Am J Cancer Res 2024;14(12):5952-5964
www.ajcr.us /ISSN:2156-6976/ajcr0160163

Original Article

Risk factors for mediastinal lymph node metastases in early-stage non-small-cell lung cancer and prediction model establishment

Yubo Tang[1*], A Garu[2*], Xiao Chen[1], Ziyun Guan[3], Xingdong Cai[4], Huaxing Huang[5], Linghu Xitao[2], Kejing Tang[6], Yong Dong[2,7]

[1]Department of Pharmacy, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou 510080, Guangdong, China; [2]Department of Orthopaedic Surgery, The Second Affiliated Hospital of Guangzhou Medical University, Guangzhou 510260, Guangdong, China; [3]Department of Emergency, The Sixth Affiliated Hospital of South China University of Technology School of Medicine, Foshan 528200, Guangdong, China; [4]Department of Pulmonary and Critical Care Medicine, The First Affiliated Hospital of Jinan University, Guangzhou 510630, Guangdong, China; [5]Department of Pulmonary and Critical Care Medicine, The Second Affiliated Hospital of Guangzhou Medical University, Guangzhou 510260, Guangdong, China; [6]Department of Pulmonary and Critical Care Medicine, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou 510080, Guangdong, China; [7]Department of Radiation Oncology, The First Affiliated Dongguan Hospital of Guangdong Medical University, Dongguan 523710, Guangdong, China. [*]Equal contributors and co-first authors.

Received August 29, 2024; Accepted December 5, 2024; Epub December 15, 2024; Published December 30, 2024

Abstract: This study aimed to explore the risk factors for mediastinal lymph node metastases (MLNM) in patients with early-stage non-small-cell lung cancer (NSCLC) and to establish a predictive model. A retrospective analysis was conducted on the clinical data from NSCLC patients treated at the Second Affiliated Hospital of Guangzhou Medical University and the First Affiliated Dongguan Hospital of Guangdong Medical University between March 2021 and March 2023. Baseline clinical data, laboratory parameters, and pathological features were collected and analyzed. Univariate and multivariate logistic regression identified several independent risk factors for MLNM, including Cyfra21-1, D-dimer (D-D), tumor size, percentage of tumor solid, and lesion location. These risk factors were incorporated into a Nomogram model to visually assess the likelihood of MLNM. The model demonstrated excellent diagnostic accuracy with an area under the curve (AUC) of 0.904, a specificity of 73.85%, and a sensitivity of 93.68%. Cyfra21-1 and D-D were particularly significant predictors of MLNM. This Nomogram model provides an effective and practical tool for assessing MLNM risk in early-stage NSCLC, aiding clinical decision-making and optimizing treatment strategies.

Keywords: Non-small-cell lung cancer, mediastinal lymph node metastasis, risk factors, prediction model, Nomogram

Introduction

Malignant tumors pose a significant public health challenge globally, accounting for 25% of all annual deaths. In China, lung cancer (LC) has the highest incidence and mortality rates among malignancies [1]. According to the data in 2022 [2], the annual incidence and mortality of LC were 2.48 million and 1.81 million, respectively. A major contributing factor is that most LC cases are diagnosed at advanced stages, missing the optimal window for treatment and resulting in suboptimal results even after clinical intervention [3]. Early-stage (ES) LC often lacks obvious symptoms and is mostly detected during routine physical examinations. Therefore, early diagnosis and treatment are crucial for improving patient prognosis. Research has shown that the computerized tomography (CT) features of ES-LC, such as spiculation, lobulation, bronchial indentation, and vascular convergence signs, are of great significance for diagnosis [4]. With growing health awareness, routine physical examina-

https://doi.org/10.62347/DIZG4944

tions are becoming a vital tool to detect diseases, enabling more ES-LC to be detected, diagnosed, and treated early.

Surgery remains a cornerstone for LC treatment, with various surgical modalities available [5]. Lymph node metastasis (LNM) is a key prognostic factor and plays a crucial role in guiding postoperative treatment strategies and surgical method selection [6]. The presence of mediastinal LNM (MLNM) is a critical determinant of postoperative prognosis in LC [7]. For patients with stage I-II non-small-cell LC (NSCLC) and tumor diameter ≤3.0 cm, the 5-year survival rate is approximately 25-73% if LNM is absent or limited to intrapulmonary nodes, but drops significantly to 2-24% in cases of MLNM [8].

Due to the low probability of LNM in ES-NSCLC, there is controversy over the optimal method for mediastinal lymph node dissection. Some scholars suggest systematic lymph node dissection to ensure precise postoperative TNM staging and to minimize residual tumor cells [9]. However, others believe that excessive lymph node dissection can significantly affect patients' postoperative quality of life and increase the risk of complications, such as chylothorax and hoarseness [10]. As a result, no consensus has been reached on the ideal lymph node dissection approach for ES-NSCLC. Studies have reported that ES-NSCLC with a tumor diameter (T) ≤2.0 cm still has a 15%-20% probability of developing LNM [8]. The overall 5-year survival rates for patients with N1 or N2 LNM are approximately 67% and 37%, respectively [11]. This highlights the importance of preoperative evaluation of MLNM, which can guide the selection of appropriate intraoperative lymph node dissection methods.

While numerous studies have explored factors associated with lung cancer prognosis, such as tumor size, histological subtype, and biomarkers, limited research has comprehensively integrated multiple laboratory parameters and pathological features to predict MLNM in early-stage NSCLC. The innovation of this study lies in combining key biomarkers such as Cyfra21-1 and D-dimer (D-D) with tumor characteristics into a Nomogram model, which provides a visual and clinically applicable tool for predicting MLNM risk. This model offers a more personalized and accurate approach to preoperative evaluation, potentially guiding lymph node dissection decisions and improving patient outcomes in early-stage NSCLC.

Methods and materials

General data

This study retrospectively included 225 NSCLC patients who received treatment at the Second Affiliated Hospital of Guangzhou Medical University and the First Affiliated Dongguan Hospital of Guangdong Medical University between March 2021 and March 2023. Among these patients, 95 were diagnosed with MLNM, and the other 130 did not.

Ethical statement

This study was approved by the Second Affiliated Hospital of Guangzhou Medical University Medical Ethics Committee.

Eligibility and exclusion criteria

Inclusion criteria: Age >18 years; single nodular lesion in the lungs; tumor diameter ≤3.0 cm with surgical indications; confirmed NSCLC by postoperative paraffin pathological sections; tumor classified as T1 based on the 8th edition of TNM staging criteria [12]; complete case data.

