Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Development of an Interpretable Machine Learning Model and Nomogram for Predicting Radiation-Induced Pericarditis Risk in Breast Cancer

Yongkai Lu

luyongkai0714@163.com

First Affiliated Hospital of Xi'an Jiaotong University

Lei Xu

First Affiliated Hospital of Xi'an Jiaotong University

Jian Zhang

First Affiliated Hospital of Xi'an Jiaotong University

Rongze Ma

First Affiliated Hospital of Xi'an Jiaotong University

Yao Wang

First Affiliated Hospital of Xi'an Jiaotong University

Shanshan Gao

First Affiliated Hospital of Xi'an Jiaotong University

Article

Keywords: Breast cancer, Radiation-induced pericarditis, Machine learning, SHAP, Nomogram

Posted Date: May 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9470275/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Radiation-induced pericarditis (RIP) is a potential toxicity of breast cancer radiotherapy. Traditional single-parameter predictions exhibit inherent limitations.

Methods

This retrospective study included 262 female breast cancer patients. A Random Forest (RF) model was developed to predict RIP risk, defined by normal tissue complication probability. The SMOTE technique addressed class imbalance. The SHapley Additive exPlanations (SHAP) framework was applied to evaluate statistical feature associations. A clinical nomogram was subsequently constructed.

Results

The RF model achieved an optimal area under the curve of 0.989 in the testing set. SHAP analysis indicated that high-dose parameters, particularly heart V40 and V30, were most strongly associated with the predicted risk, alongside spatial features like maximum heart distance. The constructed nomogram demonstrated robust discrimination (optimism-corrected C-index = 0.9487) and satisfactory calibration.

Conclusion

The interpretable RF model and clinical nomogram offer a multidimensional approach for RIP risk stratification. These tools suggest that specific dosimetric and spatial features are closely associated with predicted risk, providing quantitative references that may assist in future dose-constraint optimizations.

1. Introduction

Breast cancer remains the most frequently diagnosed malignancy among women globally, with radiotherapy serving as an indispensable component of its comprehensive treatment[1, 2]. Although radiotherapy is an effective treatment that increases the chance of survival, it inevitably places the heart at risk of collateral radiation[3]. Such exposure can cause different levels of cardiac damage, and radiation-induced pericarditis is a known clinically significant acute cardiotoxicity[4].

It is believed that the development of radiation-induced pericarditis is a complicated phenomenon which can be driven by the combined influence of patient clinical characteristics, pathological characteristics of tumors, and dosimetric parameters of treatment[5]. Most of the earlier research has focused on the relationship between single-dose measurements (i.e., the percentage of heart volume receiving 5 Gy

(V5)) and the occurrence of pericarditis[6]. The use of a single-parameter predictive pattern has its own limitations. Conventional linear statistical models often face challenges in processing high-dimensional clinical data with non-linear relationships and high multicollinearity, leading to a potential loss of predictive information[7, 8].

Machine learning methods have been shown to have a significant benefit in surmounting these statistical challenges through the successful incorporation of multidimensional heterogeneous data[7, 8]. However, current clinical use of state-of-the-art machine learning algorithms is usually hindered by their inherent black-box characteristics that make the decision-making process less transparent[9]. The emergence of the SHapley Additive exPlanations (SHAP) framework offers a feasible solution to elucidate these complicated algorithms, thus facilitating transparent clinical assessments[10, 11].

Therefore, this study aims to develop a multidimensional risk prediction model for radiotherapy-induced pericarditis in breast cancer patients based on machine learning algorithms. With the inclusion of SHAP analysis, we aim to explain the individual contribution of combined dosimetric and clinical characteristics to the risk as it is predicted. Moreover, to address the gap in complexity of algorithms and clinical practice, we will build a nomogram using the identified key features and intend to offer a convenient risk-stratification tool that could be applied individually at the bedside.

2. Materials and Methods

2.1 Study Population and Ethical Approval

This retrospective cohort study included female patients diagnosed with breast cancer who received radiotherapy at the Department of Radiation Oncology, First Affiliated Hospital of Xi'an Jiaotong University, between January 2022 and June 2023. This study was approved by the Ethics Committee of the First Affiliated Hospital of Xi'an Jiaotong University. Given the retrospective nature of the study and the use of anonymized clinical data, the requirement for written informed consent was waived by the ethics committee.

2.2 Data Collection and Outcome Definition

A total of 47 candidate variables were systematically extracted from electronic medical records and treatment planning systems. The collected clinical indicators encompassed four main categories: (1) Demographic and clinical characteristics, including age, body mass index, and history of underlying diseases such as diabetes and hypertension. (2) Molecular and pathological features, comprising estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER-2), Ki67 expression, pathological type, lesion position, and TNM staging. (3) Radiotherapy scheme parameters, covering target area, radiotherapy technique, number of fractions, and dose coverage metrics. (4) Dosimetric parameters, including heart V5, V30, V40, heart mean dose, maximum heart distance (MHD), ipsilateral and contralateral lung dose-volume parameters, and liver mean dose.

The primary outcome was the probability of radiation-induced pericarditis, calculated based on the normal tissue complication probability (NTCP) model parameters (TD50 = 50Gy, γ50 = 3). The calculated probability was dichotomized using a threshold of 3.8487×10^{-6}. Based on this predefined cutoff, the study population (N = 262) was categorized into a high-risk group (54 cases) and a low-risk group (208 cases).

2.3 Data Preprocessing and Handling of Class Imbalance

To ensure data quality, systematic preprocessing was performed. Missing values for continuous variables were imputed using the median, while categorical variables were imputed using the mode. Outlier corrections were made, including capping opposite-side (contralateral) lung V5 values exceeding 100 percent to the median. According to clinical criteria, Ki67 was dichotomized using a 20 percent cutoff. Variables with a high proportion of zero values (more than 50 percent), including total dose and gross tumor volume, were converted into binary variables.

A 7:3 stratified random sampling approach was then used to partition the dataset into a training set (183 cases) and a testing set (79 cases). The Synthetic Minority Over-sampling Technique (SMOTE) was used to overcome the imbalance of the classes between the high-risk and low-risk instances. It should be mentioned that SMOTE was restricted solely to the training set to avoid potential data leakage and ensure an objective assessment on the testing set.

