

# Determinants of In-Hospital Mortality After Percutaneous Coronary Intervention: A Machine Learning Approach

Subhi J. Al'Aref, MD;\* Gurpreet Singh, PhD;\* Alexander R. van Rosendaal, MD; Kranthi K. Kolli, PhD; Xiaoyue Ma, MSc; Gabriel Maliakal, MSc; Mohit Pandey, MSc; Benjamin C. Lee, PhD; Jing Wang, PhD; Zhuoran Xu, MD, MSc; Yiye Zhang, MSc, PhD; James K. Min, MD; S. Chiu Wong, MD; Robert M. Minutello, MD

**Background**—The ability to accurately predict the occurrence of in-hospital death after percutaneous coronary intervention is important for clinical decision-making. We sought to utilize the New York Percutaneous Coronary Intervention Reporting System in order to elucidate the determinants of in-hospital mortality in patients undergoing percutaneous coronary intervention across New York State.

**Methods and Results**—We examined 479 804 patients undergoing percutaneous coronary intervention between 2004 and 2012, utilizing traditional and advanced machine learning algorithms to determine the most significant predictors of in-hospital mortality. The entire data were randomly split into a training (80%) and a testing set (20%). Tuned hyperparameters were used to generate a trained model while the performance of the model was independently evaluated on the testing set after plotting a receiver-operator characteristic curve and using the output measure of the area under the curve (AUC) and the associated 95% CIs. Mean age was  $65.2 \pm 11.9$  years and 68.5% were women. There were 2549 in-hospital deaths within the patient population. A boosted ensemble algorithm (AdaBoost) had optimal discrimination with AUC of 0.927 (95% CI 0.923–0.929) compared with AUC of 0.913 for XGBoost (95% CI 0.906–0.919,  $P=0.02$ ), AUC of 0.892 for Random Forest (95% CI 0.889–0.896,  $P<0.01$ ), and AUC of 0.908 for logistic regression (95% CI 0.907–0.910,  $P<0.01$ ). The 2 most significant predictors were age and ejection fraction.

**Conclusions**—A big data approach that utilizes advanced machine learning algorithms identifies new associations among risk factors and provides high accuracy for the prediction of in-hospital mortality in patients undergoing percutaneous coronary intervention. (*J Am Heart Assoc.* 2019;8:e011160. DOI: 10.1161/JAHA.118.011160.)

**Key Words:** big data analytics • in-hospital mortality • machine learning • percutaneous coronary intervention

Accurate prediction of adverse events after the performance of percutaneous coronary intervention (PCI) is the hallmark of contemporary societal guidelines, because the magnitude of risk can aid in therapy selection and form the basis for a precise preprocedural informed consent practice.<sup>1–3</sup>

From the Dalio Institute of Cardiovascular Imaging, New York-Presbyterian Hospital, New York, NY (S.J.A., G.S., A.R.v.R., K.K.K., X.M., G.M., M.P., B.C.L., J.W., Z.X., J.K.M.); Division of Health Informatics, Weill Cornell Graduate School of Medical Sciences, New York, NY (Y.Z.); Division of Cardiology, Department of Medicine, Weill Cornell Medicine, New York, NY (S.C.W., R.M.M.).

Accompanying Figures S1 through S4 are available at <https://www.ahajournals.org/doi/suppl/10.1161/JAHA.118.011160>

\*Dr Al'Aref and Dr Singh contributed equally to this work.

**Correspondence to:** James K. Min, MD, Weill Cornell Medical College, Dalio Institute of Cardiovascular Imaging/New York-Presbyterian Hospital, Non-Invasive Imaging/Cardiovascular Research Foundation, 413 E. 69th St, Suite 108, New York, NY. E-mail: [jkm2001@med.cornell.edu](mailto:jkm2001@med.cornell.edu)

Received October 9, 2018; accepted December 26, 2018.

© 2019 The Authors. Published on behalf of the American Heart Association, Inc., by Wiley. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Based on preprocedural factors (typically a combination of clinical and/or angiographic variables), existing risk scores estimate the individualized risk for adverse outcomes after coronary revascularization.<sup>1,4–6</sup> Two recent risk scores using data from the New York PCI reporting system (PCIRS) have been developed aiming at predicting in-hospital and/or 30-day mortality after PCI.<sup>7,8</sup> These scores were created by assigning a specified number of points to important risk factors that were summed to obtain a per-patient predicted probability for the outcome, and possessed good discrimination (area under the receiver operating characteristics curve [AUC] of 0.886 for in-hospital mortality and 0.890 for in hospital/30-day mortality) with appropriate calibration. However, variables for logistic regression models were chosen based on the presence of a significant bivariate relationship with the primary outcome ( $P<0.10$ ), and these candidate variables were then further multivariately modeled. This approach may ignore the potential prognostic value of interactions between several unexpected weaker risk factors with the primary outcome. Secondly, continuous variables were grouped into categories, which may, albeit easy to use, induce loss of predictive information.

## Clinical Perspective

### What Is New?

- Accurate prediction of adverse events after coronary revascularization is essential for preprocedural informed consent and for appropriate therapy selection.
- This study utilized novel machine learning methodologies for the accurate prediction of in-hospital mortality following percutaneous coronary intervention, utilizing a contemporary database and without the exclusion of any patients.
- The study also showed that several variables, beyond what has been traditionally established, are important predictors of in-hospital mortality including angiographic measures of diameter stenosis, the occurrence of acute cerebrovascular events within 24 hours of percutaneous coronary intervention, as well as day of the week in which percutaneous coronary intervention was performed.

### What Are the Clinical Implications?

- This work could lead to the widespread adoption of machine learning–based algorithms for the development of accurate, precise, and generalizable risk assessment tools in clinical practice.
- Such improved risk assessment could lead to better therapy selection and lower periprocedural complications.

Machine learning (ML) is a novel field in computer science that has been increasingly utilized in clinical research in an attempt to improve predictive modeling and elucidate novel determinants of a specific outcome.<sup>9,10</sup> ML is a subset of artificial intelligence that uses algorithms that autonomously acquire knowledge by extracting patterns from data. For example, ML-based algorithms have been successfully applied to several aspects of cardiovascular research, ranging from image segmentation in automated coronary artery calcium scoring to outcomes research such as prediction of heart failure rehospitalization.<sup>11,12</sup> Given the potential for ML to analyze ever-expanding data sets and to include a large number of variables that can be tested for numerous interactions and nonlinear relationships with the outcome, ML may improve risk assessment. The current study sought to utilize ML for the prediction of in-hospital mortality by utilizing the state-mandated clinical registry cohort, the New York PCIRS, among patients undergoing PCI in New York State between 2004 and 2012.

## Methods

### New York PCIRS

In order to access data collected for this study, requests to access the data set from qualified researchers trained in human subject confidentiality protocols may be sent to the

New York State Department of Health at cardiac-data@health.ny.gov. The code used for data analysis in this study has been made publicly available at GitHub.<sup>13</sup> The New York PCI registry was initiated by the New York State Department of Health in 1992 in order to establish a clinical registry that provides information regarding quality of care provided across New York State hospitals.<sup>14</sup> The database contains detailed, de-identified information on demographics, baseline clinical characteristics, periprocedural and procedural variables of patients undergoing PCI, as well as reperfusion time intervals in acute myocardial infarction patients. The primary outcome in this investigation was in-hospital mortality, which was one of the reported variables for the discharge status.

The New York State Department of Health has multiple mechanisms in place to confirm accuracy of the data in PCIRS. Accuracy of risk-factor entries is confirmed through auditing of samples obtained from participating hospitals. Furthermore, data are matched to both New York's administrative database and the Statewide Planning and Research Cooperative System, which contain information on both inpatient and outpatient PCIs. In-hospital mortality is confirmed by matching with Statewide Planning and Research Cooperative System entries as well. The Weill Cornell Medicine Institutional Review Board approved use of the PCIRS database as well as the study protocol.

