Renal dysfunction and its impact on the duration of hospitalization in patients with heart failure based on insights from echocardiographic and laboratory parameters: A retrospective study

NURIL FARID ABSHORI[1], SYANINDITA PUSPA WARDHANI[2], MUHAMMAD NAUVAL DWI AFANDI[1], FITRI AMALIA[1], ALDIVO PRADANA[2], ANISATUL HAMIDAH[3], ACHMAD ZAINUDIN ARIF[4,5], IWAL REZA AHDI[4,5], DJANGGAN SARGOWO[6] and MUHAMMAD HABIBURRAHMAN[7,8]

[1]Faculty of Medicine and Health Sciences, Maulana Malik Ibrahim Islamic State University Malang, Malang, East Java 65144, Indonesia; [2]Faculty of Medicine, Universitas Brawijaya, Malang, East Java 65145, Indonesia; [3]Faculty of Medicine, Universitas Muhammadiyah Surabaya Indonesia, Surabaya, East Java 60113, Indonesia; [4]Department of Internal Medicine, Faculty of Medicine and Health Sciences, Maulana Malik Ibrahim Islamic State University Malang, Malang, East Java 65144, Indonesia; [5]Karsa Husada Hospital, Batu, East Java 65311, Indonesia; [6]Department of Cardiology, Faculty of Medicine, Brawijaya University, Malang, East Java 65145, Indonesia; [7]Faculty of Medicine, Imperial College London, London W12 0NN, UK; [8]Faculty of Medicine, Universitas Indonesia, Jakarta, DKI Jakarta 10430, Indonesia

Received December 29, 2025; Accepted June 3, 2026

DOI: 10.3892/wasj.2026.488

Abstract. Heart failure (HF) often coexists with renal dysfunction, which complicates the management and prolongs the length of stay (LOS) of patients in hospital. It is associated with an increased risk of in-hospital complications, higher readmission rates and increased healthcare costs, placing a significant burden on both patients and healthcare systems. The present study evaluated whether renal biomarkers or echocardiographic indices can more effectively predict LOS in hospitalized patients with HF and compared the performance of predictive models derived from these features. In the present retrospective observational single-center study, 112 adult patients admitted with HF between January, 2024 and May, 2025 were analyzed. LOS was evaluated both as a continuous variable and as a binary outcome (≤3 days vs. >3 days). Group comparisons employed the Mann-Whitney U and Chi-squared tests. Predictive models were constructed using logistic regression, linear regression and random forest algorithms. Three feature sets were assessed: i) Renal biomarkers [serum urea, creatinine and estimated glomerular filtration rate (eGFR)]; ii) echocardiographic left ventricular ejection fraction (LVEF); and iii) a combined set including renal biomarkers, LVEF, age, sex and diabetes status. The median LOS was 3 days (range, 1-10 days). Patients with prolonged periods of hospitalization (>3 days) exhibited higher urea and creatinine levels, and a lower eGFR (all P≤0.003), whereas age, sex, diabetes and LVEF exhibited no significant association. Among the predictive models, the random forest model using the combined feature set achieved the highest discrimination (AUC, 0.74; 95% CI, 0.67-0.85) and regression performance (R^2=0.26). Models incorporating LVEF alone demonstrated limited predictive capacity, and the combined model did not substantially outperform renal biomarkers alone. Renal dysfunction independently predicts prolonged LOS in HF and outperforms LVEF. Integrating renal biomarkers into admission assessments can improve early risk stratification and guide clinical management by incorporating routinely available measures, such as serum creatinine and eGFR into standardized admission protocols, enabling the early identification of high-risk patients.

Correspondence to: Dr Muhammad Habiburrahman, Faculty of Medicine, Imperial College London, 4th Floor IRDB, Hammersmith Campus, Du Cane Road, London W12 0NN, UK
E-mail: h.habiburrahman23@imperial.ac.uk

Key words: heart failure, renal dysfunction, length of stay, renal biomarkers, echocardiography, predictive modelling

Introduction

Heart failure (HF) remains one of the leading causes of morbidity, mortality and hospitalization worldwide, representing a substantial and persistent burden on healthcare systems (1). The prevalence of HF continues to rise due to the aging of the population and improved survival following cardiovascular events, leading to a growing population of patients requiring chronic management (2). This increasing prevalence underscores the urgent need for effective strategies to mitigate the associated clinical, societal and economic consequences of HF (1).

In Asia, including Indonesia, the epidemiological landscape of HF poses distinct challenges, with prevalence rates remaining high across diverse populations (3). Regional studies have reported estimates approaching 900 cases per 100,000 population, highlighting the considerable disease burden within these communities (4). This impact is further compounded by region-specific etiological factors, inequities in healthcare access, and socioeconomic disparities that influence both disease progression and treatment outcomes. Consequently, the identification of modifiable determinants of adverse outcomes, particularly those contributing to prolonged hospitalization, has become an essential public health and clinical priority (3).

Among these determinants, the cardiorenal axis has emerged as a critical pathway influencing the prognosis of patients with HF. Cardiorenal syndrome, defined by bidirectional dysfunction between the heart and kidneys, represents a common and deleterious comorbidity in patients with HF (5). Contemporary registry data confirm that renal dysfunction is highly prevalent in this population and serves as an independent predictor of poor outcomes (6). Worsening renal function during hospitalization has been consistently associated with an extended length of stay (LOS) in the hospital, increased in-hospital mortality rates and higher rates of re-hospitalization (7). In addition to renal dysfunction, right ventricular impairment has been identified as a key determinant of LOS, reflecting the multifactorial nature of congestion and organ dysfunction in patients with HF (8).

