Monitoring and Prediction of Cardiac Arrest in Pediatric ICU Patients with Machine Learning

April Yujie Yan[1,2], Sukrit Treewaree[1], Jiahui Yao[1], Jiwoo Noh[1], Sheel Tanna[1], Tamara Orduna[1],
Joseph L. Greenstein, PhD[1], Casey Overby Taylor, PhD[1,2], Timothy Ruchti, PhD[3] , James Fackler, MD[2], Jules Bergmann, MD[3]

[1] Whiting School of Engineering, Johns Hopkins University, [2] Johns Hopkins University School of Medicine, [3] Nihon Kohden Digital Health Solutions Inc

Introduction

- In-hospital **cardiac arrest** (IHCA) is a leading cause of mortality within pediatric intensive care unit (PICU), causing **~40%** of pediatric CA in the US every year.
- Studies have shown that analysis of **monitored physiological data** can provide early warnings that enable **timely interventions** and potential CA prevention.
- A pilot study demonstrated the potential of IHCA prediction using **machine learning (ML) models** trained on **heart-rate variability (HRV)** features extracted from the **electrocardiogram (ECG)**.
- However, **photoplethysmography (PPG)** morphology and precursor events (e.g., respiratory failure) that may be indicative of cardiac arrest remain unexplored.

Objectives

1. Develop a novel ML algorithm that leverages **morphological features from PPG**, as well as HRV metrics from ECG, vital signs, demographics, medications, and precursor events such as respiratory failure.
2. Assess the algorithm's performance for **accurate and timely IHCA predictions to alert clinicians and allow for potential interventions**.

Materials and Methods

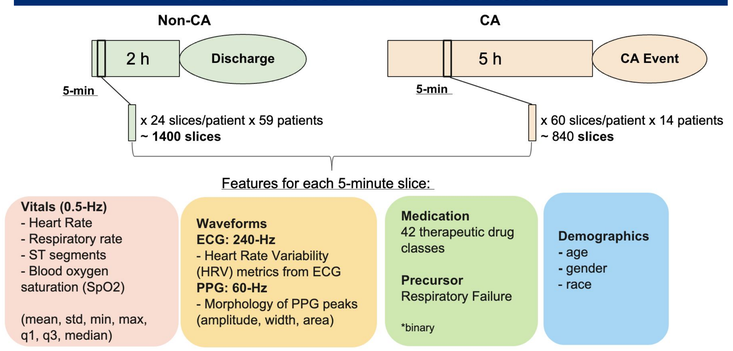

Figure 1—Feature Engineering. Non-CA (n = 59) and CA (n = 14) patients' data are organized as **5-minute slices** and includes features such as vitals, waveforms, medications, precursor events, and demographics. Non-CA patients' data were retrieved from 2 hours before discharge, and CA patients' data were retrieved 5 hours before the cardiac arrest onset. The derived metrics include **23 HRV metrics derived from ECG** waveforms, **21** summary statistics from **3 PPG morphological features, 10 vital signs (**respiratory rate, blood oxygen saturation, **principal components of ST segments** etc.), and **42** therapeutic drug classes and respiratory failure as binary.

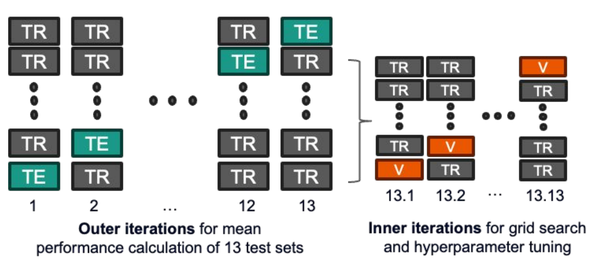

Figure 2—Nested Cross Validation. 13 CA patients with adequate denoised signals were used for training four ML models (e.g., **logistic regression, random forest, support vector machine, XGBoost**) using 13-fold nested cross validation where each fold contains data from one CA patient. We computed average performance across 13 test sets across all folds.