### Patient Population

The PCIRS database obtained from January 1, 2004 until December 31, 2012 was used in its entirety with no exclusion criteria applied. The study protocol was approved by the Institutional Review Board of Weill Cornell Medicine, and informed consent had been waived since the database obtained had no identifiable information. The total sample size was 479 804 unique patients. All patients who underwent PCI in the state of New York were included in the analysis, which comprised both elective and emergent cases (covering the spectrum of coronary artery disease presentation). Furthermore, the number of nonfederal New York State Hospitals enrolled in the registry increased from 48 hospitals in 2004 to 60 hospitals in 2012.

### Variables Examined

Members on the cardiac advisory committee determine variables included in the PCIRS database. Patients in shock at initial presentation (defined as acute hypotension with systolic blood pressure <80 mm Hg or low cardiac index [ $<2.0$  L/min per  $m^2$ ], despite pharmacologic or mechanical support) had been originally excluded from the PCIRS database. Further to our analysis, we excluded clinically

nonrelevant variables (n=4), variables with 1 value (n=1), duplicate variables (n=1), and variables with >70% missing values (n=62) (Figure S1). In the final analysis, we included a total of 49 variables (8 continuous and 41 categorical): baseline demographics and clinical characteristics (n=18), ejection fraction (n=1), baseline chemistry values (n=1), periprocedural therapy and equipment used (n=8), hemodynamic instability (n=1), invasive coronary angiographic findings (n=15), periprocedural complications and outcomes (n=3), day of the week PCI was performed (n=1), and facility type (n=1). The occurrence of postprocedural complications was defined as the occurrence of stroke, Q-wave myocardial infarction, acute occlusion in the target lesion or in a significant side branch, vascular injury at the access site requiring intervention, renal failure, emergency cardiac surgery, stent thrombosis, and coronary perforation or the need to emergently return to the catheterization laboratory for PCI.

## ML Methodology

### Data processing

For continuous variables, we performed range normalization (values ranging from 0 to 1) to eliminate the possibility of model bias caused by magnitude of the numerical values. For categorical variables, we performed 1-hot-encoding, which is defined as the process of dividing categorical values into pairs of zero and nonzero values for the goal of transforming the variables into a format that can be used for a classification algorithm.<sup>15</sup> In a large database, such as PCIRS, having missing values in more than 1 variable poses a special challenge, in terms of prediction model development and extraction of maximal information available from the data at hand. To handle this issue, missing values were imputed using Multiple Imputations by Chained Equations (MICE). MICE has emerged as one of the principal statistical approaches to dealing with missing data, which involves multiple imputations, as opposed to single imputations, in order to account for the statistical uncertainty associated with imputations. The chained equations approach can also handle variables of various types and complexities.<sup>16,17</sup>

Before model construction, variables were examined for nullity correlation (Figure S2) and value correlation (Figure S3). Nullity correlation was performed in order to examine whether certain features were correlated in terms of missingness (since several angiographic variables included in the model depended on each other: if a patient had 1 lesion then they would not have lesion #2 and so forth). The nullity correlation ranges from  $-1$  to  $1$ . A value of  $-1$  indicates that if 1 variable appears the other definitely does not. A value of "0" indicates that variables appearing or not appearing have no effect on one another, whereas a value of "1" indicates that if 1 variable

appears the other definitely also does. Value correlation was performed in order to remove variables that may contribute to numerical instability, cause model overfitting, and/or collude the interpretability of the model. In this instance, no variables were excluded based on the correlation coefficients and clinical significance.

### Attribute selection

Attribute selection was done after fine-tuning of the hyperparameters—defined as the model parameters that are given an arbitrary value before the initiation of the learning process. Attribute selection was performed using the information gain ranking method that aims at ranking features based on high information gain entropy. The attributes with information gain  $>0$  were only used for the ML approach.

### Supervised ML approach

Predictive classifiers were developed based on data from the training set using 4 supervised ML methods: (1) Adaptive Boosting (AdaBoost), (2) Extreme Gradient Boosting (XGBoost), (3) Random Forest, and (4) Logistic Regression. The following ML-based algorithms were used in the present analysis since they represent the full spectrum of analytic approach from traditional logistic regression as used with statistical analyses, to traditional ML algorithms (Random Forest), decision-tree-based adaptive boosting algorithms (AdaBoost) as well as ensemble boosting (XGBoost). Boosting has been increasingly used within ML as it involves the sequential creation of models, with each iteration attempting at correcting the errors of the first model. The first successful boosting algorithm was AdaBoost, and current state-of-the-art boosting algorithms utilize gradient boosted decision trees (XGBoost) for optimal speed and performance. In order to evaluate the efficacy of each model, we used the technique of K-Fold cross-validation on a randomly under-sampled subset from the entire data set.<sup>9,15</sup> Random undersampling was performed to eliminate the adverse effects of highly unbalanced classes (in the ratio of 2549/477 255 [0.0053] for this data set) on the model accuracy. For each model, we performed 5-fold cross-validation by randomly splitting the entire data set into 5 parts for 5 iterations. At each iteration, we chose 4 parts as training data and 1 part as a testing set. We report the averaged results for each model on the unseen 20% testing data. The overall performance of the prediction model on the test set was assessed by calculation of the AUC from the receiver operating characteristic curve and the associated 95% CI. Finally, calibration was reported for each model. Calibration provides knowledge regarding the confidence in assigning a specific class to each of the already established labels, and is commonly reported in developed clinical prediction models.

## Results

### Patient Characteristics

Table 1 lists the baseline characteristics of the study population (n=479 804). The mean age was 65.2 years (SD±11.9) and there was a predominance of female patients (68.5%). Patients were relatively healthy at baseline, with low prevalence of concomitant conditions: 8.2% with prior cerebrovascular event, 4% with history of heart failure, 2.2% with previous renal failure necessitating dialysis, 7.8% with peripheral vascular disease, and 15.5% with chronic obstructive pulmonary disease. Diabetes mellitus (33.7%), on the other hand, was the most prevalent cardiovascular risk factor. Finally, one tenth of patients presented with ST-segment elevation on ECG, while 0.5% patients were hemodynamically unstable at the time of coronary angiography (defined as a requirement for pharmacologic or mechanical support to maintain blood pressure or cardiac index). There were 2549 in-hospital deaths from 2004 to 2012 (representing an event rate of 0.5%).

**Table 1.** Baseline Characteristics of the Study Population

Variable	All Patients (n=479 804)
Age (y), mean±SD	65.2±11.9
Male sex (%)	151 349 (31.5%)
White ethnicity (%)	385 984 (80.4%)
Ejection fraction, mean±SD	50.6±14.5
BMI, kg/m <sup>2</sup>	29.4 ± (5.9)
Median CCS class (IQR)	3 [2, 4]
Previous PCI (%)	
1	115 200 (24%)
2	45 153 (9.4%)
3 or more	35 456 (7.4%)
History of cerebrovascular disease (%)	39 434 (8.2%)
History of peripheral vascular disease (%)	37 647 (7.8%)
History of heart failure (%)	19 279 (4%)
History of malignant ventricular arrhythmia (%)	2769 (0.6%)
History of COPD (%)	74 423 (15.5%)
History of diabetes mellitus (%)	161 771 (33.7%)
History of renal failure on dialysis (%)	10 456 (2.2%)
History of previous CABG (%)	79 075 (16.5%)
Hemodynamic instability (%)	2363 (0.5%)
ST-segment elevation on ECG	49 084 (10.2%)

BMI indicates body mass index; CABG, coronary artery bypass graft; CCS, Canadian Cardiovascular Society; COPD, chronic obstructive pulmonary disease; IQR, interquartile range (25th and 75th percentile); PCI, percutaneous coronary intervention.

## ML Analysis

### Variable selection

Variable importance plot was obtained after training on the training data set (80% of total cohort) using the tuned hyperparameters. Figure 1 and Figure S4 show ranking of the variables that are most significant in the prediction of in-hospital mortality in the studied cohort. Age was the most important predictor of in-hospital mortality, followed by ejection fraction, time (in days) since onset of myocardial ischemia/infarction, and body mass index. Notable variables that feature prominently are angiographically determined stenosis severity within the coronary vasculature, the occurrence of acute cerebrovascular events within 24 hours of PCI as well as day of the week on which PCI was performed.

