Journal of Translational Medicine

RESEARCH

Open Access

Development and validation of novel interpretable survival prediction models based on drug exposures for severe heart failure during vulnerable period

Yu Guo[1,2,3†], Fang Yu[1†], Fang-Fang Jiang[1,2†], Sun-Jun Yin[1†], Meng-Han Jiang[1,2], Ya-Jia Li[1,2], Hai-Ying Yang[1,2], Li-Rong Chen[1,2], Wen-Ke Cai[4*] and Gong-Hao He[1*]

Abstract

Background Severe heart failure (HF) has a higher mortality during vulnerable period while targeted predictive tools, especially based on drug exposures, to accurately assess its prognoses remain largely unexplored. Therefore, this study aimed to utilize drug information as the main predictor to develop and validate survival models for severe HF patients during this period.

Methods We extracted severe HF patients from the MIMIC-IV database (as training and internal validation cohorts) as well as from the MIMIC-III database and local hospital (as external validation cohorts). Three algorithms, including Cox proportional hazards model (CoxPH), random survival forest (RSF), and deep learning survival prediction (DeepSurv), were applied to incorporate the parameters (partial hospitalization information and exposure durations of drugs) for constructing survival prediction models. The model performance was assessed mainly using area under the receiver operator characteristic curve (AUC), brier score (BS), and decision curve analysis (DCA). The model interpretability was determined by the permutation importance and Shapley additive explanations values.

Results A total of 11,590 patients were included in this study. Among the 3 models, the CoxPH model ultimately included 10 variables, while RSF and DeepSurv models incorporated 24 variables, respectively. All of the 3 models achieved respectable performance metrics while the DeepSurv model exhibited the highest AUC values and relatively lower BS among these models. The DCA also verified that the DeepSurv model had the best clinical practicality.

Conclusions The survival prediction tools established in this study can be applied to severe HF patients during vulnerable period by mainly inputting drug treatment duration, thus contributing to optimal clinical decisions prospectively.

†Yu Guo, Fang Yu, Fang-Fang Jiang and Sun-Jun Yin contributed equally to this work and share first authorship.

*Correspondence:
Wen-Ke Cai
caiwenke002@126.com
Gong-Hao He
gonghow@hotmail.com; gonghow@dali.edu.cn

Full list of author information is available at the end of the article

Keywords Severe heart failure, Survival model, Cox proportional hazards model, Random survival forest, DeepSurv

Background

Heart failure (HF) is the leading cause of cardiovascular morbidity and mortality, which becomes a major threat to human health and social development and causes a series of serious medical burdens [1, 2]. Despite massive investments devoted to the prevention and treatment of HF worldwide [3], HF patients still had a higher rate of hospitalization and may progress to severe HF that was prone to admission to the intensive care unit (ICU) [4, 5]. Although the related symptoms of HF patients might be quickly alleviated by optimal treatment during hospitalization, patients were still reported to have up to 15% of death within the first 3 months of discharge, which was commonly called the vulnerable period of HF [6]. In addition, it was also shown that the risk of all-cause death in HF patients at this stage was increased by 4 to 6 times compared with HF patients who were not hospitalized [7]. According to these findings, an increasing number of scholars currently believe that shifting the focus appropriately to HF patients during their vulnerable post-discharge period and strengthening their risk assessment, monitoring, management, and treatment could be the most effective and cost-effective strategy to address the public health and economic burdens associated with HF [6, 7]. Therefore, considering that patients with severe HF admitted to ICU have worse prognoses during the vulnerable period, ideal tools that can precisely predict survival still need to be developed urgently.

Currently, several scoring systems, such as the Sequential Organ Failure Assessment (SOFA) and the Acute Physiology Score (APS III), are valuable for assessing the condition of ICU patients [8, 9]. However, these systems have limited predictive power and lack pertinence and sensitivity in evaluating the prognosis of severe HF patients [10, 11]. Furthermore, although several new models have been developed to predict the prognoses of severe HF patients, they either lacked the prediction of mortality risk during the vulnerable period [12, 13] or contained multiple complex parameters that might eventually overlook the significant effects of drug therapy on the prognoses of HF patients [14–17]. From the perspective of clinical application, compared to the predictors that predominantly relied on devices, drug therapy information may be more accessible and more convenient for most users. Besides, since drug therapy is an essential means to help HF patients stabilize through the vulnerable period, including drug information as predictors is very likely to reduce certain errors regarding the predicted results and might also provide a reference for clinical decision-makers to formulate medication strategies for patients after discharge. Therefore, it is necessary to construct a simple and easy-to-use death risk prediction model that includes drug parameters for severe HF patients during the vulnerable period, which, however, has not been performed so far as we know.

In this study, we aimed to develop and validate models that mainly consisted of drug exposures based on machine learning (ML) and deep learning (DL) algorithms for predicting survival during the vulnerable period in severe HF patients. Moreover, we performed interpretability analyses on the models and evaluated the contribution of variables to prediction in combination with clinical significance, hoping to provide further practical prognostic prediction tools for severe HF management.

Methods

Date source

This study was conducted based on the Medical Information Mart for Intensive Care III v1.4 and IV v2.0, a publicly available database comprising medical data of patients who admitted to ICUs of the Beth Israel Deaconess Medical Center [18, 19]. The use of the database was approved by the institutional review boards of the Massachusetts Institute of Technology (MIT), and informed consent was not required as all data were de-identified. The principal investigators accessed the databases after completing the required courses and assessments (certificate numbers: 38884075 and 44408919). We randomized eligible HF patients from MIMIC-IV database into the training and internal validation cohorts in a 7:3 ratio, with MIMIC-III as the external validation cohort A. Furthermore, we also retrospectively included HF patients admitted to the 920th Hospital of Joint Logistics Support Force data from January 2020 to March 2023 as the external validation cohort B, which was approved by the ethics committee of the 920th Hospital.

This study followed the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement for reporting multivariable prediction models development and validation [20].

Study population

The research subjects were identified according to the ninth or tenth revision of International Classification of Diseases (ICD-9/10) code. Patients with the ICD diagnoses sequence of more than 3 were excluded. For patients with multiple hospital admissions, only data from their first admission were included. Besides, additional exclusion criteria are as follows: (1) age < 18 years old, (2) length of hospital or ICU stay < 24 h, (3) suffering from

malignant cancer, and (4) unrealistic hospitalization information.

