

Original Paper

Using Machine Learning to Predict Early Onset Acute Organ Failure in Critically Ill Intensive Care Unit Patients With Sickle Cell Disease: Retrospective Study

Akram Mohammed¹, PhD; Pradeep S B Podila², PhD, MHA, MS; Robert L Davis¹, MD, MPH; Kenneth I Ataga³, MD; Jane S Hankins⁴, MD, MS; Rishikesan Kamaleswaran⁵, PhD

¹Center for Biomedical Informatics, University of Tennessee Health Science Center, Memphis, TN, United States

²Faith and Health Division, Methodist Le Bonheur Healthcare, Memphis, TN, United States

³Center for Sickle Cell Disease, University of Tennessee Health Science Center, Memphis, TN, United States

⁴Department of Hematology, St Jude Children's Research Hospital, Memphis, TN, United States

⁵Department of Biomedical Informatics, Emory University School of Medicine, Atlanta, GA, United States

Corresponding Author:

Rishikesan Kamaleswaran, PhD

Department of Biomedical Informatics

Emory University School of Medicine

WMB, 101 Woodruff Circle, Suite 4127

Atlanta, GA, 30322

United States

Phone: 1 (901) 462 6908

Email: rkamaleswaran@emory.edu

Abstract

Background: Sickle cell disease (SCD) is a genetic disorder of the red blood cells, resulting in multiple acute and chronic complications, including pain episodes, stroke, and kidney disease. Patients with SCD develop chronic organ dysfunction, which may progress to organ failure during disease exacerbations. Early detection of acute physiological deterioration leading to organ failure is not always attainable. Machine learning techniques that allow for prediction of organ failure may enable early identification and treatment and potentially reduce mortality.

Objective: The aim of this study was to test the hypothesis that machine learning physiomarkers can predict the development of organ dysfunction in a sample of adult patients with SCD admitted to intensive care units (ICUs).

Methods: We applied diverse machine learning methods, statistical methods, and data visualization techniques to develop classification models to distinguish SCD from controls.

Results: We studied 63 sequential SCD patients admitted to ICUs with 163 patient encounters (mean age 30.7 years, SD 9.8 years). A subset of these patient encounters, 22.7% (37/163), met the sequential organ failure assessment criteria. The other 126 SCD patient encounters served as controls. A set of signal processing features (such as fast Fourier transform, energy, and continuous wavelet transform) derived from heart rate, blood pressure, and respiratory rate was identified to distinguish patients with SCD who developed acute physiological deterioration leading to organ failure from patients with SCD who did not meet the criteria. A multilayer perceptron model accurately predicted organ failure up to 6 hours before onset, with an average sensitivity and specificity of 96% and 98%, respectively.

Conclusions: This retrospective study demonstrated the viability of using machine learning to predict acute organ failure among hospitalized adults with SCD. The discovery of salient physiomarkers through machine learning techniques has the potential to further accelerate the development and implementation of innovative care delivery protocols and strategies for medically vulnerable patients.

(*J Med Internet Res* 2020;22(5):e14693) doi: [10.2196/14693](https://doi.org/10.2196/14693)

KEYWORDS

multiple organ failure; sickle cell disease; machine learning; electronic medical record; hematology

Introduction

Background

Sickle cell disease (SCD), one of the most common genetic disorders, affects millions across the globe [1]. It was the first monogenic disorder to be characterized at the molecular level. It is characterized by the presence of abnormal hemoglobin S, which, under hypoxic conditions, causes sickling of red blood cells, resulting in tissue and organ damage. Among an array of complications afflicting patients with SCD, the most devastating is major organ failure, including pulmonary failure, end-stage renal disease, stroke, and heart failure [1]. A 4-decade observational study reported that, by the fifth decade of life, up to half of all patients with SCD had documented irreversible organ damage [2]. Organ dysfunction may manifest or worsen during hospitalizations, when disease complications arise. Thus, therapy supplemented by predictive analytics can potentially improve the outcomes of patients with SCD [3]. Early diagnosis of acute organ dysfunction may allow for early intervention, thereby preventing or reducing the severity of organ failure, particularly during hospitalization for acute complications.

Early recognition of organ failure may [4,5] thereby enable clinicians to provide targeted therapies to improve outcomes. Various scoring methods have been developed for qualifying organ dysfunction, including Acute Physiology and Chronic Health Evaluation [6], Multi-Organ Dysfunction Score [7], *quick* Sequential Organ Failure Assessment [8] and Sequential Organ Failure Assessment (SOFA) [9]. The SOFA is a mortality prediction score that is based on the degree of dysfunction of six organ systems. The score is calculated at admission and every 24 hours until discharge, using the worst parameters measured during the previous 24 hours. Compared with other scoring methods, the use of SOFA allowed us to retrospectively quantify both the number and severity of individual organ dysfunction.

Objectives

In this retrospective study, we used serial calculations of SOFA to identify the onset of organ failure and then used physiologic markers in machine learning models to predict organ failure for patients with SCD presenting with a severe, acute painful crisis. Our hypothesis was that physiologic markers [10] identified by machine learning methods can be used to predict organ failure.

