Research Square

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

Preoperative Risk Stratification of Lymph Node Metastasis in cT1 Lung Adenocarcinoma Using CT Radiomics and Habitat Analysis

Na Geng
The Fourth Affiliated Hospital of Soochow University (Suzhou Dushu Lake Hospital)
Pinchao Gu
The Fourth Affiliated Hospital of Soochow University (Suzhou Dushu Lake Hospital)
Haitao Ma
18862177620@163.com

The Fourth Affiliated Hospital of Soochow University (Suzhou Dushu Lake Hospital)

Research Article

Keywords: Lung adenocarcinoma, lymph node metastasis, radiomics, habitat analysis, computed tomography

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11097058/v1

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

Additional Declarations: No competing interests reported.

Abstract

Background

Accurate preoperative assessment of lymph node metastasis (LNM) is important for risk stratification and surgical planning in patients with cT1 lung adenocarcinoma. This study aimed to develop and externally validate a multi-regional computed tomography (CT) radiomics model incorporating intratumoral, peritumoral, and habitat information for preoperative prediction of LNM.

Methods

A total of 172 patients with cT1 invasive lung adenocarcinoma from one institution were retrospectively included and randomly divided into a training cohort (n = 120) and an internal validation cohort (n = 52). An independent cohort of 63 patients from a second institution was used for external validation. Radiomic features were extracted from the intratumoral region, a 3-mm peritumoral region, and two habitat subregions generated using K-means clustering. After feature selection, a multi-regional radiomic signature (Radscore) was constructed and combined with preoperative clinical variables using logistic regression. Model performance was evaluated using the area under the receiver operating characteristic curve (AUC), sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV).

Results

Nineteen radiomic features were retained in the final radiomic signature. In the internal validation cohort, the clinical, radiomic, and combined models achieved AUCs of 0.758, 0.844, and 0.852, respectively. Using a prespecified training-derived threshold of 0.4036, the combined model achieved a sensitivity of 65.0%, specificity of 87.5%, and NPV of 80.0% in the internal validation cohort. In the independent external cohort, the combined model achieved an AUC of 0.837 (95% CI: 0.731–0.930), with an accuracy of 77.8%, sensitivity of 77.3%, specificity of 78.0%, PPV of 65.4%, and NPV of 86.5%. The Radscore was significantly associated with IASLC pathological grade and visceral pleural invasion (both $P < 0.001$).

Conclusions

The multi-regional CT radiomics model incorporating intratumoral, peritumoral, and habitat information demonstrated good discrimination for preoperative LNM risk stratification in cT1 lung adenocarcinoma and maintained stable performance in an independent external cohort. The model may provide complementary quantitative information for preoperative risk assessment, although further prospective validation is warranted before clinical implementation.

Highlight box

Key findings

• A multi-regional CT radiomics model integrating intratumoral, peritumoral, and habitat features achieved an external validation AUC of 0.837 for predicting LNM in cT1 lung adenocarcinoma.

• In the independent external cohort, sensitivity, specificity, and NPV were 77.3%, 78.0%, and 86.5%, respectively.

What is known and what is new?

• Conventional imaging has limited sensitivity for occult nodal metastasis in early-stage lung adenocarcinoma.

• This study integrates intratumoral, peritumoral, and habitat radiomic information and demonstrates maintained discrimination and acceptable calibration in an independent external cohort.

What is the implication, and what should change now?

• Multi-regional radiomics may provide complementary quantitative information for preoperative LNM risk stratification.

• The model should be considered an adjunct rather than a replacement for established nodal staging and intraoperative assessment.

1 Introduction

Lung cancer is the malignant tumor with the highest morbidity and mortality worldwide, and lung adenocarcinoma is its most common histological subtype. Lymph node metastasis is a critical factor affecting the prognosis of patients with lung adenocarcinoma and serves as a key determinant for the surgical approach and the extent of lymph node dissection. For patients with cT1 stage (tumor maximum diameter ≤ 3 cm) lung adenocarcinoma, accurate preoperative assessment of lymph node metastasis risk is of paramount clinical importance. With the publication of results from large randomized controlled trials such as CALGB 140503[1] and JCOG0802, anatomic segmentectomy has been proven non-inferior for certain early-stage non-small cell lung cancers, provided that no micrometastasis is present in the systematic lymph nodes.

Currently, commonly used preoperative assessment tools have notable limitations. Conventional CT has limited capability in detecting micrometastases in normal-sized lymph nodes, PET/CT suffers from insufficient sensitivity and a high false-positive rate for small lesions, and invasive procedures like endobronchial ultrasound are complex. Consequently, there is an urgent clinical need for a non-invasive, accurate, and reproducible preoperative assessment tool. Radiomics can reveal microscopic heterogeneity information within tumors through high-throughput extraction of quantitative features from medical images[2]. However, most studies have focused solely on the primary tumor, neglecting the value of the peritumoral microenvironment and spatial heterogeneity. Research has shown that combining intratumoral and peritumoral radiomic features can improve predictive performance[3], while voxel-based clustering habitat radiomics can decode spatial heterogeneity at a deeper level[4][5]. Based on this, our study aims to integrate radiomic features from three dimensions—intratumoral, peritumoral, and habitat—and combine them with preoperative clinical variables to construct a preoperative prediction model for lymph node metastasis in cT1 stage lung adenocarcinoma, and to explore its clinical application value and biological basis[6][7].

2 Materials and Methods

2.1 Clinical Data

Clinical and imaging data from 172 patients with cT1 invasive lung adenocarcinoma who underwent surgical resection at the Department of Thoracic Surgery, The Fourth Affiliated Hospital of Soochow University, between 2016 and 2025 were retrospectively analyzed. The definition of cT1 disease was based on the 8th edition of the TNM staging system [8,9]. The inclusion criteria were as follows: (I) histopathologically confirmed invasive lung adenocarcinoma after surgery; (II) complete preoperative chest CT imaging data; (III) clinical stage cT1, defined as a maximum tumor diameter of ≤ 3 cm; and (IV) lobectomy or sublobar resection with systematic lymph node dissection or sampling. The exclusion criteria were: (I) receipt of neoadjuvant chemotherapy and/or radiotherapy before surgery; (II) poor CT image quality; (III) incomplete clinical or pathological data; and (IV) adenocarcinoma in situ (AIS) or minimally invasive adenocarcinoma (MIA). These indolent subtypes were excluded to obtain a more homogeneous cohort of invasive adenocarcinomas because they are associated with a very low risk of lymph node metastasis and could otherwise introduce substantial clinical heterogeneity.

