World Journal of Surgical Oncology

RESEARCH

Open Access

Risk factors and clinical risk stratification of distant metastasis in early-stage lung cancer in never smokers

Dongsheng Wu[1†], Xiaohu Hao[1†], Zhipeng Gong[1†], Ruichen Cui[1], Liang Xia[1] and Lunxu Liu[1*]

Abstract

Background Risk factors for distant metastasis in early-stage lung cancer in never smokers (LCINS) remain poorly understood. This study aimed to identify key risk factors and to develop a clinical risk stratification model for early-stage LCINS.

Methods We retrospectively analyzed patients diagnosed with early-stage LCINS at West China Hospital, Sichuan University, from 2015 to 2020. Univariable and multivariable Cox regression analyses were performed to identify independent risk factors for distant metastasis. A predictive model was developed and internally validated using bootstrap resampling, with performance assessed by the concordance index (C-index), area under the receiver operating characteristic curve (AUC), calibration plot, and decision curve analysis.

Results A total of 1,406 patients with pathological stage I-II LCINS were included, among whom 76 (5.41%) developed distant metastasis during follow-up. Multivariable Cox regression analysis revealed that independent risk factors included advanced pathological T and N stages, higher consolidation-to-tumor ratio, and histologic subtype, particularly solid/micropapillary predominant adenocarcinoma. Based on these predictors, a predictive model was developed, demonstrating strong discrimination with a C-index of 0.799 and AUC values of 0.809, 0.791, and 0.783 for predicting 1-, 2-, and 3-year distant metastasis, respectively. Calibration and decision curve analyses confirmed the reliability and clinical utility of the model.

Conclusions This study identified risk factors and developed a clinical risk stratification model for distant metastasis in early-stage LCINS. This validated model enables risk stratification and personalized monitoring to facilitate early detection of distant recurrence in LCINS.

Keywords Lung cancer in never smokers, Distant metastasis, Risk factors, Predictive model

Background

Lung cancer remains the leading cause of both cancer incidence and cancer-related mortality worldwide [1]. Although the majority of cases are attributable to tobacco exposure, the incidence of lung cancer in never smokers (LCINS) has been rising and is estimated as the fifth leading cause of cancer-related deaths worldwide [2]. Surgical resection is the standard treatment for early-stage lung cancer, including LCINS, and generally offers favorable survival outcomes. Nevertheless, distant metastasis remains a major clinical challenge, with approximately

†Dongsheng Wu, Xiaohu Hao and Zhipeng Gong contributed equally to this work.

*Correspondence:
Lunxu Liu
lunxu_liu@aliyun.com
[1] Department of Thoracic Surgery and Institute of Thoracic Oncology, West China Hospital, Sichuan University, No. 37, Guoxue Alley, Chengdu 610041, Sichuan, China

34% of early-stage non-small cell lung cancer (NSCLC) developing metastatic progression after curative surgery [3]. Previous studies on the early-stage NSCLC have identified risk factors such as age, pleural invasion [4], histological types, and pathological stage [5] as key contributors to distant metastasis. However, whether these risk factors are equally relevant to early-stage LCINS remains largely unknown, as limited research has focused on the distant metastatic tissue in this population.

LCINS exhibits distinct clinicopathologic characteristics and genetic alterations compared to smoking-related lung cancer, which may influence its metastatic behavior. LCINS is predominantly adenocarcinoma and occurs more frequently in women and individuals of Asian descent [6, 7]. Studies have reported that individuals with adenocarcinoma have a higher tendency for distant metastasis than those with squamous cell carcinoma [5, 8]. Genomically, LCINS is characterized by a high prevalence of oncogenic alterations, particularly EGFR mutation and ALK rearrangement [9]. Population-based studies have revealed notable differences in EGFR mutation patterns by ethnicity and smoking status, with never smokers exhibiting distinct polymorphism profiles compared to smokers [10]. On one hand, 50–60% of NSCLC patients harboring these mutations develop brain metastasis, a significantly higher proportion compared to 16–20% in the overall NSCLC population [11]. On the other hand, the use of third-generation, brain-penetrant EGFR tyrosine kinase inhibitors (TKIs) has reduced the risk of brain metastasis during treatment by approximately threefold compared to earlier TKIs [12]. In addition, ROS1 rearrangement—a genetic alteration also commonly found in LCINS—has been associated with an increased risk of distant metastasis [11, 13, 14]. Collectively, LCINS may present a different clinical and genetic metastatic profile indicating the importance of studying LCINS as a separate population to better understand its risk factors of distant metastasis.

The present study aimed to retrospectively analyze the clinical features of early-stage LCINS who underwent surgical resection, identify key risk factors for distant metastasis, and develop a clinical risk stratification model.

Methods

This study was approved by the Institutional Review Board of West China Hospital, Sichuan University (No.2268) and adhered to the principles of the Declaration of Helsinki (2013 revision). Given the retrospective design and use of de-identified patient data, the need for written informed consent was waived.

Patient selection

Patients who underwent lung cancer surgery at West China Hospital, Sichuan University, between January 2015 and September 2020, were included in this study. Inclusion criteria required patients to be: (i) 18 years or older, (ii) never-smokers, (iii) diagnosed with pathological stage I-II NSCLC based on the 8th edition of the IASLC staging system, and (iv) recipients of lobectomy with systematic lymph node dissection. Exclusion criteria included: (i) patients with non-R0 resection; (ii) those who received neoadjuvant therapy; (iii) patients with a previous diagnosis of any malignancy; (iv) pathological evidence of pre-invasive adenocarcinoma, including atypical adenomatous hyperplasia, adenocarcinoma in situ (AIS), and minimally invasive adenocarcinoma (MIA); and (v) incomplete clinical information of interest.

Clinicopathologic data and variables definition

In this study, data on clinical variables, including age ($\leq$ 60 years or $>$60 years), sex, body mass index (BMI classified as $\leq$18.5 kg/m^2, 18.5–24.9 kg/m^2, and $>$25 kg/m^2), status of smoking, preoperative comorbidities assessed by the Charlson Comorbidity Index (CCI), tumor history, treatment information (including neoadjuvant therapy, adjuvant therapy, and surgical details), tumor characteristics (including tumor location, pathological stage, and histological types) were obtained from the Western China Lung Cancer Database, which was prospectively collected.