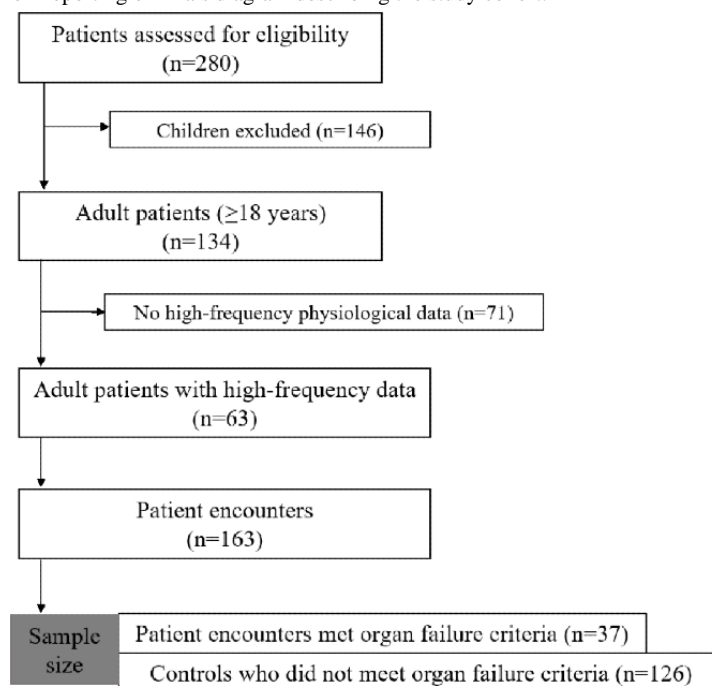
Methods

Cohort

Continuous physiologic data were collected on 134 adult subjects with SCD admitted to intensive care units (ICUs) at

Methodist Le Bonheur Healthcare hospitals, Memphis, Tennessee, United States, between June 2017 and March 2018. Patients were retrospectively identified using a discharge International Classification of Diseases, Tenth Revision (ICD-10) code of D57.*. Of the 134 unique patients, 71 patients who did not have at least 24 hours of continuous physiologic data were excluded from the analysis. We studied patients who had at least 24 hours of continuous high-frequency physiologic data available before the time of organ failure onset (identified using SOFA criteria). A total of 63 unique adult subjects who had SCD (discharge ICD-10 code of D57.*) and were admitted to the medical, surgical, neurological, and cardiac ICUs and had continuous physiologic data available were retrospectively identified and included in the study. These 63 patients had 163 encounters (Figure 1). Of the 163 encounters, 37 patient encounters corresponding to 29 unique patients met organ failure criteria. The inclusion and exclusion criteria are summarized in Figure 1. Demographic and clinical data on cases and controls were collected from the electronic medical record (EMR) using Cerner's Web Intelligence reporting module (Cerner Health Facts). Each patient admission was considered as a separate patient encounter if the interval between admissions was at least for 1 month. The principal or admit diagnosis was identified using the ICD-10 codes. Patients with organ failure at admission were excluded from the analysis.

High-frequency physiologic data were collected at the frequency of once per minute from the time of admission until discharge. A total of 5 physiologic characteristics were used in the analysis, including heart rate (HR), respiratory rate (RR), systolic blood pressure (SBP), diastolic blood pressure (DBP), and mean blood pressure (MBP). These vital signs were selected because they were always obtained in all patients admitted to ICUs, and these can be used in identifying organ failure using a minimal set of physiologic data. The main outcome of organ failure was the failure of at least one organ or system (cardiovascular, liver, respiratory, coagulation, central nervous system, or renal), and cases were defined as patients meeting an increase in a serially calculated SOFA criteria by at least one score within a 24-hour rolling window, from admission till discharge. Event time (t_{onset}) was recorded as the earliest time stamp of every occurrence of organ failure, as defined independently by SOFA for each patient. Patients without 24 hours of physiological data before the t_{onset} (for cases and controls) were excluded. To normalize our prediction time horizon, we created relative alignments of time windows, pivoted to t_{onset} . This study was approved by The University of Tennessee Health Science Center and Methodist Le Bonheur Hospital Institutional Review Boards, and it was performed in compliance with the ethical principles for medical research involving human subjects from the Declaration of Helsinki.

Figure 1. Consolidated Standards for Reporting of Trials diagram describing the study cohort.

Feature Extraction and Feature Selection

We utilized Python libraries for extracting features from each of the physiological data streams, including a combination of temporal, frequency, and statistical features ([Multimedia Appendix 1](#)) [11]. Features were derived from six overlapping 3-hour time intervals, with a stride of 1 hour, from 1 to 4 hours to 6 to 9 hours before organ failure, so that we could build predictive models at different times before organ failure. For controls (patients with SCD not developing organ failure), we extracted these same features from 3-hour intervals before a random time period (identified by time stamp) during their ICU stay. The detailed feature extraction for organ failure cases and controls is shown in [Multimedia Appendix 2](#). We performed feature selection using multiple null hypothesis testing, using the Benjamini-Yekutieli procedure and Mann-Whitney U test. Finally, we ranked the features using random forest (RF) [12] feature importance algorithm and applied various feature thresholds to select subsets of features that were most discriminatory among cases and controls.

Machine Learning Algorithms

Multilayer perceptron (MLP), support vector machine (SVM), RF, and logistic regression (LR) methods were used for building classification models. These methods were adopted because of their successful applications to medical datasets for disease classification [13-18].

