Macrothink Institute™

Journal of Biology and Life Science
ISSN 2157-6076
2026, Vol. 17, No. 2

A Machine Learning Approach for Predicting Failed First-Attempt Radial Artery Puncture Patients with Heart Failure

Huan Wang

Pulmonary and Critical Care Medicine, Dujiangyan Hospital of traditional Chinese Medicine

Chengdu, Sichuan, China, 611830

Tel: 86-182-8001-1503 E-mail: 2370279649@qq.com

Jiangbo Zhang (Corresponding author)

School of Science, Southwest Petroleum University

No. 8 Xindu avenue, Chengdu, Sichuan, 610500, China

Tel: 86-151-8438-7612 E-mail: eridalin@163.com

Received: December 14, 2025 Accepted: January 26, 2026 Published: January 28, 2026

doi:10.5296/jbls.v17i2.23426 URL: https://doi.org/10.5296/jbls.v17i2.23426

Abstract

Objective: To identify risk factors for first-attempt failure of radial arterial puncture in heart failure patients and to develop and compare predictive models using logistic regression with advanced feature engineering and ensemble learning approaches.

Method: A retrospective study was conducted involving 789 heart failure patients who underwent radial arterial puncture. Patients were divided into a training set (80%) and a test set (20%) using a stratified hold-out method. A logistic regression framework incorporating feature engineering (interaction terms, transformations, composite scores, PCA) and a dual-stage variable selection strategy (LASSO followed by stepwise selection with $p<0.1$ threshold) was employed. Four ensemble models were developed in parallel. Model performance was evaluated using area under the receiver operating characteristic curve (AUC), sensitivity, specificity, and calibration metrics through a multi-repeat validation framework.

Results: The final logistic regression model identified eleven variables, with edema degree (β

= -0.0959, OR = 0.9085, p = 0.0160) and the interaction between ejection fraction and log-transformed BNP ($\beta = 0.1507$, OR = 1.1627, $p = 0.0371$) reaching statistical significance. The model demonstrated fair discriminative ability with an average test AUC of 0.693 (±0.033), high specificity (95.59% ± 2.55%), but lower sensitivity (31.88% ± 13.98%). The ensemble learning models showed weaker discriminative performance but exhibited potential overfitting on the training set. The optimal probability cutoff for the logistic model was 0.658.

Conclusion: Both modeling approaches developed effective prediction tools. The logistic regression model provided clinically interpretable risk factors, while the ensemble learning models achieved higher discriminatory power at the cost of interpretability. These models can assist in pre-procedural identification of high-risk patients, allowing for tailored strategies to improve first-attempt success rates and reduce patient discomfort.

Keywords: heart failure, edema, BNP, feature engineering, logistic regression, ensemble learning

1. Introduction

1.1 The Problem Introduction

The increasing incidence and mortality of heart failure (HF)—a complex syndrome arising from structural, functional, or myocardial damage—have emerged as a major public health concern, driven by demographic aging and the rising burden of chronic diseases (Malik et al., 2021). Despite guideline-directed medical therapy, heart failure remains associated with a high hospitalization rate (Wintrich et al., 2020). For these patients, radial artery puncture is a routine procedure required for arterial blood gas analysis and invasive hemodynamic monitoring (Sze et al., 2021). As an invasive technique, however, unsuccessful initial attempts necessitate repeated punctures. This can lead to multiple adverse outcomes, including increased procedural difficulty, patient pain and anxiety, local bleeding, hematoma, vasospasm, diminished treatment compliance, and increased resource consumption. Consequently, maximizing the first-attempt success rate is crucial.

Prior research has analyzed factors influencing radial artery catheterization success in settings like coronary angiography (Hu, 2018; Lu et al., 2018). Building upon this foundation, this study aimed to develop and validate a prediction model for first-attempt radial artery puncture failure specifically in HF patients by synthesizing known risk factors with the unique clinical attributes of this population.

1.2 Importance of the Problem

HF constitutes a major global health challenge (Malik et al., 2021). Radial artery puncture is a key procedure for these patients, enabling critical assessment and guiding treatment (Sze et al., 2021). However, its invasive nature means first-attempt failure directly leads to repeated attempts, increasing patient discomfort, risks of local complications (e.g., hematoma, vasospasm), reduced compliance, and higher costs.

Although techniques like ultrasound-guided puncture can improve success rates (Lu et al., 2018), traditional palpation methods remain prevalent in ward settings for routine monitoring due to practical constraints. Even with standardized training, first-attempt failure rates remain appreciable. Currently, systemic risk factors for first-attempt radial artery puncture failure in HF patients are not fully elucidated, and there is a lack of bedside prediction tools tailored to this group. Therefore, identifying key failure factors and establishing an effective prediction model is clinically urgent to enable individualized procedures, reduce patient harm, and improve care quality.

