

RESEARCH

Open Access



A radiomics-based machine learning model and SHAP for predicting spread through air spaces and its prognostic implications in stage I lung adenocarcinoma: a multicenter cohort study

Yuhang Wang^{1,2†}, Xufeng Liu^{3†}, Xiaojiang Zhao², Zixiao Wang⁴, Xin Li^{1,2*} and Daqiang Sun^{1,2*}

Abstract

Background Despite early detection via low-dose computed tomography and complete surgical resection for early-stage lung adenocarcinoma, postoperative recurrence remains high, particularly in patients with tumor spread through air spaces. A reliable preoperative prediction model is urgently needed to adjust the treatment modality.

Methods In this multicenter retrospective study, 609 patients with pathological stage I lung adenocarcinoma from 3 independent centers were enrolled. Regions of interest for the primary tumor and peritumoral areas (extended by three, six, and twelve voxel units) were manually delineated from preoperative CT imaging. Quantitative imaging features were extracted and filtered by correlation analysis and Random forest Ranking to yield 40 candidate features. Fifteen machine learning methods were evaluated, and a ten-fold cross-validated elastic net regression model was selected to construct the radiomics-based prediction model. A clinical model based on five key clinical variables and a combined model integrating imaging and clinical features were also developed.

Results The radiomics model achieved accuracies of 0.801, 0.866, and 0.831 in the training set and two external test sets, with AUC of 0.791, 0.829, and 0.807. In one external test set, the clinical model had an AUC of 0.689, significantly lower than the radiomics model (0.807, $p < 0.05$). The combined model achieved the highest performance, with AUC of 0.834 in the training set and 0.894 in an external test set ($p < 0.01$ and $p < 0.001$, respectively). Interpretability analysis revealed that wavelet-transformed features dominated the model, with the highest contribution from a feature reflecting small high-intensity clusters within the tumor and the second highest from a feature representing low-intensity clusters in the six-voxel peritumoral region. Kaplan–Meier analysis demonstrated that patients with either pathologically confirmed or model-predicted spread had significantly shorter progression-free survival ($p < 0.001$).

[†]Yuhang Wang and Xufeng Liu contributed equally to this work.

*Correspondence:

Xin Li
lixinchest@tju.edu.cn
Daqiang Sun
sdqmd@tju.edu.cn

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



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Conclusion Our novel machine learning model, integrating imaging features from both tumor and peritumoral regions, preoperatively predicts tumor spread through air spaces in stage I lung adenocarcinoma. It outperforms traditional clinical models, highlighting the potential of quantitative imaging analysis in personalizing treatment. Future prospective studies and further optimization are warranted.

Keywords Lung adenocarcinoma, STAS, Radiomics, Machine learning, SHAP

Introduction

Lung cancer remains the leading cause of cancer-related mortality worldwide [1]. Among its pathological subtypes, lung adenocarcinoma (LUAD) is the most common, and research on its early detection and treatment has been ongoing [2]. With advancements in medical imaging, an increasing number of early-stage LUAD cases are being detected and treated through low-dose computed tomography (CT) screening. Despite complete surgical resection being the primary treatment for stage I LUAD, studies have shown that recurrence rates range from 20% to 50%, even after curative surgery [3]. Therefore, identifying high-risk patients at an early stage is crucial for improving patient prognosis and guiding treatment strategies.

In the 2015 WHO classification of lung cancer, the concept of spread through air spaces (STAS) was introduced. STAS is defined as the presence of tumor cells spreading beyond the tumor margin into the alveolar spaces in the form of micropapillary clusters, solid nests, or single cells [4]. Subsequent studies have demonstrated that STAS is a significant risk factor for recurrence in patients with stage I LUAD after surgical resection [5, 6]. However, STAS is typically diagnosed postoperatively through pathological examination, limiting its utility for preoperative treatment planning. Therefore, a reliable preoperative STAS prediction model is essential for identifying high-risk patients and tailoring surgical and adjuvant treatment strategies.

With the rapid development of machine learning and artificial intelligence in the medical field, researchers have begun extracting quantitative features from medical images to assist in disease diagnosis and prognostication. Radiomics is a powerful approach that aims to transform standard imaging data into high-dimensional quantitative features, capturing complex tumor characteristics beyond human visual perception [7]. Previous studies have demonstrated that both tumoral and peritumoral CT radiomics features are valuable in assisting the diagnosis, prognostication, and histological classification of LUAD [8]. However, existing studies on preoperative STAS prediction have primarily focused on either tumoral or peritumoral radiomics features, and few have comprehensively integrated these modalities to improve predictive performance [9–12].

In our previous research, we developed a clinical model for preoperative STAS prediction in stage I non-small cell lung cancer (NSCLC) patients based on demographic and imaging characteristics, such as tumor size, spiculation, vacuole sign, and carcinoembryonic antigen (CEA) levels [13]. Other studies have constructed STAS prediction models based on traditional or deep learning radiomics features [12, 14, 15]. However, most of these studies were limited to single-center datasets, and their model generalizability remains uncertain.

To address these limitations, we conducted a multi-center retrospective cohort study to develop and validate a multimodal preoperative STAS risk prediction model incorporating both tumoral and peritumoral radiomics features. We employed a multi-machine learning approach, integrating various algorithms—including LASSO, support vector machine, random forest, gradient boosting, elastic net regression, and neural networks—to optimize model performance. We combined radiomics features with key clinical variables to construct a hybrid model, enhancing predictive accuracy. Our dataset was derived from three independent medical centers, allowing us to assess model robustness and generalizability. We believe that our findings will contribute to improved risk stratification and personalized treatment strategies for early-stage LUAD patients.

The technical workflow and main findings of this study are summarized in Fig. 1.

Methods

In this multi-center retrospective cohort study, we developed and validated a multimodal preoperative STAS risk prediction model based on radiomics features extracted from preoperative CT, including both tumoral and peritumoral radiomics features. This model aims to assist clinicians in the early identification of high-recurrence-risk stage I LUAD patients, enabling timely adjustments to treatment strategies.

Clinical data collection and follow-up

In this multi-center cohort study, we included patients with pathological stage I lung adenocarcinoma (LUAD) who had undergone complete surgical resection. The study samples were collected from three centers. Specifically, patients who underwent surgery at Tianjin

