© Pol J Radiol 2026; 91: e514-e525
DOI: https://doi.org/10.5114/pjr/219046

Received: 25.01.2026
Accepted: 09.03.2026
Published: 26.10.2026

Polish Journal of Radiology

http://www.polradiol.com

Original paper

A radiomics-clinical nomogram for predicting surgeon-selected resection type in early-stage lung cancer: a single-center retrospective study

Bing Wang[1,A,B,E,F], Zhenghong Wang[1,D,E,F], Zhi Chen[2,3,B,C], Zhengjun Guo[2,A,F,G]

[1]Department of Thoracic Surgery, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China
[2]Department of Cancer Center, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China
[3]Department of Oncology and Hematology, The First People's Hospital of Longquanyi District, Chengdu, China

Abstract

Purpose: Selecting lobectomy versus parenchyma-sparing limited resection for early-stage lung cancer remains challenging because clinicians must balance oncologic adequacy with preservation of pulmonary function. Current guideline-based decisions rely largely on tumor size and clinical judgment, and individualized quantitative assessment is limited. We developed and internally validated a computed tomography (CT) radiomics-clinical nomogram to estimate the probability of lobectomy versus limited resection being selected in routine practice.

Material and methods: We retrospectively enrolled 225 patients with early-stage lung cancer (training cohort, $n = 180$; test cohort, $n = 45$). From preoperative CT, 1,834 radiomic features were extracted, and five features were selected using LASSO with 10-fold cross-validation to construct a radiomics signature. Candidate clinical variables available before surgery were assessed using multivariable logistic regression to derive a clinical signature. The radiomics and clinical signatures were then integrated into a logistic regression nomogram. Model performance was evaluated using the area under the receiver operating characteristic curve (AUC), calibration, decision curve analysis (DCA), and DeLong's test.

Results: In the test cohort, the nomogram achieved an AUC of 0.876, outperforming the radiomics-only model (AUC = 0.809, $p = 0.0008$) and the clinical-only model (AUC = 0.757, $p < 1 \times 10^{-9}$). Calibration showed good agreement between predicted and observed resection selection (Hosmer-Lemeshow $p = 0.72$). DCA indicated that the nomogram yielded a higher net benefit than the alternative models across threshold probabilities of 0-50%.

Conclusions: This radiomics-clinical nomogram provides an individualized estimate of the likelihood of lobectomy versus limited resection being selected in routine practice and may support preoperative planning and multidisciplinary discussion. External and multicenter validation is warranted before clinical implementation.

Key words: early-stage lung cancer, surgery, radiomics, nomogram, model, machine learning.

Introduction

Lung cancer remains the leading cause of cancer-related mortality, with surgical resection remaining pivotal for curative management in early-stage disease [1]. Surgical resection remains the cornerstone of curative treatment, yet the appropriate extent of resection remains debated, particularly when choosing between lobectomy and parenchyma-sparing limited resection such as segmentectomy or wedge resection [2,3]. While lobectomy has historically been the gold standard, recent randomized evidence, including JCOG0802 reveals that: for carefully selected

Correspondence address:
Bing Wang, Department of Thoracic Surgery, the Second Affiliated Hospital of Chongqing Medical University, Chongqing, China,
e-mail: 305501@hospital.cqmu.edu.cn

Authors' contribution:
A Study design · **B** Data collection · **C** Statistical analysis · **D** Data interpretation · **E** Manuscript preparation · **F** Literature search · **G** Funds collection

This is an Open Access journal, all articles are distributed under the terms of the Creative Commons Attribution-Noncommercial-No Derivatives 4.0 International (CC BY-NC-ND 4.0). License (https://creativecommons.org/licenses/by-nc-nd/4.0/).

tumors ≤ 2 cm, segmentectomy may provide non-inferior long-term survival while better preserving postoperative pulmonary function and patient-reported quality of life [4,5]. Consistent with this, limited resection has been associated with improved postoperative functional recovery and quality-of-life outcomes [6]. Nevertheless, a clinically important limitation remains: even when size-based criteria are applied, approximately 14% of patients undergoing sublobar resection develop locoregional recurrence, highlighting the shortcomings of paradigms that rely primarily on tumor diameter and TNM stage for surgical selection [7,8].

Traditional predictive tools, including TNM staging, biomarkers, and ctDNA, provide clinically useful but relatively coarse stratification and may not fully reflect intratumoral heterogeneity or microenvironmental complexity [9,10]. Radiomics, which enables high-throughput extraction of quantitative imaging features, offers a complementary approach by quantifying subvisual patterns of tumor texture, shape, and peritumoral architecture on preoperative computed tomography (CT) [11,12]. Prior studies have suggested that radiomic signatures can be associated with tumor aggressiveness and may aid in identifying occult lymphovascular invasion in lung cancer [11,12]. However, many reported radiomics models have been developed primarily from imaging features and may not fully incorporate patient-specific clinical context, such as hematologic parameters, biochemical profiles, and inflammation-derived indices [13,14].

To address this limitation, we developed and internally validated a nomogram that integrates CT-derived radiomic features, including wavelet-filtered textures and three-dimensional morphological descriptors, with preoperative clinical variables encompassing blood routine results, biochemical profiles, inflammation-derived indices, and tumor markers, together with other preoperative indicators relevant to surgical planning. Using a logistic regression-based nomogram, the model estimates the individualized probability of lobectomy versus limited resection being selected in routine practice, providing a quantitative reference to support multidisciplinary discussion and shared decision-making rather than relying solely on fixed anatomic cutoff points. We further show that the radiomics-clinical nomogram demonstrated better discrimination, calibration, and net clinical benefit compared with radiomics-only and clinical-only models, while providing readily interpretable probability estimates for clinical communication and decision support.

Material and methods

Patient selection

A total of 247 patients with histologically proven early-stage lung cancer who underwent preoperative CT screening between January 2024 and October 2025 in the hospital were included in this retrospective study. Inclusion criteria were as follows: (1) confirmed diagnosis of early-stage lung cancer (stage I or II), (2) availability of preoperative CT scans, (3) undergoing either lobectomy or limited resection, and (4) availability of clinical data. Patients with metastatic disease, incomplete data, or those who underwent other treatment modalities (e.g., chemotherapy or radiotherapy) were excluded from the study. After applying the inclusion and exclusion criteria, a total of 225 patients were enrolled in the study. The baseline clinical characteristics of the patients in our cohort are presented in Table 1. The patient recruitment workflow is presented in Figure 1.