1.3 Relevant Scholarship

Previous studies have explored the factors related to arterial puncture, especially radial artery puncture, from different perspectives. Extensive literature focuses on radial access for coronary intervention, analyzing anatomical factors like vessel diameter and tortuosity. In terms of arterial blood collection or catheterization in non-invasive wards, research has mainly focused on operational techniques, such as the comparison between ultrasound-guided and traditional palpation (Lu et al., 2018), patient demographic characteristics, such as age and gender, and some physiological indicators. For example, research has explored the risk factors and predictive model construction for the failure of radial artery catheterization (Hu, 2018), providing methodological references for understanding this clinical issue.

However, there is still a significant gap in the first radial artery puncture study for the special population of heart failure patients. The unique pathophysiology of heart failure patients, such as volume overload causing edema, low cardiac output weakening pulses, hypoalbuminemia, and respiratory distress, may constitute specific risk factors that distinguish them from other patients (Wintrich et al., 2020; Serenelli et al., 2020). The existing universal puncture studies have not fully included and quantified these core variables related to heart failure, limiting the extrapolation of their conclusions.

Methodologically, the research standards for clinical predictive models are becoming increasingly sophisticated, covering sample size calculation (Riley et al., 2020), data partitioning strategies (such as Hold Out method, k-fold cross validation) (Mabuni & Babu, 2021), and model construction and validation (Shiwakoti et al., 2020; Elliott et al., 2021). Meanwhile, in addition to traditional logistic regression models, ensemble machine learning algorithms such as random forests and bagging have shown potential in handling complex nonlinear relationships and have been applied in other medical prediction fields (Jin et al., 2020). These developments provide technical possibilities for building more robust predictive models.

In summary, although existing research provides a basic understanding and technical framework for arterial puncture, there is a lack of specialized and systematic exploration on the risk of failure in the first radial artery puncture for heart failure patients. This study aims to integrate HF clinical characteristics with established methodologies to systematically identify risk factors and construct a prediction model using both statistical and machine learning approaches.

Journal of Biology and Life Science
ISSN 2157-6076
2026, Vol. 17, No. 2

1.4 Hypotheses and Research Design

Based on the above, this study proposes the following hypotheses and research questions:

Hypotheses:

The risk of first-attempt radial artery puncture failure in HF patients is significantly associated with specific clinical indicators reflecting disease severity and pathophysiology, particularly tissue edema, respiratory rate, mean arterial pressure, and plasma albumin levels.

Compared to using traditional logistic regression methods alone, predictive models constructed using ensemble machine learning algorithms such as Bagged Trees can capture more complex interactions and nonlinear relationships between variables, potentially exhibiting better discriminative performance (measured in AUC) on the validation set.

Research question:

What clinical and laboratory indicators that can be easily obtained from the medical records of HF patients are independent predictors of first-attempt radial artery puncture failure?

What are the discriminability (AUC), calibration, and clinical applicability of a logistic regression model based on these predictors?

Do ensemble machine learning models outperform logistic regression in this prediction task?

This retrospective cohort study included data from eligible HF patients undergoing their first radial artery puncture. This design allowed systematic collection of potential predictors (including demographics, vital signs, laboratory tests, complications, etc.) linked to a clear outcome (first puncture success/failure). The sample was randomly split into training and validation sets (Rueda & Guzmán, 2020). Logistic regression was used on the training set to identify independent risk factors and build a model (Kuss & McLerran, 2007; Petoukhov & Tuukkanen, 2017), with performance evaluated on the validation set. Concurrently, multiple ensemble models were trained on the same data for comparison. This design enables testing of the specific risk factor hypotheses and exploration of different modeling methods for the same predictive problem.

2. Data and Methods

2.1 Research Object

The sample size was estimated using the logistic independent variable event method (Riley et al., 2020). Accordingly, 789 inpatients from the Department of Internal Medicine at Dujiangyan Traditional Chinese Medicine Hospital were included, from April 2019 to August 2025. The inclusion criteria were as follows: (1) age over 18 years; (2) diagnosis of HF confirmed via the first radial artery puncture after admission, supported by symptoms, signs, laboratory tests, and cardiac ultrasound. Patients were excluded if they met any of the following conditions: severe critical illness requiring ICU care, shock, sudden death, procedure failure due to patient refusal, or cases with incomplete data.

2.2 Data Collection and Predictor Screening

Data for the first puncture attempt, including nursing notes, laboratory/imaging results, and clinical documentation, were retrieved from the hospital's Health Information System (HIS). Predictor screening involved applying frequency distribution analysis, independent samples t-tests, Chi-square tests, correlation analysis, and logistic regression to filter variables associated with procedural failure. Initial predictors included: age (years), sex (0<female> /1<male>), temperature situation (0<normal>\1<lower heat>\2<middle heat>\3<high heat>), body temperature (Celsius), arterial blood collection devices, heart rate (in times/min), brain natriuretic peptide (pg/mL), pulse rate, respiratory systolic pressure, diastolic pressure, pulse pressure difference, average arterial pressure, edema or not (0<no edema>/1<edema>), degree of edema (1<mild edema>\2<moderate edema> \3<severe edema>), atrial fibrillation during blood collection (0<none>/1<atrial fibrillation>), respiratory failure, pH value, carbon dioxide partial pressure, oxygen partial pressure, lactate, total protein, albumin, BNP, and NT-proBNP.