A total of 172 patients were finally included in the development cohort, including 66 patients with lymph node metastasis (LNM) (38.4%) and 106 without LNM (61.6%). These patients were randomly divided into a training cohort (n = 120) and an internal validation cohort (n = 52) at a ratio of approximately 7:3.

For independent external validation, an additional cohort of 63 patients was retrospectively collected from the Second Affiliated Hospital of Soochow University between 2018 and 2025 using the same inclusion and exclusion criteria. The external

validation cohort included 22 patients with LNM (34.9%) and 41 patients without LNM (65.1%).

Pathological lymph node metastasis (LNM) was defined as the presence of tumor cells, including macrometastases or micrometastases (> 0.2 mm), in any examined hilar (N1) or mediastinal (N2) lymph node on postoperative histopathological evaluation. All patients underwent systematic hilar and mediastinal lymph node dissection at the time of pulmonary resection. Nodal assessment included both N1 and N2 stations, and all resected hilar, intrapulmonary, and mediastinal lymph nodes were submitted for histopathological examination. Across the study cohort, a median of 3 lymph node stations and 7 lymph nodes were examined per patient. Patients without pathological evidence of tumor cells in any examined lymph node were classified as node-negative. The overall study design and analytical workflow are illustrated in Fig. 1.

2.2 CT Image Acquisition and Preprocessing

All patients underwent preoperative thin-section chest CT plain scans using a standardized scanning protocol. The CT scanners used included GE Discovery CT750 HD, Siemens SOMATOM Force, and Philips Brilliance 64, with the following acquisition parameters: tube voltage, 120 kVp; tube current, 200–350 mAs (auto-mA adjustment applied); rotation time, 0.5 s; pitch, 0.8–1.0; and slice thickness, 0.625–1.25 mm. All images were reconstructed with a matrix size of 512 × 512 using soft-tissue reconstruction kernels (Standard for GE, B31f for Siemens, and B for Philips). All scans were performed with the patient in the supine position with arms raised above the head, during end-inspiratory breath-hold, without intravenous contrast administration. Image preprocessing was performed using Python 3.9 and the SimpleITK library, resampling all images to isotropic voxels of 1 × 1 × 1 mm^3 to ensure the stability of radiomic features [2].

2.3 Region of Interest Segmentation and Spatial Habitat Partitioning

The intratumoral region was manually delineated slice by slice by two chest imaging physicians using 3D Slicer software.The peritumoral region was defined by uniformly expanding outward 3 mm from the intratumoral region.For the three-dimensional spatial habitats, an unsupervised K-means algorithm was applied to automatically cluster voxel-level feature maps (with the number of clusters set to 2), The optimal number of clusters (k = 2) was determined using the silhouette coefficient and clinical interpretability to effectively decouple the tumor into high-density (solid) and low-density (GGO) subregions. To ensure clustering stability, the K-means algorithm was executed 10 times with different random centroid initializations, and the most stable partition was selected." partitioning the tumor into a high-density habitat (Label 1) and a low-density habitat (Label 2).

2.4 Radiomic Feature Extraction and Selection

PyRadiomics was used to extract features from the four ROIs (intratumoral, peritumoral, habitat 1, and habitat 2). To assess the reproducibility of ROI segmentation, 30 cases were randomly selected for inter-observer and intra-observer reliability analysis. Two chest radiologists independently delineated the intratumoral regions, and one of them repeated the delineation after a 2-week interval. The intraclass correlation coefficient (ICC) was calculated for all extracted radiomic features. Features with ICC > 0.75 were considered to have good reproducibility and were retained for further analysis. Feature extraction was performed using PyRadiomics (version 3.0.1), which is fully compliant with the Image Biomarker Standardization Initiative (IBSI) guidelines. A fixed bin width of 25 Hounsfield units (HU) was used for image gray-level discretization, and no voxel intensity normalization was applied to preserve the raw density characteristics of the CT images. A two-step feature selection method was employed: first, features with $P < 0.05$ were screened using a t-test; then, least absolute shrinkage and selection operator (LASSO) regression with 10-fold cross-validation was applied to select core features with non-zero coefficients. The optimal penalty parameter λ was selected as the value that minimized the cross-validated mean squared error (i.e., lambda.min). The LASSO cross-validation curves for each region of interest, with the lambda.min position indicated by a vertical dashed line.

2.5 Multi-regional Feature Fusion and Model Construction

The radiomic features selected from the intratumoral, peritumoral, Habitat 1, and Habitat 2 regions were integrated into a unified multi-regional feature matrix. After standardization using parameters derived from the training cohort, logistic regression was used to construct the multi-regional radiomic model and generate an individual radiomic score (Radscore).

Preoperative clinical variables were evaluated in the training cohort, and selected clinical predictors were used to construct the Clinical model. Finally, the standardized Radscore and selected clinical variables were jointly incorporated into a multivariable logistic regression model to construct the Combined model.

2.6 Statistical Analysis and Pathological Correlation Validation

Model performance was evaluated using AUC, accuracy, sensitivity, and specificity in the validation set. The AUCs were compared using the DeLong test, and a DCA curve was plotted to assess net clinical benefit. The preoperative Radscore was correlated with the gold standard of postoperative pathology: the Kruskal-Wallis test was used to analyze its distribution across different IASLC pathological grades (grades 1–3), and the Mann-Whitney U test was used for different VPI statuses. A two-sided $P < 0.05$ was considered statistically significant. This study is reported in accordance with the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) statement.The classification threshold for the Combined model was determined from the training data and prespecified at 0.4036. This fixed threshold was subsequently applied without modification to both the internal and external validation cohorts. No threshold optimization was performed in either validation cohort. Model performance was reported using AUC, accuracy, sensitivity, specificity, PPV, and NPV.Model calibration was assessed using calibration curves and the Brier score. Calibration intercept and calibration slope were additionally estimated using logistic recalibration, with ideal values of 0 and 1, respectively.

