Using Machine Learning for Early Prediction of Cardiogenic Shock in Patients with Acute Heart Failure

Faisal Rahman (✉ frahman480@outlook.com)
Johns Hopkins University

Noam Finkelstein
Johns Hopkins University

Anton Alyakin
Johns Hopkins University

Nisha Gilotra
Johns Hopkins University

Jeff Trost
Johns Hopkins University

Steven Schulman
Johns Hopkins University

Suchi Saria
Johns Hopkins University

Research Article

Keywords: acute decompensated heart failure, cardiogenic shock, machine learning, risk prediction,

DOI: https://doi.org/10.21203/rs.3.rs-453102/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

Objective: Despite technological and treatment advancements over the past two decades, cardiogenic shock (CS) mortality has remained between 40-60%. A number of factors can lead to delayed diagnosis of CS, including gradual onset and nonspecific symptoms. Our objective was to develop an algorithm that can continuously monitor heart failure patients, and partition them into cohorts of high- and low-risk for CS.

Methods: We retrospectively studied 24,461 patients hospitalized with acute decompensated heart failure, 265 of whom developed CS, in the Johns Hopkins Healthcare system. Our cohort identification approach is based on logistic regression, and makes use of vital signs, lab values, and medication administrations recorded during the normal course of care.

Results: Our algorithm identified patients at high-risk of CS. Patients in the high-risk cohort had 10.2 times (95% confidence interval 6.1-17.2) higher prevalence of CS than those in the low-risk cohort. Patients who experienced cardiogenic shock while in the high-risk cohort were first deemed high-risk a median of 1.7 days (interquartile range 0.8 to 4.6) before cardiogenic shock diagnosis was made by their clinical team.

Conclusions: This risk model was able to predict patients at higher risk of CS in a time frame that allowed a change in clinical care. Future studies need to evaluate if CS analysis of high-risk cohort identification may affect outcomes.

Introduction:

Mortality rates for cardiogenic shock (CS) patients have remained mostly unchanged at 40–60% over the last two decades[1–6], despite the introduction of important innovations in cardiac care such as the cardiac intensive care unit (CICU) and mechanical support devices. The task of improving outcomes for CS is increasingly urgent, as the prevalence of CS and acute decompensated heart failure (ADHF) has been increasing for decades[7, 8]. While the factors that lead to poor CS outcomes are not well-understood, there is some indication that they may be due to the difficulty of diagnosing and treating the condition early enough to reverse patients’ downward clinical trajectories[9, 10].

Early diagnosis of CS is complicated by the fact that it can present gradually, and with nonspecific symptoms. In addition, patients who develop CS in-hospital may not be under the direct care of a cardiologist. For these reasons, it has been difficult to test the hypothesis that early diagnosis would lead to improved outcomes. Data in electronic health records (EHRs) of a patient may provide a valuable source of information as to whether a patient is at high-risk for CS. However, for practical reasons clinicians cannot fully synthesize and make use of the high volume of data collected on each patient over time. On the other hand, without the aid of clinical expertise, computational algorithms using EHRs are insufficient to the task of diagnosing CS on their own.
In this work, we present a computational method for identifying patients whose EHRs indicate that they should be treated as being at high-risk for CS. Our method makes use of vital signs, lab values, and medication administrations throughout the patient's hospital stay to ascertain risk status using a logistic regression model, which will be described in detail below. Our results open the door to future study of whether early diagnosis and treatment of CS can improve outcomes.

**Methods:**

**Data**

Johns Hopkins Healthcare system is composed of multiple hospitals and outpatient centers. Our study included patients admitted to Johns Hopkins Hospital between 01/01/2016 and 03/31/2018, Howard County General Hospital between 02/01/2014 and 01/29/2019 and Johns Hopkins Bayview Medical Center between 01/01/2016 and 03/31/2018. The study was approved by the Institutional Review Board of Johns Hopkins Medicine with waiver of consent. All data is obtained from Epic Systems Corp. (Verona, WI) EHR software. Patients under 18 years of age were excluded from the cohort. All methods were carried out in accordance with relevant guidelines and regulations.