2.4 Machine Learning Modeling and Parameter Optimization

Following data preprocessing, five machine learning algorithms were constructed to predict the risk of pericarditis. The algorithms were Logistic Regression, Random Forest, eXtreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM) and Support Vector Machine (SVM). Each algorithm's hyperparameters were optimized using a grid search (GridSearchCV) strategy with 5-fold stratified cross-validation on the training set, optimizing primarily for the area under the receiver operating characteristic curve (AUC). The performance of the optimized models was also fully assessed on the independent testing set based on measures such as AUC, sensitivity, specificity, F1 score and Brier score. To evaluate statistical differences between the AUC values of the candidate models, the DeLong test was used.

2.5 Interpretability Analysis and Nomogram Construction

To elucidate the decision-making rationale of the optimal machine learning model, the SHapley Additive exPlanations (SHAP) framework was employed. In particular, SHAP values were calculated using a TreeExplainer on the whole dataset, facilitating the generation of global feature importance ranks, summary beeswarm plots, and localized waterfall plots. The partial dependence plots were also created to explore the marginal effects of the top-ranked features. According to the predominant features of the SHAP analysis, a multivariate logistic regression nomogram was created to enable clinical utility at the

bedside. Calibration and discrimination of the nomogram was internally validated through a Bootstrap resampling method that took 1000 iterations giving an optimism-corrected C-index.

3. Results

3.1 Baseline Characteristics and Data Preprocessing of the Study Population

A total of 262 female breast cancer patients who received radiotherapy were retrospectively included in this study. The main outcome was determined according to the normal tissue complication probability model, with a threshold calculated at 3.8487×10^{-6}. As a result, 54 patients (20.6 percent) were categorized into the high-risk category of radiation-induced pericarditis, and 208 patients (79.4 percent) were assigned into the low-risk category. Figure 1A depicts the overall distribution of outcomes of the study population. A systematic preprocessing plan was applied to the 47 candidate variables to guarantee the integrity of data. The median and mode were used to impute missing values of continuous and categorical variables respectively. Special outlier adjustments were performed, including capping extreme values of opposite side lung V5 to the median and dichotomizing Ki67 at a 20 percent clinical level. Visualization of age and body mass index distribution of the two cohorts was presented in Fig. 1B with a median age of 47 years and relatively similar body mass index distributions across groups, using violin plots. Figure 1C density plots showed that patients at high-risk were more likely to receive higher radiation doses to the heart. The structural distribution was further detailed by a detailed dot plot of clinical and treatment characteristics in Fig. 1D which showed possible differences in variables like position of left sided lesions. Moreover, the correlation heatmap in Fig. 1E showed numerous positive correlations among the cardiac and pulmonary dosimetric parameters, which provided a critical rationale for the subsequent choice of statistical modeling methods.

3.2 Feature Selection and Statistical Evaluation

To systematically identify potential predictors, a rigorous feature selection pipeline was implemented. First, all candidate variables were subjected to univariate logistic regression, identifying 11 parameters with a p-value of less than 0.05. Cardiac parameters (heart V40 (odds ratio 87.88) and heart V30 (odds ratio 5.93)) had significant univariate relationships with pericarditis risk, as shown in the forest plot in Fig. 2A. An overview of these significant features and their significance levels is summarized in the bubble plot in Fig. 2B. In order to establish the best clinical thresholds of continuous variables, receiver operating characteristic curves were plotted using the Youden index method, as illustrated in Fig. 2C. Thereafter, the analysis of variance inflation factor was done to assess possible multicollinearity. The bar chart in Fig. 2D shows that the cardiac dose parameters had severe multicollinearity with the variance inflation factor of heart mean dose up to 25.0. In this case, conventional stepwise regression was only able to stably retain one variable (heart V30), implying that a significant amount of predictive information was lost. In its turn, to deal with this, least absolute shrinkage and selection operator (LASSO) regression

with 10-fold cross-validation was used (Fig. 2E). Figure 2F is the coefficient path plot that illustrates the shrinkage process that chose a stable subset of five core features. The different distribution patterns of these top significant features between the risk groups are shown in the boxplots in Fig. 2G, and their different effects sizes in the risk stratifications are described in the bubble plot in Fig. 2H. Additional reasons that underscore the need to use machine learning algorithms are the limitations of linear models when dealing with highly correlated features.

3.3 Machine Learning Model Construction and Comprehensive Evaluation

The dataset was divided into a training set (183 cases) and a testing set (79 cases) using a 7:3 stratified random sampling approach to maintain the outcome distribution. To reduce the risk of model bias due to the imbalance in the classes (20.6 percent high-risk and 79.4 percent low-risk), Synthetic Minority Over-sampling Technique (SMOTE) was used only in the training set. The training and optimization of 5 different machine learning algorithms were performed with the use of 5-fold cross-validation combined with grid search. The receiver operating characteristic curves of the test set shown in Fig. 3A suggested that the Random Forest model achieved the highest area under the curve (0.989). Figure 3B is a radar chart that summarizes the multidimensional metrics in detail, demonstrating that the Random Forest model gave a balanced output in terms of sensitivity (0.938) and specificity (0.968). Figure 3C showed that the predicted probabilities using the Random Forest model were close to the actual observations and the Brier score was low (0.045). Additionally, the decision curve analysis in Fig. 3D indicated that this model could provide considerable net clinical benefit across various threshold probabilities. In Fig. 3E, precision-recall curves and in Fig. 3F, the sensitivity-specificity scatter plot, the strength of the Random Forest algorithm in processing this very dataset was further supported. Even though the DeLong test heatmap in Fig. 3G showed that the values of area under the curve did not differ statistically significantly between the five optimized models (p larger than 0.05), the Random Forest was chosen as the final predictive model because it showed better stability and more balanced metric profile. Lastly, an exceptional discriminatory ability was observed by a risk stratification bar chart using predicted probabilities in Fig. 3H with actual positive rate of 88.2 percent in the predicted high-risk group, and 0 percent in the predicted low-risk group.