Data extraction and preprocessing

The following information was extracted: demographics information, hospital and ICU length of stay (LOS), duration of medications, comorbidities, mechanical ventilation (MV), and renal replacement therapy (RRT). Some records within the first 24 h after ICU admission were collected, including vital signs, SOFA, APS III, and urine output. Drug classification was determined by the WHO-Anatomical Therapeutic and Chemical (ATC). Only systemic administration was reserved for each kind of drug, excluding local administration (eye, ear, topical, etc.). The exposure duration of medication was calculated by the start date and end date with repeated use of similar classes of drug on the same day not cumulatively counted.

The missing data were <25% and were imputed by multiple imputations. In order to eliminate the dimension and ensure the reliability of data, indicators need to be normalized. The formula $x^{*}=\frac{x\text{-}min*0.99}{max\text{-}min}$ was selected to avoid zero minimum values in partial continuous variables (age, body mass index [BMI], and LOS) [21]. For the exposure duration of drug, the formula $x^{*}=\frac{x}{\text{length of hospital stay}}$ was selected.

Model development and evaluation

The predictive factors included durations of drugs, hospital and ICU LOS, and demographics information. Three algorithms were chosen: Cox Proportional-Hazards model (CoxPH), Random Survival Forests (RSF), and Deep Learning Survival Neural Network (DeepSurv). Firstly, Spearman's correlation coefficients were calculated to exclude variables with strong correlations. The LASSO method was applied to avoid CoxPH model overfitting. The method reduces the coefficients of irrelevant variables to zero while retaining important variables [22]. Meanwhile, univariate and multivariate Cox regression were performed for reference. Each contribution of predictors of the CoxPH model was measured as the partial chi-square statistic minus the degrees of freedom of predictors.

The RSF is an ensemble method, in which multiple decision trees are trained on a random sample of observations from the study data [23]. Their prediction is combined by using a mean value or majority vote. The DeepSurv is a feed-forward neural network method based on the Cox proportional hazards model, which was proven to perform well without prior assumptions on the risk function [24]. These 2 methods had more efficient data processing capabilities; thus, variable pre-screening was not carried out. For hyperparameter tuning, the optimal parameter combination of these models was determined by random search [25].

The predictive accuracy and discriminative ability of models were determined by the area under the receiver operating characteristics curves (AUC) and cumulative/dynamic time-dependent AUC. Brier scores (BS) were evaluated for assessing calibration, which measured the mean squared error between the actual survival and the estimated probabilities. Then we integrated the BS to get the integrated Brier score (IBS). A lower score indicates better calibration and only models with scores below 0.25 are deemed useful in practice. We evaluated the clinical effectiveness of models by decision curve analysis (DCA), which was reported to have certain advantages over other evaluating tools [26].

To enable ML model interpretability, permutation importance and Shapley additive explanations (SHAP) analyses were implemented in this study. Permutation importance is calculated through the increase in the prediction error of models after randomly shuffling each feature [27]. The latter uses a concept from cooperative game theory to assign each feature importance score and rank them based on the impact of features on model predictions [28].

Statistical analysis

Continuous variables were described as mean and standard-deviation (SD) values or median and interquartile-range (IQR) values depending on whether they conformed to a normal distribution. Categorical variables were summarized as frequency and percentage values. P values of less than 0.05 were considered statistically significant. The R (version 4.1.2) software was used for data preprocessing. The models were implemented in Python (version 3.6).

Results

Baseline characteristics

After applying the inclusion criteria and data preprocessing (Fig. 1), 6873 severe HF patients were identified from the MIMIC-IV database, with 4811 and 2062 in training and internal validation cohorts, respectively. For external validation cohorts, we included 4463 and 254 participants from the MIMIC-III database and the 920th Hospital, respectively. The baseline clinical characteristics were listed in Table 1 and Supplementary Table S1. The proportions of gender were relatively balanced in the 3 cohorts. The patients were generally older, and the median ages of the 3 cohorts were 75 years (IQR 65–84 years), 73 years (IQR 62–81 years), and 72 years (IQR 62–78 years), respectively. For medication treatment during hospitalization, drugs with higher average days of use in all 3 cohorts were generally similar, mainly including β blockers, diuretics, antiplatelet agents, and THMG-CoA