Echocardiography remains the cornerstone of HF diagnosis and classification; however, reliance on left ventricular ejection fraction (LVEF) alone as a prognostic marker has been increasingly challenged (9). While LVEF provides essential information on systolic function, its capacity to predict short-term outcomes, such as LOS or re-hospitalization is limited. Additionally, there is limited comparative evidence available evaluating the relative prognostic value of renal biomarkers versus echocardiographic parameters in predicting hospital LOS among patients with HF, particularly in low to middle income settings. This limitation highlights the need to integrate additional physiological markers, particularly renal biomarkers, into risk prediction frameworks.

To address this gap, the present study aimed to directly compare the prognostic performance of routinely measured renal biomarkers with standard echocardiographic parameters in predicting hospital LOS among patients admitted with HF. Furthermore, the present study integrated conventional statistical modelling with machine learning approaches, thereby enhancing predictive accuracy and early risk stratification. Such an approach could inform individualized management, optimize resource allocation, and support data-driven strategies to improve outcomes for patients with HF (8,10).

Patients and methods

Study population. The present retrospective study cohort comprised adult patients admitted to Karsa Husada Hospital, Batu in Indonesia between January 1, 2024 and May 31, 2025 with a primary diagnosis of HF. Data were collected from June to July 2025 after ethics approval for the study had been obtained. Karsa Husada Hospital serves as a teaching hospital affiliated with the Faculty of Medicine and Health Sciences, Maulana Malik Ibrahim Islamic State University Malang.

A total of 112 patients met the eligibility criteria and were included in the final analysis. Patients were included using a consecutive sampling approach, whereby all eligible patients meeting the inclusion criteria during the study period were enrolled sequentially to minimize bias. As the present study was a retrospective study utilizing all available data within the study period, no formal sample size calculation was performed. The inclusion criteria were: i) An age ≥18 years; ii) admission with HF as the primary diagnosis; and iii) the availability of laboratory measurements for serum urea, serum creatinine and estimated glomerular filtration rate (eGFR), along with the echocardiographic assessment of LVEF. Patients were excluded if essential clinical or laboratory data were not available, if they were discharged or transferred within 24 h of admission, or if HF was not the primary reason for hospitalization. The workflow of patient selection is depicted in (Fig. 1).

The present study was approved by the Health Research Ethics Committee of Faculty of Medicine, Maulana Malik Ibrahim Islamic State University Malang, Indonesia (Approval no. 60/01/EC/KEPK-FKIK/10/2024). Only de-identified medical records were analyzed, with a waiver of written informed consent granted due to the retrospective, anonymized data use, and non-interventional design (Letter no. 009/KEPK.FKIK/KP/XI/2025).

Variables and outcome definition. Demographic variables included age, sex and diabetes mellitus status (coded as 1 for presence, 0 for absence). Laboratory parameters consisted of serum urea (mg/dl), serum creatinine (mg/dl) and eGFR (ml/min/1.73 m^2), calculated using the Modification of Diet in Renal Disease (MDRD) equation. Echocardiographic assessment included LVEF (%), measured using the biplane Simpson method in accordance with American Society of Echocardiography guidelines (11). LVEF was selected as the primary echocardiographic parameter as it is the most routinely available and standardized measure across all patients in clinical practice. Other echocardiographic parameters, such as right ventricular function or diastolic indices, were not included due to their inconsistent availability in the dataset. The primary outcome was hospital LOS, defined as the number of days from admission to discharge. LOS was analyzed both as a continuous variable and as a categorical variable dichotomized at the median value: ≤3 days vs. >3 days. Patients with a hospital LOS >3 days were classified as having a prolonged period of hospitalization for subsequent descriptive and inferential analyses.

Statistical analysis. Continuous variables are reported as the mean ± standard deviation (SD) or median with interquartile range (IQR), depending on the underlying distribution, while categorical variables are presented as frequencies and percentages. Normality was assessed using the Shapiro-Wilk test. Comparisons between patients with a short (≤3 days) and prolonged (>3 days) LOS were performed using the Mann-Whitney U test for continuous variables and the Chi-squared test for categorical variables. A two-tailed P-value <0.05 was considered to indicate a statistically significant difference.

