

## Research Article

# From Valves to Vessels: A Machine Learning Approach to Explore Heart Failure Risk in ICU Patients with Aortic Valve and Aortic Vascular Disorders

Karamo Bah<sup>1</sup>; Adama Ns Bah<sup>2</sup>; Amadou Wurry Jallow<sup>3</sup>; Musa Touray<sup>4\*</sup>

<sup>1</sup>Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, Taiwan

<sup>2</sup>Graduate Institute of Biomedical Informatics, College of Medical Science and Technology, Taipei Medical University, Taiwan

<sup>3</sup>Department of Medical Laboratory Science and Biotechnology, Taipei Medical University, Taiwan

<sup>4</sup>School of Medicine and Allied Health Sciences, University of The Gambia, The Gambia

\*Corresponding author: Dr. Musa Touray

Senior Lecturer School of Medicine and Allied Health Sciences University of the Gambia, West Africa.

Email: musatouray@utg.edu.gm

Received: October 25, 2023

Accepted: November 29, 2023

Published: December 06, 2023

## Introduction

Heart Failure (HF) is a clinical syndrome characterized by the reduced ability of the heart to pump and/or fill with blood. In 2021, a consensus on the universal definition and classification of HF was proposed, defining HF as a clinical syndrome with symptoms and/or signs caused by cardiac abnormality [1]. HF was categorized based on left ventricular Ejection Fraction (EF) into HF with reduced (HFrEF), mildly reduced (HFmrEF), and preserved EF (HFpEF). A new entity, HF with improved EF, was also introduced. HF is considered a global pandemic, affecting an estimated 64.3 million people worldwide in 2017, with prevalence expected to rise due to improved survival and longer life expectancy. The burden of HF on healthcare expenditures is significant, with projections indicating a substantial increase in costs by 2030 [2]. The prevalence of Heart Failure (HF) varies

Austin Journal of Nephrology and Hypertension  
Volume 10, Issue 2 (2023)

[www.austinpublishinggroup.com](http://www.austinpublishinggroup.com)

Touray M © All rights are reserved

## Abstract

**Background and objectives:** Aortic valve and aortic vascular disorders represent a subset of cardiovascular conditions that can lead to heart failure. Aortic valve diseases, such as aortic stenosis or regurgitation, and aortic vascular diseases, including aortic aneurysms and aortic dissection, can contribute to impaired cardiac function and increase the risk of heart failure. This study aims to investigate the risk of heart failure in patients with aortic valve and aortic vascular disorders within 30-days of admission in the Intensive Care Unit (ICU).

**Methods:** Patients from a US-based critical care database (MIM-IC-III) who developed heart failure in the ICU within 30 days of follow-up were included. Two predictive models, XGBoost and logistic regression, were developed and evaluated using ROC, sensitivity, specificity, and F1 measure. The dataset was split into training and testing samples in an 8:2 ratio.

**Results:** Out of 2,871 patients analyzed, 1,062 (37%) developed heart failure in the ICU during the 30-day follow-up. Key predictors of heart failure included creatinine, phosphate, age, COPD, INR, diabetes, CAD, magnesium, atrial fibrillation, and hyperlipidemia. The logistic regression model outperformed XGBoost with (AU-ROC, 0.78 vs. 0.77, respectively).

**Conclusions:** This study demonstrates the potential of machine learning techniques to enhance predictive modeling in critical care research. It provides valuable insights into heart failure risk in patients with aortic valve and aortic vascular disorders admitted to the ICU.

**Keywords:** Aortic disorders; Heart failure; Intensive care unit; Length of stay; Mortality; Machine learning

significantly between countries and regions, with the highest rates observed in Central Europe, North Africa, and the Middle East (ranging from 1133 to 1196 per 100,000 people) and lower rates in Eastern Europe and Southeast Asia (ranging from 498 to 595 per 100,000 people) [3]. In the next sections, we present a summary of epidemiological data on HF prevalence, focusing on various geographical areas. Heart failure is a significant and complex cardiovascular condition that poses a substantial burden on public health globally. It affects millions of people and is associated with high morbidity and mortality rates. While heart failure is commonly studied in the general population, there is a need for more focused investigations in specific patient cohorts with underlying cardiovascular disorders.

**Citation:** Bah K, NS Bah A, Jallow AW, Touray M. From Valves to Vessels: A Machine Learning Approach to Explore Heart Failure Risk in ICU Patients with Aortic Valve and Aortic Vascular Disorders. Austin J Nephrol Hypertens. 2023; 10(2): 1110.

Machine learning has made significant advancements in healthcare. AI is being used to aid in case triage and diagnoses [4], improve image scanning and segmentation [5], assist with decision-making [6], **predict disease risk [7,8], and even in neuroimaging [9]**. These applications have the potential to revolutionize healthcare and improve patient outcomes. Researchers have developed deep learning models to predict clinical conditions using Electronic Health Records (EHRs). One study utilized LSTM networks and CNNs to predict diseases like heart failure and stroke, achieving improved accuracy by incorporating both structured and unstructured data from progress and diagnosis notes [10]. In another study, a deep neural network model predicted post-stroke pneumonia with high accuracy, reaching 92.8% and 90.5% AUC for 7-day and 14-day predictions, respectively [7]. Additionally, ML-based models, such as SRML-Mortality Predictor, demonstrated the ability to predict mortality in specific conditions, like paralytic ileus, with an 81.30% accuracy rate [11]. These predictive algorithms can provide valuable insights for informed clinical decision-making.