Radiomic feature extraction

Figure 2 depicts the procedural workflow in this study. Preoperative CT scans were acquired using a standardized imaging protocol, and tumor lesions were manually delineated by experienced radiologists to define regions of interest (ROIs) via ITK-SNAP (version 4.2.2, https://www.itksnap.org). From each ROI, we extracted a total of 1,834 handcrafted radiomic features, which we grouped into three categories: geometry features captured the three-dimensional shape characteristics of the tumor, while intensity features described the first-order statistical distribution of voxel intensities within the lesion; texture features, on the other hand, characterized the spatial patterns and higher-order distributions of intensities, offering deeper insights into tumor heterogeneity. To extract texture features, several advanced methods were employed, including the gray-level co-occurrence matrix (GLCM), gray-level run length matrix (GLRLM), gray-level size zone matrix (GLSZM), and neighborhood gray-tone difference matrix (NGTDM). To further enrich our feature set, wavelet filters were applied to generate derived firstorder and texture features from eight wavelet decompositions of the original images. All feature extraction adhered to the Image Biomarker Standardization Initiative (IBSI) guidelines [15]. Prior to modeling, feature distributions were visually inspected (Supplementary Figure S1).

Radiomic feature selection

The Mann-Whitney *U* test was performed for all radiomic features, followed by feature screening. Only features with a p-value < 0.05 were retained for further analysis. To assess correlations among highly repeatable features, Spearman's rank correlation coefficient was calculated. For any two features with a correlation coefficient greater than 0.9, only one was retained. To maximize the distinctiveness of the remaining features, a greedy recursive deletion strategy was applied, wherein the most redundant feature in the current set was iteratively removed. This process ultimately resulted in the selection of 23 features.

Table 1. Baseline characteristics of training and test cohorts

Characteristics	Training ($n = 180$)	Test ($n = 45$)	Total ($N = 225$)	*p*-value
Pathology, *n* (%)				
Adenocarcinoma	161 (89.4)	40 (88.9)	201 (89.3)	0.91
Squamous	19 (10.6)	5 (11.1)	24 (10.7)	
Age (years), mean ± SD	61.88 ± 15.64	63.18 ± 7.99	62.14 ± 14.45	0.44
Sex, *n* (%)				
Male	92 (51.1)	20 (44.4)	112 (49.78)	0.42
Female	88 (48.9)	25 (55.6)	113 (50.2)	
Tumor stage, *n* (%)				
1	136 (75.56)	40 (88.89)	176 (78.22)	0.05
2	35 (19.44)	5 (11.11)	40 (17.78)	
3	9 (5.00)	0	9 (4.00)	
Node, *n* (%)				
Negative	157 (87.22)	44 (97.78)	201 (89.33)	0.10
Positive	23 (12.78)	1 (2.22)	24 (10.67)	
Stage, *n* (%)				
1	144 (80.00)	40 (88.89)	184 (81.78)	0.17
2	28 (15.56)	5 (11.11)	33 (14.67)	
3	8 (4.44)	0	8 (3.56)	
PS score, *n* (%)				
0	99 (55.00)	19 (42.22)	118 (52.44)	0.16
1	81 (45.00)	26 (57.78)	107 (47.56)	
Weight (kg), mean ± SD	61.82 ± 15.30	62.54 ± 9.56	61.96 ± 14.33	0.79
Height (cm), mean ± SD	166.72 ± 28.87	159.94 ± 9.56	165.36 ± 26.30	0.18
Smoking, *n* (%)				
Nerve	115 (63.89)	22 (48.89)	137 (60.89)	0.95
Previous	65 (36.11)	23 (51.11)	88 (39.11)	
Diabetes, *n* (%)				
Yes	15 (8.33)	4 (8.89)	19 (8.44)	0.56
No	165 (91.67)	41 (91.11)	206 (91.56)	
COPD, *n* (%)				
Yes	60 (33.33)	21 (46.67)	81 (36.00)	0.10
No	120 (66.67)	24 (53.33)	144 (64.00)	
Lung cancer side, *n* (%)				
Left	65 (36.11)	14 (31.11)	79 (35.11)	0.95
Right	115 (63.89)	31 (68.89)	146 (64.89)	
Lobe location, *n* (%)				
Upper	92 (51.11)	24 (53.33)	116 (51.56)	0.59
Middle	24 (13.33)	4 (8.89)	28 (12.44)	
Lower	64 (35.56)	17 (37.78)	81 (36.00)	
WBC (mean ± SD)	6.00 ± 2.16	5.79 ± 1.34	5.95 ± 2.02	0.61
N (mean ± SD)	3.87 ± 1.45	3.84 ± 1.19	3.86 ± 1.40	0.92
L (mean ± SD)	1.46 ± 0.48	1.39 ± 0.37	1.45 ± 0.46	0.40
M (mean ± SD)	0.41 ± 0.16	0.38 ± 0.13	0.40 ± 0.15	0.33

Table 1. Cont.

Characteristics	Training ($n = 180$)	Test ($n = 45$)	Total ($N = 225$)	*p*-value
PLT (mean ± SD)	216.30 ± 75.26	205.88 ± 53.51	214.21 ± 71.54	0.45
NLR (mean ± SD)	2.92 ± 1.60	2.93 ± 1.05	2.92 ± 1.50	0.97
LMR (mean ± SD)	4.10 ± 2.05	4.02 ± 1.54	4.08 ± 1.96	0.83
PLR (mean ± SD)	61.32 ± 24.60	57.53 ± 28.95	60.68 ± 16.67	0.44
NPR (mean ± SD)	0.02 ± 0.008	0.02 ± 0.006	0.02 ± 0.008	0.99
SII (mean ± SD)	162.24 ± 77.86	171.97 ± 121.91	167.73 ± 89.66	0.57
Hb (mean ± SD)	140.44 ± 27.97	134.79 ± 13.64	139.30 ± 25.84	0.26
hsCRP (mean ± SD)	2.96 ± 1.80	2.86 ± 1.61	2.94 ± 1.76	0.78
Total protein (mean ± SD)	70.87 ± 14.31	70.93 ± 5.55	70.88 ± 13.04	0.98
Albumin (mean ± SD)	43.36 ± 7.87	43.20 ± 2.69	43.36 ± 7.14	0.91
Globulin (mean ± SD)	29.19 ± 7.50	27.74 ± 4.60	28.90 ± 7.04	0.28
Prealbumin (mean ± SD)	261.62 ± 67.78	259.91 ± 48.02	261.28 ± 64.30	0.89
AST (mean ± SD)	23.69 ± 9.77	21.53 ± 4.60	23.25 ± 9.02	0.21
ALT (mean ± SD)	26.07 ± 14.56	22.71 ± 9.79	25.40 ± 13.80	0.21
LDH (mean ± SD)	192.33 ± 49.05	178.82 ± 24.54	189.61 ± 45.53	0.12
Lactate (mean ± SD)	2.60 ± 0.83	2.43 ± 0.60	2.57 ± 0.80	0.26
PAR (mean ± SD)	3.66 ± 2.09	3.14 ± 1.28	3.45 ± 2.05	0.36
CAR (mean ± SD)	0.07 ± 0.05	0.07 ± 0.04	0.07 ± 0.05	0.64
CLR (mean ± SD)	2.23 ± 1.62	2.22 ± 1.47	2.23 ± 1.59	0.97
CALLY index (mean ± SD)	3.07 ± 1.02	3.37 ± 2.09	3.13 ± 1.34	0.23
CEA (mean ± SD)	3.05 ± 4.66	3.51 ± 6.32	3.14 ± 5.04	0.63
CYFRA21-1 (mean ± SD)	2.90 ± 5.85	1.50 ± 1.29	2.62 ± 5.29	0.17
SF (mean ± SD)	183.85 ± 131.48	169.77 ± 101.42	181.01 ± 126.13	0.56
NSE (mean ± SD)	12.88 ± 3.86	11.90 ± 2.34	12.68 ± 3.63	0.16
SCCA (mean ± SD)	0.89 ± 0.51	0.81 ± 0.47	0.88 ± 0.50	0.40
TPA (mean ± SD)	105.97 ± 129.46	80.50 ± 45.83	100.84 ± 117.96	0.26
ProGRP (mean ± SD)	39.83 ± 46.44	42.23 ± 56.35	40.31 ± 48.60	0.80
Surgical, *n* (%)				
Lobectomy	93 (51.67)	22 (48.89)	115 (51.11)	0.74
Parenchyma-sparing limited resection	87 (48.33)	23 (51.11)	110 (48.89)	
Tumor size (cm), *n* (%)				
< 1	20 (11.11)	0	20 (8.89)	0.22
≥ 1 - < 2	67 (37.22)	18 (40.00)	85 (37.78)	
≥ 2 - < 3	49 (27.22)	14 (31.11)	63 (28.00)	
≥ 3	44 (24.44)	13 (28.89)	57 (25.33)	
Lung peripheral location, *n* (%)				
Peripheral	104 (57.78)	25 (55.56)	129 (57.33)	0.74
Non-peripheral	76 (42.22)	20 (44.44)	96 (42.67)	
Solid component, *n* (%)				
< 25%	43 (23.89)	12 (26.67)	55 (24.44)	0.11
≥ 25% - < 50%	60 (33.33)	13 (28.89)	73 (32.44)	
≥ 50%	77 (42.78)	20 (44.44)	97 (43.11)	

