Thoracic Cancer

WILEY

ORIGINAL ARTICLE OPEN ACCESS

Image-Based Deep Learning Model for Predicting Lymph Node Metastasis in Lung Adenocarcinoma With CT ≤ 2 cm

Shang Liu[1] | Zhen Gao[1] | Li-feng Shi[1] | Han Xiao[1] | Su Li[2] | Meng Li[1] | Zhong-min Peng[1]

[1]Department of Thoracic Surgery, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Shandong First Medical University, Jinan, People's Republic of China | [2]Department of Pathology, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Shandong First Medical University, Jinan, People's Republic of China

Correspondence: Meng Li (hussar101@163.com) | Zhong-min Peng (pengzhongmin@sdfmu.edu.cn)

Received: 23 January 2025 | **Revised:** 7 March 2025 | **Accepted:** 10 March 2025

Funding: This work was supported by Major Scientific and Technological Innovation Project of Shandong Province, 2019JZZY021002.

Keywords: lung adenocarcinoma | lymph node metastasis (LNM) | prediction model | proportion of solid components (PSC) | surgery

ABSTRACT

Background: Lymph node metastasis (LNM) poses a considerable threat to survival in lung adenocarcinoma. Currently, minor resection is the recommended surgical approach for small-diameter lung cancer. The accurate preoperative identification of LNM in patients with small-diameter lung cancer is important for improving patient survival and outcomes.

Methods: A total of 1740 patients with clinical early-stage lung adenocarcinoma who underwent surgical resection were enrolled in this study. The Lasso model was used to screen clinical and imaging features, and multivariate logistic regression analysis was used to analyze the relevant diagnostic factors to establish a diagnostic model for predicting LNM. Receiver operating characteristic (ROC) curve analysis, decision curve analysis (DCA) and calibration curve analysis were used to verify the clinical efficacy of the model, which was further validated with an internal validation set.

Results: The proportion of solid components (PSC), sphericity, nodule margin, entropy, and edge blur were identified as diagnostic factors that were strongly correlated with LNM in lung adenocarcinoma patients. The area under the ROC curve (AUC) in the internal training set was 0.91. Decision curve analysis revealed that the model could achieve greater benefits for patients. The calibration curve was used to further verify the applicability of the prediction model.

Conclusions: Patients with early-stage lung adenocarcinoma with LNM can be identified by typical imaging features. The diagnostic model can help to optimize surgical planning among thoracic surgeons.

1 | Introduction

Lung cancer is the malignancy with the highest global prevalence and is the most common cause of cancer-related death [1]. Adenocarcinomas account for more than 40% of lung cancer cases and are the most common histological subtype of non-small cell lung cancer (NSCLC) [2]. In recent years, with the popularization of low-dose CT scans, the detection rate of early lung adenocarcinoma has greatly improved, but CT has a high false-positive rate in the diagnosis of early-stage lung cancer [3]. The overall 5-year survival rate of patients with lung adenocarcinoma is as low as 19.4% [4]. At present, surgical resection is still the best treatment for early-stage lung cancer [5]. Even in patients with stage I NSCLC who undergo complete tumor resection, the rate of postoperative recurrence or metastasis can reach 21.7% [6]. When lymph node metastasis (LNM) occurs in pulmonary adenocarcinoma, it usually indicates a poor prognosis. The JCOG0802 study [7] showed that for lung nodules

Meng Li and Zhong-min Peng are contributed equally to this work.

with a diameter ≤2 cm and a ground-glass opacity component greater than 50%, sublobar resection is not inferior to lobectomy. However, there is no conclusive evidence regarding the optimal surgical approach for lymph node upstaging. Notably, many studies have shown that LNM is not uncommon in lung adenocarcinomas with a diameter ≤2 cm. Even if the tumor is small, it can spread to the lymphatic system; thus, even if lung adenocarcinoma is detected at an early stage, the possibility of LNM cannot be ignored. Therefore, timely and in-depth evaluation, diagnosis, and treatment are very important for the treatment of lung adenocarcinoma. Early measures to control the metastasis of lung adenocarcinoma can improve patient survival and disease prognosis [3, 6, 7]. Many previous studies have shown that systematic lymph node dissection is highly important for improving the outcomes of patients with pathological LNM, and the preoperative prediction of LNM in lung adenocarcinoma patients is highly important [8–10].

Due to the deepening understanding of lymph node metastasis in pulmonary adenocarcinoma, there is an increasing amount of research focusing on lymph node metastasis in this type of cancer. For example, the whole-body PET-CT SUVmax, histological subtypes (such as papillary, micropapillary, solid, and acinar), imaging findings, and serum tumor markers (such as CEA and SCCA) can be used to predict lymph node metastasis [11, 12]. These research protocols provide some approaches for detecting LNM, but the whole-body PET–CT examination can cause more radiation and side effects on patients. Serum oncology indicators are still affected by certain external factors, and intraoperative frozen pathology cannot accurately indicate histological subtypes. These examinations are not achievable in many hospitals at present, so it is difficult to combine the model with clinical practice. In contrast, high-resolution CT and three-dimensional reconstruction are available in hospitals. Therefore, there is an urgent need for a suitable model for predicting LNM in lung adenocarcinoma patients, which can provide a new scientific basis for surgeons to make surgical plans through simple, true, and reliable preoperative examinations and corresponding indicators.

At present, CT is still the main imaging approach for detecting LNM, and it is generally believed that a short-axis diameter of CT-detected lymph nodes ≥1 cm is considered to indicate metastasis [12]. Relying solely on the above threshold may not lead to the accurate identification of LNM, as some small-diameter lung cancer patients without imaging evidence of LNM have shown confirmed positive LNM via postoperative pathology, and even a small number of patients underwent only sublobar resection and lymph node sampling. This condition may increase the risk of recurrence or secondary surgery. Our study aimed to establish a predictive model for LNM in pulmonary adenocarcinoma patients by integrating clinical and imaging features to estimate the likelihood of LNM. We hope this model will help to predict the risk of LNM, thereby enabling clinicians to select the optimal surgical approach, improving patient survival rates and disease-free survival periods, and ultimately improving the disease prognosis. Thus, the surgical method can be determined more accurately and the operation time can be shortened. This predictive model includes both clinical and imaging features and has important application significance in the diagnosis and treatment of lung adenocarcinoma.

2 | Methods

2.1 | Patients

This retrospective study was reviewed and approved by the ethics committee of our center. The medical data and specimens analyzed herein were anonymized. This study included all patients who underwent surgical treatment for primary pulmonary adenocarcinoma at our center between March 2019 and March 2022. The exclusion criteria were as follows: 1. has a history of malignancy in other sites; 2. patients received tumor treatment before surgery; 3. the chest CT data were incomplete, including chest CT scans completed at an outside hospital and imaging data lost at our hospital; 4. preoperative three-dimensional reconstruction could not be performed for various reasons. Finally, a total of 1740 patients were included in the study.

2.2 | Assessment of Histology

All specimens were processed according to routine clinical procedures and pathological data. Two experienced pathologists specializing in thoracic tumor pathology performed histological diagnoses of the lung specimens without knowledge of the imaging information. The pathological analysis of each specimen was performed according to the WHO2021 lung adenocarcinoma classification standard, and TNM was divided into the 8th edition of the Union for International Cancer Control (UICC) lung cancer TNM classification standard issued in 2017. Finally, patients were divided into a LNM-positive group and a LNM-negative group according to whether there was LNM based on postoperative pathology.