The aim of this study is to investigate the risk of heart failure in patients with aortic valve and aortic vascular disorders within **30-days of admission in the Intensive Care Unit (ICU)**. The secondary outcomes are to assess the 30-day mortality rate, the Length of Hospital Stay (LoS) and to investigate the clinical implications of the machine learning predictions for improving patient outcomes in this population.

**Method**

**Data Source**

In this retrospective investigation, we analyzed data retrieved from the Medical Information Mart for Intensive Care (MIMIC) repositories. MIMIC databases contain extensive and anonymized health-related information of critical care patients admitted to the Beth Israel Deaconess Medical Center, a prominent tertiary medical facility in Boston, USA. The dataset encompasses various variables such as demographics, vital signs, laboratory outcomes, prescriptions, and clinical notes, providing valuable insights into critical patient profiles [12]. In this study, we conducted an analysis of the MIMIC databases, specifically focusing on the most recent version, MIMIC-III v1.4. The MIMIC-III clinical database encompasses data collected between 2001 and 2012, utilizing the MetaVision (iMDSoft, Wakefield, MA, USA) and CareVue (Philips Healthcare, Cambridge, MA, USA) systems. Notably, the original Philips CareVue system, comprising archived data from 2001 to 2008, was subsequently replaced by the advanced MetaVision data management system, which remains in active use today.

**Ethics and Data Use Agreement**

After successfully completing the mandatory online human research ethics training as mandated by PhysioNet Clinical Databases (Certification Number: 55140935), we obtained data access following the prescribed procedures. The study was conducted in accordance with the Declaration of Helsinki.

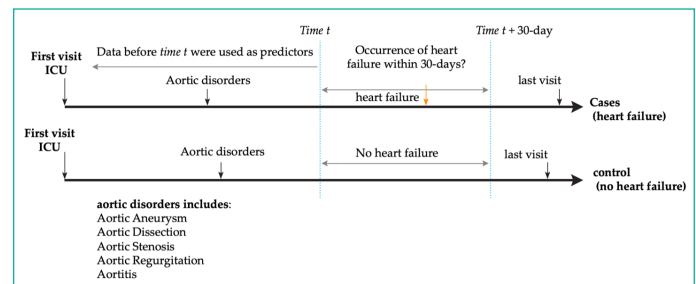
**Definition of the Outcome of Interest (cases and controls)**

In the context of the study conducted in ICU patients to explore the risk of heart failure after having aortic disorders or aortic vascular problems within a 30-day follow-up period, the definitions of cases and controls would be as follows:

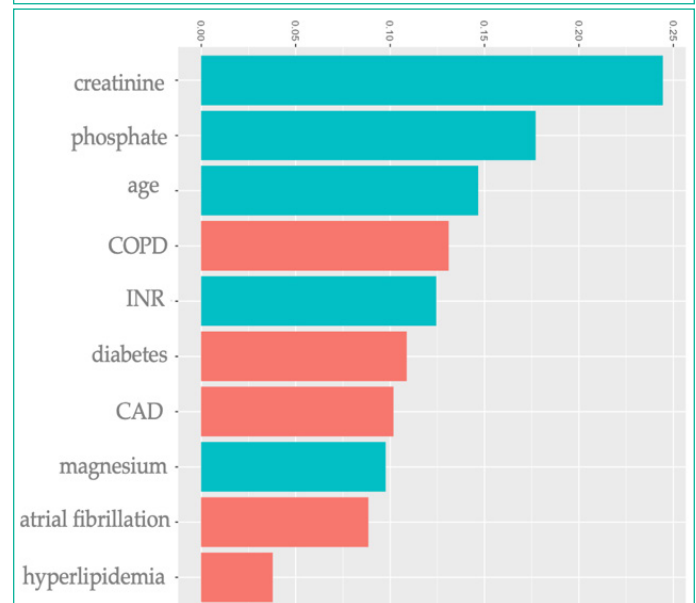
Cases in this study refer to ICU patients who were diagnosed with aortic disorders (such as aortic valve disorders or aortic

vascular problems) and subsequently developed heart failure within the 30-day follow-up period. These are individuals who experienced heart failure as an outcome of interest during their stay in the ICU. Cases were selected based on the International Classification of Diseases (ICD) procedure codes (ICD-9 CODE) from the MIMIC III data. Controls, in this study, is defined as ICU patients with aortic disorders who did not develop heart failure during the 30-day follow-up period. They are individuals who did not experience the outcome of interest (heart failure) during their ICU stay and within the 30-day time frame (Figure 1).

