WEP Warwick Evans Publishing

International Journal of Computer Science and Information Technology

ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 5, Number 2, Year 2025
DOI: https://doi.org/10.62051/ijcsit.v5n2.14
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Sex-specific Predictors of Mortality in Patients with Heart Failure in Intensive Care Units Using Machine Learning

Xinhui Tian *

Key Laboratory of Electronic and Information Engineering, State Ethnic Affairs Commission (Southwest Minzu University), Chengdu, China

* Corresponding Author Email: tianxxxxh@163.com

ABSTRACT

This retrospective cohort study analyzed congestive heart failure (HF) patients from MIMIC-IV and eICU-CRD. Statistical tests described patient characteristics, while machine learning models (CART, RF, XGBoost, LightGBM) ranked sex-specific predictors, with SHAP used for interpretation. Among 30,411 patients, 3,840 died (men: 12.7%, mean age 69.50; women: 12.6%, mean age 73.02). Systolic blood pressure (SBP) was more critical for mortality in women, while body temperature, blood urea nitrogen (BUN), and creatinine was more significant in men. This is the first machine learning study to explore sex-specific in-hospital mortality predictors in ICU HF patients, identifying both known and potentially important risk factors.

KEYWORDS

Intensive Care Units; Machine Learning Models; Risk Factors; Explainability Analysis.

1. INTRODUCTION

Heart Failure (HF) is a global cardiovascular disease that imposes a heavy burden on public health. Approximately 6 million people in the United States have HF [1], and the prevalence of HF is expected to increase as the population ages and the prevalence of risk factors such as hypertension, diabetes, and obesity increases. Despite advances in treatment, the 5-year mortality rate is still as high as 30%-50% [2]. While the overall lifetime risk of HF is comparable between men and women [3], there exist significant and under-recognized sex disparities [4–6]. Studies have found that although female patients with HF have higher hospitalization rates, their survival rates tend to be higher than those of males [7,8]. However, investigations into sex-specific predictors of mortality among HF patients have been relatively scarce. Further understanding of sex-specific variables could refine prognostic models tailored to each sex, thereby advancing both predictive accuracy and insights into the underlying pathophysiological mechanisms driving these sex-related differences. Previous studies on sex differences in HF-related predictors [9,10] have not examined large, highly correlated predictor sets, nor have they elucidated direct interactions among predictors, and they lacked a priori specification. To comprehensively understand mortality complexities in HF, further investigations using adaptable modeling approaches are warranted. Machine learning methods have already been applied to predict mortality risk among HF patients, across both small and large cohorts. However, to date, no studies have employed machine learning to explore sex-specific mortality and predictors among intensive care unit (ICU) patients with HF. Hence, this study aimed to leverage machine learning for a comprehensive analysis of mortality and the identification of predictors among both female and male ICU patients with HF. By investigating sex-related differences in mortality and

predictors, this study aimed to provide valuable insights into the management and treatment of HF in both genders within the ICU setting.

2. METHODS

2.1. Data Source

The data used in the study came from the Dryad data package (Han, Didi et al. (2022), Early prediction of in-hospital mortality in patients with congestive heart failure in intensive care unit: a retrospective observational cohort study, Dryad, Dataset, https://doi.org/10.5061/dryad.tx95x6b18) [11]. We performed a secondary analysis of the data downloaded from the Dryad digital repository. Dryad is a non-profit repository of research data in the fields of medicine, biology, and ecology to facilitate the flow of scientific data and provide researchers with high-quality data resources. Since this was a post hoc study using data from an existing study, informed consent was waived.

2.2. Study Population

The original data for HF patients came from the Medical Information Mart for Intensive Care IV (MIMIC-IV) and the eICU Collaborative Research Database (eICU-CRD). MIMIC-IV is a freely accessible contemporary electronic health record database containing data on more than 40,000 intensive care unit patients from 2008 to 2019 [12]. Another intensive care unit database, the eICU Cooperative Research Database (eICU-CRD) [13], from which we extract the same features as MIMIC-IV. Conducting the multi-center study, we combined and processed data extracted from both databases. For patients with multiple ICU admissions, demographic characteristics, and vital signs in the first 24 hours after admission were recorded, and the results of the first measurement on admission were selected as laboratory examination indicators. Exclusion criteria were age <18 years and missing outcome data. After applying these criteria, the study included 16,295 and 14,116 CHF patients in the male and female cohorts, respectively.

2.3. Statistical Analysis

The primary outcome was in-hospital mortality, determined by survival status at discharge. Heart failure patients were categorized by sex. Continuous variables (such as age and heart rate) for male and female patients were presented as mean ± standard deviation and compared using the Kruskal-Wallis test. Categorical variables were presented as total and percentage, with differences between groups compared using the $\chi 2$ test. All statistical analysis and calculation using EmpowerStats (http://www.empowerstats.com, X&Y Solutions, Inc., Boston, MA) and R 4.0.5 (HTTP: //www.R-project.org, The R Foundation). A double-tailed P value <0.05 indicates a statistical difference.

2.4. Machine Learning and Interpretability Methods

This study uses open-source libraries in Python (version 3.9.0) to implement machine-learning methods. The data were extracted from MIMIC-IV and eICU-CRD and divided into training sets and validation sets according to the ratio of 9:1 for model construction and feature analysis.

Firstly, data structures for male and female patients were initially evaluated using a classification tree model. Classification tree modeling is a non-parametric method that builds decision trees based on predictor variables and their combinations to distinguish cases from non-cases with the highest probability [14].

Secondly, we used Classification and Regression Tree (CART), Random Forest (RF), eXtreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM) in the Scikit-learn package [15] to build multiple machine learning models. The model is trained on the training dataset

and validated on the validation dataset, and the optimal model is determined by adjusting the parameters. The model performance was evaluated using receiver operating characteristic curve (ROC curve) analysis over 1000 replicates, and the area under the curve (AUC) and its 95% confidence interval (CI) were calculated to evaluate the prediction accuracy.