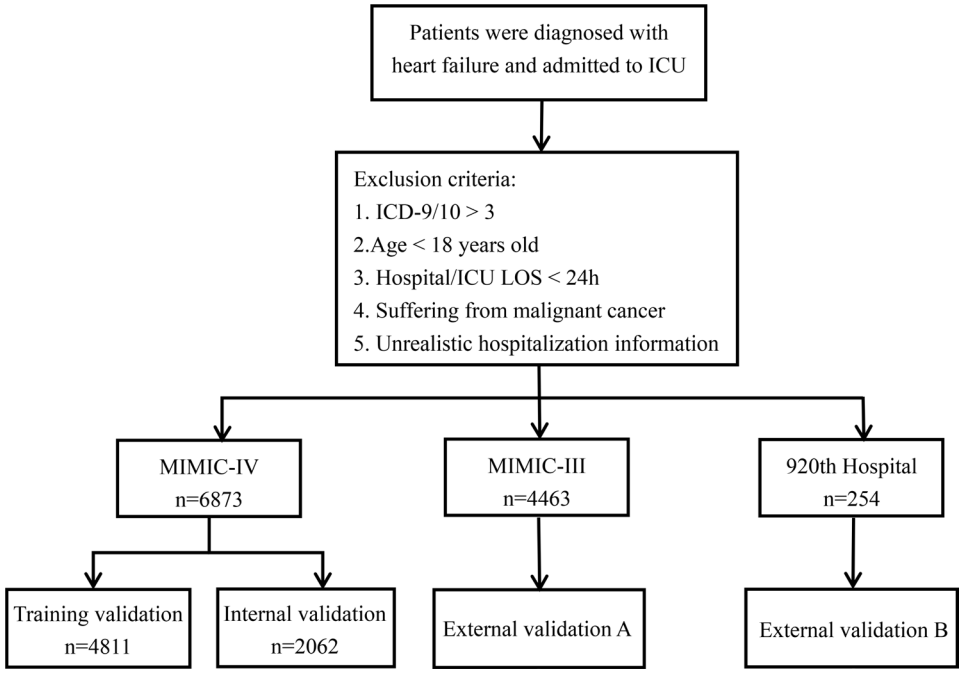

Fig. 1 Flowchart of study inclusion

Table 1 Demographic characteristics and drug exposures situation of participants

Characteristics	MIMIC-IV (Training/Internal validation) *N* = 6873	MIMIC-III (External validation A) *N* = 4463	920th Hospital (External validation B) *N* = 254
Age, median (IQR), years	75.0 (65.0–84.0)	73.0 (62.0–81.0)	72.0 (62.0–78.0)
Gender, n (%), male	3700 (53.8)	2487 (55.7)	144 (56.7)
BMI, median (IQR), kg/m^2	28.7 (24.5–33.7)	28.0 (24.2–32.8)	23.3 (20.5–25.7)
Drug exposures, mean (SD), days			
ACEIs	2.1 ± 3.6	3.0 ± 4.6	1.5 ± 3.5
ARBs	0.7 ± 2.2	0.5 ± 1.9	9.1 ± 8.3
β blockers	6.8 ± 6.4	7.3 ± 7.1	11.9 ± 8.3
Cardiac glycosides	0.8 ± 3.0	1.1 ± 3.6	1.6 ± 4.1
Diuretics	6.4 ± 6.0	6.7 ± 6.8	11.5 ± 8.4
Nitrates	1.7 ± 3.0	2.8 ± 4.3	1.3 ± 2.0
HMG-CoA reductase inhibitors	6.6 ± 7.2	5.7 ± 6.9	11.4 ± 8.9
CCBs	1.9 ± 4.1	1.7 ± 4.0	1.3 ± 3.8
PDEs inhibitors	0.3 ± 1.5	0.5 ± 1.9	1.7 ± 3.8
PPIs	4.7 ± 7.1	6.0 ± 8.2	7.6 ± 8.1
H2RAs	2.3 ± 4.3	2.4 ± 4.2	0.0 ± 0.0
Antibiotics	4.8 ± 5.8	5.8 ± 7.2	5.8 ± 6.4
Antiplatelet agents	6.9 ± 7.2	7.3 ± 7.7	10.5 ± 8.7
Insulin	5.8 ± 7.2	7.3 ± 8.1	0.6 ± 1.3
Heparin	7.0 ± 6.6	7.0 ± 7.6	4.6 ± 4.9
Adrenergic/dopaminergic agents	2.0 ± 4.0	2.4 ± 4.3	4.0 ± 6.7
Glucocorticoids	1.2 ± 3.8	1.4 ± 4.4	2.2 ± 2.6
Vitamin K antagonists	2.3 ± 4.5	2.1 ± 3.9	0.1 ± 1.1
Hospital LOS, median (IQR), days	8.0 (6.0–13.0)	9.0 (6.0–14.0)	13.0 (9.0–18.0)
ICU LOS, median (IQR), days	3.0 (2.0–5.0)	4.0 (3.0–6.0)	5.0 (3.0-9.8)

Abbreviations ACEIs, angiotensin-converting enzyme inhibitors; ARBs, angiotensin receptor blockers; CCBs, calcium channel blockers; PDEs inhibitors, phosphodiesterase inhibitors; PPIs, proton pump inhibitors; H2RAs, histamine H2 receptor antagonists; LOS, length of stay

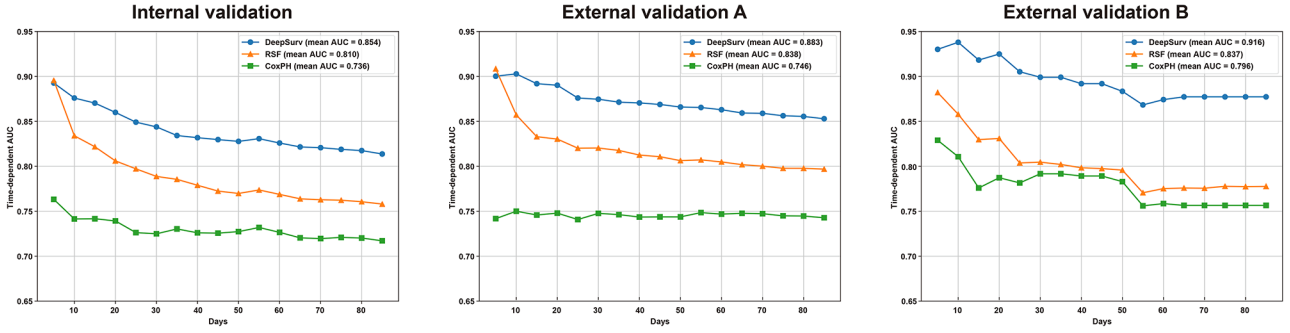

Fig. 2 Time-dependent area under the curve (AUC) for CoxPH, RSF, and DeepSurv models

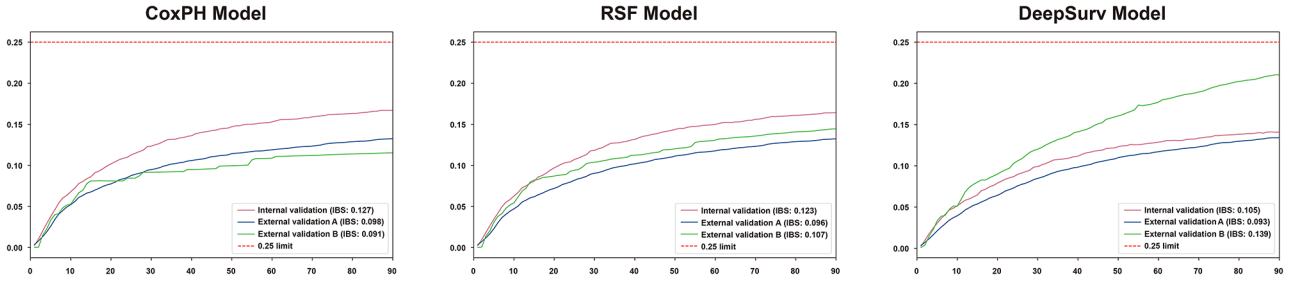

Fig. 3 Prediction error curves show the brier score for CoxPH, RSF, and DeepSurv models at each time point. As a benchmark, a useful model will have a Brier score below 0.25

reductase inhibitors. The frequencies of the use of related drugs were shown in Supplementary Table S2.