**Clinical Cohort**

Patients with ADHF were included for monitoring. However, they were excluded if they met criteria for ADHF after surgery or after development of CS or other shock. Patients that developed shock within 6 hours of admission were also excluded due to the assumption that the patient likely was in a peri-shock state at the time of admission. All patients who were discharged to hospice with end-stage cardiomyopathy were excluded if they never met the criteria for CS.

For both model development and validation, it was necessary to identify (label) patients that went into CS along with time of onset. Patients were labeled as having CS if they had hypotension (systolic blood pressure ≤ 90 mmHg or mean arterial pressure ≤ 65 mmHg) that required inotropic therapy (dopamine, dobutamine, milrinone or norepinephrine) or mechanical circulatory support, and evidence of end-organ failure due to a cardiac cause of shock\[10–12\]. The diagnosis was retrospectively adjudicated by a study team cardiologist and required the clinical team to arrive at the diagnosis of CS based on EHR documentation suggesting that the cause of shock was cardiac in origin. The time of CS onset was assumed to be the time of order of inotropic therapy or mechanical support but practically this ended up being time of inotropic therapy as all patients who received mechanical support had gotten initial inotropic therapy to stabilize. During validation, each patient in the dataset set aside for this purpose, was evaluated every 6 hours (sequentially) during hospitalization to see if their risk prediction hit threshold for high-risk.

**Real-time Flagging of Likely ADHF Patients**

An accurate diagnosis of ADHF may not be made immediately or be recorded in physician notes at the time of treatment, which most commonly involves initiation of diuretics. We therefore created a set of
criteria to use as a proxy for an ADHF diagnosis, to enable timely, automatic identification and monitoring of patients with likely ADHF (Supplemental Fig. 1; Supplemental Methods). (17) We note that this definition is different to some of the previously proposed definitions used in some studies of EHR[13–16] but a similar definition has also been tested showing a high-sensitivity[15]. This is for two reasons. First, some existing definitions rely on information not routinely captured in the EHR, and it was important that we be able to automatically identify ADHF patients. Second, for our application the cost of missing ADHF patients far exceeds the cost of including patients who do not have ADHF, as missed ADHF patients would not be monitored for CS. Thirdly, any model would need to be able to identify ADHF immediately for prospective monitoring, and therefore patient notes are limited in being able to timely identify ADHF. Patients were monitored until they developed CS, other type of shock (referred to as “other shock”), went to the operating room, or were discharged. Monitoring for these outcomes began with the first dose of intravenous diuretics.

Variables

Clinical inpatient data from admission to discharge was extracted from the EHR into our PostgreSQL database. Data extracted included demographics, vital signs, laboratory results, medications, past medical history, admitting service, hospital unit, discharge disposition, and discharge diagnoses ICD-10 codes (see Supplemental Methods).

We explored two approaches to missing data imputation. In the first approach, we used a multi-output Gaussian process model[17]. Gaussian processes are flexible regression models, and multi-output Gaussian processes can learn the correlations between the various dynamic variables. We use this model to estimate the value of the dynamic variables as a function of all previously observed data at the time of prediction.

In the second approach, we use a simple last-value-carry-forward strategy. For each dynamic feature, we use the last observed value as an estimate of the value at the time of prediction. If no observation was made, we experimented with different imputation strategies, including zeroes, medians, means, Multiple Imputation by Chained Equations (MICE) imputation strategy(20), and nonparametric k-nearest-neighbors approach. The K-nearest-neighbors’ approach provided the most consistent results and was used for imputation.

Missing data is part of real-life EHR data, and accommodating missing data using imputations methods is important for clinical usefulness of any risk-prediction model. Therefore, no patients were excluded from the analysis due to missing data. In addition, the validation of the model, was performed prospectively using K-nearest-neighbors imputation for missing data and reflects how the model would function in a real-life deployment.