Finally, the importance scores of features are obtained by training CART, RF, XGBoost, LightGBM, and other machine learning models. In addition, the SHapley Additive exPlanation (SHAP) method [16] was used to quantify the contribution of features to the model prediction results, evaluate the importance of features in patients with different genders, and analyze the impact of each feature on the model output, and visualize the Shapley value of a single feature.

3. RESULTS

3.1. Study Population

The stratification of all samples by sex is shown in Table 1. Among the 30,411 HF patients, 14,116 (46.4%) were female and 16,295 (53.6%) were male. Among these patients, 1775 (12.6%) female patients died during hospitalization, with a mean age of 73.02 years, and the mean age of the non-fatal group was 76.46 years. Among male patients, 2065 (12.7%) died during hospitalization, with a mean age of 69.50 years, and the mean age of the non-fatal group was 74.43 years. A higher percentage of female patients were black compared to men. The Sequential Organ Failure Assessment (SOFA) and Acute Physiological ScoreIII (APSIII) at the time of death during hospitalization were lower in female patients than in men.

Analysis of comorbidities showed that women who died during hospitalization were more likely to develop chronic lung obstruction and less likely to develop acute myocardial infarction, diabetes, and liver failure. During treatment, men are more likely to use drugs such as norepinephrine, dopamine, epinephrine, phenylephrine, and vasopressin. In addition, among patients who died while hospitalized, women had higher systolic blood pressure (SBP), platelet levels, respiratory rate, heart rate, blood sodium, and red blood cell distribution width (RDW). Men had higher blood urea nitrogen (BUN) levels, anion gap (AG), creatinine, hemoglobin, mean corpuscular hemoglobin concentration (MCHC), and mean corpuscular volume (MCV).

Table 1. Characteristics of the Death Cases and the Comparison Subcohort

	Men		Women	
Variable	Death (n=2065)	Comparison Subcohort (n=14230)	Death (n=1775)	Comparison Subcohort (n=12341)
Age, years	69.50 ± 13.81	74.43 ± 12.28	73.02 ± 13.55	76.46 ± 12.21
Ethnicity(%)				
White	10396 (73.06%)	1562 (75.64%)	8658 (70.16%)	1246 (70.20%)
Black	2094 (14.72%)	209 (10.12%)	2313 (18.74%)	293 (16.51%)
Other	1740 (12.23%)	294 (14.24%)	1370 (11.10%)	236 (13.30%)
COPD(%)				
No	10242 (71.97%)	1494 (72.35%)	8161 (66.13%)	1193 (67.21%)
Yes	3988 (28.03%)	571 (27.65%)	4180 (33.87%)	582 (32.79%)
AMI(%)				
No	11451 (80.47%)	1616 (78.26%)	10561 (85.58%)	1476 (83.15%)
Yes	2779 (19.53%)	449 (21.74%)	1780 (14.42%)	299 (16.85%)
Diabetes(%)				

	Men		Women	
Variable	Death (n=2065)	Comparison Subcohort (n=14230)	Death (n=1775)	Comparison Subcohort (n=12341)
No	8725 (61.31%)	1303 (63.10%)	7627 (61.80%)	1161 (65.41%)
Yes	5505 (38.69%)	762 (36.90%)	4714 (38.20%)	614 (34.59%)
HepF(%)				
No	13952 (98.05%)	1976 (95.69%)	12172 (98.63%)	1731 (97.52%)
Yes	278 (1.95%)	89 (4.31%)	169 (1.37%)	44 (2.48%)
Norepinephrine(%)				
No	12318 (86.56%)	1273 (61.65%)	11075 (89.74%)	1153 (64.96%)
Yes	1912 (13.44%)	792 (38.35%)	1266 (10.26%)	622 (35.04%)
Dopamine(%)				
No	14062 (98.82%)	1906 (92.30%)	12220 (99.02%)	1671 (94.14%)
Yes	168 (1.18%)	159 (7.70%)	121 (0.98%)	104 (5.86%)
Epinephrine(%)				
No	13868 (97.46%)	1885 (91.28%)	12170 (98.61%)	1666 (93.86%)
Yes	362 (2.54%)	180 (8.72%)	171 (1.39%)	109 (6.14%)
Phenylephrine(%)				
No	13687 (96.18%)	1753 (84.89%)	11982 (97.09%)	1506 (84.85%)
Yes	543 (3.82%)	312 (15.11%)	359 (2.91%)	269 (15.15%)
Vent(%)				
No	4149 (29.16%)	335 (16.22%)	3505 (28.40%)	318 (17.92%)
Yes	10081 (70.84%)	1730 (83.78%)	8836 (71.60%)	1457 (82.08%)
Intubated(%)				
No	11142 (78.30%)	1270 (61.50%)	10072 (81.61%)	1141 (64.28%)
Yes	3088 (21.70%)	795 (38.50%)	2269 (18.39%)	634 (35.72%)
Vasopressin(%)				
No	13851 (97.34%)	1693 (81.99%)	12150 (98.45%)	1516 (85.41%)
Yes	379 (2.66%)	372 (18.01%)	191 (1.55%)	259 (14.59%)
MC(%)				
No	13381 (94.03%)	1877 (90.90%)	11764 (95.32%)	1623 (91.44%)
Yes	849 (5.97%)	188 (9.10%)	577 (4.68%)	152 (8.56%)
SOFA	4.24 ± 3.00	7.19 ± 3.90	3.96 ± 2.95	6.47 ± 3.65
APSIII	48.51 ± 22.76	73.61 ± 29.94	47.82 ± 22.20	73.19 ± 30.57
Temperature,℃	36.59 ± 0.70	36.42 ± 1.09	36.59 ± 0.64	36.36 ± 1.12
Respiratory rate, breaths per minute	22.81 ± 10.51	25.47 ± 11.23	23.84 ± 11.02	25.73 ± 11.41
Heart rate, beats per minute	90.55 ± 24.65	97.11 ± 27.99	91.06 ± 24.95	97.50 ± 28.56
SBP, mm Hg	118.33 ± 22.28	108.98 ± 22.23	121.17 ± 22.86	111.68 ± 24.49
Chloride, mmol/L	101.83 ± 6.61	101.31 ± 7.84	101.59 ± 6.78	101.38 ± 7.76
AG	13.21 ± 4.77	15.31 ± 6.07	12.99 ± 4.66	15.03 ± 5.88
BUN, mg/dL	36.55 ± 24.56	49.66 ± 30.73	32.72 ± 22.56	42.26 ± 26.05
Creatinine, mg/dL	2.09 ± 2.04	2.39 ± 1.70	1.69 ± 1.63	1.95 ± 1.42
Potassium, mmol/L	4.30 ± 0.74	4.47 ± 0.87	4.22 ± 0.75	4.38 ± 0.89