Additionally, the characteristics of solid components of tumor nodules were assessed via preoperative high-resolution computed tomography (CT). The solid component was characterized by areas of increased opacity that concealed underlying vascular structures. Following the Fleischner Society guidelines, the consolidation-to-tumor ratio (CTR) was determined by comparing the maximum diameter of the consolidated region to that of the tumor [15]. A pure ground-glass nodule (pGGN) was defined as having a CTR of 0, a part-solid nodule (PSN) exhibited focal opacities ($0 < CTR < 1.0$), and a solid nodule (SN) was characterized by a complete solid component ($CTR = 1$). Pathological N staging was categorized as N0, N1a (N1 at a single station), and N1b (N1 at multiple stations) [16]. Additionally, lung adenocarcinoma is divided into three subtypes based on the IASLC/ATS/ERS classification system: poorly differentiated (solid or micropapillary predominant), moderately differentiated (acinar or papillary predominant), and well-differentiated (lepidic predominant) [17].

Outcome and follow-up

The primary outcome of this study was the occurrence of distant metastasis during the long-term follow-up of patients who underwent radical resection. Distant metastasis was defined according to the American College of Surgeons Oncology Group criteria, encompassing both intrathoracic (contralateral, mediastinal, or hilar lymph nodes) and extrathoracic metastasis (brain, bone, liver, adrenal gland, distant lymph nodes, and other sites) [18]. The interval from the surgery date to the first observation of metastatic spread was defined as the time to distant metastasis.

Patient follow-up was conducted through regular telephone calls or visits to the outpatient department. Status assessments were based on periodic evaluations, which included chest CT scans, upper abdominal CT scans, and brain magnetic resonance imaging (MRI) or CT scans, as appropriate. The follow-up schedule was structured to include visits every 3 to 6 months during the first two years, every 6 months for the next three years, and annually thereafter. Additionally, additional procedures such as bone scans, positron emission tomography/computed tomography (PET/CT), or biopsies were conducted as needed.

Statistical analysis

Continuous variables were categorized based on their clinical relevance. Categorical variables were reported as numbers and percentages, and statistical significance was evaluated using the Pearson Chi-squared test. The Cox regression model was performed for both univariable and multivariable analyses to uncover independent risk factors for distant metastasis in early-stage LCINS. Variables with a P value of less than 0.05 in the univariable analysis were subsequently incorporated into the multivariable analysis. The independent risk factors identified were then used to construct a predictive model for distant metastasis at 1, 2, and 3 years.

To ensure unbiased estimates and minimize the risk of overfitting, internal validation of the nomogram was conducted through 1,000 bootstrap resampling. The performance of the predictive model was evaluated in terms of both discrimination and calibration. Discrimination was measured using the Harrell concordance index (C-index), along with the receiver operating characteristic (ROC) curves and the area under the curve (AUC). Calibration was evaluated using calibration curves. Decision curve analysis (DCA) was conducted to assess the potential clinical benefit of the model. Additionally, clinical risk stratification was performed using the "surv_cutpoint" function from the R package "survminer" to determine the optimal cutoff value for the clinical risk stratification. A significance level of $P < 0.05$ was established, and all tests were two-tailed. The statistical analyses were performed using R software (version 4.3.2).

Results

Baseline characteristics

In this study, a total of 1,406 LCINS who underwent lung cancer surgery were included (Fig. 1). Among them, 76 patients (5.41%) developed distant metastasis during the follow-up period. The baseline characteristics of the cohort are summarized in Table 1. No significant differences were observed between patients with and without distant metastasis regarding age ($P = 0.501$) or sex ($P = 0.371$). Patients who developed distant metastasis were more likely to have a CCI of ≥ 1 (85.53% vs. 74.66%, $P = 0.046$) and SN on preoperative CT (90.79% vs. 58.05%, $P < 0.001$). Additionally, adjuvant chemotherapy was administered more frequently in the metastasis group (35.53% vs. 12.86%, $P < 0.001$). Tumor-related characteristics also showed significant differences. Patients with distant metastasis were more likely to present with advanced pathological stages, particularly stage IIB (34.21% vs. 5.56%, $P < 0.001$). They also had higher pT and pN stages ($P < 0.001$). In the LCINS cohort, adenocarcinoma was the predominant histologic type, accounting for 97.44% of cases, while squamous cell carcinoma comprised only 2.56%. Significant differences were also seen in the distribution of adenocarcinoma subtypes between the metastasis and non-metastasis groups ($P < 0.001$).

The cohort had a median follow-up duration of 31.1 months, with an interquartile range spanning from 19.6 to 42.5 months. Distant metastasis rates at 1, 2, and 3 years were observed to be 2.13%, 3.77%, and 4.98%, respectively. The median interval from surgery to the onset of distant metastasis was 16.9 months, with an interquartile range of 9.5 to 26.1 months.

Risk factors of distant metastasis

Both univariable and multivariable Cox analyses identified four independent predictors of distant metastasis in early-stage LCINS (Table 2): (i) advanced pT stage, (ii) higher pN stage, (iii) increased CTR, and (iv) histologic subtype especially solid/micropapillary predominant adenocarcinoma.

The univariable analysis revealed several significant associations with distant metastasis risk (Table 2). While both CCI ≥ 1 (Hazard ratio [HR]: 1.94, 95% confidence interval [CI]: 1.02–3.68, $P = 0.042$) and adjuvant chemotherapy (HR: 3.11, 95% CI: 1.94–4.97, $P < 0.001$) showed initial significance, these associations were not maintained in multivariable analysis. Radiographic features demonstrated strong associations with metastasis risk: PSN (HR: 6.53; 95% CI: 1.25–33.96; $P = 0.026$) and SN (HR: 7.47; 95% CI: 1.76–31.66; $P = 0.006$) were associated