### Prediction of in-hospital mortality

Among all 4 different ML methods (AdaBoost, XGBoost, Random Forest, and Logistic Regression), AdaBoost, which is short for Adaptive Boosting, revealed the highest performance for discrimination between survival/in-hospital mortality with an AUC of 0.927 (95% CI 0.923–0.929,  $P<0.05$ ) (Figure 2). XGBoost had similar discriminatory performance to that of logistic regression (XGBoost AUC of 0.913, 95% CI 0.906–0.919 compared with logistic regression AUC of 0.908, 95% CI 0.907–0.910,  $P=0.34$ ). Finally, Random Forest, representing a more established and traditional ML algorithm, had the lowest AUC of 0.892 (95% CI 0.889–0.896).

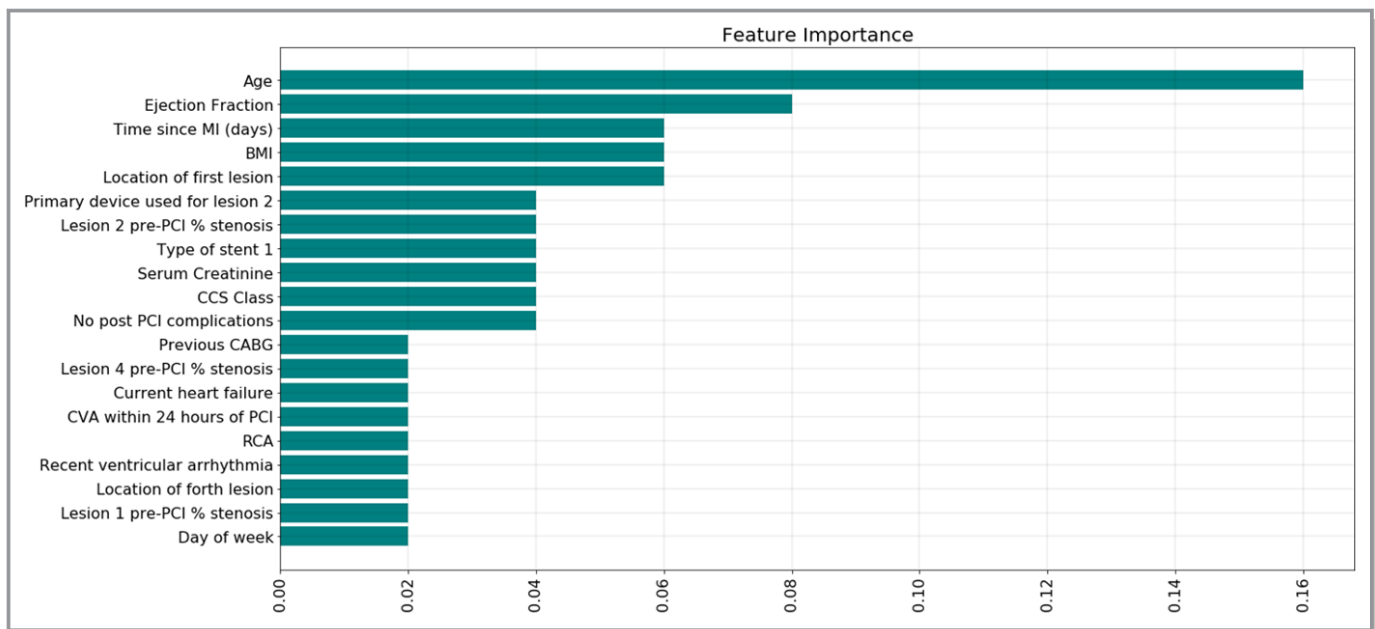
### Calibration of the prediction models

Calibration was performed on this 2-class classification task (determination of post-PCI in-hospital death or living status) in order to evaluate class-assignment probability distribution. The Brier score, measuring the accuracy of the probabilistic predictions, for AdaBoost to predict in-hospital mortality was 0.159, indicating a good fit of the ML-based model. Table 2 summarizes the Brier scores for the remainder of the models.

### Comparison of feature trends and model outcome

Figure 3 visually represents the effect of significant features on model prediction of the primary outcome (in-hospital death). Normalized continuous features (top panel) are plotted on the x-axis showing positive correlation between age (Pearson correlation coefficient of 0.31) and serum creatinine (Pearson correlation coefficient of 0.15) and negative correlation between ejection fraction (Pearson correlation coefficient  $-0.43$ ) and body mass index (Pearson correlation coefficient  $-0.11$ ). For categorical variables, box plots show the density distribution revealing the expected correlation between the occurrence of post PCI complications and the occurrence of in-hospital death ( $P<0.01$ ), as well as the





**Figure 1.** Feature importance ranking. This figure lists the relative importance of clinical and angiographic variables in the developed machine learning–based model for the prediction of in-hospital mortality after percutaneous coronary intervention (selected for the model with the highest area under the curve—AdaBoost. See Figure S4, for feature importance ranking with SD across 5-fold cross-validation). BMI indicates body mass index; CABG, coronary artery bypass grafting; CCS, Canadian Cardiovascular Society; CVA, cerebrovascular accident; MI, myocardial infarction; PCI, percutaneous coronary intervention; RCA, right coronary artery.

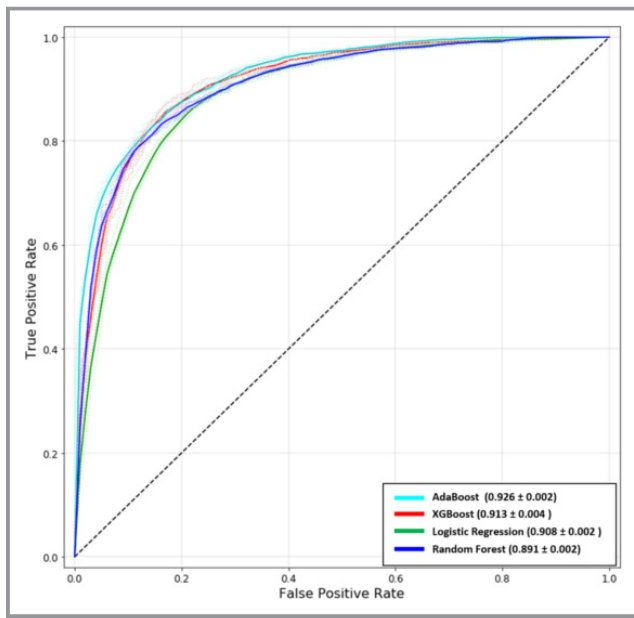
presence of heart failure at the time of PCI and Canadian Cardiovascular Society class and death ( $P<0.01$  for both). Interestingly, the occurrence of in-hospital mortality had a higher cluster on weekend days (Saturday and Sunday) compared with weekdays (Monday to Friday) ( $P<0.01$ ).

## Discussion

PCI has become one of the most common therapeutic procedures in modern cardiovascular practice. Additionally, the rapid pace of progress and increasing operator experience has resulted in a steady and sustained decline in periprocedural adverse events, resulting in excellent and comparable outcomes for coronary artery bypass surgery. Notwithstanding, differences between percutaneous and surgical revascularization exist and as a result, societal guidelines have highlighted the importance of risk stratification, which typically takes into consideration clinical and angiographic characteristics, for the administration of the appropriate therapy. To this end, numerous risk scores and prediction models have been developed using traditional statistical approaches that involve the inclusion of limited, single-center cohorts, the application of numerous exclusion criteria, inclusion of prespecified variables expected to be related to the outcome, and do not address the potential prognostic value of interactions between several unexpected weaker risk factors and the primary outcome.<sup>18</sup> In this study, we sought to

harness the power of big data analytics and ML in order to develop a ML-based prediction model for the occurrence of in-hospital mortality following PCI. Without the application of any exclusion criteria, we found that advanced ML algorithms accurately predict the occurrence of in-hospital death after PCI. We also found that several features, not typically incorporated in risk scores such as day of the week, demonstrate important prognostic value, in addition to already established variables such as body mass index, preprocedural serum creatinine, as well as several angiographic features related to lesion location and stenosis severity.