Exclusion criteria: Coexistence of other malignancies; inadequate lymph node dissection, defined as the removal of fewer than 3 mediastinal lymph nodes or fewer than 10 total lymph nodes; history of preoperative neoadjuvant chemoradiotherapy or molecular targeted therapy; previous history of malignancies.

Definition of MLNM

MLNM was determined based on preoperative imaging (e.g., CT and positron emission tomography [PET]-CT) or intraoperative pathology (lymph node biopsy or dissection). Specific criteria: preoperative CT or PET-CT showing abnormal enlargement or increased metabolic activity of mediastinal lymph nodes, suggesting possible metastasis; intraoperative or postoperative pathological examination confirming the presence of cancer cells in the mediastinal lymph nodes [13].

Data collection

Baseline, laboratory, and pathological data were collected from patients' electronic medical records. The baseline information included age, sex, body mass index (BMI), smoking history (pack-years), lesion location, and history of comorbidities, including hypertension, diabetes, cardiovascular disease, and cerebrovascular disease. Laboratory parameters analyzed included hemoglobin (HB), hematocrit (HCT), platelet count (PLT), red cell distribution width (RDW), lymphocyte count (Lym), neutrophil count (Neu), neutrophil-to-lymphocyte ratio (NLR), systemic immune-inflammation index (SII), neuron-specific enolase (NSE), carcinoembryonic antigen (CEA), cytokeratin fragment 21-1 (Cyfra21-1), D-dimer (D-D), albumin (ALB), alkaline phosphatase (ALP), and fibrinogen (FIB). Pathological data included tumor size, presence of multiple nodules, percentage of tumor solid, and lesion location. Note: The laboratory parameters were the data of the last test before surgery.

Outcome measures

1. Differences in patient baseline information, laboratory parameters, and pathological data were analyzed. 2. The diagnostic efficacy of laboratory parameters for MLNM was analyzed by receiver operating characteristic (ROC) curves. 3. Independent risk factors for MLNM were identified by Logistic regression analysis. 4. A Nomogram was developed to visually represent MLNM risk, and the diagnostic efficiency of the model was evaluated.

Statistical methods

SPSS 26.0 was used for data analysis. GraphPad Prism 9.5.1 software was used for figure rendering. All measurement data were subjected to normality and homogeneity of variance tests. Data following a normal distribution were represented by mean ± standard deviation, and the between-group comparisons employed independent sample t-tests and one-way ANOVA as appropriate; Data with non-normal distribution were represented by the median (lower and upper quartiles), with inter-group comparisons conducted using non-parametric tests for two or multiple independent samples. Count data were expressed by the number of cases, and chi-square tests were performed to identify inter-group differences. The independent factors associated with MLNM were identified using binary logistic regression. The cut-off values were determined by the optimal Youden index, the sensitivity and specificity corresponding to each cut-off value were calculated separately, and the diagnostic efficacy was assessed by the area under the curve (AUC). Nomogram plots and calibration curves were plotted using the "rms" package of R (4.3.2) to evaluate the agreement between predicted and observed probabilities. Statistical significance was defined as $P<0.05$.

Results

Comparison of baseline information

No statistical differences were observed between the metastasis and non-metastasis groups regarding age, sex, BMI, pack-year, lesion location, or history of hypertension, diabetes, cardiovascular disease, or cerebrovascular disease (all $P>0.05$, **Table 1**).

Comparison of laboratory parameters

The inter-group comparison of laboratory parameters revealed no marked differences in preoperative Hb ($P=0.086$), HCT ($P=0.581$), PLT ($P=0.068$), RDW ($P=0.671$), Lym ($P=0.141$), Neu ($P=0.524$), NLR ($P=0.945$), SII ($P=0.599$), NSE ($P=0.093$), CEA ($P=0.862$), ALB ($P=0.579$), ALP ($P=0.609$), and FIB ($P=0.546$). However, Cyfra21-1 ($P<0.001$) and D-D ($P<0.001$) were significantly higher in the metastasis group compared to those in the non-metastasis group (**Table 2**).

Comparison of pathological data

The proportion of patients with multiple nodules was not significantly differed between the two groups ($P=0.644$) (**Table 3**). However, the metastasis group had a greater number of patients with a tumor size ≥2 cm ($P=0.001$), a higher percentage of tumor solid ($P<0.001$), and central lesions ($P<0.001$) than the non-metastasis group.

Efficacy of laboratory parameters in diagnosing MLNM

ROC analysis was used to assess the diagnostic performance of laboratory parameters for

Table 1. Comparison of baseline information between the two groups

Variables	Metastasis group (n=95)	Non-metastasis group (n=130)	χ^2	P
Age				
≥60	39	48	0.395	0.530
<60	56	82		
Sex				
Male	46	68	0.332	0.565
Female	49	62		
Body mass index				
≥25 kg/m^2	30	35	0.579	0.447
<25 kg/m^2	65	95		
Pack-year				
≥400	62	95	1.589	0.207
<400	33	35		
Lesion location				
Left lung	44	55	0.358	0.550
Right lung	51	75		
History of hypertension				
With	81	118	1.628	0.202
Without	14	12		
History of diabetes				
With	7	14	0.750	0.386
Without	88	116		
History of cardiovascular disease				
With	4	9	0.742	0.389
Without	91	121		
History of cerebrovascular disease				
With	3	10	2.073	0.150
Without	92	120		

MLNM. Cyfra21-1 and D-D demonstrated superior diagnostic efficacy with an area under the curve (AUC) >0.7, indicating moderate diagnostic value. Other laboratory indicators showed limited diagnostic efficacy, with AUC values near 0.5 (**Figure 1**; **Table 4**). These results underscore the potential of Cyfra21-1 and D-D as predictive biomarkers for mediastinal lymph node metastasis in early-stage NSCLC patients.

Analysis of risk factors for MLNM

We used logistic regression analysis to analyze the risk factors of MLNM, assigning values to all the data first (**Table 5**). Univariate analysis revealed that PLT (P=0.026, OR=1.850), RDW (P=0.010, OR=2.202), NLR (P=0.031, OR=3.706), NSE (P=0.017, OR=1.929), Cyfra21-1 (P<0.001, OR=8.261), D-D (P<0.001, OR=7.786), tumor size (P=0.001, OR=2.447), percentage of tumor solid (P<0.001, OR=3.324), and lesion location (P=0.001, OR=4.056) were risk factors for MLNM in patients (**Figure 2**). Subsequently, multivariate logistic regression analysis showed that Cyfra21-1 (P<0.001, OR=11.783), D-D (P<0.001, OR=10.072), tumor size (P=0.017, OR=2.602), percentage of tumor solid (P<0.001, OR=3.770), and lesion location (P=0.003, OR=5.366) were independent risk factors for MLNM (**Figure 3**).