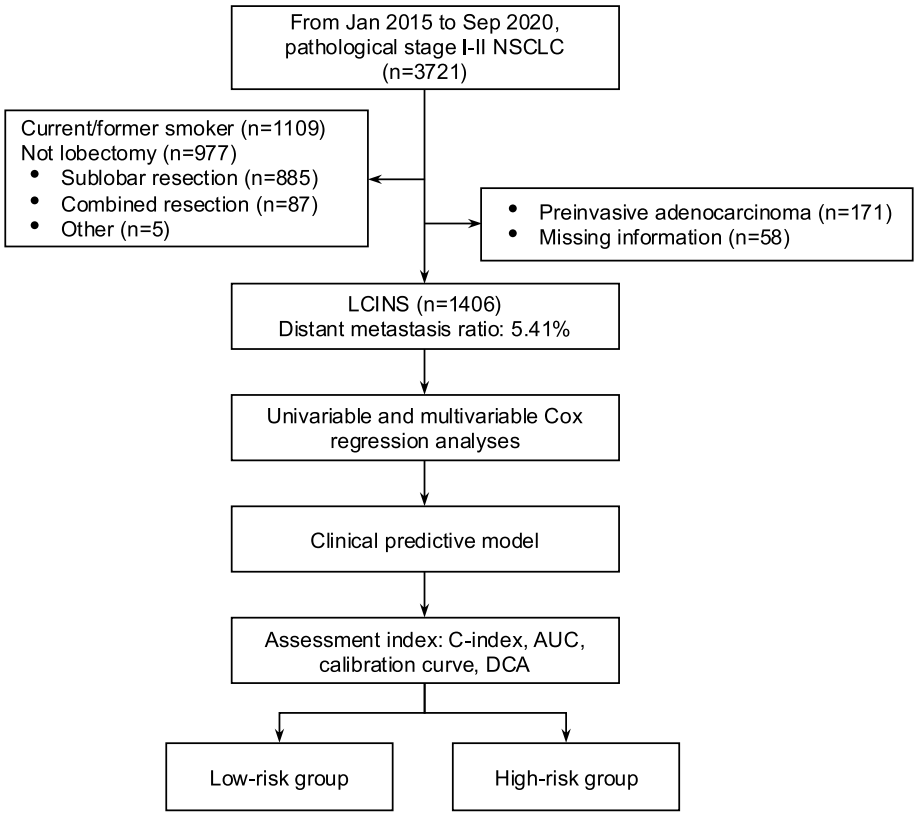

Fig. 1 Flowchart illustrating the patient selection process. Abbreviations: NSCLC, non-small cell lung cancer; LCINS, lung cancer in never smokers; C-index, Harrell concordance index; ROC, receiver operating characteristic curves; AUC, area under the curve; DCA, decision curve analysis

with significantly higher risk compared to pGGN. Among pathological variables, pT3 tumors showed the highest risk (HR: 8.86; 95% CI: 3.26–24.04; $P < 0.001$) relative to pT1 tumors. Similarly, nodal involvement was significantly associated with distant metastasis, with increased risk observed for both pN1a (HR: 4.58; 95% CI: 2.43–8.65; $P < 0.001$) and pN1b (HR: 5.63; 95% CI: 2.42–13.12; $P < 0.001$) relative to pN0. Histologically, solid/micropapillary-predominant adenocarcinomas were associated with a substantially higher risk of distant metastasis (HR: 5.56; 95% CI: 1.28–24.07; $P = 0.022$) compared to squamous cell carcinoma.

Model establishment and validation

Utilizing the independent predictors identified, we developed a predictive model for the risk of distant metastasis in early-stage LCINS, along with a corresponding nomogram (Fig. 2). The nomogram assigns weighted points to each variable: imaging features (pGGN = 0, SSN = 82, SN = 88 points), histologic subtypes (squamous/lepidic = 0, acinar/papillary = 30, solid/micropapillary = 78), nodal status (N0 = 0, N1a = 67, N1b = 78), and T stage (T1 = 0, T2 = 26, T3 = 100). This model demonstrated strong discriminatory ability, achieving a C-index of 0.799 (95% CI, 0.748–0.850). Time-dependent ROC curve analysis supported this, with AUC values of 0.809 (95% CI, 0.736–0.881) at 1 year, 0.791 (95% CI, 0.726–0.855) at 2 years, and 0.783 (95% CI, 0.726–0.840) at 3 years (Fig. 3A-C). Calibration plots indicated a close alignment between predicted and actual probabilities (Fig. 3D-F), confirming the reliability of the predictions. The DCA at 1, 2, and 3 years showed a clear net benefit, highlighting its clinical relevance (Fig. 3G-I). Internal validation using 1000 bootstrap resamples further validated its robustness, with C-index values of 0.809, 0.800, and 0.796 for predictions at 1, 2, and 3 years, respectively.

Clinical risk classification

Using the established nomogram, each individual was assigned a risk score. Based on an optimized cutoff of 138 points, individuals were classified into two distinct risk categories: a low-risk group (score range 0–138, distant metastasis rate of 2.18%) and a high-risk group (score range 138–350, distant metastasis rate of 15.14%). The analysis showed that the low-risk group experienced a significantly higher rate of distant metastasis-free survival related to the high-risk group ($P < 0.001$) (Fig. 4).