We compared the performance of advanced ML algorithms (AdaBoost and XGBoost) with traditional ML (Random Forest) and a statistical model, Logistic Regression. The main finding of the current analysis was that AdaBoost exhibited the highest discriminatory performance for the prediction of in-hospital mortality following PCI (AUC of 0.923,  $P<0.05$  compared with other models) compared with XGBoost (AUC of 0.906), Random Forest (AUC of 0.892), and Logistic Regression (AUC of 0.908). What is unique about the application of ML in this data set has been the ability to include all the patients and a significant proportion of the variables (with exclusion of nonrelevant or redundant variables) without the application of major exclusion criteria. Use of such a methodology for big data analysis in clinical research could contribute towards development of widely



**Figure 2.** Receiver operating curves. In this study, we trained 4 models: (1) AdaBoost (2), XGBoost (3), Logistic Regression, and (4) Random Forest. We performed 5-fold cross-validation on the data set for each model. The area-under-the-curve for all the models has been indicated as mean $\pm$ SD. AdaBoost was noted to have the best performance for prediction of in-hospital mortality after percutaneous coronary intervention.

applicable prediction models as well as an improvement in the ability to predict future events, which is a major focus in an era of precision medicine. Furthermore, the ability to agnostically incorporate a multitude of variables without preconceived notion of likely important predictors will ultimately help unravel novel associations between specific features and a specific outcome of interest.

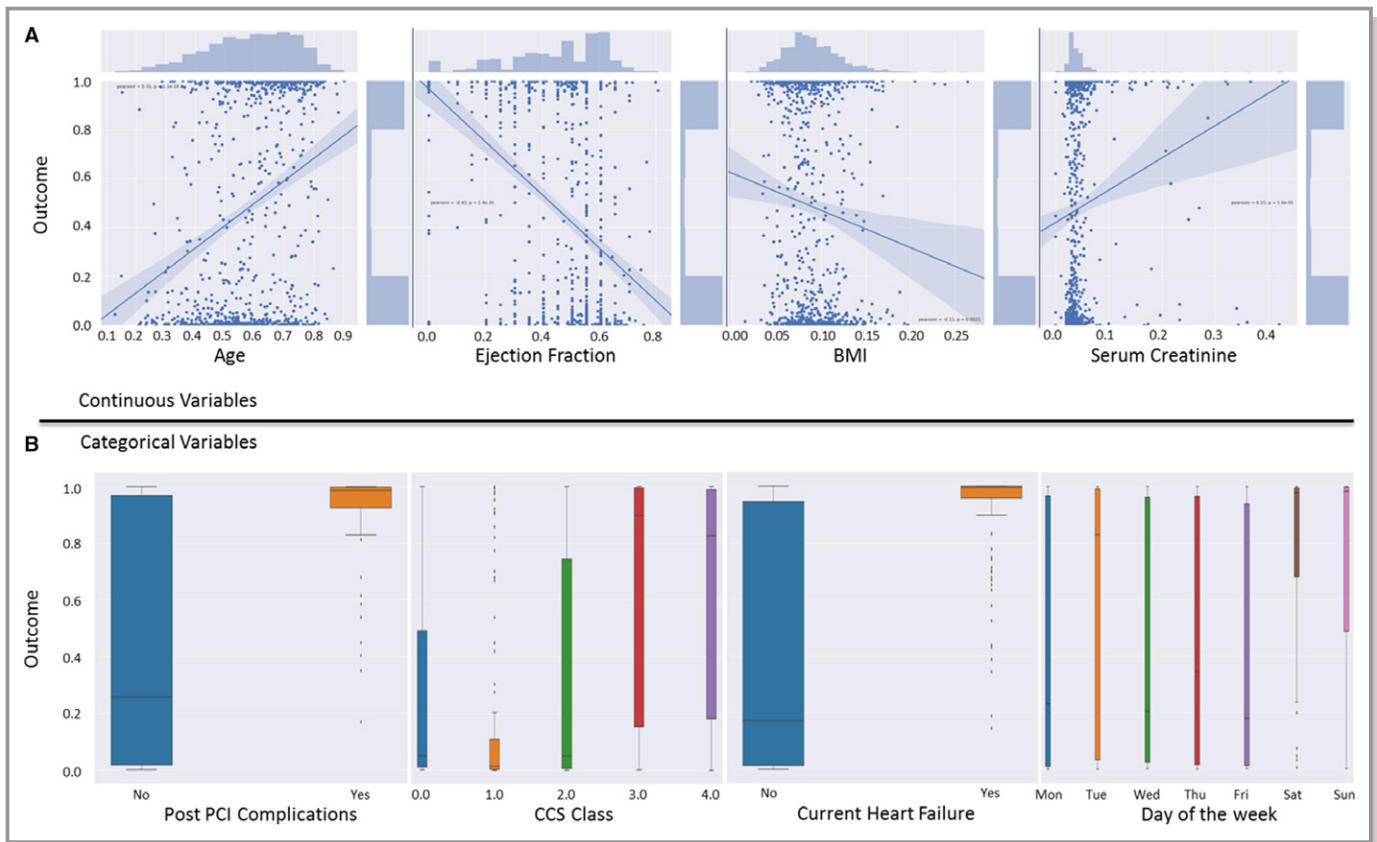
Various prediction models have been developed for in-hospital mortality following PCI and have been validated on different populations. The Mayo Clinic Risk Score, developed in 2002, was based on 5463 patients' data and predicts post PCI procedural complications using 8 clinical and angiographic variables.<sup>19</sup> The Mayo Clinic Risk Score was further

**Table 2.** Summary of the Brier Scores Evaluating the Calibration of the Machine Learning Models (AdaBoost, XGBoost, and Random Forest) as Well as That of Logistic Regression

Model	Brier Score
AdaBoost	0.159 $\pm$ 0.031
XGBoost	0.494 $\pm$ 0.091
Random Forest	0.084 $\pm$ 0.001
Logistic Regression	0.173 $\pm$ 0.045

validated using the National Heart, Lung, and Blood Institute registry, and was found to be an accurate predictor of in-hospital mortality when applied to the PCIRS registry (AUC of 0.85).<sup>20</sup> Separately, Wu et al utilized a logistic regression model to derive a prediction using 46 090 patients using the 2002 PCIRS data and subsequently performed validation of the model using the 2003 data.<sup>8</sup> The model was found to have high accuracy in discriminating events/nonevents when applied to the validation cohort (C-statistic of 0.886). The most significant variables in the model (variables with the largest  $\beta$  coefficient) were hemodynamic instability (odds ratio of 7.8), shock (odds ratio 19.91), and preprocedural myocardial infarction <24 hours with stent thrombosis (odds ratio 18.75). In this investigation, we show that ML, as a new analytic tool in outcomes research, improved the predictive accuracy even further, using a big database of patients undergoing PCI across New York State hospitals. Furthermore, we find that certain risk factors that were not included in previous risk scores are fairly important predictors of in-hospital mortality, such as serum creatinine, body mass index, and several angiographically determined coronary atherosclerotic characteristics. For instance, acute kidney injury following PCI has been consistently shown to be associated with the occurrence of both in-hospital and postdischarge adverse cardiovascular events.<sup>21–24</sup> Furthermore, the association between contrast-induced nephropathy and long-term mortality was found to be significant in patients with chronic kidney disease, as opposed to those without it.<sup>24</sup> To this effect, our results are in concordance with published literature and show that preprocedural serum creatinine was a strong predictor of the occurrence of in-hospital mortality. Additionally, numerous multi-ethnic investigations have studied the relationship between body mass index and outcomes following PCI.<sup>25–27</sup> While few studies showed increased 1- and 5-year major adverse cardiac events in obese patients,<sup>28,29</sup> there is a predominant notion that the presence of obesity is a protective factor in patients undergoing PCI, a phenomenon that has been termed the “obesity paradox.”<sup>30–34</sup> Yet, body mass index is not typically incorporated into most contemporary risk scores of in-hospital mortality.