2.3 Standardized Operation of Radial Artery Puncture

The patient was placed supine with the forearm extended and supinated, supported by a towel roll under the wrist. The puncture site was selected at the point of maximal distal radial artery pulsation. The operator stabilized the artery with gentle middle-finger pressure, while the index finger was placed laterally to prevent slippage, especially in sclerotic vessels. A disposable arterial needle was inserted at a 25°–45° angle with the bevel facing cephalad. The angle was adjusted based on blood return. Successful first-attempt puncture was defined by aspiration of ≥2 ml of arterial blood; otherwise, it was considered a failure.

2.4 Prediction Model Construction and Evaluation

Using the Hold-Out method in Matlab 2024b, the sample was stratified into a training set (80%, n=631) and a validation set (20%, n=158) (Rueda & Guzmán, 2020). A predictive model was developed by screening for risk factors associated with initial puncture failure using logistic regression on the training data. Model verification included: 1) Comparing observed vs. predicted failure rates; 2) Calculating sensitivity, specificity, and the area under the ROC curve (AUC); 3) Assessing goodness-of-fit via the Hosmer-Lemeshow test, with the optimal classification threshold determined by maximizing Youden's index (Shiwakoti et al., 2020).

2.5 Statistical Methods

Analyses were performed using Matlab 2024b (Petoukhov & Tuukkanen, 2017; Kuss & McLerran, 2007). Normally distributed continuous variables are presented as mean ± standard deviation (SD) and compared using Student's t-test. Skewed variables are presented as median (interquartile range, IQR) and compared using the Wilcoxon rank-sum test. Categorical variables are presented as counts (percentages) and compared using the Chi-square test. A logistic regression-based prediction model was built and evaluated using ROC and calibration curves, with AUC quantifying discrimination. Statistical significance was set at a two-sided P-value < 0.05.

Journal of Biology and Life Science
ISSN 2157-6076
2026, Vol. 17, No. 2

3. Results

3.1 Baseline Characteristics

A comparison of baseline characteristics between the training and validation sets revealed no statistically significant differences for all relevant predictors (P > 0.05), confirming data homogeneity. Consequently, the validation set was deemed suitable for assessing the predictive efficacy of the developed model (Shiwakoti et al., 2020). Baseline characteristics are presented in Table 1.

Table 1. Baseline comparison between the training and validation groups

Variable	training group (631 cases)	validation group (158 cases)	$Z/t/\chi^2$	*P* value
Age(years, Median(Range))	78(27, 100)	78(42, 96)	-0.219	0.827
EF(EF value, M±SD)	55.99±12.07	56.31±12.58	-0.279	0.781
BNP(type B brain natriuretic peptide, pg/ml, M±SD)	667.19±914.03	731.36±991.13	-0.589	0.556
nt-BNP(NT-proBNP, pg/ml, M±SD)	4632.31±6868.39	4706.04±6378.32	-0.080	0.937
af(atrial fibrillation during blood collection, M±SD)	0.21±0.41	0.17±0.37	1.263	0.207
alb(albumin, M±SD)	35.48±4.96	35.02±5.44	1.009	0.313
ap(mean arterial pressure, M±SD)	94.47±13.66	96.77±20.62	-1.328	0.186
breath(breathe, M±SD)	23.60±3.06	23.78±3.33	-0.636	0.525
coo(carbon dioxide partial pressure, M±SD)	45.04±16.93	44.80±15.59	0.162	0.871
dp(diastolic pressure, M±SD)	76.23±12.65	77.43±12.10	-1.072	0.284
ede(edema or not, M±SD)	0.44±0.50	0.41±0.49	0.805	0.421
eded(edema degree, M±SD)	0.69±0.91	0.56±0.79	1.853	0.065
la(lactate, M±SD)	1.62±0.97	1.63±0.88	-0.111	0.912
oo(oxygen partial pressure, M±SD)	74.37±29.56	71.64±27.48	1.052	0.293
ph(PH value, M±SD)	7.41±0.08	7.39±0.24	1.161	0.247
pp(pulse pressure, M±SD)	54.72±18.63	58.02±51.81	-0.787	0.432
rate(pulse rate, M±SD)	93.16±21.38	93.74±21.21	-0.308	0.758
rf(respiratory failure or not, M±SD)	0.43±0.49	0.47±0.50	-1.096	0.273
sbp(systolic blood pressure, M±SD)	130.96±21.71	135.45±52.65	-1.051	0.295
sex(M±SD)	0.59±0.49	0.71±0.46	-2.764	0.026
temp(body temperature, M±SD)	36.65±0.87	36.88±3.21	-0.905	0.367
tempj(body temperature degree, M±SD)	0.20±0.62	0.19±0.55	0.271	0.786
tp(Total Protein, M±SD)	63.60±24.25	62.99±9.70	0.491	0.624
y(First arterial puncture success or not, M±SD)	0.20±0.40	0.19±0.39	0.188	0.851