2.3 | CT Acquisition and Evaluation

The subjects were scanned with a Brilliance 64 CT (Philips Healthcare, Best, The Netherlands) with the following parameters: section width 0.625 mm, reconstruction interval 0.625 mm, pitch 0.984, 120 kV, and 250 mA. All images were obtained with a high-resolution 52.8-cm grayscale monitor with 2048 × 1560 pixels and a standard lung window (window width, 1500 HU; horizontal window, −700 HU) and mediastinal window (width window, 350 HU; window level, 50 HU). Finally, the relevant data were statistically analyzed. The images were analyzed and evaluated by two experienced radiologists who were blinded to the pathological subtypes.

The size, consolidation tumor ratio (CTR), location (central or peripheral), margin (smooth, lobulated, and spiculated), shape (round-like (round to ovoid), or irregular), vacuole sign, calcification, air bronchogram sign, satellite lesion, pleural retraction, vascular passage, and bronchial honeycomb cavity sign of the isolated lesions were evaluated [13]. The consolidation tumor ratio (CTR) is calculated as the ratio of the diameter of the solid component to the maximum diameter of the nodule, which includes both ground glass opacity and solid components [14]. Peripherally located tumors were defined as those completely located in the outer one-third of the lung parenchyma. Tumors were considered nonperipheral if any part of the tumor extended into the inner two-thirds of the lung. The air bronchogram sign refers to

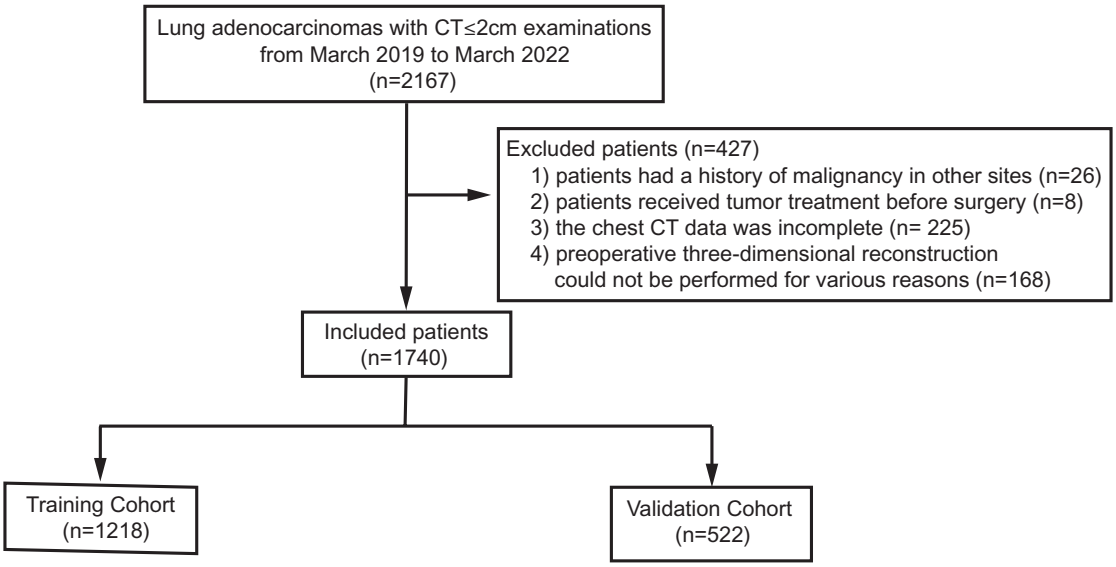

FIGURE 1 | The flow chart shows the selection and exclusion criteria for patients.

oval or circular areas of low attenuation within pulmonary nodules, masses, or consolidation areas. Satellite lesions are defined as smaller nodules located within 2 cm of the primary tumor. The honeycomb cavity sign refers to the appearance of small cavities of varying sizes and irregular shapes in the nodular solid area after the passage of deformed and distorted bronchi. The cavities represent dilated and distorted bronchi or focal emphysema rather than cavitation.

2.4 | 3D Image Construction and Evaluation

The three-dimensional spatial features of pulmonary nodules were constructed based on the three-dimensional reconstruction system of pulmonary nodules developed by deep learning [15, 16]. After obtaining the chest image data of the patient, three-dimensional image construction and deep learning of the target area were performed by the system. Finally, we calculated various three-dimensional spatial characteristics of pulmonary nodules, including surface area, volume, mass, proportion of solid components (the ratio of solid component volume to total nodule volume, and CT value > -300 HU was used to define the solid component of the nodule [17]), maximum CT value, minimum CT value, average CT value, median CT value, the ratio of the circular circumference with the same area in the tumor area to the circumference of the tumor, and the degree of chaos within the lung nodule image.

2.5 | Statistical Analysis

The chi-square test or Fisher's exact test was used for categorical variables, and Student's *t* test or the Mann–Whitney U test was used for continuous variables. Lasso machine learning was performed on the clinical and imaging features of the patients to screen and compare the highly relevant diagnostic factors of LNM. Lasso machine learning was performed with 10-fold cross validation using one λ value of the standard error at the minimum distance mean square error. Multivariate logistic regression was performed on highly relevant diagnostic factors to construct a logistic regression model and a nomogram. A receiver operating characteristic curve was constructed to evaluate the efficacy of the logistic regression model. The Hosmer–Lemeshow goodness test was used to evaluate the difference between the actual value and the predicted value of the logistic regression model, and a calibration curve was drawn. Decision curve analysis was conducted to quantitatively verify the net benefit of the logistic model under different threshold probabilities [18]. A restricted cubic spline (RCS) was used to examine the nonlinear relationship between quantitative data and intergroup changes. The RCS model with the minimum Akaike information criterion (AIC) was selected. The logistic regression model formula was as follows:

$$
\begin{aligned}
P = {} & 1/1+\exp(1.081+6.787\times PSC \\
& -4.248\times \text{Sphericity}-1.347\times \text{entropy} \\
& +(0.325\times \text{Labulation})\text{or} \\
& (0.866\times \text{Lobulation}+\text{Spiculation})\text{or} \\
& (1.486\times \text{Spiculation})+1.081\times \text{Edge blur})
\end{aligned}
$$

SPSS version 25 and R4.2.2 software were used for statistical analysis (SPSS version 25; SPSS Statistics, IBM, Chicago, Ill; R version 4.2.2, Foundation for Statistical Computing, Vienna, Austria).

3 | Results

3.1 | Demographic and Clinicopathological Findings

From March 2019 to March 2022, a total of 2167 patients with primary lung adenocarcinoma underwent surgical treatment at our hospital. Among them, 427 patients who did not meet the inclusion criteria were excluded from the study cohort (Figure 1 included specific details on the exclusion criteria). Table 1 summarizes the demographic and clinicopathological characteristics of the final study population. There were no significant differences in age, sex, smoking status, CT max, pleural invasion, STAS, vascular invasion, or the CT location of nodules between the internal train cohort ($n=1218$) and the internal test cohort ($n=522$). The

TABLE 1 | Patient demographics and baseline characteristics.