Model Development

Our model sought to identify high- and low-risk cohorts. We treat cohort membership as a binary classification problem, with ADHF patients within 24 hours of developing CS considered to be high-risk,
and ADHF patients that do not develop CS within 24 hours as low-risk.

During model development, we chose a single time point at random for each patient who never developed any form of shock. Amongst patients who developed shock, we used two sample points - one within 24 hours of the onset of shock, and the other beforehand. Using these two samples violates the independence assumption under which the model is learned but was necessary due to the limited CS cases. However, we determined that it was important to have examples of patients who would go on to develop shock both within and outside the classification window.

We considered using, and evaluated, several machine learning classification algorithms (Supplemental Table 1). The performances of all the algorithms were similar with logistic regression performing most consistently when evaluated using odds ratios, sensitivity, specificity, positive and negative predictive values, accuracy and area under the receiver operating curve (AUC/C-statistic).

For each model we evaluated, hyperparameter tuning was performed by means of a five-fold cross-validation[18]. Variable selection was performed by means of cross-validation and L1 regularization, a technique that encodes a preference for disregarding unimportant features. Ultimately, our choice of binary classification model was made based on cross-validation as well. The performance of the model was evaluated using a second validation dataset, which was separate from the development dataset and was not seen during model development.

**Method Evaluation**

We evaluated our method in two ways. First, we evaluated the statistical properties of our model on the model development dataset, and on the separate validation dataset. This evaluation provides a quantitative assessment of the different properties of the two cohorts.

Next, we examined whether and how the knowledge that a patient belongs to the high-risk cohort is actionable, i.e., whether and how treatment and diagnosis might change in the face of this knowledge. In some cases, clinicians may have already been aware the patient was high risk, in which case little would change. But in other cases, this designation may have induced a consult with a cardiologist, or a beneficial change in treatment regime.

To evaluate actionability, we randomly selected 50 patients designated as high-risk who did develop CS, and 50 who did not. These patients were selected using cross-validation techniques (19), such that none were in the model development data for the model used for their classification. We did not make changes to the model based on these evaluations.

Each true positive case was reviewed from the time the model first designated the case as high-risk to time of CS diagnosis by the clinical team based on EHR documentation. These patients were categorized into one of three groups (Supplemental Methods):
(a) Patients that may have received treatment that was potentially harmful or contributed to deterioration (Group A)

(b) Patients for whom more tailored non-invasive or invasive therapies may be available (Group B)

(c) Patients who are already known by the team to be critically ill and high-risk of CS where an high-risk identification does not provide any new information (Group C).

**Results:**

There were 24,461 cases identified as ADHF of which 265 were validated cases of CS. The in-hospital mortalities of the ADHF and CS cases were 4% and 34%, respectively. Table 1 shows patient baseline characteristics.
Table 1
Baseline characteristics of patients classified as acute decompensated heart failure and cardiogenic shock.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Acute decompensated heart failure</th>
<th>Cardiogenic shock</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>24,196</td>
<td>265</td>
</tr>
<tr>
<td>Age</td>
<td>67.6 (15.6)</td>
<td>64.3 (15.5)</td>
</tr>
<tr>
<td>Sex (male)</td>
<td>12,327 (51%)</td>
<td>118 (45%)</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>30.6 (9.3)</td>
<td>29.6 (7.9)</td>
</tr>
<tr>
<td>Mortality</td>
<td>1054 (4.3%)</td>
<td>89 (33.5%)</td>
</tr>
</tbody>
</table>

**Study site**

<table>
<thead>
<tr>
<th>Study site</th>
<th>ADHF (n)</th>
<th>CS (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Johns Hopkins Hospital</td>
<td>9242 (38.1%)</td>
<td>141 (53.2%)</td>
</tr>
<tr>
<td>Bayview Medical Center</td>
<td>5200 (21.5%)</td>
<td>82 (30.9%)</td>
</tr>
<tr>
<td>Howard County General Hospital</td>
<td>9754 (40.3%)</td>
<td>42 (15.8%)</td>
</tr>
</tbody>
</table>