3.2 Data Preprocessing

Missing values in the clinical dataset were imputed using the Expectation-Maximization (EM) algorithm. The algorithm ran for 15 iterations until convergence, generating a complete dataset, comparison reports, and data quality assessments, ensuring integrity for subsequent modeling.

3.3 Logistic Regression Prediction Model

A logistic regression framework with feature engineering and dual-stage variable selection was implemented. The initial pool had 24 variables, including original measures and engineered features (interaction terms, log-transformations, composite scores, PCA components).

The variable selection process employed a two-step strategy to balance model stability and interpretability. 1) Least Absolute Shrinkage and Selection Operator (LASSO) regression with 5-fold cross-validation performed initial feature screening by applying L1 regularization to compress coefficients of less informative variables toward zero. 2) Stepwise logistic regression with a p-value threshold of 0.1 was applied to the LASSO-selected variables, iteratively adding and removing features based on statistical significance.

The final model, which included an intercept term, comprised 11 variables. These were selected based on having an inclusion frequency of at least 60% across the top-performing iterations. The model equation is:
$\text{logit}(P) = -0.4137 + (-0.0536 \times \text{sex}) + (0.1017 \times \text{heart rate}) + (-0.0959 \times \text{edema duration}) + (0.0433 \times \text{left atrial size}) + (0.0751 \times \text{albumin}) + (0.0399 \times \log_{10}(\text{BNP}+1)) + (0.0630 \times \text{age} \times \text{BNP}) + (0.0647 \times \text{heart rate} \times \text{respiration}) + (-0.0698 \times \text{systolic-diastolic pressure ratio}) + (0.1507 \times \text{ejection fraction} \times \log_{10}\text{BNP}) + (0.0448 \times \text{respiratory})$,

where the probability of arterial puncture failure $P = \exp(\text{logit}) / [1 + \exp(\text{logit})]$. Two variables were statistically significant ($p<0.05$): edema duration ($\beta = -0.0959$, OR = 0.9085, $p = 0.0160$) and the interaction between ejection fraction and log-transformed BNP ($\beta = 0.1507$, OR = 1.1627, $p = 0.0371$). Full model details are in Table 2.

Table 2. Stepwise regression analysis of predictors for radial artery puncture failure

Incorporating variable	**Estimated value**	**Standard error**	**t value**	**Odds Ratio**	**OR lower bounds**	**OR upper bounds**	***Wald***	***P value***
Intercept	-0.4137	0.1013	-4.0828	---	0.6612	0.6612	16.6689	---
sex	-0.0536	0.1004	-0.5341	0.9478	0.7785	1.1539	0.2852	0.1250
rate	0.1017	0.2231	0.4556	1.1070	0.7149	1.7143	0.2076	0.3393
eded	-0.0959	0.1079	-0.8892	0.9085	0.7354	1.1224	0.7907	0.0160
la	0.0433	0.0973	0.4450	1.0442	0.8630	1.2635	0.1980	0.2463
alb	0.0751	0.1138	0.6599	1.0780	0.8625	1.3472	0.4355	0.0687
$\log_{10}$(BNP+1)	0.0399	0.1724	0.2316	1.0407	0.7423	1.4592	0.0536	0.5503
age×BNP	0.0630	0.1940	0.3247	1.0650	0.7281	1.5579	0.1054	0.2867
rate×breath	0.0647	0.1785	0.3623	1.0668	0.7518	1.5138	0.1313	0.5447
sbp×dp_ratio	-0.0698	0.1075	-0.6492	0.9326	0.7554	1.1513	0.4215	0.0866
EF×BNP	0.1507	0.1221	1.2345	1.1627	0.9152	1.4771	1.5239	0.0371
rf_PC1	0.0448	0.0816	0.5492	1.0458	0.8913	1.2272	0.3016	0.3231

3.4 Model Validation

Performance was evaluated via a 50-repeat holdout validation (80% training/20% testing). The model showed fair discriminative ability with an average AUC of 0.693 (±0.033), high specificity (95.59% ± 2.55%), but relatively low sensitivity (31.88% ± 13.98%). Other metrics included balanced accuracy of 79.49% (±1.68%), the positive predictive value 71.73% (±3.14%), the F1-score reached 42.82% (±12.07%), Brier score 0.194 (±0.008), and the Matthews Correlation Coefficient 0.375 (±0.080). The Hosmer-Lemeshow test indicated no significant difference between predicted and observed event rates ($p > 0.05$). The optimal classification threshold, determined by maximizing Youden's index, was identified at a predicted probability of 0.658, with a sensitivity of 84.7% and specificity of 67.8% in the training cohort. The ROC curve is shown in Figure 1. For comparison, a model excluding BNP/NT-proBNP performed worse (Figure 2).