Nomogram model development

To facilitate clinical application, a Nomogram visualization model was developed using the five independent risk factors identified by logistic regression. In this model, Cyfra21-1 was strongly correlated with MLNM; D-D, percentage of tumor solid, and lesion location were moderately related to MLNM; while tumor size

Table 2. Comparison of laboratory parameters between the two groups

Variables	Metastasis group (n=95)	Non-metastasis group (n=130)	t/Z	P
Hb (g/L)	136.53±19.64	141.05±19.07	-1.726	0.086
HCT (L/L)	0.48 [0.18, 0.76]	0.44 [0.19, 0.73]	0.553	0.581
PLT (10^9/L)	230.28±29.71	222.91±29.96	1.833	0.068
RDW (%)	12.66±0.48	12.62±0.71	0.426	0.671
Lym (10^9/L)	1.86±0.50	1.96±0.47	-1.478	0.141
Neu (10^9/L)	3.16±1.12	3.26±1.14	-0.639	0.524
NLR	1.70 [1.26, 2.16]	1.68 [1.32, 2.11]	0.070	0.945
SII	377.85 [267.68, 501.95]	367.20 [286.53, 477.04]	0.527	0.599
NSE (ng/mL)	12.36±1.98	11.93±1.84	1.686	0.093
CEA (ng/mL)	2.49 [1.83, 3.47]	2.69 [1.54, 3.66]	-0.175	0.862
Cyfra21-1 (ng/mL)	3.40±0.88	2.36±1.06	8.061	<0.001
D-D (mg/L)	0.44±0.19	0.29±0.10	6.835	<0.001
ALB (g/L)	39.11±4.93	38.76±4.05	0.556	0.579
ALP (U/L)	69.75±12.09	70.61±12.71	-0.513	0.609
FIB (g/L)	2.75±0.39	2.78±0.46	-0.604	0.546

Note: Hb, hemoglobin; HCT, hematocrit; PLT, platelet count; RDW, red cell distribution width; Lym, lymphocyte count; Neu, neutrophil count; NLR, neutrophil-to-lymphocyte ratio; SII, systemic immune-inflammation index; NSE, neuron-specific enolase; CEA, carcinoembryonic antigen; Cyfra21-1, cytokeratin fragment 21-1; D-D, D-dimer; ALB, albumin; ALP, alkaline phosphatase; FIB, fibrinogen.

Table 3. Comparison of pathological data between the two groups

Variables	Metastasis group (n=95)	Non-metastasis group (n=130)	χ^2	P
Tumor size				
≥2 cm	61	55	10.543	0.001
<2 cm	34	75		
Multiple nodules				
With	38	56	0.214	0.644
Without	57	74		
Percentage of tumor solid				
Ground-glass opacity	5	26	31.294	<0.001
Ground-glass opacity-dominant	5	32		
Solid-dominant	85	72		
Lesion location				
Central type	24	10	13.211	<0.001
Peripheral type	71	120		

was weakly linked to MLNM (**Figure 4A**). To determine the diagnostic performance of the model, ROC curves were used to further analyze the model. The results showed that the AUC of the model in diagnosing MLNM was 0.904, with a specificity of 73.85% and a sensitivity of 93.68% (**Figure 4B**). The calibration curve demonstrated that the model was well-calibrated, with a goodness-of-fit test showing a chi-square value of 9.5075 and a *p*-value of 0.3013, indicating no significant deviation from ideal calibration (**Figure 4C**). Subsequently, Delong's test revealed that the diagnostic performance of the Nomogram model was significantly superior to that of any individual risk factor (P<0.001, **Table 6**).

Discussion

This study identified key risk factors associated with mediastinal lymph node metastasis (MLNM) in early-stage NSCLC patients and con-