Variable	Men		Women	
	Death (n=2065)	Comparison Subcohort (n=14230)	Death (n=1775)	Comparison Subcohort (n=12341)
Calcium, mg/dL	8.52 ± 0.73	8.39 ± 1.05	8.61 ± 0.77	8.43 ± 1.04
Sodium, mmol/L	137.84 ± 5.07	137.73 ± 6.47	138.03 ± 5.43	137.95 ± 6.42
Hemoglobin, g/dL	10.93 ± 2.44	10.52 ± 2.35	10.40 ± 2.08	10.25 ± 2.07
MCH, pg	29.61 ± 2.75	29.72 ± 2.90	29.16 ± 2.80	29.20 ± 2.82
MCHC, g/L	32.55 ± 1.57	32.10 ± 1.68	32.15 ± 1.65	31.75 ± 1.69
MCV, fL	90.98 ± 7.30	92.62 ± 7.91	90.72 ± 7.49	92.04 ± 7.97
RDW	15.82 ± 2.39	16.85 ± 2.74	15.96 ± 2.42	16.97 ± 2.76
WBC, $\times 10^9$/L	11.26 ± 8.02	14.01 ± 12.24	11.24 ± 7.06	14.64 ± 13.04
Platelet, $\times 10^9$/L	201.72 ± 94.59	194.73 ± 107.02	225.41 ± 102.96	217.77 ± 115.23
RBC, $\times 10^9$/L	3.71 ± 0.84	3.58 ± 0.82	3.59 ± 0.74	3.53 ± 0.75
MBP	84.66 ± 41.98	75.38 ± 43.72	84.78 ± 42.63	74.98 ± 44.00

AG, anion gap; AMI, acute myocardial infarction; APSIII, Acute Physiology Score III; BUN, blood urea nitrogen; CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; HepF, hepatic failure; MC, metastatic cancer; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; MCV, mean corpuscular volume; RBC, red blood cell; RDW, red blood cell distribution width; SBP, systolic blood pressure; SOFA, Sequential Organ Failure Assessment; WBC, white blood cell count.

3.2. Classification Tree of Predictors

The highest in-hospital mortality among men was in those patients with an APSIII score higher than 62.5, treated with vasopressin, and a temperature of less than 36.65℃ (n=148; risk, 0.78). Other combinations of importance are shown in Figure 1 (AUC, 0.7918; 95%CI, 0.7908-0.7927).

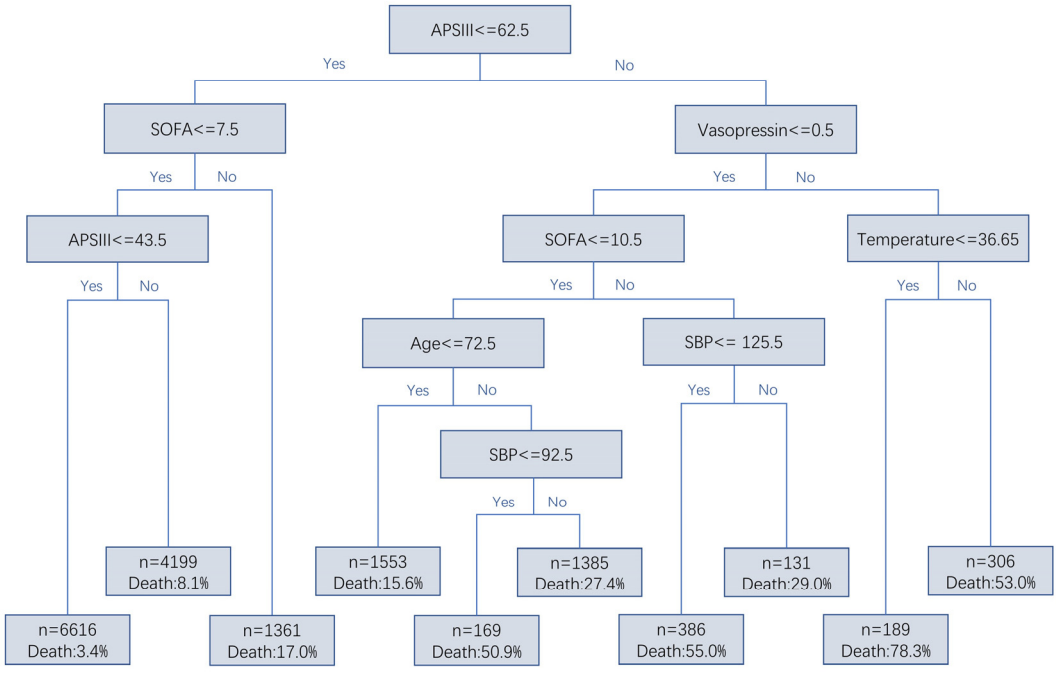

Figure 1. Classification tree of mortality predictors in male HF patients

The highest in-hospital mortality among women was in patients with an APSIII score higher than 73.5 and treated with vasopressin (n=189; risk, 0.69). The next highest was for patients who did not use antidiuretic hormone during treatment, with SOFA score higher than 5.5 and APSIII score higher

than 118.5. Other combinations of importance are shown in Figure 2 (AUC, 0.7757; 95%CI, 0.7747-0.7769).