Table 1. Cont.

Characteristics	Training ($n = 180$)	Test ($n = 45$)	Total ($N = 225$)	*p*-value
Surgery time (h), *n* (%)				
1-2	59 (32.78)	16 (35.56)	75 (33.33)	0.72
≥ 2	121 (67.22)	29 (64.44)	150 (66.67)	
Blood loss (ml), *n* (%)				
< 20	43 (23.89)	9 (20.00)	52 (23.11)	0.940
≥ 20 - < 50	59 (32.78)	8 (17.78)	67 (29.78)	
≥ 50	78 (43.33)	28 (62.22)	106 (47.11)	

ALT – alanine aminotransferase, AST – aspartate aminotransferase, CALLY – C-reactive protein-albumin-lymphocyte, CAR – C-reactive protein-to-albumin ratio, CEA – carcinoembryonic antygen, CLR – C-reactive protein-to-lymphocyte ratio, COPD – chronic obstructive pulmonary disease, CYFRA21-1 – cytokeratin 19 fragment, Hb – hemoglobin, hsCRP – high-sensitivity C-reactive protein, L – lymphocyte count, LDH – lactate dehydrogenase, LMR – lymphocyte-to-monocyte ratio, M – monocyte count, N – neutrophil count, NLR – neutrophil-to-lymphocyte ratio, NPR – neutrophil-to-platelet ratio, NSE – neuron-specific enolase, PAR – prealbumin-to-albumin ratio, PLR – platelet-to-lymphocyte ratio, PLT – platelet count, ProGRP – pro-gastrin-releasing peptide, PS – performance status, SCCA – squamous cell carcinoma antygen, SF – serum ferritin, SII – systemic immune-inflammation index, TPA – tissue polypeptide antygen, WBC – white blood cell

For feature selection and signature construction, the least absolute shrinkage and selection operator (LASSO) regression model was applied to the discovery dataset. LASSO uses a regularization parameter (λ) to shrink regression coefficients toward zero, effectively eliminating irrelevant features by setting their coefficients to zero. To determine the optimal λ, a 10-fold cross-validation approach was employed, selecting the value that minimized cross-validation error.

Radiomics signature and clinical signature construction

After LASSO feature selection, the final set of features was used as input for various machine learning models, including Logistic Regression (LR), Support Vector Machine (SVM), RandomForest (RF), ExtraTrees, eXtreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and Multi-Layer Perceptron (MLP), to construct the model predicting the appropriate surgical modality (lobectomy versus limited resection) for lung cancer cases. Features significantly correlated with surgical methods were selected through a five-fold cross-validation framework, followed by model optimization to construct the radiomics signature (Rad_sig).

Clinical parameters included demographic characteristics (age, sex, height, weight, smoking status), comorbidities (e.g., diabetes, chronic obstructive pulmonary disease), tumor-related features (TNM stage, tumor lobe location, solid component ratio), surgical parameters (operative time, blood loss), and a series of laboratory-derived hematological and inflammatory indices (complete blood counts, liver function markers, tumor markers, and derived scores such as NLR, PLR, LMR, SII, CAR, and the CALLY index). These variables were underwent univariate and multivariate logistic regression analyses to identify independent predictors of surgical approach. Then, the significant parameters were integrated to construct the clinical signature (Clinic_sig).

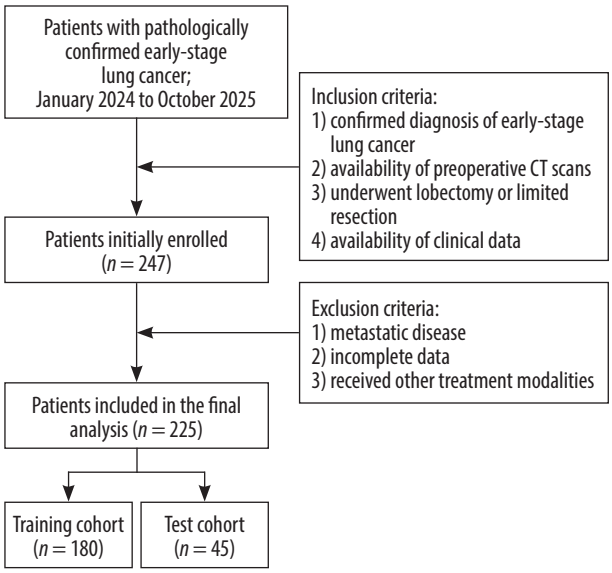


Figure 1. Patient recruitment workflow

Nomogram

A combined nomogram model integrating radiomic (Rad-sig) and clinical (Clinic-sig) signatures was developed through multivariable logistic regression analysis. This unified model integrates quantitative imaging features from Radsig and key preoperative clinical variables from Clinicsig to provide a comprehensive, interpretable tool for guiding surgical approach selection.