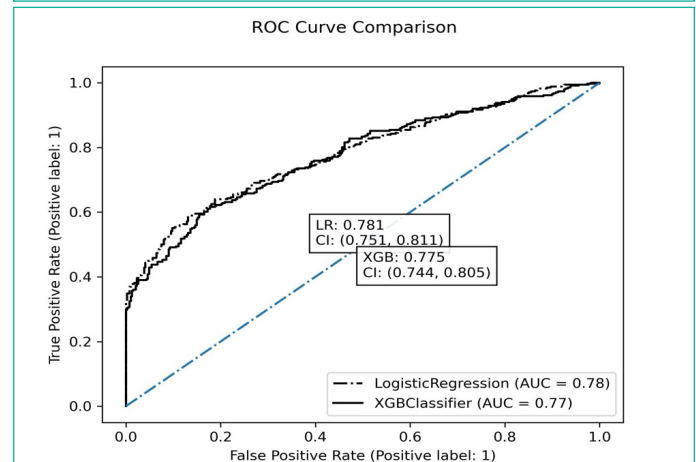
The study followed a cohort design, where a group of ICU patients with aortic disorders were observed during their stay in the ICU and for the subsequent 30 days, then the patients were categorized into cases or controls based on whether they developed heart failure within the 30-day follow-up period. By comparing the characteristics, comorbidities, and clinical outcomes between cases and controls, researchers can investigate



**Figure 1:** Timeline of study period schema.



**Figure 2:** Top 10 important features from XGBoost model.



**Figure 3:** Area under receiver operating characteristics curves for logistic regression and XGBoost model.

the association between aortic disorders and the risk of heart failure in this specific cohort of critically ill patients. This approach allows for the identification of factors associated with an increased risk of heart failure in ICU patients with aortic disorders, which can have important implications for patient care and management in critical care settings. Finally, we labeled the data as cases (ICU patients with heart failure) ( $n = 1,062$ ) and controls (ICU patients without heart failure), ( $n = 1,809$ ).

### Input Variables

In this study, we analyzed routinely collected demographic, clinical, and laboratory variables obtained during ICU admission. The candidate features included both static and dynamic information. Patient information encompassed age and sex. Laboratory measurements consisted of Complete Blood Count (CBC) features like hematocrit, MCH and platelet count, as well as chemistry measurements such as potassium, creatinine, calcium, magnesium, and phosphate. Coagulation measurements included partial international normalized ratio and comorbidities selected were diabetes, atrial fibrillation, hypertension, CAD, COPD, hyperlipidemia, and obesity.

Two machine learning models, XGBoost, and Logistic regression, were developed for analysis. To mitigate bias, variables with more than 30% missing values were excluded from further analysis. For variables with fewer missing values, multiple imputation methods were applied [13].

### Statistical Analysis

Clinical characteristics between cases and control groups were compared using either Student t test or rank-sum test as appropriate. Chi-square test or Fisher's exact test was employed to compare the differences of the categorical variables [14]. A statistically significant was measured with a  $p$  value of  $<0.05$ . A stepwise logistic regression model was used to select variables which were predictive of cases in ICU. Both forward selection and backward elimination were used, testing at each step for variables to be included or excluded. Akaike Information Criterion (AIC) was used as the selection criteria to eliminate the predictors [13].

### Extreme Gradient Boosting (XGBoost) Model

The eXtreme Gradient Boosting (XGBoost) is a powerful boosting application that combines multiple learning models to achieve superior performance [15]. In this study, we employed XGBoost with decision trees as weak learners and binary logistic objective function to predict cases versus controls [16]. XGBoost, introduced by Chen Tianqi and Carlos in 2011, has been continuously improved by researchers for subsequent studies [17]. The model utilizes a gradient descent optimization approach to minimize the loss function [18]. The boosting method iteratively refits weak classifiers (decision trees) to residuals of previous models, focusing on misclassified observations in each round of fitting [17,19,20]. A detailed information regarding the XGBoost model can be found in the literature [15,21]. We used a loop function (grid search) to select the hyperparameters for our analysis.

### Model Evaluation

We present essential evaluation metrics for assessing our machine learning models. These metrics are derived from True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) classifications.

**Sensitivity:** This metric is expressed as a probability, representing the classifier's ability to correctly predict a positive result when the corresponding ground truth is also positive. Another term for this metric is the true positive rate (TPR), and it is computed using the following formula [22]:

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

**Specificity:** The specificity, also known as the True Negative Rate (TNR), is a probability measure indicating the classifier's ability to correctly predict negative outcomes when the corresponding ground truth is also negative. It is calculated as follows [23]:

$$\text{Specificity} = \frac{TN}{TN + FP}$$

**AU-ROC:** The area under the Receiver Operating Characteristic (ROC) curve is a significant criterion for evaluating classifiers. It is derived from the plot of True Positive Rate (TPR) against False Positive Rate (FPR). The calculation of this metric is as follows [24]:

$$AU - ROC = \frac{1}{2} \left( \frac{TP}{TP + FN} + \frac{TN}{TN + FP} \right)$$

**F1-Score:** The F1-Score is a composite measure that incorporates both precision and sensitivity, represented as their weighted average. An F1-Score of 1 indicates the best performance, while an F1-Score of 0 is considered the worst. The F1-Score is calculated using the following formula [25]:

$$F1 - Score = 2 \times \left( \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \right)$$

**Accuracy:** The presented scale holds significant importance as it is commonly used for evaluating classifiers. It represents the percentage of samples correctly classified by the classifier. The calculation is as follows [26]:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

## Results

### Participants

Among the 2,871 patients with aortic disorders after ICU admission, 1,062 patients (37%) developed heart failure within 30 days after ICU admission and were categorized as cases, while 1,809 patients did not develop heart failure during the follow-up and were classified as controls. Patients with a history of heart failure before the follow-up period were excluded from the analysis.