Characteristic	Cohort		p^2
	Training cohort, $N=1218^1$	Internal test cohort, $N=522^1$	
Age			0.781
Median (IQR)	58 (52, 64)	58 (52, 64)	
Sex			0.130
Female	730 (59.9%)	333 (63.8%)	
Male	488 (40.1%)	189 (36.2%)	
Smoking			0.475
0	909 (74.6%)	398 (76.2%)	
1	309 (25.4%)	124 (23.8%)	
CT. max			0.470
Median (IQR)	1.60 (1.30, 1.90)	1.60 (1.30, 1.90)	
Location			0.747
R-U	396 (32.5%)	162 (31.0%)	
R-M	89 (7.3%)	44 (8.4%)	
R-L	221 (18.1%)	91 (17.4%)	
L-U	309 (25.4%)	144 (27.6%)	
L-L	203 (16.7%)	81 (15.5%)	
Pleural. invasion			0.413
0	1158 (95.1%)	501 (96.0%)	
1	60 (4.9%)	21 (4.0%)	
STAS			0.085
0	1157 (95.0%)	485 (92.9%)	
1	61 (5.0%)	37 (7.1%)	
Vascular. invasion			0.215
0	1203 (98.8%)	519 (99.4%)	
1	15 (1.2%)	3 (0.6%)	

[1] *n* (%).
[2] Wilcoxon rank sum test; Pearson's Chi-squared test.

distributions of the two data sets on the main variables were similar, and the *p* value showed no significant difference, indicating good comparability between the data sets.

3.2 | Screening for Typical Characteristics of Lymph Node-Positive Adenocarcinoma

Table 2 summarizes the demographic and imaging characteristics of the patients in the internal training set ($n=1218$) and internal validation set ($n=522$). Lasso machine learning simplified 25 predictor factors into 6 potential diagnostic factors (Figure 2a,b), including the PSC, sphericity, entropy, nodule margin and nodule edge blur. The correlation analyses revealed significant correlations between the PSC, CT maximum value, CT minimum value, and CT mean value (Figure 2c). A comparison of the area under the ROC curve of the abovementioned potential 6 diagnostic factors revealed that the AUC of the PSC was 0.879 (Figure 2). The Japanese Clinical Oncology Group JCOG0201 has previously used the solid component/tumor ratio as an important measure for distinguishing between noninvasive nodules and nodules with a good prognosis [14, 19]. Figure S1 shows the column diagram of CTR, and it is clear that the predictive power of CTR is lower than that of PSC. Considering the clinical practice and convenience of the model, we included the PSC, sphericity, entropy, nodule margin, and nodule edge blur as diagnostic features for positive LNM (Figure 3).

3.3 | Construction and Validation of a Nomogram for Predicting Lymph Node Metastasis in Adenocarcinoma

The results of univariate Logistic regression were shown in Table 3. Then, the PSC, sphericity, entropy, nodule margin, and nodule edge blur were included in the multivariate logistic regression, and the results are summarized in Table 4. A nomogram based on a logistic model was constructed (Figure 4), with PSC being a strong predictor of LNM. Based on the logistic regression model, the area under the ROC curve (AUC) in the internal training set was 0.91 (Figure 5a). The decision curve showed that patients with the abovementioned predictors yielded greater benefits (Figure 5b,c). The calibration curve showed that the model exhibited good discrimination ability and prediction accuracy on both the internal training set and the validation set, which indicated good model fit (Figure 3e,f).

3.4 | Clinical Value and Application of PSC

The nomogram showed that the PSC was strongly correlated with LNM (Figure 4). The model shows a significant relationship between PSC and LNM (p value <0.001). Before the inflection point (PSC <0.75), the odds ratio increases gradually. After the inflection point (PSC ≥ 0.75), the odds ratio increases more steeply (Figure 5d). The RCS plot effectively demonstrates a significant and nonlinear relationship between PSC and the odds of LNM in lung adenocarcinoma. The inflection point at 0.75 is crucial, indicating that beyond this value, the risk of LNM increases sharply.

4 | Discussion

Currently, the surgical methods for treating early-stage lung cancer are evolving. Since the 1990s, video-assisted thoracoscopic surgery (VATS) has gradually replaced traditional thoracotomy and become the mainstream surgical method [20–22]. The innovation of surgical techniques has led to greater benefits for patients, including reductions in postoperative pain, postoperative complications, and the length of hospital stays [23–26]. Standard lobectomy combined with systematic lymph node dissection is the recommended surgical resection approach for early-stage lung cancer [27]. The JCOG0802 study showed that

TABLE 2 | Patient imaging characteristics.

Characteristics	Training cohort			Internal test cohort		
	A, $N=1128$[1]	B, $N=90$[1]	p[2]	A, $N=480$[1]	B, $N=42$[1]	p[3]
CTR			<0.001			<0.001
Mean ± SD	0.55 ± 0.30	0.81 ± 0.17		0.55 ± 0.31	0.83 ± 0.19	
Median (IQR)	0.55 (0.30, 0.81)	0.85 (0.71, 0.94)		0.55 (0.28, 0.84)	0.86 (0.77, 0.95)	
Range	0.00, 1.00	0.13, 1.00		0.00, 1.00	0.08, 1.00	
PSC			<0.001			<0.001
Mean ± SD	0.51 ± 0.30	0.91 ± 0.16		0.52 ± 0.30	0.88 ± 0.22	
Median (IQR)	0.52 (0.25, 0.77)	0.98 (0.89, 1.00)		0.51 (0.25, 0.78)	0.98 (0.84, 1.00)	
Range	0.00, 1.00	0.20, 1.00		0.00, 1.00	0.08, 1.00	
Superficial. area			<0.001			0.065
Mean ± SD	1014 ± 642	1290 ± 756		988 ± 535	1172 ± 610	
Median (IQR)	957 (648, 1269)	1143 (759, 1540)		937 (602, 1267)	1003 (732, 1434)	
Range	87, 14 002	349, 4802		135, 4543	409, 2863	
Volume			0.001			0.114
Mean ± SD	1686 ± 1204	2238 ± 1727		1682 ± 1265	2086 ± 1586	
Median (IQR)	1522 (822, 2261)	1738 (1062, 2651)		1446 (760, 2296)	1708 (1008, 2698)	
Range	35, 12 211	411, 11 726		71, 9645	384, 7576	
Quality			<0.001			0.002
Mean ± SD	1336 ± 2144	2340 ± 1872		1300 ± 1145	1963 ± 1636	
Median (IQR)	1023 (589, 1640)	1880 (1077, 2711)		954 (512, 1831)	1489 (1011, 2517)	
Range	28, 60 143	411, 12 201		46, 11 158	19, 7928	
CT. value. max			0.061			0.546
Mean ± SD	269 ± 263	293 ± 176		260 ± 218	281 ± 202	
Median (IQR)	228 (127, 354)	253 (170, 371)		232 (120, 341)	257 (104, 380)	
Range	−473, 2993	55, 824		−344, 1565	20, 715	
CT. value. min			<0.001			<0.001
Mean ± SD	−609 ± 159	−345 ± 169		−604 ± 156	−351 ± 182	
Median (IQR)	−637 (−727, −521)	−325 (−450, −223)		−635 (−721, −520)	−325 (−449, −244)	
Range	−978, 3	−764, −46		−881, −48	−780, −4	
CT. mean			<0.001			<0.001
Mean ± SD	−286 ± 190	−47 ± 100		−280 ± 181	−62 ± 139	
Median (IQR)	−279 (−429, −134)	−25 (−73, 14)		−287 (−424, −129)	−31 (−88, 14)	
Range	−727, 324	−506, 90		−729, 142	−530, 127	
CT. median			<0.001			<0.001
Mean ± SD	−290 ± 219	−25 ± 108		−285 ± 211	−43 ± 148	
Median (IQR)	−284 (−473, −90)	7 (−38, 27)		−295 (−460, −75)	5 (−33, 26)	
Range	−746, 135	−591, 87		−759, 92	−579, 102	
Compactness			<0.001			0.015
Mean ± SD	0.034 ± 0.010	0.032 ± 0.005		0.036 ± 0.023	0.032 ± 0.005	