3.4 Global and Local Interpretability Analysis via SHAP Framework

Recognizing that the black-box nature of tree-based models may hinder clinical interpretation, the SHapley Additive exPlanations (SHAP) framework was introduced to provide multidimensional insights into the model's predictive behavior. Critically, it is important to observe that SHAP values are used to measure the associative values of the features to the output probability of the model and not the actual cause and effect relationships of biology. As ranked in the global feature importance bar chart in Fig. 4A, the heart V40, heart V30, and heart mean dose exhibited the highest mean absolute SHAP values,

identifying them as the strongest features associated with the model's risk predictions. Figure 4B, a beeswarm plot, helped to gain more insight into the directionality of these statistical associations. It is worth noting that maximum heart distance (MHD) had a negative association with SHAP value, which implied that larger physical distance between the heart and the radiation field is associated with a smaller risk of pericarditis prediction in this study. Figure 4C SHAP dependence plots demonstrated that the individual key dosimetric parameters had non-linear marginal relationships with the prediction outputs. In addition, the correlation heatmap of SHAP values in Fig. 4D showed the statistical interaction between the predictive weights of these features as they interacted with each other in the internal logic of the model. SHAP waterfall plots were created to shift the global trends to individual level. Figure 4E breaks down the mathematical prediction procedure of a typical true-positive high-risk patient, and shows how certain high-dose parameters progressively contributed to an increased estimated risk probability compared to the population baseline. In contrast, Fig. 4F demonstrates the pathway of probability assignment of a true-negative low-risk patient, in which the contributions of negative features were all associated with a reduced predicted risk score.

3.5 Partial Dependence and Probability Distribution Analysis

To further elucidate the complex dose-response relationships learned by the machine learning model, partial dependence plots were constructed. Figures 5A to 5D outline the marginal effects of the top four features: heart V40, heart V30, heart mean dose, and maximum heart distance. An interesting observation in Fig. 5A is that there is a possible threshold effect on heart V40 whereby the risk of pericarditis was predicted to increase steeply beyond about 2 percent of the volume. Such a quantitative observation may serve as useful data support in future clinical dose-volume constraints. Besides the isolated effects, the interaction scatter plot of SHAP, shown in Fig. 5E, indicated possible synergistic predictive interactions between heart V40 and heart V30. In order to justify the discriminative resolution of the model outputs, the density distribution of the predicted probabilities stratified by actual clinical outcomes was drawn in Fig. 5F which showed clear separation with very little overlap. Figure 5G shows the probability distributions of all five models analyzed, indicating that the Random Forest model can be confident to assign probabilities in the extremes. To combat the issue of possible overfitting in the model, a grouped bar chart in Fig. 5H was used to compare the performance measures (AUC, F1 score, sensitivity, specificity) of both cross-validation and independent test sets, indicating that there was no significant performance degradation, demonstrating consistent generalization performance.

3.6 Construction and Internal Validation of the Clinical Nomogram

While the Random Forest model exhibited superior predictive performance and could manage complex multicollinearity, its routine bedside application requires specific computational software. To simplify the complex algorithms and make them clinically useful, a nomogram with user-friendly functionality was

formulated using the results of the multivariate logistic regression with the top 8 features, provided by the SHAP analysis, as shown in Fig. 6A. In order to be rigorous in the evaluation of the calibration of this simplified tool, Bootstrap resampling with 1000 iterations was employed. Figure 6B results showed good consistency between the apparent and Bootstrap corrected predicted probabilities. The decision curve analysis in Fig. 6C indicated that the nomogram provided obvious clinical net benefit across a realistic threshold probability range of 5 to 45 percent, as compared to the treat-all or treat-none approaches. The C-index evaluation plot in Fig. 6D demonstrated an apparent C-index of 0.9647, which remained high after Bootstrap correction (optimism value of 0.016), showing strong discrimination. Figure 6E represents the nomogram corresponding receiver operating characteristic curve. The nomogram scores density and scatter distribution shown in Fig. 6F ensure that the nomogram separates risk cohorts well. The visual summary of the log-odds coefficients of the variables included is presented in the forest plot in Fig. 6G. Lastly, the Hosmer-Lemeshow calibration plot in separate decile groups in Fig. 6H ($p = 0.041$) provided additional evidence in support of the acceptable goodness-of-fit of this translated clinical instrument.

4. Discussion

Radiotherapy remains a crucial component in the comprehensive treatment of breast cancer, yet it unavoidably exposes the heart to radiation, which may potentially lead to toxicities such as radiation-induced pericarditis[12]. Our research generated an interpretable Random Forest model and a corresponding clinical nomogram, which infers that high dose cardiac dosimetric parameters like heart V40 and heart V30 have significant predictive relationships with the risk of pericarditis.

Past research has addressed this issue mostly by examining how a single dosimetric parameter, such as heart V5, is related to subsequent cardiac toxicity[13–15]. To expand on this premise, our results indicate that the parameters of a higher dose (heart V40 and V30) could be of more predictive value than the low-dose parameters in a multidimensional machine learning framework[16–18]. In addition, although the primary outcome was obtained with dose-based NTCP predictions, the Random Forest model has managed to recognize non-dosimetric spatial factors, including the maximum heart distance and the location of a lesion, as the key risk modifiers. This implies that our machine learning strategy provides a more complete one-on-one evaluation as compared to conventional single-parameter dose measures.

Mathematically, one of the issues that is often raised in dosimetric modeling is the treatment of variables that are highly correlated[19]. As indicated by the variance inflation factor analysis, extreme multicollinearity exists among cardiac parameters, which dramatically limits the stability and variable retention capacity of traditional stepwise logistic regression[20]. The Random Forest model addresses this statistical shortcoming with algorithms such as random feature subspace sampling which results in strong predictive power without excessive feature pruning[21]. Moreover, the SHAP framework was introduced to overcome the intrinsic interpretability issue of tree-based algorithms[22]. It is important to state that SHAP values are measurements of the statistical predictive relationships in the model, and not

a direct biological causality[23]. Examples include the negative predictive correlation of maximum heart distance, which could indicate that the methods to extend the physical distance between the radiation field and the heart may be associated with decreased risk of pericarditis in this particular dataset. Also, the partial dependence analysis suggests that the heart V40 could have a non-linear threshold effect at about 2 percent. Although this finding cannot be used to set an absolute biological limit, it offers a quantitative conjecture that can be used to shape future prospective dose-restraint optimizations.