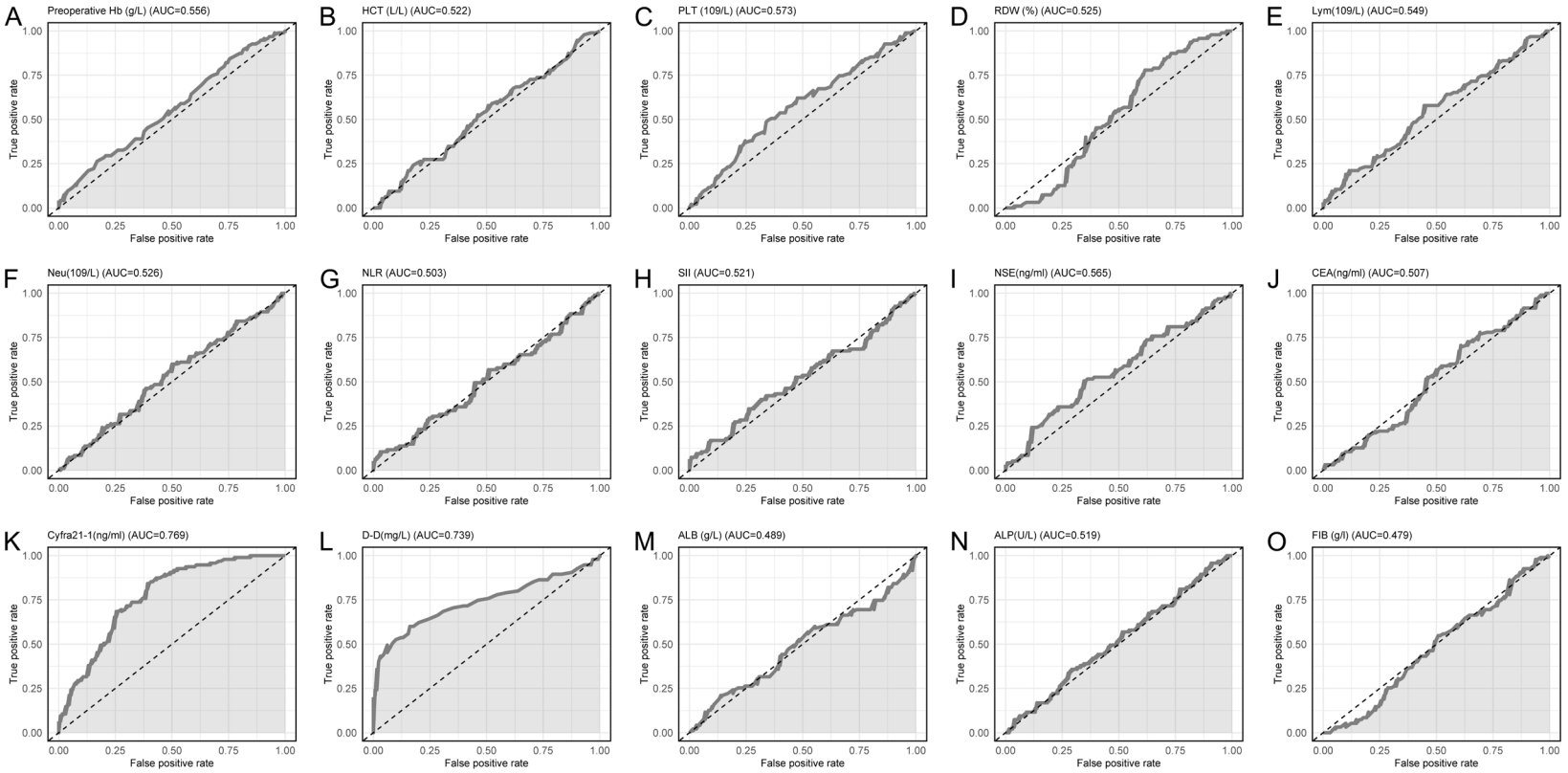

Figure 1. ROC curves of laboratory parameters in diagnosing mediastinal lymph node metastases in patients. A. Preoperative Hb: Reflects hemoglobin levels in patients before surgery. B. HCT: Represents the hematocrit level, indicating the proportion of red blood cells in blood. C. PLT: Platelet count, an indicator of blood clotting capacity. D. RDW: Red cell distribution width, reflecting the variation in red blood cell size. E. Lym: Lymphocyte count, associated with immune response. F. Neu: Neutrophil count, a marker of inflammation and infection. G. NLR: Neutrophil-to-lymphocyte ratio, often used as an inflammation marker. H. SII: Systemic immune-inflammation index, a combined indicator of immune and inflammatory status. I. NSE: Neuron-specific enolase, commonly related to neural and neuroendocrine cells. J. CEA: Carcinoembryonic antigen, a tumor marker often elevated in malignancies. K. Cyfra21-1: A cytokeratin fragment, used as a tumor marker in certain cancers. L. D-D: D-dimer, a marker of blood clot formation and breakdown. M. ALB: Albumin level, indicating nutritional and liver function status. N. ALP: Alkaline phosphatase, associated with liver and bone health. O. FIB: Fibrinogen, a key factor in blood clotting. Note: Hb, Hemoglobin; HCT, Hematocrit; PLT, Platelet count; RDW, Red cell distribution width; Lym, Lymphocyte count; Neu, Neutrophil count; NLR, Neutrophil-to-lymphocyte ratio; SII, Systemic immune-inflammation index; NSE, Neuron-specific enolase; CEA, Carcinoembryonic antigen; Cyfra21-1, Cytokeratin fragment 21-1; D-D, D-dimer; ALB, Albumin; ALP, Alkaline phosphatase; FIB, Fibrinogen.

Table 4. ROC parameters of laboratory parameters in diagnosing mediastinal lymph node metastases in patients

Marker	AUC	95% CI	Specificity	Sensitivity	Cut_off
Preoperative Hb (g/L)	0.556	0.480-0.633	83.08%	26.32%	122.5
HCT (L/L)	0.522	0.445-0.598	49.23%	57.89%	0.415
PLT (10^9/L)	0.573	0.497-0.649	65.38%	49.47%	232.5
RDW (%)	0.525	0.450-0.601	38.46%	77.89%	12.33
Lym (10^9/L)	0.549	0.473-0.626	55.38%	57.89%	1.895
Neu (10^9/L)	0.526	0.449-0.603	50.00%	60.00%	3.375
NLR	0.503	0.425-0.580	96.92%	10.53%	3.22
SII	0.521	0.443-0.599	73.85%	34.74%	460.448
NSE (ng/mL)	0.565	0.489-0.642	65.38%	50.53%	12.365
CEA (ng/mL)	0.507	0.431-0.583	39.23%	70.53%	3.205
Cyfra21-1 (ng/mL)	0.769	0.708-0.830	60.77%	84.21%	2.57
D-D (mg/L)	0.739	0.668-0.810	83.85%	60.00%	0.385
ALB (g/L)	0.489	0.411-0.568	86.15%	21.05%	34.55
ALP (U/L)	0.519	0.442-0.596	70.77%	35.79%	64.145
FIB (g/L)	0.479	0.403-0.554	11.54%	92.63%	2.145

Note: AUC, area under the curve; Hb, hemoglobin; HCT, hematocrit; PLT, platelet count; RDW, red cell distribution width; Lym, lymphocyte count; Neu, neutrophil count; NLR, neutrophil-to-lymphocyte ratio; SII, systemic immune-inflammation index; NSE, neuron-specific enolase; CEA, carcinoembryonic antigen; Cyfra21-1, cytokeratin fragment 21-1; D-D, D-dimer; ALB, albumin; ALP, alkaline phosphatase; FIB, fibrinogen.

structed a predictive Nomogram model. By analyzing clinical and laboratory data, it was found that Cyfra21-1, D-D, tumor size, percentage of tumor solid, and lesion location are significant predictors of MLNM. The constructed Nomogram exhibited excellent diagnostic accuracy, offering valuable insights into risk stratification and clinical decision-making for ES-NSCLC patients.

Cyfra21-1, a fragment of cytokeratin 19, is typically upregulated in epithelium-derived tumors and reflects tumor aggressiveness, including enhanced proliferation and invasiveness. Our findings demonstrate that elevated Cyfra21-1 is strongly associated with MLNM in early-stage NSCLC patients, which aligns with previous research. Mei et al. [14] reported a strong correlation between preoperative Cyfra21-1 and lymph node metastasis in esophageal cancer, suggesting its potential as an independent predictor. Similarly, Park et al. [15] identified Cyfra21-1 as a predictive marker for lymph node metastasis in thyroid cancer, further supporting its relevance across different tumor types. High Cyfra21-1 levels may contribute to tumor progression by promoting cell cycle progression and inhibiting apoptosis, thereby enhancing metastatic potential. These mechanisms are consistent with its established role in aggressive tumor biology, emphasizing its clinical value as a biomarker for MLNM risk assessment in NSCLC.