3 Results

3.1 Radiomic Feature Selection and ROI Segmentation

A total of four regions of interest (ROIs) were evaluated for each patient, including the intratumoral region, the 3-mm peritumoral region, and two spatial habitat subregions (high-density Habitat 1 and low-density Habitat 2) generated using unsupervised K-means clustering. Representative examples of the original CT images, manual tumor segmentation, and the derived habitat subregions are shown in Fig. 2.

Following the two-stage feature selection procedure, consisting of t-test screening followed by LASSO regression with 10-fold cross-validation, 5, 8, 9, and 5 features were retained from the intratumoral, peritumoral, Habitat 1, and Habitat 2 regions, respectively. After integrating features from the four regions and removing duplicated features, 19 unique radiomic features with non-zero coefficients were retained for construction of the final radiomic signature. The LASSO cross-validation curves for each ROI are shown in Fig. 3.

3.2 Clinical Variable Screening and Model Contribution

Univariate logistic regression analysis of preoperative clinical variables was performed in the training cohort to identify candidate predictors of lymph node metastasis. The baseline clinical and pathological characteristics of patients in the training and validation cohorts are summarized in Table 1.

Table 1
Baseline characteristics of patients in the training and validation sets

Characteristic		Total (N = 172)	Training set (N = 120)	Validation set (N = 52)	P value
Age		61.38 ± 10.72	62.41 ± 10.32	59.00 ± 11.33	0.055
SII		671.01 ± 597.56	696.14 ± 549.82	613.03 ± 697.84	0.404
Fibrinogen		3.22 ± 0.92	3.25 ± 1.00	3.16 ± 0.70	0.567
Sex	Female	80 (46.5%)	48 (40.0%)	32 (61.5%)	0.015
	male	92 (53.5%)	72 (60.0%)	20 (38.5%)	
CTR	≤ 0.5	31 (18.0%)	20 (16.7%)	11 (21.2%)	0.626
	> 0.5	141 (82.0%)	100 (83.3%)	41 (78.8%)	
LNM Status	Positive	66 (38.4%)	46 (38.3%)	20 (38.5%)	0.985
	Negative	106 (61.6%)	74 (61.7%)	32 (61.5%)	
Types of Nodules	Pure Ground-Glass	7 (4.1%)	4 (3.3%)	3 (5.8%)	0.668
	Mixed Ground-Glass	46 (26.7%)	31 (25.8%)	15 (28.8%)	
	Solid	119 (69.2%)	85 (70.8%)	34 (65.4%)	
Tumor diameter	0 < size ≤ 1	13 (7.6%)	7 (5.8%)	6 (11.5%)	0.0399
	1 < size ≤ 2	75 (43.6%)	47 (39.2%)	28 (53.8%)	
	2 < size ≤ 3	84 (48.8%)	66 (55.0%)	18 (34.6%)	

Nodule type (P = 0.0003), consolidation-to-tumor ratio (CTR; P = 0.0036), and tumor diameter (P = 0.0433) were significantly associated with lymph node metastasis in univariate analysis. Fibrinogen (P = 0.0835) did not reach statistical significance and was therefore not included in the final model. Sex (P = 0.0558) was retained because of its potential clinical relevance to lung cancer risk stratification. Tumor diameter was not included in the final combined model to reduce potential redundancy with CT-derived radiomic information. Consequently, nodule type, sex, and CTR were used as the clinical predictors for construction of the Clinical and Combined models.

In the final combined clinical-radiomics model, the multi-regional Radscore showed the largest standardized coefficient among the included predictors, suggesting that the radiomic signature contributed substantially to the overall predictive signal, whereas the clinical variables provided additional complementary information.

3.3 Model Predictive Performance

The performance of the Clinical, Radiomic, and Combined models in the training and internal validation cohorts is summarized in Table 2.

Table 2
Performance comparison of clinical, radiomic, and combined models in the training and validation cohorts, with the CV-optimized threshold for the combined model

Model	Cohort	TP/TN/FP/FN	AUC (95% CI)	Accuracy(%)	Sensitivity(%)	Specificity(%)	PPV (%)	NPV(%)
Clinical	Training	30/52/22/16	0.734 (0.647–0.814)	68.3	65.2	70.3	57.7	76.5
	Validation	8/26/6/12	0.758 (0.629–0.866)	65.4	40.0	81.3	57.1	68.4
Radiomic (Radscore)	Training	28/58/16/18	0.846 (0.767–0.906)	71.7	60.9	78.4	63.6	76.3
	Validation	10/30/2/10	0.844 (0.729–0.935)	76.9	50.0	93.8	83.3	75.0
Combined (Threshold = 0.5)	Training	31/58/16/15	0.849 (0.772–0.907)	74.2	67.4	78.4	66.0	79.5
	Validation	10/29/3/10	0.852 (0.742–0.944)	75.0	50.0	90.6	76.9	74.4
Combined (Threshold = 0.404)	Validation	13/28/4/7	0.852 (0.742–0.944)	78.8	65.0	87.5	76.5	80.0
Combined (Threshold = 0.404)	External Validation	17/32/9/5	0.837 (0.731–0.930)	77.8	77.3	78.0	65.4	86.5

In the internal validation cohort, the standalone Clinical model achieved an AUC of 0.758 (95% CI: 0.629–0.866), with a sensitivity of 40.0%, specificity of 81.3%, and negative predictive value (NPV) of 68.4%. The standalone Radiomic model achieved an AUC of 0.844 (95% CI: 0.729–0.935), with a sensitivity of 50.0%, specificity of 93.8%, and NPV of 75.0%. The Combined model achieved an AUC of 0.852 (95% CI: 0.742–0.944), with an accuracy of 75.0%, sensitivity of 50.0%, specificity of 90.6%, and NPV of 74.4% at the default threshold of 0.5 (Fig. 4, Table 2). The DeLong test showed no statistically significant difference in AUC between the Combined and Radiomic models (P = 0.676) or between the Combined and Clinical models (P = 0.107).