Development and validation of the survival prediction models

In this study, we used training cohort to generate and train 3 models (CoxPH, RSF, and DeepSurv) for survival prediction. A total of 24 variables, including 6 personal information variables (age, gender, BMI, hospital and ICU LOS, and ethnicity) and 18 types of drug exposures situations, were selected for constructing models. Spearman's analysis demonstrated no strong correlation (the absolute value of correlation coefficient <0.7), and the corresponding heatmap was shown in Supplementary Figure S1. For variable-screening, 10 variables were selected by the LASSO analysis and the related plots were displayed in Supplementary Figure S2. Meanwhile, the multivariate Cox analysis included 18 variables with $P<0.1$ in univariate analysis (Table S3 in Supplement). Based on these results, 10 independent predictive variables, including age, BMI, ICU LOS, and exposure duration of angiotensin-converting enzyme inhibitors (ACEIs), angiotensin receptor blockers (ARBs), β blockers, nitrates, antibiotics, Vitamin K antagonists, and histamine H2 receptor antagonists (H2RAs), eventually entered the CoxPH model. For the development of RSF and DeepSurv models, 24 features were fully incorporated. After a 100-repeated random search with 5-fold cross-validation, we chose those parameters showing the highest average C-index in cross-validation as the optimal parameters.

The validation results showed that all the 3 models exhibited high AUC values for predicting the survival of severe HF patients during the vulnerable period regardless of the cohort (Fig. 2). The time-dependent AUCs exceeded 0.7, indicating that these 3 models had dependable abilities at predicting long-term and short-term survival status. In internal and external validation cohort A and B, the DeepSurv model (mean AUC=0.854, 0.883, and 0.916) outperformed RSF (mean AUC=0.810, 0.838, and 0.837) and CoxPH (mean AUC=0.736, 0.746, and 0.796) models at most time points. The confusion matrix and several relevant performance indicators for 30-, 60 -, and 90-day survival predictions of the 3 models were shown in Supplementary Table S4, Figure S3 and S4. The corresponding prediction error curves representing the BS (all less than 0.25) over time were depicted in Fig. 3, demonstrating that the 3 models had good discrimination abilities. After integration, the IBS of DeepSurv model was also superior to that of CoxPH and RSF model in the internal and external validation cohort A (Fig. 3). For clinical utility, the net benefit of the 3 models was greater over the threshold probabilities range in 2 external validation cohorts, with the DeepSurv model having the greatest net benefit (Figure S5 in Supplement). As the DeepSurv model exhibited the most effective predictive performance, we then also conducted subgroup analyses to evaluate its model performance, which further confirmed that this model had ideal predictive performances

in various kinds of subpopulations (Supplementary materials Table S5).

Feature importance

The contribution of each predictor of the completed CoxPH model was visualized by scatter diagram (Supplementary Figure S6). Assessing variable permutation importance identified features important to prediction accuracy of DeepSurv and RSF models, with a more than 1% mean reduction in C-index with permutation of age, ICU and hospital LOS, BMI, and exposure duration of adrenergic/dopaminergic agents, ACEIs, nitrates, antibiotics, β blockers, H2RAs, Vitamin K antagonists, HMG-CoA reductase inhibitors, heparin, diuretics, and antiplatelet agents (Fig. 4). The feature importance ranking based on this method for the 2 models was listed in Supplementary Table S4. Furthermore, SHAP was used to explain the best DeepSurv model, which provided additional information regarding the prognosis direction of features. The summarized plots of positive or negative impact of factors were provided and clinical features were ranked by the average absolute value of SHAP in Fig. 5. It was observed that age was the top-ranked predictive variable in the 3 models, followed by ICU LOS, and that these 2 features had negative SHAP values, which drove the prediction toward survival and were strongly consistent with clinical consensus that advanced age was a well-known risk factor for cardiovascular events [29]. Likewise, ICU LOS is also usually used as a key reference to judge the condition of severe patients and therefore has a great prediction value for survival prognoses in practice. On the contrary, some factors with positive SHAP values also supported survival prediction, such as longer hospital LOS and exposure duration of ACEIs, β blockers, and H2RAs, which are all conducive to the better improvement of HF symptoms. Overall, all models had comparable variable importance profiles, suggesting that these models mainly based on drug exposures were reasonable and promising.

Discussion

To the best of our knowledge, this is the first study to develop and validate a series of drug-based clinical prediction models for predicting the survival of severe HF patients. The extensive sample size and reliable artificial intelligence algorithm enabled these models to be more universal and more accurate in predicting the prognoses of patients. Our results showed that the models constructed separately by the 3 algorithms had satisfactory prediction performance. Moreover, the drug-based DeepSurv model was significantly better than the CoxPH and RSF models in terms of clinical utility. Overall, our study innovatively explored the main use of drug regimens to predict prognoses for severe HF patients and provided a new reference path for physicians to make clinical decision management.

As a particular period of HF, the vulnerable phase is a transition stage from acute decompensated HF (hospitalization) to chronic stable HF and hence the treatment process in and out of hospital may remarkably impact the disease outcome at this stage. Besides, previous studies showed that the underlying pathophysiological mechanism for HF patients with poor prognoses in vulnerable phase was typically related to hemodynamic congestion [30]. In this regard, the application of drugs during hospitalization is one of the primary treatments for early improvement of related hemodynamic symptoms. In fact, the 2021 ESC HF Guidelines have already recommended that oral medication should be given to hospitalized HF patients as early as possible and optimized for multi-path therapy based on evidence to relieve persistent signs of congestion [31]. As a result, relevant drug exposures in hospitals are of great value in reducing the occurrence of adverse events during the vulnerable period and it is also feasible to use them as the main predictors to evaluate prognoses of HF patients. In addition, comparing with models that consisted of laboratory indicators or imaging data, models constructed by hospitalized drug exposures are not susceptible to various factors (e.g., instrument and human, etc.), which will avoid prediction errors caused by instability of certain indicators. Meanwhile, this kind of model will also significantly save time and cost, especially for patients in ICU who have relatively more mobility difficulties and higher medical costs. Therefore, the establishment of this series of models may be a kind of breakthrough of traditional models, which can assist clinical decision-makers in optimizing drug treatment strategies and will also provide novel ideas for predicting prognoses of HF as well as other diseases.