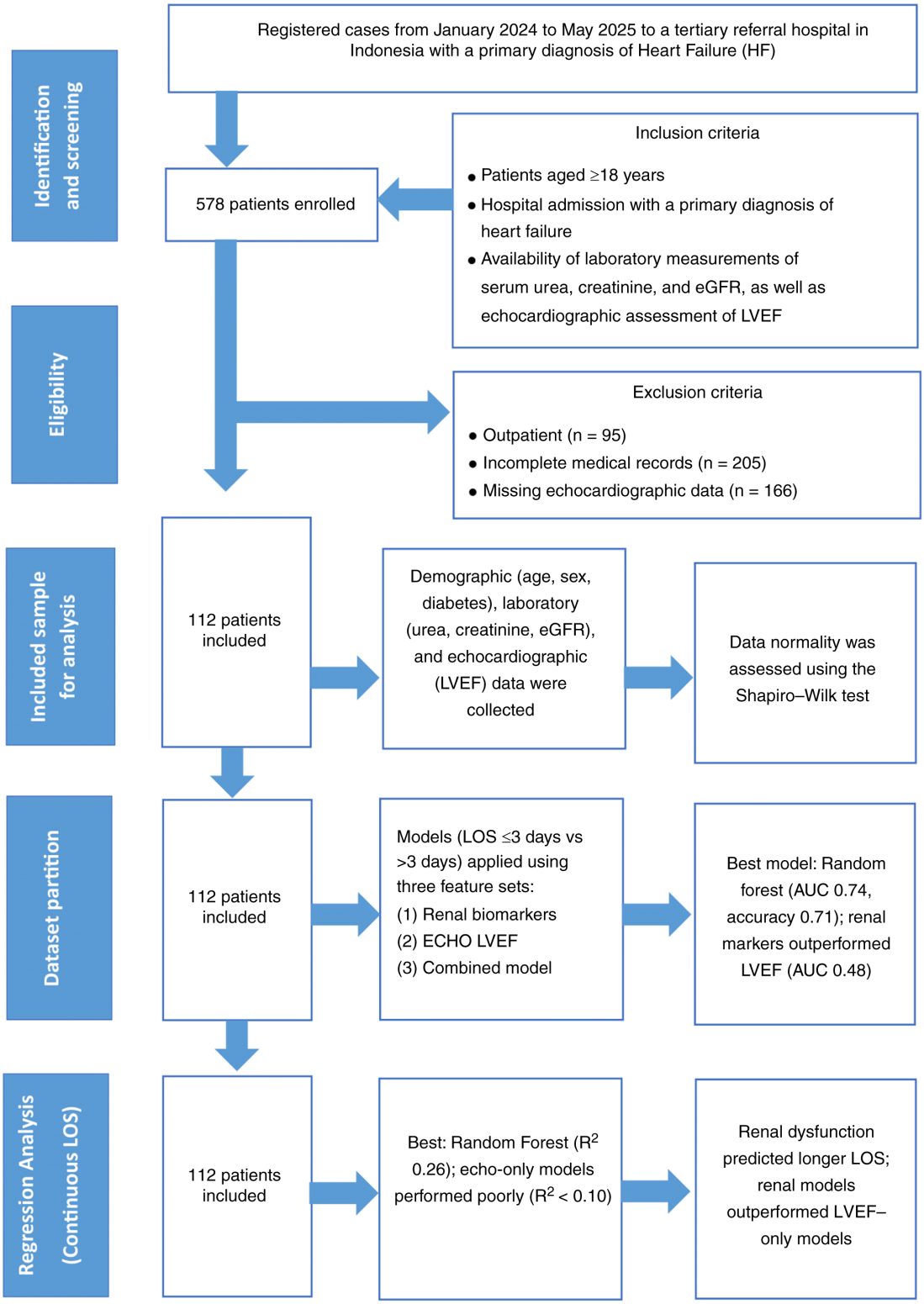

Figure 1. Flowchart of patient selection and analytic workflow. Flow diagram illustrating the identification, screening, and inclusion of patients hospitalized with heart failure. Of 280 records initially assessed, 112 patients met the inclusion criteria based on the availability of renal and echocardiographic data and were retained for the final analysis. eGFR, estimated glomerular filtration rate; LVEF, left ventricular ejection fraction.

For predictive modeling, three feature sets were evaluated: i) Renal biomarkers alone (serum urea, creatinine and eGFR); ii) echocardiographic variable alone (LVEF); and iii) a combined feature set incorporating renal biomarkers, LVEF, age, sex and diabetes status. Missing values were imputed using the median of each variable. Prior to modeling, the class distribution of the outcome variable (LOS ≤3 days vs. >3 days) was assessed. As the proportion of the two groups was relatively balanced, no resampling techniques, such as oversampling or undersampling were applied. Additionally, synthetic data generation methods were not used to avoid potential overfitting given the relatively small sample size. The dataset was randomly split

Table I. Baseline characteristics of patients with heart failure stratified by hospital LOS.

Characteristic	≤3 days (mean ± SD)	>3 days (mean ± SD)	P-value
Age (years)	61.14±10.91	63.43±10.94	0.28
Sex, n (%)			
Male	38 (65.5%)	37 (68.5%)	0.74
Female	20 (34.5%)	17 (31.5%)	
Diabetes, n (%)			
Yes	13 (22.4%)	11 (20.4%)	0.80
No	45 (77.6%)	43 (79.6%)	
Urea (mg/dl)	30.59±18.93	53.91±39.27	0.003
Creatinine (mg/dl)	0.88±0.52	1.39±1.16	0.001
eGFR (ml/min/1.73 m^2)	100.89±44.80	69.89±33.63	0.001
LVEF (%)	46.09±15.26	44.26±17.04	0.49

Data are presented as the mean ± SD for continuous variables and number (percentage) for categorical variables. Patients were categorized into two groups: LOS ≤3 days (n=58) and LOS >3 days (n=54). LOS, length of stay; SD, standard deviation; eGFR, estimated glomerular filtration rate; LVEF, left ventricular ejection fraction.

into training (70%) and testing (30%) subsets. To enhance model robustness and reduce variability associated with a single data split, k-fold cross-validation (k=5) was performed on the training set. Logistic and linear regression models were used as baseline approaches due to their interpretability, while ensemble methods, including random forest and gradient boosting, were selected to capture non-linear associations and interactions among variables. Hyperparameter optimization for random forest and gradient boosting models was conducted using grid search within the cross-validation framework to identify optimal parameter combinations. For binary classification of prolonged LOS, logistic regression (with feature standardization) and random forest classifiers were applied. Model performance was quantified using accuracy, area under the receiver operating characteristic curve (AUC), precision, recall, and F1-score.

Continuous LOS prediction was performed using linear regression and random forest regression, with performance metrics including mean absolute error (MAE), mean squared error (MSE) and the coefficient of determination (R^2). Receiver operating characteristic (ROC) curves were generated to visualize classifier performance, and feature importance values were derived from random forest models. For regression analyses, scatter plots comparing predicted vs. observed LOS were constructed to evaluate model fit. All statistical analyses and machine learning workflows were implemented in Python (pandas, scikit-learn, scipy, matplotlib). To improve reproducibility, all machine learning models were implemented with a fixed random_state. Random forest models were constructed using 100 trees (n_estimators=100) with default depth settings. Continuous variables were standardized using z-score normalization prior to logistic regression modelling. Model performance was evaluated on a held-out test set (30% of the data), while the remaining 70% was used for training. Overfitting was assessed by comparing model performance between training and testing datasets. Given the relatively small sample size, a parsimonious modelling approach was adopted to balance predictive performance and generalizability.