Table 1 presents the differences in characteristics between the cases and control groups. The average time for aortic disorder patients to develop heart failure in the ICU after admission was 4 days ( $SD \pm 4.5$ ). Those who developed the outcome were older, with a median age of 77 years. In both cases and controls, there were more males than females (650, 61% vs. 1,120, 62%, respectively;  $p$ -value  $<0.000$ ).

### Assessment of Clinical Outcomes

Clinical outcomes including duration of hospital stay and mortality were assessed in two groups of patients (Table 2). Although the above outcomes seemed to be higher in the cases group, there was no evident difference ( $P < 0.05$ ).

### The XGBoost Model Model and Feature Importance

We used specific settings (hyperparameters) for our analysis, like learning rate, minimum loss reduction, maximum tree depth, subsample, and number of trees, which we deter-

**Table 1:** Baseline characteristics between cases and controls

Variables	Cases (n=1,062)	Controls (n=1,809)	p-value
Days to HF in ICU; mean (SD)	4 (± 4.5)		
Age (yrs.), median (min– max)	77 (22 - 88)	73 (20 – 88)	< 0.000*
Gender (Male) n (%)	650 (61)	1120 (62)	0.944
<b>Laboratory measures</b>			
Creatinine (mg/dL)	1.17 (0.8 – 1.7)	0.94 (0.7 – 1.2)	0.000*
MCH (pg)	30.3 (28.9)	30.5 (29.4 – 31.4)	0.028*
Platelet count (K/uL)	192.8 (143.8 – 248.4)	180.0 (141.3 – 232.6)	0.104
Hematocrit (%)	29.7 (27.7 – 31.1)	29.6 (27.6 – 32.0)	0.437
INR	1.35 (1.2 – 1.6)	1.32 (1.2 – 1.5)	0.002*
Potassium (mEq/L)	4.19 (3.9 – 4.4)	4.18 (3.9 – 4.3)	0.004*
Calcium (mg/dL)	8.47 (8.2 – 8.8)	8.45 (8.1 – 8.7)	0.219
Magnesium (mEq/L)	2.11 (1.9 – 2.2)	2.00 (1.9 – 2.2)	0.004*
Phosphate (mg/dL)	3.56 (3.2 – 4.0)	3.44 (2.9 – 3.7)	0.000*
<b>Comorbidities</b>			
Diabetes (yes)	367 (35.0)	443 (24.4)	< 0.000*
Obesity (yes)	94 (8.9)	140 (7.7)	0.128
Atrial fibrillation (yes)	593 (55.8)	828 (45.8)	0.012*
Hypertension ((yes)	785 (73.9)	1339 (74.0)	0.029*
Hyperlipidemia (yes)	492 (46.3)	928 (51.3)	0.004*
CAD (yes)	364 (34.3)	446 (24.7)	< 0.000*
COPD (yes)	259 (24.4)	267 (14.7)	< 0.000*

CAD: Coronary Artery Diseases; INR: International Normalized Ratio; MCH: Mean Corpuscular Hemoglobin; mg/dl milligrams per deciliter; pg, picograms; IU/L International units per litre; cm, centimeter (cm); yrs. Years; HF, heart failure; S.D Standard Deviation, K/uL thousand per microliter, m/uL million per microliter, % percentage, mEq/L milliequivalents per liter. Continuous values that are normally distributed were recorded as mean (S.D) and others input as median (IQR), and categorical values (absolute numbers and percentages). The Chi-square test was used for the comparison of categorical variables and the two-sample t-test for continuous variables. All p values were two-sided. Statistical significance was defined as p < 0.05.

**Table 2:** Assessment of clinical outcomes

Variables	Cases (n = 1,062)	Controls (n = 1,809)	p-value
Length of Stay (days) median (IQR)	6 (2 - 8)	4 (1 - 6)	0.134
Mortality n (%)	119 (11.2)	98 (5.4)	0.098

mined through a grid search. The hyperparameters used in our analysis were as follows (determine by grid search): learning rate=0.1, maximum tree depth=7, subsample=0.6, and number of trees=200. With these settings, we created a strong machine learning model to predict heart failure in patients with aortic disorders. To understand which factors are most important in our model's predictions, we calculated "feature importance." This tells us how much each factor contributes to making accurate predictions. In our model, the top 10 important factors are creatinine, phosphate, age, COPD, INR, diabetes, CAD, magnesium, atrial fibrillation, and hyperlipidemia. **With this information, we can better understand and predict heart failure in patients with aortic disorders, which can help doctors provide better care and treatment.**

