

# Development and Validation of a Novel Nomogram for Predicting Tumor-Distant-Metastasis in Patients with Early T1-2 Stage Lung Adenocarcinoma


This article was published in the following Dove Press journal:  
*Therapeutics and Clinical Risk Management*

WeiGuo Gu<sup>1,2</sup>

MingBin Hu<sup>1,2</sup>

Weijia Wang<sup>1</sup>

Chao Shi<sup>3</sup>

JinHong Mei<sup>1</sup> 

<sup>1</sup>Department of Pathology, The First Affiliated Hospital of Nanchang University, Nanchang, Jiangxi, People's Republic of China; <sup>2</sup>First Clinical Medical College, Nanchang University, Nanchang, Jiangxi, People's Republic of China;

<sup>3</sup>Department of Oncology, The First Affiliated Hospital of Nanchang University, Nanchang, Jiangxi, People's Republic of China

**Background:** Distant metastasis in early T1-2 (diameter≤5 cm) stage lung adenocarcinoma (ET-LUAD) patients largely affect treatment strategies in clinical practice. However, the associated mechanism remains unclear and related studies is less. This study aimed to establish and validate a novel nomogram to predict the risk of distant metastasis in ET-LUAD.

**Methods:** A total of 258 patients diagnosed with ET-LUAD and not receiving any treatment were recruited into this study. The patients were randomly divided into a training cohort and validation cohort in a ratio of 1:2. Univariate and multivariate logistic regression analysis was used to select the most significant predictive risk factors associated with distant metastasis in the training cohort. The established nomogram was validated by the consistency index (C-index), calibration curve, and decision curve analysis (DCA).

**Results:** There were 124 patients with confirmed distant metastasis and 134 patients with non-distant metastases ET-LUAD were enrolled in the study. Multivariate logistic hazards regression analysis identified independent risk factors associated with distant metastasis to include platelet-to-lymphocyte ratios (PLR), lactate dehydrogenase (LDH), neural-specific enolase (NSE), carcinoembryonic antigen (CEA) and cytokeratin 19 fragments (Cyfra211), which were included in the establishment of the nomogram. The nomogram achieved a high consistency (C-index=0.792), good calibration, and high clinical application value in the validation cohort.

**Conclusion:** The established nomogram can be used to predict distant metastasis in high-risk ET-LUAD nonmetastasis patients and can also be used by doctors to guide preventive and individualized treatment for ET-LUAD patients.

**Keywords:** lung adenocarcinoma, distant metastasis, nomogram

Correspondence: JinHong Mei  
Department of Pathology, The First Affiliated Hospital of Nanchang University, No. 17 Yong Wai Street, Nanchang, Jiangxi 330006, People's Republic of China  
Email mjhdoctor@126.com

Chao Shi  
Department of Oncology, The First Affiliated Hospital of Nanchang University, No. 17 Yong Wai Street, Nanchang, Jiangxi 330006, People's Republic of China  
Email 43286143@qq.com

## Introduction

Lung cancer is one of the leading causes of cancer-related morbidity and mortality, worldwide.<sup>1</sup> Non-Small Cell Lung Cancer (NSCLC) accounts for over 85% of all lung cancer, and lung adenocarcinoma (LUAD) is the most common histologic type of NSCLC and accounts for 40%-50%.<sup>2</sup> With the recent developments in diagnostic imaging, more NSCLC patients are diagnosed in the early stages, however, due to lack of specific symptoms in early NSCLC, nearly 50%-60% of the patients are finding out distant metastasis at their first diagnosis in east China.<sup>3</sup> In advanced NSCLC, the most common sites of metastasis are bone, brain, liver, and the adrenal gland.<sup>3-5</sup> The 5-year survival rate of advanced NSCLC is less than 4%, and nearly 90% of the patients die from distant metastasis.<sup>6,7</sup>

The serum laboratory indicators, clinical-pathological characteristics or inflammation-related factors may be useful biomarkers to predict distant metastasis in NSCLC and the effect of treatment.<sup>8–10</sup> Darlix et al<sup>11</sup> showed that serum neural-specific enolase (NSE) level and *HER2* gene status were independent prognostic factors in breast cancer patients with brain metastasis. Another study also showed that the serum CEA expression in lung cancer patients with metastasis was higher than non-metastasis patients. Besides, subgroup analysis revealed that serum CEA level is related to the location and number of metastasis, and it is considered an important evaluation index of distant metastasis in lung cancer.<sup>12</sup> Furthermore, numerous studies also show that inflammation-related factors play an important role in promoting distant metastasis in lung cancer.<sup>13,14</sup> The systemic immune inflammation index (SII) composed of lymphocytes, neutrophils, and platelets was an independent prognostic factor for LUAD with brain metastasis and EGFR gene mutation.<sup>15</sup>