Recently, ML algorithms, especially DL algorithms, were widely used in survival prediction models with superior modeling capability and prediction performance [32–34]. Compared with the traditional CoxPH model, these models are able to identify complex feature networks more comprehensively and maximize the value of each factor during model construction, which is hence not prone to erroneous prediction results [35]. Given this, the RSF and DeepSurv models were constructed in the present study, which respectively contained 24 parameters for primary personal data and medication information. The comparison results of the 3 models showed that DeepSurv model had the best discrimination and calibration in internal and external validations, followed by RSF model, which further proved the robustness of the abovementioned learning algorithms and suggested that both models might have a specific scope of application. However, given the poor clinical utility of RSF model in external validation cohorts A and B,

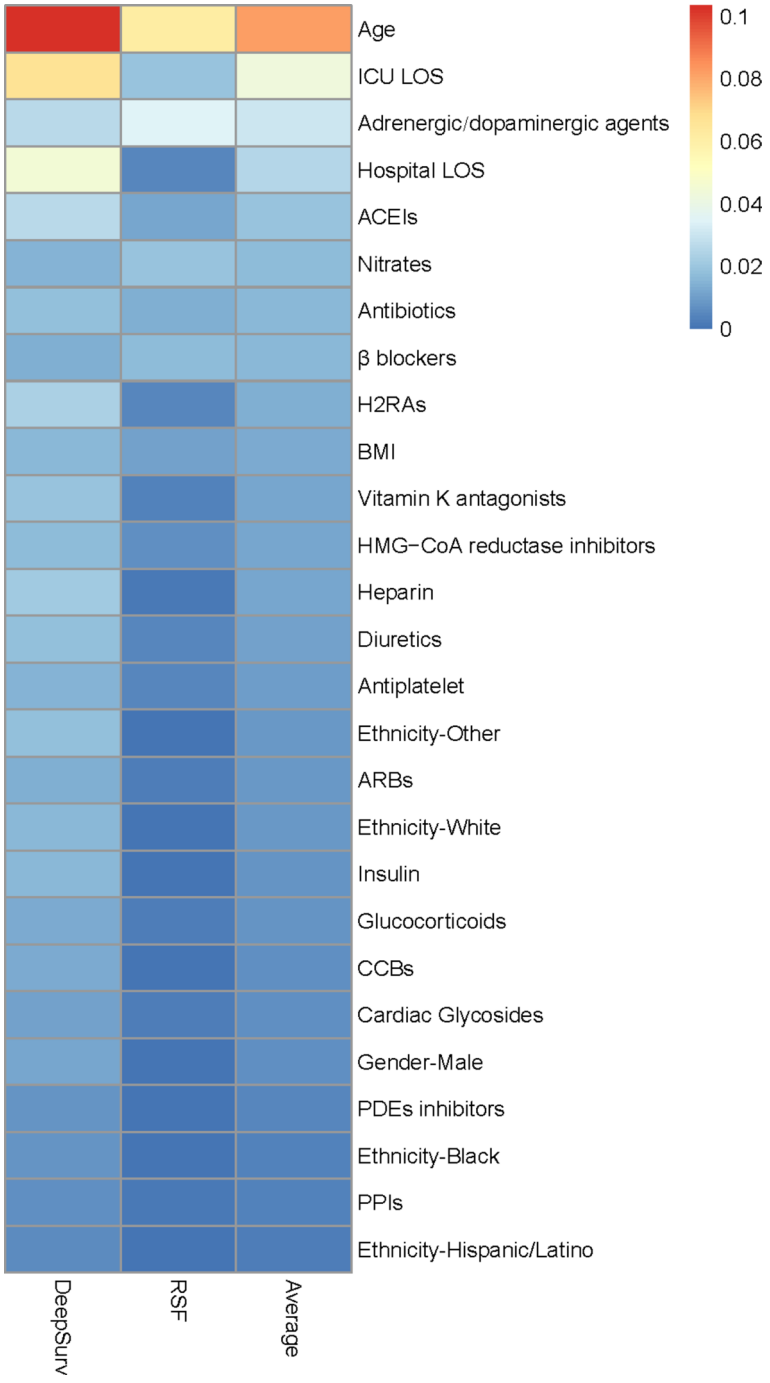

Fig. 4 Heatmap of feature importance for RSF and DeepSurv models. The values are expressed as a percentage reduction in the C-index after the value of a feature has been replaced by random numbers. Higher values suggest that a feature is more important in influencing the predictive accuracy of the corresponding models

this kind of model was more likely to be susceptible to the structural characteristics of the population. Therefore, the DeepSurv model would be more preferentially recommended for patients meeting their conditions of use. While for the current CoxPH model, although only 10 parameters were included, moderate predictive accuracy ($0.7 < AUC < 0.8$) was still achieved in different populations with disparate sociodemographic characteristics and risk profiles, which would also have a particular range of application. In this regard, CoxPH model may be recommended for survival prediction with patients with severe situations or limited available drug exposure information. In brief, these models are respectively applicable to different patient groups, which provide a new