To further assess the potential clinical utility of the Combined model, threshold optimization was performed using 5-fold cross-validation within the training cohort, with the optimal threshold determined by maximizing the Youden index. The mean optimal threshold across the five folds was 0.4036. When this pre-specified threshold was applied to the internal validation cohort, sensitivity increased to 65.0% and accuracy to 78.8%, while specificity remained 87.5% and NPV increased to 80.0% (Table 2). The number of false-negative cases decreased from 10 to 7, corresponding to a 30% reduction in missed nodal metastases, while false-positive cases increased from 3 to 4. These findings suggest that adjustment of the decision threshold may allow the model to be adapted to different clinical priorities, particularly when greater emphasis is placed on reducing missed lymph node metastases.

For external validation, the Combined model was evaluated in an independent cohort of 63 patients from a separate institution using the same prespecified threshold of 0.4036. The external cohort included 22 patients with LNM and 41 patients without LNM. The model achieved an AUC of 0.837 (95% CI: 0.731–0.930). At the prespecified threshold, the model correctly identified

17 true-positive and 32 true-negative cases, with 9 false-positive and 5 false-negative cases. The corresponding accuracy, sensitivity, specificity, PPV, and NPV were 77.8%, 77.3%, 78.0%, 65.4%, and 86.5%, respectively. These findings indicate that the discriminative performance of the model was maintained in the independent external cohort, with a relatively balanced sensitivity and specificity profile.

Calibration analysis showed reasonable agreement between predicted and observed probabilities. In the internal validation cohort, the Brier score was 0.167, with a calibration intercept of 0.569 and a calibration slope of 1.097. In the independent external validation cohort, the corresponding values were 0.158, 0.068, and 0.834, respectively (Fig. 6B,C).

Decision curve analysis suggested potential net clinical benefit of the Combined model across a range of clinically relevant threshold probabilities (Fig. 5). To facilitate clinical interpretation of the model, a logistic regression coefficient plot was generated to display the coefficients of the Radscore and clinical predictors included in the Combined model (Fig. 6A).

3.4 Pathological Correlation Validation

To validate the biological plausibility of the radiomic signature, we examined the association between the Radscore and postoperative pathological indicators of tumor aggressiveness. The Radscore showed a highly significant difference among IASLC pathological grades ($H = 34.12$, $P < 0.001$), exhibiting a stepwise increase with higher grades (Grade 1 vs. Grade 2 vs. Grade 3, all pairwise $P < 0.05$). The median Radscore of visceral pleural invasion (VPI)-positive patients (0.59) was significantly higher than that of VPI-negative patients (0.24, $P < 0.001$) (Fig. 7). These findings demonstrated significant associations between the Radscore and pathological indicators of tumor aggressiveness, supporting the pathological plausibility of the radiomic signature. However, these associations should be interpreted as exploratory pathological correlations rather than direct evidence of underlying biological mechanisms.

4 Discussion

This study integrated intratumoral, peritumoral, and voxel-level habitat radiomic features to construct a clinical-radiomics model for preoperative prediction of LNM in cT1 lung adenocarcinoma, with independent external validation and exploratory pathological correlation analysis.

4.1 Spatial Microenvironmental Value of Multi-Regional Radiomics

Traditional radiomics studies often focus solely on the intratumoral region, neglecting the peritumoral microenvironment that critically influences tumor invasion and metastasis[10][11]. In this study, we introduced a 3-mm peritumoral expansion to precisely capture information from the tumor invasive margin, where early lymphatic invasion is most likely to occur. More importantly, through unsupervised K-means clustering-based habitat analysis, we achieved spatial decoupling of "ground-glass opacity subregions" and "solid dense subregions" without prior knowledge, significantly reducing the redundancy of intratumoral heterogeneity information[12][13]. This voxel-level habitat approach enabled the Radscore to more purely reflect the invasive potential of the tumor, which may explain why the radiomic model alone achieved a validation AUC of 0.844—comparable to or even exceeding many previously published models.

4.2 Clinical Decision-Making Implications

Against the backdrop of landmark studies such as JCOG0802 and JCOG1211 establishing the role of sublobar resection for selected patients with early-stage lung cancer [6][7], accurate preoperative assessment of nodal status remains an important component of surgical planning and risk stratification. In the present study, the combined clinical-radiomics model achieved an AUC of 0.852 in the internal validation cohort. At the default threshold of 0.5, the model showed a specificity of 90.6%, whereas application of the prespecified cross-validation-derived threshold of 0.4036 increased sensitivity from 50.0% to 65.0% while maintaining a specificity of 87.5%. This threshold adjustment reduced the number of false-negative cases from 10 to 7, corresponding to a 30% reduction in missed nodal metastases, with only one additional false-positive case (from 3 to 4). These findings illustrate the potential value of threshold selection in balancing the competing clinical consequences of missed

nodal metastases and unnecessary additional nodal evaluation or treatment [14]. Notably, the combined model did not demonstrate a statistically significant improvement in AUC compared with the standalone radiomic model (DeLong test, P > 0.05), suggesting that the multi-regional radiomic signature captured a substantial proportion of the discriminatory information, while clinical variables provided complementary information for individualized risk assessment. Importantly, in the expanded independent external validation cohort of 63 patients, the combined model maintained good discriminative performance, achieving an AUC of 0.837 (95% CI: 0.731–0.930). Using the same prespecified threshold of 0.4036 without recalibration, the model achieved a sensitivity of 77.3%, specificity of 78.0%, accuracy of 77.8%, positive predictive value of 65.4%, and negative predictive value of 86.5%. The relatively balanced sensitivity and specificity observed in the external cohort provide additional support for the transportability of the model across institutions, although further prospective validation in larger and more heterogeneous populations remains necessary. Therefore, the proposed model should be regarded as an exploratory adjunctive tool for preoperative LNM risk stratification rather than a replacement for established clinical assessment, conventional nodal staging, surgical judgment, or intraoperative pathological evaluation [15]. Its potential clinical value lies in providing an additional layer of quantitative imaging information that may complement existing clinical and radiological assessments and support more individualized preoperative decision-making regarding surgical strategy and lymph node management.