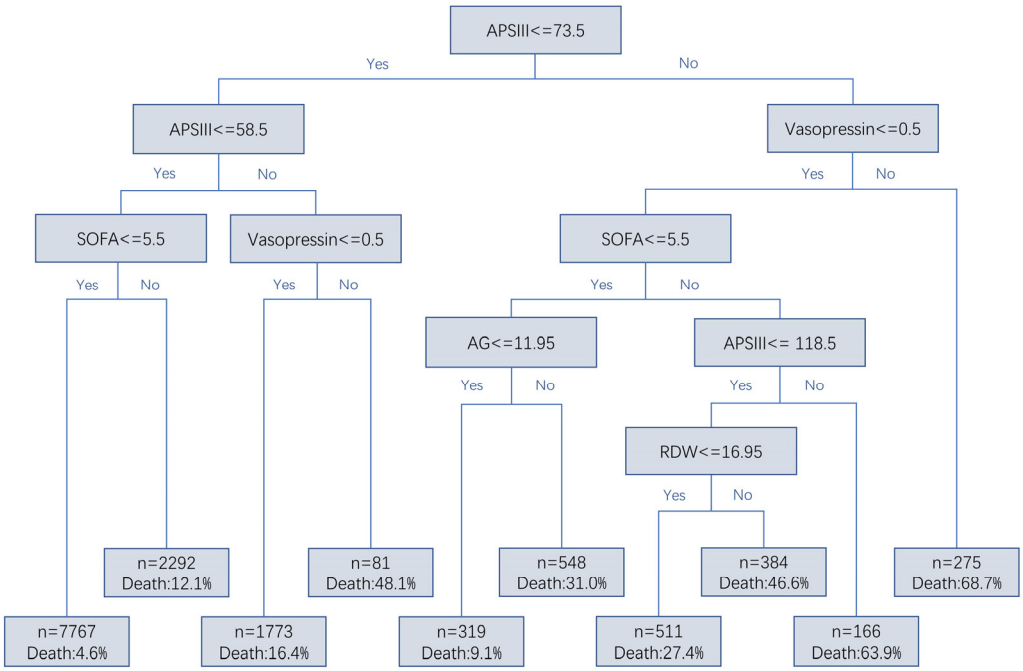

Figure 2. Classification tree of mortality predictors in female HF patients

3.3. Importance Ranking of Model Features

We used four classifiers (CART, RF, XGBoost, and LightGBM) to predict in-hospital mortality for male and female patients separately, ranking predictors by importance. Figure 3 presents the best-performing model (XGBoost, AUC: 0.8465; 95% CI: 0.8458–0.8471) and its predictor rankings. We compared the top ten features across models and calculated their median rankings (Table S1).

For women, APSIII score and SBP consistently ranked in the top ten across all models, with APSIII score ranked highest. Other features, including antidiuretic hormone, norepinephrine, phenylephrine, SOFA score, age, body temperature, WBC, and hemoglobin, appeared in the top ten in some models, while respiratory rate, mechanical ventilation, calcium, heart rate, and platelets appeared in only one model.

For men, APSIII score and SBP ranked in the top ten across all models, with APSIII score ranked highest. Other features, including antidiuretic hormone, norepinephrine, SOFA score, age, body temperature, respiratory rate, and BUN, appeared in the top ten in some models, whereas mechanical ventilation, phenylephrine, anion gap, heart rate, MCV, MCHC, and creatinine were top-ranked in only one model.

Overall, SBP was among the top ten predictors across all models and ranked higher in women. Phenylephrine was more important in the female model, while hemoglobin, intubation, and calcium were unique to female models. In contrast, age ranked higher for men. WBC was more important for men but appeared in only one model. Dopamine, BUN, MCV, MCHC, and creatinine were among the top ten only in men. However, APSIII score, SOFA score, antidiuretic hormone, norepinephrine, respiratory rate, mechanical ventilation, RDW, and platelets showed no significant sex-based differences.

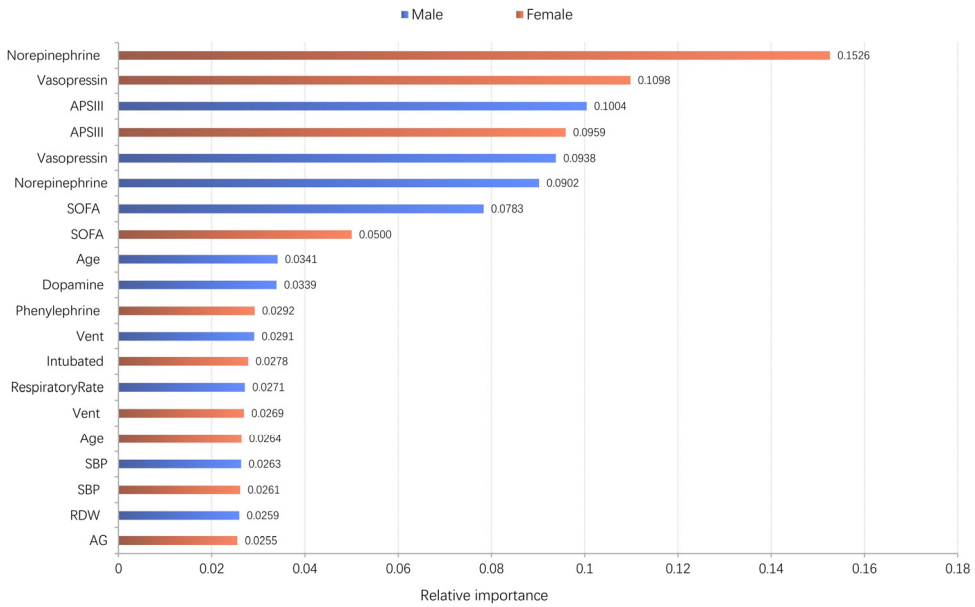

Figure 3. Importance of death factors predicted by XGBoost model in HF patients of different genders