As a result, the increasing application of artificial intelligence in cardiovascular research has been done in an attempt to circumvent the current limitations of existent approaches. Artificial intelligence has revolutionized society through innovations in various sectors of technology. ML, which is a subset of artificial intelligence, is a code-based algorithm that autonomously learns patterns within data and applies that knowledge to tasks that are provided. ML provides novel frameworks and a new approach to image interpretation and data analysis that is beyond what is provided with traditional statistical approaches.<sup>10</sup> ML has already been used to aid with detection in screening for breast cancer on mammography



**Figure 3.** Trend comparison between model outcome and clinically important features. Clinically important categorical and continuous features were plotted to understand their underlying trends in relation to the model outcome. **A**, Continuous variables have been plotted in a joint scatter and regression plot. The underlying trend between variables and the model outcome is shown for each variable. Normalized values for each variable are plotted on the x-axis. **B**, Categorical variables have been plotted using box plot. BMI indicates body mass index; CCS, Canadian Cardiovascular Society; PCI, percutaneous coronary intervention.

and to help with prediction of risk, echocardiographic image analysis, and electrocardiography interpretation.<sup>35–38</sup> Furthermore, the performance of ML has already been adopted to leverage big data in order to minimize biases and to accurately assess hospital performance after PCI.<sup>39</sup>

Despite the highlighted advantages of the proposed approach, there are several limitations worth mentioning. First, the developed model has not been externally validated on a separate cohort. Second, one of the challenges associated with ML is to avoid overfitting, a limitation that occurs when a predictive model is to perfectly fit the derivation cohort without accounting for generalizability, thereby producing a model that performs with high accuracy on the training tasks, while lacking such performance in the general population. To mitigate these issues in an unbiased manner, we performed k-fold cross-validation and reported the results as mean values (with SDs and *P* values). Nevertheless, it would still be imperative to validate the model on an external cohort. A third issue that comes into play is associated with an inherent limitation of existent databases. A large database such as the New York State PCI

registry defines certain variables in a specific way, which likely differs from other single- or multicenter databases. For instance, the PCIRS database defines postprocedural complications as a conglomerate of adverse events, encompassing vascular and nonvascular events (as defined in the Methods section), while the National Cardiovascular Data Registry breaks down post-PCI complications into several distinct categories. Such inconsistency across databases may limit the applicability and generalizability of developed predictive models, since every algorithm is limited by the quality of the ground truth that is being used for training and testing purposes. One potential solution would be to avoid categorization of variables, thereby providing more predictive power and limiting the subjectivity in variable definition (for instance, instead of different cutoff values for renal failure, one could include serum creatinine in a model as a continuous variable).

## Conclusion

In summary, this study sought to elucidate the determinants of in-hospital mortality in a large cohort of patients

undergoing PCI across New York State between 2004 and 2012, utilizing advanced ML algorithms that offer a novel approach to predictive modeling. We found that ML produced high discriminatory performance (specifically AdaBoost) while utilizing the cohort in its entirety and without the application of major exclusion criteria. Such findings, coupled with previously published work, could highlight the utility of using such an approach for the development of more precise and generalizable risk assessment in an era where risk stratification is becoming essential for the identification of optimal revascularization strategies.

## Sources of Funding

This work was supported by a generous fund from the Michael J. Wolk Foundation.

## Disclosures

Dr Min serves on the scientific advisory board of Arineta, on the speaker's bureau of GE Healthcare, and owns equity in Cleerly. The remaining authors have no disclosures to report.

