

Interpretable Machine Learning Using Normal Lung Radiomics and Dosiomics to Predict Radiation Pneumonitis After Lung Cancer Radiotherapy

Hao Qiu

Siyang Hospital

Qianjin Shi

Siyang Hospital

Fei Chen

Affiliated Hospital of Jiangsu University

Kang Shen

Siyang Hospital

Miao Wang

Siyang Hospital

Lei Ding

Siyang Hospital

Hu Chen

2982019188@qq.com

Siyang Hospital of Traditional Chinese Medicine

Article

Keywords: Lung cancer, Radiation pneumonitis, Radiomics, Dosiomics, Machine learning

Posted Date: September 29th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9913203/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

Objective

This study aimed to develop and interpret machine learning models based on radiomics and dosiomics features extracted from normal lung tissue after excluding the planning target volume for predicting grade ≥ 2 radiation pneumonitis within 6 months after radiotherapy.

Methods

A total of 232 patients with lung cancer who received radiotherapy between February 2023 and June 2025 were retrospectively enrolled and divided into a training set ($n = 162$) and a validation set ($n = 70$). The region of interest was defined as the whole lung excluding the planning target volume. Radiomics features were extracted from pretreatment contrast-enhanced simulation CT images, and dosiomics features were extracted from three-dimensional RT Dose files using the same normal lung region. After preprocessing, LASSO regression and random forest were used for feature selection, and their intersecting features were retained. Eleven machine learning models were trained using five-fold cross-validation and evaluated in the validation set. Model performance was assessed using AUC, classification metrics, confusion matrix, and decision curve analysis. Kernel SHAP was applied for global model interpretation.

Results

Among all patients, 97 developed grade ≥ 2 radiation pneumonitis, with an overall incidence of 41.8%. From 214 extracted features, 18 radiomics and dosiomics features were selected for modeling. Validation-set AUCs ranged from 0.725 to 0.812. LASSO and linear discriminant analysis achieved the highest AUCs, both 0.812; the 95% CI for LASSO was 0.708–0.917. The final LASSO model showed an accuracy of 74.3%, sensitivity of 58.1%, and specificity of 87.2%. SHAP analysis indicated that first-order, shape, and texture features jointly contributed to model prediction.

Conclusion

Radiomics and dosiomics features from normal lung tissue may provide complementary information for predicting grade ≥ 2 radiation pneumonitis after lung cancer radiotherapy. The LASSO model showed favorable discrimination and interpretability in the validation set, but external validation is needed to further assess its stability and generalizability.

Introduction

Lung cancer is one of the major indications for thoracic radiotherapy. With the widespread adoption of Intensity-Modulated Radiotherapy (IMRT) and Volumetric Modulated Arc Therapy (VMAT), target dose coverage and normal tissue sparing have been substantially improved; however, Radiation-Induced Lung Injury (RILI) remains an important factor affecting treatment safety^[1–3]. Radiation Pneumonitis (RP) is one of the most common adverse events following thoracic radiotherapy. In particular, grade 2 or higher RP is usually associated with clinically evident symptoms and may require medical intervention, thereby impairing patients' quality of life and potentially disrupting subsequent anticancer treatment^[4–5]. Therefore, identifying patients at high risk of RP before radiotherapy or during treatment planning is of considerable clinical value for optimizing individualized treatment strategies, refining follow-up management, and enabling early intervention.

At present, risk assessment for RP primarily relies on clinical factors and conventional dosimetric parameters^[6–7]. Clinical variables such as age, pre-existing pulmonary disease, and concurrent chemoradiotherapy may influence the risk of RP, while parameters derived from the Dose-Volume Histogram (DVH), including Mean Lung Dose (MLD) and the Lung Volume Receiving 20 Gy or More (V20), are commonly used in radiotherapy plan evaluation^[8]. Yan et al. developed a clinical prediction model for severe RP in patients with lung cancer receiving thoracic radiotherapy and reported that subclinical interstitial lung disease, Diffusing Capacity of the Lung for Carbon Monoxide (DLCO), and ipsilateral lung V20 were associated with the occurrence of severe RP^[9]. However, the development of RP varies substantially among individuals, and clinical factors alone may not fully explain differences in normal lung tissue susceptibility to radiation injury^[10–12]. Moreover, conventional DVH-based parameters reduce the three-dimensional dose distribution to a limited set of summary statistics, which may inadequately capture the spatial heterogeneity of dose deposition within lung tissue. This limitation may be particularly relevant in complex IMRT plans^[13–15].

Radiomics and dosiomics provide additional quantitative information for risk assessment of RP^[16–17]. Radiomics enables the extraction of intensity-, shape-, and texture-based features from pretreatment Computed Tomography (CT) images, thereby characterizing potential imaging heterogeneity within normal lung tissue. Dosiomics extends radiomic analytical approaches to three-dimensional dose distributions and may further capture variations in dose intensity, spatial configuration, and textural patterns^[18–19]. Jiang et al. extracted radiomic features from different Regions of Interest (ROIs) on treatment-planning CT images in patients with lung cancer undergoing concurrent chemoradiotherapy and developed prediction models incorporating dosimetric factors. Their findings suggested that radiomic features could complement conventional dosimetric parameters and improve the prediction of grade 2 or higher RP^[20]. Given that RP predominantly occurs in irradiated normal lung tissue, feature extraction from the whole-lung region after exclusion of the Planning Target Volume (PTV) may more closely reflect the underlying biological and dosimetric basis of RILI.

However, omics-derived features are typically high-dimensional and may exhibit substantial redundancy, while different Machine Learning (ML) models may vary in how they learn from and utilize these complex

feature sets ^[21-22]. Relying solely on predictive performance is therefore insufficient to fully elucidate the basis of model decision-making and may limit the interpretability and credibility required for clinical translation. Accordingly, integrating model explanation methods during predictive model development may help identify key predictive features and provide further insight into how individual variables contribute to patient-specific risk estimation.

Based on this rationale, the present study included patients undergoing radiotherapy for lung cancer and extracted radiomic and dosiomic features from normal lung tissue defined as the whole-lung region after exclusion of the PTVRP. Multiple ML models were developed and compared to predict the risk of grade 2 or higher RP. In addition, Shapley Additive Explanations (SHAP) were applied to interpret the final model and evaluate the contribution of major predictive features to model decision-making. This study aimed to provide a clinically relevant reference for individualized risk assessment of RP after radiotherapy for lung cancer.

Methods

Study population and outcome definition

Patient data were retrospectively collected from individuals who underwent radiotherapy for lung cancer at our institution between February 2023 and June 2025. The inclusion criteria were as follows: (1) pathological or clinical diagnosis of lung cancer; (2) receipt of thoracic radiotherapy with completion of the prescribed treatment course; (3) availability of complete simulation CT images, Radiotherapy Dose (RT Dose) files, and contouring information for the target volumes and lung tissue; and (4) documented follow-up records for RP within 6 months after completion of radiotherapy. The exclusion criteria were as follows: (1) missing imaging or dose files; (2) Region of Interest (ROI) delineation that did not meet the requirements for feature extraction; (3) insufficient follow-up data or inability to determine the RP grade; and (4) a history of prior thoracic radiotherapy. According to the study design, all eligible patients were divided into a training cohort and a validation cohort at an approximate ratio of 7:3. The training cohort was used for feature selection, model development, and parameter optimization, whereas the validation cohort was used exclusively for model performance evaluation. This study was approved by the Medical Ethics Committee of Siyang Hospital, approval number KS2026001. The requirement for written informed consent was waived by the Medical Ethics Committee of Siyang Hospital owing to the retrospective nature of the study.

