

Modeling In-Hospital Mortality Among Patients Undergoing Percutaneous Coronary Intervention with Acute Myocardial Infarction Complicated by Cardiogenic Shock Receiving Mechanical Circulatory Support

Nathan C Hurley¹, Bobak J Mortazavi¹⁻³, Sanket S Dhruva^{4,5}, Joseph S Ross^{3,6,7},
Che G Ngufo^{8,9}, Jephtha P Curtis^{3,10}, Harlan M Krumholz^{3,7,10}, Nihar R Desai^{3,10}

¹Department of Computer Science and Engineering, Texas A&M University, College Station, TX, USA; ²Center for Remote Health Technologies and Systems, Texas A&M University, College Station, TX, USA; ³Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, New Haven, CT, USA; ⁴University of California, San Francisco School of Medicine, San Francisco, CA, USA; ⁵Section of Cardiology, Department of Medicine, San Francisco Veterans Affairs Medical Center, San Francisco, CA, USA; ⁶Section of General Internal Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA; ⁷Department of Health Policy and Management, Yale School of Public Health, New Haven, CT, USA; ⁸Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, MN, USA; ⁹Department of Artificial Intelligence and Informatics, Mayo Clinic, Rochester, MN, USA; ¹⁰Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, CT, USA

Correspondence: Bobak J Mortazavi, Email bobakm@tamu.edu

Introduction: Acute myocardial infarction complicated by cardiogenic shock (AMI-CS) is a heterogeneous clinical syndrome associated with substantial morbidity and mortality. We developed a machine learning-based mortality model to identify features and patient subgroups associated with the largest change in mortality risk when evaluating treatment with Impella devices or intra-aortic Balloon Pump (IABP).

Methods: Our cohort comprised 369 sites and 15,796 patient visits to the cardiac catheterization laboratory from the National Cardiovascular Data Registry. We developed XGBoost decision tree models to model mortality with and without treatment variable included (Impella or IABP).

Results: The estimated population mean excess mortality effect of treatment with Impella devices vs IABP was $10.4 \pm 0.8\%$. However, we identified clinical subgroups of 282 patients for whom a decreased risk of mortality was associated with use of Impella as compared with IABP. Those patients were on average younger, presented with higher systolic blood pressure, higher rate of salvage percutaneous coronary intervention, higher initial creatinine, and lower hemoglobin.

Discussion: While Impella devices were associated with higher mortality risk overall, certain clinical profiles were associated with lower risk, illustrating heterogeneity of treatment effects.

Keywords: mechanical circulatory support, intra-aortic balloon pump, impella, machine learning, National Cardiovascular Data Registry

Introduction

Acute myocardial infarction (AMI) complicated by cardiogenic shock (AMI-CS) is associated with substantial morbidity and mortality. However, clinical outcomes among patients with AMI complicated by cardiogenic shock (AMI-CS) have not substantially improved over time,¹⁻³ with estimated mortality as high as 60%.⁴ AMI-CS is a heterogeneous clinical syndrome, encompassing heterogeneous patients with differing degrees of clinical presentation and symptom or condition severity, which may potentially be responsive to differing interventions and



strategies.^{5,6} However, many prior studies may have failed to adequately account for this clinical heterogeneity. The absence of a clinical benefit of a particular strategy or intervention may have been related to suboptimal risk adjustment. Correctly identifying these subgroups may therefore allow for identification of interventions leading to benefit in those subgroups.

One intervention used for these patients is mechanical circulatory support (MCS) devices. Approximately 40% of patients with AMI-CS are treated with MCS devices,⁷ most commonly intra-aortic balloon pumps (IABPs) or Impella devices, which are intravascular microaxial left ventricular assist devices (LVADs).⁷ However, clinical trials comparing IABP and Impella devices with each other have been small, underpowered, and provided inconsistent results.^{8,9} Multiple large observational studies found Impella devices were might be associated with significantly increased mortality risk compared to IABP, and these studies have had disagreement upon risks and benefits.^{1,10–12} More recently, however, the DanGer Shock trial found consistent improvement with microaxial flow pumps. Investigators found that microaxial flow pumps lowered risk of death in STEMI-related cardiogenic shock over standard care, even with adverse events occurring more frequently in this same group of patients.¹³ The investigators of that trial found that their patient population was more homogeneous, likely leading to the identified success of the microaxial flow pump in reducing risk of death. This suggests that patient heterogeneity has been a leading factor towards inconsistent results.

An improved understanding of heterogeneity of patients and associated risk and benefit may be a critical factor to determine the patient subgroups for whom various MCS devices may be associated with improved clinical outcomes among patients with AMI-CS. Understanding this heterogeneity and accounting for it in modeling may help develop hypotheses for prospective randomized evaluation of MCS devices in clinical scenarios where they may be more likely to demonstrate reduced mortality risk. We hypothesized that adjusting for patient characteristics would find the homogeneous subgroups of patients for which risk adjustment models would provide more tailored decision-making about MCS device selection. Accordingly, we developed a machine learning-based mortality model to identify features and patient subgroups that were associated with the largest change in risk when evaluating treatment with specific MCS devices.