## References

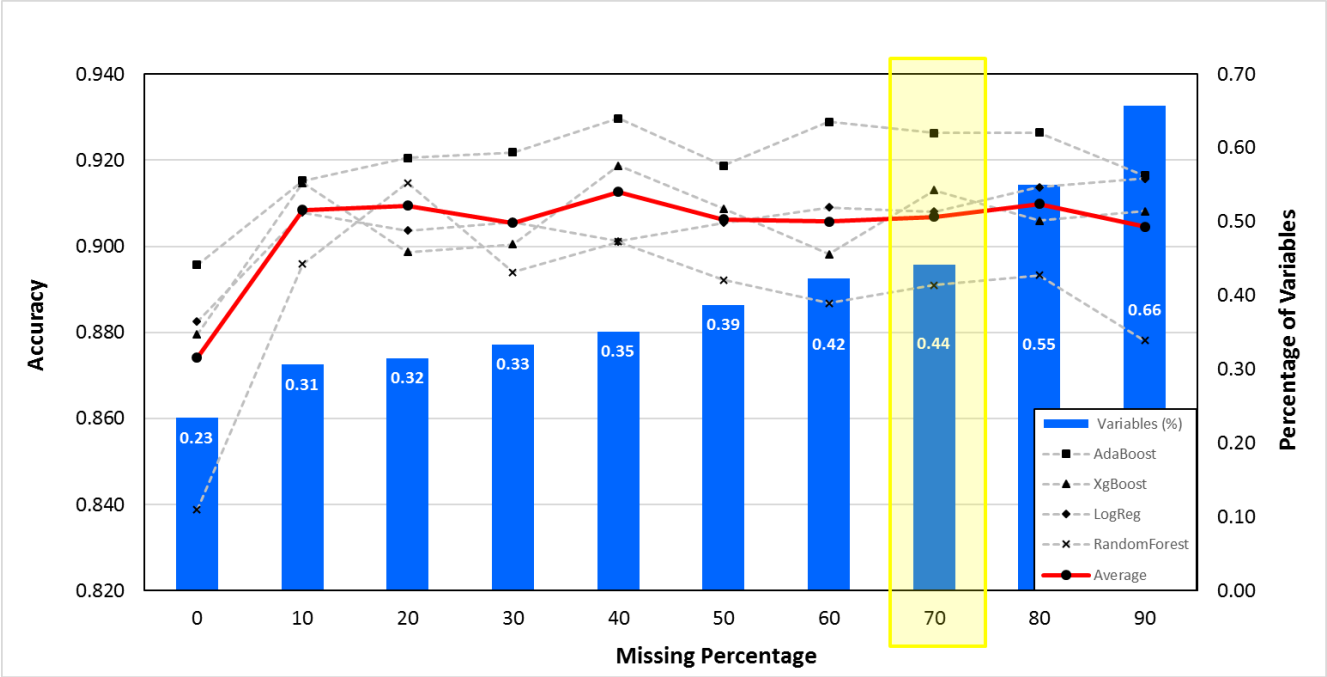
- Windecker S, Kolh P, Alfonso F, Collet JP, Cremer J, Falk V, Filippatos G, Hamm C, Head SJ, Jüni P, Kappetein AP, Kastrati A, Knuuti J, Landmesser U, Laufer G, Neumann FJ, Richter DJ, Schauerte P, Sousa Uva M, Stefanini GG, Taggart DP, Torracca L, Valgimigli M, Wijns W, Witkowski A. 2014 ESC/EACTS guidelines on myocardial revascularization: the Task Force on Myocardial Revascularization of the European Society of Cardiology (ESC) and the European Association for Cardio-Thoracic Surgery (EACTS). Developed with the special contribution of the European Association of Percutaneous Cardiovascular Interventions (EAPCI). *Eur Heart J*. 2014;35:2541–2619.
- Amsterdam EA, Wenger NK, Brindis RG, Casey DE Jr, Ganiats TG, Holmes DR Jr, Jaffe AS, Jneid H, Kelly RF, Kontos MC, Levine GN, Liebson PR, Mukherjee D, Peterson ED, Sabatine MS, Smalling RW, Zieman SJ. 2014 AHA/ACC guideline for the management of patients with non-ST-elevation acute coronary syndromes: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014;130:e344–e426.
- Fihn SD, Blankenship JC, Alexander KP, Bittl JA, Byrne JG, Fletcher BJ, Fonarow GC, Lange RA, Levine GN, Maddox TM, Naidu SS, Ohman EM, Smith PK. 2014 ACC/AHA/AATS/PCNA/SCAI/STS focused update of the guideline for the diagnosis and management of patients with stable ischemic heart disease: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines, and the American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. *Circulation*. 2014;130:1749–1767.
- Serruys PW, Morice MC, Kappetein AP, Colombo A, Holmes DR, Mack MJ, Stähle E, Feldman TE, van den Brand M, Bass EJ, Van Dyck N, Leadley K, Dawkins KD, Mohr FW; SYNTAX Investigators. Percutaneous coronary intervention versus coronary-artery bypass grafting for severe coronary artery disease. *N Engl J Med*. 2009;360:961–972.
- Yadav M, Palmerini T, Caixeta A, Madhavan MV, Sanidas E, Kirtane AJ, Stone GW, Généreux P. Prediction of coronary risk by SYNTAX and derived scores: synergy between percutaneous coronary intervention with taxus and cardiac surgery. *J Am Coll Cardiol*. 2013;62:1219–1230.
- Weintraub WS, Grau-Sepulveda MV, Weiss JM, DeLong ER, Peterson ED, O'Brien SM, Kolm P, Klein LW, Shaw RE, McKay C, Ritzenthaler LL, Popma JJ, Messenger JC, Shahian DM, Grover FL, Mayer JE, Garratt KN, Moussa ID, Edwards FH, Dangas GD. Prediction of long-term mortality after percutaneous coronary intervention in older adults: results from the National Cardiovascular Data Registry. *Circulation*. 2012;125:1501–1510.
- Hannan EL, Farrell LS, Walford G, Jacobs AK, Berger PB, Holmes DR Jr, Stamato NJ, Shama S, King SB III. The New York State risk score for predicting in-hospital/30-day mortality following percutaneous coronary intervention. *JACC Cardiovasc Interv*. 2013;6:614–622.
- Wu C, Hannan EL, Walford G, Ambrose JA, Holmes DR Jr, King SB III, Clark LT, Katz S, Sharma S, Jones RH. A risk score to predict in-hospital mortality for percutaneous coronary interventions. *J Am Coll Cardiol*. 2006;47:654–660.
- Singh G, Al'Aref SJ, Van Assen M, Kim TS, van Rosendaal A, Kolli KK, Dwivedi A, Maliakal G, Pandey M, Wang J, Do V, Gummalla M, De Cecco CN, Min JK. Machine learning in cardiac CT: basic concepts and contemporary data. *J Cardiovasc Comput Tomogr*. 2018;12:192–201.
- Krittana Wong C, Zhang H, Wang Z, Aydar M, Kitai T. Artificial intelligence in precision cardiovascular medicine. *J Am Coll Cardiol*. 2017;69:2657–2664.
- Takx RA, de Jong PA, Leiner T, Oudkerk M, de Koning HJ, Mol CP, Viergever MA, Išgum I. Automated coronary artery calcification scoring in non-gated chest CT: agreement and reliability. *PLoS One*. 2014;9:e91239.
- Frizzell JD, Liang L, Schulte PJ, Yancy CW, Heidenreich PA, Hernandez AF, Bhatt DL, Fonarow GC, Laskey WK. Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: comparison of machine learning and other statistical approaches. *JAMA Cardiol*. 2017;2:204–209.
- Al'Aref SJ, Singh G, Rosendaal ARV, Kolli KK, Ma X, Maliakal G, Pandey M, Lee BC, Wang J, Xu Z, Zhang Y, Min JK, Wong SC, Minutello RM. Determinants of in-hospital mortality after percutaneous coronary intervention: a machine learning approach. GitHub repository. 2018. Available at: <https://github.com/Gurpreethgnis/NYS-PCI>. Accessed December 26, 2018.
- Hannan EL, Cozzens K, King SB III, Walford G, Shah NR. The New York State cardiac registries: history, contributions, limitations, and lessons for future efforts to assess and publicly report healthcare outcomes. *J Am Coll Cardiol*. 2012;59:2309–2316.
- Al'Aref SJ, Anouchou K, Singh G, Slomka PJ, Kolli KK, Kumar A, Pandey M, Maliakal G, van Rosendaal AR, Beecy AN, Berman DS, Leipsic J, Nieman K, Andreini D, Pontone G, Schoepf UJ, Shaw LJ, Chang HJ, Narula J, Bax JJ, Guan Y, Min JK. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *Eur Heart J*. 2018. Available at: <https://academic.oup.com/eurheartj/advance-article/doi/10.1093/eurheartj/ehy404/5060564>. Accessed January 17, 2019.
- Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons; 1987.
- Rubin D. Multiple imputation after 18 years. *J Am Stat Assoc*. 1996;91:473–489.
- Mallett S, Royston P, Dutton S, Waters R, Altman DG. Reporting methods in studies developing prognostic models in cancer: a review. *BMC Med*. 2010;8:20.
- Singh M, Lennon RJ, Holmes DR Jr, Bell MR, Rihal CS. Correlates of procedural complications and a simple integer risk score for percutaneous coronary intervention. *J Am Coll Cardiol*. 2002;40:387–393.
- Kumar SS, Negassa A, Monrad ES, Srinivas VS. The Mayo Clinic Risk Score predicts in-hospital mortality following primary angioplasty. *J Invasive Cardiol*. 2005;17:522–526.
- Valle JA, McCoy LA, Maddox T, Rumsfeld JS, Ho PM, Casserly IP, Nallamothu BK, Roe MT, Tsai TT, Messenger JC. Longitudinal risk of adverse events in patients with acute kidney injury after percutaneous coronary intervention: insights from the National Cardiovascular Data Registry. *Circ Cardiovasc Interv*. 2017;10:e004439.
- Narula A, Mehran R, Weisz G, Dangas GD, Yu J, Généreux P, Nikolsky E, Brenner SJ, Witzensbichler B, Guagliumi G, Clark AE, Fahy M, Xu K, Brodie BR, Stone GW. Contrast-induced acute kidney injury after primary percutaneous coronary intervention: results from the HORIZONS-AMI substudy. *Eur Heart J*. 2014;35:1533–1540.
- Fox CS, Muntner P, Chen AY, Alexander KP, Roe MT, Wiviott SD. Short-term outcomes of acute myocardial infarction in patients with acute kidney injury: a report from the National Cardiovascular Data Registry. *Circulation*. 2012;125:497–504.
- Abe M, Morimoto T, Akao M, Furukawa Y, Nakagawa Y, Shizuta S, Ehara N, Taniguchi R, Doi T, Nishiyama K, Ozasa N, Saito N, Hoshino K, Mitsuoka H, Toma M, Tamura T, Haruna Y, Kita T, Kimura T. Relation of contrast-induced nephropathy to long-term mortality after percutaneous coronary intervention. *Am J Cardiol*. 2014;114:362–368.
- Ellis SG, Elliott J, Horrigan M, Raymond RE, Howell G. Low-normal or excessive body mass index: newly identified and powerful risk factors for death and other complications with percutaneous coronary intervention. *Am J Cardiol*. 1996;78:642–646.
- Holroyd EW, Sirker A, Kwok CS, Kontopantelis E, Ludman PF, De Belder MA, Butler R, Cotton J, Zaman A, Mamas MA. The relationship of body mass index to percutaneous coronary intervention outcomes: does the obesity paradox