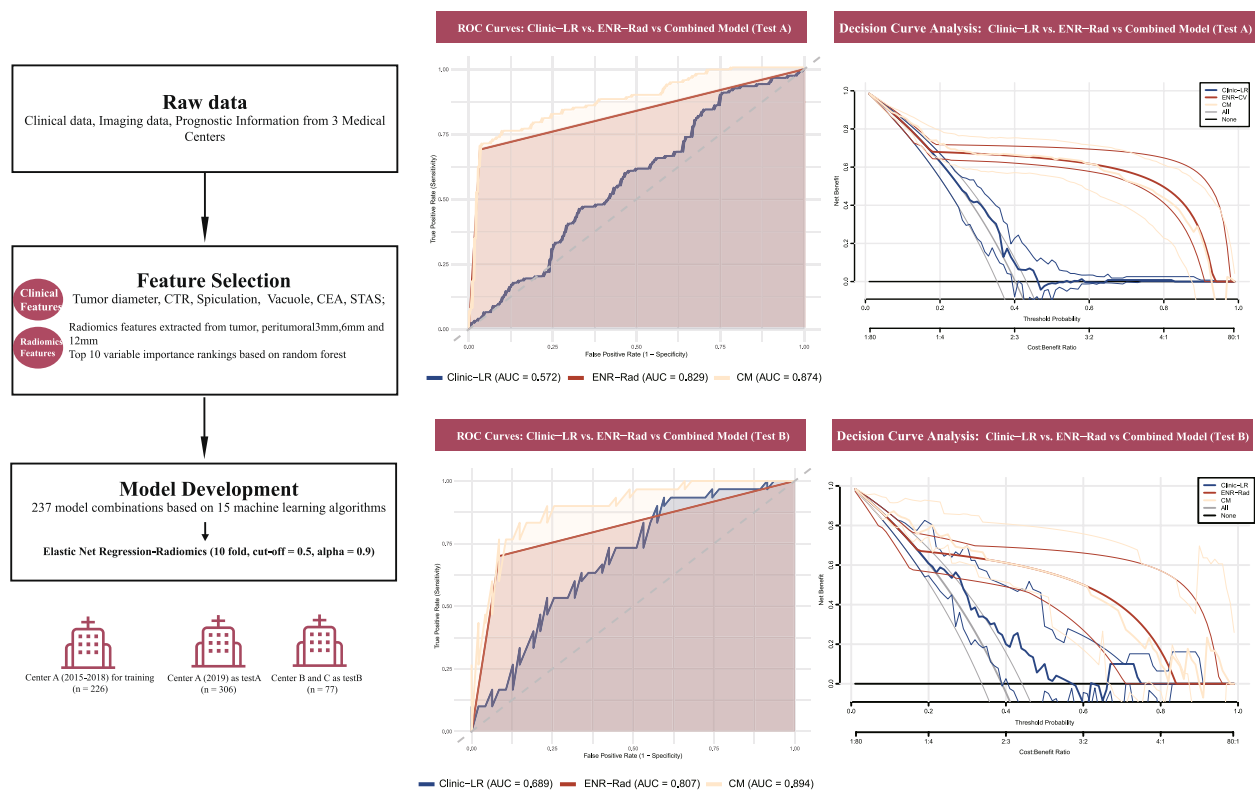


Fig. 1 Overview of the data processing, feature selection, model development, and evaluation workflow. Data sources include three medical centers: Center A (2015–2018) for training set ($n=226$), Center A (2019) for test A ($n=306$), and Centers B and C for test B ($n=77$). Radiomics features selection utilizes the top 10 most important variables from Random forest Ranking. The 237 model combinations were developed using 15 machine learning algorithms. The final model, ENR-Rad, was evaluated using ROC curves (Test A: AUC Clinic-LR = 0.572, ENR-Rad = 0.829, Combined = 0.874; Test B: AUC Clinic-LR = 0.689, ENR-Rad = 0.807, CM = 0.894) and Decision Curve Analysis. STAS: Spread Through Air Spaces; CTR: Consolidation-to-Tumor Ratio; CEA: Carcinoembryonic Antigen; Clinic-LR: Clinical Logistic Regression Model; ENR-rad: Elastic Net Regression-Radiomics Model; CM: Combined Model; ROC: Receiver Operating Characteristic

Chest Hospital (Center A) between January 1, 2015, and December 31, 2018, were assigned to the training set. Patients who underwent surgery at the same institution from January 1, 2019, to December 31, 2019, were included in test set A. Additionally, patients who underwent surgery at Tianjin Binhai New Area Haibin People's Hospital (Center B) and Qinhuaogdao First Hospital (Center C) between January 1, 2019, and December 31, 2020, were included in test set B. The detailed inclusion and exclusion criteria are illustrated in Fig. 2.

In the training set, STAS status was re-evaluated based on pathological slides (Supplement Figure1), whereas in Test Set A and Test Set B, the STAS status was obtained from pathology reports. Based on our previous research, several clinical features were identified as significantly associated with STAS status in stage I non-small cell lung cancer (NSCLC), including maximum tumor diameter (Tdmax), consolidation-to-tumor ratio (CTR), spiculation, vacuole sign, and carcinoembryonic antigen (CEA) levels. Therefore, only these five clinical features were

incorporated to compare the performance of the radiomics-based STAS risk prediction model with a clinical feature-based model. These clinical features were collected only in the training set, test set A and test set B.

Patients from Tianjin Chest Hospital (Center A) were followed up for prognosis assessment, primarily evaluating tumor progression. The last follow-up date was January 1, 2025.

Image acquisition and preprocessing

Preoperative thin-slice chest CT scans of patients were retrieved from the Picture Archiving and Communication System (PACS) in Digital Imaging and Communications in Medicine (DICOM) format.

Given the multi-center nature of this study and variations in CT scanners, all CT images were resampled to standardize voxel size to $1\text{ mm} \times 1\text{ mm} \times 1\text{ mm}$, ensuring a uniform slice thickness of 1 mm across all scans. Additionally, window width and window level were standardized to 1600 HU and -500 HU , respectively. ITK-SNAP

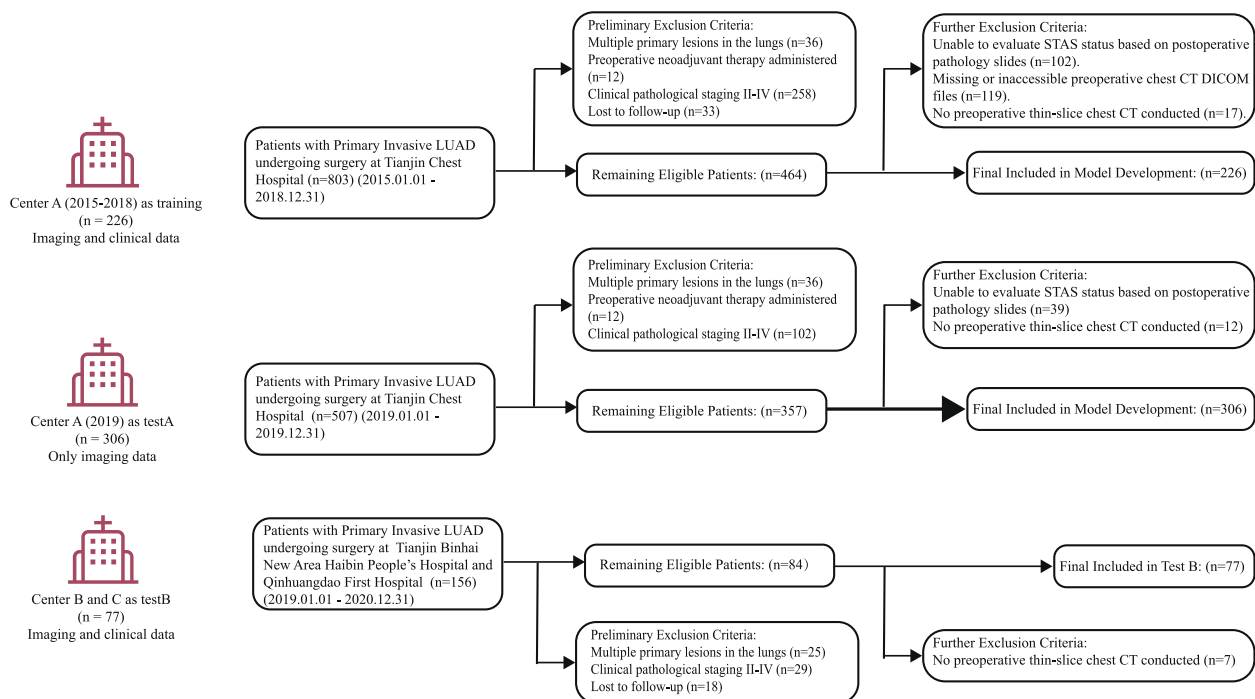


Fig. 2 Flow diagram of patient selection. Patients with primary invasive lung adenocarcinoma were retrospectively recruited from multiple centers. From Tianjin Chest Hospital between January 2015 and December 2018 (Center A), 226 patients were finally included for model development (training set). In addition, 306 patients from the same centers between January 2019 and December 2019 (Center A) were selected as Test A, and 77 patients from Tianjin Jinnan Hospital and Qinhuangdao First Hospital between January 2019 and December 2020 (Centers B and C) were included as Test B

open-source software was used for CT image parameter adjustments, while Python was employed for batch processing.