Results

Features	CA (n = 14)	non-CA (n = 59)
Gender: Female*	6 (42.9%)	34 (57.6%)
Gender: Male*	8 (57.1%)	25 (42.4%)
Age (year)*	3.9 ± 5.6	3.7 ± 5.6
Heart Rate (bpm)	129.3 ± 3.9	128.9 ± 6.9
Oxygen Saturation (Spo2)*	91.7 ± 1.1	96.9 ± 1.0
Respiratory Rate (/min)*	29.2 ± 15.4	41.7 ± 16.7
Respiratory Failure*	12 (85.7%)	4 (6.8%)
Medication		
Autonomic Drugs*	9 (64.3%)	2 (3.4%)
Blood-derived Products*	6 (42.9%)	2 (3.4%)
Diagnostic Agents*	4 (28.6%)	2 (3.4%)
Heart Rate Variability (HRV)		
Mean NN Interval	495.1 ± 134.5	497.0 ± 137.7
NN Interval Variability (SDNN)*	20.1 ± 21.6	29.6 ± 23.7
Low Frequency*	109.5 ± 257.6	209.3 ± 312.9
High Frequency*	62.0 ± 169.4	165.5 ± 418.0
Cardiac Vagal Index*	3.14 ± 0.86	3.65 ± 0.64
Triangular index*	3.51 ± 2.81	6.16 ± 4.28
PPG Morphology		
Width-50*	15.8 ± 11.1	17.4 ± 12.0
Amplitude*	1.60 ± 0.17	1.63 ± 0.21
Area*	0.52 ± 0.19	0.56 ± 0.24

Table 1— Characteristics of Patient Cohort
Table 1 demonstrates a subset of characteristics of the patient cohort, including demographics (gender, age), vital signs (e.g., heart rate, oxygen saturation, respiratory rate), the precursor event (respiratory failure), three prominent medications (autonomic drugs, blood, diagnostics), six prominent HRV metrics, and PPG morphological features. **All features showed statistically significant differences (p<0.05)** between distributions in CA and non-CA patients, except heart rate (p=0.678) and mean NNI (p=0.777).

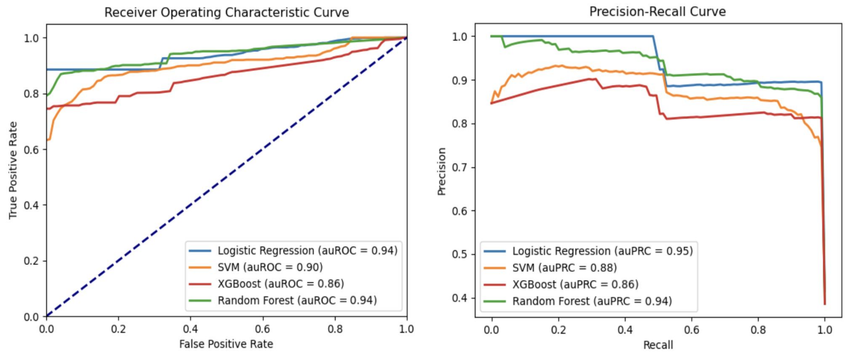

Figure 3—Evaluation of Models Trained using All Features
Figure 3 presents **receiver operating characteristic (ROC)** curves and **precision-recall (PR)** curves for the four models trained using all features extracted. Training various models with all features revealed that random forest has the best performance results across different feature combinations: **auPRC (area under PRC) = 0.937, auROC (area under ROC) = 0.940, accuracy = 0.848, specificity = 0.970, sensitivity/recall = 0.658, positive-predictive-value (PPV/precision) = 0.873, and negative-predictive-value (NPV) = 0.840.**