The primary endpoint was the occurrence of grade 2 or higher RP within 6 months after completion of radiotherapy. RP was graded according to the Common Terminology Criteria for Adverse Events Version 5.0 (CTCAE 5.0) of the National Cancer Institute (NCI). The occurrence and grade of RP were determined through a comprehensive review of clinical symptoms, imaging findings, and follow-up medical records.

Image acquisition, treatment planning, and ROI segmentation

All patients underwent contrast-enhanced thoracic simulation CT before radiotherapy. Patients were positioned supine and immobilized using a thermoplastic mask during image acquisition. CT images were obtained with a slice thickness of 3 mm. The simulation CT images were used for radiotherapy target delineation, treatment planning, and subsequent radiomic feature extraction.

Radiotherapy plans were generated using the Monaco treatment planning system. All patients were treated with either Dynamic Multileaf Collimator (DMLC) technique or VMAT, with a prescribed dose of 56 Gy or 60 Gy. The PTVRP was delineated by radiation oncologists according to relevant clinical guidelines and institutional radiotherapy protocols. Whole-lung contours were initially generated using the RT-Mind automatic contouring system and were subsequently reviewed on a case-by-case basis by clinical physicians, with manual correction performed when necessary to ensure the consistency and accuracy of lung segmentation.

In this study, the ROI was defined as normal lung tissue obtained by excluding the PTVRP from the whole-lung contour. Specifically, the PTV was subtracted from the whole-lung segmentation to generate the normal lung ROI used for feature extraction. Radiomic features were extracted from the simulation CT images, whereas dosiomic features were extracted from the three-dimensional RT Dose files. Both feature sets were derived from the same normal lung ROI to ensure spatial consistency between imaging information and dose-distribution information.

Radiomics and dosiomics feature extraction

Radiomic features were extracted from contrast-enhanced simulation CT images acquired before radiotherapy, whereas dosiomic features were extracted from three-dimensional RT Dose files. For both feature types, the ROI was defined as normal lung tissue derived from the whole-lung contour after exclusion of the PTVRP, ensuring that imaging phenotypic information and spatial dose-distribution information were obtained from the same normal lung tissue region.

Before radiomic feature extraction, the CT images were first subjected to intensity truncation within a range of - 1250 to 250 Hounsfield Units (HU) to reduce the influence of extreme gray-level values on feature calculation. Intensity normalization was then performed using the Z-score method. After the RT Dose files were imported, the dose distribution was converted into a physical dose map in Gray (Gy) according to the dose scaling factor. The dose images were subsequently resampled using linear interpolation to match the spatial coordinates, slice thickness, and matrix size of the normal lung ROI.

All features were extracted using PyRadiomics. The extracted feature categories included first-order statistical features, shape features, Gray Level Co-occurrence Matrix (GLCM) features, Gray Level Run Length Matrix (GLRLM) features, Gray Level Size Zone Matrix (GLSZM) features, Gray Level Dependence Matrix (GLDM) features, and Neighboring Gray Tone Difference Matrix (NGTDM) features. Gray-level and dose discretization parameters were set according to the default configuration of PyRadiomics and were

kept consistent during radiomic and dosiomic feature extraction. In total, 214 radiomic and dosiomic features were extracted for subsequent feature selection and model development.

Feature preprocessing and selection

To minimize the influence of extreme values and differences in feature scale on model training, the extracted radiomic and dosiomic features were first preprocessed. For each feature, the mean and standard deviation were calculated using the training cohort. Values exceeding the range of mean $\pm$ 3 standard deviations were defined as extreme outliers and were treated as missing values. Missing values were subsequently imputed using Multiple Imputation (MI). After imputation, feature standardization was performed using the Z-score method. All preprocessing parameters were derived exclusively from the training cohort, and the validation cohort was transformed according to the same rules established in the training cohort.

Feature selection was performed exclusively in the training cohort, with no involvement of the validation cohort. First, Least Absolute Shrinkage and Selection Operator (LASSO) regression was applied for dimensionality reduction, and the optimal penalty parameter was determined through cross-validation. Features with non-zero coefficients were retained. A Random Forest (RF) model was then constructed to rank and select variables according to feature importance. To improve the stability of features entered into the predictive models, the intersection of features selected by LASSO regression and RF was used as the candidate feature set in this study.

A total of 18 radiomic and dosiomic features were ultimately retained for subsequent ML model development. Correlation analysis was further performed on these 18 features to evaluate potential collinearity among candidate predictors. The correlation threshold was set at $|r| > 0.90$.

Machine learning model development and performance evaluation

Based on the final set of 18 selected radiomic and dosiomic features, ML prediction models were developed in the training cohort and subsequently evaluated in the validation cohort. The occurrence of grade 2 or higher RP was defined as the positive outcome, whereas the absence of grade 2 or higher RP was defined as the negative outcome.

In this study, 11 ML models were developed and compared, including RF, Support Vector Machine (SVM), Logistic Regression (LR), K-Nearest Neighbors (KNN), Partial Least Squares (PLS), Gradient Boosting Machine (GBM), Neural Network (NN), Naïve Bayes (NB), Linear Discriminant Analysis (LDA), LASSO, and Adaptive Boosting (AdaBoost). During model training, five-fold cross-validation was used for hyperparameter tuning and internal evaluation. All models were trained exclusively on the training cohort, and the validation cohort was not involved in model training or parameter selection.

Model discrimination was assessed using the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC), with corresponding 95% Confidence Intervals (CIs) calculated for AUC values. A classification threshold of 0.5 was used to convert predicted probabilities into binary outcomes. Model performance was further evaluated by calculating accuracy, sensitivity, specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), F1 score, and the Youden index. Confusion matrices were used to present the classification results of each model in the validation cohort.

To assess the potential clinical utility of the models, Decision Curve Analysis (DCA) was performed to calculate the net benefit across a range of threshold probabilities and to compare model-based strategies with the “treat-all” and “treat-none” strategies. The final model was selected primarily according to the AUC in the validation cohort, while also considering model simplicity and interpretability. When multiple models showed comparable predictive performance, the model with a relatively simple structure and greater interpretability was preferentially selected for subsequent analyses.

SHAP-based model interpretability analysis

To improve the interpretability of model predictions, the Kernel Shapley Additive Explanations (Kernel SHAP) method was applied to the final model selected according to the predefined criteria. SHAP values were used to quantify the contribution of each feature to the model output, with the magnitude and direction of each value reflecting the degree and direction of its influence on the predicted risk.

Model interpretability was evaluated from a global perspective. First, the overall contribution of each feature was quantified using the mean absolute SHAP value, and a SHAP feature-importance bar plot was generated to illustrate the average impact of different features on model predictions. Second, a SHAP beeswarm plot was used to further examine the relationship between feature values and the direction of model predictions.