ROI delineation

Following CT image preprocessing, manual delineation of the regions of interest (ROIs) was performed (Fig. 3A). Based on the tumoral ROI (ROI-tumoral), the peritumoral ROI (ROI-peritumoral) was generated by outward expansion. After reviewing the literature [16–22], we selected peritumoral regions extending 3, 6, and 12 voxel units from the tumor boundary (Fig. 3A). To prevent extraction errors, areas of the peritumoral ROIs that extended beyond the lung parenchyma were removed [23] (Fig. 3B). Both the manual delineation and region expansion were executed using ITK-SNAP.

Radiomics feature extraction and selection

Radiomics features were extracted from CT images using the Pyradiomics package in Python from both tumoral and peritumoral regions. The extracted features included first-order features, shape features (2D and 3D), and gray-level features—namely, gray-level co-occurrence matrix (GLCM), gray-level size zone matrix (GLSZM), gray-level run length matrix (GLRLM), neighboring gray tone

difference matrix (NGTDM), and gray-level dependence matrix (GLDM)—in addition to wavelet features.

For the four feature groups, an initial screening was performed in R using Spearman correlation analysis, whereby one feature from each pair with a correlation coefficient greater than 0.9 was Randomly excluded. Subsequently, a Random forest regression algorithm was applied to rank the remaining features by variable importance. The top 10 STAS-related features from each feature group were selected, resulting in a total of 40 features that were used for subsequent model development. This approach ensured that each region of interest contributed representative features with high predictive value, based on their importance rankings. The workflow for radiomics feature extraction and selection is illustrated in Fig. 3C.

Machine learning model development

We employed 15 machine learning algorithms, including LASSO, support vector machine, random forest, gradient boosting, elastic net regression, and neural networks. By optimizing hyperparameters and combining different models, we developed multiple radiomics models for STAS prediction. To avoid selecting overfitted models, we included the performance of

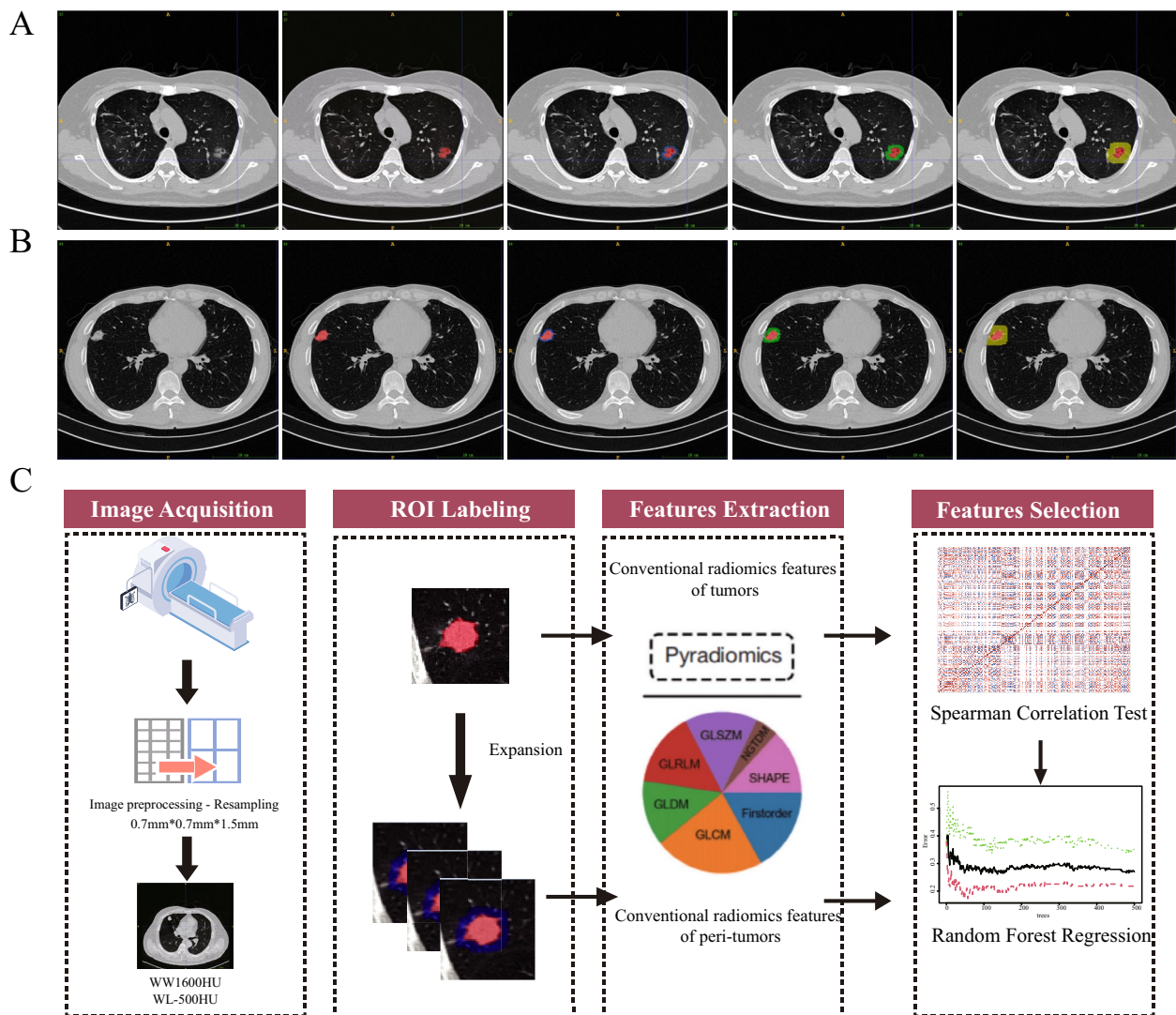


Fig. 3 Examples diagram of the region of interest is sketched. **A.** ITK-SNAP was used to delineate the region of interest, and the peritumoral area of 3 mm, 6 mm, and 12 mm was expanded, and the radiomic features of the tumor and peritumoral were extracted, respectively. **B.** For areas beyond the lung parenchyma, manually wipe off the excess portion. **C.** For the four feature groups, an initial screening was performed in R using Spearman correlation analysis, whereby one feature from each pair with a correlation coefficient greater than 0.9 was Randomly excluded. Subsequently, a Random forest regression algorithm was applied to rank the remaining features by variable importance. The top 10 STAS-related features from each feature group were selected, resulting in a total of 40 features that were used for subsequent model development

each candidate model across training and test sets during model selection. Additionally, we constructed a clinical prediction model for STAS risk (Clinic-LR) using five clinical features: Tdmax, CTR, spiculation, vacuole sign, and CEA. Furthermore, we integrated the radiomics-predicted STAS label with these clinical features to form a combined logistic regression model (Combined Model) for comparison. Finally, the models were evaluated in both the training and test sets in terms of accuracy, discrimination ability, and clinical benefit. SHapley Additive exPlanations (SHAP)

explainability analysis is used to interpret machine learning models [24].

Results

Baseline characteristics of enrolled patients

Table 1 summarizes the data of 609 patients included in this study. Among these, 226 patients from Center A (2015–2018) were used as the training set, for which both imaging features and prognostic follow-up were collected, the median follow-up was 2330 days, with a maximum follow-up of 10 years. In contrast, 306