Model validation

Model performance was assessed using discrimination, calibration, and clinical utility. Discrimination was evaluated by receiver operating characteristic (ROC) curves and the area under the ROC curve (AUC). Calibration was assessed using calibration plots comparing predicted probabilities with observed event rates across risk

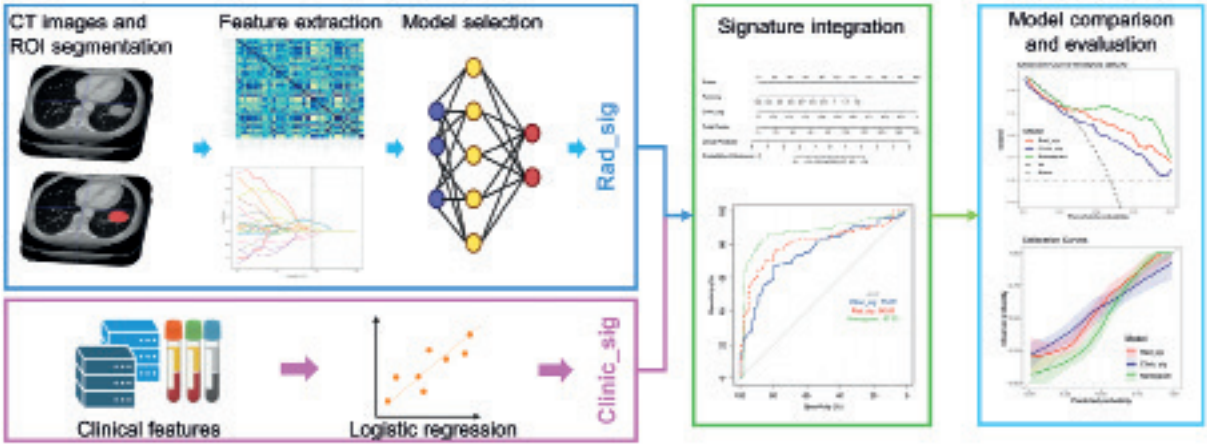


Figure 2. Study workflow

strata. Clinical utility was evaluated using decision curve analysis (DCA) to quantify the net benefit of each model across a range of threshold probabilities. Differences in AUCs between models were compared using DeLong's test. Together, these analyses characterize model performance in estimating the probability of lobectomy versus limited resection selection.

Statistical analysis

We assessed the normality of clinical features using the Shapiro-Wilk test. Continuous variables were evaluated for significance using the *t*-test or the Mann-Whitney *U* test, depending on their distribution. Categorical variables were analyzed with chi-square (χ^2) tests. The baseline characteristics of all cohorts are presented in Table 1. Radiomic feature extraction and data analyses were conducted using Python (version 3.7.12, https://www.python.org). Subsequent statistical analyses, including nomogram analysis, ROC curve analysis, DCA, calibration curve and DeLong's test were performed in R (version 4.3.3, https://www.r-project.org/). Statistical significance in all analyses was defined as a two-tailed *p*-value or adjusted *p*-value < 0.05.

Results

Patient characteristics

The surgical approach was defined as a binary outcome (lobectomy vs parenchyma-sparing limited resection) and served as the prediction target. The study cohort comprised 225 patients (training, *n* = 180; test, *n* = 45). Overall, 201 patients (89.3%) had adenocarcinoma and 24 (10.7%) had squamous cell carcinoma. The mean age at diagnosis was 62.14 ± 14.45 years, and 49.8% of patients were male (*n* = 112). Most tumors were T1 (176/225, 78.2%), whereas T3 tumors were observed only in the training cohort (5.0%, *n* = 9) and not in the test cohort. Positive lymph node metastasis (N+) was present in 10.7% of patients overall (*n* = 24), including 12.8% (*n* = 23) in the training cohort and 2.2% (*n* = 1) in the test cohort; all patients were M0. Pathologic stage distribution was 81.8% stage I (*n* = 184), 14.7% stage II (*n* = 33), and 3.6% stage III (*n* = 8). ECOG performance status was 0 in 52.44% (*n* = 118) and 1 in 47.56% (*n* = 107). The cohort was nearly balanced by surgical procedure, with 51.11% (*n* = 115) undergoing lobectomy and 48.89% (*n* = 110) receiving parenchyma-sparing limited resection. Tumor size was < 1 cm in 8.9% (*n* = 20), 1 – < 2 cm in 37.8% (*n* = 85), 2 – < 3 cm in 28.09% (*n* = 63), and ≥ 3 cm in 25.3% (*n* = 57), and 57.33% (*n* = 129) of tumors were peripherally located.

Routine laboratory indices (including WBC, neutrophil, lymphocyte, monocyte, and platelet counts), liver function tests (AST, ALT, LDH, albumin), and inflammatory indices derived from blood counts (NLR, PLR, LMR, SII, and CALLY index) did not differ meaningfully between the training and test cohorts (all $p > 0.05$). Tumor biomarkers (CEA, SF, SCCA, TPA) were likewise comparable across cohorts. Surgical variables (operative time and estimated blood loss) and the imaging-derived solid-component ratio showed no significant between-cohort differences (all $p > 0.05$). No statistically significant between-cohort differences were observed in height (training cohort, 166.7 ± 28.9 cm; test cohort, 159.9 ± 9.6 cm; $p = 0.18$) or nodal status ($p = 0.10$). Baseline characteristics of the training and test cohorts are summarized in Table 1. Overall, baseline clinical, laboratory, surgical, and imaging characteristics were broadly comparable between cohorts, supporting internal model development and independent evaluation on the held-out test set.

Radiomic feature extraction and model construction

A total of 1834 radiomic features were extracted from each ROI, following the Image Biomarker Standardisation Initiative (IBSI) guidelines [15]. To improve robustness and reduce redundancy, features were filtered prior to modeling using a predefined pipeline, removal of near zero variance features, and correlation-based redundancy