(Continues)

TABLE 2 | (Continued)

Characteristics	Training cohort			Internal test cohort		
	A, $N=1128$[1]	B, $N=90$[1]	p[2]	A, $N=480$[1]	B, $N=42$[1]	p[3]
Median (IQR)	0.035 (0.031, 0.038)	0.033 (0.029, 0.035)		0.035 (0.031, 0.038)	0.033 (0.029, 0.036)	
Range	0.012, 0.328	0.021, 0.043		0.017, 0.373	0.021, 0.041	
Sphericity			< 0.001			0.019
Mean ± SD	0.74 ± 0.07	0.72 ± 0.07		0.74 ± 0.07	0.72 ± 0.08	
Median (IQR)	0.75 (0.70, 0.80)	0.72 (0.68, 0.76)		0.75 (0.70, 0.79)	0.73 (0.65, 0.77)	
Range	0.37, 0.88	0.53, 0.87		0.47, 0.88	0.53, 0.85	
Entropy			< 0.001			< 0.001
Mean ± SD	5.10 ± 0.35	4.80 ± 0.40		5.09 ± 0.33	4.82 ± 0.36	
Median (IQR)	5.14 (4.89, 5.35)	4.80 (4.54, 5.09)		5.13 (4.90, 5.33)	4.83 (4.61, 5.08)	
Range	3.68, 6.64	3.69, 5.65		3.86, 5.93	3.98, 5.53	
Margin			< 0.001			0.003
Smooth	258 (23%)	8 (9%)		125 (26%)	3 (7%)	
Lobulation	314 (28%)	15 (17%)		123 (26%)	7 (17%)	
Spiculation	64 (6%)	9 (10%)		30 (6%)	3 (7%)	
Lobulation + Spiculation	492 (44%)	58 (64%)		202 (42%)	29 (69%)	
Shape			0.101			0.066
Circular	664 (59%)	45 (50%)		287 (60%)	19 (45%)	
Irregularity	464 (41%)	45 (50%)		193 (40%)	23 (55%)	
Pseudocavit			0.048			0.141
0	892 (79%)	79 (88%)		350 (73%)	35 (83%)	
1	236 (21%)	11 (12%)		130 (27%)	7 (17%)	
Calcification			0.193			> 0.999
0	1119 (99%)	88 (98%)		478 (100%)	42 (100%)	
1	9 (1%)	2 (2%)		2 (0%)	0 (0%)	
Edge blur			< 0.001			< 0.001
0	986 (87%)	65 (72%)		428 (89%)	28 (67%)	
1	142 (13%)	25 (28%)		52 (11%)	14 (33%)	
Air. bronchogram			0.022			0.599
0	573 (51%)	57 (63%)		254 (53%)	24 (57%)	
1	555 (49%)	33 (37%)		226 (47%)	18 (43%)	
Satellite. lesions			0.460			0.533
0	1121 (99%)	89 (99%)		472 (98%)	41 (98%)	
1	7 (1%)	1 (1%)		8 (2%)	1 (2%)	
Pleural. retraction			0.034			0.138
0	414 (37%)	23 (26%)		181 (38%)	11 (26%)	
1	714 (63%)	67 (74%)		299 (62%)	31 (74%)	
Blood. vessel. across			0.108			0.692
0	215 (19%)	11 (12%)		104 (22%)	8 (19%)	

(Continues)

TABLE 2 | (Continued)

Characteristics	Training cohort			Internal test cohort		
	A, $N=1128$[1]	B, $N=90$[1]	p[2]	A, $N=480$[1]	B, $N=42$[1]	p[3]
1	913 (81%)	79 (88%)		376 (78%)	34 (81%)	
Honeycomb			0.541			>0.999
0	1018 (90%)	83 (92%)		429 (89%)	38 (90%)	
1	110 (10%)	7 (8%)		51 (11%)	4 (10%)	
Outer.1.3			0.004			0.192
0	324 (29%)	39 (43%)		137 (29%)	16 (38%)	
1	804 (71%)	51 (57%)		343 (71%)	26 (62%)	

[1]n (%).
[2]Wilcoxon rank sum test; Pearson's Chi-squared test; Fisher's exact test.
[3]Wilcoxon rank sum test; Fisher's exact test; Pearson's Chi-squared test.

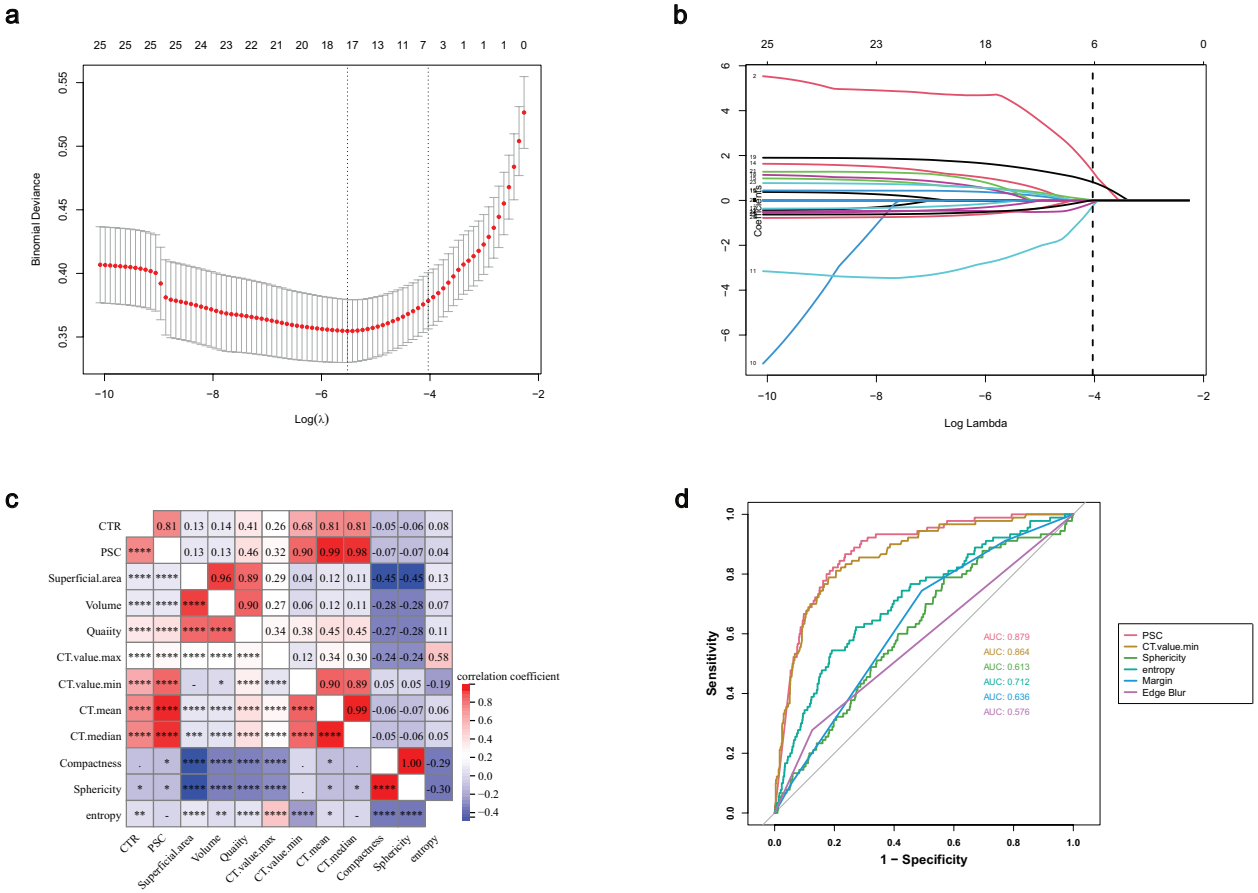

FIGURE 2 | Signature selection (a) Decision tree machine learning. (b) Cross-validation plot of the Lasso regression, with 10-fold cross-validation using the minimum criterion, and the resulting variable using the λ value with the minimum mean square error. (c) Correlation analysis between each diagnostic factor. (d) Area under the receiver operating characteristic curve of each diagnostic factor.