3.4. ROC Curve and SHAP Interpretability Analysis

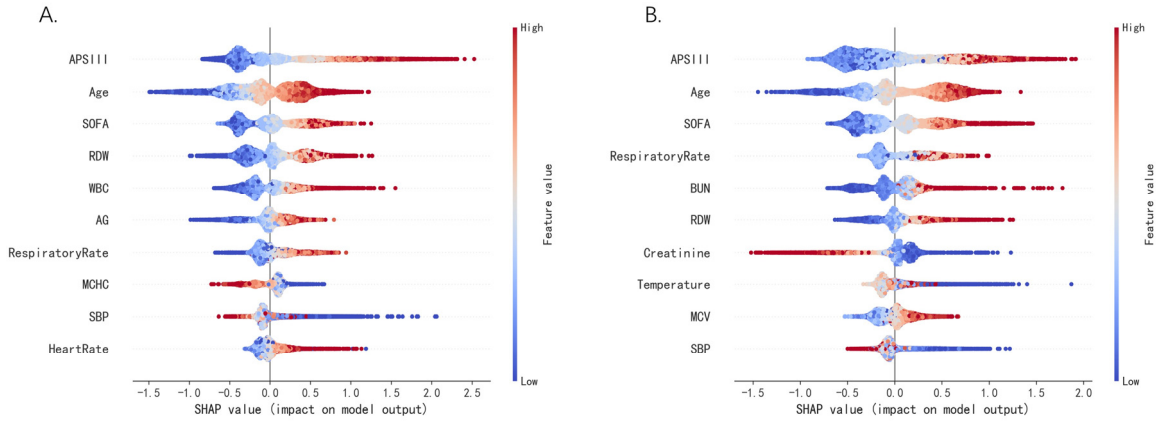

Figure 4. Top 10 most important risk factors affecting mortality in HF patients. The SHAP value (a value representing the influence of the feature on the model output) and the color corresponding feature value explain the feature contribution in the prediction. The midpoint of the colony plot represents the eigenvalue. A. Female patients. B. Male patients

To enhance the reliability of our results, we used SHAP, a game theory-based interpretive method, to rank sex differences in predictors. SHAP calculates the contribution of each feature to the predictions, helping clinicians understand individual predictions and risk factors.

We used 35 features to predict in-hospital mortality for male and female patients separately and evaluated the models' performance using AUC. The XGBoost model performed best (AUC: 0.8465; 95% CI: 0.8458–0.8471), as shown in Figure S2. SHAP TreeExplainer was then applied to interpret

the risk factors in the XGBoost model, and the top ten predictors for each gender are displayed in Figure 4.

SHAP analysis revealed that for women, APSIII score, age, SOFA score, RDW, and WBC were positively correlated with mortality, while SBP and MCHC were inversely associated. For men, BUN and MCV were positively correlated with mortality, while creatinine and body temperature were negatively correlated.

Table S2 compares the top ten predictors across models, showing higher contributions of SBP, phenylephrine, WBC, MCHC, heart rate, and hemoglobin in the female model, while BUN, MCV, creatinine, body temperature, and chloride had higher contributions in the male model. APSIII score, age, SOFA score, respiratory rate, and RDW showed minimal gender differences in importance.

4. DISCUSSION

This study focused on sex-specific predictors of mortality in patients with heart failure in the intensive care unit. We used multiple machine learning models to predict the risk of death during hospitalization in patients with HF and explored sex-specific predictors. The results showed that age, vital signs (body temperature, respiratory rate, SBP), and laboratory indicators (BUN, creatinine, WBC, RDW, hemoglobin) may be different between genders. Age is a traditional risk factor for HF, which can cause changes in the cardiovascular system, which has also been verified in this study. In addition, we observed that there may be potential differences between the sexes in other measures (such as MCV) and some drug treatments.

4.1. Clinical Scores and Vital Signs

4.1.1. Clinical Scores

Our study combines two predictors ranking methods (shown in Tables S1 and S2). We found that Acute Physiological Score III (APSIII) and Sequential Organ Failure Assessment (SOFA) have a high potential to predict in-hospital mortality in ICU patients with heart failure. Previous studies have proposed APSIII score as a part of APACHE equations to predict the outcome of ICU patients [17]. In addition, SOFA score has also been used to predict long-term all-cause mortality in patients with heart failure [18].

4.1.2. Vital Signs

In our study (see also Tables S1-S2), SBP was strongly associated with sex prediction models in ICU patients with heart failure. In contrast, SBP was a more powerful predictor in female patients. According to previous studies, similar results have been found in general hospitalized patients with heart failure, and compared with men, women have a stronger association between SBP and HF risk [19,20]. Pathophysiologically, these findings may be due to a variety of factors, including sex hormone factors, chromosomal factors, and sex-biased nonchromosomal gene expression differences. Even factors such as social environment and life experience can cause differences in vascular biology and physiology between men and women (for example, women have a smaller coronary artery diameter than men) [21]. Body temperature and respiratory rate were more important predictors for male patients than for women (see also Tables S1-S2). HF patients are often accompanied by sleep-disordered breathing, resulting in decreased respiratory rate or apnea, which increases the pressure on the sympathetic nervous system and negatively affects cardiovascular function [22]. Studies have shown that in patients < 50 years old, the respiratory rate of men is lower than that of women, and the respiratory rate of men ≥50 years old tends to be the same as that of women [23]. In addition, thermoregulation disorder is a recognized feature of advanced heart failure. Studies have revealed that hypothermia is associated with adverse outcomes in both hospitalized and ICU patients with heart failure [24,25]. Other studies have shown that the average body temperature of men is slightly lower

than that of women [26]. Therefore, the sex-specific appearance of body temperature and respiratory rate may be due to physiological differences between men and women.