Results

Cohort, Model, and Covariates

Our cohort comprised 369 sites and 15,796 patient visits to the cardiac catheterization laboratory from the American College of Cardiology's National Cardiovascular Data Registry CathPCI and Chest Pain-MI registries, as previously described.^{1,7} In-hospital mortality risk was estimated using machine learning. Specifically, we employed a stratified five-fold cross-validation procedure to train the extreme gradient boosting machine (XGBoost)¹⁴ algorithm to predict in-hospital mortality. We used this model because of how the various decision trees trained lead to a divide-and-conquer strategy of patient populations, modeling in each subgroup with multiple interaction effect terms, which has been shown to improve risk prediction in registry data for other models.^{15–17} Model hyperparameter (learning rate, maximum depth of each learner, and number of learners) tuning was conducted through grid search on the training set within the cross-validation. For the final models, the hyperparameter tuning lead to using a learning rate of 0.3, maximum depth of each tree estimator of 6, and with 100 trees. The initial set of candidate features was provided in a prior study and described there.^{1,7} After feature selection, the base model included patient age, sex, percutaneous coronary intervention (PCI) status (elective, urgent, emergent, salvage), recent history of cardiac arrest, initial creatinine, initial hemoglobin, stenosis present in left main coronary artery (percent), stenosis present in proximal left anterior descending coronary artery (percent), presence of multivessel disease, systolic blood pressure at first medical contact, recent history of thrombolytic therapy, and findings on first electrocardiogram (ECG) (no ST elevation, ST elevation, new left bundle branch block, or isolated posterior MI). The cohort creation is shown in [Figure 1](#) and The cohort characteristics are noted in [Table 1](#) and their association with mortality and major bleeding in [Table 2](#). This study was considered exempt by the Human Research Protection Program at Yale University under IRB #0607001639.

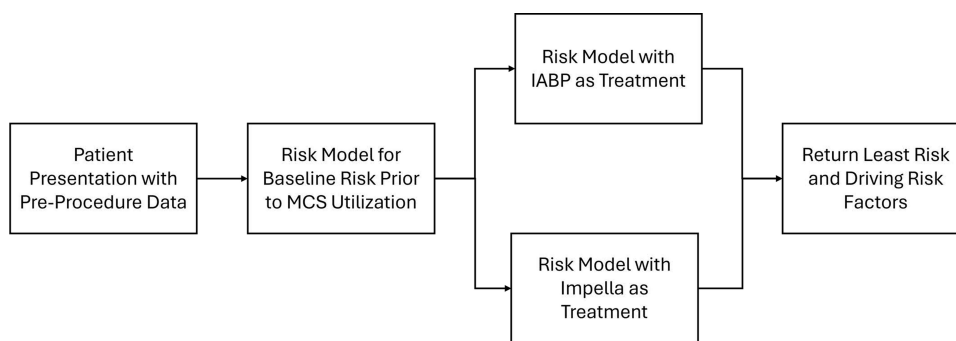


Figure 1 Flowchart of Risk Models and Treatment Decision Making. This figure illustrates the model development process and how a final risk score is calculated.

Model Performance

We measured model performance by the area under the receiver operating characteristic curve (AUROC), for evaluating model discrimination, the average precision score, which calculates the area under the precision recall curve for evaluating model positive predictive performance, and the model Brier Score, for evaluating model calibration.^{18,19} The base model predicted mortality with a mean area under the receiver operating characteristic curve (AUROC) of 0.747 ± 0.002 (standard deviation) (Table 3) and the average precision score of 0.517 ± 0.04 . Adding the MCS device

Table 1 Patient Characteristics

Characteristic	All Patients (n=15,796)
Age (years)	65.7 (\pm 12.6)
Male Sex	10,618 (67.2%)
SBP at First Medical Contact (mmHg)	114.8 (\pm 48.2)
Any Cardiac Arrest	3731 (23.6%)
PCI Status	
Elective	102 (0.6%)
Urgent	2265 (14.3%)
Emergent	12,170 (77.0%)
Salvage	1259 (8.0%)
Initial Creatinine (mg/dL)	1.44 (\pm 1.28)
Left Main Coronary Artery Stenosis (%)	12.2% (\pm 25.1%)
Multivessel Disease	8751 (55.4%)
Initial Hemoglobin (g/dL)	13.7 (\pm 2.2)
Proximal LAD Stenosis (%)	46.5% (\pm 43.5%)
Findings on Initial ECG	
No STEMI	3210 (20.3%)
STEMI	12,258 (77.6%)
LBBB	181 (1.1%)
Posterior MI	147 (0.9%)
Thrombolytic Therapy	556 (3.5%)
MCS Device Utilization	
IABP Only	4919 (31.1%)
Impella devices Only	1686 (10.7%)
Medical Therapy Only	9191 (58.2%)
Major Bleeding	2212.0 (14.0%)
Mortality	4042 (25.6%)

Note: All fields are count (n) and percentage, unless otherwise noted as mean and standard deviation.