Statistical analysis

Radiomic and dosiomic feature extraction was performed using Python 3.11, whereas data preprocessing, statistical analysis, feature selection, model development, and SHAP analysis were conducted using R 4.5.2. Baseline patient characteristics were summarized using descriptive statistics. Age was reported as the median and range, and categorical variables were presented as counts and percentages. Before model construction, radiomic and dosiomic features underwent outlier processing, missing-value imputation, and Z-score standardization. Missing values were imputed using ML with the R package mice. Model discrimination was evaluated using the AUC of the ROC curve and its corresponding 95% Confidence Interval (CI). Classification performance was assessed using accuracy, sensitivity, specificity, PPV, NPV, F1 score, and the Youden index. DCA was used to evaluate the net

benefit of the model across different threshold probabilities. Model interpretability was assessed using the Kernel SHAP method.

Results

Baseline Patient Characteristics

A total of 232 patients who underwent radiotherapy for lung cancer were included in this study, comprising 162 patients in the training cohort and 70 patients in the validation cohort. Within 6 months after completion of radiotherapy, 97 patients developed grade 2 or higher RP, with an overall incidence of 41.8%.

In the training cohort, 66 patients developed grade 2 or higher RP, corresponding to an incidence of 40.7%; among them, 52 patients had grade 2 RP and 14 had grade 3 RP. In the validation cohort, 31 patients developed grade 2 or higher RP, with an incidence of 44.3%; among them, 24 patients had grade 2 RP and 7 had grade 3 RP. The outcome distribution was generally comparable between the training and validation cohorts.

Baseline patient characteristics are summarized in Table 1. The median age was 63 years in the training cohort and 64 years in the validation cohort, and both cohorts were predominantly composed of male patients. In both cohorts, the prescribed dose was mainly 60 Gy, and DMLC was the most commonly used treatment technique. Regarding pathological subtype, Small-Cell Lung Cancer (SCLC) accounted for the largest proportion in both cohorts. The proportions of patients receiving concurrent chemoradiotherapy were 64.8% in the training cohort and 64.3% in the validation cohort, respectively.

Table 1
Baseline characteristics of patients

Variable	Training set (n = 162)	Validation set (n = 70)
Age	63 (42–82)	64 (45–80)
Sex		
Male	119 (73.5%)	50 (71.4%)
Female	43 (26.5%)	20 (28.6%)
Radiation pneumonitis grade		
Grade 0	64 (39.5%)	26 (37.1%)
Grade 1	32 (19.8%)	13 (18.6%)
Grade 2	52 (32.1%)	24 (34.3%)
Grade 3	14 (8.6%)	7 (10.0%)
RP classification		
No grade $\geq$ 2 RP	96 (59.3%)	39 (55.7%)
Grade $\geq$ 2 RP	66 (40.7%)	31 (44.3%)
Prescription dose		
56 Gy	48 (29.6%)	21 (30.0%)
60 Gy	114 (70.4%)	49 (70.0%)
Pathology type		
Squamous cell carcinoma	45 (27.8%)	19 (27.1%)
Adenocarcinoma	39 (24.1%)	17 (24.3%)
Small cell lung cancer	78 (48.1%)	34 (48.6%)
Treatment technique		
DMLC	130 (80.2%)	56 (80.0%)
VMAT	32 (19.8%)	14 (20.0%)
Concurrent chemoradiotherapy		
Yes	105 (64.8%)	45 (64.3%)
No	57 (35.2%)	25 (35.7%)

Feature Selection and Correlation Analysis

A total of 214 radiomic and dosiomic features were extracted in this study. After feature preprocessing, feature selection was performed in the training cohort using LASSO regression and RF, and the intersection of features selected by both methods was used as the final variable set for subsequent model development. Ultimately, 18 features were retained, encompassing different categories of information, including first-order statistical features, shape features, and texture features.

To further evaluate the potential redundancy among the final model-input features, correlation analysis was performed for the 18 retained features, and a correlation heatmap was generated, as shown in Fig. 1. Overall, the correlations among these features were weak, with most feature pairs showing correlation coefficients close to 0. These findings suggested that no evident high collinearity was present among the final features included in the models.

Comparison of Predictive Performance Among Machine Learning Models

Based on the 18 selected radiomic and dosiomic features, 11 ML models were developed and compared in the training cohort, and their predictive performance was evaluated in the validation cohort. The AUC values of the models in the validation cohort ranged from 0.725 to 0.812, as shown in Fig. 2. The corresponding classification performance metrics are presented in Table 2.

Table 2
Classification performance of different machine learning models in the validation set

Model	Accuracy	Sensitivity	Specificity	PPV	NPV	F1 score	Youden index
Linear discriminant analysis	0.743	0.645	0.821	0.741	0.744	0.69	0.466
LASSO	0.743	0.581	0.872	0.783	0.723	0.667	0.452
Logistic regression	0.729	0.613	0.821	0.731	0.727	0.667	0.433
Support vector machine	0.729	0.613	0.821	0.731	0.727	0.667	0.433
Naive Bayes	0.714	0.581	0.821	0.720	0.711	0.643	0.401
Neural network	0.743	0.677	0.795	0.724	0.756	0.7	0.472
Partial least squares	0.743	0.613	0.846	0.76	0.733	0.679	0.459
Random forest	0.686	0.419	0.897	0.765	0.660	0.542	0.317
Gradient boosting	0.686	0.516	0.821	0.696	0.681	0.593	0.337
Adaptive boosting	0.686	0.581	0.769	0.667	0.698	0.621	0.35
K-nearest neighbors	0.700	0.645	0.744	0.667	0.725	0.656	0.389
PPV, positive predictive value; NPV, negative predictive value; F1 score, harmonic mean of precision and sensitivity; Youden index, sensitivity + specificity – 1. All classification metrics were calculated using a probability threshold of 0.5.							

In the validation cohort, LASSO and LDA achieved the highest AUC values, both reaching 0.812, with 95% CIs of 0.708–0.917 and 0.710–0.915, respectively. The AUC values of LR, SVM, and NB were 0.808, 0.807, and 0.806, respectively, indicating predictive performance comparable to that of the best-performing models. KNN showed the lowest AUC, with a value of 0.725.

Further comparison of classification performance metrics indicated a degree of trade-off between sensitivity and specificity across different models. The NN model showed relatively high sensitivity, F1 score, and Youden index, whereas the LASSO model achieved the highest PPV and demonstrated relatively high specificity. Considering the validation-cohort AUC, model simplicity, and interpretability, LASSO was selected as the final predictive model for subsequent analyses.

Classification Performance and Decision Curve Analysis of the Final Model in the Validation Cohort

The confusion matrix in the validation cohort showed that the LASSO model correctly classified 52 patients, yielding an accuracy of 74.3%. Specifically, 34 patients who did not develop grade 2 or higher RP were correctly predicted as negative, while 18 patients who developed grade 2 or higher RP were correctly predicted as positive. In addition, the model generated 5 false-positive and 13 false-negative results (Fig. 3A). In the validation cohort, the LASSO model achieved a sensitivity of 58.1% and a specificity of 87.2%, indicating relatively high specificity, although its ability to identify positive cases remained comparatively limited.

DCA in the validation cohort showed that the net benefit varied among different ML models across threshold probabilities (Fig. 3B). Within certain ranges of threshold probability, the curves of several models, including the final model, lay above those of the “treat-all” and “treat-none” strategies, suggesting that these models may provide potential benefit for clinical decision-making.

Model Interpretability Analysis Based on Shapley Additive Explanations

Global interpretability analysis based on SHAP showed that different radiomic and dosiomic features contributed unequally to the output of the final LASSO model (Fig. 4). Ranked by the mean absolute SHAP value, the most influential features mainly included first-order statistical features, shape features, and texture features, with first-order statistical and shape-related features accounting for a large proportion of the top-ranked predictors.