for small NSCLC (diameter < 2 cm, CTR > 0.5), segmentectomy is not inferior to lobectomy, but segmentectomy is associated with a higher risk of local recurrence. For patients with lymph node upstaging, more aggressive surgical approaches may be necessary. A retrospective study revealed that lobectomy remains the best surgical approach for treating lung adenocarcinoma with LNM [8]. Therefore, effective identification of LNM before operation and selection of reasonable surgical methods are particularly important for lung cancer, especially for NSCLC patients with LNM.

Our study revealed that the prevalence of LNM among patients with small adenocarcinoma was relatively low, approximately 6.6%. The JCOG0802 study also indicated a metastasis rate of 6.1% in this population, which is similar to our findings. The restricted cubic spline (RCS) plot illustrates a notable association between the primary tumor's prognostic scoring component (PSC) and the probability of LNM in lung adenocarcinoma. The critical inflection point is identified at 0.75, beyond which there is a marked increase in the risk of LNM. This insight is instrumental for clinicians in evaluating the likelihood of lymph node

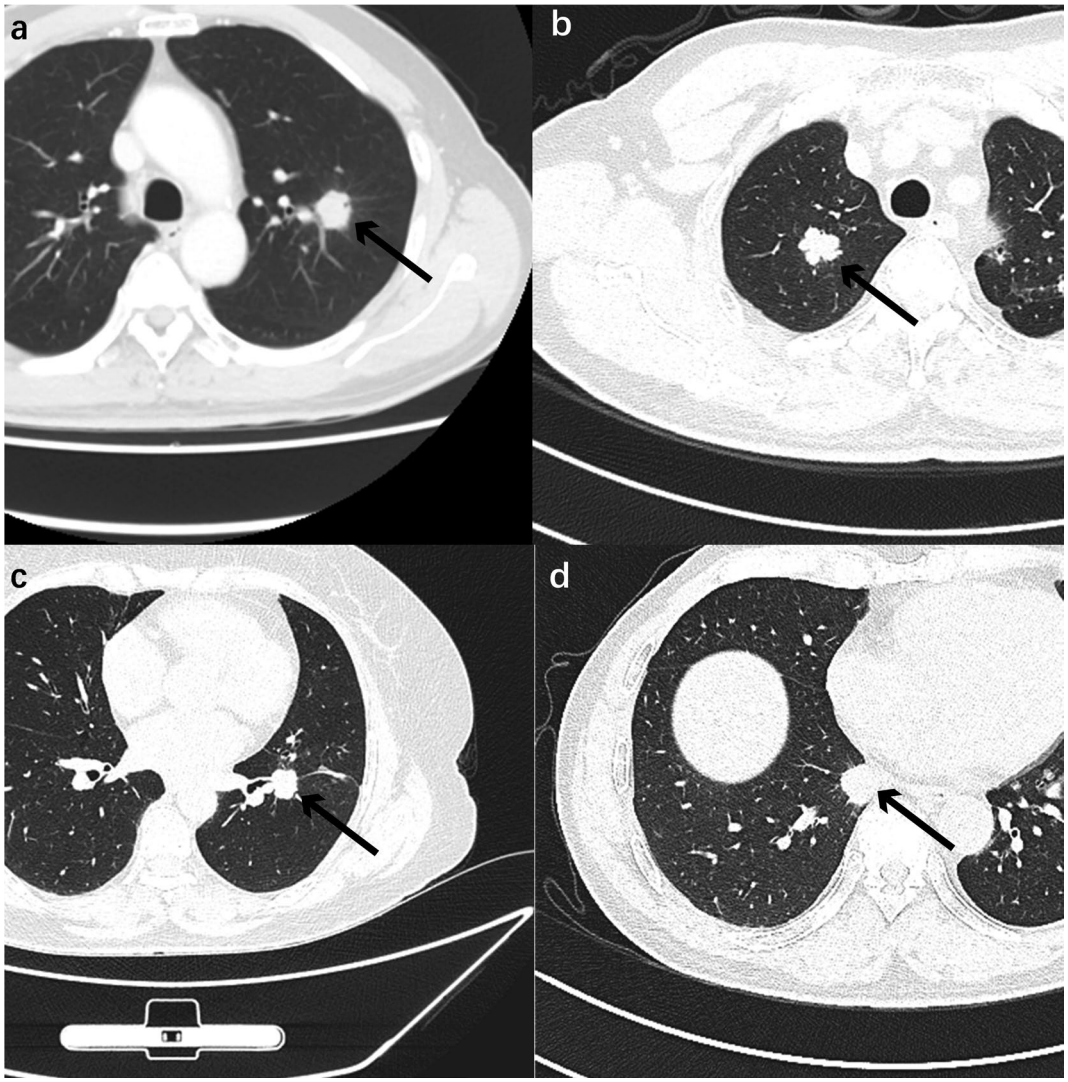

FIGURE 3 | (a) Tumor with extremely high PSC; (b) tumor with marked lobulation; (c) tumor with lobulation and spiculation; (d) tumor with marked spiculation.

TABLE 3 | Results of univariate ogistic regression.

Characteristic	*N*	Event *N*	OR	95% CI	*p*
PSC	1218	90	2135.90	422.51, 14043.97	< 0.001
Sphericity	1218	90	0.01	0.00, 0.18	0.001
Entropy	1218	90	0.15	0.09, 0.25	< 0.001
Margin					
Smooth	266	8	—	—	
Lobulation	329	15	1.54	0.66, 3.88	0.332
Spiculation	73	9	4.54	1.67, 12.53	0.003
Lobulation + Spiculation	550	58	3.80	1.89, 8.73	< 0.001
Edge blur					
0	1051	65	—	—	
1	167	25	2.67	1.61, 4.33	< 0.001

Abbreviations: CI = confidence interval; OR = odds ratio.

TABLE 4 | Results of multivariate logistic regression for training cohort.

Characteristic	*N*	Event *N*	OR	95% CI	*p*
PSC	1218	90	886.19	158.07, 7188.83	<0.001
Sphericity	1218	90	0.01	0.00, 0.37	0.010
Entropy	1218	90	0.26	0.12, 0.55	<0.001
Margin					
Smooth	266	8	—	—	
Lobulation	329	15	1.38	0.52, 3.92	0.523
Spiculation	73	9	4.42	1.35, 14.81	0.014
Lobulation + Spiculation	550	58	2.38	1.04, 6.05	0.051
Edge blur					
0	1051	65	—	—	
1	167	25	6.90	3.48, 13.86	<0.001

Abbreviations: CI = confidence interval; OR = odds ratio.