Table 1 Clinicopathological characteristics of patients with early-stage LCINS

Characteristics	Total ($n = 1406$)	No distant metastasis ($n = 1330$)	Distant metastasis ($n = 76$)	*P* value
Age, years				0.501
< 60	715 (50.85)	673 (50.60)	42 (55.26)	
≥ 60	691 (49.15)	657 (49.40)	34 (44.74)	
BMI, kg/m^2				0.611
< 18.5	66 (4.97)	62 (4.94)	4 (5.41)	
18.5–24.9	929 (69.90)	874 (69.64)	55 (74.32)	
≥ 25	334 (25.13)	319 (25.42)	15 (20.27)	
Sex				0.371
Male	303 (21.55)	283 (21.28)	20 (26.32)	
Female	1103 (78.45)	1047 (78.72)	56 (73.68)	
CCI				0.046
0	348 (24.75)	337 (25.34)	11 (14.47)	
≥ 1	1058 (75.25)	993 (74.66)	65 (85.53)	
Findings on HRCT				< 0.001
pGGN	412 (29.30)	410 (30.83)	2 (2.63)	
PSN	153 (10.88)	148 (11.13)	5 (6.58)	
SN	841 (59.82)	772 (58.05)	69 (90.79)	
Tumor location				0.616
Left	508 (36.13)	478 (35.94)	30 (39.47)	
Right	898 (63.87)	852 (64.06)	46 (60.53)	
Adjuvant radiotherapy				1
No	1404 (99.86)	1328 (99.85)	76 (100.00)	
Yes	2 (0.14)	2 (0.15)	0	
Adjuvant chemotherapy				< 0.001
No	1208 (85.92)	1159 (87.14)	49 (64.47)	
Yes	198 (14.08)	171 (12.86)	27 (35.53)	
Pathogenic stage				< 0.001
IA	601 (42.75)	590 (44.36)	11 (14.47)	
IB	680 (48.36)	645 (48.50)	35 (46.05)	
IIA	25 (1.78)	21 (1.58)	4 (5.26)	
IIB	100 (7.11)	74 (5.56)	26 (34.21)	
pT stage				< 0.001
T1	617 (43.88)	604 (45.41)	13 (17.11)	
T2	765 (54.41)	709 (53.31)	56 (73.68)	
T3	24 (1.71)	17 (1.28)	7 (9.21)	
pN stage				< 0.001
N0	1324 (94.17)	1273 (95.71)	51 (67.11)	
N1a	64 (4.55)	47 (3.53)	17 (22.37)	
N1b	18 (1.28)	10 (0.75)	8 (10.53)	
Histologic type				< 0.001
Squamous cell carcinoma	36 (2.56)	33 (2.48)	3 (3.95)	
Lepidic	831 (59.10)	813 (61.13)	18 (23.68)	
Acinar/papillary	528 (37.55)	478 (35.94)	50 (65.79)	
Solid/micropapillary	11 (0.78)	6 (0.45)	5 (6.58)	

Abbreviations: BMI Body mass index, *CCI* Charlson comorbidity index, *pGGN* Pure ground-glass nodule, *PSN* Part-solid nodule, *SN* Solid nodule

Table 2 Univariable and multivariable Cox regression analyses for patients with early-stage LCINS

Variables	Univariable			Multivariable		
	HR	95% CI	*P value*	HR	95% CI	*P value*
Age, years						
< 60	Ref					
≥ 60	0.83	(0.53–1.31)	0.427			
BMI, kg/m^2						
< 18.5	Ref					
18.5–24.9	0.95	(0.34–2.62)	0.921			
≥ 25	0.74	(0.24–2.22)	0.586			
Sex						
Male	Ref					
Female	0.70	(0.42–1.17)	0.170			
CCI						
0	Ref			Ref		
≥ 1	1.94	(1.02–3.68)	0.042	1.64	(0.86–3.15)	0.136
Solid component						
pGGN	Ref			Ref		
PSN	8.25	(1.60–42.56)	0.012	6.53	(1.25–33.96)	0.026
SN	16.42	(4.02–66.99)	< 0.001	7.47	(1.76–31.66)	0.006
Tumor location						
Left	Ref					
Right	0.87	(0.55–1.37)	0.538			
Adjuvant chemotherapy						
No	Ref			Ref		
Yes	3.11	(1.94–4.97)	< 0.001	0.85	(0.48–1.52)	0.587
pT stage						
T1	Ref					
T2	3.13	(1.71–5.73)	< 0.001	1.79	(0.96–3.33)	0.067
T3	16.44	(6.56–41.21)	< 0.001	8.86	(3.26–24.04)	< 0.001
pN stage						
N0	Ref			Ref		
N1a	7.49	(4.32–12.97)	< 0.001	4.58	(2.43–8.65)	< 0.001
N1b	10.92	(5.18–23.01)	< 0.001	5.63	(2.42–13.12)	< 0.001
Histologic type						
Squamous cell carcinoma	Ref			Ref		
Lepidic	0.26	(0.08–0.88)	0.030	0.93	(0.25–3.45)	0.917
Acinar/papillary	1.03	(0.32–3.29)	0.964	1.81	(0.53–6.20)	0.346
Solid/micropapillary	6.37	(1.52–26.66)	0.011	5.56	(1.28–24.07)	0.022

Abbreviations: HR Hazard ratio, *CI* Confidence interval, *Ref* Reference, *BMI* Body mass index, *CCI* Charlson comorbidity index, *pGGN* Pure ground-glass nodule, *PSN* part-solid nodule, *SN* Solid nodule

Discussion

Distant metastasis is a crucial factor affecting the prognosis of early-stage LCINS patients and significantly influences postoperative management. LCINS has distinct clinicopathologic features and genetic drivers compared to smoking-related lung cancer that may impact distant metastasis risk, highlighting the need to study LCINS separately to understand its unique risk factors. This study firstly identified differences in clinicopathologic characteristics between early-stage LCINS patients with and without distant metastasis. Significant risk factors were then confirmed through univariable and multivariable Cox regression analysis. Based on these predictors, we developed and validated a visual nomogram model to predict distant metastasis in early-stage LCINS. The predictive model demonstrated strong discriminatory ability and clinical utility.