Table 1 Baseline characteristics of patients in the model development cohort

	Train set			Test setA			Test setB		
	Negative (N=133)	Positive (N=93)	All (N=226)	Negative (N=183)	Positive (N=123)	All (N=306)	Negative (N=47)	Positive (N=30)	All (N=77)
Tdmax_cm									
Mean (SD)	1.96 (0.747)	2.45 (0.812)	2.16 (0.810)	2.60 (0.962)	2.85 (1.13)	2.70 (1.04)	1.71 (0.675)	1.89 (0.619)	1.78 (0.656)
Median [Min, Max]	1.87 [0.700, 4.50]	2.37 [0.990, 5.28]	2.07 [0.700, 5.28]	2.59 [0.757, 5.70]	2.77 [0.845, 7.05]	2.69 [0.757, 7.05]	1.70 [0.730, 3.31]	1.91 [0.770, 3.10]	1.80 [0.730, 3.31]
CTR									
Mean (SD)	0.614 (0.310)	0.769 (0.220)	0.678 (0.286)	0.730 (0.391)	0.864 (0.274)	0.784 (0.354)	0.478 (0.439)	0.556 (0.407)	0.508 (0.426)
Median [Min, Max]	0.638 [0, 1.00]	0.806[0.136, 1.00]	0.707 [0, 1.00]	1.00 [0, 1.00]	1.00 [0, 1.00]	1.00 [0, 1.00]	0.400 [0, 1.00]	0.480 [0, 1.00]	0.450 [0, 1.00]
spiculation									
No	69 (51.9%)	26 (28.0%)	95 (42.0%)	89 (48.6%)	40 (32.5%)	129 (42.2%)	36 (76.6%)	8 (26.7%)	44 (57.1%)
Yes	64 (48.1%)	67 (72.0%)	131 (58.0%)	94 (51.4%)	83 (67.5%)	177 (57.8%)	11 (23.4%)	22 (73.3%)	33 (42.9%)
vacuole									
No	90 (67.7%)	56 (60.2%)	146 (64.6%)	142 (77.6%)	84 (68.3%)	226 (73.9%)	44 (93.6%)	19 (63.3%)	63 (81.8%)
Yes	43 (32.3%)	37 (39.8%)	80 (35.4%)	41 (22.4%)	39 (31.7%)	80 (26.1%)	3 (6.4%)	11 (36.7%)	14 (18.2%)
CEA									
Mean (SD)	3.36 (3.80)	5.06 (7.58)	4.06 (5.71)	3.69 (4.71)	4.22 (7.32)	3.90 (5.89)	2.37 (1.43)	4.81 (4.83)	3.32 (3.40)
Median [Min, Max]	2.80 [0.210, 34.0]	2.81 [0.760, 54.9]	2.81 [0.210, 54.9]	2.50 [0.300, 45.0]	2.18 [0.410, 50.9]	2.39 [0.300, 50.9]	2.20 [0.660, 8.50]	3.04 [0.720, 25.5]	2.56 [0.660, 25.5]
PFS.status									
No-pro- gress	122 (91.7%)	57 (61.3%)	179 (79.2%)	176 (96.2%)	107 (87.0%)	283 (92.5%)	-	-	-
Progress	11 (8.3%)	36 (38.7%)	47 (20.8%)	7 (3.8%)	16 (13.0%)	23 (7.5%)	-	-	-
PFS.time									
Mean (SD)	2570 (482)	2000 (895)	2330 (736)	1990 (184)	1870 (334)	1940 (262)	-	-	-
Median [Min, Max]	2490 [752, 3650]	2340 [246, 3460]	2440 [246, 3650]	2010 [943, 2190]	1950 [695, 2190]	1990 [695, 2190]	-	-	-

Tdmax Tumor diameter max, CTR Consolidation-to-tumour ratio, CEA Carcinoembryonic antigen, PFS Progression free survival

patients from Center A (2019) formed Test Set A, with only prognostic follow-up data available, the median follow-up was 1940 days. Additionally, 77 patients from Centers B and C comprised Test Set B, where imaging features were collected without prognostic follow-up. Missing values in CEA were imputed using the median value.

Feature extraction and selection

For each ROI, Pyradiomics extracted 1,133 features. After Spearman correlation analysis, 258 features from ROI-tumoral, 232 from ROI-peritumor_3mm, 236 from ROI-peritumor_6mm, and 236 from ROI-peritumor_12mm were retained. Random forest was then used for further feature selection within each ROI, retaining the top 10 features based on variable importance ranking (Fig. 4). Ultimately, these 40 selected features were incorporated into the radiomics model.

Model development and evaluation

A total of 237 optimized models were constructed. Based on the average accuracy across the training set and two test sets (Fig. 5A, Supplementary Figure 2), the best-performing model was selected: a tenfold cross-validated elastic net regression model (ENR-CV: tenfold, cut-off=0.5, alpha=0.9), referred to as ENR-Rad. The accuracy of ENR-Rad was 0.801 in the training set, 0.866 in Validation Set A, and 0.831 in Validation Set B. The receiver operating characteristic (ROC) curves and corresponding area under the curve (AUC) values were calculated for all three datasets (train=0.791, test A=0.829, test B=0.807) (Fig. 5B). In addition to AUC and accuracy, the F1-score and recall of the final model in the training set were 0.761 and 0.774, respectively, indicating favorable balance between sensitivity and precision, particularly for the STAS-positive class.

Additionally, a logistic regression model based on clinical imaging features (Clinic-LR) was developed for comparison. The diagnostic performance of ENR-Rad,

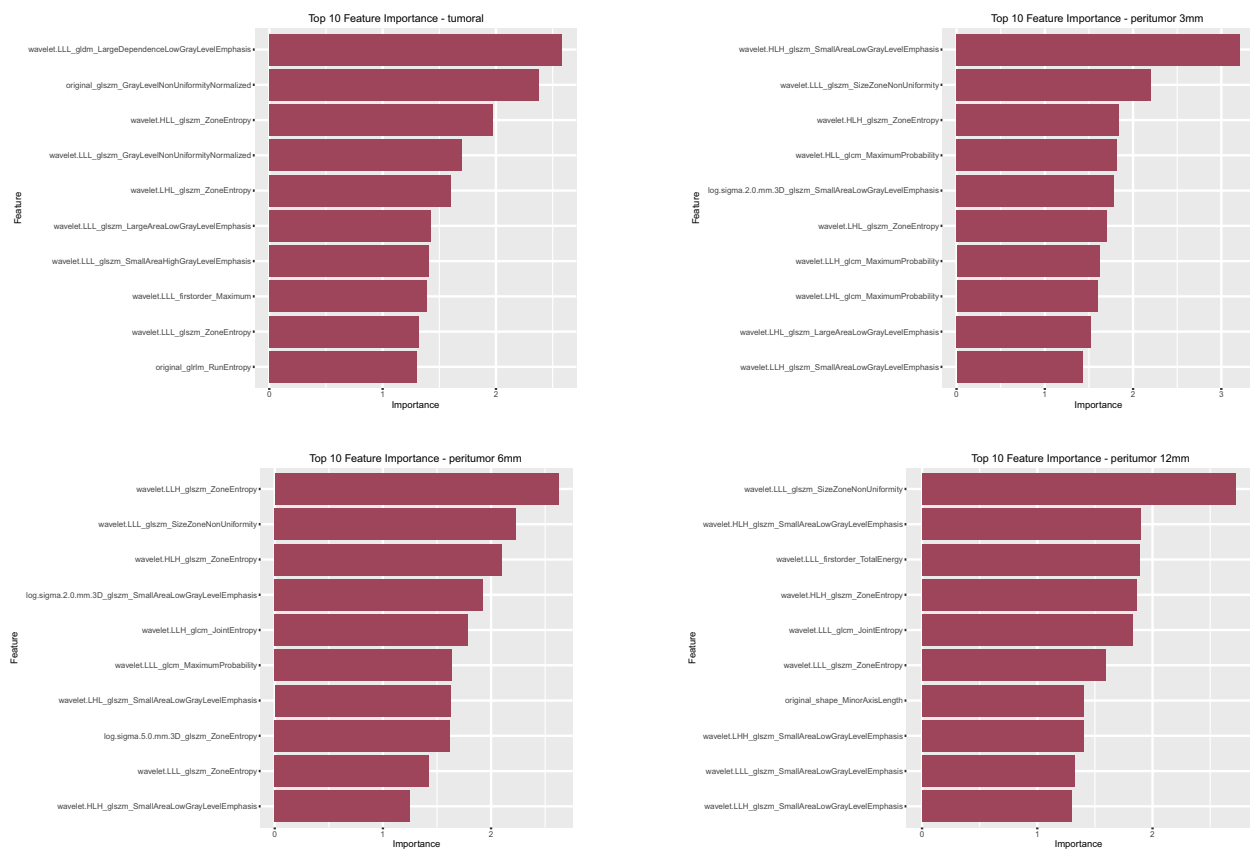


Fig. 4 Random forest-based feature selection for radiomics. The figure displays the top 10 features ranked by importance for each region of interest: peritumoral regions at 3 mm, 6 mm, and 12 mm from the tumor boundary, as well as for the tumoral region. For each panel, the horizontal bars represent the importance scores derived from the random forest algorithm, and the vertical axis lists the feature names (including various wavelet, log-sigma, and original shape/first-order metrics)

(See figure on next page.)