Regarding the study design, it is crucial to address how the class imbalance was managed and why models were simplified. To counter the uneven distribution of outcomes, we strictly applied the SMOTE oversampling technique only within the training group[24]. It is critically essential to avoid data leakage and make sure the evaluation metrics on the test set are indicative of actual generalization capability[25]. We developed a logistic regression-based nomogram with the highest-ranked features selected by SHAP to enable clinical translation[26]. Although translating a complex non-linear model into a linear nomogram involves a degree of simplification and naturally inherits some multicollinearity among the dosimetric variables, its main aim is pragmatic bedside risk stratification, as opposed to causal inference. The Bootstrap internal validation and Hosmer-Lemeshow test showed that the nomogram retained a satisfactory discriminative ability and goodness-of-fit, and provided a reasonable and clinically acceptable balance between the complexity of the algorithm and its practical use.

Although these are encouraging findings, this study has two major limitations. To begin with, the study is a retrospective single-center study, and thus the data can have inherent selection bias and institution-specific treatment patterns. It is not clear how the proposed models can be generalized to a wider population and additional external validation in more independent centers is necessary to assess their wider applicability. Second, the primary outcome of pericarditis risk was derived from NTCP model predictions, as opposed to longitudinal clinical follow-up data. This methodological option can be subject to theoretical variations of actual clinical frequency, making it necessary to include future research that will entail the use of prospectively acquired data on clinical outcomes.

5. Conclusion

In conclusion, this study developed an interpretable Random Forest model and a clinical nomogram to predict the risk of radiation-induced pericarditis in breast cancer patients. Our findings suggest that high-dose cardiac dosimetric parameters, especially heart V40 and heart V30, as well as spatial characteristics, such as maximum heart distance, are strongly correlated with the estimated risk. The incorporation of a SHAP framework offered useful model transparency, which emphasized the possibility of non-linear dose-response relationships. Moreover, the developed nomogram showed good discrimination and calibration, which suggests that it might be used as a viable bedside tool to offer personalized risk stratification. Nevertheless, the study has a retrospective and single-center design and is based on NTCP-derived outcomes, therefore, these results should be interpreted with caution. Multicenter prospective studies with the inclusion of long-term clinical follow-ups in the future are

justified in order to determine further the generalizability and clinical applicability of these predictive tools.

Declarations

Author Contributions: Yongkai Lu: Writing – original draft, Methodology, Validation, Software, Project administration, Methodology, Data curation, Conceptualization. **Lei Xu:** Writing – review & editing, Methodology, Conceptualization, Validation. **Jian Zhang:** Methodology, Conceptualization, Validation acquisition. **Rongze Ma:** Visualization, Conceptualization, Validation. **Yao Wang:** Visualization, Software. **Shanshan Gao:** Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Funding: This program was supported by the Natural Science Basic Research Program of Shaanxi (2025JC-YBQN-1093).

Ethics Statement: This study was approved by the Ethics Committee of the First Affiliated Hospital of Xi'an Jiaotong University. The research was conducted in accordance with the principles of the Declaration of Helsinki. Written informed consent was obtained from all participants prior to their inclusion in the study.

Clinical Trial Number: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to confidentiality.

Conflicts of Interest: The authors declare no competing financial interests.

References

1. Liang, Y. et al. The emerging roles of metabolism in the crosstalk between breast cancer cells and tumor-associated macrophages. *Int. J. Biol. Sci.* **19**, 4915–4930 (2023).
2. Luo, N., Zhang, K., Li, X., Hu, Y. & Guo, L. Tanshinone IIA destabilizes SLC7A11 by regulating PIAS4-mediated SUMOylation of SLC7A11 through KDM1A, and promotes ferroptosis in breast cancer. *J. Adv. Res.* **69**, 313–327 (2025).
3. Liu, Q. et al. Recent death early warning value of ECG changes in patients with NSCLC. *Med. (Baltim).* **102**, e35698 (2023).
4. Diaz-Gavela, A. A. et al. Breast Radiotherapy-Related Cardiotoxicity. When, How, Why. Risk Prevention and Control Strategies. *Cancers (Basel)* ;13. (2021).
5. Liu, X. et al. Analysis of the factors affecting the safety of robotic stereotactic body radiation therapy for hepatocellular carcinoma patients. *Onco Targets Ther.* **10**, 5289–5295 (2017).

6. Wu, W. et al. Dosimetric analysis of half-field-based VMAT with the deep inspiration breath-hold technique for left breast cancer patients following breast-conserving surgery. *Front. Oncol.* **14**, 1418723 (2024).
7. Talwar, A., Lopez-Olivo, M. A., Huang, Y., Ying, L. & Aparasu, R. R. Performance of advanced machine learning algorithms overlogistic regression in predicting hospital readmissions: A meta-analysis. *Explor. Res. Clin. Soc. Pharm.* **11**, 100317 (2023).
8. Tong, R., Zhu, Z. & Ling, J. Comparison of linear and non-linear machine learning models for time-dependent readmission or mortality prediction among hospitalized heart failure patients. *Heliyon* **9**, e16068 (2023).
9. Li, Y. et al. Observation on prefrontal cortex activation in patients with Parkinson's disease: a fNIRS study. *Front. Aging Neurosci.* **17**, 1560315 (2025).
10. Topaz, M., Peltonen, L. M. & Zhang, Z. Beyond human ears: navigating the uncharted risks of AI scribes in clinical practice. *NPJ Digit. Med.* **8**, 569 (2025).
11. Zhang, L. et al. Interpretable machine learning for coronary heart disease risk stratification in patients with carotid atherosclerosis: A retrospective cross-sectional study. *Med. (Baltim).* **105**, e47203 (2026).
12. Bai, T. & Wu, C. Association of cardiovascular disease on cancer: observational and mendelian randomization analyses. *Sci. Rep.* **14**, 28465 (2024).
13. Eber, J. et al. [Impact of thoracic irradiation on cardiac structures]. *Cancer Radiother*. **26**, 526–536 (2022).
14. Jin, Z., Sun, X., Zhou, C., Yang, H. & Zhou, S. Cardiac substructures dosimetric predicts cardiac toxicity and prognosis in esophageal squamous cell cancer treated by radiotherapy. *Neoplasia* **48**, 100969 (2024).
15. Shao, Y. et al. Cardiac substructure-guided radiotherapy planning for esophageal cancer: a novel strategy for cardiotoxicity reduction. *Eur. J. Med. Res.* **31**, 205 (2026).
16. Ji, M. C. et al. Dosimetric risk factors for radiation esophagitis in patients with breast cancer following regional nodal radiation. *World J. Clin. Cases*. **12**, 2995–3003 (2024).
17. McWilliam, A., Abravan, A., Banfill, K., Faivre-Finn, C. & van Herk, M. Demystifying the Results of RTOG 0617: Identification of Dose Sensitive Cardiac Subregions Associated With Overall Survival. *J. Thorac. Oncol.* **18**, 599–607 (2023).
18. Li, Y. et al. Innovative regression model-based decision support tool for optimizing radiotherapy techniques in thoracic esophageal cancer. *Front. Oncol.* **14**, 1370293 (2024).
19. Moisoglou, I. et al. Association Between Workplace Gaslighting and Perceived Quality of Care, Patient Safety and Quiet Quitting: A Cross-Sectional Study Among Nurses in Greece. *Healthc. (Basel)* ;14. (2026).
20. Yuksel, G., Ardic, M. L., Sumbul, H. E. & Koc, M. SCORE2 and SCORE2-OP Assessment in the Predicting of Cardiovascular Diseases and AF Recurrence in Hypertensive AF Patients Who Underwent Catheter Ablation. *J. Clin. Med.* ;**15**. (2025).