The SHAP beeswarm plot demonstrated distinct distribution patterns between feature values and SHAP directions across different feature categories. High values of several first-order statistical and shape-related features were predominantly located in the positive SHAP value region, whereas high values of certain texture features and homogeneity-related features were mainly distributed in the negative SHAP value region. Overall, the major contributing features of the final model spanned multiple feature categories, suggesting that the model predictions may be jointly influenced by multidimensional information derived from both the imaging phenotype of normal lung tissue and the spatial distribution of radiation dose.

Discussion

In this study, radiomic and dosiomic features were extracted from normal lung tissue defined as the whole-lung region after exclusion of the PTVRP in patients receiving radiotherapy for lung cancer. These features were used to develop ML models for predicting the risk of grade 2 or higher RP within 6 months after completion of radiotherapy. The results showed that multiple models based on the selected omics features demonstrated acceptable discriminatory performance in the validation cohort, among which the LASSO model achieved a favorable balance between AUC, model simplicity, and interpretability. Further analysis using SHAP indicated that the model predictions were not driven by a single feature type, but were jointly influenced by multiple categories of omics-derived information, including first-order statistical features, shape-related features, and texture features.

The development of RP may be associated with the baseline condition of individual lung tissue, the spatial distribution of RT Dose, and treatment-related factors [23–25]. Bi et al. investigated dosimetric predictors of RP in a multi-institutional lung cancer cohort of patients previously treated with immunotherapy and found that MLD was a key DVH parameter associated with RP risk, supporting the close relationship between RT Dose distribution and the occurrence of RP [26]. Jang et al. reported that dosimetric factors, including MLD, V20, Lung Volume Receiving 30 Gy or More (V30), and Lung Volume Receiving 40 Gy or More (V40), were associated with grade 2 or higher RP in patients with Non-Small Cell Lung Cancer (NSCLC) receiving concurrent chemoradiotherapy combined with Immune Checkpoint Inhibitors (ICIs), suggesting that both dose distribution and the immunotherapy context may influence RP risk [27].

However, conventional risk assessment mainly relies on clinical variables and DVH parameters, which remain limited in their ability to characterize the spatial heterogeneity of lung tissue. Radiomics can extract quantitative phenotypic information from pretreatment CT images of normal lung tissue, whereas dosiomics can further characterize the spatial features of three-dimensional dose distributions [28–30]. Previous studies have suggested the potential value of radiomics and dosiomics in RP prediction [31–33]. Huang et al. extracted dosiomic and deep learning radiomic features from three-dimensional dose distributions in patients with NSCLC undergoing Stereotactic Body Radiotherapy (SBRT) and developed a predictive model for RP, indicating that spatial dose information and omics-derived features may improve RP risk prediction [34]. Ai et al. integrated Positron Emission Tomography/Computed Tomography (PET/CT) radiomics, dosiomics, and deep learning features to predict RP in patients with lung cancer treated with VMAT, demonstrating the predictive potential of multimodal omics information [35]. The findings of the present study are broadly consistent with these previous reports, suggesting that imaging phenotypes and spatial dose-distribution information derived from normal lung regions may provide complementary information for RP risk prediction.

Unlike conventional risk assessment approaches based primarily on whole-lung DVH parameters, the present study restricted feature extraction to normal lung tissue after exclusion of the PTVRP from the whole-lung contour. This strategy may help reduce the influence of target-related information on model development, allowing the model to focus more specifically on the baseline imaging characteristics and dose-distribution patterns of irradiated normal lung tissue [36–37]. In the present results, no evident high collinearity was observed among the final model-input features, suggesting that the selected features retained, to some extent, complementary information across different dimensions.

In terms of model performance, the LASSO model achieved a relatively high AUC in the validation cohort; however, its sensitivity and specificity were not fully balanced. The validation-cohort confusion matrix showed that, at a classification threshold of 0.5, the model had high specificity but relatively limited sensitivity, suggesting that it was more inclined to identify patients who did not develop grade 2 or higher RP. Future studies may further explore the classification performance of the model across different risk thresholds to improve the identification of high-risk patients. DCA demonstrated a positive net benefit

within certain threshold probability ranges, indicating that the model predictions may have potential clinical reference value under specific risk thresholds. Nevertheless, these findings require further confirmation through external validation and prospective studies.

SHAP analysis provided further insight into the model predictions. The features with relatively high contributions were mainly distributed across first-order statistical, shape-related, and texture feature categories, suggesting that the model may simultaneously incorporate information on the overall intensity distribution, spatial morphology, and local heterogeneity of the normal lung region. The relatively high ranking of first-order statistical and shape-related features may be associated with the irradiated extent of normal lung tissue and the spatial characteristics of dose distribution, whereas texture features may provide complementary information reflecting heterogeneity in lung imaging phenotypes or dose-distribution patterns. However, SHAP results indicate the contribution of individual features to model output and do not establish a causal relationship between these features and RP. Therefore, the clinical interpretation of omics-derived features should remain cautious and requires further validation in larger samples, external cohorts, and reproducibility analyses.

This study has several limitations. First, this was a single-center retrospective study with a relatively limited sample size. Although a validation cohort was established, independent external validation was not performed, and the generalizability of the model requires further evaluation. Second, clinical variables were not incorporated into the final model, and no direct comparison model based on conventional DVH parameters was constructed. Therefore, the incremental value of omics-derived features over clinical factors or routine dosimetric indicators remains unclear. Third, grade 2 or higher RP was analyzed as a binary endpoint, and the predictive performance for different severity grades of RP was not further assessed. Fourth, radiomic and dosiomic features may be influenced by image acquisition protocols, dose grids, ROI delineation, and preprocessing parameters; therefore, standardization and stability validation are needed in future studies. Finally, although SHAP analysis improved model interpretability, its explanations remain model-based and should not be directly interpreted as biological mechanisms.

In conclusion, this study developed an interpretable ML model based on radiomic and dosiomic features extracted from normal lung tissue to predict the risk of grade 2 or higher RP after radiotherapy for lung cancer. The findings suggest that integrated omics-derived features may provide useful information for RP risk prediction. The LASSO model demonstrated favorable discriminatory performance in the validation cohort, and SHAP analysis revealed the contribution of multiple feature categories to model predictions. Further multicenter studies with independent external validation are warranted to evaluate the stability and clinical applicability of the proposed model.

Conclusion

This study extracted radiomic and dosiomic features from normal lung tissue, defined as the whole-lung region after exclusion of the PTVRP, in patients undergoing radiotherapy for lung cancer, and developed

ML models to predict the risk of grade 2 or higher RP within 6 months after radiotherapy. The results indicated that the selected omics-derived features provided useful information for RP risk prediction. Among the evaluated models, the LASSO model demonstrated favorable discriminatory performance in the validation cohort and offered certain advantages in terms of model simplicity and interpretability.

SHAP analysis showed that the predictions of the final model were jointly influenced by multiple categories of omics-derived information, including first-order statistical features, shape-related features, and texture features, suggesting that the model did not rely on a single feature source for risk estimation. Overall, the interpretable ML model based on radiomic and dosiomic features extracted from normal lung tissue may provide a useful reference for individualized risk assessment of RP after radiotherapy for lung cancer. Further studies with larger sample sizes and independent external validation cohorts are needed to evaluate the stability and generalizability of the model.