4.3 pathological correlation Validation

Radiomics is frequently characterized as an uninterpretable "black box". To address this challenge, our study establishes a robust statistical association between the preoperative Radscore and postoperative pathological gold standards. The stepwise increase of the Radscore with IASLC pathological grade and its significant elevation in patients with visceral pleural invasion indicate that the radiomic signature indeed captures tumor biological features such as high proliferative capacity, cellular dedifferentiation, and tissue invasiveness. This pathological correlation provides a solid biological foundation for the clinical applicability of the model and enhances its credibility among thoracic surgeons.

4.4 Limitations and Future Directions

Several limitations should be acknowledged. First, this was a retrospective study with a relatively modest sample size. Although the independent external validation cohort was expanded to 63 patients, including 22 LNM-positive cases, larger prospective multicenter cohorts are still required to confirm model transportability and clinical applicability.

Second, the prevalence of LNM in the development cohort was relatively high for cT1 lung adenocarcinoma, which may reflect referral patterns, patient selection, or the exclusion of adenocarcinoma in situ and minimally invasive adenocarcinoma. Because predictive values are influenced by disease prevalence, PPV and NPV should be interpreted in the context of the populations studied.

Third, differences in CT scanners and acquisition protocols between institutions may introduce variability in radiomic features. Although all CT images were resampled to a uniform voxel size before feature extraction, further harmonization and validation across additional imaging platforms are warranted.

Fourth, the 3-mm peritumoral expansion and the predefined number of habitat clusters (k = 2) were selected within the current analytical framework and may not represent the optimal configuration for all patients. Future studies should systematically evaluate different peritumoral distances and habitat partitioning strategies.

Fifth, some clinical predictors, including nodule type and CTR, were also derived from CT images and may partially overlap with information captured by radiomic features. This may partly explain why the Combined model did not significantly outperform the standalone Radiomic model.

Finally, although the model demonstrated stable discrimination in the external cohort, it should not currently be used as a standalone basis for surgical de-escalation or omission of conventional nodal assessment. Prospective studies are needed to determine whether integration of the model into clinical workflows provides incremental benefit beyond established clinical and radiological assessment.

5 Conclusion

The multi-regional CT radiomics model integrating intratumoral, peritumoral, and habitat features with clinical variables demonstrated good discriminative performance for preoperative LNM risk stratification in cT1 lung adenocarcinoma. Its performance was maintained in an independent external cohort of 63 patients, with an AUC of 0.837 and balanced sensitivity and specificity. The observed associations between the Radscore and pathological indicators of tumor aggressiveness further support the plausibility of the imaging signature. The model may serve as an adjunctive quantitative tool for preoperative risk assessment, although prospective multicenter validation is required before clinical implementation.

Declarations

Ethics approval and consent to participate:

This retrospective study was conducted using anonymized CT imaging and clinical data from patients who underwent surgical resection for cT1 lung adenocarcinoma at the Fourth Affiliated Hospital of Soochow University and the Second Affiliated Hospital of Soochow University. The study did not involve any direct patient intervention, did not cause harm to human subjects, and did not involve sensitive personal information or commercial interests. In accordance with Article 32 of the Measures for Ethical Review of Life Science and Medical Research Involving Humans (National Health Commission [2023] No. 4), the Institutional Review Board of the Fourth Affiliated Hospital of Soochow University reviewed the protocol and confirmed that this study met the criteria for exemption from ethical review, including waiver of written informed consent. All patient data were de-identified and analyzed anonymously. All methods were performed in accordance with the relevant guidelines and regulations of the Declaration of Helsinki.

Consent for publication:

Not applicable.

Availability of data and materials:

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

Competing interests:

The authors declare that they have no competing interests.

Funding:

None.

Authors' contributions:

NG and PG performed the data analysis and drafted the manuscript. HM designed the study and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements:

None.

References

1. Altorki N, Wang X, Kozono D, et al. Lobectomy versus Sublobar Resection for Small-Sized Peripheral Non-Small-Cell Lung Cancer. N Engl J Med. 2023;388(6):489–98.

2. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer. 2012;48(4):441–6.
3. Beig N, Khorrami M, Alilou M, et al. Peritumoral and Intratumoral Radiomics Features on Clinical CT Images Predict Lobe-Specific Recurrence-Free Survival of Non-Small Cell Lung Cancer. Radiology. 2019;290(3):783–92.
4. Wu J, Cao G, Sun X, et al. Intratumor Heterogeneity Characterized by Reconstruction of Phase Survival Habitats on CT Images Associates with Metastasis and Survival in Non-Small Cell Lung Cancer. Clin Cancer Res. 2018;24(21):5293–301.
5. Moreira AL, Ocampo PSS, Xia Y, et al. A Grading System for Invasive Pulmonary Adenocarcinoma Proposed by the International Association for the Study of Lung Cancer Pathology Committee. J Thorac Oncol. 2020;15(10):1599–610.
6. Saji H, Okada M, Tsuboi M, et al. Segmentectomy versus lobectomy in small-sized peripheral non-small-cell lung cancer (JCOG0802/WJOG4607L). Lancet. 2022;399(10335):1607–17.
7. Aokage K, Suzuki K, Saji H, et al. Segmentectomy for ground-glass-dominant lung cancer with a tumour diameter of 3 cm or less (JCOG1211). Lancet Respir Med. 2023;11(6):540–9.
8. Suzuki K, Koike T, Asakawa T, et al. A prospective radiological study of thin-section computed tomography to predict pathological noninvasiveness in peripheral clinical IA lung cancer (JCOG0201). J Thorac Oncol. 2011;6(4):751–6.
9. Goldstraw P, Chansky K, Crowley J, et al. The IASLC Lung Cancer Staging Project: Proposals for the Revision of the TNM Stage Groupings in the Forthcoming (Eighth) Edition. J Thorac Oncol. 2016;11(1):39–51.
10. Wang X, Ji X, Sen L, et al. Can peritumoral radiomics increase the efficiency of the prediction for lymph node metastasis in clinical stage T1 lung adenocarcinoma on CT? Eur Radiol. 2019;29(11):6084–90.
11. Huang X, Xue Y, Deng B, et al. Predicting pathological grade of stage I pulmonary adenocarcinoma: a CT radiomics approach. Front Oncol. 2024;14:1355–64.
12. Moreira AL, Ocampo PSS, et al. Integrating IASLC grading and radiomics for predicting postoperative outcomes in stage IA invasive lung adenocarcinoma. Med Phys. 2024;51(8):5138–51.
13. Altorki N, Wang X, Kozono D, et al. Habitat Imaging With Tumoral and Peritumoral Radiomics for Prediction of Lung Adenocarcinoma Invasiveness on Preoperative Chest CT: A Multicenter Study. AJR Am J Roentgenol. 2024;223(2):1061–70.
14. Wu S, Jin Z, et al. Evaluating peritumoral and intratumoral radiomics signatures for predicting lymph node metastasis in surgically resectable non-small cell lung cancer. Front Oncol. 2024;14:1046–58.
15. Cao G, Sun X, et al. CT-based machine learning model integrating intra- and peri-tumoral radiomics features for predicting occult lymph node metastasis in peripheral lung cancer. BMC Med Imaging. 2025;25:1688–98.