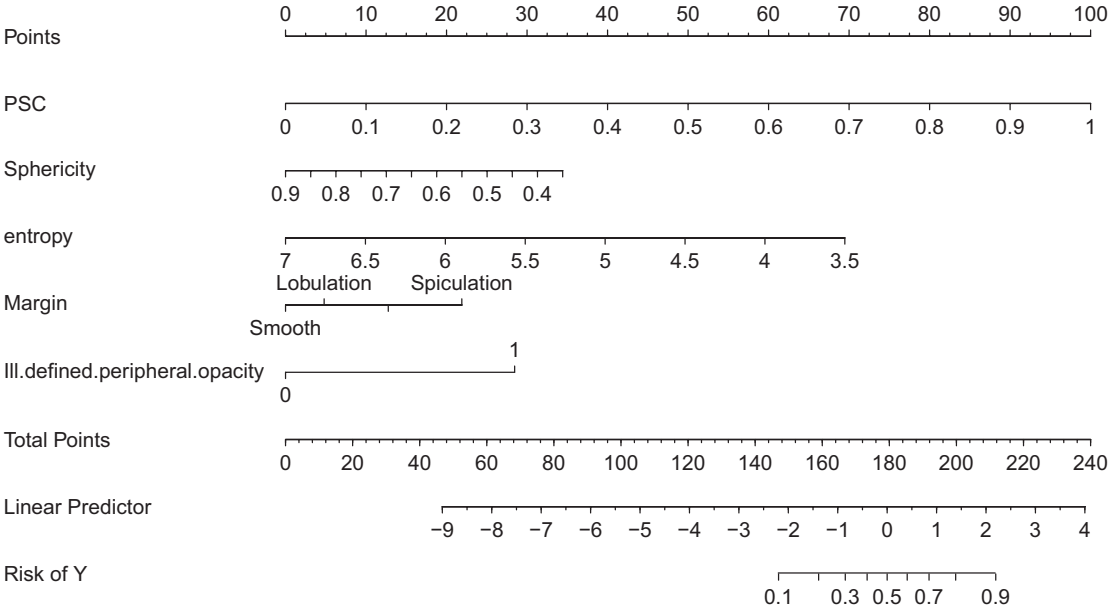

FIGURE 4 | Construction and validation of diagnosis nomogram.

involvement using this prediction model, thus facilitating more informed clinical decisions for patient care.

In recent years, research on LNM in lung adenocarcinoma has been increasing. The use of EBUS-TBNA has provided a diagnostic basis for N2 lymph node metastasis [28]. However, as an invasive examination, it is impractical to perform this examination in every patient, and N1 LNM cannot be assessed. Keiju Aokage et al. used tumor diameter and the CTR to predict lymph node metastasis, but our study used the PSC to predict LNM in lung adenocarcinoma patients. The PSC is based on the total sum of the solid component ratios in different two-dimensional layers in a three-dimensional context. Theoretically, this approach is more accurate, as indicated by our findings. The SUVmax of the primary lung tumor is an important factor for the prognosis of lung adenocarcinoma, but the whole body PET-examination may cause more radiation and side effects on patients and increase costs for patients. In addition, compared with whole-body PET-CT, chest CT and three-dimensional reconstruction are easier to obtain, and the radiation hazard to the patient is relatively small. In our opinion, we can use this model to predict the likelihood of LNM and then decide whether a patient needs to undergo PET-CT, which would benefit most patients. In our study, a clinical prediction model based on CT imaging features (2D and 3D) of lung adenocarcinoma was established, which was easier to implement and achieved satisfactory results. By combining clinical features and imaging features, the feasibility and effectiveness of this method provide a more convenient and accurate method for the diagnosis and treatment of lung adenocarcinoma patients. Many studies have also shown that three-dimensional features of nodules are more accurate than two-dimensional features

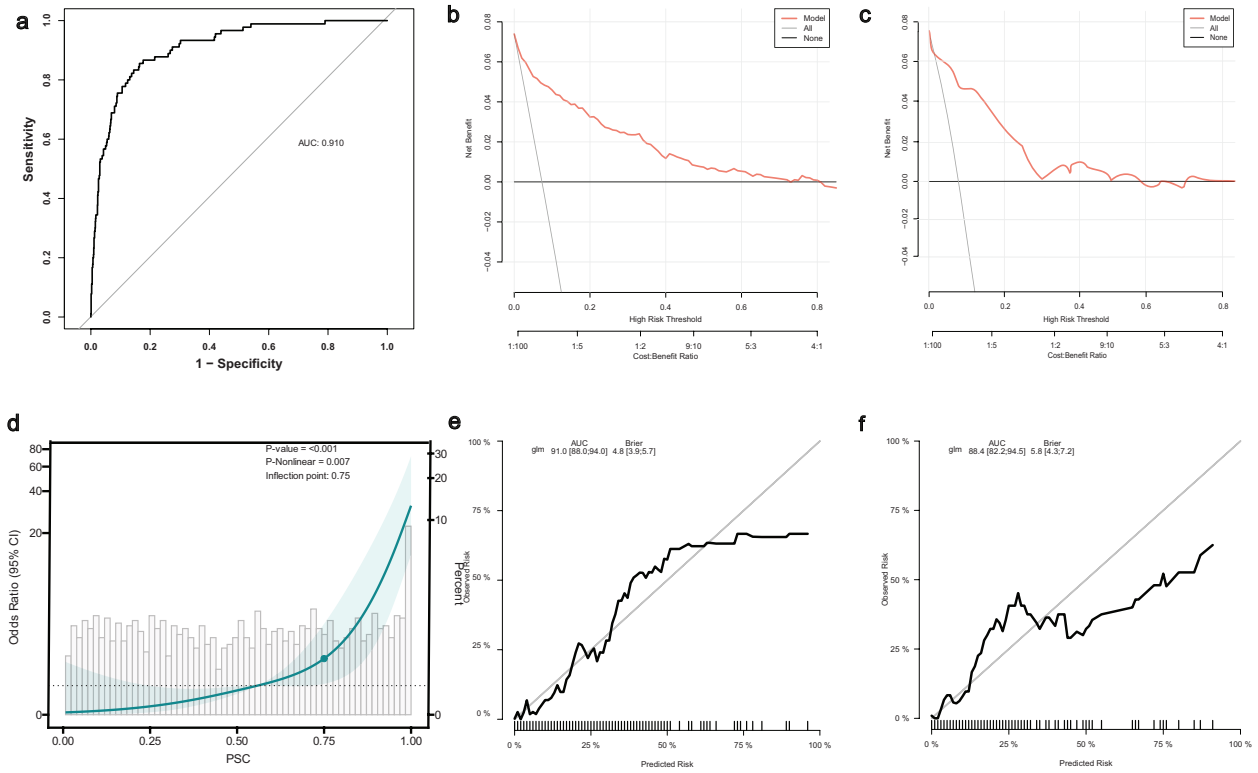

FIGURE 5 | (a) ROC curve of the diagnostic model in the training group. (b) Training the queue model of decision-making curve analysis. (c) Validating the queue model of decision-making curve analysis. (d) RCS was used to analyze the nonlinear relationship between prediction probability and LNM. The *y*-axis represents the OR value of any predicted probability value to the current group compared to an individual with a predicted probability of 75%. (e) Calibration curve for the training of the nomogram queue. (f) Calibration curve for the validating of the nomogram queue.

[29, 30]. In the process of tumor growth, it is often difficult to achieve regular growth, so the morphology and characteristics of tumors may be relatively complex and variable. The overall morphology and structure of the tumor can be revealed more comprehensively by analyzing the 3D features of the tumor, which are derived from the segmentation and reconstruction of each 2D level. Therefore, tumor analysis and diagnostic methods based on three-dimensional features are more accurate and comprehensive. Sphericity is an important morphological parameter commonly used to assess the shape characteristics and growth tendencies of tumors. The evaluation of sphericity can provide a reference for patients with LNM of lung adenocarcinoma. The tumor margin refers to the boundary or limit between the nodule and surrounding tissue. The marginal features of nodules have certain value in predicting LNM in small-diameter lung adenocarcinoma.