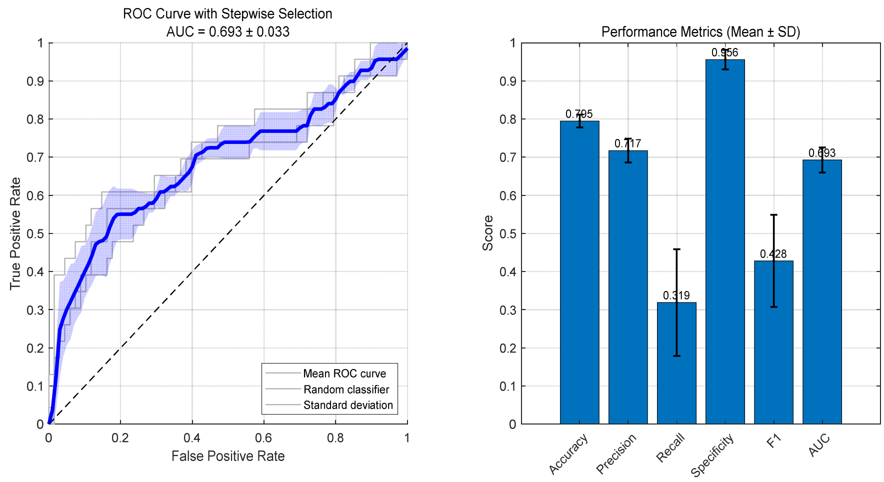

Figure 1. ROC curves of the final logistic regression prediction model

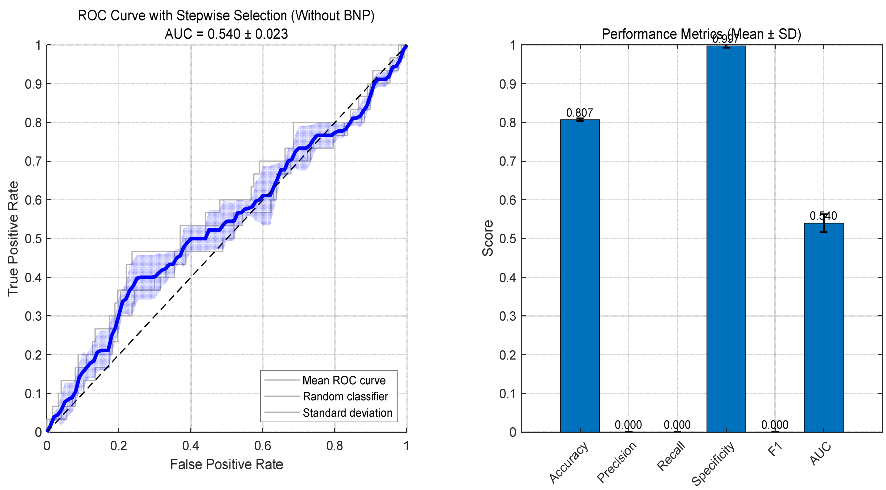

Figure 2. ROC curves of a model excluding BNP and ntBNP variables

3.5 Ensemble Learning Models

To improve predictive performance beyond conventional statistical approaches, we employed multiple ensemble machine learning algorithms such as Random Forest, AdaBoost, Bagged Trees, and RUSBoost. The study dataset comprised 456 clinical samples with 22 predictor variables, demonstrating significant class imbalance with 340 negative cases (class 0) and 116 positive cases (class 1) for the target outcome.

The ensemble models were implemented using MATLAB's classification ensemble functions (TreeBagger for Random Forest and fitcensemble for others). Model configurations varied:

Random Forest employed 100 trees with bootstrap aggregation, AdaBoost utilized 200 estimators with a learning rate of 0.1, Bagged Trees implemented 100 cycles with full bootstrap resampling, and RUSBoost incorporated 200 cycles with random under-sampling to address class imbalance. The dataset was partitioned into a training set (70%, n=320) for model development and an independent test set (30%, n=136) for unbiased validation, with a fixed random seed ensuring methodological reproducibility.

Comparative analysis revealed that Bagged Trees achieved the highest discriminative performance (AUC = 0.6533), followed by Random Forest (AUC = 0.6214), RUSBoost (AUC = 0.6 192), and AdaBoost (AUC = 0.5750). The Bagged Trees model demonstrated moderate accuracy (73.53%) on the test set, though its sensitivity (5.56%) and F1-score (0.1000) were suboptimal, primarily due to challenges in identifying positive cases within this imbalanced clinical dataset. While these ensemble methods did not exhibit severe overfitting typically associated with complex algorithms on limited clinical data, their modest AUC values (0.575-0.653) suggest room for improvement in predictive capability.