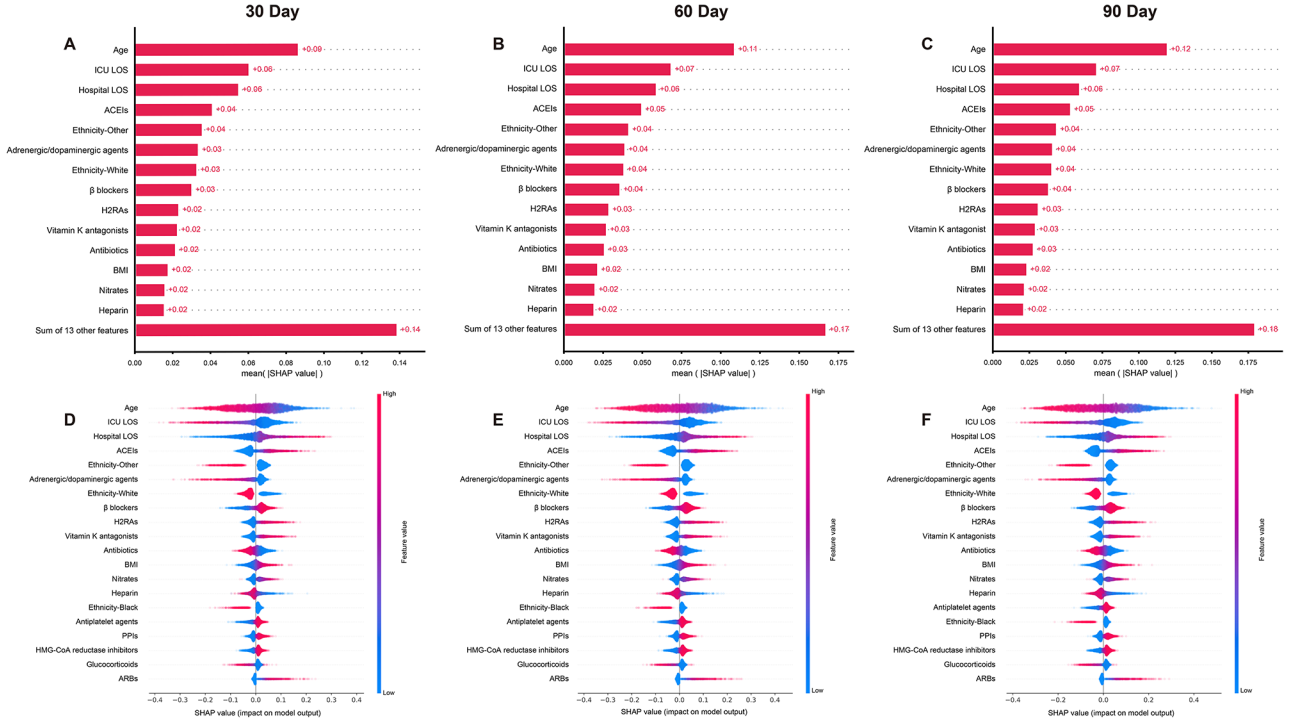

Fig. 5 Interpreting the results of DeepSurv model using SHAP explainer. Bar plots of mean absolute SHAP values (**A**-**C**): ranking of feature importance indicated by SHAP. The matrix plot depicts the importance of each covariate in the development of the final predictive model. SHAP summary plots for the top 20 clinical features (**D**-**F**): the higher the SHAP value of a feature, the higher the probability of survival development. Each line represents a feature, and the abscissa is the SHAP value. Red dots represent higher feature values, and blue dots represent lower feature values

approach for precise management of HF patients. Considering the convenience of users, web page and software presentation will be one of our future research directions after the model is well calibrated.

In the following interpretability analyses, we observed that, apart from certain basic personal characteristics, features regarding traditional anti-HF drug exposures were of crucial importance in artificial intelligence models, which further confirmed the rationality of the present models as well as our viewpoint regarding the necessity to integrate drug exposures parameters into construction of survival prediction models. Of note, we also observed that H2RAs, which were generally supposed to be unrelated to HF treatment, still obtained moderate weight and were even relatively comparable to β receptor antagonists in both artificial intelligence models (Figs. 4 and 5). These interesting findings strongly indicated that H2RAs might play important roles in the treatment of HF. This is quite reasonable as cardiac histamine H_2 receptor shares a common downstream signaling pathway with β1-receptor and has long been suggested to exert non-benign or even negative effects on cardiovascular system according to both previous experimental and clinical studies [36–39]. Moreover, our recent investigations further demonstrated that H2RAs exposure was associated with a lower mortality in patients with various kinds of cardiovascular diseases (including HF) and also confirmed the safety profile of H2RAs in cardiovascular system [40–43]. In this regard, considering the relatively significant predictive value of H2RAs exposure on HF survival prediction in the present study and the relatively strong adverse reactions and contraindications of β receptor antagonists, more attention should be paid to the potential treatment effect of H2RAs on HF patients in future studies.

The present study has several intrinsic limitations. First, given that this study was a retrospective analysis, the authenticity of the results might be influenced by inherent wrong and missing data. Second, it was impossible to collect the latest features that influenced HF mortality, such as NT-proBNP and Sacubitril/valsartan, due to the early establishment of the database. Therefore, further studies should consider adding the latest guidance-recommended drug to the model to enhance predictive power. Third, the dosages of drug were not further extracted because of the lack of comprehensive drug records in the MIMIC database. Fourth, MIMIC-III and IV collected patients from the same hospital at different times, which might somewhat reduce the generalizability of the present models. Although cohort from a local hospital was employed for the second external validation, the included sample size was still relatively small. These limitations should be addressed in future investigations.

Conclusions

In conclusion, the results of this study show that 3 different algorithms models based on drug exposures accurately identify the survival status of patients with severe HF during vulnerable period and that drug exposures as significant predictors are feasible and have considerable predictive value. Additionally, interpretive analysis comprehensively reveals the contribution of drug use to HF outcome prediction. Our research innovatively generates convenient and practical predictive tools for clinics, which will contribute to improving the individualized treatment of severe HF patients and provide direction for subsequent research on potential relationship between related drugs and HF.

Abbreviations

HF	Heart failure
ICU	Intensive care unit
SOFA	Sequential Organ Failure Assessment
APS III	Acute Physiology Score
ML	Machine learning
DL	Deep learning
MIT	Massachusetts Institute of Technology
TRIPOD	Individual Prognosis or Diagnosis
ICD	International Classification of Disease
LOS	Length of stay
MV	Mechanical ventilation
RRT	Renal replacement therapy
ATC	Anatomical Therapeutic and Chemical
BMI	Body mass index
CoxPH	Cox proportional hazards model
RSF	Random survival forest
DeepSurv	Deep learning survival prediction
AUC	Area under the receiver operating characteristics curves
BS	Brier scores
IBS	Integrated Brier score
DCA	Decision curve analysis
SHAP	Shapley additive explanations
SD	Standard-deviation
IQR	Interquartile-range
ACEI	Angiotensin-converting enzyme inhibitor
ARBs	Angiotensin receptor blockers
H2RAs	Histamine H2 receptor antagonists

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-024-05544-6.

Supplementary Material 1

Acknowledgements

The authors wish to acknowledge the MIMIC database and the 920th Hospital for providing clinical data for this study. The interpretation and reporting of these data are the sole responsibility of the authors.

Author contributions

YG, FY, FFJ, and GHH contributed to the conception or design of the work. FY, SJY, WKC, and YJL contributed to the acquisition and analysis of data for the work. MHJ, HYY, and LRC are responsible for the interpretation of data. YG and FFJ drafted the manuscript. YG and GHH critically revised the manuscript. All authors read and approved the final manuscript.