Fig. 5 Model construction and evaluation. **A:** By combining 15 machine learning methods and adjusting parameters, a total of 237 parameter optimization models were constructed, and their accuracy on the training set and test set was shown as the basis for selecting the model. Finally, based on the average accuracy of the model in the three data sets, ENR – CV:tenfold (cutoff:0.5, alpha:0.9) was selected and named ENR-Rad; **B:** ROC curve of ENR-Rad in the training set and test set; **C:** The ENR-Rad model is compared with the LR model based only on clinical characteristics, and then a Combined Model combining the ENR-Rad and clinical models is constructed. The ROC curve is drawn based on the training set data. It can be seen that CM has the best AUC (0.834); **D:** The DCA curves of the three models are drawn based on the training set data, with a bootstrap of 50. It can be seen that CM has the best clinical benefit rate. **E:** Finally, 12 meaningful features were screened out in the ENR-Rad model, and SHAP interpretability analysis was performed on these features, and a SHAP importance bar chart was drawn based on the SHAP value; **F:** Single sample SHAP force diagram. STAS: Spread Through Air Spaces; Clinic-LR: Clinical Logistic Regression Model; ENR-rad: Elastic Net Regression-Radiomics Model; CM: Combined Model; ROC: Receiver Operating Characteristic; AUC: Area Under the Curve;

Clinic-LR, and a combined model (CM) integrating ENR-Rad outcome with clinical features was evaluated (Fig. 5C). In the training set, ENR-Rad (AUC=0.791) demonstrated higher diagnostic performance than Clinic-LR (AUC=0.738), though the difference was not statistically significant (DeLong test $p=0.1246$). However, CM (AUC=0.834) exhibited significantly better performance than ENR-Rad (DeLong test $p=0.0089$). We then compared the predictive performance of the

clinical model (Clinic-LR), the radiomics model (ENR-Rad), and the combined model (CM) in Test set A (Fig. 1). The AUCs for Clinic-LR, ENR-Rad, and CM were 0.572, 0.829, and 0.874, respectively. Delong test results demonstrated that ENR-Rad significantly outperformed Clinic-LR ($p<0.001$), and CM significantly outperformed ENR-Rad ($p=0.0029$). The superiority of the radiomics-based model was further highlighted in Test set B (Fig. 1), where ENR-Rad (AUC=0.807) significantly

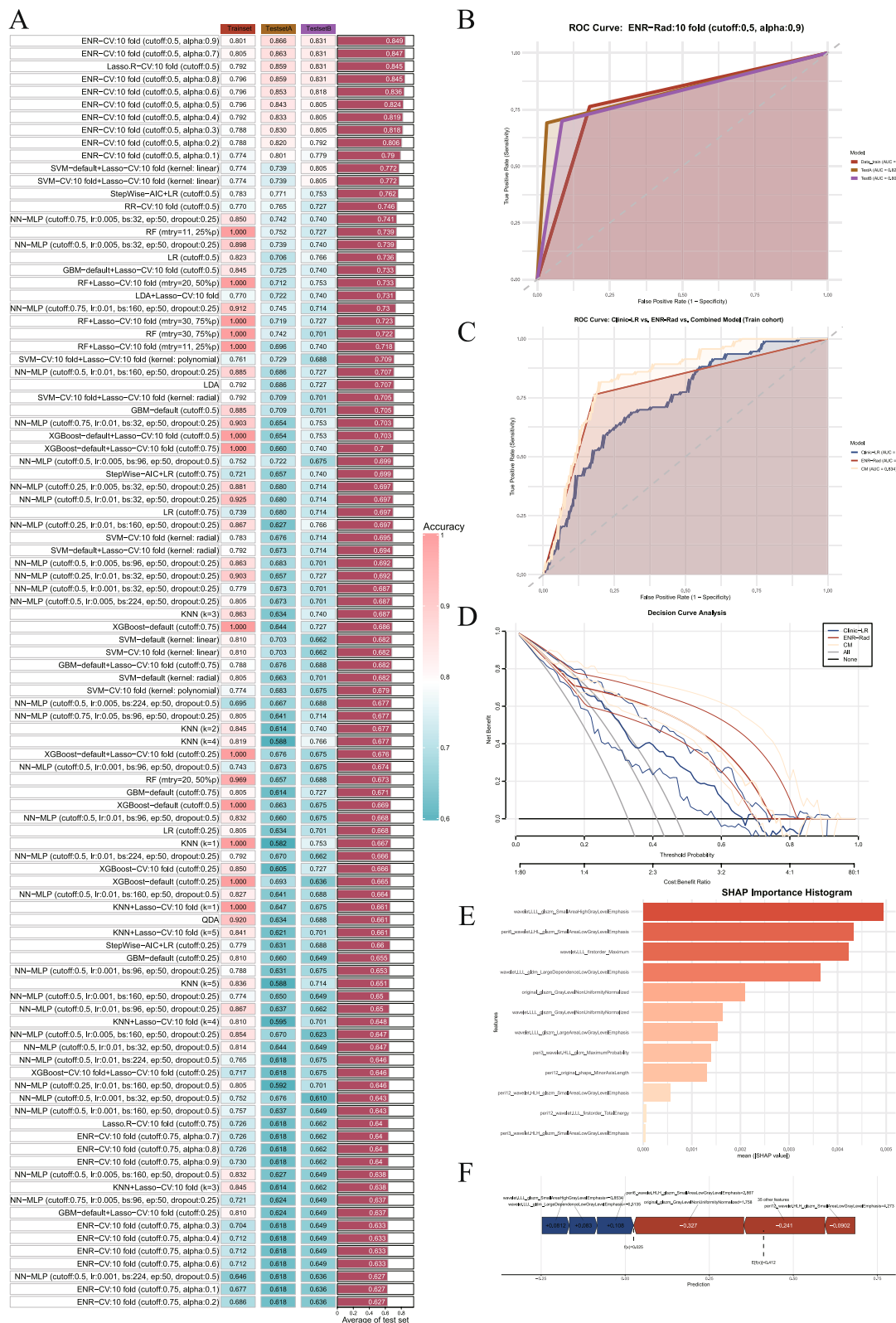


Fig. 5 (See legend on previous page.)