Figures

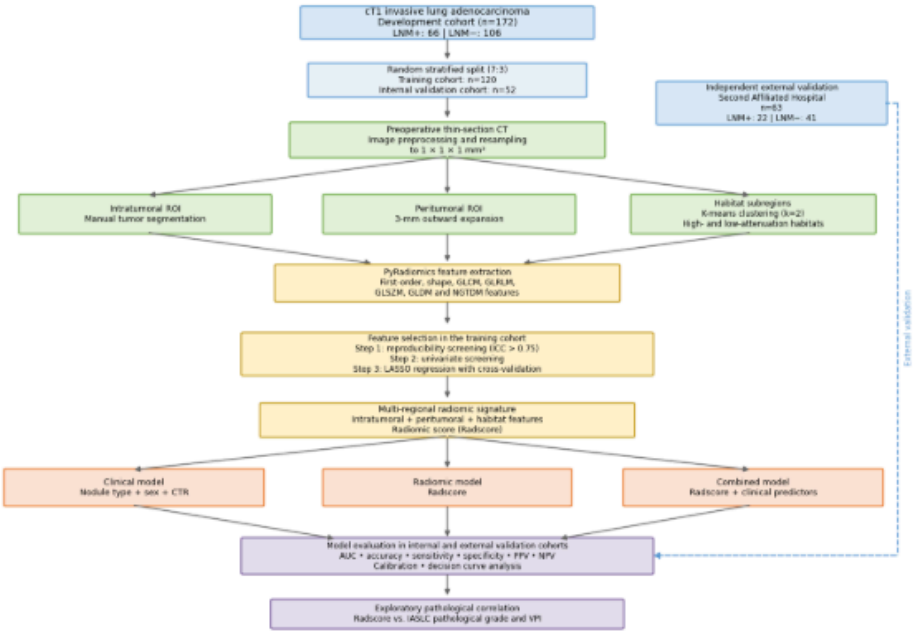


Figure 1

Study flowchart.

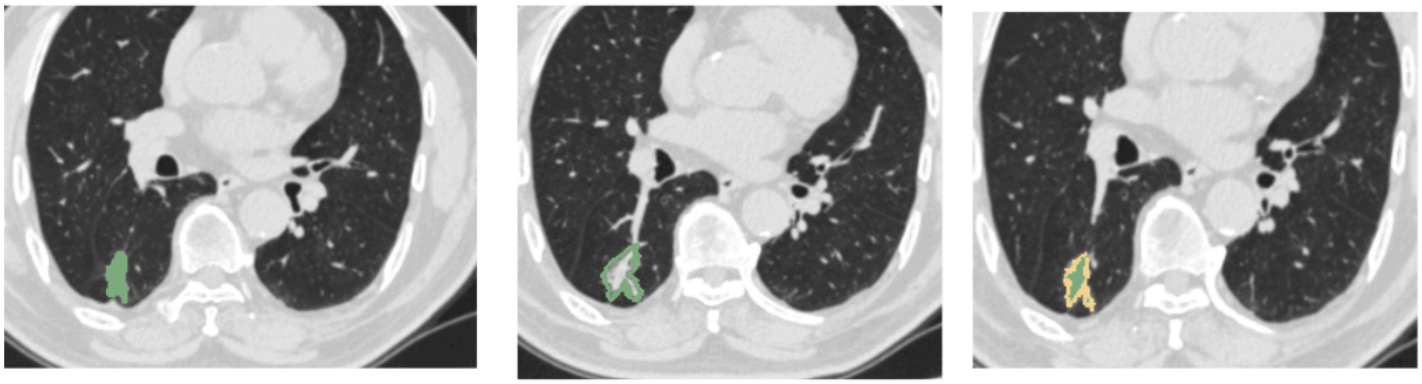


Figure 2. Representative CT images showing ROI segmentation and habitat subregions.

Figure 2

Representative CT images showing ROI segmentation and habitat subregions.