Table 2 Raw Association of Patient Characteristics with In-Hospital Major Bleeding and Mortality

Characteristic	Bleeding (n=2212)	No Bleeding (n=13,584)	Mortality (n=4042)	No Mortality (n=11,754)
Age (Years)	67.3 (± 12.3)	65.5 (± 12.6)	69.6 (± 12.7)	64.4 (± 12.3)
Male Sex	1352 (61.1%)	9266 (68.2%)	2578 (63.8%)	8040 (68.4%)
SBP at First Medical Contact (mmHg)	114.4 (± 52.1)	114.9 (± 47.5)	101.4 (± 53.8)	119.4 (± 45.2)
Any Cardiac Arrest	632 (28.6%)	3099 (22.8%)	1494 (37.0%)	2237 (19.0%)
PCI Status				
Elective	11 (0.5%)	91 (0.7%)	16 (0.4%)	86 (0.7%)
Urgent	334 (15.1%)	1931 (14.2%)	409 (10.1%)	1856 (15.8%)
Emergent	1635 (73.9%)	10,535 (77.6%)	2845 (70.4%)	9325 (79.3%)
Salvage	232 (10.5%)	1027 (7.6%)	772 (19.1%)	487 (4.1%)
Initial Creatinine (mg/dL)	1.51 (± 1.34)	1.43 (± 1.26)	1.71 (± 1.50)	1.35 (± 1.17)
Left Main Coronary Artery Stenosis (%)	13.7% (± 26.7%)	12.0% (± 24.8%)	16.6% (± 30.0%)	10.7% (± 23.0%)
Multivessel Disease	1294 (58.5%)	7457 (54.9%)	2516 (62.2%)	6235 (53.0%)
Initial Hemoglobin (g/dL)	13.4 (± 2.4)	13.7 (± 2.2)	13.1 (± 2.4)	13.9 (± 2.2)
Proximal LAD Stenosis (%)	49.1% (± 43.7%)	46.1% (± 43.4%)	51.9% (± 44.3%)	44.7% (± 43.1%)
Findings on Initial ECG				
No STEMI	472 (21.3%)	2738 (20.2%)	758 (18.8%)	2452 (20.9%)
STEMI	1690 (76.4%)	10,568 (77.8%)	3158 (78.1%)	9100 (77.4%)
LBBB	28 (1.3%)	153 (1.1%)	78 (1.9%)	103 (0.9%)
Posterior MI	22 (1.0%)	125 (0.9%)	48 (1.2%)	99 (0.8%)
Thrombolytic Therapy	88 (4.0%)	468 (3.4%)	121 (3.0%)	435 (3.7%)
MCS Device Utilization				
IABP Only	731 (33.0%)	4188 (30.8%)	1465 (36.2%)	3454 (29.4%)
Impella devices Only	529 (23.9%)	1157 (8.5%)	759 (18.8%)	927 (7.9%)
Medical Therapy Only	952 (43.0%)	8239 (60.7%)	1818 (45.0%)	7373 (62.7%)
Major Bleeding	2212 (100.0%)	0 (0.0%)	754 (18.7%)	1458 (12.4%)
Mortality	754 (34.1%)	3288 (24.2%)	4042 (100.0%)	0 (0.0%)

Table 3 Model AUROCs

Model	AUROC	AUPRC	Brier Score
Base Model	0.747 ± 0.024	0.517 ± 0.043	-0.164 ± 0.009
Base Model + MCS Utilization	0.754 ± 0.025	0.519 ± 0.047	-0.164 ± 0.010
Base Model + In-Hospital Major Bleeding	0.748 ± 0.022	0.511 ± 0.043	-0.165 ± 0.008
Base Model + MCS Utilization + In-Hospital Major Bleeding	0.757 ± 0.022	0.522 ± 0.037	-0.163 ± 0.009

utilization variable improved performance to 0.754 ± 0.025 AUROC and 0.519 ± 0.047 average precision score. Adding the major bleeding outcome as a risk factor for mortality resulted in an AUROC of 0.748 ± 0.022 and an average precision score of 0.511 ± 0.043 . Adding both bleeding and MCS device utilization achieved a performance of 0.761 ± 0.022 AUROC and an average precision score of 0.522 ± 0.037 .

Estimated Mortality Risk Difference with MCS Utilization

Using the MCS device utilization model, we calculated the risk differences for Impella devices vs IABP (Figure 2 and Table 4). This figure identifies the population-level estimated risk difference in mortality for patients treated with Impella devices vs IABP. The estimated population mean excess mortality effect of treatment with Impella devices vs IABP was $10.4 \pm 0.8\%$.

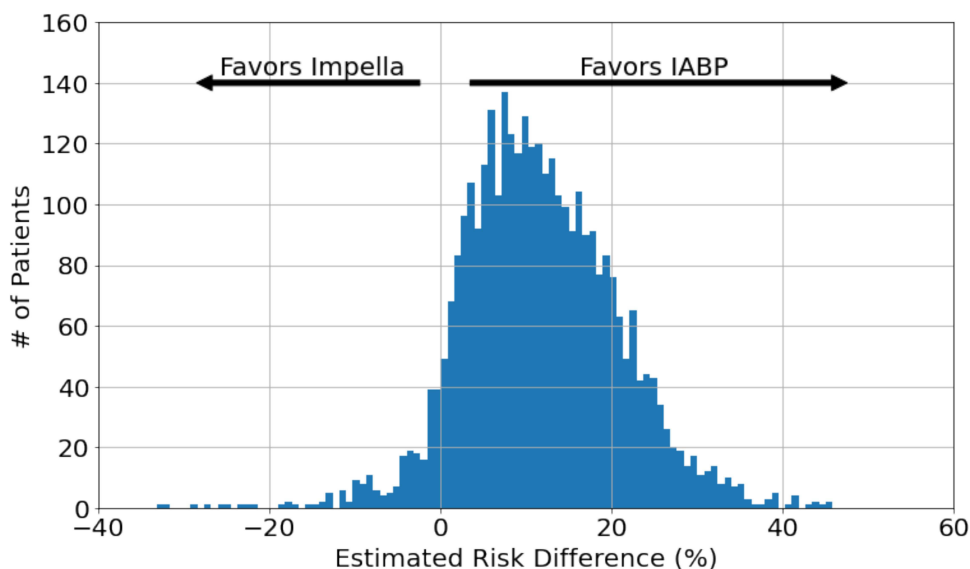


Figure 2 Estimated Mortality Risk Difference of Using IABP vs Impella Devices. This figure illustrates the subset of patients in which the use of either IABP or Impella devices is associated with a more favorable risk of mortality estimate.