Results

Group comparisons. Baseline demographic and laboratory characteristics stratified by hospital LOS are summarized in (Table I). Patients were evenly distributed between those discharged within 3 days and those hospitalized for longer durations. Age (61.1±10.9 vs. 63.4±10.9 years; P=0.28), sex distribution (65.5 vs. 68.5% male; P=0.74) and the prevalence of diabetes (22.4 vs. 20.4%; P=0.80) were comparable between the groups. By contrast, significant differences were observed in renal function parameters. Patients with prolonged periods of hospitalization exhibited markedly higher blood urea concentrations (53.9±39.3 vs. 30.6±18.9 mg/dl; P=0.003) and serum creatinine levels (1.39±1.16 vs. 0.88±0.52 mg/dl; P=0.001), indicating greater renal dysfunction. LVEF did not differ significantly between the groups (44.3±17.0 vs. 46.1±15.3%; P=0.49), suggesting that renal dysfunction, rather than systolic cardiac performance, was more strongly associated with prolonged hospitalization.

Classification performance. The data for predictive model performance on the held-out test sets are presented in Table II. Using only renal biomarkers, the logistic regression demonstrated modest discrimination (accuracy, 0.56; AUC, 0.68; 95% CI, 0.58-0.78) with limited recall (0.38) and F1 score (0.44). By contrast, tree-based ensemble approaches demonstrated superior discrimination: The random forest classifier improved accuracy to 0.71 with recall (0.63) and F1 score (0.67) alongside good discrimination (AUC, 0.74; 95% CI, 0.65-0.83) outperforming both logistic regression and gradient boosting. Gradient boosting achieved a slightly higher AUC of 0.76 (95% CI, 0.67-0.85) , although with lower recall (0.56) and F1 score (0.62). Models trained exclusively on LVEF features exhibited a poor predictive ability (AUC, 0.48;

Table II. Performance of predictive models for prolonged hospital LOS in patients with heart failure using different feature sets.

Feature set	Model	Accuracy	AUC (95% CI)	Precision	Recall	F1
Renal	Logistic regression	0.56	0.68 (0.58-0.78)	0.55	0.38	0.44
	Random forest	0.71	0.74 (0.65-0.83)	0.71	0.63	0.67
	Gradient boosting	0.68	0.76 (0.67-0.85)	0.69	0.56	0.62
Echocardiography (LVEF)	Logistic regression	0.50	0.48 (0.38-0.58)	0.44	0.25	0.32
	Random forest	0.53	0.52 (0.42-0.62)	0.50	0.81	0.62
Combined (renal + LVEF + age + sex + diabetes)	Logistic regression	0.56	0.64 (0.54-0.74)	0.54	0.44	0.48
	Random forest	0.71	0.68 (0.58-0.78)	0.71	0.63	0.67

Models were evaluated using accuracy, AUC, precision, recall and F1 score. LOS, length of stay; AUC, area under the receiver operating characteristic curve; CI, confidence interval; eGFR, estimated glomerular filtration rate; LVEF, left ventricular ejection fraction.

Table III. Regression model performance for predicting hospital LOS in patients with heart failure.

Feature set	Model	MAE (days)	MSE (days^2)	R^2 (95% CI)
Renal	Linear regression	1.16	3.28	0.10 (0.00-0.21)
	Random forest regressor	1.30	3.50	0.04 (0.00-0.14)
Echocardiography (LVEF)	Linear regression	1.26	3.51	0.04 (0.00-0.14)
	Random forest regressor	1.27	3.34	0.09 (0.00-0.20)
Combined (renal + LVEF + age + sex + diabetes)	Linear regression	1.24	3.32	0.09 (0.00-0.20)
	Random forest regressor	1.18	2.70	0.26 (0.11-0.41)

Predictive accuracy was evaluated using MAE, MSE and R^2. LOS, length of stay; MAE, mean absolute error; MSE, mean squared error; R^2, coefficient of determination; CI, confidence interval; eGFR, estimated glomerular filtration rate; LVEF, left ventricular ejection fraction.

95% CI, 0.38-0.58) and low classification metrics (accuracy, 0.50; F1 score, 0.32). Although random forest improved recall substantially (0.81), overall discrimination remained limited, indicating LVEF alone provides insufficient prognostic information for LOS. Incorporating additional clinical covariates, including LVEF, age, sex and diabetes, did not substantially enhance model performance. Logistic regression exhibited minimal improvement (accuracy, 0.56; AUC, 0.64; 95% CI, 0.54-0.74) while random forest delivered the most robust and consistent performance (accuracy, 0.71; AUC, 0.68; 95% CI, 0.58-0.78; F1 score, 0.67), underscoring its superior ability to leverage multimodal clinical information. From a clinical perspective, an AUC of 0.74 indicates acceptable discriminative ability, indicating that the model can correctly distinguish between patients with prolonged and non-prolonged LOS ~74% of the time. This level of performance suggests potential utility for early risk stratification in hospital settings, although it may still be insufficient as a standalone decision-making tool and should be interpreted in conjunction with clinical judgment and other patient-specific factors.