4.2. Laboratory Index

4.2.1. Renal Function Index

According to HF guidelines, kidney injury and deterioration of renal function are common in HF patients and are associated with adverse outcomes in HF patients, and blood urea nitrogen (BUN) and creatinine are both common markers of renal function (and injury) [27]. Studies have found that BUN and creatinine are associated with the mortality of critically ill patients admitted to ICU, and BUN has an additive value in patients with HF [28]. In patients with HF, BUN reflects hemodynamic and neurohormonal changes, and elevated BUN levels may be due to insufficient arterial filling in patients with HF, which activates neurohormones, leading to increased urea reabsorption [29]. Previous studies have shown that BUN and creatinine levels differ between the sexes, and men are higher than women [30], which is confirmed in our study. We found that BUN and creatinine, as predictors of HF mortality, were stronger in the male model (see also Tables S1-S2) and that creatinine was inversely associated with HF mortality, with increased creatinine levels frequently observed in HF patients. Higher creatinine levels are associated with worse outcomes and higher mortality.

4.2.2. Cardiovascular Index

In our study, white blood cell count (WBC), red blood cell distribution width (RDW), and hemoglobin also showed potential to predict mortality in women (see Tables S1-S2). Previous studies [31] have shown that high white blood cell counts are associated with increased in-hospital mortality in ICU patients with heart failure, and our results show that this association is even more pronounced in women. Physiologically, changes in the WBC count in women are attributed to changes in the number of neutrophils, and neutrophil counts appear to change with changes in estrogen levels during the female menstrual cycle [32]. RDW, as an indicator of changes in red blood cell size, is strongly associated with the risk of adverse heart failure outcomes and increases with age, especially in women, who have slightly higher RDW than men [33]. In addition, studies have shown that hemoglobin also shows a sex difference in predicting in-hospital mortality in ICU patients with heart failure, which may be due to biological differences between men and women. Previous studies have shown that the average hemoglobin level of healthy venous blood is different between men and women, and the average hemoglobin level of women is lower than that of men [34], which may provide a physiological basis for our findings.

4.3. Limitations

Study limitations include the use of only four machine learning classifiers, which may overlook other possibilities. Despite cross-validation, redundancy in reference data could lead to biased results, and observed gender differences may stem from modeling variations. Thus, these findings are exploratory and require further validation. Additionally, CART model risks may not accurately reflect real-world populations, so causality should not be inferred. Future research should expand model selection and incorporate more clinical features to better understand gender differences in HF patients. The study used a U.S.-based heart failure database, limiting generalizability, but similar findings from other countries suggest broader relevance [35,36]. Future studies should include diverse population data.

5. CONCLUSION

Machine learning may allow the development of predictive models that simultaneously evaluate many relevant predictors of mortality in patients with HF. This study used a variety of machine learning models to analyze the risk of death during hospitalization in HF patients in intensive care

units and explored sex differences in predictors. We found predictors consistent with the known risk of death in patients with HF but also identified some potentially important factors that deserve to be the focus of future research.

REFERENCES

[1] Alper AB, Campbell RC, Anker SD et al. A propensity-matched study of low serum potassium and mortality in older adults with chronic heart failure. Int J Cardiol 2009;137:1–8.

[2] Braunwald E. The war against heart failure: the Lancet lecture. Lancet (London, England) 2015;385:812–24.

[3] Arata A, Ricci F, Khanji MY et al. Sex Differences in Heart Failure: What Do We Know? J Cardiovasc Dev Dis 2023;10, DOI: 10.3390/JCDD10070277.

[4] Timmis A, Vardas P, Townsend N et al. European Society of Cardiology: cardiovascular disease statistics 2021. Eur Heart J 2022;43:716–99.

[5] Nguyen QD, Peters E, Wassef A, Desmarais P, Rémillard-Labrosse D, Tremblay-Gravel M. Evolution of Age and Female Representation in the Most-Cited Randomized Controlled Trials of Cardiology of the Last 20 Years. Circ Cardiovasc Qual Outcomes 2018;11, DOI: 10.1161/CIRCOUTCOMES.118.004713.

[6] Agarwala A, Michos ED, Samad Z, Ballantyne CM, Virani SS. The Use of Sex-Specific Factors in the Assessment of Women's Cardiovascular Risk. Circulation 2020;141:592–9.

[7] Roger VL. Epidemiology of Heart Failure. Circ Res 2013;113:646–59.

[8] Cannatà A, Fabris E, Merlo M et al. Sex Differences in the Long-term Prognosis of Dilated Cardiomyopathy. Can J Cardiol 2020;36:37–44.

[9] Mansur A de P, Pereira-Barretto AC, del Carlo CH et al. Sex Differences in Prognosis of Heart Failure Due to Ischemic and Nonischemic Cardiomyopathy. J Clin Med 2023;12, DOI: 10.3390/JCM12165323.

[10] Raafs A, Verdonschot J, Ferreira JP et al. Identification of sex-specific biomarkers predicting new-onset heart failure. ESC Hear Fail 2021;8:3512–20.

[11] Han D, Xu F, Zhang L et al. Early prediction of in-hospital mortality in patients with congestive heart failure in intensive care unit: a retrospective observational cohort study. BMJ Open 2022;12:1–11.

[12] Johnson AEW, Bulgarelli L, Shen L et al. Author Correction: MIMIC-IV, a freely accessible electronic health record dataset. Sci Data 2023 101 2023;10:1–1.