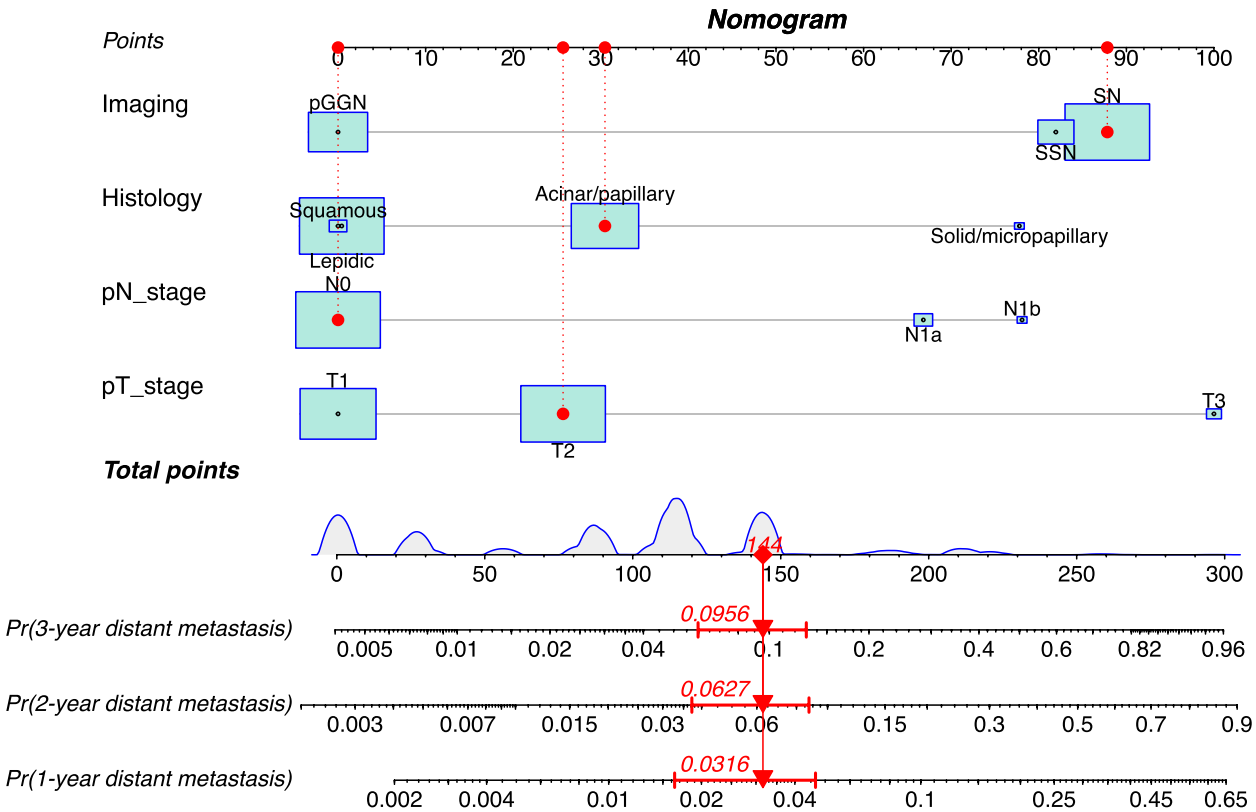

Fig. 2 Nomogram model for predicting distant metastasis in patients with early-stage LCINS. The patient #6 is illustrated in the nomogram by mapping its values to the covariate scales. The probability of distant metastasis in 1-, 2-, and 3-year follow-up are estimated to be 3.16%, 6.27%, and 9.56%, respectively. Abbreviation: LCINS, lung cancer in never smokers

The nomogram developed for predicting distant metastasis included four critical risk factors: advanced pT and pN stages, a high CTR in tumor nodules, and specific histological subtypes, particularly the presence of solid/micropapillary-predominant adenocarcinoma. These factors align closely with findings from previous clinical studies [4, 5, 19]. This suggests that physicians should prioritize these high-risk patients and imaging approach remains a reliable and convenient method for identifying patients with high risk of distant metastasis [20]. Specifically, we identified pathological T and N stages as independent risk indicators for distant metastasis in early-stage LCINS after radical resection, two crucial elements within the TNM staging system for guiding treatment strategies and prognosis. Earlier studies by Wang et al. [4] and Tian et al. [19] highlighted T stage as a key factor in distant metastasis risk among early-stage NSCLC patients, drawing from data in the Surveillance, Epidemiology, and End Results (SEER) database. Lymph node involvement emerged as a key predictor of distant metastasis in our model, aligning with previous studies that emphasized its role in distant recurrence [21, 22]. According to Yang et al. [23], patients with M1b NSCLC exhibited higher proportion of N1, N2, and N3 involvement compared to those with M1a disease, and higher N stages were related with an elevated risk of multiorgan metastasis.

In the revised 8th edition TNM classification by the IASLC, survival analysis of N descriptors indicated that N1 involvement across multiple nodal stations is associated with a significantly worse prognosis than single-station N1 involvement [16]. Building on these insights, our study closely examined how the number of metastatic lymph node stations impacts the risk of distant metastasis in early-stage LCINS. Results showed that pathological N1 multiple-station involvement significantly raised the risk of distant metastasis over N1 single-station cases, consistent with previous studies that explored novel pathological N-stage classifications based on factors such as the location of metastatic lymph nodes and number of involvement stations [24, 25].

Regarding the histological type, our study found that lung adenocarcinoma was associated with a significantly higher risk of distant metastasis than squamous cell carcinoma. Indeed, our previous study reported that adenocarcinoma is the most common subtype to develop

