between and within the clusters. The distinctive patterns could be used in cardiovascular (CVS) computational models and CVS state estimation.

## Methods

We analyzed 38 patients with repeated PE administrations (173 events) from MOVER database (https://mover.ics.uci.edu/). Time series of MAP values within a 360-pulse window after annotated PE administration were smoothed using a moving average (window size = 10), and K-Shape clustering was applied. We computed the following metrics from the original (unsmoothed) time series: MAP average of the 120 pulses before PE ($MAP_{pre}$), MAP average of the 360 pulses after PE administration ($MAP_{post}$), and the change between them ($MAP_{change} = MAP_{post} - MAP_{pre}$). We tested differences in these metrics between K-Shape clusters, and among patients within the same cluster, using t-tests. A p-value of $< 0.05$ was considered statistically significant.

## Results

K-Shape clustering, directed by the silhouette index, identified two distinct response patterns. Cluster center 1 showed a sharp increase around beat 90 gradually decreasing toward the end of the window. Cluster center 2 exhibited a mostly flat trend with a pronounced peak between pulses 200 and 250. Within patients of the Cluster 1, $MAP_{pre}$ to $MAP_{post}$ increased significantly ($p = 0.04$), a change not observed in those of Cluster 2. No significant differences in the explored metrics were found between clusters.

## Conclusions

Clustering revealed two distinct temporal patterns, without significant differences in the metrics explored between them. In the Cluster 1, the significant increase of $MAP_{pre}$ to $MAP_{post}$ suggests that a rapid upward trend around pulse 90 coincides with an overall MAP increase in the analyzed segment and could be used as an indicator of an immediate positive response.

doi:10.1016/j.jcrc.2025.155383

---

# Development of a photoplethysmography-derived cardiovascular age metric for trauma and intensive care risk stratification

Ryan B. Felix[b,c,*], Kathalyn Urquizo[c], Meagan Watkins[c], William Gu[c], Bradford Burdette[c], Shiming Yang[c], Jamie Podell[a,d], John P. Fisher[b], William Teeter[e,f], Neeraj Badjatia[a,d], Peter Hu[a,c]

[a]*Program in Trauma, University of Maryland School of Medicine, Baltimore, USA*
[b]*Fischell Department of Bioengineering, University of Maryland, College Park, USA*
[c]*Department of Anesthesiology, University of Maryland School of Medicine, Baltimore, USA*
[d]*Department of Neurology, University of Maryland School of Medicine, Baltimore, USA*
[e]*Department of Emergency Medicine, University of Maryland School of Medicine, Baltimore, USA*
[f]*Charles McC Mathias National Study Center for Trauma & Emergency Medical Systems, Baltimore, USA*
*Corresponding author.
*E-mail address:* rfelix12@umd.edu (R.B. Felix)

## Introduction

Despite advances in artificial intelligence (AI) for clinical decision support (CDS), most validated models focus on electronic health records (EHR) and static predictors, with limited incorporation of real-time physiological waveforms. Many studies have shown the predictive power of heart-rate variability (HRV) metrics for mortality or deterioration; however, cardiovascular fitness measures are underrepresented in ICU CDS systems. Such gaps limit the ability to capture dynamic cardiovascular resilience – a key predictor of outcome in trauma and critical illness. Here we develop a machine learning pipeline using continuous photoplethysmography (PPG) waveforms to estimate physiological age, integrating heart rate variability, dicrotic notch morphology, and vascular stiffness proxies. We aim to use these physiological markers to improve risk stratification in trauma patients, moving beyond chronological age and static demographics.

## Methods

From a retrospective cohort of 29,929 adult trauma patients admitted to a level one trauma center between 2009–2018, we collected continuous PPG signals from their first hour after arrival. We extracted HRV metrics, features of dicrotic notch morphology, vectorized PPG waveform morphologies, and also used PPG waveform data as a non-decomposed time-series variable. We performed modeling for physiological age determination using three approaches: 1) binary classification of young ($\leq 38$) vs. old ($\geq 60$) age using supervised machine learning on tabular variables, 2) unsupervised clustering of waveform morphologies via vectorized dissimilarities and 3) direct regression for age prediction using time-series models.

## Results

Our binary classifiers achieved excellent performance for classifying young vs. old age (AUROC 0.89–0.91). Our vectorized dissimilarity approach showed clear morphological separation of PPG waveforms by age groups. Results from our physiological age regression modeling using foundational models and extrapolated associations between physiological age and adverse events are in progress.