### The Logistic Regression Model

The results of logistic regression model are shown in (Table 3). Age (OR 1.67) An OR of 1.67 for age means that for every one-year increase in age, the odds of the outcome (e.g., heart failure in this case) increase by 67% when all other variables are kept constant, **An OR of 1.06 for gender indicates that being male is associated with a 6% increase in the odds of the outcome compared to being female, while other factors are un-**

**Table 3:** Multivariable logistic regression model

Variables	OR (95% CI)	p-value
Age	1.67 (1.360,2.040)	< 0.001*
Gender	1.06 (0.844,1.198)	0.945
Creatinine (mg/dL)	1.19 (1.090,1.300)	< 0.001*
MCH (pg)	0.96 (0.920,1.000)	0.028*
Platelet count (K/uL)	1.00 (0.999,1.001)	0.105
Hematocrit (%)	1.08 (0.987,1.029)	0.438
Diabetes	1.55 (1.290,1.850)	< 0.001*
Obesity	1.26 (0.940,1.700)	0.130
Atrial fibrillation	1.24 (1.050,1.470)	0.012*
Hypertension	0.81 (0.660,0.980)	0.030*
Hyperlipidemia	0.78 (0.660,0.930)	0.005*
CAD	1.49 (1.250,1.780)	< 0.001*
COPD	1.67 (1.360,2.040)	< 0.001*
Potassium (mEq/L)	0.68 (0.530,0.890)	0.005*
Calcium (mg/dL)	1.10 (0.940,1.290)	0.223
Magnesium (mEq/L)	1.59 (1.150,2.200)	0.004*
Phosphate (mg/dL)	1.25 (1.110,1.410)	< 0.001*
INR	1.29 (1.100,1.530)	0.002*

CAD: Coronary Artery Diseases; INR: International Normalized Ratio; MCH: Mean Corpuscular Hemoglobin; OR: Odd Ratio; CI: Confidence Interval; mg/dl: Milligrams per Deciliter; pg: Picograms; IU/L: International Units per Litre; cm: Centimeter (cm); yrs. Years; HF: Heart Failure; S.D: Standard Deviation, K/uL thousand per microliter, m/uL million per microliter, % percentage, mEq/L milliequivalents per liter. An OR value greater than 1 indicates that the presence of a variable or increase in a continuous variable is associated with higher probability of case occurrence.

**Table 4:** Model performance in the testing dataset.

Model	Precision	F1 score	AUROC	Sensitivity	Specificity
XGBoost	0.69	0.70	0.77	0.70	0.68
Logistic Regression	0.73	0.72	0.78	0.80	0.64

changed. An OR of 1.19 for creatinine suggests that for every one-unit increase in creatinine levels, the odds of the outcome (heart failure) increase by 19% when all other variables remain the same. Similarly, for hematocrit, diabetes, obesity, atrial fibrillation, Coronary Artery Disease (CAD) and COPD, the ORs of 1.08, 1.55, 1.26, 1.24, 1.49 and 1.67, respectively, indicate the percentage increase in the odds of the outcome associated with one-unit increases in each corresponding predictor variable. Likewise, calcium and magnesium with OR 1.10 and 1.59 respectively. However, MCH (OR 0.96), hypertension (OR 0.81), hyperlipidemia (OR 0.78), and potassium (OR 0.68) were associated with a decrease in the likelihood of the outcome (heart failure).

### Model Performance

Model discrimination was assessed using the area under receiver operating characteristic curve (AU-ROC). The logistic regression has little greater AU-ROC than the XGBoost model (AU-ROC, 0.781; 95% CI, 0.751 to 0.811 vs. 0.775; 95% CI, 0.744 to 0.805, respectively; Figure 3. Table 4 describes the classification evaluation metrics for the two models. **We used model evaluation metrics to see how well our models performed. We looked at the XGBoost and LR models. The logistic regression has little higher discrimination capability of 78% AU-ROC and precision score of 0.73, f1 score 0.72, sensitivity 0.80 and specificity of 0.64. On the other hand, the XGBoost model had an AU-ROC of 77% on the testing set, with a sensitivity of 70% and specificity of 68%, an F1 score of 70% and precision score of 69%. Overall, logistic regression performed much better than XGBoost model in our evaluation.**

## Discussion

In this research, we found that certain clinical factors are more linked to heart failure in patients with aortic disorders in the Intensive Care Unit (ICU). By using advanced machine learning methods, we were able to identify important factors associated with heart failure, such as creatinine, phosphate, age, COPD, INR, diabetes, CAD, magnesium, atrial fibrillation, and hyperlipidemia.

Our research revealed that diabetes, obesity, Coronary Artery Disease (CAD), and Chronic Obstructive Pulmonary Disease (COPD) were also associated with a higher risk of heart failure in these patients. Recent efforts to improve heart failure outcomes have focused not only on the main disease but also on related health issues. Chronic Obstructive Pulmonary Disease (COPD) is becoming more common [27]. Up to one-third of patients with stable heart failure also have COPD, mainly due to the shared risk factor of smoking and cumulative smoking exposure [28]. It's important to note that obesity-related factors are estimated to be responsible for about 11% of heart failure cases in men and 14% in women. Heart failure is often caused by a condition called Coronary Artery Disease (CAD), which is the most common reason for this condition. CAD happens when fatty deposits build up in the arteries, making them narrower and reducing blood flow. This can eventually lead to a heart attack. About two-thirds of heart failure cases are linked to CAD [29]. In people with heart failure, having CAD has been shown in many studies to be independently connected to a poorer long-term outlook [30]. Obesity can impact the heart's function by altering blood flow and affecting the heart muscle, which may contribute to the development of heart failure [31]. Diabetes mellitus is commonly found in patients with heart failure, particularly in those with Heart Failure and Preserved Ejection Fraction (HFpEF). Many studies have shown that diabetes mellitus is closely linked to the development of heart failure, and the risk is more than doubled in men and more than quintupled in women [32,33].