Regression performance. The results of regression analysis are summarized in Table III. Using the renal feature set, the linear regression model predicted LOS with an MAE of 1.16 days and explained ~10% of the variance (R^2=0.10). The random forest model did not yield substantial improvement, demonstrating a slightly higher MAE of 1.30 days and a lower R^2 of 0.04. Models based solely on echocardiographic (LVEF) features performed poorly, with both linear and random forest regressions explaining minimal variance (R^2=0.04 and 0.09, respectively) and achieving comparable MAE values (1.26-1.27 days). By contrast, incorporating additional clinical variables in the combined feature set (renal + LVEF + age + sex + diabetes) modestly enhanced predictive performance. Linear regression achieved an MAE of 1.24 days (R^2=0.09; 95% CI, 0.00-0.20), whereas the random forest model improved to an MAE of 1.18 days and R^2 of 0.26 (95% CI, 0.11-0.41). Although the overall variance explained remained modest, the random forest model using combined features demonstrated the best balance of predictive accuracy and generalizability across all tested configurations.

The comparative performance of predictive models and their graphical summaries are presented in (Fig. 2). The ROC curve for the logistic regression model trained on the combined clinical, renal, and echocardiographic feature set (Fig. 2IA) yielded an AUC of 0.64, reflecting modest discriminative ability. The curve, which plots the true positive rate against the false positive rate across varying thresholds, lies only slightly above the diagonal reference line denoting chance-level classification. This indicates that the model captured limited but discernible signal within the data, consistent with partial discrimination between patients with prolonged versus shorter

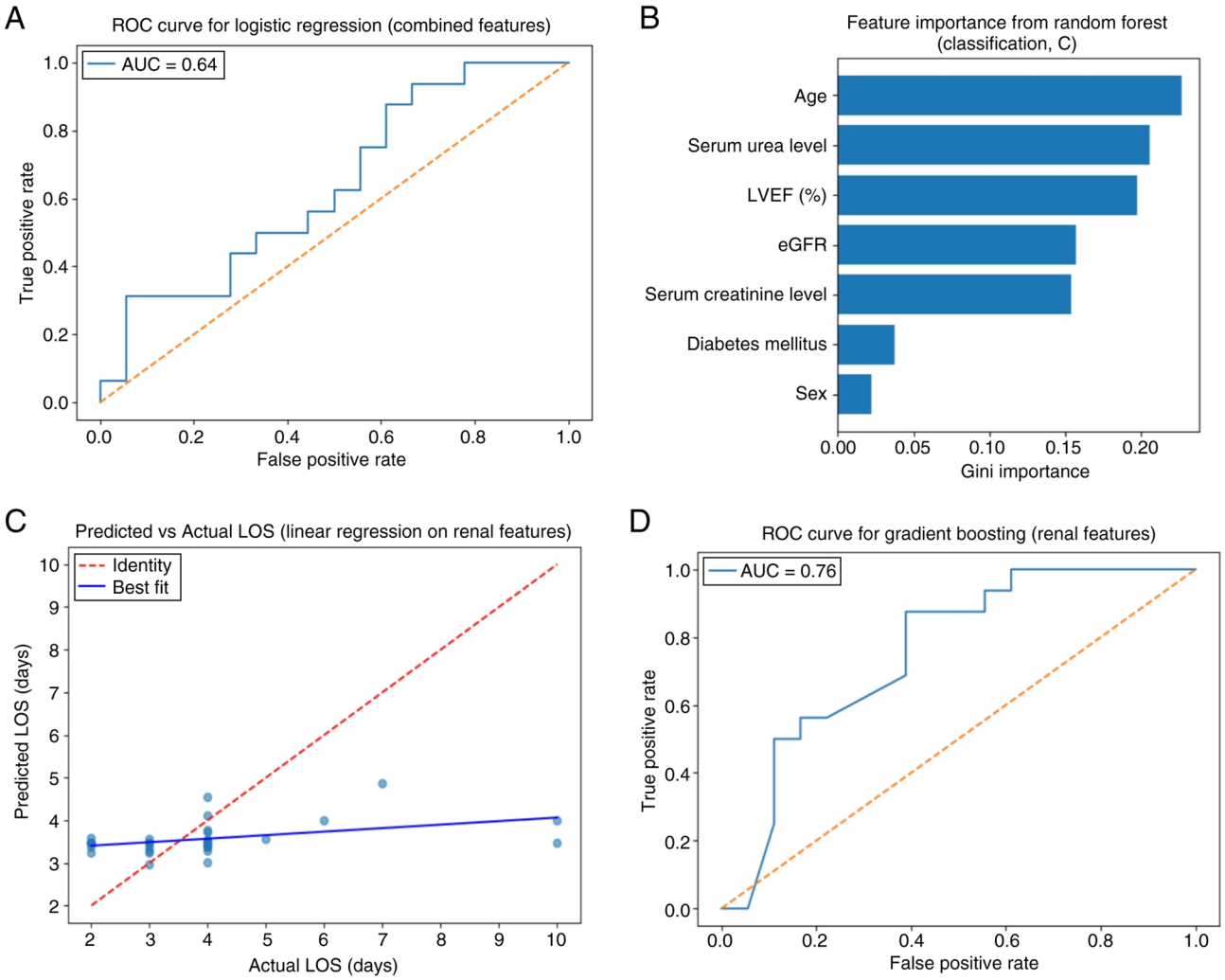

Figure 2. Model performance and feature importance across clinical, renal, and echocardiographic predictors. Comparative performance of predictive models using different feature sets. (A) ROC curve for the logistic regression model combining clinical, renal, and echocardiographic variables, showing modest discrimination (AUC, 0.64). This suggests limited ability of conventional linear models to reliably distinguish patients at risk of prolonged hospitalization. (B) Feature importance from the random forest model, identifying age and serum urea as dominant predictors of prolonged hospitalization. This highlights the greater prognostic relevance of renal biomarkers compared with echocardiographic parameters. (C) Association between predicted and observed LOS derived from the linear regression model using renal biomarkers, demonstrating weak overall correlation, indicates that renal function alone cannot fully explain variability in LOS. (D) ROC curve for the gradient boosting model trained on renal biomarkers, showing the strongest discrimination (AUC, 0.76), suggests that advanced models capturing nonlinear relationships may improve early identification of high-risk patients. ROC, receiver operating characteristic; LOS, length of stay.

hospital stays. This implies traditional linear models using routinely collected variables may have limited utility as standalone tools for early risk stratification at admission.