Abbreviations

Abbreviation

Full term (English)

AdaBoost

Adaptive Boosting

AUC

Area Under the Curve

CI

Confidence Interval

CT

Computed Tomography

CTCAE 5.0

Common Terminology Criteria for Adverse Events Version 5.0

DCA

Decision Curve Analysis

DLCO

Diffusing Capacity of the Lung for Carbon Monoxide

DMLC

Dynamic Multileaf Collimator

DVH

Dose-Volume Histogram

GBM

Gradient Boosting Machine

GLDM

Gray Level Dependence Matrix

GLCM

Gray Level Co-occurrence Matrix

GLRLM
Gray Level Run Length Matrix
GLSZM
Gray Level Size Zone Matrix
Gy
Gray
HU
Hounsfield Units
ICIs
Immune Checkpoint Inhibitors
IMRT
Intensity-Modulated Radiotherapy
KNN
K-Nearest Neighbors
LASSO
Least Absolute Shrinkage and Selection Operator
LDA
Linear Discriminant Analysis
LR
Logistic Regression
MI
Multiple Imputation
ML
Machine Learning
MLD
Mean Lung Dose
NB
Naïve Bayes
NCI
National Cancer Institute
NGTDM
Neighboring Gray Tone Difference Matrix
NN
Neural Network
NPV
Negative Predictive Value
NSCLC
Non-Small Cell Lung Cancer
PET/CT
Positron Emission Tomography/Computed Tomography

PLS
Partial Least Squares
PPV
Positive Predictive Value
PTV
Planning Target Volume
RF
Random Forest
RILI
Radiation-Induced Lung Injury
ROI
Region of Interest
ROC
Receiver Operating Characteristic
RP
Radiation Pneumonitis
RT Dose
Radiotherapy Dose
SBRT
Stereotactic Body Radiotherapy
SCLC
Small-Cell Lung Cancer
SHAP
Shapley Additive Explanations
SVM
Support Vector Machine
V20
Lung Volume Receiving 20 Gy or More
V30
Lung Volume Receiving 30 Gy or More
V40
Lung Volume Receiving 40 Gy or More
VMAT
Volumetric Modulated Arc Therapy

Declarations

Competing Interests:

The authors declare they have no conflict of interest.

Ethics approval and consent to participate

All research procedures involving human participants adhered to the 1964 Declaration of Helsinki and its subsequent amendments or equivalent ethical standards. This study was approved by the Medical Ethics Committee of Siyang Hospital, approval number KS2026001. The requirement for written informed consent was waived by the Medical Ethics Committee of Siyang Hospital owing to the retrospective nature of the study.

Consent to participate:

Not applicable. Consent to publish: Not applicable.

Declarations

Generative AI and AI-assisted technologies were NOT used in the preparation of this work

Funding:

This study was jointly funded by Siyang Hospital and the Suqian Municipal Health Commission, as part of the 2024 Suqian Municipal Health Commission Medical Research Project (ZD202409). Zhenjiang science and technology (social development)projectsNo.SH2024011

Author Contribution

Conceptualization: Hao Qiu, Qianjin Shi, Hu Chen; Data curation: Fei Chen, Miao Wang, Lei Ding; Formal analysis: Hao Qiu; Methodology: Hao Qiu, Kang Shen; Software: Kang Shen; Supervision: Hu Chen, Qianjin Shi; Writing - original draft: Hao Qiu; Writing - review & editing: Qianjin Shi, Fei Chen, Kang Shen, Miao Wang, Lei Ding, Hu Chen. All authors have read and approved the manuscript.

Acknowledgements

None.

Data Availability

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

References

1. Simone, C. B. 2 et al. ASTRO Radiation Therapy Summary of the ASCO Guideline on Management of Stage III Non-Small Cell Lung Cancer. *Pract. Radiat. Oncol.* **13** (3), 195–202. 10.1016/j.prro.2023.01.005 (2023).
2. Li, C. et al. Dosimetric comparison of four radiotherapy techniques for stage III non-small cell lung cancer. *Oncol. Lett.* **26** (2), 347. 10.3892/ol.2023.13933 (2023).
3. Wang, S., Xu, D., Xiao, L., Liu, B. & Yuan, X. Radiation-induced lung injury: from mechanism to prognosis and drug therapy. *Radiat. Oncol.* **20** (1), 39. 10.1186/s13014-025-02617-8 (2025).
4. Shintani, T. et al. Incidence and Risk Factors of Symptomatic Radiation Pneumonitis in Non-Small-Cell Lung Cancer Patients Treated with Concurrent Chemoradiotherapy and Consolidation Durvalumab. *Clin. Lung Cancer.* **22** (5), 401–410. 10.1016/j.clcc.2021.01.017 (2021).
5. Kuipers, M. E., van Doorn-Wink, K. C. J., Hiemstra, P. S. & Slat, A. M. Predicting Radiation-Induced Lung Injury in Patients With Lung Cancer: Challenges and Opportunities. *Int. J. Radiat. Oncol. Biol. Phys.* **118** (3), 639–649. 10.1016/j.ijrobp.2023.10.044 (2024).
6. Huang, B. T., Lin, P. X., Wang, Y. & Luo, L. M. Developing a Prediction Model for Radiation Pneumonitis in Lung Cancer Patients Treated With Stereotactic Body Radiation Therapy Combined With Clinical, Dosimetric Factors, and Laboratory Biomarkers. *Clin. Lung Cancer.* **24** (8), e323. e331.e2 (2023).
7. Kita, N. et al. Clinical and dosimetric factors for symptomatic radiation pneumonitis after stereotactic body radiotherapy for early-stage non-small cell lung cancer. *Clin. Transl Radiat. Oncol.* **41**, 100648. 10.1016/j.ctro.2023.100648 (2023).
8. Bensenane, R. et al. Optimizing lung cancer radiation therapy: A systematic review of multifactorial risk assessment for radiation-induced lung toxicity. *Cancer Treat. Rev.* **124**, 102684. 10.1016/j.ctrv.2024.102684 (2024).
9. Yan, Y. et al. Clinical predictors of severe radiation pneumonitis in patients undergoing thoracic radiotherapy for lung cancer. *Transl Lung Cancer Res.* **13** (5), 1069–1083. 10.21037/tlcr-24-328 (2024).
10. Liu, X., Shao, C. & Fu, J. Promising Biomarkers of Radiation-Induced Lung Injury: A Review. *Biomedicines* **9** (9), 1181. 10.3390/biomedicines9091181 (2021).
11. Ma, L. et al. A Study on the Radiosensitivity of Radiation-Induced Lung Injury at the Acute Phase Based on Single-Cell Transcriptomics. *Front. Immunol.* **13**, 941976. 10.3389/fimmu.2022.941976 (2022).
12. Bourbonne, V. et al. Combination of Radiomics Features and Functional Radiosensitivity Enhances Prediction of Acute Pulmonary Toxicity in a Prospective Validation Cohort of Patients with a Locally Advanced Lung Cancer Treated with VMAT-Radiotherapy. *J. Pers. Med.* **12** (11), 1926. 10.3390/jpm12111926 (2022).