Funding

This study was supported by Grants from the National Science Foundation of China (No. 81960664), the Applied Basic Research Program Yunnan Province of China (Joint Special Project of Kunming Medical University) (No. 202101AY070001-300 and 202401AY070001-367), and the 2024 Young and Middle-aged Academic and Technical Leaders Reserve Talent Project of Yunnan Province (No. 202405AC350037).

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The MIMIC database was approved by the Massachusetts Institute of Technology (Cambridge, MA) and Beth Israel Deaconess Medical Center (Boston, MA), and consent was obtained for the original data collection. Additionally, the institutional review board (IRB) of 920th Hospital also approved our study and waived the need for informed consent due to the retrospective nature of this study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Clinical Pharmacy, The 920th Hospital of Joint Logistics Support Force, 212 Daguan Rd, Kunming 650032, China
[2]College of Pharmacy, Dali University, Dali 671000, China
[3]Yunnan Baiyao Group Limited Ltd, Kunming 650500, China
[4]Department of Cardiothoracic Surgery, The 920th Hospital of Joint Logistics Support Force, 212 Daguan Rd, Kunming 650032, China

Received: 8 January 2024 / Accepted: 29 July 2024
Published online: 06 August 2024

References

1. 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(15):1682–90. https://doi.org/10.1093/eurjpc/zwaa147.
2. Ambrosy AP, Fonarow GC, Butler J, Chioncel O, Greene SJ, Vaduganathan M, et al. The global health and economic burden of hospitalizations for heart failure: lessons learned from hospitalized heart failure registries. J Am Coll Cardiol. 2014;63(12):1123–33. https://doi.org/10.1016/j.jacc.2013.11.053.
3. Cook C, Cole G, Asaria P, Jabbour R, Francis DP. The annual global economic burden of heart failure. Int J Cardiol. 2014;171(3):368–76. https://doi.org/10.1016/j.ijcard.2013.12.028.
4. Metkus TS, Lindsley J, Fair L, Riley S, Berry S, Sahetya S, et al. Quality of heart failure care in the Intensive Care Unit. J Card Fail. 2021;27(10):1111–25. https://doi.org/10.1016/j.cardfail.2021.08.001.
5. Van Esbroeck ACM, Varga ZV, Di X, van Rooden EJ, Tóth VE, Onódi Z, et al. Activity-based protein profiling of the human failing ischemic heart reveals alterations in hydrolase activities involving the endocannabinoid system. Pharmacol Res. 2020;151:104578. https://doi.org/10.1016/j.phrs.2019.104578.
6. Greene SJ, Fonarow GC, Vaduganathan M, Khan SS, Butler J, Gheorghiade M. The vulnerable phase after hospitalization for heart failure. Nat Rev Cardiol. 2015;12(4):220–9. https://doi.org/10.1038/nrcardio.2015.14.
7. Solomon SD, Dobson J, Pocock S, Skali H, McMurray JJ, Granger CB, et al. Candesartan in Heart failure: Assessment of reduction in mortality and morbidity (CHARM) investigators. Influence of nonfatal hospitalization for heart failure on subsequent mortality in patients with chronic heart failure. Circulation. 2007;116(13):1482–7. https://doi.org/10.1161/CIRCULATIONAHA.107.696906.
8. Vincent JL, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, et al. The SOFA (Sepsis-related Organ failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-related problems of the European Society of Intensive Care Medicine. Intensive Care Med. 1996;22(7):707–10. https://doi.org/10.1007/BF01709751.