Based on our prediction model, the possibility of LNM can be determined preoperatively, and then the treatment or surgical plan can be changed. Intraoperative frozen section (FS) pathology is considered an important tool for selecting the optimal surgical strategy [31]. Fortunately, its accuracy in distinguishing between invasive and preinvasive lesions is relatively high [32, 33]. Therefore, if surgical resection is planned, based on our predictive model and contingent upon the intraoperative frozen pathology confirmation of lung adenocarcinoma, surgical resection, specifically lobectomy combined with a systematic mediastinal lymphadenectomy, can be undertaken in patients diagnosed with high-risk adenocarcinoma exhibiting potential nodal metastases. Many studies have also shown that lobectomy can prolong the OS of patients; however, it may increase the risk of postoperative complications [34–36]. The model showed good discrimination and generalization ability in both the training set and the internal test set, which provided a new reference for the early diagnosis of lung adenocarcinoma. Therefore, for patients who do not experience LNM and are confirmed by intraoperative frozen section pathology, selective lymph node dissection or simple lymph node sampling can be an effective treatment strategy. This approach can help reduce intraoperative bleeding and surgical complications and prevent patient overtreatment. This prediction model can be used to evaluate the risk of LNM in patients more accurately and provides a scientific basis for the selection of treatment options. Therefore, this personalized treatment strategy can benefit the vast majority of patients.

This study aimed to establish an imaging predictive model for predicting LNM in patients with lung adenocarcinoma (≤ 2 cm in diameter) before surgery based on clinical characteristics and imaging features (in two-dimensional and three-dimensional structures) and to provide guidance for the selection of surgical methods and treatment. The PSC, sphericity, nodule margin, and nodule location were included in the multivariate logistic regression. Based on the logistic regression model, the AUC was 0.91 in the internal training set, and the prediction accuracy of our model varied according to the risk cut-off used to define a positive result. The inflection point at 0.75 suggests a critical value of PSC where the odds of lymph node metastasis change significantly. Based on the model prediction results, combined with intraoperative frozen section (FS) pathology, doctors can develop a more accurate

surgical plan. If the model predicts high-risk LNM and intraoperative frozen pathology confirms lung adenocarcinoma, the physician may consider more aggressive surgery, such as lobectomy combined with systemic mediastinal lymph node dissection. For patients with low risk of LNM predicted by the model and no LNM confirmed by intraoperative frozen pathology, selective lymph node dissection or simple lymph node sampling may be a more appropriate surgical modality to reduce intraoperative bleeding and surgical complications and avoid overtreatment.

At present, as a predictive model for preoperative diagnosis, our study has certain limitations. First, we lacked survival rates and prognostic data related to patients who developed LNM and those who did not. This will be further improved in our future research. Second, as a single-center retrospective study, there may be some bias in the data collected. We aim to conduct multicenter studies in the future to obtain more convincing findings, thereby contributing to the health of patients and the development of public health.

5 | Conclusions

In conclusion, we developed an imaging diagnostic model for LNM in lung adenocarcinoma patients using three-dimensional and two-dimensional imaging features of lung nodules. The model can accurately assess preoperatively whether patients will develop LNM, thereby providing guidance for preoperative treatment or surgical approaches. This can help clinical physicians develop more precise and beneficial plans for patients, thus reducing the risk of surgical complications and improving patient outcomes, ultimately benefiting patients.

Author Contributions

Shang Liu: conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing – original draft, writing – review and editing. **Zhen Gao:** data curation, formal analysis, investigation, project administration, resources; writing – review and editing. **Li-feng Shi:** conceptualization, investigation, methodology, resources, software, supervision. **Han Xiao:** data curation, investigation, validation, visualization. **Su Li:** formal analysis, investigation, supervision, validation. **Meng Li:** conceptualization, investigation, methodology, project administration, resources, supervision, validation, writing – review and editing. **Zhong-min Peng:** conceptualization, data curation, formal analysis, funding acquisition, project administration, resources, supervision, writing – review and editing.

Acknowledgments

We would like to express our gratitude to Dr. Feng-ying Du for method guidance and professional statistical advice.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The original data of this study are available upon reasonable request from the corresponding author.

References

1. B. C. Bade and C. S. Dela Cruz, "Lung Cancer 2020 Epidemiology, Etiology, and Prevention," *Clinics in Chest Medicine* 41, no. 1 (2020): 1–24.

2. H. Sung, J. Ferlay, R. L. Siegel, et al., "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA: A Cancer Journal for Clinicians* 71, no. 3 (May 2021): 209–249.

3. F. C. Detterbeck, P. J. Mazzone, D. P. Naidich, and P. B. Bach, "Screening for Lung Cancer: Diagnosis and Management of Lung Cancer, 3rd Ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines," *Chest* 143, no. 5 Suppl (May 2013): e78S–e92S, https://doi.org/10.1378/chest.12-2350.

4. A. N. Giaquinto, K. D. Miller, K. Y. Tossas, R. A. Winn, A. Jemal, and R. L. Siegel, "Cancer Statistics for African American/Black People 2022," *CA: A Cancer Journal for Clinicians* 72, no. 3 (May 2022): 202–229.

5. W. G. Cahan, "Radical Lobectomy," *Journal of Thoracic and Cardiovascular Surgery* 39 (1960): 555–572.

6. C. Wang, Y. Wu, J. Shao, D. Liu, and W. Li, "Clinicopathological Variables Influencing Overall Survival, Recurrence and Post-Recurrence Survival in Resected Stage I Non-Small-Cell Lung Cancer," *BMC Cancer* 20, no. 1 (February 2020): 150.

7. H. Saji, M. Okada, M. Tsuboi, 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* 399, no. 10335 (April 2022): 1607–1617.

8. T. Fujimoto, S. D. Cassivi, P. Yang, et al., "Completely Resected N1 Non-Small Cell Lung Cancer: Factors Affecting Recurrence and Long-Term Survival," *Journal of Thoracic and Cardiovascular Surgery* 132, no. 3 (September 2006): 499–506.

9. A. Demir, A. Turna, C. Kocaturk, et al., "Prognostic Significance of Surgical-Pathologic N1 Lymph Node Involvement in Non-Small Cell Lung Cancer," *Annals of Thoracic Surgery* 87, no. 4 (2009): 1014–1022, https://doi.org/10.1016/j.athoracsur.2008.12.053.

10. F. Gester, A. Paulus, A. L. Sibille, B. Duysinx, and R. Louis, "Prognostic Factors in non-Small Cell Lung Cancer," *Revue Médicale de Liège* 71, no. 1 (January 2016): 34–39.

11. B. Ye, M. Cheng, W. Li, et al., "Predictive Factors for Lymph Node Metastasis in Clinical Stage IA Lung Adenocarcinoma," *Annals of Thoracic Surgery* 98, no. 1 (July 2014): 217–223.

12. J. J. Hung, Y. C. Yeh, Y. C. Wu, T. Y. Chou, and W. H. Hsu, "Prognostic Factors in Completely Resected Node-Negative Lung Adenocarcinoma of 3 Cm or Smaller," *Journal of Thoracic Oncology* 12, no. 12 (December 2017): 1824–1833, https://doi.org/10.1016/j.jtho.2017.07.009.