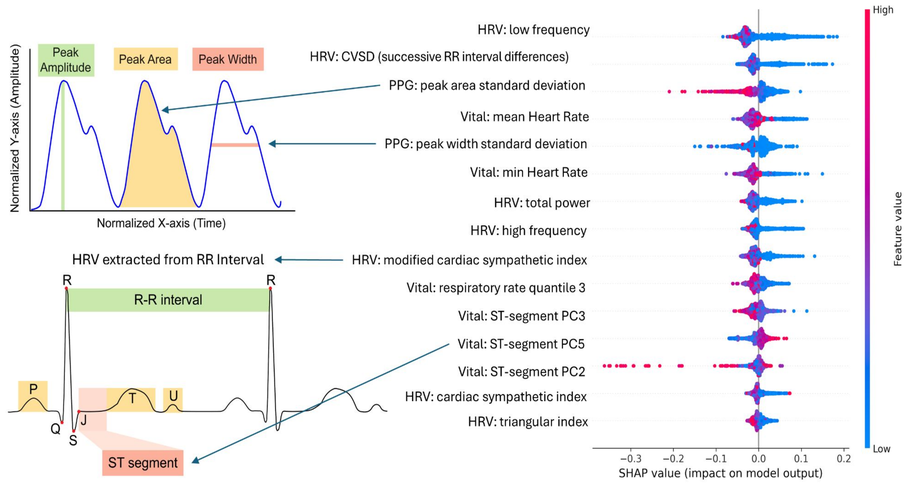

Figure 4—SHAP Analysis for Feature Interpretation on Real-time Signals
Figure 4 demonstrates top important features based on SHAP (SHapley Additive exPlanations). Features that were used in SHAP analysis does not include the precursor events (i.e., respiratory failure) and medication because they are known to be highly correlated with the outcome and may induce bias. We are more interested in how the model trained with only real-time monitored data (e.g., ECG, PPG, vital signs) could perform in the prediction of cardiac arrest. In the beeswarm plot, data points with **positive SHAP values contribute to higher probability of the occurrence of cardiac arrest.**

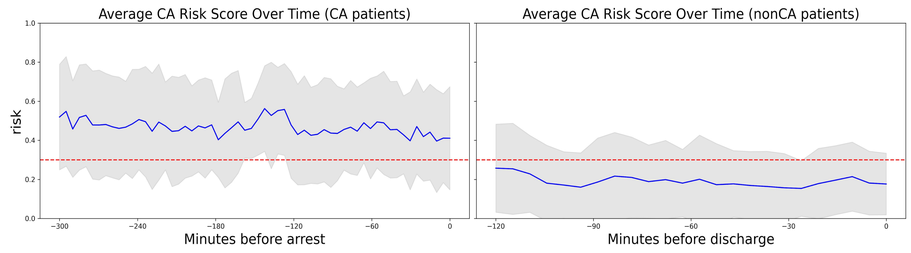

Figure 5—Risk Prediction Over Time by Random Forest Trained using Real-time Signals
Figure 5 demonstrates the predicted risks of cardiac arrest for CA patients (left) and non-CA patients (right) over time. Risk represents the **positive predicted probability** of the outcome of a certain **5-min slice.**
In both panels, the blue line represents average risk of all CA or non-CA patients and the gray area represents standard deviation. In general, CA patients have higher risks prior to the onset of cardiac arrest. Based on the overall trend of average risks, we can select **a risk threshold** (e.g., **p = 0.3, since red dashed line clearly separates these two states: CA vs. non-CA) for raising alarms to clinician.** We plan to make inference on data from the whole admission (e.g., 12-hour data prior to CA) to see the trend of risk progression over time.

Discussion

- This project demonstrates promising results in ML-based prediction of pediatric IHCA by leveraging **real-time monitored data** (e.g., ECG, PPG, physiological time series) and **static data** (e.g., demographics). The developed ML models achieve and demonstrate **actionable early warning of impending IHCA** in pediatric patients using multimodal signals and electronic health record data that are collected routinely in the PICU.
- Moving forward, we plan to bolster our dataset by implementing time series data of 12-hour prior to cardiac arrest, thereby providing a more comprehensive prediction of cardiac arrest in the PICU.