Electrolyte abnormalities are common in heart failure [34]. In our study using a multiple regression model, we discovered a connection between phosphate, magnesium and calcium levels and an increased risk of heart failure in patients with aortic disorder in the ICU. In individuals who have aortic disorders, like aortic stenosis, elevated levels of serum phosphate have been linked to a higher risk of Cardiovascular Disease (CVD) mortality, as indicated by previous studies. This association has also been observed in patients with a history of Myocardial Infarction (MI) [35]. Magnesium plays a crucial role as a co-factor in various enzymatic reactions that contribute to stable cardiovascular function and heart rhythm. Its deficiency is common and can be linked to risk factors and complications of heart failure [36]. Low phosphate levels, known as hypophosphatemia, can impact multiple organ systems, including the cardiovascular system [37]. A depletion of phosphate may lead to ventricular arrhythmias and reduced ATP synthesis, causing temporary heart dysfunction. Studies have also shown that high-normal serum phosphate levels can be associated with vascular and valvular calcification [34].

In our study, we found that patients with heart failure experienced a longer stay in the ICU, with a median of 6 days, and a higher number of mortality cases compared to those without heart failure. Regardless of the presence of other health conditions and risk factors for cardiovascular diseases, a prolonged stay in the ICU during heart failure hospitalization is associated

with unfavorable clinical outcomes. This includes an increased risk of life-threatening medical complications, higher readmission rates, and elevated mortality [38,39]. Numerous studies have reported higher lengths of stay in heart failure hospitalization in the ICU, with estimated median ranges varying from 7 to 21 days [40,41]. For example, the Sub-Saharan Africa Survey of Heart Failure [42], which included Ethiopian patients, reported a median length of stay of 7 days.

This study presents both strengths and limitations. The utilization of the XGBoost modeling technique is a unique approach in critical care research, offering promising applications. Previous successful implementations of XGBoost in complex scenarios, such as predicting treatment failure for parapneumonic empyema, demonstrated superior predictive accuracy compared to logistic regression model [43]. The XGBoost model's ability to capture intricate data relationships without the need for explicit specification of high-order interactions and non-linear functions is advantageous [44]. Moreover, the model's built-in cross-validation and regularization mechanisms effectively combat overfitting concerns [45]. In addition, limiting the analysis to a 30-day observation window allows for a more focused investigation of the short-term risk of aortic disorders in heart failure patients. These findings highlight the potential of XGBoost to enhance critical care epidemiological studies in the future. The limitation includes, the size of the study cohort was just 2,871, and this could impact the statistical power and generalizability of the results. Incomplete or missing data in electronic health records may affect the accuracy and completeness of the analysis. And finally, as the study focuses on ICU patients, the generalizability of the findings to non-ICU settings or other patient populations may be limited.

### Potential Impacts on Clinical Utility

The study results provide valuable insights for clinicians, enabling informed discussions with patients and families about heart failure risk in the context of aortic valve and aortic vascular disorders. This knowledge supports shared decision-making, informed consent, and similarly supports an emphasis on the importance of adhering to treatment plans and lifestyle changes. Clinicians will be better able to tailor treatment approaches based on factors associated with increased heart failure risk, optimizing medications, and considering timely surgical interventions.

Understanding length of hospital stay and 30-day mortality rates will assist healthcare administrators to allocate resources for healthcare administrators, optimizing patient care, staffing, and facility requirements. Additionally, the study's outcomes may drive further research in heart failure risk assessment and management for patients with aortic valve and aortic vascular disorders, potentially leading to novel advancements in biomarkers, imaging techniques, and therapeutic interventions.

Patient advocacy groups can utilize the study findings to raise awareness about heart failure risk in individuals with these disorders, supporting the development of patient education materials and support services for better heart health management. Overall, the study's clinical usefulness has the potential to positively impact patient care, leading to improved outcomes and reduced heart failure-related morbidity and mortality.

### Conclusion

Machine learning models show promising predictive accuracy in identifying patients at higher risk of heart failure within

30-days of ICU admission. These models leverage various clinical and demographic variables, enabling early detection and intervention for those at higher risk. Overall, the use of machine learning in this study represents a significant advancement in cardiovascular research and patient care, with the potential to enhance risk assessment and improve clinical outcomes for individuals with aortic valve and aortic vascular disorders. However, further validation and implementation in clinical practice are necessary to fully realize the clinical benefits of this approach.