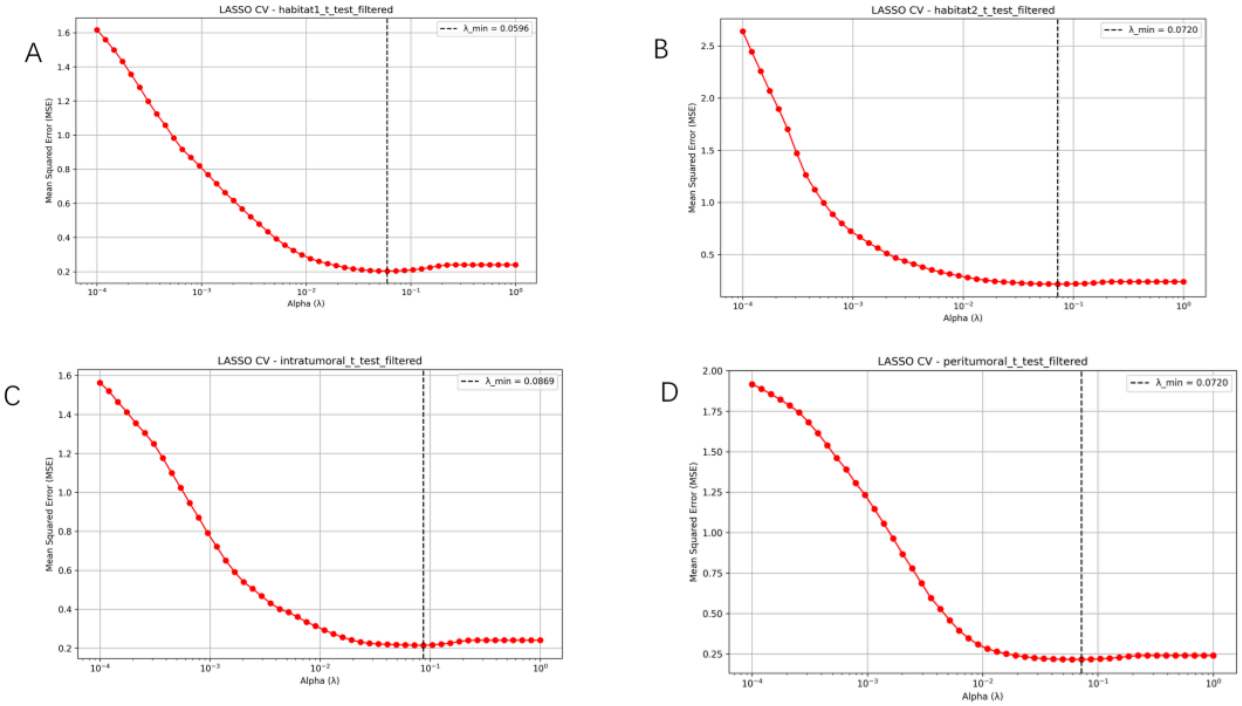


Figure 3. Feature selection using the least absolute shrinkage and selection operator (LASSO) regression with 10-fold cross-validation for each region of interest

(A) Habitat 1 (high-density subregion, 9 features). (B) Habitat 2 (low-density subregion, 5 features)(C) Peritumoral region (8 features). (D Intratumoral region (5 features). The vertical dashed line in each panel indicates the optimal λ selected by the lambda.min criterion, i.e., the value that minimizes the cross-validated mean squared error.

Figure 3

Feature selection using the least absolute shrinkage and selection operator (LASSO) regression with 10-fold cross-validation for each region of interest.

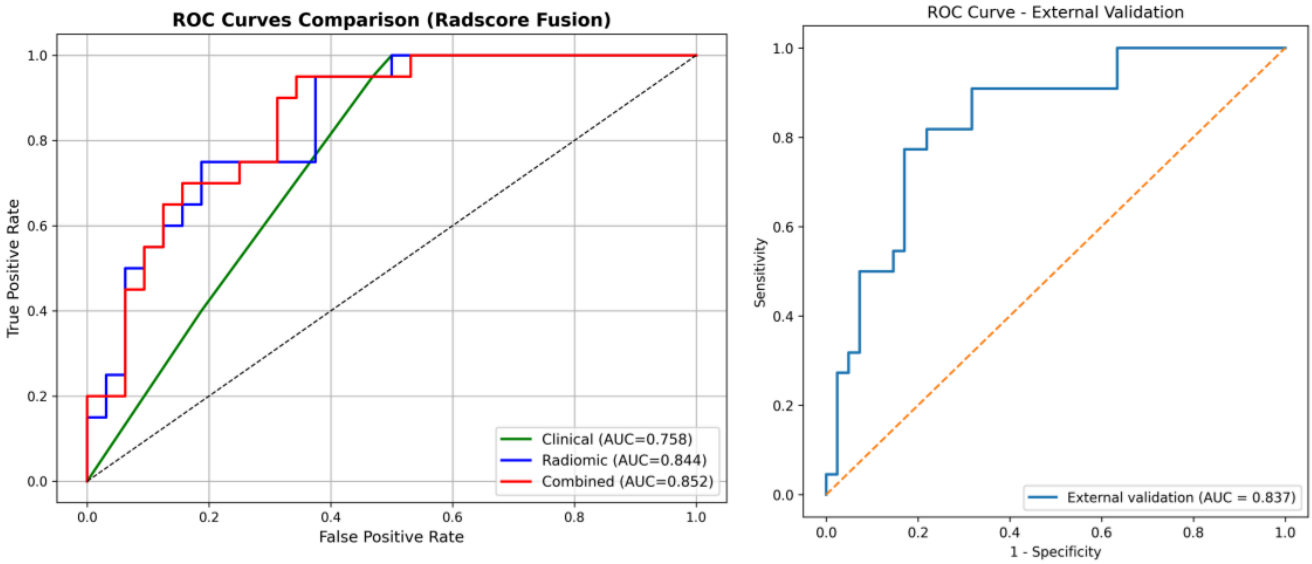


Figure 4. Receiver operating characteristic (ROC) curves of the prediction models in the internal and external validation cohorts.

(A) ROC curves of the Clinical, Radiomic, and Combined models in the internal validation cohort, with AUCs of 0.758, 0.844, and 0.852, respectively. (B) ROC curve of the Combined model in the independent external validation cohort (n=63), with an AUC of 0.837 (95% CI: 0.731–0.930).

Figure 4

ROC curves of the clinical, radiomic, and combined models in the internal validation cohort, with the ROC curve for the combined model in the external validation cohort.

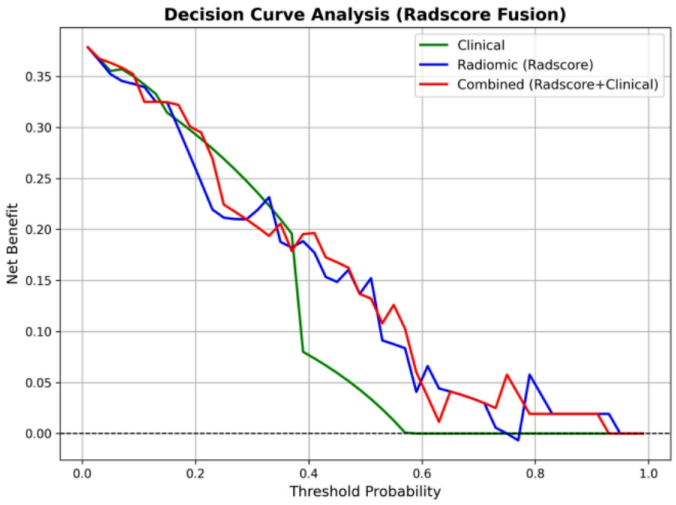


Figure 5. Decision curve analysis (DCA) for the clinical model, radiomic model, and combined （rad-score+clinical） model in the validation set.