13. R. Perez-Johnston, J. A. Araujo-Filho, J. G. Connolly, et al., "CT-Based Radiogenomic Analysis of Clinical Stage I Lung Adenocarcinoma With Histopathologic Features and Oncologic Outcomes," *Radiology* 303, no. 3 (June 2022): 664–672, https://doi.org/10.1148/radiol.211582.

14. H. Asamura, T. Hishida, K. Suzuki, et al., "Radiographically Determined Noninvasive Adenocarcinoma of the Lung: Survival Outcomes of Japan Clinical Oncology Group 0201," *Journal of Thoracic and Cardiovascular Surgery* 146, no. 1 (July 2013): 24–30.

15. K. Li, K. Liu, Y. Zhong, et al., "Assessing the Predictive Accuracy of Lung Cancer, Metastases, and Benign Lesions Using an Artificial Intelligence-Driven Computer Aided Diagnosis System," *Quantitative Imaging in Medicine and Surgery* 11, no. 8 (August 2021): 3629–3642.

16. P. Yu, H. Zhang, H. Kang, W. Tang, C. W. Arnold, and R. Zhang, "RPLHR-CT Dataset and Transformer Baseline for Volumetric Super-Resolution From CT Scans," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Cham: Springer, 2022).

17. E. T. Scholten, C. Jacobs, B. van Ginneken, et al., "Detection and Quantification of the Solid Component in Pulmonary Subsolid Nodules by Semiautomatic Segmentation," *European Radiology* 25, no. 2 (February 2015): 488–496.

18. A. J. Vickers, A. M. Cronin, E. B. Elkin, and M. Gonen, "Extensions to Decision Curve Analysis, a Novel Method for Evaluating Diagnostic Tests, Prediction Models and Molecular Markers," *BMC Medical Informatics and Decision Making* 8 (November 2008): 53.

19. K. Suzuki, T. Koike, T. Asakawa, et al., "A Prospective Radiological Study of Thin-Section Computed Tomography to Predict Pathological Noninvasiveness in Peripheral Clinical IA Lung Cancer (Japan Clinical Oncology Group 0201)," *Journal of Thoracic Oncology* 6, no. 4 (2011): 751–756, https://doi.org/10.1097/JTO.0b013e31821038ab.

20. R. J. Landreneau, S. R. Hazelrigg, P. F. Ferson, et al., "Thoracoscopic Resection of 85 Pulmonary Lesions," *Annals of Thoracic Surgery* 54, no. 3 (September 1992): 415–419.

21. G. Roviaro, C. Rebuffat, F. Varoli, C. Vergani, C. Mariani, and M. Maciocco, "Videoendoscopic Pulmonary Lobectomy for Cancer," *Surgical Laparoscopy & Endoscopy* 2, no. 3 (1992): 244–247.

22. J. A. Howington, M. G. Blum, A. C. Chang, A. A. Balekian, and S. C. Murthy, "Treatment of Stage I and II Non-Small Cell Lung Cancer: Diagnosis and Management of Lung Cancer, 3rd Ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines," *Chest* 143, no. 5 Suppl (May 2013): e278S–e313S.

23. B. J. Champagne, H. T. Papaconstantinou, S. S. Parmar, et al., "Single-Incision Versus Standard Multiport Laparoscopic Colectomy: A Multicenter, Case-Controlled Comparison," *Annals of Surgery* 255, no. 1 (January 2012): 66–69, https://doi.org/10.1097/SLA.0b013e3182378442.

24. A. Chow, S. Purkayastha, and P. Paraskeva, "Appendicectomy and Cholecystectomy Using Single-Incision Laparoscopic Surgery (SILS): The First UK Experience," *Surgical Innovation* 16, no. 3 (2009): 211–217.

25. J. D. Raman, A. Bagrodia, and J. A. Cadeddu, "Single-Incision, Umbilical Laparoscopic Versus Conventional Laparoscopic Nephrectomy: A Comparison of Perioperative Outcomes and Short-Term Measures of Convalescence," *European Urology* 55, no. 5 (2009): 1198–1204.

26. K. M. Reavis, M. W. Hinojosa, B. R. Smith, and N. T. Nguyen, "Single-Laparoscopic Incision Transabdominal Surgery Sleeve Gastrectomy," *Obesity Surgery* 18, no. 11 (2008): 1492–1494.

27. R. J. Ginsberg and L. V. Rubinstein, "Randomized Trial of Lobectomy Versus Limited Resection for T1 N0 Non-Small Cell Lung Cancer. Lung Cancer Study Group," *Annals of Thoracic Surgery* 60, no. 3 (September 1995): 615–622.

28. M. R. Vial, O. J. O'Connell, H. B. Grosu, et al., "Diagnostic Performance of Endobronchial Ultrasound-Guided Mediastinal Lymph Node Sampling in Early Stage Non-Small Cell Lung Cancer: A Prospective Study," *Respirology* 23, no. 1 (January 2018): 76–81.

29. Y. S. Song, C. M. Park, S. J. Park, S. M. Lee, Y. K. Jeon, and J. M. Goo, "Volume and Mass Doubling Times of Persistent Pulmonary Subsolid Nodules Detected in Patients Without Known Malignancy," *Radiology* 273, no. 1 (2014): 276–284.

30. J. Li, T. Xia, X. Yang, et al., "Malignant Solitary Pulmonary Nodules: Assessment of Mass Growth Rate and Doubling Time at Follow-Up CT," *Journal of Thoracic Disease* 10, no. Suppl 7 (2018): S797–S806, https://doi.org/10.21037/jtd.2018.04.25.

31. S. Liu, R. Wang, Y. Zhang, et al., "Precise Diagnosis of Intraoperative Frozen Section Is an Effective Method to Guide Resection Strategy for Peripheral Small-Sized Lung Adenocarcinoma," *Journal of Clinical Oncology* 34, no. 4 (February 2016): 307–313.

32. Y. Zhang, C. Deng, F. Fu, et al., "Excellent Prognosis of Patients With Invasive Lung Adenocarcinomas During Surgery Misdiagnosed as Atypical Adenomatous Hyperplasia, Adenocarcinoma In Situ, or Minimally Invasive Adenocarcinoma by Frozen Section," *Chest* 159, no. 3 (2021): 1265–1272, https://doi.org/10.1016/j.chest.2020.10.076.

33. E. Zhu, H. Xie, C. Dai, et al., "Intraoperatively Measured Tumor Size and Frozen Section Results Should Be Considered Jointly to Predict the Final Pathology for Lung Adenocarcinoma," *Modern Pathology* 31, no. 9 (September 2018): 1391–1399, https://doi.org/10.1038/s41379-018-0056-0.

34. Y. Sano, H. Shigematsu, M. Okazaki, et al., "Hoarseness After Radical Surgery With Systematic Lymph Node Dissection for Primary Lung Cancer," *European Journal of Cardio-Thoracic Surgery* 55, no. 2 (February 2019): 280–285.

35. C. A. Kutlu, A. Sayar, G. Olgac, et al., "Chylothorax: A Complication Following Lung Resection in Patients With NSCLC - Chylothorax Following Lung Resection," *Thoracic and Cardiovascular Surgeon* 51, no. 6 (2003): 342–345, https://doi.org/10.1055/s-2003-45423.

36. O. Laccourreye, D. Malinvaud, B. Delas, et al., "Early Unilateral Laryngeal Paralysis After Pulmonary Resection With Mediastinal Dissection for Cancer," *Annals of Thoracic Surgery* 90, no. 4 (October 2010): 1075–1078, https://doi.org/10.1016/j.athoracsur.2010.06.035.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.