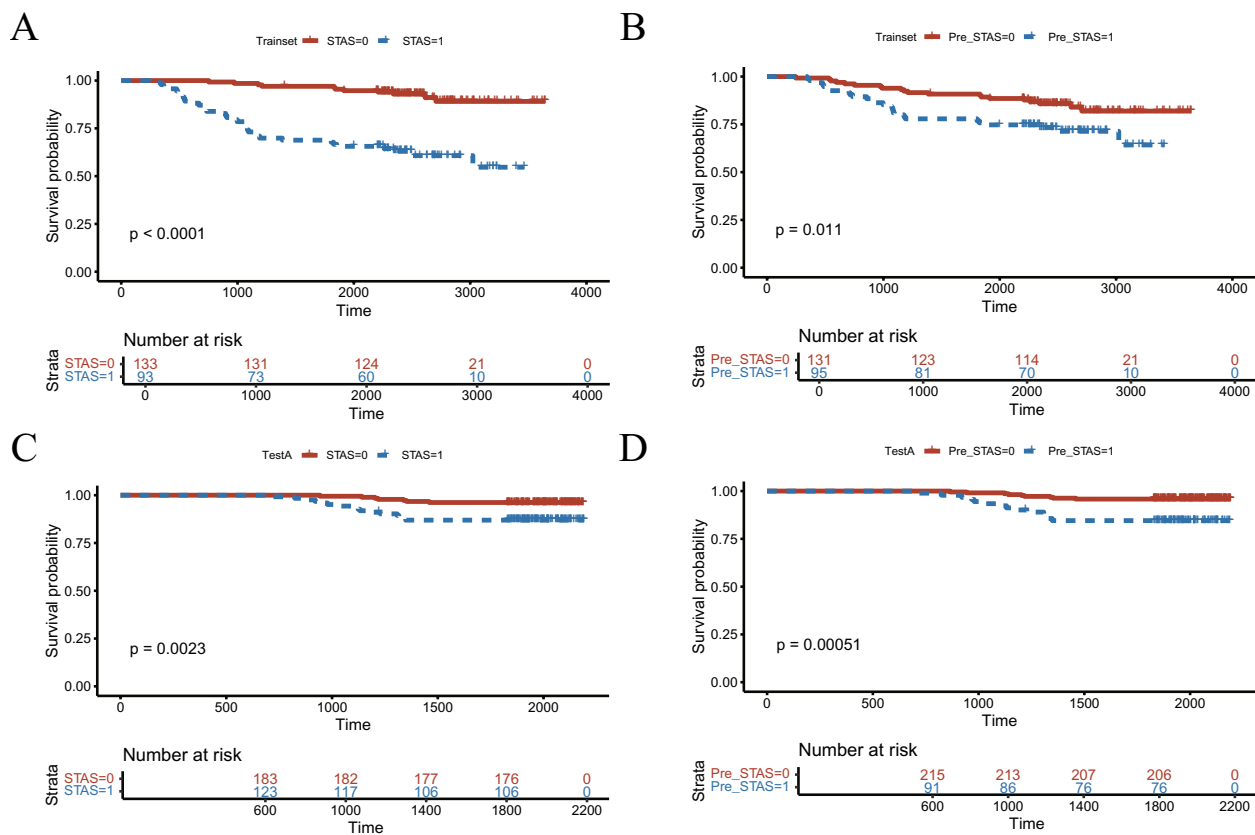


Fig. 6 Evaluate the prognostic relevance of ENR-Rad. **A**: Recurrence-free survival of the STAS positive group and negative group in the training set; **B**: Recurrence-free survival of the STAS positive group and negative group predicted by ENR-Rad in the training set; **C**: Recurrence-free survival of the STAS positive group and negative group in the testB; **D**: Recurrence-free survival of the STAS positive group and negative group predicted by ENR-Rad in the testB. STAS: Spread Through Air Spaces; ENR-rad: Elastic Net Regression-Radiomics Model;

outperformed Clinic-LR (AUC=0.689) (DeLong test $p < 0.05$). Additionally, CM (AUC=0.894) showed a significant improvement over ENR-Rad (DeLong test $p < 0.001$).

Decision curve analysis (DCA) in the training set indicated that CM provided the highest net clinical benefit, followed by ENR-Rad, with Clinic-LR ranking last (Fig. 5D). A similar trend was observed in test set A and B (Fig. 1).

SHAP interpretability analysis

In the ENR-Rad model, only 12 features are actually retained when $\lambda = \lambda_{\min}$, i.e., these 12 features really play a role in the model (Supplementary Figure 3). SHAP interpretability analysis of ENR-Rad showed that wavelet.LLL_glszm_SmallAreaHighGray-LevelEmphasis made the most contributions to the model (Fig. 5E). Figure 5F is a single-sample SHAP force plot that shows the extent to which each feature contributes to the model. Blue indicates positive contribution and red indicates negative contribution. The baseline expected

value $E[f(x)] = 0.412$ is the overall prediction mean of the model. The current sample prediction value $f(x) = 0.025$ is significantly lower than the baseline, indicating that its feature combination tends to suppress the prediction result, that is, STAS negative.

Prognostic relevance of ENR-Rad

As a newly identified pathway of lung cancer dissemination, STAS has been proven to be closely associated with tumor prognosis. To evaluate the prognostic value of ENR-Rad, Kaplan–Meier survival curves were plotted to compare progression-free survival (PFS) between high-risk and low-risk groups predicted by the model.

In the training set, STAS-positive patients exhibited significantly worse PFS than STAS-negative patients ($p < 0.001$) (Fig. 6A). Similarly, patients predicted by ENR-Rad to be at high risk for STAS had significantly poorer PFS than those predicted to be at low risk ($p = 0.011$) (Fig. 6B). This finding was further validated in Test A, where STAS-negative patients had a significantly longer PFS ($p = 0.002$) (Fig. 6C), and those classified as low risk

by ENR-Rad also showed a prolonged PFS ($p < 0.001$) (Fig. 6D). These results reinforce the potential clinical utility of the ENR-Rad model, offering valuable insights for improving diagnostic accuracy and personalizing treatment strategies.

Discussion

Studies in recent Years have shown that sublobectomy has disease-free survival and overall survival that are not inferior to lobectomy for hilar and mediastinal lymph node-negative early-stage lung cancer with a diameter of less than 2 cm [25–27]. Currently, there are no clear international guidelines for surgical strategies for adenocarcinoma of the lung (LUAD) nodules ≤ 3 cm in diameter and positive STAS (airway spread). However, several studies have suggested that the option of lobectomy may provide a better prognosis for patients with STAS-positive early-stage lung adenocarcinoma [28–32].

Regarding postoperative adjuvant therapy, the 2023 NCCN guidelines recommend adjuvant chemotherapy as a standard treatment for early-stage NSCLC patients with high-risk factors, such as highly invasive tumors [33]. According to the Guidelines for Clinical Diagnosis and Treatment of Lung Cancer issued by the Chinese Medical Association in 2024 [34], adjuvant chemotherapy is recommended for patients with early-stage non-small cell lung cancer (NSCLC) with high-risk factors, including STAS-positive patients. In certain clinical trials, such as the KEYNOTE-091 trial [35], immune checkpoint inhibitors (ICIs) have been shown to significantly improve disease-free survival in early-stage (IB) high-risk NSCLC patients. Sum up, STAS-positive patients may derive additional clinical benefits from immunotherapy. However, there are currently no clinical trials that specifically target STAS-positive patients as a separate subgroup, highlighting the need for further investigation.

Based on our previous research [13], we developed a clinical feature-based model (Clinic-LR) using five clinical factors: Tdmax, CTR, burr sign, vacuole sign, and CEA, and compared its performance with a radiomics-based model (ENR-Rad) and a combined model (CM). Our results demonstrated that the radiomics-based model (ENR-Rad) significantly outperformed the clinical feature-based model (Clinic-LR) in the external validation cohort (Test Set A and B). Furthermore, the combined model (CM) incorporating both radiomic and clinical features achieved the best predictive performance. These findings suggest that radiomic features serve as a powerful complement to traditional clinical factors in improving STAS prediction. The superiority of the CM model further indicates that while radiomic features provide strong predictive capabilities, integrating

clinical information can enhance model stability and generalizability.

Kaplan–Meier survival analysis revealed that STAS-positive patients had significantly shorter progression-free survival (PFS) than STAS-negative patients, further confirming STAS as a high-risk factor for recurrence. Additionally, when patients were stratified according to STAS status predicted by the radiomics model, those predicted as STAS-positive exhibited worse survival outcomes. This finding suggests that our radiomics model not only serves as a tool for STAS risk assessment but may also function as a surrogate indicator for recurrence risk, thereby assisting in personalized treatment decision-making.