## Conclusions

While HRV-based models are well established, no prior study has combined stiffness and aging clocks derived from PPG into predictive models for trauma or critical care CDDS. By focusing on real-time, physiological metrics, we aim to improve risk stratification, moving the field toward CDS systems that capture dynamic cardiovascular resilience rather than static demographic or comorbidity factors.

doi:10.1016/j.jcrc.2025.155384

---

# Learning and evaluating improved reinforcement learning-based policies for sepsis treatment on MIMIC-IV

Mohammad MahdiAzarbeik[a,b,c,*], Lorenz Kapral[a,b,c], Richard Weiss[b], Clemens Heitzinger[b], Oliver Kimberger[a,c]

[a]*Ludwig Boltzmann Institute for Digital Health and Patient Safety, Vienna, Austria*
[b]*Department of Computer Science, Technical University of Vienna, Vienna, Austria*
[c]*Department of Anaesthesia, Intensive Care Medicine and Pain Medicine, Clinical Division of General Anaesthesia and Intensive Care Medicine, Medical University of Vienna, Vienna, Austria*
*Corresponding author.
*E-mail address:* mohammad.azarbeik@lbg.ac.at (M. MahdiAzarbeik)

## Introduction

Sepsis remains a leading cause of mortality among critically ill patients, imposing a substantial burden on healthcare systems globally[1]. Timely and effective clinical decision support is essential for the management of intravenous fluids and vasopressors in sepsis treatment, and reinforcement learning (RL) has emerged as a promising approach for developing dynamic and personalized treatment strategies[2]. However, existing RL-algorithms in this domain often employ policy evaluation techniques such as Weighted Importance Sampling (WIS), which can exhibit limitations in reliability in complex environments[3,4,5]. Despite the progress, developing RL models that robustly attain clinician-level performance in sepsis management remains challenging. This study proposes solutions to the identified limitations and is evaluated using the latest version of the Medical Information Mart for Intensive Care (MIMIC) database, MIMIC-IV v3.1, which contains data from 94,458 unique ICU stays[6,7].

## Methods

We implement two widely used model-free RL algorithms in the field: Q-learning[8], following the approach of Komorowski et al.[3], and the more advanced Dueling Deep Q-Network (DuelingDQN), which uses neural network architectures. To define meaningful patient states, we adopt a weighted clustering approach that emphasizes clinically important features. Traditional clustering methods in high-dimensional medical data can be susceptible to the curse of dimensionality[9]. By weighting features based on their relevance to sepsis treatment, we aim to create more clinically meaningful state representations.

While Importance Sampling methods such as Weighted Importance Sampling (WIS) are commonly used for off-policy evaluation (OPE) in the clinical domain[3,4,5], they can produce high variance estimates, potentially leading to inaccurate conclusions. To overcome this limitation, we employ a set of more robust OPE methods, including tabular Fitted Q-Evaluation (FQE)[10], and stationary-distribution correction methods, such as DualDICE[11], GenDICE[12], and GradientDICE[13]. FQE is recognized for its effectiveness in evaluating policies from historical data, although it is computationally intensive[14].

Consistent with prior work[3], we adopt a sparse binary reward structure: +100 for survival, -100 for death, and 0 otherwise. Recent studies have explored incorporating intermediate rewards (e.g., vital sign stabilization or lactate clearance), which can potentially yield higher policy values[15,16]. However, modifying the reward function fundamentally alters the optimization target, which hinders cross-study comparisons of policy values.

## Results

When evaluating using tabular FQE and DICE-based methods, Q-learning underperforms relative to human clinicians. However, it outperforms them when using WIS. This observation suggests that more sophisticated algorithms are needed to reach the level of human medical professionals in the complex domain of sepsis treatment. This is demonstrated by the DuelingDQN agent performing at the same level as human clinicians, with slight improvements observed when using more robust evaluation metrics. The discrepancy with WIS-based evaluations, which tended to overestimate the performance of RL agents, underscores the limitations of WIS in this setting and highlights the value of employing more robust OPE methods.

## Conclusions

This study highlights the need for methodological rigor in both RL algorithm development and evaluation strategies in the clinical domain. While DuelingDQN approaches human-level performance, surpassing the level of human clinicians remains a substantial challenge. Future research could explore more advanced RL architectures, enhanced state representation strategies beyond basic clustering, more effective exploration techniques, or the integration of domain knowledge into the learning process to achieve superior performance.

Equally important is the choice of evaluation method. Our findings demonstrate that reliance on simplistic techniques like WIS can lead to overly optimistic conclusions. Advanced OPE techniques such as FQE and DICE variants offer more accurate assessments of RL policy performance. Building on these findings, improvements can help to create the methodological basis for the transition from retrospective research into prospective clinical studies.

doi:10.1016/j.jcrc.2025.155385