ms), and Lozano linear estimate (89 ms). Image (c) Emory University, CC-BY-SA.

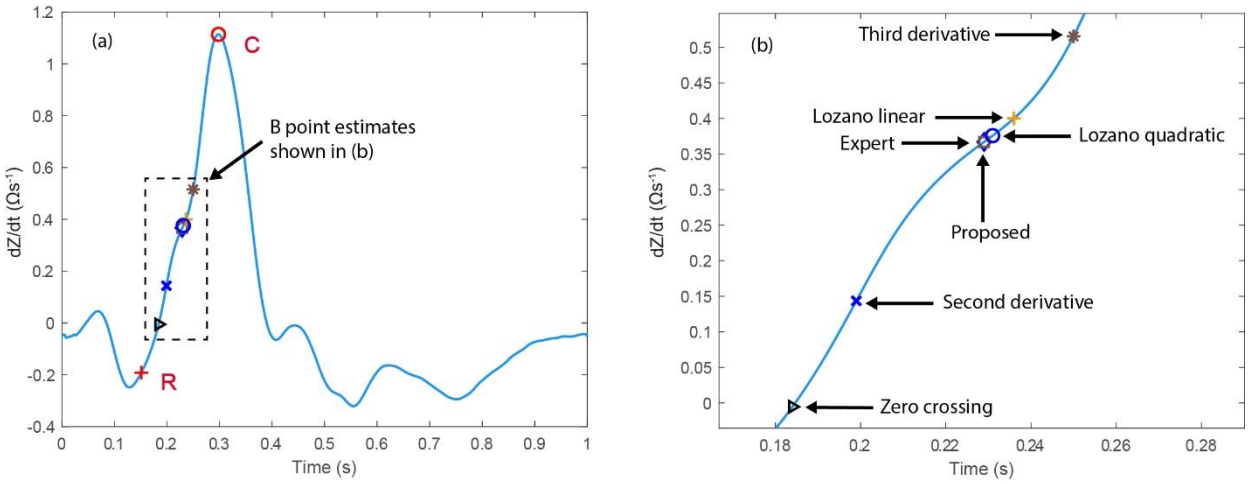

Figure E4: Performance comparison of the B point detection using proposed and state-of-the-art algorithms for the "inflection" morphology. Figure E4 (a) presents the ICG cycle with R peak, C point, and B point estimates using different algorithms. The RC interval for this ICG cycle is 147 ms. Figure E4 (b) represents the zoom-in of B point estimates. Proposed algorithm was successful in identifying the "inflection" morphology of the B point with an accurate PEP estimation (78 ms). The B point estimates were under-estimated by zero crossing (33 ms), and second derivative rule (47 ms). Other algorithms overestimated the B point: Lozano quadratic estimate (81 ms), Lozano linear estimate (85 ms), and third derivative rule (99 ms). Image (c) Emory University, CC-BY-SA.