21. Castronuovo, G., Favia, G., Telesca, V. & Vammacigno, A. Analyzing the Interactions between Environmental Parameters and Cardiovascular Diseases Using Random Forest and SHAP Algorithms. *Rev. Cardiovasc. Med.* **24**, 330 (2023).
22. Zhang, Y., Weng, Y. & Lund, J. Applications of Explainable Artificial Intelligence in Diagnosis and Surgery. *Diagnostics (Basel)* ;12. (2022).
23. Baptista, M. L., Goebel, K. & Henriques, E. M. P. Relation between prognostics predictor evaluation metrics and local interpretability SHAP values. *Artif. Intell.* ;306. (2022).
24. Jyothi, B. & Gladence, L. M. Modified AlexNet for accurate lung adenocarcinoma prediction with enhanced UMAP and optimized feature selection. *Discover Comput.* ;**29**. (2026).
25. Lu, R. et al. Development of a prediction model for lower limb deformity in patients with hereditary multiple exostoses based on interpretable models and nomogram. *BMC Med. Inf. Decis. Mak.* **25**, 355 (2025).
26. Shi, W. et al. Predicting psychological distress in informal caregivers of older people with dementia: development and interpretation of a nomogram and XGBoost model. *BMC Nurs.* **24**, 1311 (2025).

Figures

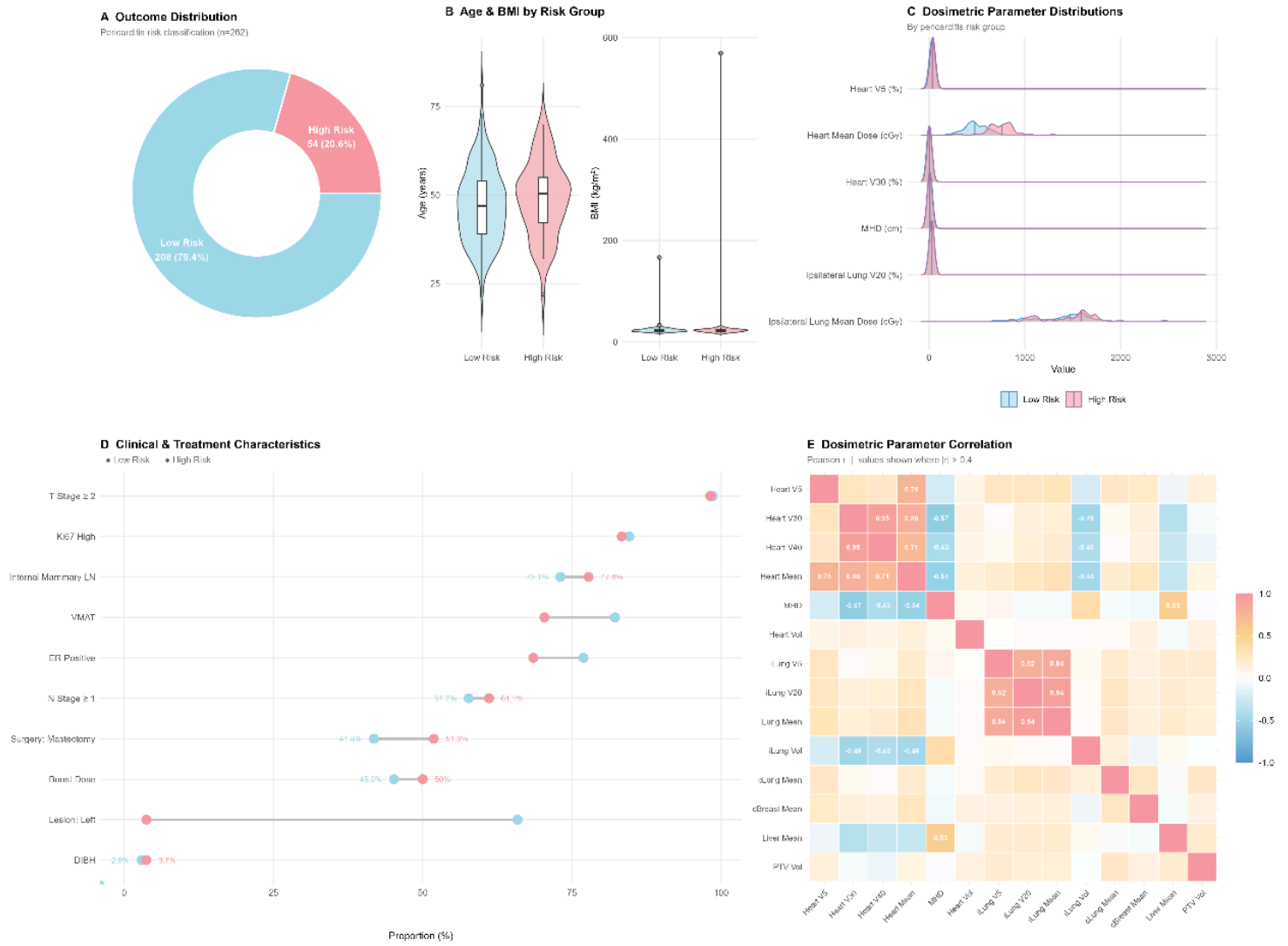


Figure 1

Baseline clinical and dosimetric characteristics of the study population.(A) Distribution of radiation-induced pericarditis (RIP) risk among the included patients. (B) Violin plots illustrating the distribution of age and BMI between high- and low-risk groups. (C) Density plots displaying the distributions of key dosimetric parameters. (D) Dot plot showing the proportions of clinical and treatment characteristics stratified by risk group. (E) Heatmap representing the correlation matrix among various dosimetric parameters.