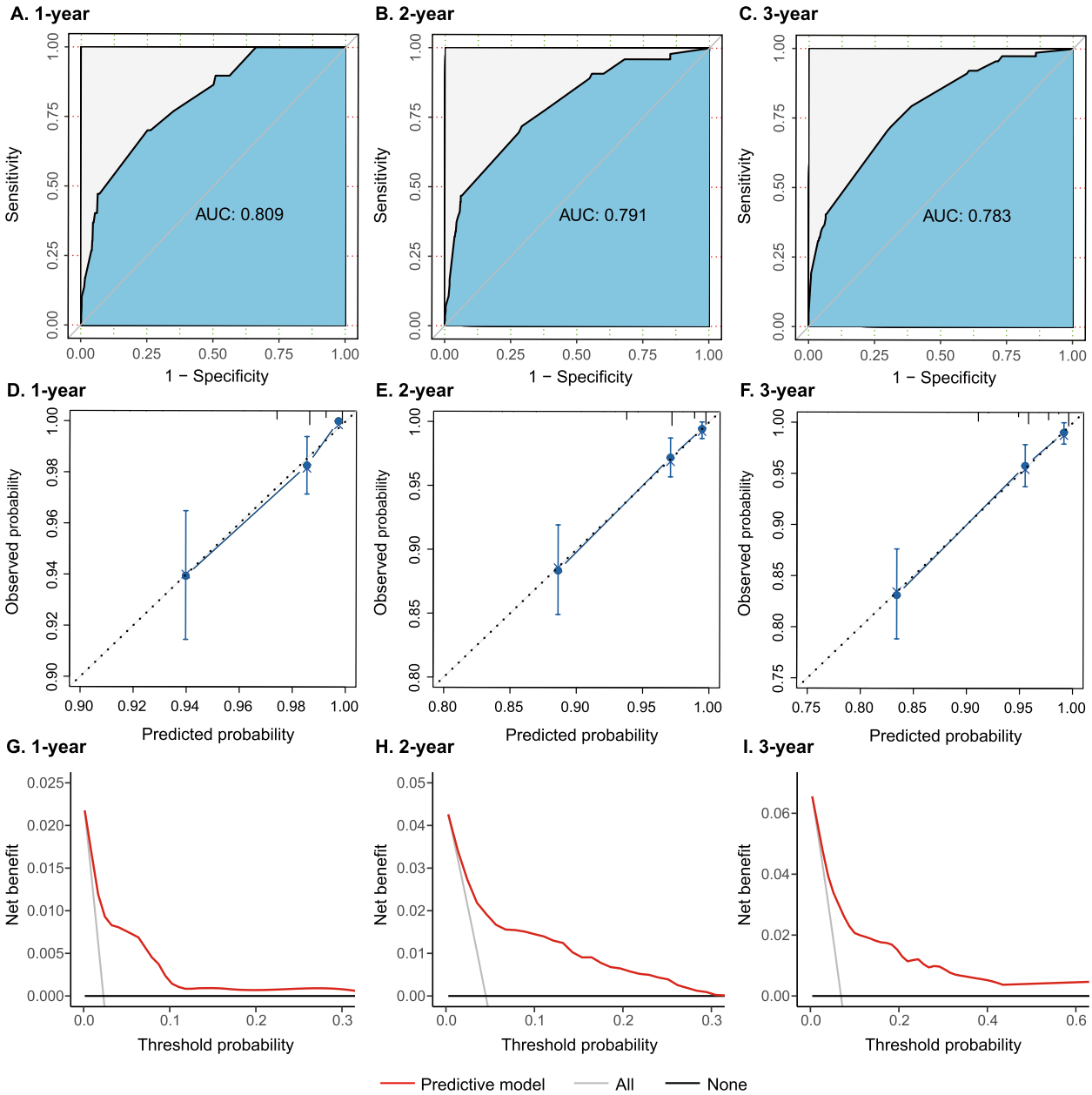

Fig. 3 **A**-**C** Time-dependent ROC curves for predicting the probability of distant metastasis at 1, 2, and 3 years, respectively. **D**-**F** Calibration curves showing predicted versus observed probabilities of distant metastasis at 1, 2, and 3 years, respectively. **G**-**I** DCA illustrating the clinical utility of the model at 1, 2, and 3 years, respectively

bone metastasis (87.04%), whereas squamous cell carcinoma presents with significantly lower bone metastasis risk (12.96%) [5]. Additionally, another clinical investigation revealed that more than 50% of lung cancer patients with bone metastasis had adenocarcinoma [26], while a tumor registry study in Sweden reported a bone metastasis incidence of up to 39% among adenocarcinoma cases [27]. Moreover, patients with adenocarcinoma were observed to have a risk of brain metastasis that was 2.86 times greater than that of individuals with non-adenocarcinoma NSCLC [22]. We further examined the specific adenocarcinoma subtypes, as classified by the IASLC/ATS/ERS system: poorly differentiated (solid or micropapillary predominant), moderately differentiated (acinar or papillary predominant), and well-differentiated (lepidic predominant) [17]. Results indicated that the solid/micropapillary subtype was significantly associated with the highest risk of distant

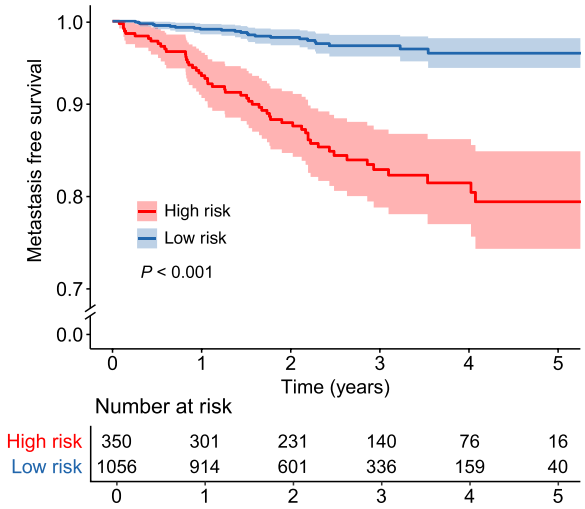

Fig. 4 Kaplan–Meier curves comparing distant metastasis outcomes between low- and high-risk groups for early-stage LCINS. Abbreviation: LCINS, lung cancer in never smokers

metastasis, followed by acinar/papillary types, while the lepidic subtype exhibited the lowest metastasis risk. This aligns with prior findings regarding the prognostic distribution of adenocarcinoma subtypes, indicating substantial differences in distant metastasis risk across these classifications [17, 28].

Our analysis identified tumor nodules with high CTR of as the most significant predictor for distant metastasis in early-stage LCINS. Previous studies have indicated that an elevated CTR is associated with more aggressive tumor behavior and a worse prognosis. Specifically, CTR has been extensively researched as a predictor of lymph node metastasis, particularly in Asian populations [29, 30]. Chen et al. [31] reported that CTR not inferior to primary tumor Standardized Uptake Value max (SUVmax) in its preoperative predictive value for lymphatic metastasis in lung cancer patients with pGGNs. Furthermore, Lin et al. [32] found that patients in the higher CTR subgroup exhibited more invasive adenocarcinomas and a greater incidence of visceral pleural invasion than those in the lower CTR subgroup. In a study of lung cancer patients who received surgery, all cases with lymph node metastasis had a CTR exceeding 60% [31]. These findings collectively support that a high CTR on preoperative CT scans may serve as a valuable indicator for distant metastasis risk in early-stage LCINS.

Routine postoperative surveillance for early-stage LCINS primarily relies on clinical assessment and chest CT, while advanced imaging (e.g., brain MRI, bone scintigraphy) is typically reserved for symptomatic patients due to cost and radiation concerns. Although studies suggest early detection of asymptomatic metastases may improve outcomes [33, 34], current guidelines do not recommend routine systemic imaging for resected stage I-II NSCLC [35], creating a critical need for risk-stratified approaches. Our model addresses this gap by providing the first validated LCINS-specific risk quantification tool (C-index 0.799), enabling personalized surveillance strategies: high-risk patients (score >138, 15.14% metastasis rate) may benefit from intensified protocols (e.g., annual brain MRI/PET-CT), while low-risk patients (score ≤138, 2.18% metastasis rate) can avoid unnecessary procedures. Notably, the model identifies imaging/histologic predictors that outperform conventional TNM staging alone, reflecting LCINS's unique biology. By integrating these features into an accessible nomogram, we facilitate earlier detection and targeted interventions, optimizing resource utilization without compromising oncologic outcomes in this distinct population.