Our study reinforces the significant association between D-D levels and MLNM in early-stage NSCLC. D-D, a degradation product of fibrin, is a key marker of coagulation and fibrinolysis, processes that are often heightened in the context of tumor-related angiogenesis and cellular invasion. In the setting of NSCLC, increased D-D levels likely indicate a more aggressive tumor phenotype with a higher potential for metastasis through both hematogenous and lymphatic routes. Tumor cells secrete procoagulant factors and cytokines, activating the coagulation cascade and leading to microthrombosis formation, which facilitates metastasis. This mechanism aligns with our findings, suggesting that D-D levels can be used as an indirect marker of tumor invasiveness and metastatic potential. For instance, Song et al. [16] reported significantly higher plasma D-D levels in NSCLC patients compared to those with benign pulmonary nodules, emphasizing its role in malignancy-driven lymph node metastasis. Similarly, Chen et al. [17] identified preoperative D-D as a predictor of lymph node metas-

Table 5. Assignment table

Variables	Assignment content
Age	≥60 = 1, <60 = 0
Sex	Male = 1, female = 0
BMI	≥25 kg/m^2 = 1, <25 kg/m^2 = 0
Pack-year	≥400 = 1, <400 = 0
Lesion location	Left lung = 1, right lung = 0
History of hypertension	With = 0, without = 1
History of diabetes	With = 0, without = 1
History of cardiovascular disease	With = 0, without = 1
History of cerebrovascular disease	With = 0, without = 1
Tumor size	≥2 cm = 1, <2 cm = 0
Multiple nodules	Yes = 1, no = 0
Percentage of tumor solid	Ground-glass opacity = 1, ground-glass opacity-dominant = 2, and solid-dominant = 3
Lesion location	Central type = 1, peripheral type = 0
Preoperative Hb (g/L)	≤122.5 = 0, >122.5 = 1
HCT (L/L)	≤0.415 = 0, >0.415 = 1
PLT (10^9/L)	≤232.5 = 0, >232.5 = 1
RDW (%)	≤12.33 = 0, >12.33 = 1
Lym (10^9/L)	≤1.895 = 0, >1.895 = 1
Neu (10^9/L)	≤3.375 = 0, >3.375 = 1
NLR	≤3.22 = 0, >3.22 = 1
SII	≤460.448 = 0, >460.448 = 1
NSE (ng/mL)	≤12.365 = 0, >12.365 = 1
CEA (ng/mL)	≤3.205 = 0, >3.205 = 1
Cyfra21-1 (ng/mL)	≤2.57 = 0, >2.57 = 1
D-D (mg/L)	≤0.385 = 0, >0.385 = 1
ALB (g/L)	≤34.55 = 0, >34.55 = 1
ALP (U/L)	≤64.145 = 0, >64.145 = 1
FIB (g/L)	≤2.145 = 0, >2.145 = 1

Note: BMI, body mass index; Hb, hemoglobin; HCT, hematocrit; PLT, platelet count; RDW, red cell distribution width; Lym, lymphocyte count; Neu, neutrophil count; NLR, neutrophil-to-lymphocyte ratio; SII, systemic immune-inflammation index; NSE, neuron-specific enolase; CEA, carcinoembryonic antigen; Cyfra21-1, cytokeratin fragment 21-1; D-D, D-dimer; ALB, albumin; ALP, alkaline phosphatase; FIB, fibrinogen.

tasis in intrahepatic cholangiocarcinoma, further validating D-D's broader relevance across tumor types. Thus, D-D serves as a valuable biomarker for predicting MLNM in NSCLC patients, reinforcing its potential as part of a diagnostic panel for assessing metastasis risk.

Tumor size is an independent risk factor for MLNM in early-stage NSCLC, with larger tumors having a higher likelihood of metastasis due to their greater malignant potential and aggressiveness. Larger tumors contain more tumor cells, increasing the likelihood of invasion into the lymphatic system. Jin et al. [18] have shown that NSCLC patients with tumors ranging from 1-2 cm in diameter are at an increased risk of LNM in both the lung lobes and mediastinum. Larger tumors not only contain more tumor cells but also foster a microenvironment conducive to metastasis, with more active angiogenesis and lymph-angiogenesis facilitating tumor cell spread [19]. Additionally, larger tumors secrete higher levels of pro-invasive enzymes that degrade the extracellular matrix, enabling tissue infiltration and metastasis. Including tumor size in the predictive model enhances its ability to assess MLNM risk, providing a more comprehensive evaluation for early-stage NSCLC.

Tumors with a higher solid content are typically more malignant, exhibiting active growth, strong invasiveness, and a higher cell density, all of which make them more prone to LNM. Cho et al. [20] demonstrated that solid components larger than 1.5 cm were significantly associated with LNM, underscoring the relevance of solid tumor components in predicting