An MLP is a deep, feed-forward neural network comprising an input layer, an output layer, and at least two or more hidden layers [19]. MLP has been used in a variety of applications, including electroencephalogram signal classification [15], heart disease diagnosis [20], ovarian tumor classification [21], and continuous speech recognition [22]. The MLP architecture used in this study comprised five hidden layers of 512, 256, 128, 64, and 16 neurons. We applied batch normalization [23] before activation, using the rectified linear unit. To avoid overfitting,

we further imposed a dropout [24] ratio of 0.3. The output layer performed binary classification using the sigmoid activation, and our loss function used the Adam optimizer.

The SVM is a multivariate machine learning approach for classifying samples through a pattern recognition analysis [25]. SVM aims to find the best hyperplane that separates all data points of one class compared with those of another class. We used a radial basis function as the kernel parameter for model building.

The RF classifier is well suited to the classification of medical data because of the following advantages: (1) it performs embedded feature selection, (2) it incorporates interactions between predictors, (3) it allows the algorithm to learn both simple and complex classification functions accurately, and (4) it is applicable to both binary and multiclass classification tasks [12]. On the basis of the out-of-bag error [26], we identified 500 trees in the RF models as the optimal number of trees.

LR can be used as a machine learning method used to predict the value of a binary variable based on its relationship with predictor variables [27]. The P value for statistical testing of variable significance for *inclusion-in* and *exclusion-from* the model was set to $P=.05$ and $P=.10$, respectively, and LIBLINEAR solver was used for the optimization function.

Statistical Analysis and Machine Learning Framework

Python scikit-learn machine learning library [28] was used for calculating descriptive statistical measures, for feature selection, and for building machine learning classifiers. Bivariate LR, bootstrap, and Bayesian bootstrap (adjusted for weights) were used to assess the predictability of the features generated for predicting organ failure [29]. We used nonparametric Kruskal-Wallis statistical tests to analyze the difference among the five physiological signals (DBP, SBP, MBP, HR, and RR).

Each of these five signals had six feature measurements, namely, mean, energy ratio, fast Fourier transform, linear trend, quantile, and continuous wavelet transform.

Cross Validation

Models were developed from the distinct time intervals and tested on patients who were not included in the training of the model. For model selection and accuracy estimation, we used 5-fold cross validation [30]. This technique divides data into five equal and discrete folds and uses four folds for model generation, whereas predictions are generated and evaluated using the remaining single fold. This step is subsequently repeated five times, so each fold is tested against the other four

fold. We further ran each of these 5-fold cross-validation models 10 times by shuffling the data in each iteration and averaged the performance metrics from all iterations to reduce bias.

Results

Patient Characteristics

Tables 1 and 2 outline the descriptive-level characteristics of demographics and clinical characteristics of 163 encounters. Four patients with organ failure at admission were omitted from further analysis and their data are not shown in Table 1.

Table 1. Encounter-level demographics and principal diagnosis of patients in the overall cohort (n=163).

Variable ^a	Total cohort	Organ failure (yes)	Organ failure (no)	P value
Total sample, n (%)	163 (100.0)	37 (22.7)	126 (77.3)	N/A ^b
Age (years), mean (SD)	30.7 (9.8)	35.2 (12.9)	29.3 (8.3)	.01 ^c
Female, n (%)	87 (53.4)	24 (64.9)	63 (50.0)	.11
African American, n (%)	163 (100.0)	37 (100.0)	126 (100.0)	N/A
Admit diagnosis, n (%)^d				
Vaso-occlusive event (pain or acute chest syndrome)	130 (79.8)	23 (62.2)	107 (84.9)	.003 ^c
Nonvaso-occlusive crises pain	7 (4.3)	4 (10.8)	3 (2.4)	.05 ^c
Infection/sepsis	7 (4.3)	2 (5.4)	5 (4.0)	.66
Other ^e	19 (11.6)	8 (21.6)	11 (8.7)	.03 ^c

^aFor continuous variables, independent *t* test was used; for categorical variables, Chi-square test of independence was used. Fisher exact test was used for variables with cell counts of less than 5.

^bN/A: Not applicable.

^cStatistically significant at $P=.05$.

^dThe admit diagnoses are based on the International Classification of Diseases, Tenth Revision, Clinical Modification codes at the admission time.

^eOther category includes respiratory distress, sickle cell disease without crisis, diabetes complications (diabetic ketoacidosis/hyperglycemia), pneumonia, myocardial infarction, hematemesis, cough, and deep venous thrombosis.

The mean age of the patient encounters in the cohort was 30.7 years (9.8 years); all patients were African American, and there were more females, 53.4% (87/163), than males, 46.6% (76/163). Admit diagnoses of vaso-occlusive event (pain or acute chest syndrome; 130/163, 79.8%), nonvaso-occlusive crises pain (7/163, 4.3%), and infection/sepsis (7/163, 4.3%)

were common (Table 1). Both vaso-occlusive and nonvaso-occlusive events significantly altered between the patient with *organ failure* and *no organ failure* groups. Patients with organ failure had longer hospital stays (3.2 additional hospital days; $P=.03$) than controls, and they had higher severity of illness ($P<.001$) and risk of mortality ($P<.001$; Table 2).