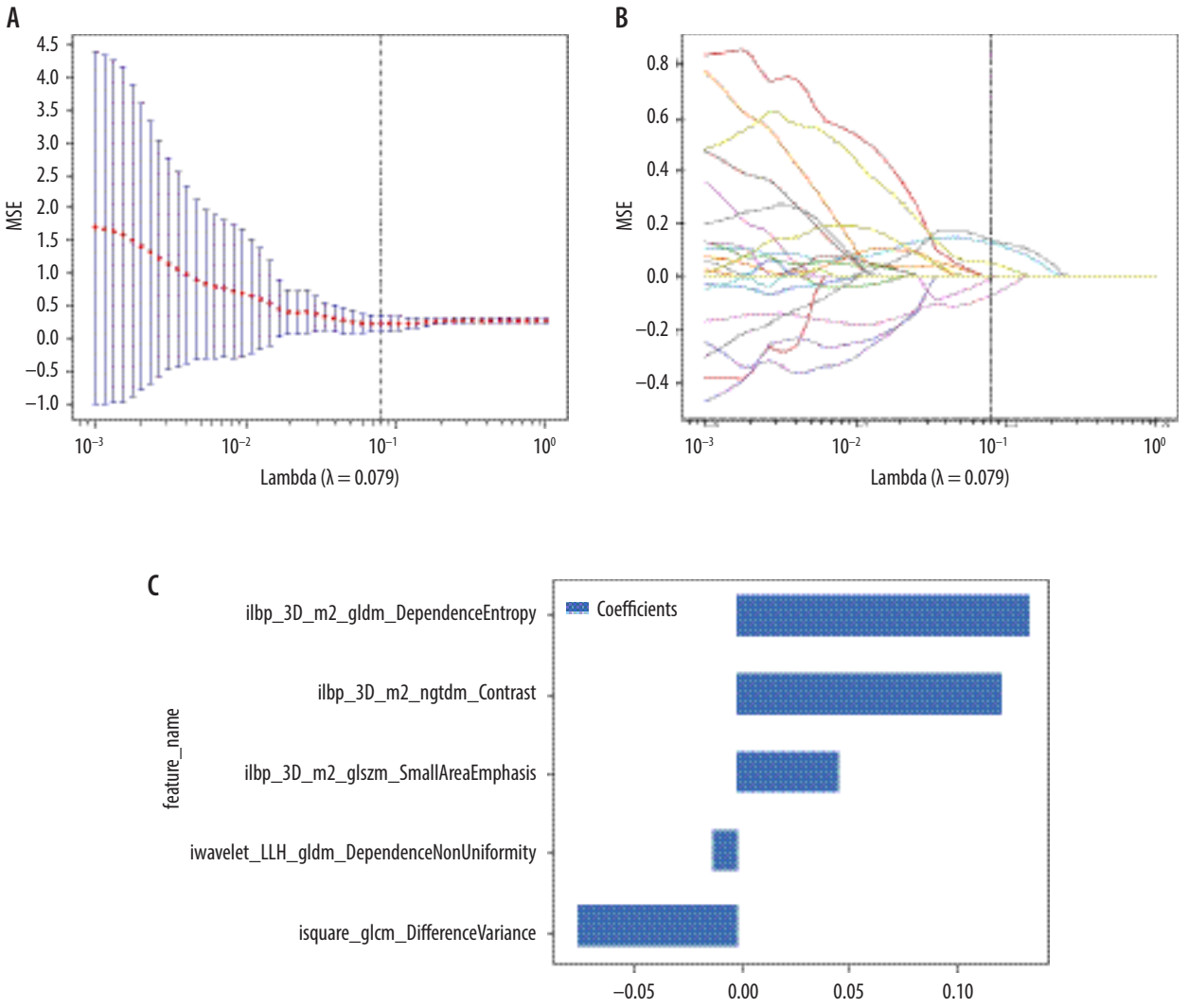


Figure 3. LASSO feature selection for radiomic analysis. A) Cross-validation path plot. B) Coefficient path plot. C) The five most predictive radiomic features identified

Table 2. Univariate multivariable logistic regression analyses of preoperative clinical variables for predicting resection selection

Feature	Univariate analysis		Multivariate analysis	
	OR (95% CI)	*p*-value	OR (95% CI)	*p*-value
T	5.333 (1.691-23.611)	0.010*	6.811 (1.898-33.722)	0.007*
Lobe location	1.440 (1.037-2.019)	0.031*	1.398 (0.972-2.032)	0.073
LMR	1.173 (1.023-1.376)	0.034*	1.212 (1.046-1.444)	0.018*
SCCA	0.456 (0.224-0.901)	0.026*	0.500 (0.226-1.064)	0.077
Blood loss	0.633 (0.429-0.92)	0.019*	0.708 (0.466-1.064)	0.100

*$p < 0.05$.

reduction (Supplementary Figure S2). After filtering, 60 candidate radiomic features were retained for subsequent model development.

We then applied least absolute shrinkage and selection operator (LASSO) logistic regression with 10-fold cross-validation within the training cohort to select features associated with the resection type and to reduce overfitting (Figure 3A). The selected features and their coefficients are shown in Figure 3B. Five radiomic features were ultimately retained to construct the radiomics signature (Figure 3C). Multicollinearity among the selected features was assessed using the variance inflation factor (VIF), and no clinically meaningful collinearity was observed (all VIF < 5).

To assess robustness, we performed an exploratory sensitivity analysis comparing multiple machine-learning classifiers (LR, SVM, RF, ExtraTrees, XGBoost, LightGBM,