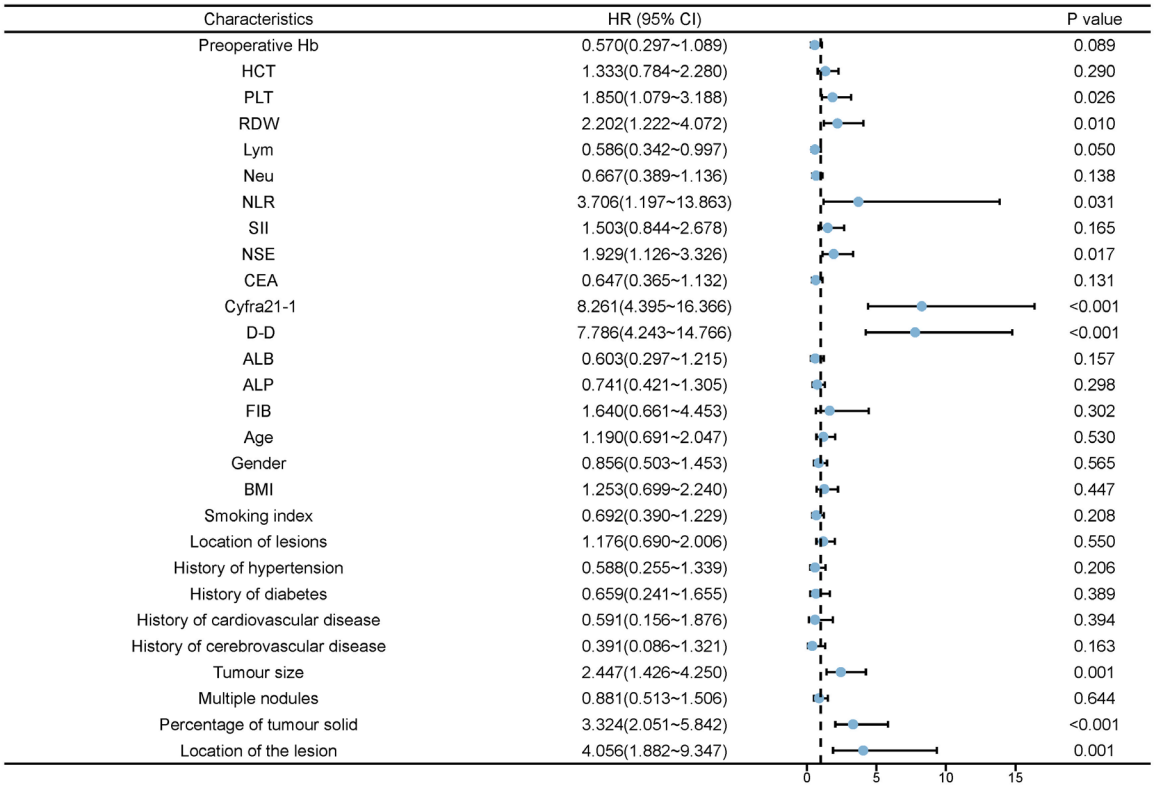

Figure 2. Univariate logistic regression analysis of risk factors for mediastinal lymph node metastasis. Note: BMI, body mass index; Hb, hemoglobin; HCT, hematocrit; PLT, platelet count; RDW, red cell distribution width; Lym, lymphocyte count; Neu, neutrophil count; NLR, neutrophil-to-lymphocyte ratio; SII, systemic immune-inflammation index; NSE, neuron-specific enolase; CEA, carcinoembryonic antigen; Cyfra21-1, cytokeratin fragment 21-1; D-D, D-dimer; ALB, albumin; ALP, alkaline phosphatase; FIB, fibrinogen.

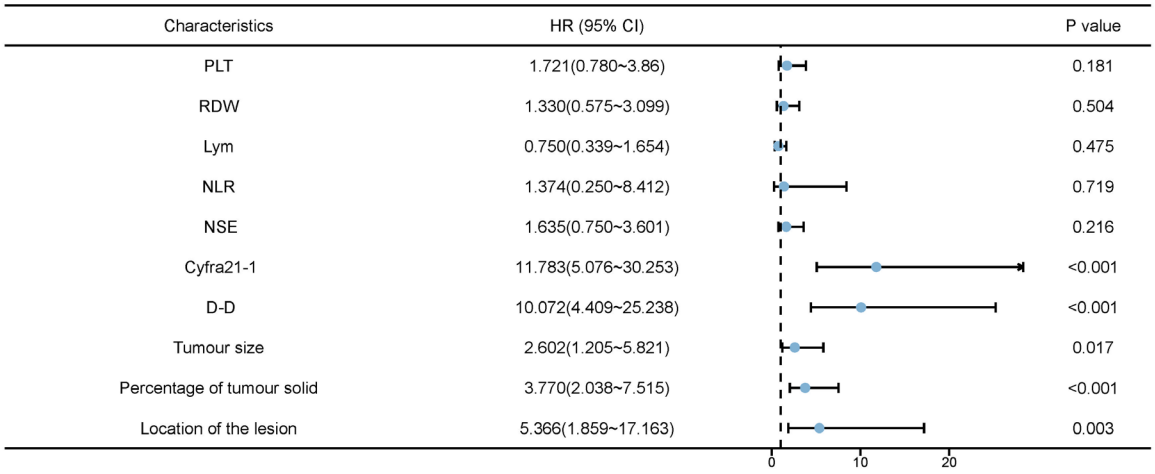

Figure 3. Multivariate logistic regression analysis of independent risk factors for mediastinal lymph node metastasis. Note: PLT, platelet count; RDW, red cell distribution width; Lym, lymphocyte count; NLR, neutrophil-to-lymphocyte ratio; NSE, neuron-specific enolase; Cyfra21-1, cytokeratin fragment 21-1; D-D, D-dimer.

metastasis. Similarly, Nakahashi [21] demonstrated that the solid-part tumor volume doubling time is a predictor of LNM in stage IA NSCLC. Liu et al. [22] identified increased consolidation and mass as risk factors for LNM in early-stage lung adenocarcinoma. These asso-