Taking a single test fold, there were 282 (7.2%) patients who had an estimated decreased risk of mortality when hypothetically managed with Impella devices compared to IABP (Table 5). The models performed similarly across all folds as demonstrated by the confidence intervals on the findings, so we present a single fold, without loss of generality. The additional folds are illustrated in [Supplementary Figures 1-10](#). Table 5 also presents data from a cohort of equal size from the opposite tail: the 282 patients with the largest increase in mortality risk given Impella devices vs IABP, to identify cases associated with lower risk based upon MCS device choice.

Among these two clinical subgroups of $n=282$ patients each, MCS utilization, bleeding, and mortality rates were not significantly different. Patients with a benefit with IABP had a significantly higher rate of emergent PCI (87.2% vs 39.0%, $p < 0.001$) and a significantly higher rate of STEMI present on initial ECG (85.1% vs 50.0%, $p < 0.001$). Patients with a benefit with Impella use were on average younger (66.6 ± 14.9 years vs 70.7 ± 10.0 years, $p < 0.001$), presented with higher systolic blood pressure (130.6 ± 58.4 mmHg vs 102.0 ± 52.5 mmHg, $p < 0.001$), higher rate of salvage PCI (12.4% vs 6.0%, $p < 0.001$), higher initial creatinine (2.41 vs 1.25 mg/dL, $p < 0.001$) and more likely to present with NSTEMI. This population tended to have a higher degree of proximal LAD stenosis (52.1% vs 26.3%, $p < 0.001$). Finally, this population had significantly higher incidence of NSTEMI (48.6% vs 13.1%, $p < 0.001$).

Interpretation of Factors Associated with Risk

While the risk differences provide a prospective view, model interpretation was performed on the retrospective models with SHapley Additive exPlanations (SHAP) values and plots in order to understand risk factors and model performance^{20,21} and this analysis is explained in [Supplementary materials](#) and [Supplementary Figures 1-13](#)). In brief, age was the most strongly

Table 4 Summary Statistics of Estimated Risk Differences

Summary Statistic	Estimated Mortality Risk Increase Given Impella Devices vs IABP
Mean	10.4% \pm 0.8%
Standard Deviation	9.6% \pm 0.3%
Minimum	-37.3% \pm 10.6%
25%ile	3.9% \pm 0.5%
50%ile (Median)	9.5% \pm 0.9%
75%ile	16.4% \pm 1.2%
Maximum	49.5% \pm 5.7%

Table 5 Descriptive Statistics of Populations with Largest Risk Shifts Given Impella Devices vs IABP

Characteristic	Most Likely to be Associated with In-Hospital Mortality Benefit from Impella Devices Compared to IABP (n=282)	Most Likely to be Associated with In-Hospital Mortality Benefit from IABP Compared to Impella Devices (n=282)	P Value
Age (Years)	66.6 (± 14.9)	70.7 (± 10.0)	< 0.001
Male Sex	174 (61.7%)	171 (60.6%)	0.86
SBP at First Medical Contact (mmHg)	130.6 (± 58.4)	102.0 (± 52.5)	< 0.001
Any Cardiac Arrest	81 (28.7%)	67 (23.8%)	0.21
PCI Status			
Elective	9 (3.2%)	0 (0.0%)	< 0.001
Urgent	128 (45.4%)	19 (6.7%)	
Emergent	110 (39.0%)	246 (87.2%)	
Salvage	35 (12.4%)	17 (6.0%)	
Initial Creatinine (mg/dL)	2.41 (± 2.59)	1.25 (± 0.69)	< 0.001
Left Main Coronary Artery Stenosis (%)	15.4% (± 28.8%)	13.1% (± 26.1%)	0.26
Multivessel Disease	163 (57.8%)	157 (55.7%)	0.67
Initial Hemoglobin (g/dL)	12.9 (± 2.2)	13.7 (± 2.3)	< 0.001
Proximal LAD Stenosis (%)	52.1% (± 39.6%)	26.3% (± 40.8%)	< 0.001
Findings on Initial ECG			
No STEMI	137 (48.6%)	37 (13.1%)	< 0.001
STEMI	141 (50.0%)	240 (85.1%)	
LBBB	1 (0.4%)	3 (1.1%)	
Posterior MI	3 (1.1%)	2 (0.7%)	
Thrombolytic Therapy	6 (2.1%)	6 (2.1%)	1
MCS Device Utilization			
IABP Only	87 (30.9%)	97 (34.4%)	0.94
Impella devices Only	30 (10.6%)	31 (11.0%)	
Medical Therapy Only	165 (58.5%)	154 (54.6%)	
Major Bleeding	44 (15.6%)	40 (14.2%)	0.72
Mortality	88 (31.2%)	84 (29.8%)	0.78

informative feature, as ranked by SHAP, followed by initial creatinine and history of cardiac arrest. Higher age, creatinine, and history of cardiac arrest were all associated with increased risk. SHAP contributions for the model incorporating MCS device utilization found that MCS device utilization was the fourth most informative feature.