While our model offers practical and clinically relevant insights, several limitations should be noted. Firstly, as a retrospective study, it is susceptible to selection bias, and being conducted at a single center further limits the generalizability of our findings. Secondly, although factors such as EGFR mutations, ALK rearrangement, and lactate dehydrogenase, which are highly prevalent in LCINS, are potentially associated with distant metastasis [36–38], we were unable to include these variables due to limited data availability in our dataset. Thirdly, while the absence of metastatic events in excluded pre-invasive cases aligns with the indolent nature of AIS/MIA, further investigation through larger-scale studies with extended follow-up periods would be valuable to comprehensively evaluate metastatic potential in this population. Fourthly, our follow-up period was limited, potentially leading to an underestimation of the true incidence of distant metastasis. Nevertheless, our median follow-up time exceeded the median time to distant metastasis observed in our cohort, supporting the reliability of our findings. Additional confirmation of our findings in wider populations requires validation through prospective multicenter studies.

Conclusions

In summary, distant metastasis in early-stage LCINS is a significant concern, and we identified several independent risk factors, including advanced pT and pN stages, tumor with higher CTR, and specific histological subtypes. Our predictive model demonstrated robust performance in stratifying patients by their risk of distant metastasis. This model holds potential for enhancing clinical decision-making and helping for personalized surveillance, improving the management and prognosis of early-stage LCINS.

Abbreviations

LCINS	Lung cancer in never smokers
NSCLC	Non-small cell lung cancer
TKIs	Tyrosine kinase inhibitors
CCI	Charlson Comorbidity Index
CT	Computed tomography
CTR	Consolidation-to-tumor ratio
pGGN	Pure ground-glass nodule
SN	Solid nodule
MRI	Magnetic resonance imaging
DCA	Decision curve analysis
C-index	Harrell concordance index
ROC	Receiver operating characteristic
AUC	Area under the curve

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12957-025-03892-1.

Supplementary Material 1.

Acknowledgements
Not available.

Authors' contributions
D.S.W. and X.H.H. conceived and designed the study; D.S.W. and Z.P.G carried out experiments; X.H.H. and R.C.C. analyzed the data; D.S.W., X.H.H., and Z.P.G made the figures; R.C.C. and L.X. drafted and revised the paper; L.X. and L.X.L. reviewed the paper; all authors approved the final version of the manuscript.

Funding
This work was supported by 1.3.5 Project for Disciplines of Excellence, West China Hospital, Sichuan University (ZYGD23010 to Dr. Lunxu Liu), and Postdoctor Research Fund of West China Hospital, Sichuan University (2023HXBH045 to Dr. Xiaohu Hao).

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

Declarations

Ethics approval and consent to participate
This is an observational study. The Institutional Review Board of West China Hospital, Sichuan University has confirmed that no ethical approval is required.

Consent for publication
All authors gave consent for the publication of this study.

Competing interests
The authors declare no competing interests.

Received: 26 March 2025 Accepted: 8 June 2025
Published online: 18 June 2025

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(3):229–63.
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021;71(3):209–49.
3. Kelsey CR, Marks LB, Hollis D, Hubbs JL, Ready NE, D'Amico TA, et al. Local recurrence after surgery for early stage lung cancer: an 11-year experience with 975 patients. Cancer. 2009;115(22):5218–27.
4. Wang T, Wang Z, Zhou J, Jie H, Liao H, Mei J, et al. A nomogram predicting the risk of extrathoracic metastasis at initial diagnosis of T(≤3cm)N(0) lung cancer. Transl Lung Cancer Res. 2024;13(8):1841–50.
5. Zhou J, Wu D, Zheng Q, Wang T, Mei J. Development of a predictive model to predict postoperative bone metastasis in pathological I-II non-small cell lung cancer. Transl Lung Cancer Res. 2024;13(5):998–1009.
6. Devarakonda S, Li Y, Martins Rodrigues F, Sankararaman S, Kadara H, Goparaju C, et al. Genomic Profiling of Lung Adenocarcinoma in Never-Smokers. J Clin Oncol. 2021;39(33):3747–58.
7. Daylan AEC, Miao E, Tang K, Chiu G, Cheng H. Lung Cancer in Never Smokers: Delving into Epidemiology, Genomic and Immune Landscape, Prognosis, Treatment, and Screening. Lung. 2023;201(6):521–9.
8. Wang X, Wang Z, Pan J, Lu ZY, Xu D, Zhang HJ, et al. Patterns of Extrathoracic Metastases in Different Histological Types of Lung Cancer. Front Oncol. 2020;10:715.
9. Planchard D, Besse B. Lung cancer in never-smokers. Eur Respir J. 2015;45(5):1214–7.
10. Obradović J, Ništević-Lazović J, Sekeruš V, Milašin J, Perin B, Jurisic V. Investigating the frequencies of EGFR mutations and EGFR single nucleotide polymorphisms genotypes and their predictive role in NSCLC patients in Republic of Serbia. Mol Biol Rep. 2025;52(1):350.
11. LoPiccolo J, Gusev A, Christiani DC, Janne PA. Lung cancer in patients who have never smoked - an emerging disease. Nat Rev Clin Oncol. 2024;21(2):121–46.
12. Reungwetwattana T, Nakagawa K, Cho BC, Cobo M, Cho EK, Bertolini A, et al. CNS Response to osimertinib versus standard epidermal growth factor receptor tyrosine kinase inhibitors in patients with untreated EGFR-mutated advanced non-small-cell lung cancer. J Clin Oncol. 2018;36(33):3290–7.
13. Gainor JF, Tseng D, Yoda S, Dagogo-Jack I, Friboulet L, Lin JJ, et al. Patterns of metastatic spread and mechanisms of resistance to crizotinib in ROS1-positive non-small-cell lung cancer. JCO Precis Oncol. 2017;(1):1–13.
14. Patil T, Smith DE, Bunn PA, Aisner DL, Le AT, Hancock M, et al. The Incidence of Brain Metastases in Stage IV ROS1-Rearranged Non-Small Cell Lung Cancer and Rate of Central Nervous System Progression on Crizotinib. J Thorac Oncol. 2018;13(11):1717–26.
15. MacMahon H, Naidich DP, Goo JM, Lee KS, Leung ANC, Mayo JR, et al. Guidelines for Management of Incidental Pulmonary Nodules Detected on CT Images: From the Fleischner Society 2017. Radiology. 2017;284(1):228–43.
16. Asamura H, Chansky K, Crowley J, Goldstraw P, Rusch VW, Vansteenkiste JF, et al. The International Association for the Study of Lung Cancer Lung Cancer Staging Project: Proposals for the Revision of the N Descriptors in the Forthcoming 8th Edition of the TNM Classification for Lung Cancer. J Thorac Oncol. 2015;10(12):1675–84.
17. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger KR, Yatabe Y, et al. International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. J Thorac Oncol. 2011;6(2):244–85.
18. Fernando HC, Landreneau RJ, Mandrekar SJ, Nichols FC, Hillman SL, Heron DE, et al. Impact of brachytherapy on local recurrence rates after sublobar resection: results from ACOSOG Z4032 (Alliance), a phase III randomized trial for high-risk operable non-small-cell lung cancer. J Clin Oncol. 2014;32(23):2456–62.
19. Tian Y, He Y, Li X, Liu X. Novel nomograms to predict lymph node metastasis and distant metastasis in resected patients with early-stage non-small cell lung cancer. Ann Palliat Med. 2021;10(3):2548–66.
20. Mok TSK, Wu YL, Kudaba I, Kowalski DM, Cho BC, Turna HZ, et al. Pembrolizumab versus chemotherapy for previously untreated, PD-L1-expressing, locally advanced or metastatic non-small-cell lung cancer (KEYNOTE-042): a randomised, open-label, controlled, phase 3 trial. Lancet. 2019;393(10183):1819–30.
21. Zhang J, Xu J, Jin S, Gao W, Guo R, Chen L. The development and validation of a nomogram for predicting brain metastases in lung squamous cell carcinoma patients: an analysis of the Surveillance, Epidemiology, and End Results (SEER) database. J Thorac Dis. 2021;13(1):270–81.