Figure 2

Feature selection and statistical evaluation. (A) Forest plot of univariate logistic regression for candidate variables. (B) Bubble plot providing an overview of significant features. (C) Receiver Operating Characteristic (ROC) curves suggesting optimal cutoff points for continuous variables. (D) Variance Inflation Factor (VIF) bar chart evaluating multicollinearity among features. (E-F) Cross-validation and coefficient path plots for feature selection using the LASSO regression model. (G) Boxplots showing the

distribution of top significant features between risk groups. (H) Bubble plot illustrating the effect sizes of significant features stratified by risk status.

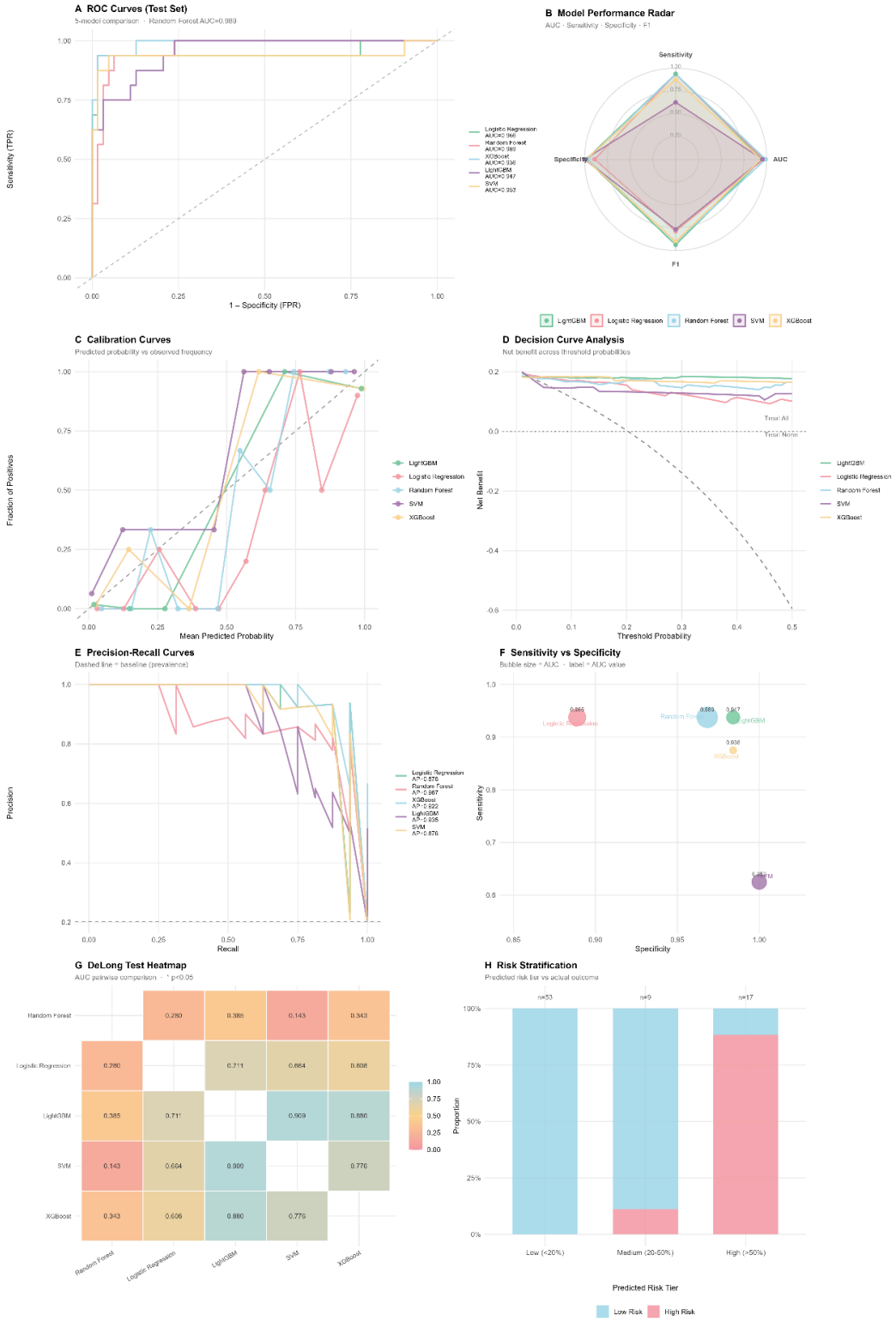


Figure 3

Performance evaluation of different machine learning models. (A) ROC curves of the candidate models in the test set. (B) Radar chart summarizing comprehensive evaluation metrics (AUC, sensitivity,

specificity, etc.). (C) Calibration curves evaluating the agreement between predicted probabilities and actual observations. (D) Decision Curve Analysis (DCA) assessing the potential clinical net benefit across threshold probabilities. (E) Precision-Recall (PR) curves. (F) Scatter plot of sensitivity versus specificity for model comparison. (G) Heatmap of DeLong test p-values comparing the AUCs among different models. (H) Bar chart displaying the results of risk stratification based on predicted probabilities.