**Comorbidities**

<table>
<thead>
<tr>
<th>Comorbidity</th>
<th>ADHF (n)</th>
<th>CS (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypertension</td>
<td>15111 (62.5%)</td>
<td>132 (49.8%)</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>8000 (33.1%)</td>
<td>92 (34.7%)</td>
</tr>
<tr>
<td>Coronary artery disease</td>
<td>4920 (20.3%)</td>
<td>68 (25.7%)</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>4744 (19.6%)</td>
<td>87 (32.8%)</td>
</tr>
<tr>
<td>Chronic renal disease</td>
<td>2777 (11.4%)</td>
<td>41 (15.4%)</td>
</tr>
<tr>
<td>Cardiomyopathy</td>
<td>7557 (31.2%)</td>
<td>134 (50.6%)</td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease</td>
<td>4523 (18.7%)</td>
<td>45 (17.0%)</td>
</tr>
</tbody>
</table>

Variables presented as mean (standard deviation) or number (percentage)

The quantitative performance of the model is shown in Fig. 1, Supplemental Fig. 2 and Table 2. Based on cross-validation performance, we chose to classify an individual as high-risk when the output of the logistic model was greater than 0.1, balancing the sensitivity, specificity, positive predictive value and negative predictive value. At this threshold, 58% of cardiogenic shock patients were in the high-risk cohort (i.e. there was a 58% sensitivity), and 88% of ADHF patients who did not experience cardiogenic shock were in the low-risk cohort (i.e. an 88% specificity). These numbers correspond to 10.2 times (95% confidence interval 6.1–17.2) higher prevalence of CS in the high-risk cohort than in the low-risk cohort.
Table 2
Risk prediction model sensitivity, specificity, negative and positive predictive values, and diagnostic odds ratio at different thresholds.

<table>
<thead>
<tr>
<th>Model Development Dataset</th>
<th>Threshold</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>NPV</th>
<th>PPV</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.03</td>
<td>0.78</td>
<td>0.79</td>
<td>0.99</td>
<td>0.04</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.70</td>
<td>0.87</td>
<td>0.99</td>
<td>0.05</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>0.06</td>
<td>0.65</td>
<td>0.89</td>
<td>0.99</td>
<td>0.06</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>0.19</td>
<td>0.39</td>
<td>0.97</td>
<td>0.99</td>
<td>0.13</td>
<td>0.97</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Validation Dataset</th>
<th>Threshold</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>NPV</th>
<th>PPV</th>
<th>Accuracy</th>
<th>Diagnostic OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.02</td>
<td>0.88</td>
<td>0.55</td>
<td>0.99</td>
<td>0.02</td>
<td>0.55</td>
<td>9.1 (4.1–20.1)</td>
</tr>
<tr>
<td></td>
<td>0.06</td>
<td>0.75</td>
<td>0.80</td>
<td>0.99</td>
<td>0.04</td>
<td>0.80</td>
<td>11.9 (6.7–21.7)</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>0.58</td>
<td>0.88</td>
<td>0.99</td>
<td>0.05</td>
<td>0.88</td>
<td>10.2 (6.1–17.2)</td>
</tr>
<tr>
<td></td>
<td>0.20</td>
<td>0.34</td>
<td>0.95</td>
<td>0.99</td>
<td>0.07</td>
<td>0.94</td>
<td>9.7 (5.6–16.8)</td>
</tr>
<tr>
<td></td>
<td>0.30</td>
<td>0.27</td>
<td>0.98</td>
<td>0.99</td>
<td>0.11</td>
<td>0.97</td>
<td>14.6 (8.1–26.6)</td>
</tr>
</tbody>
</table>

NPV = negative predictive value; PPV = positive predictive value; OR = odds ratio; CI = confidence interval

Age, systolic blood pressure, heart rate, temperature, blood urea nitrogen, sodium, oxygen saturation, venous pH, hemoglobin, white blood cell count, hydralazine use, trend of respiratory rate and trend of systolic blood pressure were associated with risk of developing CS (Table 3 and Supplemental Table 2).
### Table 3