13. Bourbonne, V. et al. Development and prospective validation of a spatial dose pattern based model predicting acute pulmonary toxicity in patients treated with volumetric arc-therapy for locally advanced lung cancer. *Radiother Oncol.* **164**, 43–49. 10.1016/j.radonc.2021.09.008 (2021).
14. Bourbonne, V. et al. VMAT-Based Planning Allows Sparing of a Spatial Dose Pattern Associated with Radiation Pneumonitis in Patients Treated with Radiotherapy for a Locally Advanced Lung Cancer. *Cancers (Basel)*. **14** (15), 3702. 10.3390/cancers14153702 (2022).
15. Chen, Z. et al. Predicting radiation pneumonitis in lung cancer using machine learning and multimodal features: a systematic review and meta-analysis of diagnostic accuracy. *BMC Cancer*. **24** (1), 1355. 10.1186/s12885-024-13098-5 (2024).
16. Reuter, L. M. et al. Prediction of symptomatic radiation pneumonitis in lung cancer patients: a radiomics and dosiomics machine learning approach using the prospective multicenter RTOG 0617 and REQUITE trials. *Int. J. Radiat. Oncol. Biol. Phys.* 2026 Feb **18**:S0360 -3016(26)00365–2. 10.1016/j.ijrobp.2026.01.031
17. Zhang, Z. et al. Radiomics and Dosiomics Signature From Whole Lung Predicts Radiation Pneumonitis: A Model Development Study With Prospective External Validation and Decision-curve Analysis. *Int. J. Radiat. Oncol. Biol. Phys.* **115** (3), 746–758. 10.1016/j.ijrobp.2022.08.047 (2023).
18. Puttanawarut, C. et al. Radiomic and Dosiomic Features for the Prediction of Radiation Pneumonitis Across Esophageal Cancer and Lung Cancer. *Front. Oncol.* **12**, 768152. 10.3389/fonc.2022.768152 (2022).
19. Li, B. et al. Function-Wise Dual-Omics analysis for radiation pneumonitis prediction in lung cancer patients. *Front. Pharmacol.* **13**, 971849. 10.3389/fphar.2022.971849 (2022).
20. Zhou, L. et al. Machine learning-based multiomics prediction model for radiation pneumonitis. *J. Oncol.* **2023**, 5328927. 10.1155/2023/5328927 (2023).
21. Demircioğlu, A. Evaluation of the dependence of radiomic features on the machine learning model. *Insights Imaging*. **13** (1), 28. 10.1186/s13244-022-01170-2 (2022).
22. Hong, S. et al. Development of a flexible feature selection framework in radiomics-based prediction modeling: Assessment with four real-world datasets. *Sci. Rep.* **14** (1), 29297. 10.1038/s41598-024-80863-8 (2024).
23. Jiang, A., Lu, H., Zhao, R., Yuan, J. & Chen, D. The impact of pre-existing interstitial lung disease on radiation and checkpoint inhibitor pneumonitis in lung cancer patients: a systematic review and meta-analysis. *Ther. Adv. Med. Oncol.* **17**, 17588359251338624. 10.1177/17588359251338624 (2025).
24. Saha, A., Dickinson, P., Shrimali, R. K., Salem, A. & Agarwal, S. Is Thoracic Radiotherapy an Absolute Contraindication for Treatment of Lung Cancer Patients With Interstitial Lung Disease? A Systematic Review. *Clin. Oncol. (R Coll. Radiol)*. **34** (12), e493–e504. 10.1016/j.clon.2022.01.043 (2022).
25. Han, C. et al. Pneumonitis Risk After Chemoradiotherapy With and Without Immunotherapy in Patients With Locally Advanced Non-Small Cell Lung Cancer: A Systematic Review and Meta-Analysis. *Int. J. Radiat. Oncol. Biol. Phys.* **119** (4), 1179–1207. 10.1016/j.ijrobp.2024.01.217 (2024).

26. Bi, J. et al. Dosimetric predictors of radiation pneumonitis in patients with prior immunotherapy exposure: A multi-institutional analysis. *Radiother Oncol.* **190**, 110040. 10.1016/j.radonc.2023.110040 (2024).
27. Jang, J. Y. et al. Radiation pneumonitis in patients with non-small-cell lung cancer receiving chemoradiotherapy and an immune checkpoint inhibitor: a retrospective study. *Radiat. Oncol.* **16** (1), 231. 10.1186/s13014-021-01930-2 (2021).
28. Jiao, Y. et al. Integrating radiomics and dosiomics with lung biologically equivalent dose for predicting symptomatic radiation pneumonitis after lung SBRT: A dual-center study. *Radiother Oncol.* **212**, 111129. 10.1016/j.radonc.2025.111129 (2025).
29. Niu, L. et al. A multiomics approach-based prediction of radiation pneumonia in lung cancer patients: impact on survival outcome. *J. Cancer Res. Clin. Oncol.* **149** (11), 8923–8934. 10.1007/s00432-023-04827-7 (2023).
30. Wang, W., Wang, Y., Wang, X., Yu, J. & Meng, X. Predicting radiation pneumonitis with dose-segmented radiomics in locally advanced non-small cell lung cancer patients undergoing consolidative immunotherapy post-concurrent chemoradiotherapy. *Front. Immunol.* **16**, 1684629. 10.3389/fimmu.2025.1684629 (2025).
31. Sheen, H. et al. Radiomics-based hybrid model for predicting radiation pneumonitis: A systematic review and meta-analysis. *Phys. Med.* **123**, 103414. 10.1016/j.ejmp.2024.103414 (2024).
32. Kraus, K. M. et al. Dosiomics and radiomics-based prediction of pneumonitis after radiotherapy and immune checkpoint inhibition: The relevance of fractionation. *Lung Cancer.* **189**, 107507. 10.1016/j.lungcan.2024.107507 (2024).
33. Su, W. et al. Multi-omics deep learning for radiation pneumonitis prediction in lung cancer patients underwent volumetric modulated arc therapy. *Comput. Methods Programs Biomed.* **254**, 108295. 10.1016/j.cmpb.2024.108295 (2024).
34. Huang, Y. et al. Radiation pneumonitis prediction after stereotactic body radiation therapy based on 3D dose distribution: dosiomics and/or deep learning-based radiomics features. *Radiat. Oncol.* **17** (1), 188. 10.1186/s13014-022-02154-8 (2022).
35. Ai, Y. et al. Integrating deep learning and multi-omics features in radiation pneumonitis prediction for lung cancer patients using PET/CT. *BMC Med. Imaging.* **25** (1), 426. 10.1186/s12880-025-01971-z (2025).
36. Jiang, W., Song, Y., Sun, Z., Qiu, J. & Shi, L. Dosimetric Factors and Radiomics Features Within Different Regions of Interest in Planning CT Images for Improving the Prediction of Radiation Pneumonitis. *Int. J. Radiat. Oncol. Biol. Phys.* **110** (4), 1161–1170. 10.1016/j.ijrobp.2021.01.049 (2021).
37. Hou, L. et al. CT-based different regions of interest radiomics analysis for acute radiation pneumonitis in patients with locally advanced NSCLC after chemoradiotherapy. *Clin. Transl Radiat. Oncol.* **48**, 100828. 10.1016/j.ctro.2024.100828 (2024).