### Author Statements

#### Author Contributions

K Bah, Ns Bah AW Jallow and Dr. M Touray conceived the study. K Bah responsible for the methodology; K Bah, Ns Bah, and AW Jallow managed the software; K Bah, Ns Bah, AW Jallow and Dr. M Touray were responsible for validation; K Bah, Ns Bah, **AW Jallow and Dr. M Touray conducted the formal analysis**; K Bah, Ns Bah and AW Jallow conducted the investigation; K Bah was responsible for data curation; K Bah wrote the original draft; K Bah, Ns Bah, AW Jallow and Dr. M Touray reviewed and edited the draft. All authors have read and agreed to the published version of the manuscript.

#### Funding

This research received no external funding.

#### Informed Consent Statement

Patient consent was waived due to the data used is anonymous from MIMIC III database.

#### Data Availability Statement

The data that was generated and/or analyzed during the current study are not publicly available due to the MIMIC III data policy and rules but are available from the corresponding author upon reasonable request.

#### Conflicts of Interest

The authors declare no conflict of interest.

### References

- GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018; 392: 1789-858.
- Heidenreich PA, Albert NM, Allen LA, Bluemke DA, Butler J, Fonarow GC, et al. Forecasting the impact of heart failure in the United States: a policy statement from the American Heart Association. *Circ Heart Fail*. 2013; 6: 606-19.
- Bragazzi NL, Zhong W, Shu J, Abu Much A, Lotan D, Grupper A, et al. Burden of heart failure and underlying causes in 195 countries and territories from 1990 to 2017. *Eur J Prev Cardiol*. 2021; 28: 1682-90.
- Liang Z, Zhang G, Huang JX, Hu QV. Deep learning for healthcare decision making with EMRs. In: *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. Vol. 2014. IEEE Publications; 2014.
- Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
- Rao SR, Desroches CM, Donelan K, Campbell EG, Miralles PD, Jha AK. Electronic health records in small physician practices: availability, use, and perceived benefits. *J Am Med Inform Assoc*. 2011; 18: 271-5.
- Ge Y, Wang Q, Wang L, Wu H, Peng C, Wang J, et al. Predicting post-stroke pneumonia using deep neural network approaches. *Int J Med Inform*. 2019; 132: 103986.
- Nguyen BP, Pham HN, Tran H, Nghiem N, Nguyen QH, Do TTT, et al. Predicting the onset of type 2 diabetes using wide and deep learning with electronic health records. *Comput Methods Programs Biomed*. 2019; 182: 105055.
- Faturrahman M, Wasito I, Hanifah N, Mufidah R. Structural MRI classification for Alzheimer's disease detection using deep belief network. In: *11<sup>th</sup> International Conference on Information & Communication Technology and System (ICTS)*. Vol. 2017. IEEE Publications; 2017.
- Liu J, Zhang Z, Razavian N. Deep ehr: chronic disease prediction using medical notes. In: *Machine Learning for Healthcare Conference*. 2018.
- Ahmad FS, Ali L, Raza-Ul-Mustafa, Khattak HA, Hameed T, Wahajat I, et al. A hybrid machine learning framework to predict mortality in paralytic ileus patients using electronic health records (EHRs). *J Ambient Intell Hum Comput*. 2021; 12: 3283-93.
- Johnson A, et al. MIMIC-III, a freely accessible critical care database. *Sci. Data*. 2016; 3: 1.
- Zhang Z. Multiple imputation with multivariate imputation by chained equation (MICE) package. *Ann Transl Med*. 2016; 4: 30.
- Zhang Z, Gayle AA, Wang J, Zhang H, Cardinal-Fernández P. Comparing baseline characteristics between groups: an introduction to the CBCgrps package. *Ann Transl Med*. 2017; 5: 484.
- Lv CX, An SY, Qiao BJ, Wu W. Time series analysis of hemorrhagic fever with renal syndrome in mainland China by using an XGBoost forecasting model. *BMC Infect Dis*. 2021; 21: 839.
- Fang ZG, Yang SQ, Lv CX, An SY, Wu W. Application of a data-driven XGBoost model for the prediction of COVID-19 in the USA: a time-series study. *BMJ Open*. 2022; 12: e056685.
- Li W, Yin Y, Quan X, Zhang H. Gene expression value prediction based on XGBoost algorithm. *Front Genet*. 2019; 10: 1077.
- Junling L, Zhang Z, Fu Y, Rao F. Time series prediction of COVID-19 transmission in America using LSTM and XGBoost algorithms. *Results Phys*. 2021; 27: 104462.
- Paliari I, Karanikola A, Kotsiantis S. A comparison of the optimized LSTM, XGBOOST and Arima in Time Series forecasting. In: *12th International Conference on Information, Intelligence, Systems & Applications (IISA)*. Vol. 2021. IEEE Publications; 2021.
- Noureldin NA, Aboelghar MA, Saady HS, Ali AM. Rice yield forecasting models using satellite imagery in Egypt. *Egypt J Remote Sens Space Sci*. 2013; 16: 125-31.
- Rahman MS, Chowdhury AH, Amrin M. Accuracy comparison of Arima and XGBoost forecasting models in predicting the incidence of COVID-19 in Bangladesh. *PLOS Glob Public Health*. 2022; 2: e0000495.
- Schuld M, Sinayskiy I, Petruccione F. An introduction to quantum machine learning. *Contemp Phys*. 2015; 56: 172-85.
- Char DS, Abràmoff MD, Feudtner C. Identifying ethical considerations for machine learning healthcare applications. *Am J Bioeth*. 2020; 20: 7-17.