In the real world clinical practice, many early T1-2 stages of LUAD patients (ET-LUAD) already have distant metastasis when they are first diagnosed. Even with the use, the modern advanced imaging technology, many micrometastasis sites are still difficult to detect. Moreover, the specific mechanism of ET-LUAD in distant metastasis is complex and unknown, and few studies have reported which the high-risk factors of a clinical biomarker can predict it. Therefore, based on the relevant clinical indicators of serum laboratory indicators and clinical-pathological features there are important reference values to judge the possibility of distant metastasis. In this study, correlation analysis between the clinicopathological characteristics and serum laboratory indicators of NSCLC patients with or without distant metastasis was performed and used to establish a nomogram to predict distant metastasis, and further analyze and verify the key clinical factors for identifying high-risk patients for clinical doctors to provide early treatment intervention.

## Materials and Methods

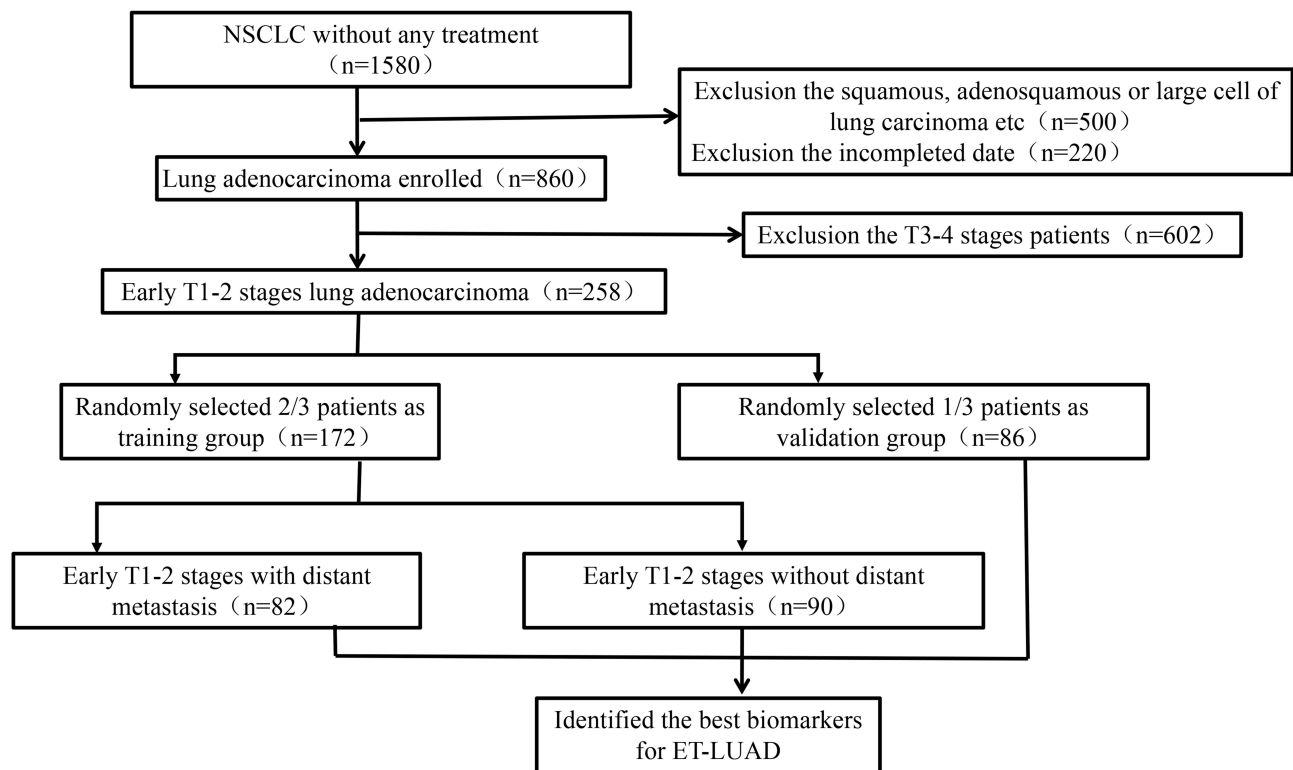
### Patients and Data Collection

We retrospectively identified patients who underwent histologically confirmed LUAD and without any pre-treatment, between January 2015 and December 2019. This study was approved by the Ethics Committee of First Affiliated Hospital of Nanchang University, Nanchang, China, ethics number is (2020) medicine