Figures

Correlation Heatmap of Selected Radiomics and Dosiomics Features

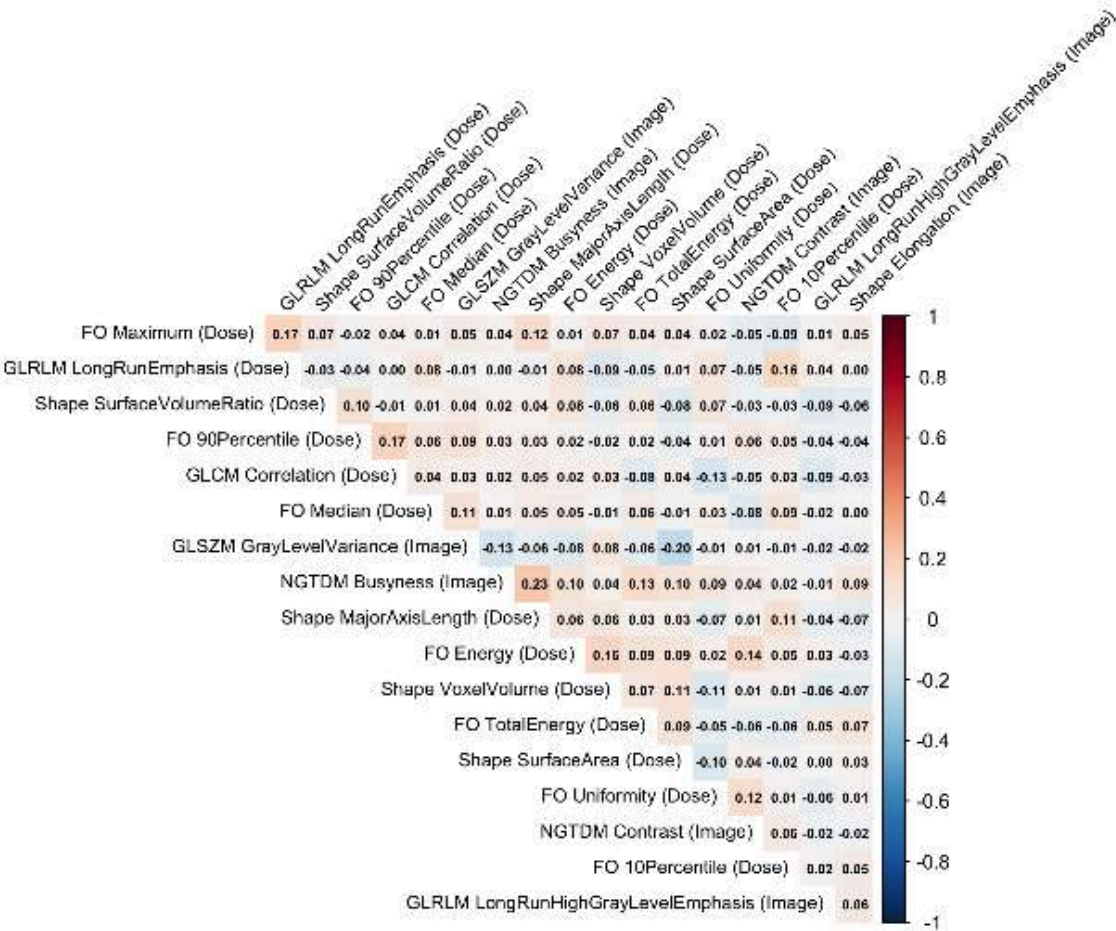


Figure 1

Correlation heatmap of the selected radiomics and dosiomics features. Pairwise correlations among the 18 selected features are shown. No feature pair had an absolute correlation coefficient greater than 0.90.

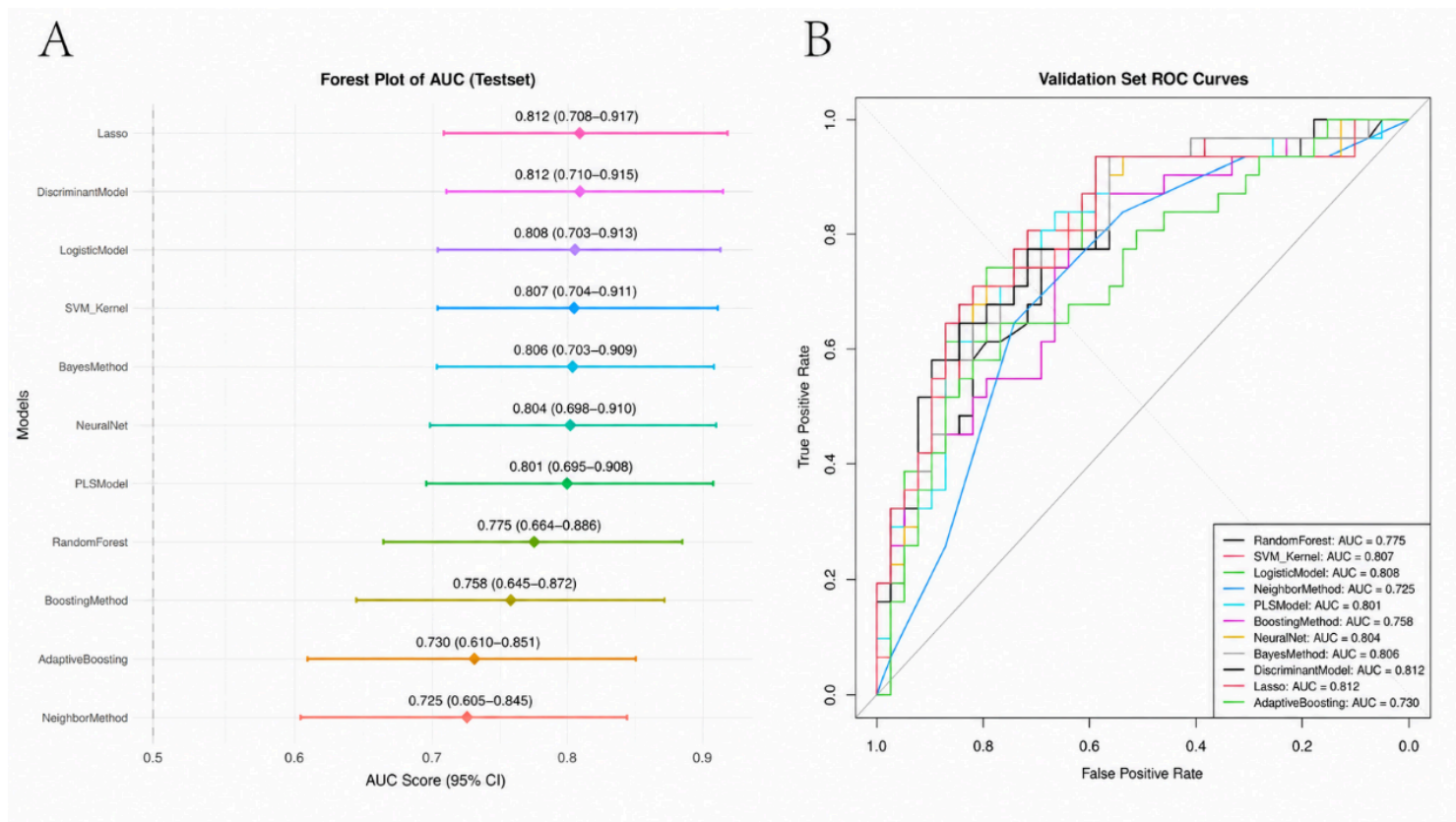


Figure 2

Comparison of predictive performance among machine learning models in the validation set. (A) AUC forest plot with 95% confidence intervals. (B) ROC curves of the 11 machine learning models. AUC, area under the receiver operating characteristic curve; ROC, receiver operating characteristic.