[13] Pollard TJ, Johnson AEW, Raffa JD, Celi LA, Mark RG, Badawi O. The eICU collaborative research database, a freely available multi-center database for critical care research. Sci Data 2018;5:1–13.

[14] Gradus JL, Rosellini AJ, Horváth-Puhó E et al. Prediction of Sex-Specific Suicide Risk Using Machine Learning and Single-Payer Health Care Registry Data from Denmark. JAMA Psychiatry 2020;77:25–34.

[15] Pedregosa F, Varoquaux G, Gramfort A et al. Scikit-learn: Machine Learning in {P}ython. J Mach Learn Res 2011;12: 2825--2830.

[16] Murdoch WJ, Singh C, Kumbier K, Abbasi-Asl R, Yu B. Definitions, methods, and applications in interpretable machine learning. Proc Natl Acad Sci U S A 2019;116:22071–80.

[17] Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med 1985;13:818–29.

[18] Aoyama D, Morishita T, Uzui H et al. Sequential organ failure assessment score on admission predicts long-term mortality in acute heart failure patients. ESC Hear Fail 2020;7:245–53.

[19] Butler J, Kalogeropoulos AP, Georgiopoulou V V. et al. Systolic blood pressure and incident heart failure in the elderly. The Cardiovascular Health Study and the Health, Ageing and Body Composition Study. Heart 2011; 97: 1304–11.

[20] Ji H, Niiranen TJ, Rader F et al. Sex Differences in Blood Pressure Associations With Cardiovascular Outcomes. Circulation 2021;143:761–3.

[21] Ji H, Kim A, Ebinger JE et al. Sex Differences in Blood Pressure Trajectories Over the Life Course. JAMA Cardiol 2020;5:255–62.

[22] Lévy P, Naughton MT, Tamisier R, Cowie MR, Bradley TD. Sleep apnoea and heart failure. Eur Respir J 2022;59, DOI: 10.1183/13993003.01640-2021.

[23] Natarajan A, Su HW, Heneghan C, Blunt L, O'Connor C, Niehaus L. Measurement of respiratory rate using wearable devices and applications to COVID-19 detection. NPJ Digit Med 2021;4, DOI: 10.1038/S41746-021-00493-6.

[24] Payvar S, Spertus JA, Miller AB et al. Association of low body temperature and poor outcomes in patients admitted with worsening heart failure: a substudy of the Efficacy of Vasopressin Antagonism in Heart Failure Outcome Study with Tolvaptan (EVEREST) trial. Eur J Heart Fail 2013;15:1382–9.

[25] Xu H, Xie Y, Sun X, Feng N. Association between first 24-h mean body temperature and mortality in patients with diastolic heart failure in intensive care unit: A retrospective cohort study. Front Med 2022;9, DOI: 10.3389/ FMED. 2022. 1028122/FULL.

[26] Waalen J, Buxbaum JN. Is Older Colder or Colder Older? The Association of Age With Body Temperature in 18,630 Individuals. Journals Gerontol Ser A Biol Sci Med Sci 2011;66A:487.

[27] Van Veldhuisen DJ, Ruilope LM, Maisel AS, Damman K. Biomarkers of renal injury and function: diagnostic, prognostic and therapeutic implications in heart failure. Eur Heart J 2016;37:2577–85.

[28] Arihan O, Wernly B, Lichtenauer M et al. Blood Urea Nitrogen (BUN) is independently associated with mortality in critically ill patients admitted to ICU. PLoS One 2018;13, DOI: 10.1371/JOURNAL.PONE.0191697.

[29] Hong C, Zhu H, Zhou X et al. Association of Blood Urea Nitrogen with Cardiovascular Diseases and All-Cause Mortality in USA Adults: Results from NHANES 1999-2006. Nutrients 2023;15, DOI: 10.3390/NU15020461.

[30] Xu Z-R, Zhang Q, Chen L-F et al. Characteristics of hyperuricemia in older adults in China and possible associations with sarcopenia. Aging Med 2018;1:23.

[31] Huang AA, Huang SY. Dendrogram of transparent feature importance machine learning statistics to classify associations for heart failure: A reanalysis of a retrospective cohort study of the Medical Information Mart for Intensive Care III (MIMIC-III) database. PLoS One 2023;18, DOI: 10.1371/JOURNAL.PONE.0288819.

[32] Rathod KS, Kapil V, Velmurugan S et al. Accelerated resolution of inflammation underlies sex differences in inflammatory responses in humans. J Clin Invest 2017;127:169.

[33] Lippi G, Salvagno GL, Guidi GC. Red blood cell distribution width is significantly associated with aging and gender. Clin Chem Lab Med 2014;52:e197–9.

[34] Murphy WG. The sex difference in haemoglobin levels in adults — Mechanisms, causes, and consequences. Blood Rev 2014;28:41–7.

[35] Cai A, Qiu W, Xia S et al. Sex-specific characteristics and outcomes in hospitalized heart failure with preserved ejection fraction: the China Cardiovascular Association Database-Heart Failure Center Registry. Eur Heart J 2023;44:4715–8.

[36] Suthahar N, Lau ES, Blaha MJ et al. Sex-Specific Associations of Cardiovascular Risk Factors and Biomarkers With Incident Heart Failure. J Am Coll Cardiol 2020;76:1455.