research (78), all patients provided written informed consent and compliance with the declaration of Helsinki. The patients' major baseline characteristics including, age, sex, *EGFR* mutation, serum clinical laboratory indicators, and stage of LUAD (American Joint Committee on Cancer Staging Manual, 8th edition) at initial diagnosis were collected. The inclusion criteria were as follows: (1) all patients were diagnosed with stage T1-2 LUAD by histological or cytological analysis, (2) had not received any pre-treatment, included chemoradiotherapy or surgical operation, and all the serum laboratory indicators were acquired in their preliminary diagnosis, (3) the size of distant metastasis lesion could be measured by computer tomography (CT) or magnetic resonance imaging (MRI), (4) patients free from infection and use of drugs affecting the hematopoietic function of the bone marrow before treatment and without hematological diseases. The exclusion criteria were as follows: (1) patients with previous blood system diseases (leukemia or lymphoma), infection before treatment, or antibiotic use, (2) incomplete clinical data of the patient, (3) squamous cell, adenosquamous, large cell or small cell lung cancer, or LUAD combined with other malignancies.

### Establishment and Validation of the Nomogram

A total of 258 patients in early T1-2 stages of LUAD were enrolled in this study, patients with N3 stage positive and/or M stage positive as the distant metastasis group (n=124) and those with N0-2 and M stage negative patients as the non-metastasis group (n=134). The patients were randomly divided in a ratio of 1:2 into the training cohort group (n=172) and validation cohort group (n=86) by the R package "rms" (Figure 1). In the training cohort, univariate and multivariate logistic regression analysis was used to identify the independent risk factors and those variables (only the factors with a *P*-value <0.05) were used to establish nomogram model. In the validation group, the consistency index (C-index) was used to evaluate the model's prediction accuracy that ranged from 0.5 (no discriminative) to 1 (perfect discrimination), and a high C-index meant a good predicted model. The calibration curve was used to evaluate the prediction compliance and the decision curve analysis (DCA) was used to evaluate the model's clinical application value and threshold probabilities.



**Figure 1** Flow chart of the research.

## Statistical Analysis

Continuous variables were transformed into categorical variables and Chi-square or Fisher's exact test used for analysis. Univariate and multivariate (using the forward-LR method) logistic regression analysis was used to calculate the odds ratio (OR) and 95% confidence interval (CI) to identify the independent risk factors. Correlation analysis was carried out by the Spearman correlation coefficient for independent risk factors. Statistical analysis was performed using IBM SPSS 22.0 software (SPSS Inc., Chicago, IL, USA) and the figures were used by GraphPad Prism version 8.0 software (Inc., La Jolla, CA, USA), and nomogram formulation were performed in R statistical software version 4.0.0 (<http://www.R-project.org>). The cut-off points of peripheral serum routine blood test were used as mean value, serum LDH and tumor biomarkers of cut-off points were used relevant assay kits,<sup>9</sup> and all those factors divided into high and low groups for statistical analysis.  $P < 0.05$  was considered to be statically significant.

## Results

### Correlation Between Distant Metastasis and Clinicopathological Characteristics

The clinicopathological characteristics of the 258 patients in ET-LUAD (124 with distant metastasis and 134 non-distant

metastasis) are summarized in Table 1. Serum inflammatory markers included white blood cell (WBC), lymphocyte count, neutrophils count, monocyte count, platelet (PLT), neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), monocyte-to-lymphocyte ratio (MLR), neutrophil-to-monocyte ratio (NMR) and lactate dehydrogenase (LDH). Serum tumor biomarkers included CEA, NSE, carbohydrate antigen 125 (CA125), and cytokeratin 19 fragments (Cyfra211). The median age of the patients was 59 years (range 23–82). Among the 258 patients, 132 were males and 126 females. Tumors of the N stage were analyzed (103 with N0, 20 with N1, 107 with N2 and 28 with N3). Among 124 patients with distant metastasis, there were 73 with multiple metastasis (metastasis sites  $\geq 3$  or number of metastatic tumors was more than 5),<sup>3,4,16</sup> 53 with brain metastasis, 81 with bone metastasis, 9 with liver metastasis, 3 with adrenal gland metastasis, 24 with pleural dissemination metastasis. There were 97 EGFR mutations, among which 49 were EGFR wild and 112 unknown.