9. Moreno RP, Metnitz PG, Almeida E, Jordan B, Bauer P, Campos RA, et al. SAPS 3 investigators. SAPS 3–From evaluation of the patient to evaluation of the intensive care unit. Part 2: development of a prognostic model for hospital mortality at ICU admission. Intensive Care Med. 2005;31(10):1345–55. https://doi.org/10.1007/s00134-005-2763-5.
10. Rapsang AG, Shyam DC. Scoring systems in the intensive care unit: a compendium. Indian J Crit Care Med. 2014;18(4):220–8. https://doi.org/10.4103/0972-5229.130573.
11. Sekulic AD, Trpkovic SV, Pavlovic AP, Marinkovic OM, Ilic AN. Scoring systems in assessing survival of critically ill ICU patients. Med Sci Monit. 2015;21:2621–9. https://doi.org/10.12659/MSM.894153.
12. Ma H, Li D, Zhao J, Li W, Fu J, Li C. HR-BGCN: Predicting readmission for heart failure from electronic health records. Artif Intell Med. 2024;150:102829. https://doi.org/10.1016/j.artmed.2024.102829.
13. Pishgar M, Theis J, Del Rios M, Ardati A, Anahideh H, Darabi H. Prediction of unplanned 30-day readmission for ICU patients with heart failure. BMC Med Inf Decis Mak. 2022;22(1):117. https://doi.org/10.1186/s12911-022-01857-y.
14. Gao Z, Liu X, Kang Y, Hu P, Zhang X, Yan W, et al. Improving the Prognostic Evaluation Precision of Hospital Outcomes for Heart failure using admission notes and clinical Tabular Data: Multimodal Deep Learning Model. J Med Internet Res. 2024;26:e54363. https://doi.org/10.2196/54363.
15. Zhou Z, Liu Q, Zheng M, Zuo Z, Zhang G, Shi R, et al. Comparative study on the predictive value of TG/HDL-C, TyG and TyG-BMI indices for 5-year mortality in critically ill patients with chronic heart failure: a retrospective study. Cardiovasc Diabetol. 2024;23(1):213. https://doi.org/10.1186/s12933-024-02308-w.
16. Luo C, Zhu Y, Zhu Z, Li R, Chen G, Wang Z. A machine learning-based risk stratification tool for in-hospital mortality of intensive care unit patients with heart failure. J Transl Med. 2022;20(1):136. https://doi.org/10.1186/s12967-022-03340-8.
17. Li F, Xin H, Zhang J, Fu M, Zhou J, Lian Z. Prediction model of in-hospital mortality in intensive care unit patients with heart failure: machine learning-based, retrospective analysis of the MIMIC-III database. BMJ Open. 2021;11(7):e044779. https://doi.org/10.1136/bmjopen-2020-044779.
18. Yang J, Li Y, Liu Q, Li L, Feng A, Wang T, et al. Brief introduction of medical database and data mining technology in big data era. J Evid Based Med. 2020;13(1):57–69. https://doi.org/10.1111/jebm.12373.
19. Johnson AEW, Bulgarelli L, Shen L, Gayles A, Shammout A, Horng S, et al. MIMIC-IV, a freely accessible electronic health record dataset. Sci Data. 2023;10(1):1. https://doi.org/10.1038/s41597-022-01899-x.
20. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ. 2015;350:g7594. https://doi.org/10.1136/bmj.g7594.
21. Guo F, Zhu X, Wu Z, Zhu L, Wu J, Zhang F. Applications of machine learning in the survival prediction and classification of sepsis: coagulation and heparin usage matter. J Transl Med. 2022;20(1):265. https://doi.org/10.1186/s12967-022-03469-6.
22. Tibshirani R. Regression shrinkage and selection via the LASSO. J Royal Statist Soc. 1996;73:273–82. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.
23. Taylor JM. Random Survival forests. J Thorac Oncol. 2011;6(12):1974–5. https://doi.org/10.1097/JTO.0b013e318233d835.
24. Katzman JL, Shaham U, Cloninger A, Bates J, Jiang T, Kluger Y. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. BMC Med Res Methodol. 2018;18(1):24. https://doi.org/10.1186/s12874-018-0482-1.
25. Park SJ, Cho KJ, Kwon O, Park H, Lee Y, Shim WH, et al. Development and validation of a deep-learning-based pediatric early warning system: a single-center study. Biomed J. 2022;45(1):155–68. https://doi.org/10.1016/j.bj.2021.01.003.
26. Mogensen UB, Ishwaran H, Gerds TA. Evaluating Random Forests for Survival Analysis using Prediction Error curves. J Stat Softw. 2012;50(11):1–23. https://doi.org/10.18637/jss.v050.i11.
27. Altmann A, Toloşi L, Sander O, Lengauer T. Permutation importance: a corrected feature importance measure. Bioinformatics. 2010;26(10):1340–7. https://doi.org/10.1093/bioinformatics/btq134.
28. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to Global understanding with explainable AI for trees. Nat Mach Intell. 2020;2(1):56–67. https://doi.org/10.1038/s42256-019-0138-9.
29. Fonarow GC, Abraham WT, Albert NM, Stough WG, Gheorghiade M, Greenberg BH, et al. OPTIMIZE-HF investigators and hospitals. Age- and gender-related differences in quality of care and outcomes of patients hospitalized with heart failure (from OPTIMIZE-HF). Am J Cardiol. 2009;104(1):107–15. https://doi.org/10.1016/j.amjcard.2009.02.057.
30. Zile MR, Bennett TD, St John Sutton M, Cho YK, Adamson PB, Aaron MF, et al. Transition from chronic compensated to acute decompensated heart failure: pathophysiological insights obtained from continuous monitoring of intracardiac pressures. Circulation. 2008;118(14):1433–41. https://doi.org/10.1161/CIRCULATIONAHA.108.783910.
31. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, et al. 2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: developed by the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). With the special contribution of the Heart Failure Association (HFA) of the ESC. Eur J Heart Fail. 2022;24(1):4–131. https://doi.org/10.1002/ejhf.2333.
32. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol. 2017;2(4):230–43. https://doi.org/10.1136/svn-2017-000101.
33. Miller DD, Brown EW. Artificial Intelligence in Medical Practice: the question to the answer? Am J Med. 2018;131(2):129–33. https://doi.org/10.1016/j.amjmed.2017.10.035.
34. Wu WT, Li YJ, Feng AZ, Li L, Huang T, Xu AD, et al. Data mining in clinical big data: the frequently used databases, steps, and methodological models. Mil Med Res. 2021;8(1):44. https://doi.org/10.1186/s40779-021-00338-z.
35. Deo RC. Machine learning in Medicine. Circulation. 2015;132(20):1920–30. https://doi.org/10.1161/CIRCULATIONAHA.115.001593.
36. He G, Hu J, Li T, Ma X, Meng J, Jia M, et al. Arrhythmogenic effect of sympathetic histamine in mouse hearts subjected to acute ischemia. Mol Med. 2012;18(1):1–9. https://doi.org/10.2119/molmed.2011.00225.
37. Saheera S, Potnuri AG, Guha A, Palaniyandi SS, Thandavarayan RA. Histamine 2 receptors in cardiovascular biology: a friend for the heart. Drug Discov Today. 2022;27(1):234–45. https://doi.org/10.1016/j.drudis.2021.08.008.
38. He GH, Cai WK, Meng JR, Ma X, Zhang F, Lu J, et al. Relation of polymorphism of the histidine decarboxylase gene to chronic heart failure in Han Chinese. Am J Cardiol. 2015;115(11):1555–62. https://doi.org/10.1016/j.amjcard.2015.02.062.
39. Leary PJ, Tedford RJ, Bluemke DA, Bristow MR, Heckbert SR, Kawut SM, et al. Histamine H2 receptor antagonists, left ventricular morphology, and heart failure risk: the MESA Study. J Am Coll Cardiol. 2016;67(13):1544–52. https://doi.org/10.1016/j.jacc.2016.01.045.
40. Huang YH, Cai WK, Yin SJ, Wang P, Li ZR, Yang Q, et al. Histamine H2 receptor antagonist exposure was related to decreased all-cause mortality in critical ill patients with heart failure: a cohort study. Eur J Prev Cardiol. 2022;29(14):1854–65. https://doi.org/10.1093/eurjpc/zwac122.
41. Meng R, Chen LR, Zhang ML, Cai WK, Yin SJ, Fan YX, et al. Effectiveness and safety of Histamine H2 receptor antagonists: an Umbrella review of Meta-analyses. J Clin Pharmacol. 2023;63(1):7–20. https://doi.org/10.1002/jcph.2147.
42. Zhang XS, Cai WK, Wang P, Xu R, Yin SJ, Huang YH, et al. Histamine H2 receptor antagonist exhibited comparable all-cause mortality-decreasing effect as β-blockers in critically ill patients with heart failure: a cohort study. Front Pharmacol. 2023;14:1273640. https://doi.org/10.3389/fphar.2023.1273640.
43. Pan JM, Guo Y, Jiang FF, Xu R, Zhang X, Cai WK et al. Effect of Histamine H2 receptor antagonists on all-cause mortality in critically ill patients with essential hypertension: a retrospective cohort study. J Clin Pharmacol. https://doi.org/10.1002/jcph.2445

Publisher's Note