Selection of variables that are statistically significant in cardiogenic shock prediction model (see supplemental table for all variables)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.015 (-0.028 to -0.002)</td>
<td>0.028</td>
</tr>
<tr>
<td>Systolic blood pressure</td>
<td>-0.050 (-0.060 to -0.041)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Heart rate</td>
<td>0.020 (0.011 to 0.028)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Temperature</td>
<td>-0.0789 (-0.152 to -0.005)</td>
<td>0.035</td>
</tr>
<tr>
<td>Blood urea nitrogen</td>
<td>0.032 (0.023 to 0.040)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Sodium</td>
<td>-0.059 (-0.090 to -0.028)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Oxygen saturation (SpO₂)</td>
<td>0.078 (0.021 to 0.135)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Venous pH</td>
<td>3.653 (1.040 to 6.266)</td>
<td>0.006</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>0.108 (0.027 to 0.188)</td>
<td>0.009</td>
</tr>
<tr>
<td>White blood cell count</td>
<td>-0.051 (-0.096 to -0.006)</td>
<td>0.025</td>
</tr>
<tr>
<td>Hydralazine</td>
<td>0.826 (0.367 to 1.285)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Respiratory rate trend</td>
<td>0.341 (0.104 to 0.577)</td>
<td>0.005</td>
</tr>
<tr>
<td>Systolic blood pressure trend</td>
<td>0.031 (0.000 to 0.062)</td>
<td>0.049</td>
</tr>
</tbody>
</table>

During validation, CS patients who were categorized as high-risk were identified at a median time of 1.7 days (interquartile range 0.8 to 4.6 days, Fig. 2) prior to CS diagnosis.

On review of true positive cases, 12% of patients had possible inappropriate therapy (Group A), 50% of patients had more tailored therapy options (Group B), and 38% of patients were in Group C. See supplement for an example Group A case ([Supplemental Fig. 3](#)).

The patient location at the time of high-risk categorization and CS diagnosis are shown in [Supplemental Fig. 4](#).

**Discussion:**

In the present study, an algorithm to identify patients that may be at a higher risk of developing future CS was developed. The model can continuously monitor complex patient trajectories in the background. Patients identified as high risk are substantially more likely to develop CS than those in the low-risk cohort. Moreover, high risk patients were identified with lead-time that future studies could use to evaluate if it would alter the course of their treatment and therefore outcomes.
Our method follows in the tradition of identifying risk categories for cardiac patients. Risk prediction models from the Acute Decompensated Heart Failure Registry\cite{19}, Get with the Guidelines-Heart Failure\cite{20}, and other similar retrospective data\cite{6, 21, 22} predict the risk of in-hospital and out-of-hospital mortality at the time of admission. Since these models do not track the dynamic variables of a hospitalized patient with ADHF but rather provide a single snapshot, they are challenging to use to change management or outcomes. Additionally, they are not automated but require clinicians to calculate a score, which increases work burden. Some of these issues are overcome by early warning systems (EWS)\cite{23, 24} such as the modified EWS, Targeted Real-time EWS (TREWS)\cite{24}, which can provide continuous monitoring for adverse outcomes such as cardiac arrest, ICU admission or mortality. Unfortunately, these systems have had variable performance, and alert in close proximity to the outcome event which may not be sufficient to change the trajectory of a patient’s clinical outcome\cite{24}.

We also demonstrate that age, systolic blood pressure, heart rate, temperature, blood urea nitrogen, sodium, oxygen saturation, venous pH, hemoglobin, hydralazine use, trend of respiratory rate and trend of systolic blood pressure are each individually associated with increased risk of developing CS in our model. These predictive variables cannot be directly compared with previous risk prediction models as those typically used single time point measurements at admission and evaluated inpatient or post-hospitalization mortality\cite{19–22, 6}. Whereas certain variables such as hemoglobin, sodium and systolic blood pressure have previously been associated with adverse outcomes in ADHF, to our knowledge this is first time that temperature and trends of vital signs have been shown to be associated with risk of developing CS. In our study, individuals who develop CS are more likely to have a lower temperature which may be related to the peripheral vasoconstriction seen in a low output state. Just as importantly, we also demonstrate that most trends (such as heart rate) are not as pertinent for risk prediction.