Table 2. Encounter-level clinical characteristics of patients in the overall cohort (n=163).

Variable ^a	Total cohort	Organ failure (yes)	Organ failure (no)	P value
Encounters, n (%)	163 (100.0)	37 (22.7)	126 (77.3)	__ ^b
Encounter through emergency department, n (%)	134 (82.2)	33 (89.2)	101 (80.2)	.33
Length of stay (days), mean (SD)	5.3 (4.7)	7.7 (8.4)	4.5 (2.5)	.03 ^c
APR-DRG^d severity of illness, n (%)				<.001
Minor	52 (37.1)	7 (19.4)	45 (43.2)	
Moderate	40 (28.6)	7 (19.4)	33 (31.7)	
Major	40 (28.6)	15 (41.7)	25 (24.0)	
Extreme	8 (5.7)	7 (19.4)	1 (1.0)	
APR-DRG risk of mortality, n (%)				<.001
Minor	93 (66.4)	14 (38.9)	79 (76.0)	
Moderate	26 (18.6)	6 (16.7)	20 (19.2)	
Major	12 (8.6)	8 (22.2)	4 (3.9)	
Extreme	9 (6.4)	8 (22.2)	1 (1.0)	
Discharge disposition, n (%)				<.001
Home	147 (90.2)	24 (64.9)	123 (97.6)	
Hospice or home health services	7 (4.3)	5 (13.5)	2 (1.6)	
Expired	5 (3.1)	5 (13.5)	0 (0.0)	
Other	4 (2.5)	3 (8.1)	1 (0.8)	

^aFor continuous variables, independent *t* test was used; for categorical variables, Chi-square test of independence was used. Fisher exact test was used for variables with cell counts of less than 5.

^bNot available.

^c*P*<.05.

^dAPR-DRG: all patient refined-diagnosis related group.

Feature Selection

Feature selection was performed to reduce the number of features, and the reduced feature set was fed into each of the

classifiers. The sample distribution for each of the six datasets for 3-hour observational periods is given in [Table 3](#). The number of patients varies with the availability of data during each time window.

Table 3. Sample distribution and number of features for each dataset using organ failure.

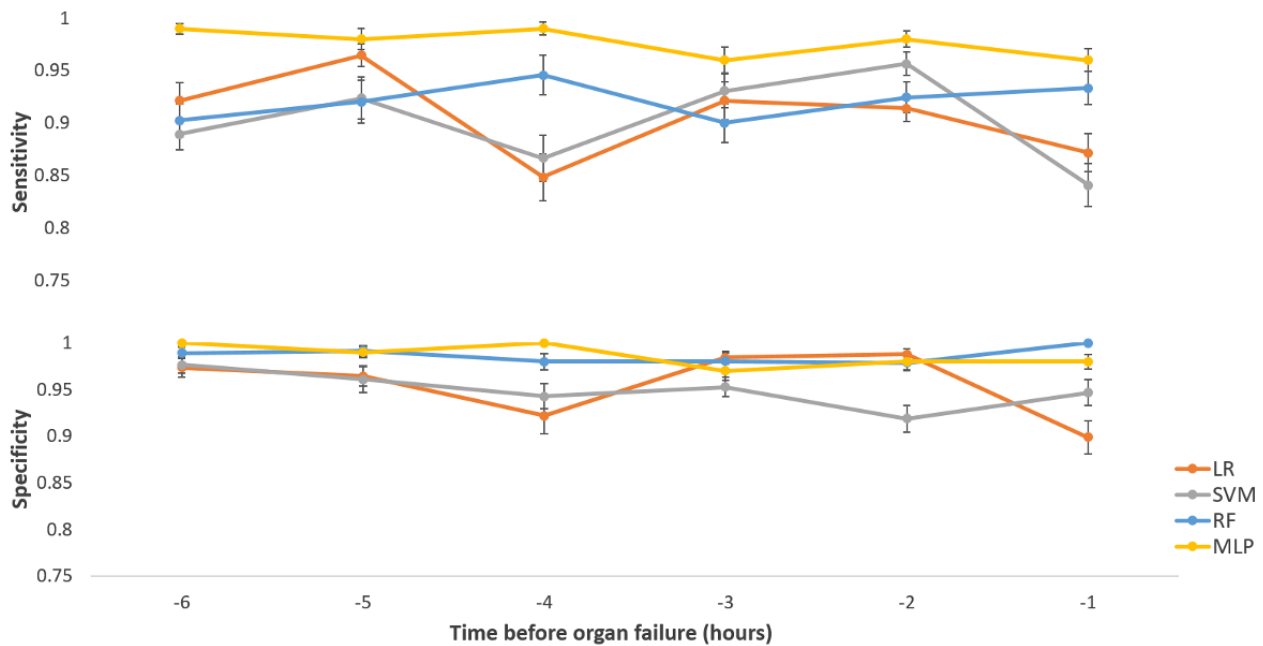
Interval before organ failure onset (hours)	Organ failure events, n	Control events, n
6-9	27	97
5-8	22	90
4-7	22	89
3-6	29	83
2-5	29	88
1-4	29	79

Model Performance

The average sensitivity and specificity from all models for each of the six time periods are given in [Figure 2](#). The MLP model achieved an average sensitivity and specificity of 96% and 98%,

respectively, an hour before organ failure ([Figure 2](#) and [Multimedia Appendix 3](#)). Among the four classifiers, MLP performed better than SVM, LR, and RF in predicting SCD with organ failure.