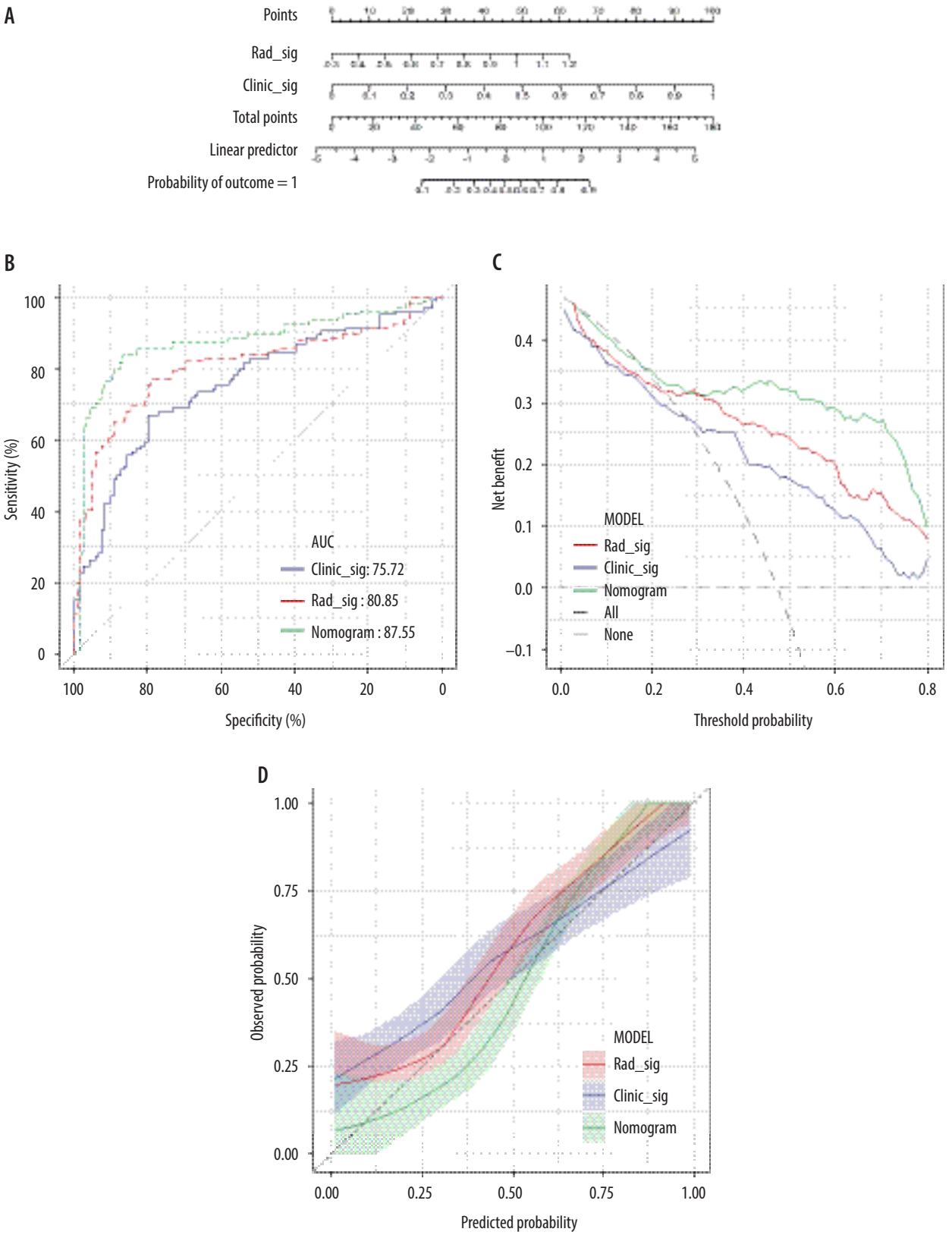


Figure 4. Development and evaluation of a nomogram integrating radiomic and clinical features for predicting surgical approach in lung cancer patients. **A**) Nomogram incorporating selected radiomic features (Rad_sig) and clinical features (Clinic_sig) to predict surgical approach. **B**) ROC curves comparing the predictive performance AUC of the Rad_sig, Clinic_sig, and the combined nomogram. **C**) Calibration curves assessing the agreement between predicted probabilities and actual outcomes for Rad_sig, Clinic_sig, and the nomogram. **D**) Decision curve analysis (DCA) evaluating the clinical utility (net benefit) across threshold probabilities for Rad_sig, Clinic_sig, and the nomogram

Table 3. Diagnostic performance of the models

Model	AUC	95% CI	*p*-value	Sensitivity	Specificity	Youden index
Rad_sig	0.809	0.748-0.869	1.84×10^{-6}	0.833	0.911	0.744
Clinic_sig	0.757	0.693-0.821	7.95×10^{-6}	0.718	0.778	0.496
Nomogram	0.876	0.824-0.927	3.77×10^{-7}	0.840	0.866	0.705

CI – confidence interval

and MLP) using cross-validation in the training cohort. Although tree-based ensemble models showed higher apparent performance during training, generalization to the independent test set was variable (see Supplementary Table S1). Given our primary objective of producing an interpretable clinical tool and to reduce the risk of optimistic estimates in a modestly sized dataset, we therefore constructed the radiomics signature and the integrated nomogram using logistic regression.

The formula for calculating the radiomic signature (Rad-sig) was as follows: label = 0.4761904761904758 + 0.135259 × ilbp_3D_m2_gldm_DependenceEntropy + 0.047721 × ilbp_3D_m2_glszm_SmallAreaEmphasis + 0.123061 × ilbp_3D_m2_ngtdm_Contrast -0.073607 × isquare_glcm_DifferenceVariance – 0.010780 × iwavelet_LLH_gldm_DependenceNonUniformity

Univariable and multivariable analysis of clinical features

Univariable logistic regression was performed for candidate preoperative clinical variables to identify factors associated with resection selection, and odds ratios (ORs) with 95% confidence intervals (CIs) and *p*-values are reported in Table 2 and Supplementary Table S2. Five variables reached statistical significance on univariable testing: T stage (OR = 5.333, $p = 0.010$), lobe location (OR = 1.440, $p = 0.031$), LMR (OR = 1.173, $p = 0.034$), SCCA (OR = 0.456, $p = 0.026$), and blood loss (OR = 0.633, $p = 0.019$).

These variables were then entered into a multivariable logistic regression model. Two features remained independently associated with resection selection: T stage (OR = 6.811, $p = 0.007$) and LMR (OR = 1.212, $p = 0.018$). On the basis of the multivariable model, we derived a clinical signature (Clinic_sig) by combining T stage and LMR into a single logistic-regression score. Clinic_sig subsequently served as the clinical component of the integrated radiomics-clinical nomogram and was used for downstream performance comparison.

Radiomics nomogram construction and model performance comparison

Based on multivariable analysis, T stage and LMR comprised the clinical signature (Clinic_sig, Table 2). The radiomics signature (Rad_sig) was derived from the selected radiomic features, and an integrated nomogram was subsequently developed by combining Clinic_sig and Rad_sig using logistic regression (Figure 4A).

Model discrimination was assessed using ROC curves (Figure 4B). The nomogram achieved the highest AUC (0.876), outperforming both the radiomics model (AUC = 0.809) and the clinical model (AUC = 0.757). Calibration plots demonstrated good agreement between predicted probabilities and observed resection selection for the nomogram, with the calibration curve closely approximating the 45-degree reference line (Figure 4C). By contrast, the Clinic_sig and Rad_sig models showed modest deviations from the reference line, indicating minor miscalibration across parts of the probability range.

Decision curve analysis was performed to compare the net clinical utility of the 3 models across a range of threshold probabilities (Figure 4D). The nomogram yielded higher net benefit than the radiomics only and clinical only models across most thresholds between 0 and 0.5, supporting its potential value as a quantitative aid for preoperative planning and multidisciplinary discussion. Detailed operating characteristics for each model at the prespecified cutoff are provided in Table 3. Pairwise DeLong tests showed that the nomogram significantly outperformed the radiomics model ($p = 0.0008$) and the clinical model ($p < 1 \times 10^{-9}$), and that Rad_sig also outperformed Clinic_sig ($p = 0.0002$; Supplementary Table S3). Collectively, these results indicate that integrating radiomics and clinical information improves discrimination, calibration, and net benefit for estimating the probability of lobectomy versus limited resection selection.