Discussion

While prior findings (and this work) noted that overall in the population, Impella devices were associated with higher mortality risk compared to IABP,^{1,10–12} some patients with certain clinical profiles were, in fact, associated with lower risk and different variables had different impact on this risk, illustrating significant patient heterogeneity. Specifically, our analysis identified 282 patients with a decreased risk of mortality when hypothetically managed with Impella devices as compared with IABP. Those patients were on average younger, presented with higher systolic blood pressure, higher rate of salvage percutaneous coronary intervention, higher initial creatinine, and lower hemoglobin. While Impella devices were associated with higher mortality risk overall, certain clinical profiles were associated with lower risk, illustrating patient heterogeneity.

These findings align with recent trials identifying subpopulations where MCS device utilization led to improved outcomes.¹³ This analysis suggests that patients with certain clinical features, or a combination of factors, may have a larger benefit or a harm with different MCS strategies, highlighting the need for personalized estimation of risk and benefit to support tailored decision-making. Once the XGBoost model has been trained, it is able to produce both an estimate risk of mortality and an estimation for how that risk would change given different MCS treatment choices.

SHAP can display the factors that lead to these individual risks, and an implementation tool could allow for model transparency to enable researchers and clinicians to understand what nonlinear interactions of factors drive the overall estimated risk.

It is interesting to note that when comparing the two populations (which represent those with potential benefit from Impella devices compared to IABP and an equal number with the highest risk from Impella devices compared to IABP) in Table 5, no significant difference is seen with outcomes relative to MCS utilization. This implies that the factors that the model uses to estimate mortality may not be fully understood in clinical practice. If a clear difference in utilization between these groups were present, it would indicate that the model's training primarily focused on clinician treatment patterns. However, since no difference in MCS utilization can be seen between these groups, it is more plausible that the model is instead identifying a previously unrecognized interaction between patient factors. Decision tree methods are capable of exploring subpopulations and identifying similar patients based upon measured factors. While limited, the completeness of the registry data allows for comprehensive exploration, similar to work that explores subpopulations in more general clinical settings.²²

It is possible this is a reflection of the challenges clinicians face integrating multiple features in real-time with an unstable patient, as well as a need for additional risk factors measured for understanding the association with treatment and risk (discussed in limitations). For the former, a further investigation is needed to test the findings of these models and these hypotheses on patients associated with lowered risk, to refine understanding of the relationship between MCS utilization and mortality in AMI-CS patients in real-time while treatment decisions are made.

Limitations

This work has several limitations. First, this is a retrospective analysis that is subject to unmeasured confounding. However, the tails of the specific distribution in Figure 2 may help inform randomized control trials to determine if and which patients benefit from IABP, Impella devices, or no MCS device. We could not distinguish specific models of MCS devices used (such as Impella device models) nor did we consider groups that involved combinations of various MCS devices. While the sites selected were those that prescribed both IABP and Impella, it did not adjust for such clinical sites that strongly preferred one to the other. Therefore, there is a potential for selection bias in the findings, given the limited focus on those that generally prescribed both MCS devices, not including any locations that may have prescribed more or multiple MCS devices not considered in this analysis. Finally, the SHAP feature importance plots show individual feature rankings and associations with outcomes but do not clearly illustrate the higher order interaction between the terms. While figures were provided that visualize two-term interactions, trees of depth 6 allow for potentially 6 terms to interact together to better estimate risk of mortality. This association, however, is challenging to identify and illustrate.

Conclusion

Advanced machine learning models, which can classify patients into subgroups of risk as driven by particular factors, better model risk of mortality in heterogeneous patient populations. These risk factors, and their associated importance in mortality prediction when interacting with MCS device utilization, identified cohorts of individuals where selection of particular MCS utilization for treatment may be associated with improved risk of mortality. This information could help guide potential cohort selection for future randomized control trial design in order to better identify best practices for treatment decisions in the heterogeneous group of patients with AMI-CS.

Online Methods

Data Sources

Our data sources were linked American College of Cardiology's National Cardiovascular Data Registry CathPCI and Chest Pain-MI registries, as previously described.^{1,7} The linkage provided full episode of care detail because of complementary data from the CathPCI registry and detailed MCS utilization data from the Chest Pain-MI registry. These registries have high data quality standards, including requiring data completeness.²³ From the matched cohort, we eliminated episodes of care from any hospital without at least one of each of Impella device or IABP within the AMI-CS

cohort, to ensure cases were at sites with capability to use any of these therapies. Additionally, we eliminated episodes of care where a given patient received an MCS other than Impella devices or IABP or where a given patient received multiple MCS devices. Our cohort selection is illustrated in [Figure 1](#).