### Univariate and Multivariate Logistic Analyses of ET-LUAD in the Training Cohort

The patients were randomly divided in a ratio of 1:2 into the training cohort group (n=172) and validation cohort

**Table I** Demographic and Clinical-Pathological Characteristics of the Training Cohort and Validation Cohort

Characteristic	Groups	All Cohort (n=258, %)	Training Cohort (n=172, %)	Validation Cohort (n=86, %)
Gender	Male Female	132(51.2) 126(48.8)	94(54.7) 78(45.3)	38(44.2) 48(55.8)
Age (years)	<60 ≥60	121(46.9) 137(53.1)	85(49.4) 87(50.6)	36(41.9) 50(58.1)
Smoking history	No Yes	173(67.1) 85(32.9)	107(62.2) 65(37.8)	66(76.7) 20(23.3)
T stages	T1a T1b T1c T2a T2b	5(2) 59(22.9) 56(21.7) 71(27.5) 67(25.9)	4(2.3) 37(21.5) 39(22.7) 50(29.1) 42(24.4)	1(1.2) 22(25.6) 17(19.8) 21(24.4) 25(29.1)
N stages	N0 N1 N2 N3	103(39.9) 20(7.7) 107(41.5) 28(10.9)	70(40.7) 16(9.3) 64(37.2) 22(12.8)	33(38.4) 4(4.7) 43(50) 6(6.9)
TNM stages	I II III IV	55(21.3) 26(10.1) 61(23.6) 116(45)	40(23.3) 20(11.6) 38(22.1) 74(43)	15(17.5) 6(7) 23(26.7) 42(48.8)
EGFR mutation	Mutation Wild Unknown	97(37.6) 49(19) 112(43.4)	64(37.2) 33(19.2) 75(43.6)	33(38.4) 16(18.6) 37(43)
WBC( $10^9/L$ ) <sup>a</sup>	≤ 6.6 >6.6	159(61.6) 99(38.4)	108(62.8) 64(37.2)	51(59.3) 35(40.7)
RBC( $10^{12}/L$ ) <sup>a</sup>	≤ 4.4 > 4.4	129(50) 129(50)	92(53.5) 80(46.5)	37(43) 49(57)
HB(g/L) <sup>a</sup>	≤ 130 >130	126(48.8) 132(51.2)	87(50.6) 85(49.4)	39(45.3) 47(54.7)
PLT( $10^9/L$ ) <sup>a</sup>	≤ 242 >242	106(41.1) 152(58.9)	69(40.1) 103(59.9)	37(43) 49(57)
Lymphocyte( $10^9/L$ ) <sup>a</sup>	≤1.6 >1.6	142(55) 116(45)	96(55.8) 76(44.2)	46(53.5) 40(46.5)
Neutrophils( $10^9/L$ ) <sup>a</sup>	≤4.4 >4.4	152(58.9) 106(41.1)	103(59.9) 69(40.1)	49(57) 37(43)
Monocyte( $10^9/L$ ) <sup>a</sup>	≤0.47 >0.47	168(65.1) 90(34.9)	115(66.9) 57(33.1)	53(61.6) 33(38.4)
NLR <sup>a</sup>	≤3.3 >3.3	168(65.1) 90(34.9)	113(65.7) 59(34.3)	55(63.9) 31(36.1)
PLR <sup>a</sup>	≤172.6 >172.6	159(61.6) 99(38.4)	106(61.6) 66(38.4)	53(61.6) 33(38.4)
MLR <sup>a</sup>	≤0.34 >0.34	194(75.2) 64(24.8)	119(46.1) 53(53.9)	75(87.2) 11(12.8)

(Continued)

Table 1 (Continued).