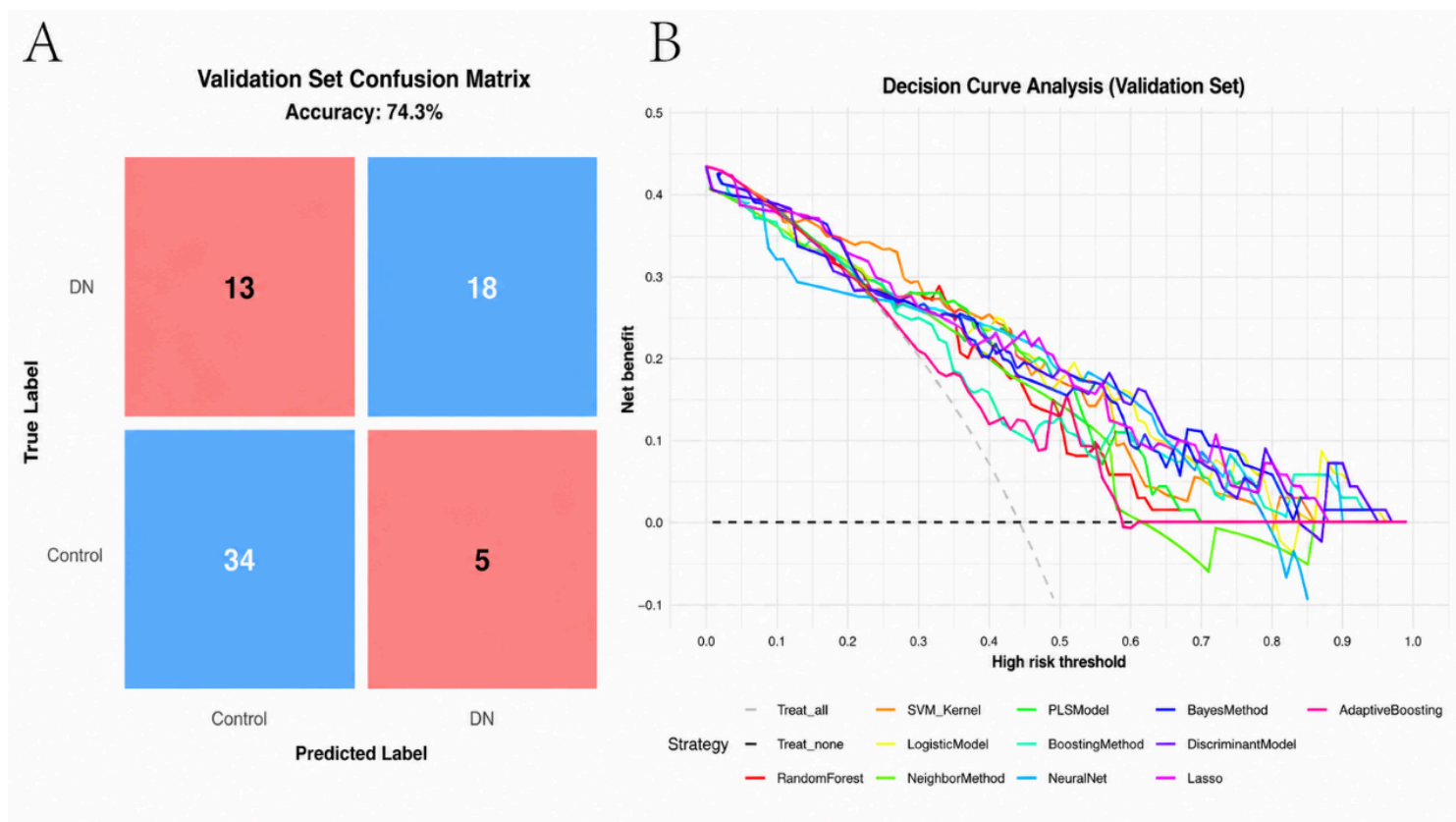


Figure 3

Validation-set confusion matrix and decision curve analysis. (A) Confusion matrix of the final LASSO model in the validation set. (B) Decision curve analysis of different machine learning models in the validation set. Control indicates patients without grade ≥ 2 radiation pneumonitis, whereas DN indicates patients who developed grade ≥ 2 radiation pneumonitis.

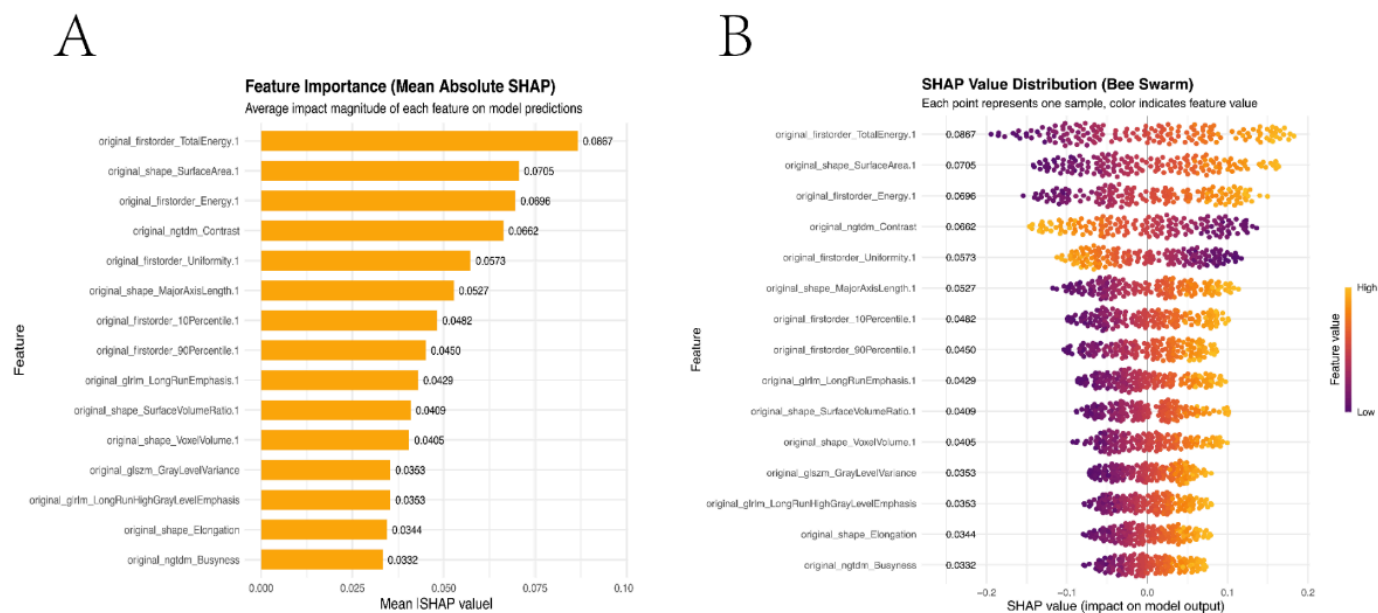


Figure 4

SHAP-based global interpretability analysis of the final LASSO model. (A) Feature importance ranking based on mean absolute SHAP values. (B) SHAP beeswarm plot showing the distribution of feature effects on model output. Each point represents one patient, and the color indicates the feature value. SHAP, Shapley additive explanations.