Outcomes

The primary outcome was in-hospital mortality. A secondary outcome was in-hospital major bleeding. Major bleeding was defined as in prior work^{1,7} as a decline in hemoglobin of at least 3 g/dL; transfusion of whole blood or packed red blood cells; procedural intervention/surgery at bleeding site to treat the bleeding; or documented or suspected retro-peritoneal bleed, gastrointestinal bleed, genitourinary bleed, or a bleed in a location not specified elsewhere.²⁴

Statistical Analysis

We trained machine learning models to estimate risk of in-hospital mortality using the extreme gradient boosting machine (XGBoost).¹⁴ XGBoost is an efficient implementation of the gradient boosting machine (GBM) algorithm²⁵ with several algorithmic and hardware enhancements leading to better scalability, speed, and performance accuracy. An important enhancement in XGBoost is the implementation of the least absolute shrinkage and selection operator (LASSO) and the ridge (Ridge) regularization techniques,²⁶ which can penalize large and complex models to prevent overfitting. The XGBoost model has been shown to improve model performance in the NCDR registries.^{15,16} We used a nested approach for model development, hyperparameter tuning, and evaluation. In the primary evaluation (for model performance), we conducted a stratified, five-fold cross-validation, with each fold splitting the data 80% for model training and development and 20% for model testing and analysis. For that 80% selected for training, we conducted an inner evaluation for hyperparameter tuning, through a stratified twenty-fold cross-validation. Here, 95% of the data were used to train and 5% to test to select optimal model hyperparameters. The hyperparameters of our model were the learning rate (searched from 0.1 to 0.3), the maximum depth of each tree (1,2,3,6), and the number of trees used in the final model (50, 100, 500, 1000). Other parameters were left at their default values. Further details are provided in the [Supplementary Materials](#). Model performance was assessed using the area under the receiver operating characteristic curve (AUROC), with a 95% confidence interval.

Covariates and Feature Selection

To develop a parsimonious and interpretable risk model, we followed a stepwise process which included both clinically driven and data-driven steps for feature elimination. We began with a large set of variables used for propensity matching in prior research.¹ These included patient demographics, medical history, clinical presentation, laboratory values, administered medications, procedural characteristics, and coronary anatomic data. Binary variables with missing data were coded as “no”, while missing categorical variable levels were coded as “no” or “other”, and continuous variables imputed by the mean, though rates of missing data were extremely low.²³

We then began our process of eliminating variables. We first eliminated left ventricular ejection fraction because this measure is not always available at the time of treatment decision. We eliminated peri-procedural medication utilization to avoid biasing the model based on treatment decisions. We then methodically eliminated features using backwards selection: we trained a model on all features, and then eliminated the least informative variables as ranked by XGBoost’s internal importance ranking. We continued this removal process until removal of the least-important feature resulted in an AUROC below the 95% confidence interval surrounding the original model’s AUROC. This was done in order to both minimize model overfitting and to allow improved generalization with a more parsimonious model.

Staging of Models

We then built additional models that considered the selection of MCS utilization for treatment. While the first model used only pre-treatment decision factors, the second model added MCS utilization, either Impella device or IABP. This model was used to interpret the association of MCS utilization with both mortality and changes in feature importance of variables. This design allows for the comparison of associations of each covariate with in-hospital mortality both with and without knowledge of MCS device utilization.

Estimated Differences in Risk of Mortality

We next used a simulated prospective approach to estimate risk changes with different MCS treatment choices. This was accomplished using the second-tier model (Baseline + MCS Utilization) and comparing differences in estimated risk for all test patients. The change in the estimated risk reflects how the treatment effect of MCS utilization is heterogeneous or varies across individuals and their characteristics. For each individual patient, a set of risk estimates (the probability of an outcome given MCS utilization and covariates) was generated. This set of estimates was generated by imputing MCS utilization to each of its two possible values, independent of the intervention the patient actually received. Two risk estimates are therefore generated for each subject: 1) an estimate for if that subject received Impella device only, and 2) an estimate if that same subject received IABP only illustrated in [Figure 2](#). The risk difference for two interventions was estimated by subtracting these values. The risk difference for Impella device vs IABP was estimated as:

$$\text{Risk Difference}_{\text{Impella vs IABP}} = P(\text{mortality} = \text{True} | \text{MCS} = \text{Impella}) - P(\text{mortality} = \text{True} | \text{MCS} = \text{IABP})$$

Here a positive estimated risk difference indicates a given patient's features are associated with higher risk of mortality if treated with Impella device rather than IABP. The distribution of estimated risk differences was visualized using a histogram. To compare the strongest changes in predicted outcome, the tails of this distribution were compared. Among the populations of these tails, patient characteristics were compared with Mann–Whitney *U*-tests for continuous variables and Fisher's Exact Test for categorical variables.

Interpretation of Risk Factors

While the risk differences provide a prospective view, model interpretation was performed on the retrospective models with SHapley Additive exPlanations (SHAP) values and plots in order to understand risk factors and model performance^{20,21} and this analysis is explained in the [Supplementary Materials](#), including SHAP plots for interactive effects of risk factors in [Supplementary Figures 11–13](#).

All data analyses was performed in Python 3.6 and R4.1, with the packages XGBoost 1.3.3,¹⁴ SHAP 0.38.1,²⁷ and scikit learn (sklearn) 0.24.1.²⁸

Acknowledgments

This paper has been uploaded to MedRxiv as a preprint: Modeling In-Hospital Mortality Among Patients Undergoing Percutaneous Coronary Intervention with Acute Myocardial Infarction Complicated by Cardiogenic Shock Receiving Mechanical Circulatory Support | [medRxiv](#).

Funding

This work was supported by the Food and Drug Administration (FDA) of the US Department of Health and Human Services (HHS) as part of a financial assistance award [FD005938] totaling \$92,921 with 100% funded by FDA/HHS. The contents are those of the authors and do not necessarily represent the official views of, nor an endorsement by, FDA/HHS, or the US Government.