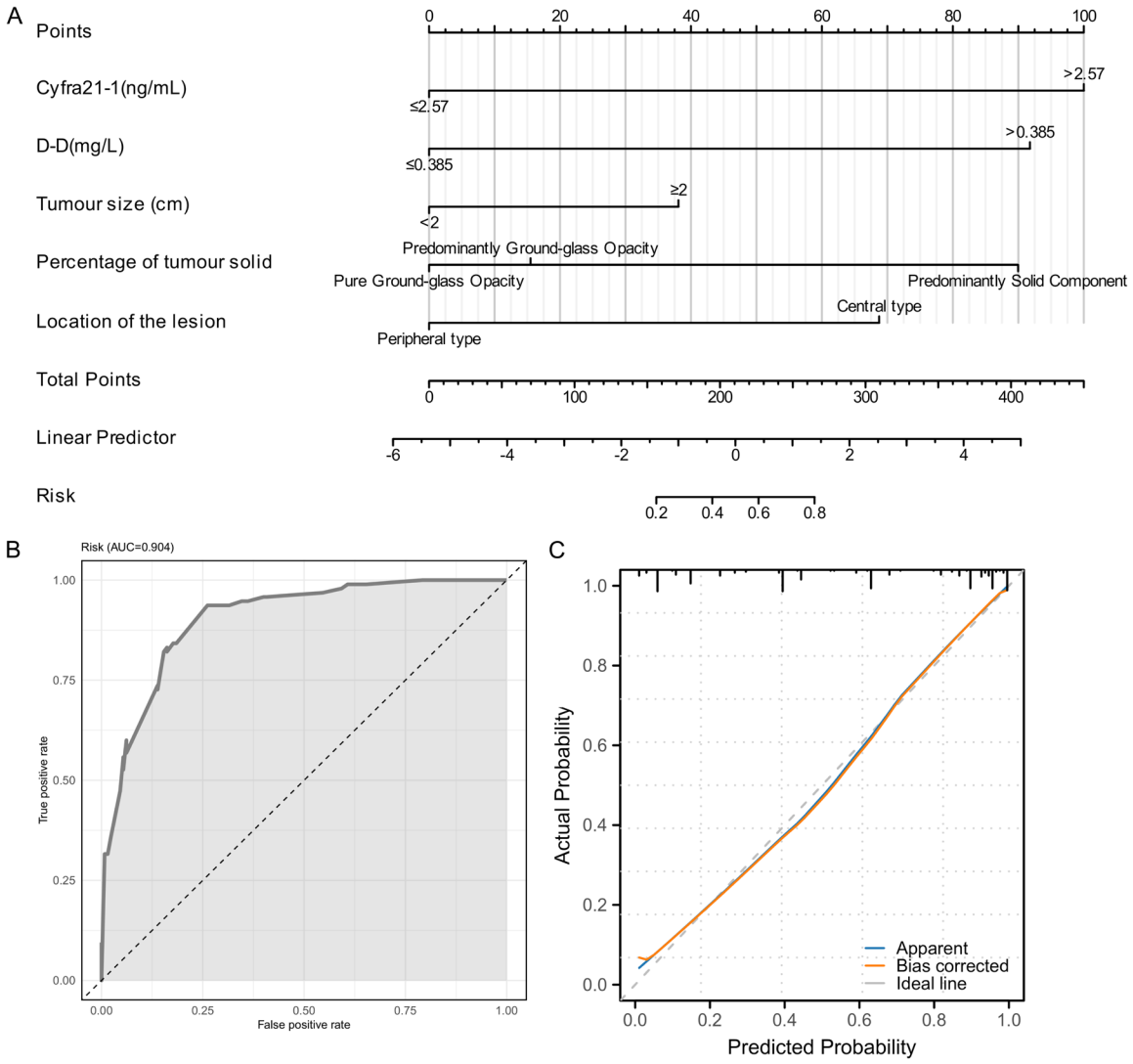

Figure 4. Nomogram model construction. A. A Nomogram model constructed based on the five risk factors. B. Effectiveness of the Nomogram model in diagnosing mediastinal lymph node metastases by ROC curves. C. Good model calibration indicated by Calibration curve. Note: ROC, receiver operating characteristic.

Table 6. Comparison of AUCs between the risk model and individual indicators by the Delong test

Marker 1	Marker 2	Z	P	AUC difference	95% CI
Cyfra21-1	Risk	-6.912	<0.001	-0.179	-0.230 - -0.128
D-D	Risk	-6.561	<0.001	-0.185	-0.240 - -0.130
Tumor size	Risk	-8.682	<0.001	-0.294	-0.361 - -0.228
Percentage of tumor solid	Risk	-8.253	<0.001	-0.235	-0.291 - -0.179
Lesion location	Risk	-11.197	<0.001	-0.316	-0.371 - -0.261

Note: AUC, area under the curve; Cyfra21-1, cytokeratin fragment 21-1; D-D, D-dimer.

ciations likely arise from the fact that solid tumors tend to have stronger intercellular adhesion and invasiveness, facilitating tumor cell infiltration and metastasis. Tumors with a high solid component may also possess enhanced anti-apoptotic capabilities and a higher proliferation rate, further driving their metastatic potential.

Tumor location also significantly affects lymphatic drainage and the likelihood of lymphatic spread [23]. Central tumors have a higher risk of metastasis due to their proximity to major lymphatic drainage channels. When located centrally in the lung, tumors have easier access to mediastinal lymph nodes, increasing the likelihood of metastasis [24]. Variations in tumor location can influence the tumor microenvironment and lymphatic drainage pathways, further affecting metastatic risk. These results not only enhance our understanding of the mechanisms underlying MLNM in NSCLC but also provide a reliable predictive tool for clinical practice, contributing to the early identification of high-risk patients and aiding in the optimization of treatment strategies to improve patient prognosis.

At the end of the study, we built a Nomogram model to improve the intuitiveness and accuracy of MLNM prediction. This model integrates multiple independent risk factors to provide individualized risk assessment, enabling clinicians to more accurately evaluate the LNM risk in patients and optimize treatment strategies. Our analysis showed that the Nomogram model performed well in the diagnosis of MLNM, with an AUC of 0.904, a specificity of 73.85%, and a sensitivity of 93.68%, demonstrating high predictive ability and clinical application value. Previously, Pak et al. [25] developed a decision tree model for predicting MLNM in NSCLC, achieving a sensitivity of 50% and a specificity of 99.28%. Similarly, Zhong et al. [26] utilized radiomics features to construct a model with an AUC of 0.972, sensitivity and specificity of 94.8% and 92%, respectively. These findings indicate that our Nomogram model achieves a balanced sensitivity and specificity, highlighting its potential for clinical application.

However, this study has several limitations. First, the small sample size and data sourced from a single medical center may limit the generalizability of the results, necessitating validation through multi-center, large-scale studies. Second, the retrospective design has selection bias, and prospective research are needed to provide stronger evidence. Third, the analysis was restricted to a limited set of biomarkers, potentially overlooking other important predictors. Future research should include a broader range of biomarkers to improve the model's accuracy. Lastly, the lack of long-term follow-up data prevents the evaluation of the model's effectiveness in predicting long-term prognosis. Hence, long-term follow-up data should be collected in future research to verify the model's effectiveness.

Conclusion

Conclusively, Cyfra21-1 and D-D are reliable biomarkers in predicting MLNM in ES-NSCLC patients. The constructed Nomogram model demonstrates high diagnostic efficiency and serves as a valuable tool for clinical evaluation and decision-making.

Acknowledgements

The Grants from the National Natural Science Foundation of China (82160577, 82073762); Key Program for Regional Joint Funds of Natural Science Foundation of Guangdong Province (No. 2023A1515030091, 2020B1515120094, 2021B1515120053); The Program for Science and Technology Plan of Guangzhou (No. 2023A03J0405; 202102010071). Guangdong Medical Science and Technology Research Fund (No. A2023205).

Disclosure of conflict of interest

None.

Address correspondence to: Yong Dong, Department of Orthopaedic Surgery, The Second Affiliated Hospital of Guangzhou Medical University, Guangzhou 510260, Guangdong, China. Tel: +86-0769-22896149; E-mail: dy1234560319@163.com; Kejing Tang, Department of Pulmonary and Critical Care Medicine, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou 510080, Guangdong, China. Tel: +86-020-87755766; E-mail: tangkj@mail.sysu.edu.cn

References

[1] Ju W, Zheng R, Zhang S, Zeng H, Sun K, Wang S, Chen R, Li L, Wei W and He J. Cancer statistics in Chinese older people, 2022: current burden, time trends, and comparisons with the US, Japan, and the Republic of Korea. Sci China Life Sci 2023; 66: 1079-1091.

[2] Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I and Jemal A. 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.