**How does our work translate?**

Unlike previous approaches, our method focuses specifically on identifying patients at high risk for CS by continuously monitoring their EHRs over the course of their stay. This method is one of a new generation of technologies that seek to take advantage of the tremendous stores of information in EHRs by combining the strengths of computers and clinicians. Computers have unlimited attention, and can monitor vast quantities of data without tiring, while clinicians have the wisdom to act appropriately when important information is made available to them. We find that a patient characterized as high-risk has a greater than 10 times the odds of developing CS.

Although, this model at a threshold of 0.1 had a positive predictive value of 5%, which at first glance would be concerningly low. However, it is not unusual for certain risk prediction tools to have low positive predictive value\cite{25}, for example the CHA\textsubscript{2}DS\textsubscript{2}-VASc score for atrial fibrillation has been used with a risk score of 2 as significant, which equates to an estimated annual stroke risk of 2.2% per year\cite{26}. Similarly, any risk prediction model for CS is not designed to be used in isolation but provide clinicians an aid in addition to their clinical evaluation.
We do not believe a model like ours would be implemented as a best practice alert. Instead, we see the future of medicine involving a dashboard for clinicians where algorithms monitor patients for risk stratification for poor outcomes during continuous surveillance. This dashboard would act to aid the clinician during their daily decision making for patients.

To that extent, we ensured that our model could provide timely, automated method of entrance for likely ADHF by using intravenous diuretics as a surrogate to diagnose ADHF. Although this is a very sensitive approach, the specificity for ADHF is reduced and therefore reduces the positive predictive value for CS, e.g., if diuretics are administered incorrectly or for other diagnoses such as renal or hepatic disease. However, we believe that in addition to a clinician evaluation, particularly when the trajectory of a patient is unclear, a risk stratification by an algorithm could aid a clinician to vary their level of concern for CS. For example, does a rise in creatinine suggest overdiuresis or low-output state. In such cases, the risk-cohort classification of a patient may prompt the clinical team to seek a cardiology consult, or otherwise alter patient care.

We do believe that our model requires continual improvement and further development. For example, future work by ourselves and other groups could include other variables such as echocardiogram data or other imaging tools, ordering of certain tests etc. could help improve the model further. In addition, a suite of algorithms in a clinician dashboard for presentations such as shortness of breath may evaluate for patients at risk of respiratory failure requiring intubation, renal failure requiring renal replacement therapy, and cardiac failure at risk of CS. These approaches will require collaboration between computer science and clinical teams to help us reach the goal of precision medicine leveraging big data and greater computing skills available in the 21st century.

Risk prediction tools are common in the scientific literature, but few are regularly used and implemented into daily practice[27]. Early risk prediction tools are even more relevant as providers spend increasing amounts of time on administrative and documentation activities with frequent interruptions to patient care, and experience increasing volumes of patient data[28, 29]. As our study shows, the overall inpatient mortality for ADHF is low (4% in our study) unless the patient deteriorates into CS, which increased the mortality by more than 8-fold, demonstrating the need to improve care for these patients.

Although mortality after CS has remained stagnant[1, 2, 4, 5, 3, 6], recent evidence in CS after acute myocardial infarction demonstrates that early, coordinated, aggressive treatment improves inpatient mortality[9]. We believe that similar early, tailored treatment in patients identified to have a high-risk of developing CS may improve outcomes. Using the expert consensus from The Society of Cardiovascular Angiography and Interventions, our algorithm attempts to identify patients in Stage B or early in Stage C of shock in comparison to later stages where outcomes are worse[10]. As recently demonstrated, the later the stage of CS the significantly higher the mortality[10, 30]. To that extent, we demonstrate a model with an AUC of 0.77 where positive cases are identified at a median 1.7 days before the diagnosis of CS was made by the clinical team. This early identification of high-risk patients would provide clinicians time to
initiate individualized, tailored treatment strategies to proactively prevent further decompensation instead of reactively responding after CS has already developed.