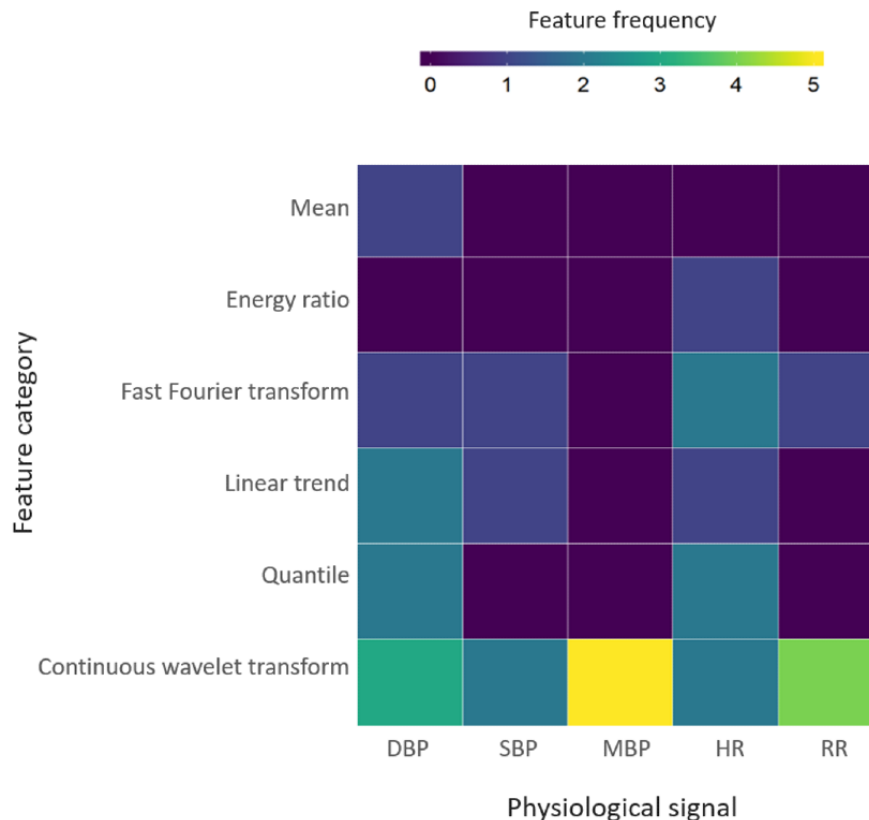
Figure 2. Average sensitivity, specificity for support vector machine, random forest, logistic regression, multilayer perceptron, and sickle cell disease models using each of the six 3-hour datasets. LR: logistic regression; SVM: support vector machine; RF: random forest; MLP: multilayer perceptron.



RF classifier identified the continuous wavelet transform generated from MBP time series physiologic variable as the most important feature, followed by continuous wavelet transform feature generated from respiratory rate of the time series data. Figure 3 shows the frequency of the top 30 important features generated from the five physiologic signals. Each box in the heat map represents the frequency of a feature (y-axis) generated from the physiological signal (x-axis). The dark purple

color in Figure 3 represents the absence of the feature, whereas the dark yellow color represents the most frequently present feature. The list of features ranked by their importance is given in Multimedia Appendix 4. The description of each feature is presented in Multimedia Appendix 1. The Kruskal-Wallis statistical test found no difference among the features extracted from physiological signals (H statistic=5.029; P=.28, possibly reflecting the relatively small sample size).

Figure 3. Features derived from physiologic signals up to six hours before organ failure. DBP: diastolic blood pressure; SBP: systolic blood pressure; MBP: mean blood pressure; HR: heart rate; RR: respiratory rate.



Discussion

Principal Findings

Acute organ failure is a major challenge in people with SCD, especially among adults experiencing an acute disease complication. The ability of predictive algorithms to identify patients at high risk for organ deterioration by using routinely collected physiological data can provide important early warnings of impending physiological deterioration. Such information can aid clinical decision making and may even eventually be useful in guiding early goal-directed therapy. In this retrospective study, we demonstrated that the machine learning–based prediction models could accurately distinguish patients with SCD at risk for developing organ failure up to 6 hours before the onset. The classifiers and the selected physiologic features may facilitate accurate, unbiased SCD diagnosis and effective treatment, ultimately improving prognosis.

The selection of relevant features involved in SCD with organ failure remains a challenge [31,32]. Therefore, we wanted to find a subset of physiologic features that are sufficiently informative to distinguish between patients with SCD at risk of developing organ failure and those who are not at risk. To extract useful information from continuous physiologic data of patients with SCD and to reduce dimensionality, feature-selection algorithms were systematically investigated. As we have demonstrated in the results, selecting smaller subsets of features allowed for the high performance of our classification models. Salient physiologic markers (such as fast Fourier transform, energy, and continuous wavelet transform) derived from the physiological signals, such as blood pressure, HR, and RR, may precede acute organ failure in patients with SCD, as suggested by the results in this study. A shortcoming of machine learning is that these physiologic markers are neither observable by physicians nor readily interpretable; instead, their benefit is primarily toward the early prediction of impending physiologic deterioration, as well as alerting health care providers of that fact. Further research is needed to understand how to use these alerts to guide the personalized care of patients with SCD.