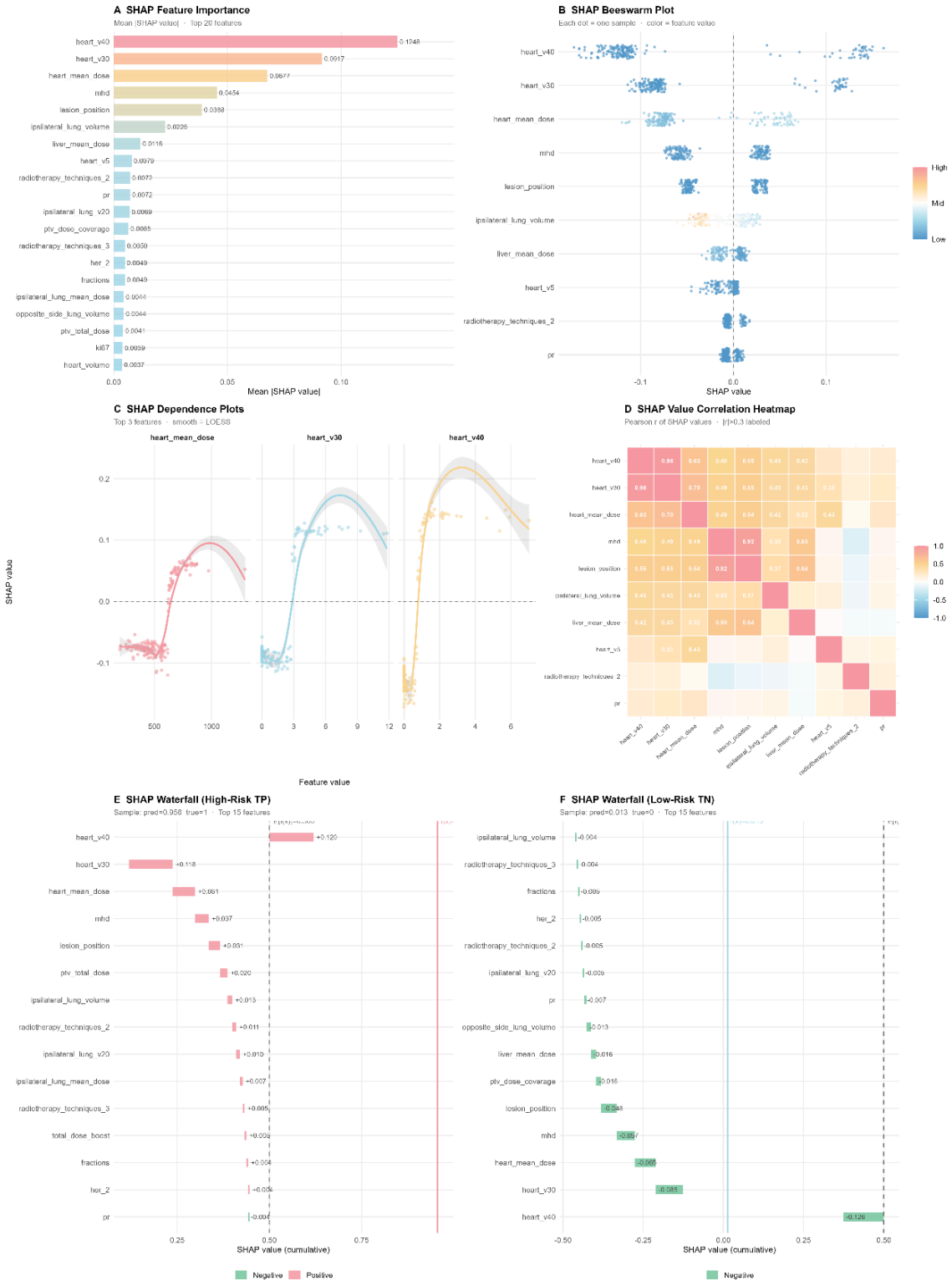


Figure 4

SHAP (SHapley Additive exPlanations) interpretability analysis of the optimal model. (A) Bar chart ranking the global feature importance based on mean absolute SHAP values. (B) Beeswarm plot illustrating the distribution of SHAP values and the directional impact of feature values on the model output. (C) SHAP dependence plots showing the marginal effect of key dosimetric features. (D) Heatmap of SHAP value correlations among top features. (E-F) SHAP waterfall plots providing local explanations for (E) a representative high-risk patient (True Positive) and (F) a representative low-risk patient (True Negative).