24. Datta S, Barua R, Das J. Application of artificial intelligence in modern healthcare system. Alginates-recent uses of this natural polymer. 2019.
25. Elsebakhi E, Lee F, Schendel E, Haque A, Kathireason N, Pathare T, et al. Large-scale machine learning based on functional networks for biomedical big data with high performance computing platforms. *J Comp Sci.* 2015; 11: 69-81.
26. Pattnayak P, Panda AR. Innovation on machine learning in healthcare services—an introduction. In: Technical advancements of machine learning in healthcare. 2021: 1-30.
27. Ramadan WH, Sarkis AT. Patterns of use of dry powder inhalers versus pressurized metered-dose inhalers devices in adult patients with chronic obstructive pulmonary disease or asthma: an observational comparative study. *Chronic Respir Dis.* 2017; 14: 309-20.
28. Hawkins NM, Petrie MC, Jhund PS, Chalmers GW, Dunn FG, McMurray JJ. Heart failure and chronic obstructive pulmonary disease: diagnostic pitfalls and epidemiology. *Eur J Heart Fail.* 2009; 11: 130-9.
29. Suleiman M, Khatib R, Agmon Y, Mahamid R, Boulos M, Kapeiovich M, et al. Early inflammation and risk of long-term development of heart failure and mortality in survivors of acute myocardial infarction: predictive role of C-reactive protein. *J Am Coll Cardiol.* 2006; 47: 962-8.
30. Smith Jr SC, Blair SN, Bonow RO, Brass LM, Cerqueira MD, Dracup K, et al. AHA/ACC Scientific Statement: AHA/ACC guidelines for preventing heart attack and death in patients with atherosclerotic cardiovascular disease: 2001 update: a statement for healthcare professionals from the American Heart Association and the American College of Cardiology. *Circulation.* 2001; 104: 1577-9.
31. Bahrami H, Kronmal R, Bluemke DA, Olson J, Shea S, Liu K, et al. Differences in the incidence of congestive heart failure by ethnicity: the multi-ethnic study of atherosclerosis. *Arch Intern Med.* 2008; 168: 2138-45.
32. Stratton IM, Adler AI, Neil HA, Matthews DR, Manley SE, Cull CA, et al. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. *BMJ.* 2000; 321: 405-12.
33. Metra M, Zacà V, Parati G, Agostoni P, Bonadies M, Ciccone M, et al. Cardiovascular and noncardiovascular comorbidities in patients with chronic heart failure. *J Cardiovasc Med (Hagerstown).* 2011; 12: 76-84.
34. Urso C, Brucculeri S, Caimi G. Acid–base and electrolyte abnormalities in heart failure: pathophysiology and implications. *Heart Fail Rev.* 2015; 20: 493-503.
35. Foley RN. Phosphorus comes of age as a cardiovascular risk factor. *Arch Intern Med.* 2007; 167: 873-4.
36. Leier CV, Dei Cas L, Metra M. Clinical relevance and management of the major electrolyte abnormalities in congestive heart failure: hyponatremia, hypokalemia, and hypomagnesemia. *Am Heart J.* 1994; 128: 564-74.
37. Gaasbeek A, Meinders AE. Hypophosphatemia: an update on its etiology and treatment. *Am J Med.* 2005; 118: 1094-101.
38. Philbin EF, Roerden JB. Longer hospital length of stay is not related to better clinical outcomes in congestive heart failure. *Am J Manag Care.* 1997; 3: 1285-91.
39. Omar HR, Guglin M. Longer-than-average length of stay in acute heart failure: determinants and outcomes. *Herz.* 2018; 43: 131-9.
40. Moriyama H, Kohno T, Kohsaka S, Shiraishi Y, Fukuoka R, Nagatomo Y, et al. Length of hospital stay and its impact on subsequent early readmission in patients with acute heart failure: a report from the WET-HF Registry. *Heart Vessels.* 2019; 34: 1777-88.
41. Salam A, Sulaiman K, Suwaidi JA, Asaad N, AlHabib K, Almahmeed W, et al. Preventable precipitating factors of hospitalizations with heart failure and prolonged hospital length of stay: observations from the gulf care study. *J Am Coll Cardiol.* 2015; 65: A1032.
42. Damasceno A, Mayosi BM, Sani M, Ogah OS, Mondo C, Ojji D, et al. The causes, treatment, and outcome of acute heart failure in 1006 Africans from 9 countries. *Arch Intern Med.* 2012; 172: 1386-94.
43. Khemasuwan D, Sorensen J, Griffin DC. Predictive variables for failure in administration of intrapleural tissue plasminogen activator/deoxyribonuclease in patients with complicated parapneumonic effusions/empyema. *Chest.* 2018; 154: 550-6.
44. Friedman JH. Stochastic gradient boosting. *Comp Stat Data Anal.* 2002; 38: 367-78.
45. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining; 2016: 785-94.