Feature importance analysis derived from the Random Forest model (Figure 3) identified diastolic pressure (dp, 0.3194), heart rate (rate, 0.2940), edema (ede, 0.1911), oxygenation (oo, 0.1769), and respiratory rate (breath, 0.1700) as the most influential predictors. These physiological parameters align with established clinical knowledge regarding cardiovascular and respiratory determinants of patient outcomes, though their moderate predictive importance scores (0.016-0.319) suggest that outcome variability is distributed across multiple clinical factors rather than dominated by single predictors.

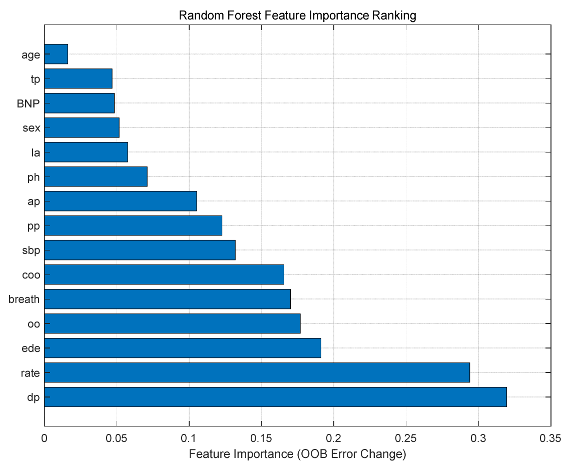

Figure 3. Feature importance ranking from the Random Forest model

Table 3. Predictive performance of ensemble learning models on the validation set

Model	Accuracy	AUC	F1-Score	Precision
Random Forest	0.7206	0.6214	0.0500	0.2500
AdaBoost	0.6985	0.5750	0.1961	0.3333
Bagged Trees	0.7353	0.6533	0.1000	0.5000
RUSBoost	0.6324	0.6192	0.4444	0.3704

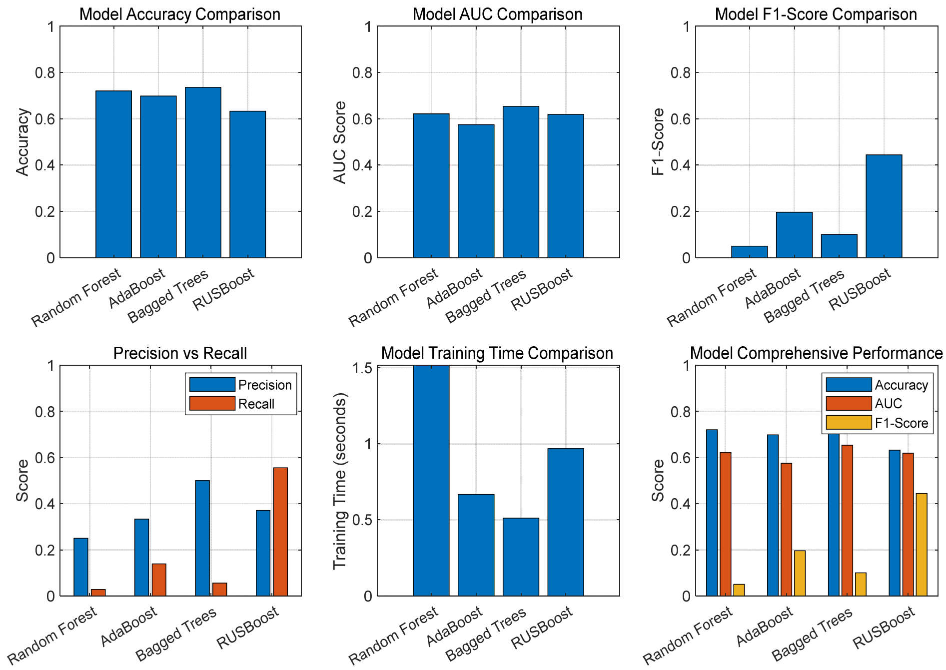

Figure 4. Summary of performance metrics for the four ensemble learning models

4. Discussion

This study developed and compared models to predict first-attempt radial artery puncture failure in HF patients. The logistic regression model identified clinically interpretable risk factors and demonstrated fair predictive performance. In contrast, the best-performing ensemble model (Bagged Trees) showed slightly higher discriminative ability on the validation set but exhibited signs of overfitting and very low sensitivity. This suggests that while ensemble machine learning methods hold potential for capturing complex relationships, their clinical utility in this specific context with the current dataset may be limited compared to more interpretable traditional models.

4.1 Identified Risk Factors and Pathophysiological Links

The risk factors identified are closely tied to HF pathophysiology, providing plausible mechanisms for their impact:

4.1.1 Edema Degree

Tissue edema is a common sign in heart failure patients, and its severity (OR = 0.9085) was associated with an increased risk of puncture failure. Tissue edema from volume overload can obscure anatomical landmarks and weaken arterial pulsation, increasing the difficulty of arterial puncture.