Feature importance derived from the random forest classifier trained on the same combined feature set is shown in Fig. 2B. Age emerged as the most influential predictor (Gini importance >0.2), followed by serum urea concentration and LVEF. Other renal markers, including eGFR and creatinine, also contributed meaningfully, albeit with lower relative weights. These results indicate that renal biomarkers and age primarily drive model predictions, while echocardiographic indices provide comparatively limited incremental value, reinforces the clinical relevance of renal function assessment at admission

The association between the predicted and observed LOS for the linear regression model based on renal biomarkers is displayed in Fig. 2C. The red dashed line represents the line of identity, whereas the blue line denotes the model's best-fit regression line. The dispersion of points around the identity line indicates that the model captured the general direction of the association but with weak correlation, underscoring that renal parameters alone explain only a fraction of the variability in LOS among patients with HF, hence is insufficient to fully explain variability in LOS, highlighting the multifactorial nature of HF outcomes.

By contrast, the gradient boosting classifier trained exclusively on renal biomarkers demonstrated markedly enhanced discrimination (AUC, 0.76) (Fig. 2D). The ROC curve rises substantially above the diagonal reference line, indicating improved sensitivity and specificity in identifying patients at risk of prolonged hospitalization. This suggests that more advanced models capable of capturing nonlinear relationships may provide clinically meaningful improvements in identifying high-risk patients early during hospitalization.

Compared with logistic regression and random forest models, Gradient boosting more effectively captured nonlinear interactions among renal variables, resulting in superior predictive accuracy. Collectively, these findings highlight the prognostic relevance of renal dysfunction and support the utility of ensemble-based learning approaches for clinical risk stratification in HF.

Discussion

The present study demonstrated that renal dysfunction is a key determinant of prolonged periods of hospitalization in patients with HF, with renal biomarkers consistently outperforming echocardiographic parameters in predicting LOS. These findings are consistent with previous evidence indicating that worsening renal function (WRF) is common in acute HF and correlates with longer LOS. Emmens *et al* (12) reported that LOS increased from 2.2 to 4.2 days in patients with WRF. Similarly, a cross-sectional study found that 33.7% of patients with HF with reduced ejection fraction (HFrEF) developed WRF, which was associated with longer hospitalization and higher mortality (13). Collectively, these data highlight renal function as a critical determinant of hospital stay duration.

The association between renal dysfunction and prolonged hospital LOS involves the complex pathophysiology of the cardiorenal interaction. In HF, reduced cardiac output and elevated venous congestion lead to decreased renal perfusion and increased renal interstitial pressure, impairing glomerular filtration. The activation of the neurohormonal system, including the renin-angiotensin-aldosterone system and sympathetic nervous system, further exacerbates sodium and water retention, promoting volume overload and diuretic resistance. Additionally, systemic inflammation and endothelial dysfunction contribute to microvascular impairment within the kidneys. These mechanisms create a self-perpetuating cycle in which cardiac dysfunction worsens renal function and vice versa, ultimately prolonging decongestion time and LOS (5,14).

In the present study, predictive modeling further demonstrated that renal biomarkers provided stronger discriminatory power than echocardiographic parameters such as LVEF. By contrast, LVEF demonstrated a limited predictive value for LOS. These findings are in agreement with those of previous research demonstrating that WRF combined with persistent congestion contributes to longer hospitalizations and poorer outcomes (15). While LVEF reflects left ventricular systolic function, it does not adequately capture the hemodynamic complexity of HF, particularly congestion status, right ventricular dysfunction, or filling pressures, which are more directly related to hospitalization duration (16,17). AlShammari *et al* (18) also reported that 27% of patients with HF experienced WRF, which was associated with prolonged periods of hospitalization and increased complications. Taken together, these results suggest that renal function on admission reflects acute pathophysiological changes more accurately than LVEF alone.

Herein, the addition of demographic or clinical variables such as age, sex and diabetes did not meaningfully improve model performance. LVEF, despite its traditional role in HF stratification, exhibited limited predictive value. Schmitt *et al* (16) similarly observed that LVEF categories, reduced ejection fraction, mildly reduced ejection fraction, or preserved ejection fraction, did not consistently predict LOS or short-term mortality. Tarekegn *et al* (19) also found that demographic factors had minimal influence, whereas clinical features including edema, hepatomegaly and pleural effusion were stronger predictors. These findings reinforce that structural parameters alone may not adequately capture hospitalization risk.

The application of machine learning models in the present study further highlights the prognostic value of renal biomarkers. Tree-based ensemble methods, particularly random forest and gradient boosting, demonstrated superior performance compared to logistic regression, likely due to their ability to capture non-linear association and complex feature interaction. These findings are in line with those of recent studies which have demonstrated improved predictive performance when incorporating multidimensional clinical data, particularly laboratory parameters, in machine learning (20,21). To the best of our knowledge, the present study is among the first to directly compare the predictive performance of renal biomarkers and echocardiographic parameters using both conventional statistical methods and machine learning approaches in the Indonesian population. This is particularly important given regional differences in patient characteristics, healthcare access and disease prevention, which may limit the generalizability of findings from high-income settings.