The high-frequency data were captured at 1-min intervals, and we only studied patients who had at least 24 hours of continuous high-frequency physiologic data available before the time of organ failure onset (identified using SOFA criteria). Patients admitted to the ICU with organ failure were excluded, as were patients without a full 24 hour of preceding data. It is also

possible that some patients who were too sick may not have been connected to the monitors, which may have introduced the selection bias. There is a need for future research to focus on developing models that rely on less data before organ failure.

Other limitations are also important to mention. First, we developed the machine learning model on a small subset of patients, specific to the Mid-South of the United States, potentially reducing generalizability. Moreover, the data were highly imbalanced, with more non–organ failure cases compared with organ failure cases, making data-driven approaches difficult to implement. Missing data elements identified in the data were a major hindrance for model validation; thus, these may have contributed to poor validation in some of the cross-validation folds. Although we included admission/encounter in the machine learning model building if the intervals between admissions were at least 1 month, a patient may be more likely to have organ failure in the subsequent encounter. With larger patient data, in the future, we can restrict events to a single event per patient.

For the purposes of this particular study, the SOFA scores were used only to classify cases and controls and to determine the time of organ failure onset. The machine learning models to distinguish cases and controls were built using a limited set of continuously streaming physiological data. There is an inherent difference in how SOFA data are collected and used versus how data for machine learning were collected and used. Moreover, the time of SOFA scores is inherently delayed, to some unknown degree, and this leads to noise in any predictive model whose goal is dependent on the timing of an event (such as organ failure). As data to compute SOFA scores may be delayed in being entered into the EMR, our proposed machine learning model could be used as an alternative for timely diagnosis. Therefore, additional data are required to develop a more robust and generalizable model.

Conclusions

In conclusion, we showed, as a proof of principle, that machine learning can accurately predict the development of organ failure in ICU patients with SCD up to 6 hours before onset. This finding is significant because it may optimize the early recognition of serious disease complications and allow for the implementation of early interventions. As future plans, we would like to extend this study to develop a multiclass machine learning classification model to predict the type of organ failure from each of the six organ systems as we collect sufficient data from each organ system.

Acknowledgments

The authors wish to acknowledge Brian Williams, Michael Younker, and Don MacMillan for their assistance in the data collection. The authors would also like to thank the Office of Scientific Writing at the University of Tennessee Health Science Center for copyediting and proofreading the paper. The authors gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp used for this research.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Temporal, frequency, and statistical features.

[\[XLSX File \(Microsoft Excel File\), 17 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Illustration of feature extraction for cases and controls. Each dashed box represent 3-hour physiological data. For cases, the green and blue arrow represent data from one and two-hour prior to organ failure onset, respectively. For controls, gray arrows represent a random 3-hour physiological data.

[\[PNG File , 79 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Average sensitivity, specificity for multi-layer perceptron (MLP), support vector machine (SVM), random forest (RF), and logistic regression (LR).

[\[XLSX File \(Microsoft Excel File\), 11 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

The list of features ranked by their importance.

[\[XLSX File \(Microsoft Excel File\), 10 KB-Multimedia Appendix 4\]](#)