APPENDIX

1. Figure S1

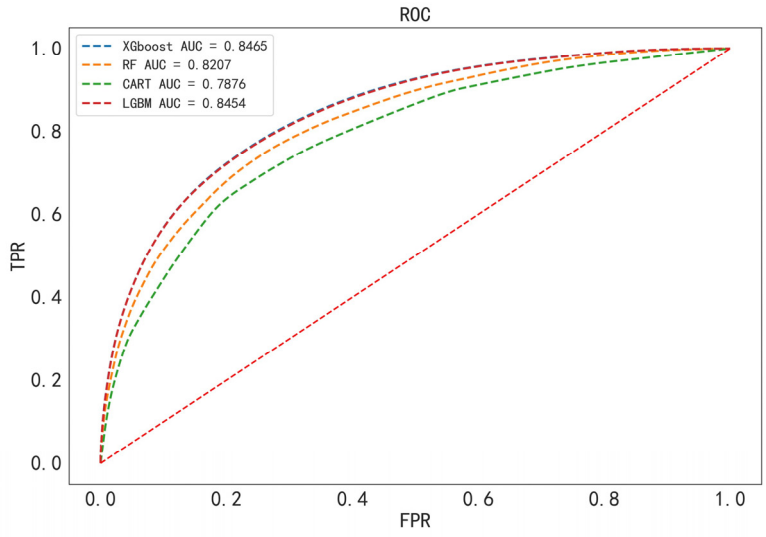

Figure S1. Comparison of multiple models. By drawing the ROC curves of four machine learning models (subjects compare model performance). AUC is defined as the area under the ROC curve, AUC∈[0,1]. The closer the value is to 1, the better the classification ability of the model.

2. Tables S1-S2

We used four machine learning models to analyze feature importance in different gender groups, recording the top 10 features. The final ranking of each feature is based on the median of the four model results, with features ranked above 10 labeled as 10. The detailed ranking is shown in Online.

We applied the SHAP interpretability method to enhance four machine learning models and analyzed feature importance for different genders, recording the top 10 features. The final ranking is based on the median of the four models, with ranks above 10 denoted as 10. The results are shown in Table S2.

Table S1. Ranking of Different Model Variables

	Male					Female				
	XGBoost	RF	CART	LightGBM	median	XGBoost	RF	CART	LightGBM	median
variable	Rank	Rank	Rank	Rank	Rank	Rank	Rank	Rank	Rank	Rank
Norepinephrine	3	4	-	-	7	1	4	5	-	4.5
Vasopressin	2	3	2	-	2.5	2	2	2	-	2
APSIII	1	1	1	1	1	3	1	1	3	2
Dopamine	6	10	-	-	10	-	-	-	-	-
SOFA	4	2	3	-	3.5	4	3	3	-	3.5
Age	5	-	4	2	4.5	8	-	7	1	7.5
RespiratoryRate	8	-	7	-	9	-	-	-	9	10
Vent	7	-	-	-	10	7	-	-	-	10
Intubated	-	-	-	-	-	6	-	-	-	10
RDW	10	-	10	4	10	-	10	9	6	9.5
Phenylephrine	-	8	-	-	10	5	5	-	-	7.5
AG	-	7	-	-	10	10	9	6	-	9.5
SBP	9	6	5	7	6.5	9	6	4	2	5
BUN	-	5	-	10	10	-	-	-	-	-
Temperature	-	9	6	8	8.5	-	7	-	4	8.5
MCV	-	-	8	-	10	-	-	-	-	-
Creatinine	-	-	-	9	10	-	-	-	-	-
WBC	-	-	-	3	10	-	8	8	5	8
Platelet	-	-	9	-	10	-	-	-	8	10
Hb	-	-	-	-	-	-	-	10	10	10
MCHC	-	-	-	5	10	-	-	-	-	-
HeartRate	-	-	-	6	10	-	-	-	7	10
Calcium	-	-	-	-	-	-	-	5	-	10

AG, anion gap; APSIII, Acute Physiology Score III; BUN, blood urea nitrogen; Hb, hemoglobin; MCHC, mean corpuscular hemoglobin concentration; MCV, mean corpuscular volume; RDW, red blood cell distribution width; SBP, systolic blood pressure; SOFA, Sequential Organ Failure Assessment; WBC, white blood cell count.

Table S2. Ranking of Different Model Variables (SHAP)

	Male					Female				
	Xgboost	RF	CART	LightGBM	median	Xgboost	RF	CART	LightGBM	median
variable	Rank	Rank	Rank	Rank	Rank	Rank	Rank	Rank	Rank	Rank
Norepinephrine	-	3	-	-	10	-	3	-	10	10
Vasopressin	-	6	4	9	7.5	-	4	3	-	7
APSIII	1	1	1	1	1	1	1	1	1	1
SOFA	3	2	2	3	2.5	3	2	4	3	3
Age	2	5	3	2	2.5	2	-	8	2	5
RespiratoryRate	4	8	7	4	5.5	7	-	-	6	8.5
RDW	6	-	-	6	8	4	8	10	4	6
Phenylephrine	-	-	-	-	-	-	7	-	-	10
AG	-	9	-	6	9.5	6	-	6	8	7
SBP	10	7	5	8	7.5	9	5	2	9	7
BUN	5	4	-	5	5	-	9	-	-	10
Temperature	8	-	6	9	8.5	-	-	-	-	-
Chloride	-	-	9	-	10	-	-	-	-	-
MCV	9	-	-	8	9.5	-	-	-	-	-
Creatinine	7	10	10	7	8.5	-	10	-	-	10
WBC	-	-	-	-	-	5	6	5	5	5
Platelet	-	-	-	-	-	-	-	-	-	-
MCHC	-	-	-	-	-	8	-	-	7	9
HeartRate	-	-	-	-	-	10	-	-	-	10
Calcium	-	-	8	-	10	-	-	9	-	10
Vent	-	-	-	10	10	-	-	-	-	-
Hb	-	-	-	-	-	-	-	7	-	10

AG, anion gap; APSIII, Acute Physiology Score III; BUN, blood urea nitrogen; Hb, hemoglobin; MCHC, mean corpuscular hemoglobin concentration; MCV, mean corpuscular volume; RDW, red blood cell distribution width; SBP, systolic blood pressure; SOFA, Sequential Organ Failure Assessment; WBC, white blood cell count.