Discussion

In this study, we developed and validated a radiomics-clinical nomogram integrating CT-derived radiomic features and clinical parameters to estimate the probability of lobectomy versus parenchyma-sparing limited resection selection in early-stage lung cancer. The integrated model showed improved discrimination, good calibration, and higher net benefit on decision curve analysis compared with radiomics-only and clinical-only models. These findings are consistent with prior reports indicating that combined radiomics-clinical nomograms can achieve better discrimination and clinical utility than single-modality models [16,17]. In our analysis, the integrated nomogram similarly yielded higher net benefit than alternative strategies across clinically relevant threshold

probabilities, supporting its potential value as a quantitative aid for preoperative planning and multidisciplinary discussion [16,17]. By quantitatively combining CT imaging signatures with patient clinical features, the nomogram provides individualized probability estimates that can facilitate communication and shared decision-making when considering lobectomy versus limited resection. Importantly, we framed the model output as an estimate of resection selection probability rather than a recommendation of the optimal procedure, which aligns the prediction target with the retrospective label definition and supports transparent interpretation.

Selecting an appropriate surgical strategy remains an important determinant of perioperative risk and long-term outcomes in lung cancer [18]. Lobectomy has long been the standard of care for early-stage non-small cell lung cancer, with sublobar resection such as segmentectomy or wedge resection traditionally reserved for small, indolent tumors or patients with limited pulmonary reserve [4]. Recent randomized trials have begun to challenge this paradigm. The JCOG0802/WJOG4607 trial reported that segmentectomy yielded comparable or even superior five-year survival compared with lobectomy in peripheral non-small cell lung cancer measuring 2 cm or smaller [4]. Tumor biology and radiologic phenotype can further influence the selection of surgical extent. Post hoc analyses have suggested heterogeneous effects across subgroups, including potential differences in outcomes for pure-solid tumors [19]. Histologic subtypes such as adenocarcinoma in situ or minimally invasive adenocarcinoma may be appropriate candidates for limited resection, whereas fully invasive tumors often require a wider resection to ensure oncologic adequacy [20,21]. Conversely, spread through air spaces is a recognized risk factor for recurrence, and sublobar resection in spread-through-air-spaces positive tumors has been associated with higher locoregional recurrence than lobectomy [22,23]. Because spread through air spaces is typically confirmed on pathology, recent work has explored whether radiomics might help estimate this risk preoperatively to better inform surgical planning [24]. In routine practice, decisions often rely on stage, location, and pulmonary function, but these criteria may not fully capture patient and tumor heterogeneity. Our nomogram aims to complement existing decision frameworks by integrating imaging-derived heterogeneity with patient-specific clinical features to provide a quantitative reference for multidisciplinary discussion. From an imaging perspective, this approach leverages information embedded in standard-of-care preoperative CT without requiring additional imaging or invasive testing, which may facilitate clinical feasibility if externally validated.

Radiomics transforms medical images into high-dimensional quantitative descriptors of tumor phenotype, including texture, shape, and intensity patterns that extend beyond visual assessment [24]. Radiomics-based models have shown potential for noninvasive phenotyping and prognostic assessment in lung cancer [11,12,24]. However, radiomics-only modeling faces several challenges. Radiomic pipelines can generate hundreds of correlated features, increasing the risk of overfitting and optimistic performance estimates, particularly in studies with limited sample sizes [25]. Variability in CT acquisition, segmentation practices, and feature definitions can further limit reproducibility. In addition, many radiomic features lack direct biological interpretability, which can impede clinical adoption [25]. In contrast, clinical models based on established variables such as TNM stage and inflammatory markers are often more intuitive and easier to implement. For example, a nomogram incorporating TNM stage and systemic inflammatory indices has been reported to predict outcomes in non-small cell lung cancer [26]. We previously developed a scoring system based on clinical features and inflammatory markers to predict recovery from acute radiation enteritis in cervical cancer patients [27]. While such clinical models are generally interpretable and practical, they may not leverage complementary information embedded in imaging [28]. By fusing radiomic and clinical data, our approach seeks to harness the strengths of both modalities. The observed improvement in discrimination and net benefit suggests that the two sources of information may provide complementary value. Moreover, the nomogram format enables transparent interpretation of how individual predictors contribute to the estimated probability of resection selection, which may facilitate clinical communication and interpretation of the model. To strengthen reproducibility for imaging readership, future work should report key IBSI-aligned preprocessing parameters (such as resampling strategy and gray-level discretization) and provide evidence of feature stability, for example through assessment of interobserver reproducibility and robustness to acquisition variability.

Limitations

Several limitations should be acknowledged. First, this was a retrospective single-center study with a modest sample size and an internal hold-out test cohort, which may limit generalizability and result in optimistic performance estimates. Specifically, the model was trained on 180 patients and tested on a held-out set of 45 patients, and the lack of external validation remains a central limitation to generalization beyond our institution. External validation using independent, multicenter, or temporally distinct cohorts is necessary to assess model robustness and generalizability. Second, radiomics performance can be influenced by CT acquisition parameters and segmentation variability. Future work should further standardize imaging preprocessing and, where feasible, report feature stability measures such as interobserver reproducibility or test-retest robustness. Where scanner heterogeneity exists, harmonization strategies and sensitivity analyses should

be considered to evaluate robustness across acquisition settings. Third, our outcome was the resection type selected in routine practice rather than oncologic or functional endpoints. Because the label reflects clinical choice, the model may partially learn institution-specific decision patterns; therefore, any clinical deployment should be contingent on external validation and, ideally, linkage to postoperative outcomes. Prospective studies should evaluate how model-based probability estimates relate to postoperative outcomes such as recurrence, survival, and pulmonary function recovery to clarify potential clinical utility. Finally, additional performance gains may be possible by incorporating complementary data sources, including histopathologic variables, genomic information, or circulating biomarkers, and by exploring more advanced fusion approaches once external validation has been established.

Conclusions

In summary, we developed a radiomics-clinical nomogram that integrates preoperative CT-derived radiomic features with clinical parameters to estimate the probability of lobectomy versus parenchyma-sparing limited resection being selected in early-stage lung cancer. The integrated nomogram demonstrated better discrimination and greater net benefit on decision curve analysis than models based on radiomics or clinical data alone and provides interpretable, individualized probability estimates that may support preoperative planning and multidisciplinary discussion. External validation in independent multicenter cohorts, together with standardized imaging preprocessing and reproducibility assessment, is warranted before broader clinical application.

Disclosures

1. Institutional review board statement: This single-center retrospective analysis was approved by the Ethics Committee of the Second Affiliated Hospital of Chongqing Medical University (Approval No: IRB-KY-2024393).
2. Assistance with the article: None.
3. Financial support and sponsorship: None.
4. Conflicts of interest: None.