Figure 5

Decision curve analysis (DCA) for the clinical model, radiomic model, and combined (rad-score+clinical) model in the validation set.

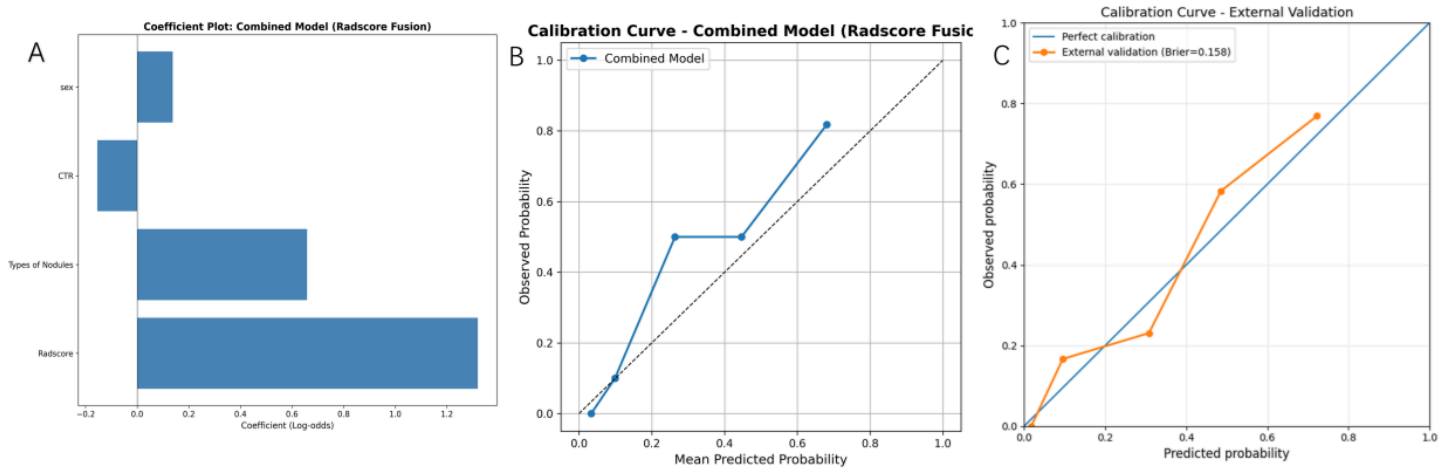


Figure 6. Model coefficients and calibration performance of the Combined model.
(A) Logistic regression coefficient plot showing the coefficients of the Radscore and clinical predictors in the Combined model. (B) Calibration curve in the internal validation cohort; the Brier score, calibration intercept, and calibration slope were 0.167, 0.569, and 1.097, respectively. (C) Calibration curve in the independent external validation cohort; the corresponding values were 0.158, 0.068, and 0.834, respectively. The diagonal dashed line represents perfect calibration.

Figure 6

Feature contribution analysis and calibration of the combined model.

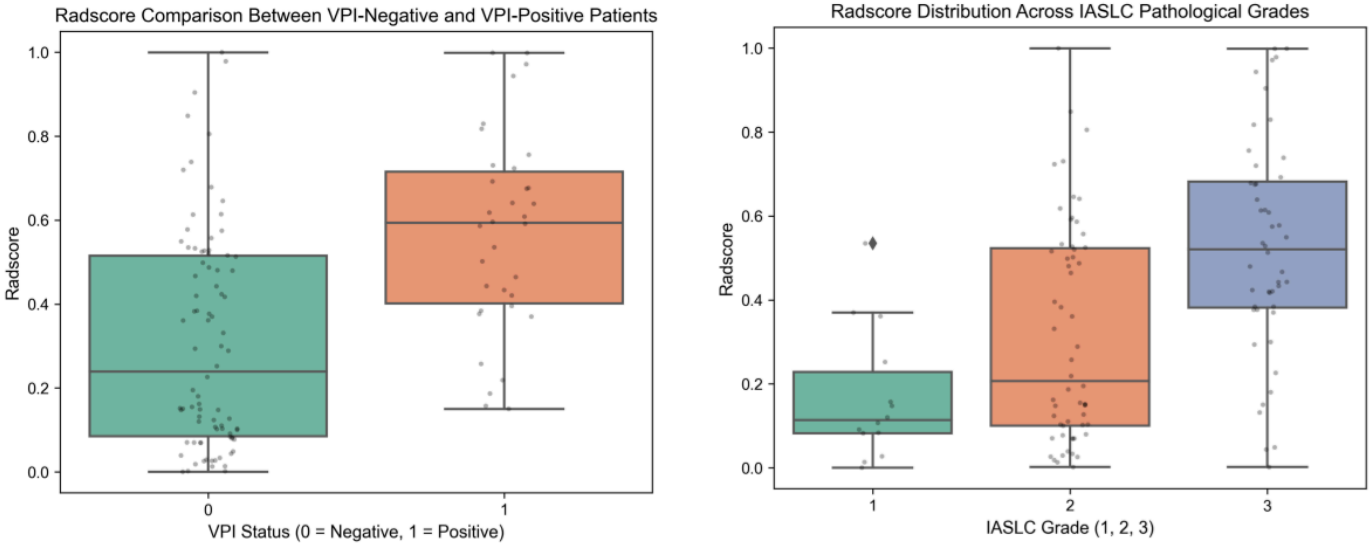


Figure 7. Association between the radiomics score (Radscore) and postoperative histopathological aggressiveness markers.

Figure 7

Association between the radiomics score (Radscore) and postoperative histopathological aggressiveness markers.