4.1.2 Rapid Respiration

A clinical dilemma arises from the need for a supine position during arterial puncture in heart failure patients with tachypnea (OR = 0.918), a marker of decompensation. This position contradicts the patient's instinct to sit up to ease respiratory distress. The increase in cardiac preload and resultant involuntary muscle tension and movement associated with lying flat compromise operator stability and success.

4.1.3 Decreased Mean Arterial Pressure

Lower blood pressure (OR = 1.073) may reflect reduced cardiac output or circulating blood volume. Under these conditions, radial artery filling decreases and the lumen narrows, increasing the difficulty of accurately entering the vascular space with the puncture needle. The cardiac output, effective blood volume, and peripheral resistance are all related to blood pressure (Serenelli et al., 2020).

4.1.4 Reduced Plasma Albumin

Hypoalbuminemia (OR = 1.078) is often associated with a state of malnutrition in heart failure patients. Hypoalbuminemia, common in HF due to malnutrition and cachexia, lowers plasma oncotic pressure, exacerbating tissue edema at the puncture site and increasing difficulty (Wada et al., 2019). The plasma albumin decreased, and the subcutaneous tissue edema at the puncture site increased accordingly, which increased the difficulty of the puncture.

4.1.5 Interaction Between EF And Log(BNP)

This significant interaction suggests a complex interplay between cardiac systolic function (EF) and neurohormonal activation (BNP). The combined effect may influence vascular tone, fluid status, and patient stability in ways that affect puncture success more than either factor alone.

4.2 Comparison of Modeling Approaches

This study highlights the trade-off between interpretability and potential predictive power. The logistic regression model offers clear odds ratios and confidence intervals, facilitating clinical understanding and potential integration into simple decision rules. The ensemble models, particularly Bagged Trees, achieved a marginally higher AUC but functioned as "black boxes" with poor sensitivity. Their performance, coupled with the risk of overfitting on limited clinical data, suggests that for this specific problem and dataset, the gain in discrimination may not justify the loss of interpretability and reliability. Future work with larger datasets might better realize the potential of these complex algorithms.

4.3 Clinical Translation and Application Prospects of the Models

The core value of this predictive model lies in enabling pre-procedural risk assessment. For patients identified as high-risk by the model, clinicians can implement interventions in advance, such as: optimizing heart failure treatment to improve the patient's overall condition (e.g., reducing edema, raising blood pressure); providing appropriate sedation or reassurance before puncture to reduce anxiety; or directly considering alternative puncture sites (e.g., femoral artery) to avoid the complications and patient distress associated with first-attempt radial artery puncture failure. This contributes to advancing radial artery puncture from an experience-based procedure towards a more precise and individualized one.

This study is retrospective. The present study has limitations in three key aspects. The constructed model does not account for local anatomical factors such as vessel compliance and perivascular fat. Furthermore, its development relies on a dataset of limited size without external validation, which restricts its potential for widespread implementation (Jin et al., 2020). Regarding the analytical approach, the random split-sample validation used may be less reliable than resampling techniques, and the data processing in Matlab did not facilitate the creation of visual model aids for clinicians (Zhang et al., 2021).

To advance this work, prospective and multi-center investigations with larger cohorts are warranted. Future models will seek to incorporate novel predictors such as arteriosclerosis metrics and evaluate the role of TCM pulse diagnosis. The validation framework will be strengthened by comparing various data partitioning techniques and by adding an essential external validation step (Liu et al., 2020). Furthermore, we intend to employ advanced statistical software (e.g., R) to create user-friendly visualizations of the model, such as nomograms or digital tools (Zhang et al., 2021). These concerted efforts are expected to refine the model's accuracy (sensitivity/specificity) and its applicability in personalizing clinical procedures.

5. Conclusion

This study identified edema degree, rapid respiration, lower blood pressure, hypoalbuminemia, and an interaction between EF and BNP as key factors associated with first-attempt radial artery puncture failure in HF patients. A logistic regression model based on these factors provided a fair and clinically interpretable prediction tool. Ensemble learning models showed modest potential but without clear superiority in this context. The developed model can aid in pre-procedural risk stratification, paving the way for more individualized and precise clinical procedures to improve patient comfort and procedural efficiency.

Acknowledgments

Not applicable.

Authors contributions

Huan Wang and Jiangbo Zhang were responsible for study design and revising. Huan Wang was responsible for data collection. Huan Wang and Jiangbo Zhang drafted the manuscript and Jiangbo Zhang revised it. All authors read and approved the final manuscript.

Funding

This study was supported by Project of Chengdu Municipal Health Commission [project number 202303023185].

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Macrothink Institute.

The journal's policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Data sharing statement

No additional data are available.