Feature importance analysis confirmed that renal biomarkers, particularly urea, creatinine and eGFR, along with age, were dominant predictors of LOS, while LVEF contributed minimally. This underscores the importance of assessing renal function at admission for early risk stratification. This finding was consistent with the findings of the study by Ignatavičiūtė *et al* (22), in which an eGFR ≤50 ml/min/1.73 m^2 independently predicted LOS >9 days in acute HF. Elmaghawry *et al* (23) also demonstrated that the WATCH-DM score, which includes renal function and age, accurately predicted longer hospitalization among patients with HF and diabetes (AUC 0.93). Integrating renal parameters into risk assessment therefore provides prognostic value beyond conventional measures.

Although renal biomarker-based models exhibited optimal performance, their predictive power remained modest (R^2<0.30), suggesting that other factors contribute to LOS. These may include hemodynamic instability, congestion severity, right ventricular dysfunction and comorbidities. Ignatavičiūtė *et al* (22) identified elevated N-terminal pro-B-type natriuretic peptide (NT-proBNP), low systolic pressure and significant tricuspid regurgitation as independent predictors for a prolonged LOS. Similarly, Tarekegn *et al* (19) found that edema, hepatomegaly and pleural effusion were stronger determinants than demographic characteristics. Thus, while renal dysfunction plays a key role, it forms part of a multifactorial pathway influencing hospitalization duration.

Clinically, these findings emphasize the importance of early renal function assessment as part of a comprehensive risk stratification of patients with HF. Routine measurements of urea, creatinine and eGFR may help identify patients who are at a higher risk for prolonged hospitalization and guide

timely interventions. Future multicenter prospective studies are required to evaluate whether interventions targeting renal dysfunction, such as optimized diuretic therapy, NT-proBNP-guided management, or multiparametric risk scores such as WATCH-DM, can shorten LOS and improve outcomes (23). Including socioeconomic factors, treatment adherence and advanced echocardiographic parameters, such as right ventricular function and filling pressures may further enhance predictive accuracy. Integrating renal biomarkers with clinical and social determinants could support a more holistic approach to modern HF management (18).

The present study has limitations which should be mentioned. The sample size was modest, derived from a single center, and no formal sample size calculation was performed which may limit generalizability and may introduce potential selection bias, as the study population may not fully represent the broader spectrum of patients with HF in different healthcare settings. Furthermore, the relatively small sample size may have limited the robustness and generalizability of the machine learning models. Its retrospective design precludes causal inference, and unmeasured confounders could have influenced the results. Echocardiographic variables beyond LVEF, such as right ventricular function or filling pressures, were unavailable and may have restricted the ability to fully assess the contribution of cardiac structure and hemodynamics to LOS. In addition, the absence of external validation limits the ability to confirm the generalizability of the predictive models across different populations. Despite imputation, missing data could have introduced bias. In addition, although machine learning models were applied, the relatively small sample size limited the use of more complex optimization strategies, and model performance should be interpreted with caution. External validation in larger multicenter cohorts is warranted.

In conclusion, in the present study, renal dysfunction was found to be significantly associated with prolonged periods of hospitalization among patients with HF. Predictive modeling demonstrated that renal biomarkers outperformed echocardiographic indices such as LVEF in predicting LOS. Among the evaluated models, Random forest using renal features achieved the strongest overall performance (AUC, 0.74; R^2=0.26). Collectively, these findings highlight the importance of renal function markers in refining risk stratification for LOS and should be prioritized during hospital admission for patients with HF to enable early risk stratification and guide targeted clinical management. While these findings highlight the clinical utility of renal biomarkers, further multicenter validation is required before widespread implementation.

Acknowledgements

The authors wish to express their sincere gratitude to the healthcare workers at Karsa Husada General Hospital and the faculty members of the Faculty of Medicine and Health Sciences, Maulana Malik Ibrahim Islamic State University Malang, for their valuable assistance and support throughout the research.

Funding

No funding was received.

Availability of data and materials

The data that support the findings of this study are not publicly available due to confidentiality concerns, as they contain patient identities and hospital medical records. Access to the data is restricted to the corresponding author and the institutional medical data committee. Data may be made available to other parties only upon formal request and with written approval from the hospital ethics or data committee. Requests should be directed to the corresponding author via email.

Authors' contributions

NFA led the conceptualization, data curation, formal analysis, investigation, methodology development, software implementation, validation, visualization and drafting of the manuscript. AZA contributed to formal analysis, methodology, project administration and resources. IRA participated in the conceptualization of the study, methodology development, project administration, data interpretation, resource provision and in the revision of the manuscript. AH contributed to data interpretation, methodology development and critical revision of the manuscript. DS contributed to the conceptualization of the study, data investigation, validation and in the critical revision of the manuscript. MNDA, FA and AP assisted with data curation, visualization and the drafting of the initial manuscript. SPW contributed to data investigation and study supervision. MH oversaw the formal analysis, methodology, funding acquisition, supervision, validation, and the critical reviewing and editing of the manuscript. NFA and MH confirm the authenticity of all the raw data. All authors have read and approved the final version of the manuscript.

Ethics approval and consent to participate

The present study was approved by the Health Research Ethics Committee of Faculty of Medicine, Maulana Malik Ibrahim Islamic State University Malang, Indonesia (Approval no. 60/01/EC/KEPK-FKIK/10/2024). Only de-identified medical records were analyzed, with a waiver of written informed consent granted due to the retrospective, anonymized data use, and non-interventional design (Letter No. 009/KEPK.FKIK/KP/XI/2025).

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Savarese G, Becher PM, Lund LH, Seferovic P, Rosano GMC and Coats AJS: Global burden of heart failure: A comprehensive and updated review of epidemiology. Cardiovasc Res 118: 3272-3287, 2022.
2. Vaduganathan M, Claggett BL, Inciardi RM, Fonarow GC, McMurray JJV and Solomon SD: Estimating the benefits of combination medical therapy in heart failure with mildly reduced and preserved ejection fraction. Circulation 145: 1741-1743, 2022.