Disclosure

Dr. Mortazavi currently receives research support through the National Science Foundation (IIS 2014475), and the National Institutes of Health (1 R01 HL167858-01A1, 1 R01 DK136414, and 1 R01 HL151240-01). Dr. Mortazavi is an expert witness at the Kirkland and Ellis Law Firm for matters unrelated to this manuscript. Dr. Mortazavi was an expert witness at the request of Abbott's attorneys, the McAndrews Held and Malloy Law Firm, in a case that was settled in December of 2024. He currently serves as an Associate Editor for the *Journal of the American College of Cardiology*. Dr. Dhruva reports research support from the Department of Veterans Affairs and Arnold Ventures. He reports serving on the Medicare Evidence Development and Coverage Advisory Committee and the Institute for Clinical and Economic Review California Technology Assessment Forum. The views expressed in the manuscript do not reflect those of the Department of Veterans Affairs or the United States government. Dr. Ross currently receives research support through Yale University from Johnson and Johnson to develop methods of clinical trial data sharing, from the Food and Drug

Administration for the Yale-Mayo Clinic Center for Excellence in Regulatory Science and Innovation (CERSI) program (U01FD005938), from the Agency for Healthcare Research and Quality (R01HS022882), and from Arnold Ventures; formerly received research support from the Medical Device Innovation Consortium as part of the National Evaluation System for Health Technology (NEST) and from the National Heart, Lung and Blood Institute of the National Institutes of Health (NIH) (R01HS025164, R01HL144644); and in addition, Dr. Ross was an expert witness at the request of Relator's attorneys, the Greene Law Firm, in a qui tam suit alleging violations of the False Claims Act and Anti-Kickback Statute against Biogen Inc. that was settled September 2022. He also reports grants from Greenwall Foundation and Arnold Ventures, outside the submitted work. Dr. Ngufor reports that, in the past 36 months, he has received research support through the Department of Defence (DOD) Prostate Cancer Research Program (PCRP) Health Equity Research and Outcomes Improvement Consortium (HEROICA) Award W81XWH2210968. No other conflicts of interest exist. Dr. Curtis has a contract with the American College of Cardiology for his role as Chief Scientific Advisor of the American College of Cardiology's National Cardiovascular Data Registry. He holds equity interest in Medtronic. Dr. Krumholz, in the past three years, has received options for Element Science and Identifeye and payments from F-Prime for advisory roles. He was a co-founder of and held equity in Hugo Health. He is a co-founder of and holds equity in Refactor Health and ENSIGHT-AI. He is associated with research contracts through Yale University from Janssen, Kenvue, Novartis, and Pfizer. He is the Editor-in-Chief of the *Journal of the American College of Cardiology*. Dr. Desai reports grants and personal fees from Amgen, Novartis, Merck, and Arrowhead, outside the submitted work. The authors report no other conflicts of interest in this work.

References

1. Dhruva SS, Ross JS, Mortazavi BJ, et al. Association of use of an intravascular microaxial left ventricular assist device vs intra-aortic balloon pump with in-hospital mortality and major bleeding among patients with acute myocardial infarction complicated by cardiogenic shock. *JAMA*. 2020;323(8):734–745. doi:10.1001/jama.2020.0254
2. Sandhu A, McCoy LA, Negi SI, et al. Use of mechanical circulatory support in patients undergoing percutaneous coronary intervention: insights from the national cardiovascular data registry. *Circulation*. 2015;132(13):1243–1251. doi:10.1161/CIRCULATIONAHA.114.014451
3. Khera R, Cram P, Lu X, et al. Trends in the use of percutaneous ventricular assist devices: analysis of national inpatient sample data, 2007 through 2012. *JAMA Internal Med*. 2015;175(6):941–950. doi:10.1001/jamainternmed.2014.7856
4. Shaefi S, O'Gara B, Kociol RD, et al. Effect of cardiogenic shock hospital volume on mortality in patients with cardiogenic shock. *J Am Heart Assoc*. 2015;4(1):e001462. doi:10.1161/JAHA.114.001462
5. Baran DA, Grines CL, Bailey S, et al. SCAI clinical expert consensus statement on the classification of cardiogenic shock: this document was endorsed by the American College of Cardiology (ACC), the American Heart Association (AHA), the Society of Critical Care Medicine (SCCM), and the Society of Thoracic Surgeons (STS) in April 2019. *Catheterization Cardiovasc Interventions*. 2019;94(1):29–37. doi:10.1002/ccd.28329
6. Naidu SS, Baran DA, Jentzer JC, et al. SCAI SHOCK Stage Classification Expert Consensus Update: a Review and Incorporation of Validation Studies: this statement was endorsed by the American College of Cardiology (ACC), American College of Emergency Physicians (ACEP), American Heart Association (AHA), European Society of Cardiology (ESC) Association for Acute Cardiovascular Care (ACVC), International Society for Heart and Lung Transplantation (ISHLT), Society of Critical Care Medicine (SCCM), and Society of Thoracic Surgeons (STS) in December 2021. *J Am College Cardiol*. 2022;79(9):933–946. doi:10.1016/j.jacc.2022.01.018
7. Dhruva SS, Ross JS, Mortazavi BJ, et al. Use of mechanical circulatory support devices among patients with acute myocardial infarction complicated by cardiogenic shock. *JAMA Network Open*. 2021;4(2):e2037748–e. doi:10.1001/jamanetworkopen.2020.37748
8. Seyfarth M, Sibbing D, Bauer I, et al. A randomized clinical trial to evaluate the safety and efficacy of a percutaneous left ventricular assist device versus intra-aortic balloon pumping for treatment of cardiogenic shock caused by myocardial infarction. *J Am College Cardiol*. 2008;52(19):1584–1588. doi:10.1016/j.jacc.2008.05.065
9. Ouweneel DM, Eriksen E, Sjaauw KD, et al. Percutaneous mechanical circulatory support versus intra-aortic balloon pump in cardiogenic shock after acute myocardial infarction. *J Am College Cardiol*. 2017;69(3):278–287. doi:10.1016/j.jacc.2016.10.022
10. Amin AP, Spertus JA, Curtis JP, et al. The evolving landscape of Impella use in the United States among patients undergoing percutaneous coronary intervention with mechanical circulatory support. *Circulation*. 2020;141(4):273–284. doi:10.1161/CIRCULATIONAHA.119.044007
11. Kim Y, Shapero K, Ahn SS, Goldsweig AM, Desai N, Altin SE. Outcomes of mechanical circulatory support for acute myocardial infarction complicated by cardiogenic shock. *Catheterization Cardiovasc Interven*. 2022;99(3):658–63.
12. Miller PE, Bromfield SG, Ma Q, et al. Clinical outcomes and cost associated with an intravascular microaxial left ventricular assist device vs intra-aortic balloon pump in patients presenting with acute myocardial infarction complicated by cardiogenic shock. *JAMA Internal Med*. 2022;182(9):926–933. doi:10.1001/jamainternmed.2022.2735
13. Møller JE, Engström T, Jensen LO, et al. Microaxial flow pump or standard care in infarct-related cardiogenic shock. *New Engl J Med*. 2024;390(15):1382–1393. doi:10.1056/NEJMoa2312572
14. Chen T, Guestrin C, editors. Xgboost: a scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining; 2016.
15. Huang C, Murugiah K, Mahajan S, et al. Enhancing the prediction of acute kidney injury risk after percutaneous coronary intervention using machine learning techniques: a retrospective cohort study. *PLoS Med*. 2018;15(11):e1002703. doi:10.1371/journal.pmed.1002703