Open access

Copyrights

References

Elliott, C. F., Lambert, J. W., Stromberg, A. J., Caffo, B. S., & Arheart, K. L. (2021). Feasibility as a mechanism for model identification and validation. *Journal of Applied Statistics, 48*(11), 2022-2041. https://doi.org/10.1080/02664763.2020.1753029

Hu, T. (2018). *Risk factors and prediction model construction of patients with failed radial artery catheterization* [Unpublished doctoral dissertation]. PLA Naval Medical University, Shanghai, China.

Jin, L., Sun, H., Zhou, L., Chen, Y., & Shi, B. (2020). Prediction model of more than 5 central lymph nodes metastases in clinically node-negative ipsilateral papillary thyroid carcinoma with tumor size 1 to 4 cm. *Medicine, 99*(16), e19809. https://doi.org/10.1097/MD.0000000000019809

Kuss, O., & McLerran, D. (2007). A note on the estimation of the multinomial logistic model with correlated responses in SAS. *Computer Methods and Programs in Biomedicine*, *87*(3), 262-269. https://doi.org/10.1016/j.cmpb.2007.05.004

Liu, L., Yin, C., Zhi, Y., Cao, Y., Zhang, L., & Wang, X. (2020). Preliminary establishment of weaning prediction model. *Zhonghua Wei Zhong Bing Ji Jiu Yi Xue, 32*(2), 171-176. http://doi.org/10.3760/cma.j.cn121430-20191015-00032

Lu, T., Fei, W., & Yuxiang, L. (2018). Ultrasound guidance for radial artery catheterization: An updated meta-analysis of randomized controlled trials. *Public Library of Science ONE, 13*(11), e0206312. https://doi.org/10.1371/journal.pone.0206312

Mabuni, D., & Babu, S. A. (2021). High accurate and a variant of k-fold cross validation technique for predicting the decision tree classifier accuracy. *International Journal of Innovative Technology and Exploring Engineering*, *10*(3), 105-10.

Malik, A., Brito, D., & Chhabra, L. (2021). Congestive heart failure. *StatPearls*. StatPearls Publishing.

Petoukhov, M. V., & Tuukkanen, A. (2017). SAS-based structural modelling and model validation. In *Data mining in clinical medicine* (pp. 87-105). Springer Singapore. https://doi.org/10.1007/978-981-10-4476-4_7

Riley, R. D., Ensor, J., Snell, K. I. E., Harrell, F. E., Jr., Martin, G. P., Reitsma, J. B., … & van Smeden, M. (2020). Calculating the sample size required for developing a clinical prediction model. *British Medical Journal, 368*, m441. https://doi.org/10.1136/bmj.m441

Rueda-Bayona, J. G., & Guzmán, A. (2020). Genetic Algorithm model and data files to assess JONSWAP spectra coefficients: MATLAB code. *Data in Brief*, *32*, 106196. https://doi.org/10.1016/j.dib.2020.106196

Serenelli, M., Böhm, M., Inzucchi, S. E., Køber, L., Kosiborod, M. N., Martinez, F. A., … & Jhund, P. S., for the DAPA-HF Trial Investigators and Committees. (2020). Effect of dapagliflozin according to baseline systolic blood pressure in the Dapagliflozin and Prevention of Adverse Outcomes in Heart Failure trial (DAPA-HF). *European Heart Journal*, *41*(36), 3402-3418. https://doi.org/10.1093/eurheartj/ehaa496

Shiwakoti, E., Song, J., Li, J., Tang, B., & Ma, W. (2020). Prediction model for anastomotic leakage after laparoscopic rectal cancer resection. *Journal of International Medical Research*, *48*(9), 0300060520944596. https://doi.org/10.1177/0300060520944596

Sze, S., Pellicori, P., Zhang, J., Weston, J., & Clark, A. L. (2021). Effect of frailty on treatment, hospitalization, and death in patients with chronic heart failure. *Clinical Research in Cardiology, 110*(8), 1249-1258. https://doi.org/10.1007/s00392-021-01858-3

Wada, Y., Izumi, H., Shimizu, T., & Hasegawa, T. (2019). A more oxidized plasma albumin redox state and lower plasma HDL particle number reflect low-protein diet ingestion in adult rats. *The Journal of Nutrition, 150*(2), 256-266. https://doi.org/10.1093/jn/nxz236

Wintrich, J., Kindermann, I., & Böhm, M. (2020). Neues zur Herzinsuffizienz [Update on heart failure]. *Herz: Cardiovascular Diseases, 45*(2), 158-169. https://doi.org/10.1007/s00059-020-04899-y

Zhang, W., Zeng, Y., Chang, B., Wang, X., Jia, H., & Li, C. (2021). Prognostic nomogram for microvascular decompression-treated trigeminal neuralgia. *Neurosurgical Review, 44*(1), 571-577. https://doi.org/10.1007/s10143-020-01246-x