- exist in contemporary percutaneous coronary intervention cohorts? Insights from the British Cardiovascular Intervention Society Registry. *JACC Cardiovasc Interv.* 2017;10:1283–1292.
27. Numasawa Y, Kohsaka S, Miyata H, Kawamura A, Noma S, Suzuki M, Nakagawa S, Momiyama Y, Naito K, Fukuda K. Impact of body mass index on in-hospital complications in patients undergoing percutaneous coronary intervention in a Japanese real-world multicenter registry. *PLoS One.* 2015;10:e0124399.
  28. Sarno G, Garg S, Onuma Y, Buszman P, Linke A, Ischinger T, Klaus V, Eberli F, Corti R, Wijns W, Morice MC, di Mario C, van Geuns RJ, Eerdmans P, Garcia-Garcia HM, van Es GA, Goedhart D, de Vries T, Jüni P, Meier B, Windecker S, Serruys P. The impact of body mass index on the one year outcomes of patients treated by percutaneous coronary intervention with Biolimus- and Sirolimus-eluting stents (from the LEADERS Trial). *Am J Cardiol.* 2010;105:475–479.
  29. Sarno G, Räber L, Onuma Y, Garg S, Brugaletta S, van Domburg RT, Pilgrim T, Pfäffli N, Wenaweser P, Windecker S, Serruys P. Impact of body mass index on the five-year outcome of patients having percutaneous coronary interventions with drug-eluting stents. *Am J Cardiol.* 2011;108:195–201.
  30. Angerås O, Albertsson P, Karason K, Råmunddal T, Matejka G, James S, Lagerqvist B, Rosengren A, Omerovic E. Evidence for obesity paradox in patients with acute coronary syndromes: a report from the Swedish Coronary Angiography and Angioplasty Registry. *Eur Heart J.* 2013;34:345–353.
  31. Gruberg L, Weissman NJ, Waksman R, Fuchs S, Deible R, Pinnow EE, Ahmed LM, Kent KM, Pichard AD, Suddath WO, Satler LF, Lindsay J Jr. The impact of obesity on the short-term and long-term outcomes after percutaneous coronary intervention: the obesity paradox? *J Am Coll Cardiol.* 2002;39:578–584.
  32. Mak KH, Bhatt DL, Shao M, Haffner SM, Hamm CW, Hankey GJ, Johnston SC, Montalescot G, Steg PG, Steinhubl SR, Fox KA, Topol EJ. The influence of body mass index on mortality and bleeding among patients with or at high-risk of atherothrombotic disease. *Eur Heart J.* 2009;30:857–865.
  33. Kaneko H, Yajima J, Oikawa Y, Tanaka S, Fukamachi D, Suzuki S, Sagara K, Otsuka T, Matsuno S, Funada R, Kano H, Uejima T, Koike A, Nagashima K, Kirigaya H, Sawada H, Aizawa T, Yamashita T. Obesity paradox in Japanese patients after percutaneous coronary intervention: an observation cohort study. *J Cardiol.* 2013;62:18–24.
  34. Sharma A, Vallakati A, Einstein AJ, Lavine CJ, Arbab-Zadeh A, Lopez-Jimenez F, Mukherjee D, Lichstein E. Relationship of body mass index with total mortality, cardiovascular mortality, and myocardial infarction after coronary revascularization: evidence from a meta-analysis. *Mayo Clin Proc.* 2014;89:1080–1100.
  35. Gilbert FJ, Astley SM, Gillan MG, Agbaje OF, Wallis MG, James J, Boggis CRM, Duffy CW. Single reading with computer-aided detection for screening mammography. *N Engl J Med.* 2008;359:1675–1684.
  36. Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. *N Engl J Med.* 2016;375:1216–1219.
  37. Tajik AJ. Machine learning for echocardiographic imaging: embarking on another incredible journey. *J Am Coll Cardiol.* 2016;68:2296–2298.
  38. Afsar FA, Arif M, Yang J. Detection of ST segment deviation episodes in ECG using KLT with an ensemble neural classifier. *Physiol Meas.* 2008;29:747–760.
  39. Spertus JV, T Normand SL, Wolf R, Cioffi M, Lovett A, Rose S. Assessing hospital performance after percutaneous coronary intervention using big data. *Circ Cardiovasc Qual Outcomes.* 2016;9:659–669.

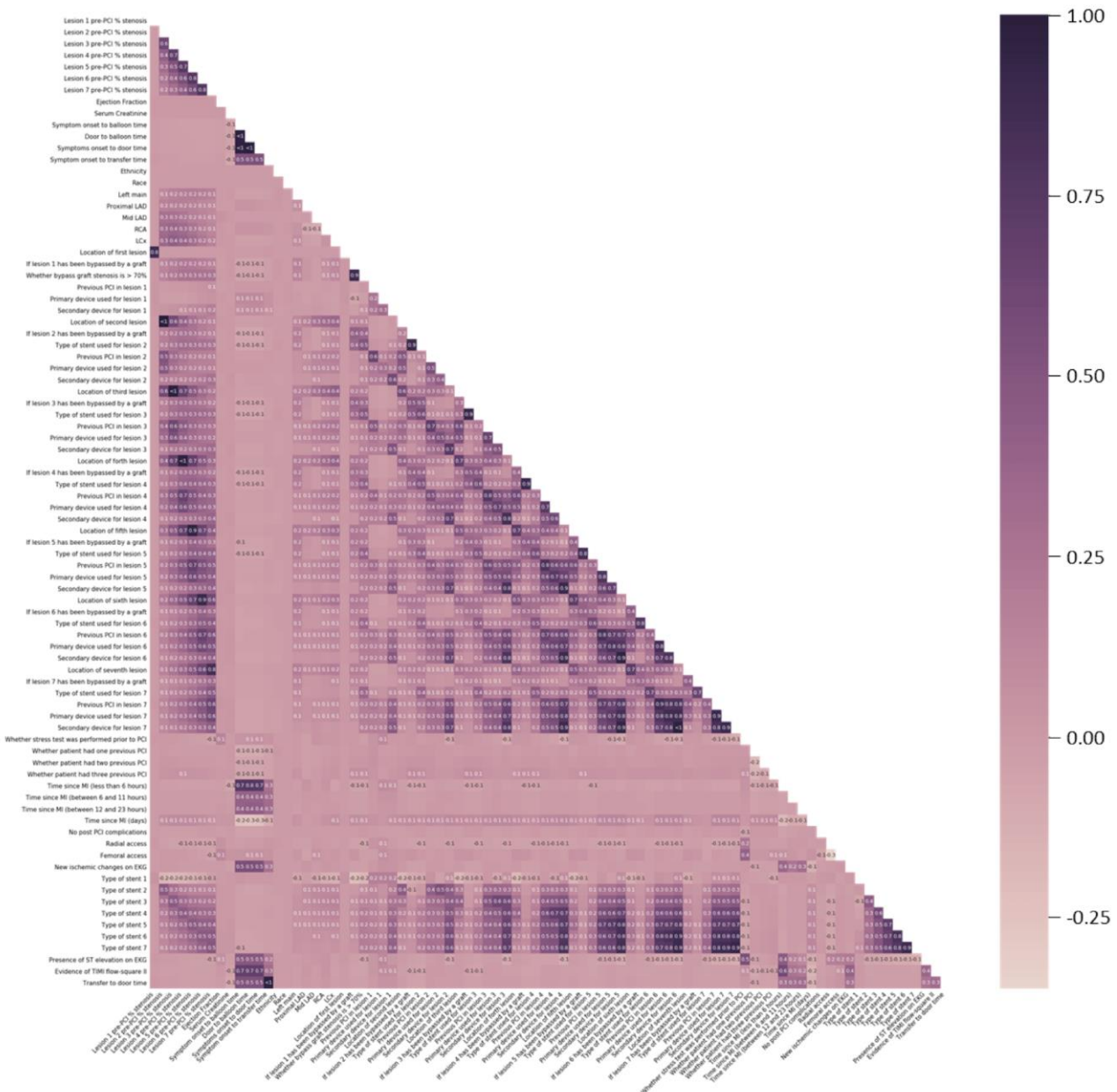
# **Supplemental Material**

Figure S1. Comparison between model accuracy and variable missingness.



Blue bars represent the proportion of variables meeting inclusion criteria stratified by missingness percentage (i.e. when variables with  $\leq 90\%$  missing values are included in the analysis, 66% of the variables meet the inclusion criteria). In the present analysis, 70% missingness was used as the cutoff value since it represented a balance between overall model performance as well inclusion of a larger proportion of variables (yellow highlight).