Recent studies have employed SHapley Additive exPlanations (SHAP) for feature selection in predicting spread through air spaces (STAS) in lung adenocarcinoma (LUAD), highlighting the significance of multi-resolution texture information and both tumor and peritumoral characteristics [36–38]. In our study, SHAP-based feature selection results showed that wavelet-transformed features played a dominant role, indicating that multi-resolution texture information is crucial for STAS prediction [39, 40]. Among the 12 selected features, tumor and peritumoral features contributed equally, suggesting that both intrinsic tumor characteristics and peritumoral tissue properties are essential for STAS identification. Features from peri3, peri6, and peri12 regions were all highly ranked, underscoring the significance of both proximal and distant peritumoral characteristics in STAS prediction. Several GLSZM and GLDM-based features (e.g., SmallAreaHighGrayLevelEmphasis and LargeDependenceLowGrayLevelEmphasis) were among the highest-ranking SHAP features, indicating that gray-level heterogeneity is a key factor in distinguishing STAS-positive cases. The top-ranked feature, wavelet.LLL_glszm_SmallAreaHighGrayLevelEmphasis, represents small high-intensity clusters within the tumor, suggesting that high-density heterogeneous regions are associated with STAS positivity. The second-ranked feature, peri6_wavelet.LHL_glszm_SmallAreaLowGrayLevelEmphasis, derived from the 6 mm peritumoral region, describes small low-intensity clusters, emphasizing the role of peritumoral microstructural alterations in STAS dissemination. The integration of SHAP analysis enhances the model's clinical interpretability and provides novel insights into the biological mechanisms of STAS.

Based on our findings, we propose a possible treatment framework for ≤ 3 cm pulmonary nodules, as illustrated in Fig. 7. This is a conceptual schema, not a clinical guideline. It is intended to inspire future prospective trials exploring the role of preoperative STAS prediction in surgical decision-making. Patients

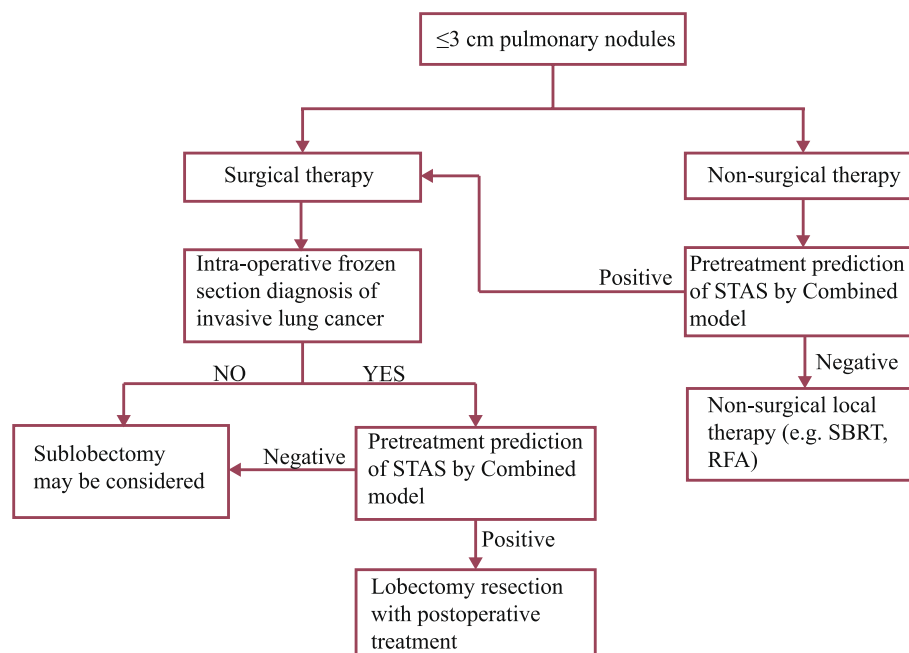


Fig. 7 A new treatment framework for ≤ 3 cm pulmonary nodules based on Combined Model. For patients with intraoperative frozen pathology considered invasive carcinoma, preoperative STAS evaluation should be performed by the model first, and sublobectomy can be performed if STAS is considered negative. If STAS is considered positive, lobectomy should be performed, and adjuvant therapy should be given postoperatively. Note: This framework is intended for conceptual demonstration only and should not be interpreted as a clinical recommendation. STAS: Spread Through Air Spaces;

identified as high-risk for STAS through the CM model can undergo adjusted surgical planning preoperatively to reduce local recurrence risk. Additionally, patients confirmed to be STAS-positive via postoperative pathology should receive aggressive adjuvant therapy, including chemotherapy, targeted therapy, or immunotherapy, depending on individual clinical characteristics.

Our study has several limitations. First, this is a retrospective study, and while it includes survival analysis, the extended follow-up period limited our ability to perform prospective validation. Second, although this is a multi-center study, the dataset consists exclusively of Chinese northern populations, and given potential population heterogeneity, future studies should incorporate samples from diverse geographic regions for broader validation. Third, radiomic feature extraction is sensitive to imaging parameters and segmentation variability. Despite standardized preprocessing, subtle differences in imaging protocols or manual segmentation procedures may affect feature reproducibility and, consequently, model performance. Therefore, developing automated segmentation models would be beneficial for future applications.

Conclusion

In this study, we developed a novel radiomics-based machine learning model incorporating both tumor and peritumoral features to preoperatively predict STAS status in stage I LUAD. The model outperformed traditional clinical feature-based models, demonstrating the potential application of radiomics in personalized surgical decision-making. Future prospective validation studies and model optimization will be crucial for the clinical translation of this technology.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40644-025-00935-4>.

Supplementary Material 1.
Supplementary Material 2.
Supplementary Material 3.

Acknowledgements

We thank all authors for their contributions to this manuscript. We also thank OnekeyAI for providing part of the technical support for this study.

Authors' contributions

Conceptualization, Yuhang Wang and Xufeng Liu; Data curation, Yuhang Wang and Xin Li; Formal analysis, Xiaojiang Zhao; Funding acquisition, Xin Li and Daqiang Sun; Investigation, Zixiao Wang and Xufeng Liu; Methodology, Yuhang Wang and Xiaojiang Zhao; Project administration, Daqiang Sun;

Resources, Zixiao Wang, Xufeng Liu and Xin Li; Software, Yuhang Wang and Xiaojiang Zhao; Supervision, Daqiang Sun; Writing – original draft, Yuhang Wang; Writing – review & editing, Daqiang Sun.

Funding

The research was funded by Tianjin Key Medical Discipline (Specialty) Construction Project(TJYXZDXK-018A).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. This study was approved by the ethical review committee of Tianjin Chest Hospital (2025YS-036-01).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Department of Thoracic Surgery, Tianjin Chest Hospital, No.261, Taierzhuang South Road, Jinnan District, Tianjin, China. ²Tianjin Chest Hospital of Tianjin University, Tianjin, China. ³Department of Cardiothoracic Surgery, Tianjin Binhai New Area Haibin People's Hospital, Tianjin, China. ⁴Department of Thoracic Surgery, Qinhuangdao First Hospital, Hebei Province, China.