Further, we demonstrate that from the time of high-risk categorization by our model to clinical diagnosis of CS approximately 1/8 patients received therapy that may have worsened their clinical status likely due to lack of recognition of worsening ADHF. In addition, in half of patients there was time and opportunity to undertake early tailored treatment, which may include simple approaches such as discontinuation of negative inotropes or appropriate diuretic adjustments, to more invasive options such as pulmonary artery catheter insertion for hemodynamic guided therapy, or early initiation of inotropes that may reduce length of hospitalization, and mortality. In addition, the patients identified as high-risk while in the emergency department or medicine services may be selectively referred for admission to a cardiology team or referred for an early cardiology consult. Although these treatment options exist, it is unclear if identification of cases ~ 1.7 days earlier would change a patient’s clinical trajectory. What requires further study is to evaluate if such an approach reduces hospital morbidity, mortality and length of stay; as demonstrated in the sepsis literature[31]. Therefore, we propose a new paradigm where clinical acumen is combined with our prediction model to risk stratify patients for individualized, patient-specific treatment (Fig. 3).

Our study has several strengths including the several years of data from three hospitals and the use of novel machine learning tools for risk prediction. Limitations include the moderate number of CS cases that limits the number of variables and learning that is possible. Additionally, we use data from one healthcare system; further studies using data from other systems will be required to improve model generalizability and robustness. As discussed above, several variables are not included such as echocardiogram data as these were not accurately available. Finally, EHR data does have several weaknesses including missing data, non-systematic way of data collection for each patient, and erroneous recording by providers. However, this reflects real-world practice and any model that is to be clinically useful needs to account for these weaknesses[32].

In this study, we present a risk prediction tool that can use continuous EHR data monitoring in patients with ADHF to help identification of patients at a high-risk of CS. Early identification of at-risk patients is essential to allow for enough time to change disease trajectory. CS is a state of not only decompensated cardiac failure but also end-organ dysfunction and is therefore a multi-organ disease. Current approaches in CS have been inadequate in improving outcomes. Future intervention studies are needed using this model to observe how early identification and potential effects on treatment strategies may alter patient outcomes.

**Abbreviations**

EHR = electronic health records; CICU = cardiac intensive care unit; ADHF = acute decompensated heart failure, CS = cardiogenic shock;
Declarations

Sources of Funding

Faisal Rahman was supported by a NIH T32 grant (5T32HL007024-43). Suchi Saria has grants from Gordon and Betty Moore Foundation, the National Science Foundation, the Defense Advanced Research Projects Agency (DARPA), and the American Heart Association. Noam Finkelstein was supported by an American Heart Association grant (18A1ML34300000) during this work.

Disclosures

Suchi Saria is the scientific founder and/or advisory board member for Bayesian Health and PatientPing

Author Contributions:

FR: Project conception, development of algorithm, statistical analysis, manuscript preparation

NM: Development of algorithm, statistical analysis, manuscript preparation

AA: Development of algorithm, statistical analysis, manuscript preparation

SSch: Project conception, manuscript preparation

NG: Project conception, manuscript preparation

JT: Project conception, manuscript preparation

SSar: Project conception, development of algorithm, manuscript preparation

References


**Figures**

![Development Dataset](image1)

**AUC = 0.82**

![Validation Dataset](image2)

**AUC = 0.77**

**Figure 1**

Receiver operating characteristic curves for prediction model during model development crossvalidation and with validation dataset. AUC = area under the curve
Figure 2

The boxplot shows the distribution of the difference between the time the model identified the patient at high-risk of cardiogenic shock and the time that the diagnosis of cardiogenic shock was made by the primary clinical team.

Figure 3
Using real-time risk prediction model may offer opportunity for early, aggressive, individualized treatment for high-risk patients thereby potentially reducing mortality associated with cardiogenic shock.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- Supplement3.2021.docx