22. Zhang F, Zheng W, Ying L, Wu J, Wu S, Ma S, et al. A Nomogram to Predict Brain Metastases of Resected Non-Small Cell Lung Cancer Patients. Ann Surg Oncol. 2016;23(9):3033–9.
23. Yang J, Peng A, Wang B, Gusdon AM, Sun X, Jiang G, et al. The prognostic impact of lymph node metastasis in patients with non-small cell lung cancer and distant organ metastasis. Clin Exp Metastasis. 2019;36(5):457–66.
24. Aksoy Y, Citak N, Obuz C, Metin M, Sayar A. Does anatomical location-based metastatic lymph node density affect prognosis in lung cancer patients? Interact Cardiovasc Thorac Surg. 2021;33(4):541–9.
25. Wang Z, Yang Z, Li S, Zhang J, Xia L, Zhou J, et al. A comprehensive comparison of different nodal subclassification methods in surgically resected non-small-cell lung cancer patients. Ann Surg Oncol. 2022;29(13):8144–53.
26. Oliveira MB, Mello FC, Paschoal ME. The relationship between lung cancer histology and the clinicopathological characteristics of bone metastases. Lung Cancer. 2016;96:19–24.
27. Riihimäki M, Hemminki A, Fallah M, Thomsen H, Sundquist K, Sundquist J, et al. Metastatic sites and survival in lung cancer. Lung Cancer. 2014;86(1):78–84.
28. Moreira AL, Ocampo PSS, Xia Y, Zhong H, Russell PA, Minami Y, et al. A Grading System for Invasive Pulmonary Adenocarcinoma: A Proposal From the International Association for the Study of Lung Cancer Pathology Committee. J Thorac Oncol. 2020;15(10):1599–610.
29. Suzuki K, Koike T, Asakawa T, Kusumoto M, Asamura H, Nagai K, 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). J Thorac Oncol. 2011;6(4):751–6.
30. Asamura H, Hishida T, Suzuki K, Koike T, Nakamura K, Kusumoto M, et al. Radiographically determined noninvasive adenocarcinoma of the lung: survival outcomes of Japan Clinical Oncology Group 0201. J Thorac Cardiovasc Surg. 2013;146(1):24–30.
31. Chen YC, Lin YH, Chien HC, Hsu PK, Hung JJ, Huang CS, et al. Preoperative consolidation-to-tumor ratio is effective in the prediction of lymph node metastasis in patients with pulmonary ground-glass component nodules. Thorac Cancer. 2021;12(8):1203–9.
32. Lin B, Wang R, Chen L, Gu Z, Ji C, Fang W. Should resection extent be decided by total lesion size or solid component size in ground glass opacity-containing lung adenocarcinomas? Transl Lung Cancer Res. 2021;10(6):2487–99.
33. Nishikawa T, Ueba T, Kawashima M, Kajiwara M, Iwata R, Kato M, et al. Early detection of metachronous brain metastases by biannual brain MRI follow-up may provide patients with non-small cell lung cancer with more opportunities to have radiosurgery. Clin Neurol Neurosurg. 2010;112(9):770–4.
34. Patchell RA, Tibbs PA, Walsh JW, Dempsey RJ, Maruyama Y, Kryscio RJ, et al. A randomized trial of surgery in the treatment of single metastases to the brain. N Engl J Med. 1990;322(8):494–500.
35. Schneider BJ, Ismaila N, Aerts J, Chiles C, Daly ME, Detterbeck FC, et al. Lung Cancer Surveillance After Definitive Curative-Intent Therapy: ASCO Guideline. J Clin Oncol. 2020;38(7):753–66.
36. Jurišić V, Obradovic J, Pavlović S, Djordjevic N. Epidermal Growth Factor Receptor Gene in Non-Small-Cell Lung Cancer: The Importance of Promoter Polymorphism Investigation. Anal Cell Pathol (Amst). 2018;2018:6192187.
37. Jurisic V, Radenkovic S, Konjevic G. The Actual Role of LDH as Tumor Marker, Biochemical and Clinical Aspects. Adv Exp Med Biol. 2015;867:115–24.
38. Xing W, Li X, Zhou Y, Li M, Zhu M. Lactate metabolic pathway regulates tumor cell metastasis and its use as a new therapeutic target. Explor Med. 2023;4(4):542–60.

Publisher's Note