Received: 31 March 2025 Accepted: 5 September 2025

Published online: 29 September 2025

References

- Siegel RL, Kratzer TB, Giaquinto AN, Sung H, Jemal A. Cancer statistics, 2025. *CA Cancer J Clin*. 2025;75(1):10–45.
- Chansky K, Detterbeck FC, Nicholson AG, et al. The IASLC lung cancer staging project: external validation of the revision of the TNM stage groupings in the eighth edition of the TNM classification of lung cancer. *J Thorac Oncol*. 2017;12(7):1109–21.
- Hung JJ, Jeng WJ, Hsu WH, et al. Prognostic factors of postrecurrence survival in completely resected stage I non-small cell lung cancer with distant metastasis. *Thorax*. 2010;65(3):241–5.
- Travis WD, Brambilla E, Burke AP, Marx A, Nicholson AG. Introduction to the 2015 World Health Organization classification of tumors of the lung, pleura, thymus, and heart. *J Thorac Oncol*. 2015;10(9):1240–2.
- Yanagawa N, Shiono S, Endo M, Ogata SY. Tumor spread through air spaces is a useful predictor of recurrence and prognosis in stage I lung squamous cell carcinoma, but not in stage II and III. *Lung Cancer*. 2018;120:14–21.
- Dai C, Xie H, Su H, et al. Tumor spread through air spaces affects the recurrence and overall survival in patients with lung adenocarcinoma >2 to 3 cm. *J Thorac Oncol*. 2017;12(7):1052–60.
- Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278(2):563–77.
- Yip SS, Aerts HJ. Applications and limitations of radiomics. *Phys Med Biol*. 2016;61(13):R150–66.
- Wu L, Lou X, Kong N, Xu M, Gao C. Can quantitative peritumoral CT radiomics features predict the prognosis of patients with non-small cell lung cancer? A systematic review. *Eur Radiol*. 2022. <https://doi.org/10.1007/s00330-022-09174-8>.
- Mu W, Jiang L, Zhang J, et al. Non-invasive decision support for NSCLC treatment using PET/CT radiomics. *Nat Commun*. 2020;11(1):5228. Published 2020 Oct 16.
- Chen D, She Y, Wang T, et al. Radiomics-based prediction for tumour spread through air spaces in stage I lung adenocarcinoma using machine learning. *Eur J Cardiothorac Surg*. 2020;58(1):51–8.
- Liao G, Huang L, Wu S, et al. Preoperative CT-based peritumoral and tumoral radiomic features prediction for tumor spread through air spaces in clinical stage I lung adenocarcinoma. *Lung Cancer*. 2022;163:87–95.
- Ding Y, Chen Y, Wen H, et al. Pretreatment prediction of tumour spread through air spaces in clinical stage I non-small-cell lung cancer. *Eur J Cardiothorac Surg*. 2022;62(3):ezac248.
- Zhuo Y, Feng M, Yang S, et al. Radiomics nomograms of tumors and peritumoral regions for the preoperative prediction of spread through air spaces in lung adenocarcinoma. *Transl Oncol*. 2020;13(10):100820.
- Qi L, Li X, He L, et al. Comparison of diagnostic performance of spread through airspaces of lung adenocarcinoma based on morphological analysis and perinodular and intranodular radiomic features on chest CT images. *Front Oncol*. 2021;11:654413.
- Tang X, Huang H, Du P, Wang L, Yin H, Xu X. Intratumoral and peritumoral CT-based radiomics strategy reveals distinct subtypes of non-small-cell lung cancer. *J Cancer Res Clin Oncol*. 2022;148(9):2247–60.
- Ran J, Cao R, Cai J, Yu T, Zhao D, Wang Z. Development and validation of a nomogram for preoperative prediction of lymph node metastasis in lung adenocarcinoma based on radiomics signature and deep learning signature. *Front Oncol*. 2021;11:585942.
- Guo QK, Yang HS, Shan SC, et al. A radiomics nomogram prediction for survival of patients with “driver gene-negative” lung adenocarcinomas (LUAD). *Radiol Med*. 2023;128(6):714–25.
- Zhao M, Kluge K, Papp L, et al. Multi-lesion radiomics of PET/CT for non-invasive survival stratification and histologic tumor risk profiling in patients with lung adenocarcinoma. *Eur Radiol*. 2022;32(10):7056–67.
- Wu L, Lou X, Kong N, Xu M, Gao C. Can quantitative peritumoral CT radiomics features predict the prognosis of patients with non-small cell lung cancer? A systematic review. *Eur Radiol*. 2023;33(3):2105–17.
- Chen Q, Shao J, Xue T, et al. Intratumoral and peritumoral radiomics nomograms for the preoperative prediction of lymphovascular invasion and overall survival in non-small cell lung cancer. *Eur Radiol*. 2022(2). <https://doi.org/10.1007/s00330-022-09109-3>.
- Liu K, Li K, Wu T, et al. Improving the accuracy of prognosis for clinical stage I solid lung adenocarcinoma by radiomics models covering tumor per se and peritumoral changes on CT. *Eur Radiol*. 2022;32(2):1065–77.
- Tunali I, Hall LO, Napel S, et al. Stability and reproducibility of computed tomography radiomic features extracted from peritumoral regions of lung cancer lesions. *Med Phys*. 2019;46(11):5075–85.
- Rodríguez-Pérez R, Bajorath J. Interpretation of machine learning models using shapley values: application to compound potency and multi-target activity predictions. *J Comput Aided Mol Des*. 2020;34(10):1013–26.
- Zhang J, Bai W, Guo C, et al. Postoperative Short-term Outcomes Between Sublobar Resection and Lobectomy in Patients with Lung Adenocarcinoma. *Cancer Manag Res*. 2020;12:9485–93.
- Altorki NK, Wang X, Wigle D, et al. Perioperative mortality and morbidity after sublobar versus lobar resection for early-stage non-small-cell lung cancer: post-hoc analysis of an international, randomised, phase 3 trial (CALGB/Alliance 140503). *Lancet Respir Med*. 2018;6(12):915–24.
- Altorki N, Wang X, Kozono D, et al. Lobar or sublobar resection for peripheral stage IA non-small-cell lung cancer. *N Engl J Med*. 2023;388(6):489–98.
- Eguchi T, Kameda K, Lu S, et al. Lobectomy is associated with better outcomes than sublobar resection in spread through air spaces (STAS)-positive T1 lung adenocarcinoma: a propensity score-matched analysis. *J Thorac Oncol*. 2019;14(1):87–98.
- Toki MI, Harrington K, Syrigos KN. The role of spread through air spaces (STAS) in lung adenocarcinoma prognosis and therapeutic decision making. *Lung Cancer*. 2020;146:127–33.
- Kadota K, Nitadori JI, Sima CS, et al. Tumor spread through air spaces is an important pattern of invasion and impacts the frequency and location of recurrences after limited resection for small stage I lung adenocarcinomas. *J Thorac Oncol*. 2015;10(5):806–14.
- Yang Y, Xie X, Wang Y, et al. A systematic review and meta-analysis of the influence of STAS on the long-term prognosis of stage I lung adenocarcinoma. *Transl Cancer Res*. 2021;10(5):2428–36.

32. Kagimoto A, Tsutani Y, Okada M. Segmentectomy for spread through air spaces-positive lung adenocarcinoma. *Ann Thorac Surg.* 2022;114(5):1989–90.
33. Ettinger DS, Wood DE, Aisner DL, et al. NCCN guidelines® insights: non-small cell lung cancer, version 2.2023. *J Natl Compr Canc Netw.* 2023;21(4):340–50.
34. Oncology Society of Chinese Medical Association. *Zhonghua Yi Xue Za Zhi.* 2024;104(34):3175–213.
35. O'Brien M, Paz-Ares L, Marreaud S, et al. Pembrolizumab versus placebo as adjuvant therapy for completely resected stage IB–IIIA non-small-cell lung cancer (PEARLS/KEYNOTE-091): an interim analysis of a randomised, triple-blind, phase 3 trial. *Lancet Oncol.* 2022;23(10):1274–86.
36. Liu C, Meng A, Xue XQ, et al. Prediction of early lung adenocarcinoma spread through air spaces by machine learning radiomics: a cross-center cohort study. *Transl Lung Cancer Res.* 2024;13(12):3443–59.
37. Zhang Z, Zhao Y, Ma YJ, et al. Prediction of STAS in lung adenocarcinoma with nodules ≤ 2 cm using machine learning: a multicenter retrospective study. *BMC Cancer.* 2025;25(1):417 Published 2025 Mar 7.
38. Ye G, Wu G, Li Y, et al. Advancing presurgical non-invasive spread through air spaces prediction in clinical stage IA lung adenocarcinoma using artificial intelligence and CT signatures. *Front Surg.* 2025;11:1511024 Published 2025 Jan 14.
39. Jiang Z, Yin J, Han P, et al. Wavelet transformation can enhance computed tomography texture features: a multicenter radiomics study for grade assessment of COVID-19 pulmonary lesions. *Quant Imaging Med Surg.* 2022;12(10):4758–70.
40. Huang K, Aviyente S. Wavelet feature selection for image classification. *IEEE Trans Image Process.* 2008;17(9):1709–20.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.