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024; 74: 229-263.
2. Spicer JD, Cascone T, Wynes MW, Ahn MJ, Dacic S, Felip E, et al. Neoadjuvant and adjuvant treatments for early stage resectable NSCLC: consensus recommendations from the International Association for the Study of Lung Cancer. J Thorac Oncol 2024; 19: 1373-1414.
3. Bertolaccini L, Tralongo AC, Del Re M, Facchinetti F, Ferrara R, Franchina T, et al. Segmentectomy vs. lobectomy in stage IA non-small cell lung cancer: a systematic review and meta-analysis of perioperative and survival outcomes. Lung Cancer 2024; 197: 107990. DOI: 10.1016/j.lungcan.2024.107990.
4. Saji H, Okada M, Tsuboi M, Nakajima R, Suzuki K, Aokage K, et al. Segmentectomy versus lobectomy in small-sized peripheral non-small-cell lung cancer (JCOG0802/WJOG4607L): a multicentre, open-label, phase 3, randomised, controlled, non-inferiority trial. Lancet 2022; 399: 1607-1617.
5. Lei X, Zhou N, Zhang H, Li T, Ren F, Zhang B, et al. Lobe-specific analysis of sublobar lung resection for NSCLC patients with tumors ≤ 2 cm. Cancers (Basel) 2022; 14: 3265. DOI: 10.3390/cancers14133265.
6. Lim E, Harris RA, McKeon HE, Batchelor TJ, Dunning J, Shackcloth M, et al. Impact of video-assisted thoracoscopic lobectomy versus open lobectomy for lung cancer on recovery assessed using self-reported physical function: VIOLET [Internet]. Health Technol Assess 2022; 26: 1-162.
7. Altorki N, Wang X, Damman B, Mentlick J, Landreneau R, Wigle D, et al. Lobectomy, segmentectomy, or wedge resection for peripheral clinical T1aN0 non-small cell lung cancer: a post hoc analysis of CALGB 140503 (Alliance). J Thorac Cardiovasc Surg 2024; 167: 338-347e1. DOI: 10.1016/j.jtcvs.2023.07.008.
8. Tasoudis P, Loufopoulos G, Manaki V, Doerr M, Agala CB, Long JM, et al. Long term outcomes after lobar versus sublobar resection for patients with non-small cell lung cancer: systematic review and individual patient data meta-analysis. Lung Cancer 2024; 195: 107929. DOI: 10.1016/j.lungcan.2024.107929.
9. Shen H, Jin Y, Zhao H, et al. Potential clinical utility of liquid biopsy in early-stage non-small cell lung cancer. BMC Med 2022; 20: 480. DOI: 10.1186/s12916-022-02681-x.
10. Davies MPA, Sato T, Ashoor H, Hou L, Liloglou T, Yang R, et al. Plasma protein biomarkers for early prediction of lung cancer. EBioMedicine 2023; 93: 104686. DOI: 10.1016/j.ebiom.2023.104686.
11. Avanzo M, Stancanello J, Pirrone G, Sartor G. Radiomics and deep learning in lung cancer. Strahlenther Onkol 2020; 196: 879-887.
12. Chen M, Copley SJ, Viola P, Lu H, Aboagye EO. Radiomics and artificial intelligence for precision medicine in lung cancer treatment. Semin Cancer Biol 2023; 93: 97-113.
13. Gillies RJ, Schabath MB. Radiomics improves cancer screening and early detection. Cancer Epidemiol Biomarkers Prev 2020; 29: 2556-2567.
14. Wu YJ, Wu FZ, Yang SC, Tang EK, Liang CH. Radiomics in early lung cancer diagnosis: from diagnosis to clinical decision support and education. Diagnostics (Basel) 2022; 12: 1064. DOI: 10.3390/diagnostics12051064.
15. Whybra P, Zwanenburg A, Andrearczyk V, Schaer R, Apte AP, Ayotte A, et al. The image biomarker standardization initiative: standardized convolutional filters for reproducible radiomics and enhanced clinical insights. Radiology 2024; 310: e231319. DOI: 10.1148/radiol.231319.

16. Liu C, Zhao W, Xie J, Lin H, Hu X, Li C, et al. Development and validation of a radiomics-based nomogram for predicting a major pathological response to neoadjuvant immunochemotherapy for patients with potentially resectable non-small cell lung cancer. Front Immunol 2023; 14: 1115291. DOI: 10.3389/fimmu.2023.1115291.
17. Peng L, Zhang X, Liu J, Liu Y, Huang J, Chen J, et al. MRI-radiomics-clinical-based nomogram for prenatal prediction of the placenta accreta spectrum disorders. Eur Radiol 2022; 32: 7532-7543.
18. Miller M, Hanna N. Advances in systemic therapy for non-small cell lung cancer. BMJ 2021; 375: n2363. DOI: 10.1136/bmj.n2363.
19. Hattori A, Suzuki K, Takamochi K, Wakabayashi M, Sekino Y, Tsutani Y, et al. Segmentectomy versus lobectomy in small-sized peripheral non-small-cell lung cancer with radiologically pure-solid appearance in Japan (JCOG0802/WJOG4607L): a post-hoc supplemental analysis of a multicentre, open-label, phase 3 trial. Lancet Respir Med 2024; 12: 105-116.
20. Lee JH, Park CM, Kim H, Hwang EJ, Park J, Goo JM. Persistent part-solid nodules with solid part of 5 mm or smaller: Can the 'follow-up and surgical resection after interval growth' policy have a negative effect on patient prognosis? Eur Radiol 2017; 27: 195-202.
21. Zhang Y, Fu F, Chen H. Management of ground-glass opacities in the lung cancer spectrum. Ann Thorac Surg 2020; 110: 1796-1804.
22. Eguchi T, Kameda K, Lu S, Bott MJ, Tan KS, Montecalvo J, 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: 87-98.
23. Ren Y, Xie H, Dai C, She Y, Su H, Xie D, et al. Prognostic impact of tumor spread through air spaces in sublobar resection for 1A lung adenocarcinoma patients. Ann Surg Oncol 2019; 26: 1901-1908.
24. Bassi M, Russomando A, Vannucci J, Ciardiello A, Dolciami M, Ricci P, et al. Role of radiomics in predicting lung cancer spread through air spaces in a heterogeneous dataset. Transl Lung Cancer Res 2022; 11: 560-571.
25. Demircioglu A. Reproducibility and interpretability in radiomics: a critical assessment. Diagn Interv Radiol 2025; 31: 321-328.
26. Xu W, Liu X, Yan C, Abdurahmane G, Lazibiek J, Zhang Y, et al. The prognostic value and model construction of inflammatory markers for patients with non-small cell lung cancer. Sci Rep 2024; 14: 7568. DOI: 10.1038/s41598-024-57814-4.
27. Zeng C, Ji J, Huang Y, Peng Y, Zhang X, Yang Z, et al. A novel scoring system to predict acute radiation enteritis recovery in cervical cancer patients undergoing concurrent chemoradiotherapy: a Southwest China cohort study. Int J Gen Med 2024; 17: 5907-5919.
28. Ji Z, Cui Y, Peng Z, Gong J, Zhu HT, Zhang X, et al. Use of radiomics to predict response to immunotherapy of malignant tumors of the digestive system. Med Sci Monit 2020; 26: e924671. DOI: 10.12659/MSM.924671.