16. Mortazavi BJ, Bucholz EM, Desai NR, et al. Comparison of machine learning methods with national cardiovascular data registry models for prediction of risk of bleeding after percutaneous coronary intervention. *JAMA Network Open*. 2019;2(7):e196835–e. doi:10.1001/jamanetworkopen.2019.6835
17. Khera R, Haimovich J, Hurley NC, et al. Use of machine learning models to predict death after acute myocardial infarction. *JAMA Cardiol*. 2021;6(6):633. doi:10.1001/jamacardio.2021.0122
18. Stevens LM, Mortazavi BJ, Deo RC, Curtis L, Kao DP. Recommendations for reporting machine learning analyses in clinical research. *Circulation*. 2020;13(10):e006556. doi:10.1161/CIRCOUTCOMES.120.006556
19. Huang C, Li S-X, Caraballo C, et al. Performance metrics for the comparative analysis of clinical risk prediction models employing machine learning. *Circulation*. 2021;14(10):e007526. doi:10.1161/CIRCOUTCOMES.120.007526
20. Lundberg SM, Erion GG, Lee S-I. Consistent individualized feature attribution for tree ensembles. arXiv preprint arXiv:180203888 2018.
21. Lundberg SM, Nair B, Vavilala MS, et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nat Biomed Eng*. 2018;2(10):749–760. doi:10.1038/s41551-018-0304-0
22. Khera R, Mortazavi BJ, Sangha V, et al. Accuracy of computable phenotyping approaches for SARS-CoV-2 Infection and COVID-19 hospitalizations from the electronic health record. *medRxiv*. 2021;2021:1.
23. Messenger JC, Ho KK, Young CH, et al. The National Cardiovascular Data Registry (NCDR) data quality brief: the NCDR data quality program in 2012. *J Am College Cardiol*. 2012;60(16):1484–1488. doi:10.1016/j.jacc.2012.07.020
24. American College of Cardiology National Cardiovascular Data Registry. What each registry collections [cited 2021 August]. Available from: <https://cvquality.acc.org/NCDR-Home/Data-Collection/What-Each-Registry-Collects>. Accessed February 5, 2025.
25. Friedman JH. Greedy function approximation: a gradient boosting machine. *Annals Stat*. 2001;2001:1189–1232.
26. Zou H, Hastie T. Regularization and variable selection via the elastic net. *J Royal Stat Soc*. 2005;67(2):301–320. doi:10.1111/j.1467-9868.2005.00503.x
27. Lundberg SM, Erion G, Chen H, et al. From local explanations to global understanding with explainable AI for trees. *Nat Machine Intell*. 2020;2(1):56–67. doi:10.1038/s42256-019-0138-9
28. Pedregosa F, Varoquaux G, Gramfort A, et al. Machine learning in Python. *J Machine Learn Res*. 2011;12:2825–2830.

Medical Devices: Evidence and Research

Dovepress
Taylor & Francis Group

Publish your work in this journal

Medical Devices: Evidence and Research is an international, peer-reviewed, open access journal that focuses on the evidence, technology, research, and expert opinion supporting the use and application of medical devices in the diagnosis, monitoring, treatment and management of clinical conditions and physiological processes. The identification of novel devices and optimal use of existing devices which will lead to improved clinical outcomes and more effective patient management and safety is a key feature of the journal. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/medical-devices-evidence-and-research-journal>