References

1. Powars DR. Sick cell anemia and major organ failure. *Hemoglobin* 1990;14(6):573-598. [doi: [10.3109/03630269009046967](https://doi.org/10.3109/03630269009046967)] [Medline: [2101835](https://pubmed.ncbi.nlm.nih.gov/2101835/)]
2. Thein MS, Igbineweka NE, Thein SL. Sick cell disease in the older adult. *Pathology* 2017 Jan;49(1):1-9. [doi: [10.1016/j.pathol.2016.10.002](https://doi.org/10.1016/j.pathol.2016.10.002)] [Medline: [27914684](https://pubmed.ncbi.nlm.nih.gov/27914684/)]
3. Powars DR, Chan LS, Hiti A, Ramicone E, Johnson C. Outcome of sickle cell anemia: a 4-decade observational study of 1056 patients. *Medicine (Baltimore)* 2005 Nov;84(6):363-376 [FREE Full text] [doi: [10.1097/01.md.0000189089.45003.52](https://doi.org/10.1097/01.md.0000189089.45003.52)] [Medline: [16267411](https://pubmed.ncbi.nlm.nih.gov/16267411/)]
4. Johnson AE, Ghassemi MM, Nemati S, Niehaus KE, Clifton DA, Clifford GD. Machine learning and decision support in critical care. *Proc IEEE Inst Electr Electron Eng* 2016 Feb;104(2):444-466 [FREE Full text] [doi: [10.1109/JPROC.2015.2501978](https://doi.org/10.1109/JPROC.2015.2501978)] [Medline: [27765959](https://pubmed.ncbi.nlm.nih.gov/27765959/)]
5. Kamaleswaran R, Akbilgic O, Hallman MA, West AN, Davis RL, Shah SH. Applying artificial intelligence to identify physiometers predicting severe sepsis in the PICU. *Pediatr Crit Care Med* 2018 Oct;19(10):e495-e503. [doi: [10.1097/PCC.0000000000001666](https://doi.org/10.1097/PCC.0000000000001666)] [Medline: [30052552](https://pubmed.ncbi.nlm.nih.gov/30052552/)]
6. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med* 1985 Oct;13(10):818-829. [Medline: [3928249](https://pubmed.ncbi.nlm.nih.gov/3928249/)]
7. Marshall JC, Cook DJ, Christou NV, Bernard GR, Sprung CL, Sibbald WJ. Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome. *Crit Care Med* 1995 Oct;23(10):1638-1652. [doi: [10.1097/00003246-199510000-00007](https://doi.org/10.1097/00003246-199510000-00007)] [Medline: [7587228](https://pubmed.ncbi.nlm.nih.gov/7587228/)]
8. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The third international consensus definitions for sepsis and septic shock (sepsis-3). *J Am Med Assoc* 2016 Feb 23;315(8):801-810 [FREE Full text] [doi: [10.1001/jama.2016.0287](https://doi.org/10.1001/jama.2016.0287)] [Medline: [26903338](https://pubmed.ncbi.nlm.nih.gov/26903338/)]
9. Vincent JL, Moreno R, Takala J, Willatts S, de Mendonça A, Bruining H, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. *Intensive Care Med* 1996 Jul;22(7):707-710. [doi: [10.1007/bf01709751](https://doi.org/10.1007/bf01709751)] [Medline: [8844239](https://pubmed.ncbi.nlm.nih.gov/8844239/)]
10. van Wyk F, Khojandi A, Mohammed A, Begoli E, Davis RL, Kamaleswaran R. A minimal set of physiometers in continuous high frequency data streams predict adult sepsis onset earlier. *Int J Med Inform* 2019 Feb;122:55-62. [doi: [10.1016/j.ijmedinf.2018.12.002](https://doi.org/10.1016/j.ijmedinf.2018.12.002)] [Medline: [30623784](https://pubmed.ncbi.nlm.nih.gov/30623784/)]
11. Christ M, Braun N, Neuffer J, Kempa-Liehr AW. Time Series Feature Extraction on basis of Scalable Hypothesis tests (tsfresh – A Python package). *Neurocomputing* 2018;307:72-77. [doi: [10.1016/j.neucom.2018.03.067](https://doi.org/10.1016/j.neucom.2018.03.067)]
12. Bishop CM. *Pattern Recognition and Machine Learning*. Cham: Springer; 2007.
13. Statnikov A, Wang L, Aliferis CF. A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. *BMC Bioinformatics* 2008 Jul 22;9:319 [FREE Full text] [doi: [10.1186/1471-2105-9-319](https://doi.org/10.1186/1471-2105-9-319)] [Medline: [18647401](https://pubmed.ncbi.nlm.nih.gov/18647401/)]
14. Tiwari AK. Machine learning based approaches for prediction of Parkinson's disease. *Mach Learn Appl An Int J* 2016;3(2):33-39. [doi: [10.5121/mlaj.2016.3203](https://doi.org/10.5121/mlaj.2016.3203)]