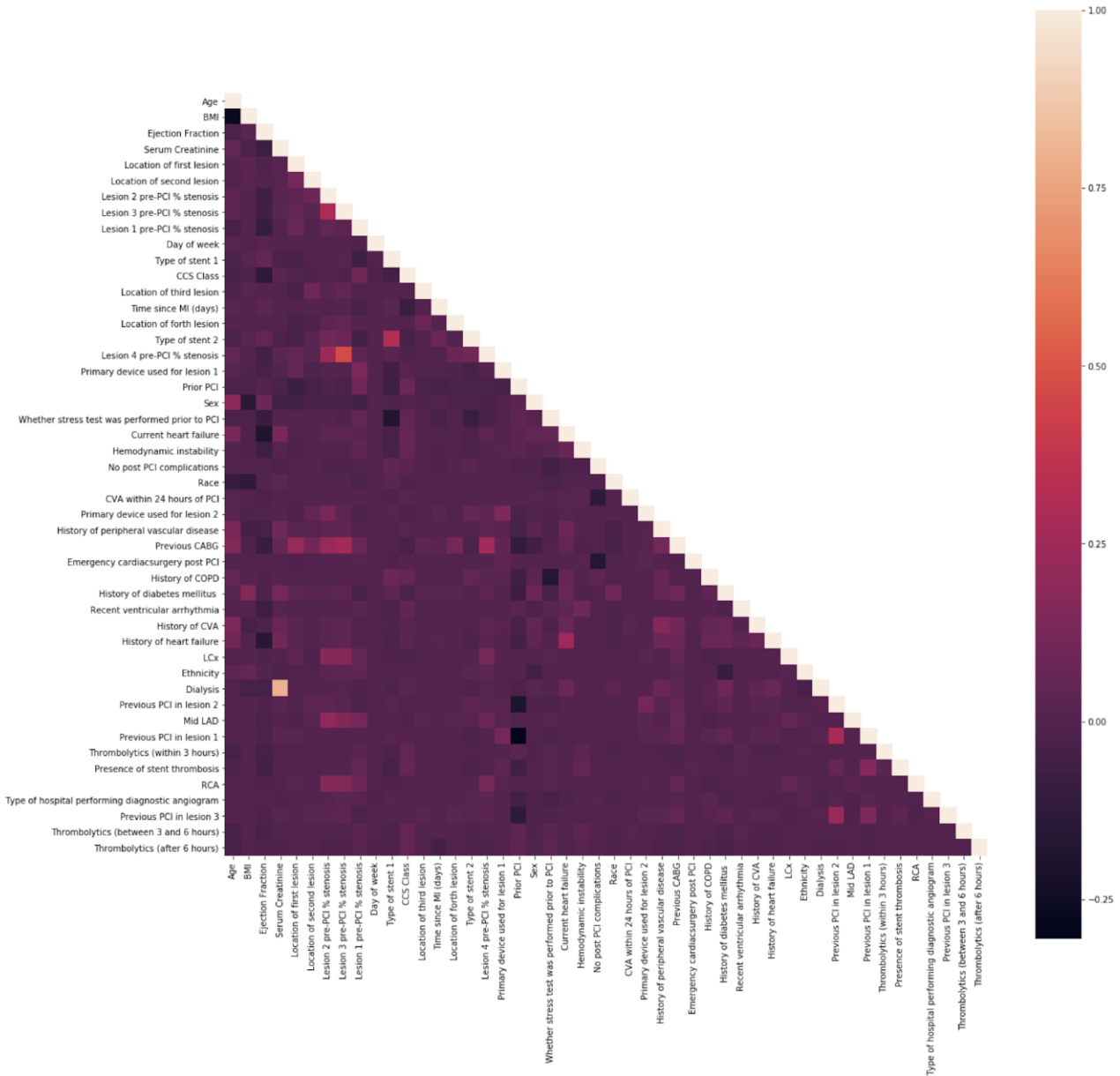
Figure S2. Nullity correlation Heat map.



The nullity correlation ranges from -1 to 1. A value of -1 indicates that if one variable appears the other definitely does not. A value of '0' indicates that variables appearing or not appearing have no effect on one another whereas a value of '1' indicates that if one variable appears the other definitely also does. PCI: percutaneous coronary intervention; LAD: left anterior descending; RCA: right coronary artery; LCx: left circumflex; MI: myocardial infarction; EKG: electrocardiogram; TIMI: thrombolysis in myocardial infarction.

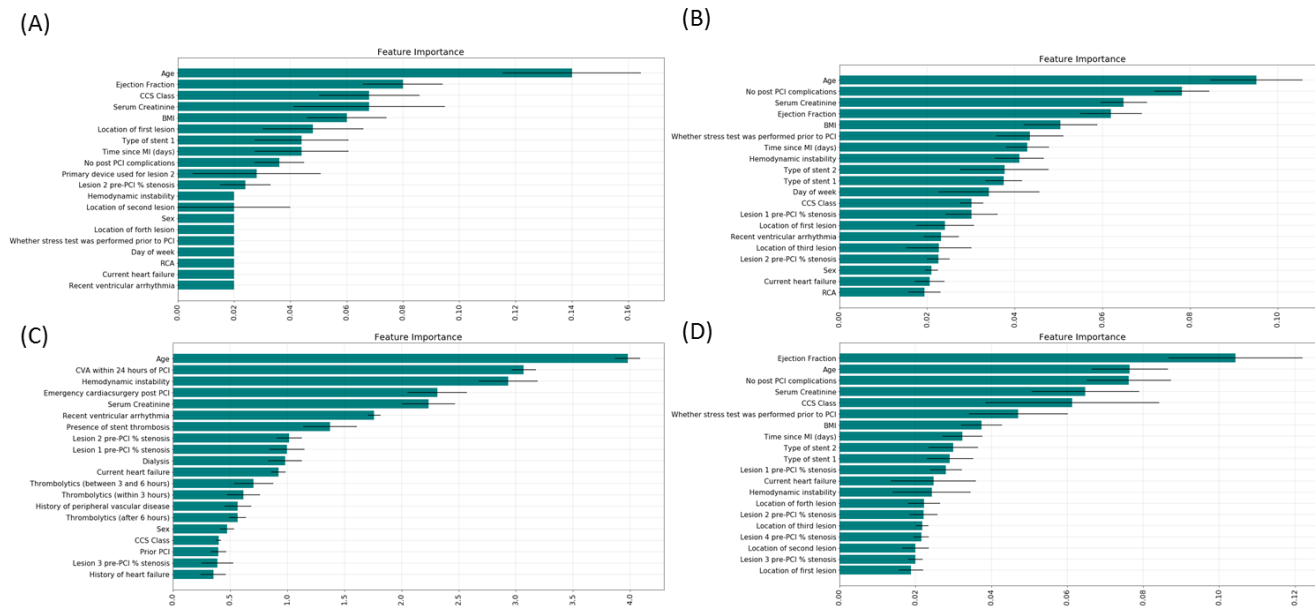


**Figure S3. Correlation analysis of the variables in the dataset.**



A simpler model is always more stable and hence preferred. Inclusion of highly correlated variables could contribute to numerical instability, mask interactions between different features, collude the interpretability of a machine learning model and also may cause overfitting. Thus, it is always preferred to exclude one of the two correlated variables. The heat map shown above indicates that dataset does not contain many correlated variables. BMI: body mass index; PCI: percutaneous coronary intervention; CCS: Canadian Cardiovascular Society; MI: myocardial infarction; CVA: cerebrovascular accident; CABG: coronary artery bypass grafting; COPD: chronic obstructive pulmonary disease; Lx: left circumflex; LAD: left anterior descending; RCA: right coronary artery.

**Figure S4. Feature importance ranking and the association standard deviations across 5-fold cross validation.**



Standard deviation, in black bars, is provided along with the feature importance ranking for **(A)** AdaBoost **(B)** XGBoost **(C)** Logistic Regression and **(D)** Random Forest. CCS: Canadian Cardiovascular Society; BMI: body mass index; MI: myocardial infarction; PCI: percutaneous coronary intervention; RCA: right coronary artery; CVA: cerebrovascular accident.