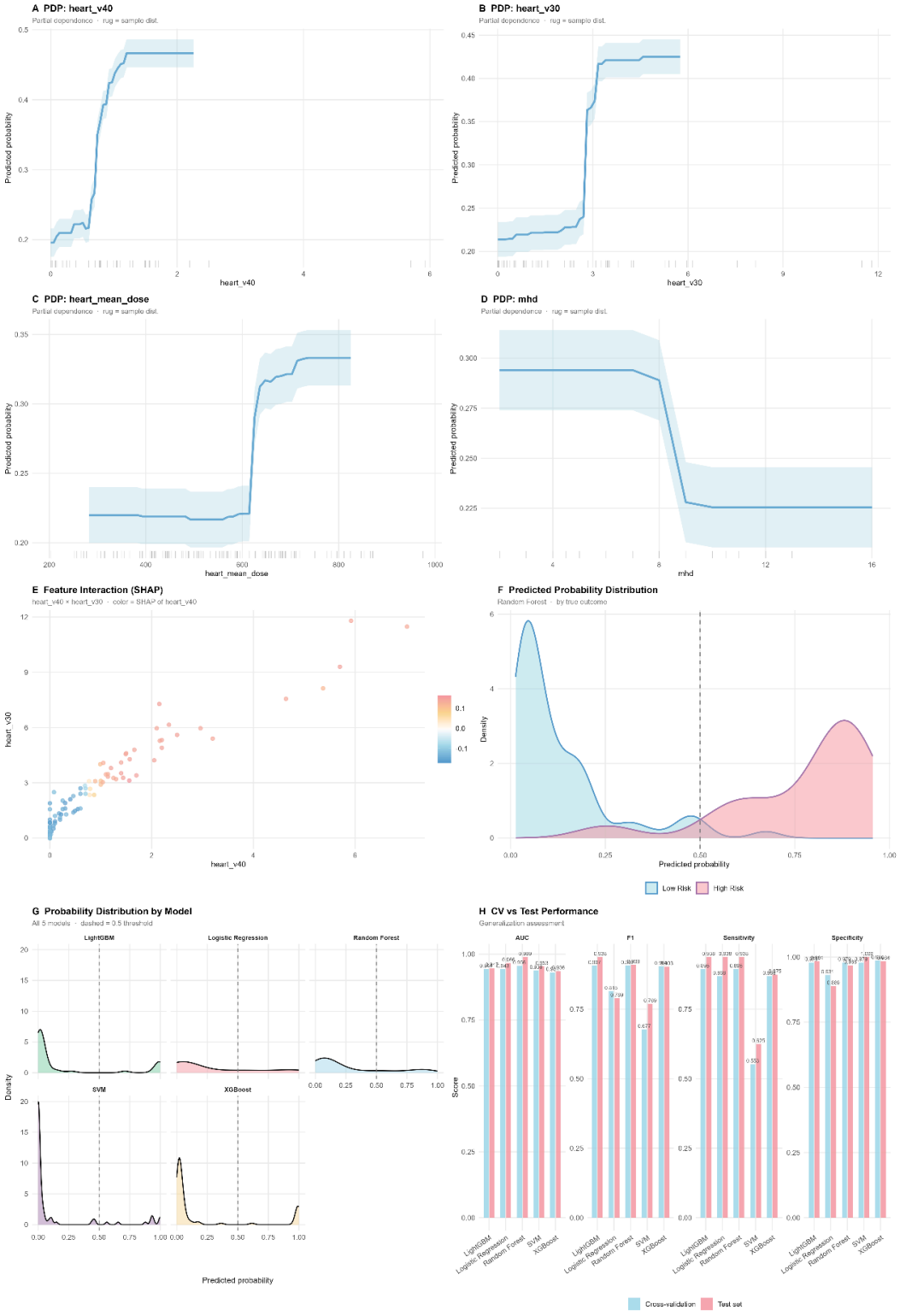


Figure 5

Partial dependence and predicted probability distribution analysis. (A-D) Partial Dependence Plots (PDP) demonstrating the marginal effects of top features (heart_v40, heart_v30, heart_mean_dose, and mhd) on the predicted probability of pericarditis. (E) Scatter plot showing the SHAP feature interaction between heart_v40 and heart_v30. (F) Density plot of predicted probabilities stratified by true outcomes. (G) Density distribution of predicted probabilities across the five evaluated models. (H) Grouped bar

chart comparing comprehensive performance metrics (AUC, F1, Sensitivity, Specificity) between cross-validation and test sets.

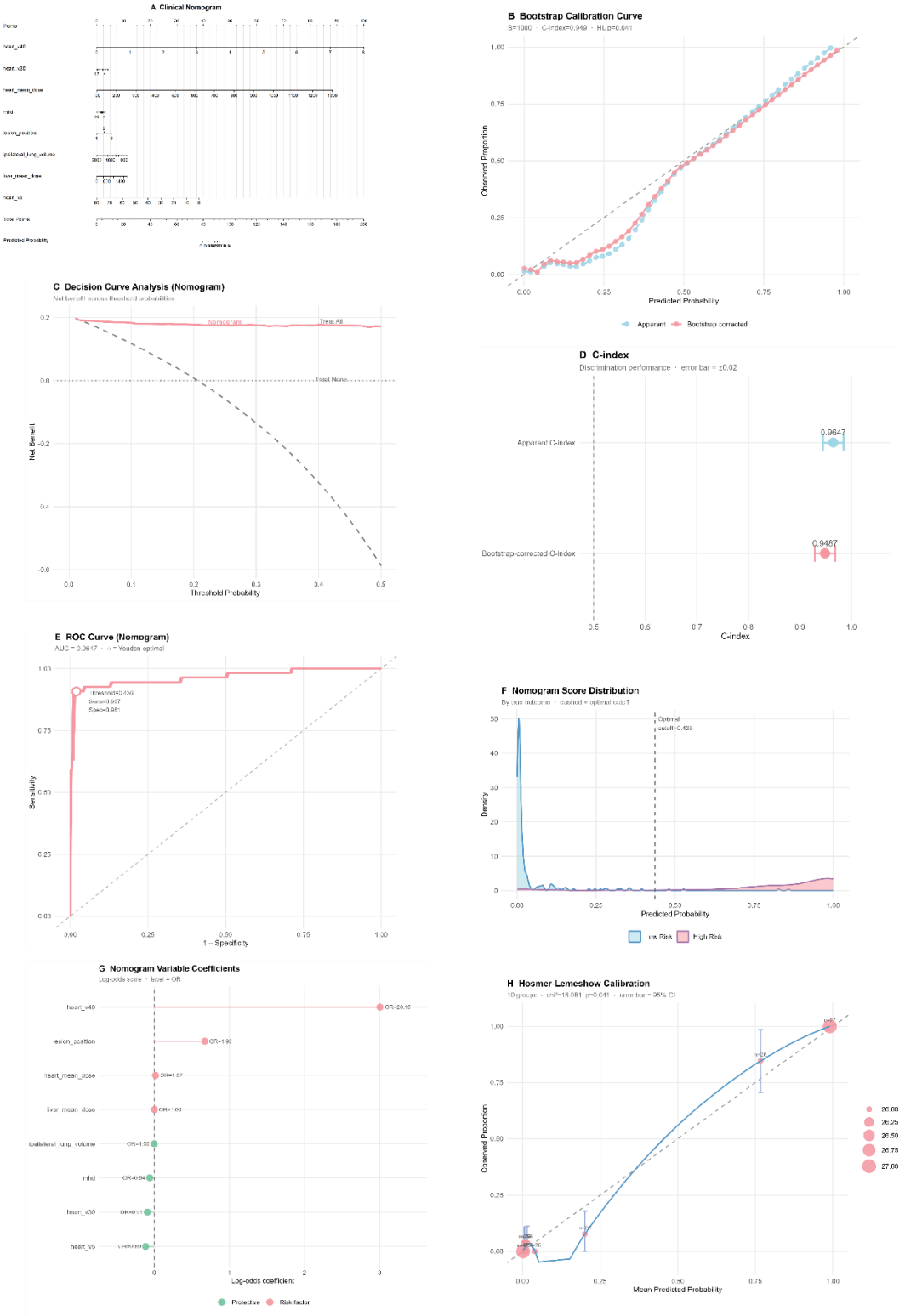


Figure 6

Construction and evaluation of the clinical nomogram. (A) The clinical nomogram developed for predicting the probability of radiation-induced pericarditis based on the selected features. (B) Bootstrap

calibration curve displaying the agreement between apparent and bootstrap-corrected predicted probabilities. (C) Decision Curve Analysis (DCA) presenting the clinical net benefit of the nomogram across different threshold probabilities. (D) Plot illustrating the apparent and bootstrap-corrected C-index values. (E) Receiver Operating Characteristic (ROC) curve of the nomogram model. (F) Density distribution plot of the predicted probabilities stratified by actual risk outcomes. (G) Forest plot showing the log-odds variable coefficients of the features incorporated into the nomogram. (H) Hosmer-Lemeshow calibration plot evaluating the goodness-of-fit across patient groups.