15. Orhan U, Hekim M, Ozer M. EEG signals classification using the K-means clustering and a multilayer perceptron neural network model. *Expert Syst Appl* 2011;38(10):13475-13481. [doi: [10.1016/j.eswa.2011.04.149](https://doi.org/10.1016/j.eswa.2011.04.149)]
16. Salem AM, Revett K, El-Dahshan EA. Machine Learning in Electrocardiogram Diagnosis. In: Proceedings of the 2009 International Multiconference on Computer Science and Information Technology. 2009 Presented at: IMCSIT'09; October 12-14, 2009; Mragowo, Poland. [doi: [10.1109/IMCSIT.2009.5352689](https://doi.org/10.1109/IMCSIT.2009.5352689)]
17. Mena L, Gonzalez JA. Symbolic one-class learning from imbalanced datasets: application in medical diagnosis. *Int J Artif Intell Tools* 2009;18(2):273-309. [doi: [10.1142/s0218213009000135](https://doi.org/10.1142/s0218213009000135)]
18. Magoulas GD, Prentza A. Machine learning in medical applications. In: Paliouras G, Karkaletsis V, Spyropoulos CD, editors. *Machine Learning and Its Applications*. Berlin, Heidelberg: Springer; 2001:300-307.
19. Itoh N, Hirashiki KI, Terada T, Kikuta M, Ishizuka SI, Koto T, et al. High sensitivity 900-MHz ISM band transceiver. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* 2005 Feb 1;E88-A(2):498-506. [doi: [10.1093/ietfec/e88-a.2.498](https://doi.org/10.1093/ietfec/e88-a.2.498)]
20. Yan H, Jiang Y, Zheng J, Peng C, Li Q. A multilayer perceptron-based medical decision support system for heart disease diagnosis. *Expert Syst Appl* 2006 Feb;30(2):272-281. [doi: [10.1016/j.eswa.2005.07.022](https://doi.org/10.1016/j.eswa.2005.07.022)]
21. Antal P, Fannes G, Timmerman D, Moreau Y, de Moor B. Bayesian applications of belief networks and multilayer perceptrons for ovarian tumor classification with rejection. *Artif Intell Med* 2003 Sep;29(1-2):39-60. [doi: [10.1016/s0933-3657\(03\)00053-8](https://doi.org/10.1016/s0933-3657(03)00053-8)]
22. Morgan N, Bourlard H. Continuous Speech Recognition Using Multilayer Perceptrons With Hidden Markov Models. In: Proceedings of the International Conference on Acoustics, Speech, and Signal Processing. 1990 Presented at: ICASSP'90; April 3-6, 1990; Albuquerque, NM, USA, USA. [doi: [10.1109/ICASSP.1990.115720](https://doi.org/10.1109/ICASSP.1990.115720)]
23. Ren J, Shen X, Lin Z, Mech R, Foran DJ. Personalized Image Aesthetics. In: Proceedings of the 2017 IEEE International Conference on Computer Vision. 2017 Presented at: ICCV'17; October 22-29, 2017; Venice, Italy. [doi: [10.1109/ICCV.2017.76](https://doi.org/10.1109/ICCV.2017.76)]
24. Agarwal A, Negahban S, Wainwright MJ. Noisy matrix decomposition via convex relaxation: Optimal rates in high dimensions. *Ann Statist* 2012;40(2):1171-1197. [doi: [10.1214/12-aos1000](https://doi.org/10.1214/12-aos1000)]
25. Zhang H, Gu C. CiteSeerX. 2006. Support Vector Machines versus Boosting URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.117.343&rep=rep1&type=pdf> [accessed 2020-02-25]
26. Kajdanowicz T, Kazienko P. Boosting-based multi-label classification. *J Univers Comput Sci* 2013;19(4):502-520. [doi: [10.3217/jucs-019-04-0502](https://doi.org/10.3217/jucs-019-04-0502)]
27. Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. *J Biomed Inform* 2002;35(5-6):352-359 [FREE Full text] [doi: [10.1016/s1532-0464\(03\)00034-0](https://doi.org/10.1016/s1532-0464(03)00034-0)] [Medline: [12968784](https://pubmed.ncbi.nlm.nih.gov/12968784/)]
28. Martinkova N, Nová P, Sablina OV, Graphodatsky AS, Zima J. Karyotypic relationships of the Tatra vole (*Microtus taticus*). *Folia Zool* 2004;53(3):279-284 [FREE Full text]
29. Rubin DB. The Bayesian Bootstrap. *Ann Statist* 1981 Jan;9(1):130-134. [doi: [10.1214/aos/1176345338](https://doi.org/10.1214/aos/1176345338)]
30. Statnikov A, Aliferis CF, Tsamardinos I, Hardin D, Levy S. A comprehensive evaluation of multicategory classification methods for microarray gene expression cancer diagnosis. *Bioinformatics* 2005 Mar 1;21(5):631-643. [doi: [10.1093/bioinformatics/bti033](https://doi.org/10.1093/bioinformatics/bti033)] [Medline: [15374862](https://pubmed.ncbi.nlm.nih.gov/15374862/)]
31. Yeruva SL, Paul Y, Oneal P, Nourai M. Renal failure in sickle cell disease: prevalence, predictors of disease, mortality and effect on length of hospital stay. *Hemoglobin* 2016 Sep;40(5):295-299 [FREE Full text] [doi: [10.1080/03630269.2016.1224766](https://doi.org/10.1080/03630269.2016.1224766)] [Medline: [27643740](https://pubmed.ncbi.nlm.nih.gov/27643740/)]
32. Al Khawaja SA, Ateya ZM, Al Hammam RA. Predictors of mortality in adults with Sickle cell disease admitted to intensive care unit in Bahrain. *J Crit Care* 2017 Dec;42:238-242. [doi: [10.1016/j.jcrc.2017.07.032](https://doi.org/10.1016/j.jcrc.2017.07.032)] [Medline: [28797896](https://pubmed.ncbi.nlm.nih.gov/28797896/)]

Abbreviations

- DBP:** diastolic blood pressure
- EMR:** electronic medical record
- HR:** heart rate
- ICU:** intensive care unit
- LR:** logistic regression
- MBP:** mean blood pressure
- MLP:** multilayer perceptron
- RF:** random forest
- RR:** respiratory rate
- SBP:** systolic blood pressure
- SCD:** sickle cell disease
- SOFA:** Sequential Organ Failure Assessment
- SVM:** support vector machine
- t_{onset}:** event time

Edited by M Focsa; submitted 13.05.19; peer-reviewed by S Creary, S Curtis, C Gao; comments to author 02.06.19; revised version received 18.08.19; accepted 28.01.20; published 13.05.20

Please cite as:

Mohammed A, Podila PSB, Davis RL, Ataga KI, Hankins JS, Kamaleswaran R

Using Machine Learning to Predict Early Onset Acute Organ Failure in Critically Ill Intensive Care Unit Patients With Sickle Cell Disease: Retrospective Study

J Med Internet Res 2020;22(5):e14693

URL: <https://www.jmir.org/2020/5/e14693>

doi: [10.2196/14693](https://doi.org/10.2196/14693)

PMID: [32401216](https://pubmed.ncbi.nlm.nih.gov/32401216/)

©Akram Mohammed, Pradeep S B Podila, Robert L Davis, Kenneth I Ataga, Jane S Hankins, Rishikesan Kamaleswaran. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 13.05.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.