3. Feng J, Zhang Y and Zhang J: Epidemiology and burden of heart failure in asia. JACC Asia 4: 249-264, 2024.
4. Tromp J and Teng THK: Regional differences in the epidemiology of heart failure. Korean Circ J 54: 591, 2024.
5. Rangaswami J, Bhalla V, Blair JEA, Chang TI, Costa S, Lentine KL, Lerma EV, Mezue K, Molitch M, Mullens W, *et al*: Cardiorenal syndrome: Classification, pathophysiology, diagnosis, and treatment strategies: A scientific statement from the american heart association. Circulation 139: E840-E878, 2019.
6. Tang WHW, Bakitas MA, Cheng XS, Fang JS, Fedson SE, Fiedler AG, Martens P, McCallum WI, Ogunniyi MO, Rangaswami J, *et al*: Evaluation and management of kidney dysfunction in advanced heart failure: A scientific statement from the American heart association. Circulation 150: e280-e295, 2024.
7. Janse RJ, Milders J, Rotmans JI, Caskey FJ, Evans M, Torino C, Szymczak M, Drechsler C, Wanner C, Pippias M, *et al*: Predicting hospitalization and related outcomes in advanced chronic kidney disease: A systematic review, external validation, and development study. Kidney Med 7: 101016, 2025.
8. Palazzuoli A, Dini FL, Pugliese NR, Ruocco G, Severino P, Vizza D, Carbonara R, Passantino A and Carluccio E: Right ventricular–pulmonary artery uncoupling and strain in acute heart failure. J Am Heart Association 14: e039858, 2025.
9. Adamczak DM, Oduah MT, Kiebalo T, Nartowicz S, Bęben M, Pochylski M, Ciepłucha A, Gwizdała A, Lesiak M and Straburzyńska-Migaj E: Heart failure with preserved ejection fraction-a concise review. Curr Cardiol Rep 22: 82, 2020.
10. van der Stam JA, Bouwmeester S, van Loon SLM, vanRiel NAW, Dekker LR, Boer AK, Houthuizen P and Scharnhorst V: Prognostic value of combined biomarkers in patients with heart failure: The heartmarker score. Ann Lab Med 43: 253-262, 2023.
11. Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, Flachskampf FA, Foster E, Goldstein SA, Kuznetsova T, *et al*: Recommendations for cardiac chamber quantification by echocardiography in adults: An update from the American society of echocardiography and the european association of cardiovascular imaging. J Am Soc Echocardiogr 28: 1-39.e14, 2015.
12. Emmens JE, ter Maaten JM, Matsue Y, Figarska SM, Sama IE, Cotter G, Cleland JGF, Davison BA, Felker GM, Givertz MM, *et al*: Worsening renal function in acute heart failure in the context of diuretic response. Eur J Heart Fail 24: 365-374, 2022.
13. Asbagh AG, Sadeghi MT, Ramazanilar P and Nateghian H: Worsening renal function in hospitalized patients with systolic heart failure: Prevalence and risk factors. ESC Heart Fail 10: 2837-2842, 2023.
14. Damman K and Testani J: Cardiorenal interactions in heart failure: Insights from recent therapeutic advances. Cardiovasc Res 120: 1372-1384, 2024.
15. Cooper L, DeVore A, Cowger J, Pinney S, Baran D, DeWald TA, Burt T, Pietzsch TB, Walton A, Aaronson K and Shah P: Patients hospitalized with acute heart failure, worsening renal function, and persistent congestion are at high risk for adverse outcomes despite current medical therapy. Clin Cardiol 46: 1163-1172, 2023.
16. Schmitt A, Behnes M, Weidner K, Abumayyaleh M, Reinhardt M, Abel N, Lau F, Forner J, Ayoub M, Mashayekhi K, *et al*: Prognostic impact of prior LVEF in patients with heart failure with mildly reduced ejection fraction. Clin Res Cardiol 114: 570-588, 2025.
17. Epelde F: Heterogeneity in heart failure with preserved ejection fraction: A systematic review of phenotypic classifications and clinical implications. J Clin Med 14: 4820, 2025.
18. AlShammari A, Magdits M, Majdalani R, Devarajan S, Hughes A, McCann L, Ionescu N and Raiszadeh F: Heart failure and worsening renal function: Prevalence, risk factors, and impact on hospital readmissions in an urban minority population. J Clin Med 14, 2025.
19. Tarekegn GY, Wondm SA, Tamene FB, Anberbr SS, Moges TA, Dagnew SB, Zeleke TK and Dagnew FN: Determinant factors of prolonged hospitalization in acute heart failure patients at jimma medical center, southwest ethiopia. Sci Rep 15: 11670, 2025.
20. Xie Y, Zhang L, Sun W, Zhu Y, Zhang Z, Chen L, Xie M and Zhang L: Artificial intelligence in diagnosis of heart failure. J Am Heart Association 14: e039511, 2025.
21. Gautam N, Mueller J, Alqaisi O, Gandhi T, Malkawi A, Tarun T, Alturkmani HJ, Zulqarnain MA, Pontone G and Al'Aref SJ: Machine learning in cardiovascular risk prediction and precision preventive approaches. Curr Atheroscler Rep 25: 1069-1081, 2023.
22. Ignatavičiūtė E, Žaliaduonytė D and Zabiela V: Prognostic factors for prolonged in-hospital stay in patients with heart failure. Medicina (Kaunas) 59: 930, 2023.
23. Elmaghawry ML, Mabrouk RRA, Shereef SA and El-Shetry SM: Length of hospital stay in acute heart failure with type 2 diabetes mellitus: Forecasting Using WATCH-DM Score. Heart Failure Journal of India 2: 90-96, 2024.