[3] Song KS, Nimse SB, Warkad SD, Kim JH, Kim HJ and Kim T. Detection and quantification of Tp53 and p53-Anti-p53 autoantibody immune complex: promising biomarkers in early stage lung cancer diagnosis. Biosensors (Basel) 2022; 12: 127.

[4] Sullivan FM, Mair FS, Anderson W, Armory P, Briggs A, Chew C, Dorward A, Haughney J, Hogarth F, Kendrick D, Littleford R, McConnachie A, McCowan C, McMeekin N, Patel M, Rauchhaus P, Ritchie L, Robertson C, Robertson J, Robles-Zurita J, Sarvesvaran J, Sewell H, Sproule M, Taylor T, Tello A, Treweek S, Vedhara K and Schembri S; Early Diagnosis of Lung Cancer Scotland (ECLS) Team. Earlier diagnosis of lung cancer in a randomised trial of an autoantibody blood test followed by imaging. Eur Respir J 2021; 57: 2000670.

[5] Zhu S, Yang C, Mei W, Kang L, Li T, Li J and Li L. Caregiver burden for informal caregivers of patients after surgical treatment of early-stage lung cancer. J Clin Nurs 2023; 32: 859-871.

[6] Tian Y, He Y, Li X and 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: 2548-2566.

[7] Biswas A, Jantz MA and Mehta HJ. Pretreatment identification of micro-metastasis in mediastinal lymph node by endobronchial ultrasound-guided transbronchial needle aspiration for early-stage non-small cell lung cancer-is it time yet? J Thorac Dis 2019; 11: 4096-4100.

[8] Woodard GA, Jones KD and Jablons DM. Lung cancer staging and prognosis. Cancer Treat Res 2016; 170: 47-75.

[9] Huang X, Wang J, Chen Q and Jiang J. Mediastinal lymph node dissection versus mediastinal lymph node sampling for early stage non-small cell lung cancer: a systematic review and meta-analysis. PLoS One 2014; 9: e109979.

[10] Chen Q, Li W, Cai N, Chen W, Zhao X and Huang X. Comparison of postoperative complications in mediastinal lymph node dissection versus mediastinal lymph node sampling for early stage non-small cell lung cancer: protocol for a systematic review and meta-analysis. PLoS One 2024; 19: e0298368.

[11] Andersson S, Ilonen I, Järvinen T, Rauma V, Räsänen J and Salo J. Surgically treated unsuspected N2-positive NSCLC: role of extent and location of lymph node metastasis. Clin Lung Cancer 2018; 19: 418-425.

[12] Feng SH and Yang ST. The new 8th TNM staging system of lung cancer and its potential imaging interpretation pitfalls and limitations with CT image demonstrations. Diagn Interv Radiol 2019; 25: 270-279.

[13] Sayan M, Satir Turk M, Celik A, Cuneyt Kurul I and Irfan Tastepe A. Surgical outcomes of early-stage small-cell lung cancer: single-center experience. Asian Cardiovasc Thorac Ann 2019; 27: 187-191.

[14] Mei X, Zhu X, Zuo L, Wu H, Guo M and Liu C. Predictive significance of CYFRA21-1, squamous cell carcinoma antigen and carcinoembryonic antigen for lymph node metastasis in patients with esophageal squamous cancer. Int J Biol Markers 2019; 34: 200-204.

[15] Park J, An S, Kim K, Kim JS, Jung CK and Bae JS. Diagnostic utilities of washout CYFRA 21-1 combined with washout thyroglobulin for metastatic lymph nodes in thyroid cancer: a prospective study. Sci Rep 2024; 14: 7599.

[16] Song X, Wang F, Shen H, Li J, Hu T, Yang Z, Zhou Y and Shi Q. Correlation between plasma D-dimer count and features of non-small cell lung cancer. Zhongguo Fei Ai Za Zhi 2019; 22: 151-156.

[17] Chen Q, Zheng Y, Zhao H, Cai J, Wang L, Zhao J, Bi X, Li Z, Huang Z, Zhang Y, Zhou J and Wu J. The combination of preoperative D-dimer and CA19-9 predicts lymph node metastasis and survival in intrahepatic cholangiocarcinoma patients after curative resection. Ann Transl Med 2020; 8: 192.

[18] Jin T, He Z, Li Z, Tang J, Xu J, Wu W and Chen L. Risk factors and sampling range evaluation of lymph node metastasis for non-small cell lung cancer with diameter ≤2 cm. Zhongguo Fei Ai Za Zhi 2023; 26: 507-514.

[19] Zhang W, Mu G, Huang J, Bian C, Wang H, Gu Y, Xia Y, Chen L, Yuan M and Wang J. Lymph node metastasis and its risk factors in T1 lung adenocarcinoma. Thorac Cancer 2023; 14: 2993-3000.

[20] Cho JY, Leem CS, Kim Y, Kim ES, Lee SH, Lee YJ, Park JS, Cho YJ, Lee JH, Lee CT and Yoon HI. Solid part size is an important predictor of nodal metastasis in lung cancer with a subsolid tumor. BMC Pulm Med 2018; 18: 151.

[21] Nakahashi K, Shiono S, Nakatsuka M and Endo M. Prediction of lymph node metastasis of clinical stage IA non-small cell lung cancer based on the tumor volume doubling time. Surg Today 2022; 52: 1063-1071.

[22] Liu M, Mu J, Song F, Liu X, Jing W and Lv F. Growth characteristics of early-stage (IA) lung adenocarcinoma and its value in predicting lymph node metastasis. Cancer Imaging 2023; 23: 115.

[23] Fang C, Xiang Y and Han W. Preoperative risk factors of lymph node metastasis in clinical N0 lung adenocarcinoma of 3 cm or less in diameter. BMC Surg 2022; 22: 153.

[24] Tao H, Cai Y, Shi L, Tang J, Liu Z, Wang Z, Bai L and Liu Z. Analysis of clinical characteristics

and prognosis of patients with anaplastic lymphoma kinase-positive and surgically resected lung adenocarcinoma. Thorac Cancer 2017; 8: 8-15.

[25] Pak K, Kim K, Kim MH, Eom JS, Lee MK, Cho JS, Kim YS, Kim BS, Kim SJ and Kim IJ. A decision tree model for predicting mediastinal lymph node metastasis in non-small cell lung cancer with F-18 FDG PET/CT. PLoS One 2018; 13: e0193403.

[26] Zhong Y, Yuan M, Zhang T, Zhang YD, Li H and Yu TF. Radiomics approach to prediction of occult mediastinal lymph node metastasis of lung adenocarcinoma. AJR Am J Roentgenol 2018; 211: 109-113.