Characteristic	Groups	All Cohort (n=258, %)	Training Cohort (n=172, %)	Validation Cohort (n=86, %)
NMR <sup>a</sup>	≤11.6 >11.6	181(70.2) 77(29.8)	119(69.2) 53(30.8)	62(72.1) 24(27.9)
LDH (U/L) <sup>b</sup>	≤250 >250	165(63.9) 93(36.1)	111(64.5) 61(35.5)	54(62.8) 32(37.2)
CEA (ng/mL) <sup>b</sup>	≤6.5 >6.5	140(54.3) 118(45.7)	97(56.4) 75(43.6)	43(50) 43(50)
CA-125 (U/mL) <sup>b</sup>	≤35 >35	158(61.2) 100(38.8)	102(61) 70(39)	56(65.1) 30(34.9)
NSE (ng/mL) <sup>b</sup>	≤16.3 >16.3	132(51.2) 126(48.8)	89(51.7) 83(48.3)	43(50) 43(50)
Cyfra211 (ng/mL) <sup>b</sup>	≤3.3 >3.3	141(54.7) 117(45.3)	90(52.3) 82(47.7)	51(59.3) 35(40.7)
Distant metastasis	Brain Bone Liver Adrenal gland Pleural dissemination	53(20.6) 81(31.4) 9(3.5) 3(1.2) 24(9.3)	41(23.8) 52(30.2) 6(3.5) 2(1.3) 14(8.2)	12(14) 29(33.7) 3(3.5) 1(1.2) 10(11.6)

**Notes:** <sup>a</sup>The cut-off points was used mean value. <sup>b</sup>The cut-off points was used relevant assay kits, and all those factors divided into high and low groups for statistical analysis.  
**Abbreviations:** EGFR, epidermal growth factor receptor; WBC, white blood cell; HB, hemoglobin; RBC, red blood cell; PLT, platelet; PLR, platelet-to-lymphocyte ratio; NLR, neutrophil-to-lymphocyte ratio; MLR, monocyte-to-lymphocyte ratio; NMR, neutrophil-to-monocyte ratio; LDH, lactate dehydrogenase; CEA, carcinoembryonic antigen; NSE, neural-specific enolase; CA125, carbohydrate antigen 125; Cyfra211, cytokeratin 19 fragment.

group (n=86). In the training cohort, there are 75 patients with distant metastasis and 97 without distant metastasis. The univariate analysis showed a significant relationship between N stage, lymphocyte count, neutrophils count, PLR, NLR, LDH, CEA, NSE, CA125, and Cyfra211 and distant metastasis in ET-LUAD ( $P<0.05$ ) (Table 2). All significant factors in the univariate analysis were included in multivariate logistic regression analysis. The result showed that PLR $>172.6$ , serum LDH $>250$  U/L, CEA $>6.5$  ng/mL, NSE $>16.3$  ng/mL, and Cyfra211 $>3.3$  ng/mL were independent risk factors associated with distant metastasis in ET-LUAD ( $P<0.05$ ) (Table 3).

## The Relationship Between Independent Risk Factors with Distant Metastasis in ET-LUAD

All clinical independent risk factors in multivariate logistic regression analysis were included in Spearman correlation analysis in the training cohort, and also included the lymphocyte count. The results showed that serum lymphocyte count ( $r=-0.169$ ,  $P=0.027$ ), PLR ( $r=0.222$ ,  $P=0.004$ ), LDH( $r=0.451$ ,  $P<0.001$ ),

CEA ( $r=0.409$ ,  $P<0.001$ ), NSE ( $r=0.394$ ,  $P<0.001$ ) and Cyfra211 ( $r=0.381$ ,  $P<0.001$ ) were significantly associated with distant metastasis in ET-LUAD. The higher the serum lymphocyte count, the less the possibility of distant metastasis, however, the higher the PLR ratio, LDH, CEA, NSE, and Cyfra211, the higher risk to develop distant metastasis (Table 4 and Figure 2).

## Construction of the Nomogram

The univariate and multivariate logistic hazards regression analysis identified independent risk factors in the training cohort and incorporated those five variables to establish the prediction of nomogram model, including PLR, LDH, CEA, NSE, and Cyfra211. Each variable is represented by a horizontal line, and the patient's information is marked on the coordinates. The regression coefficient of each predictive variable corresponds to the score value in the range of 0–100. The higher the weight of the variable, the more the scores. Furthermore, the total score of the five variables was obtained by adding the corresponding scores, and the metastasis possibility obtained for ET-LUAD (Figure 3).

**Table 2** Univariate Logistic Proportional Hazards Regression Analysis in the Training Cohort

Characteristic		Training Cohort		OR (95% CI)	P value
		T1-2 Stages without Metastasis	T1-2 Stages with Metastasis		
Gender	Male	52	42	Ref 0.908(0.495–1.664)	0.755
	Female	45	33		
Age (years)	<60	47	38	Ref 0.915(0.501–1.673)	0.773
	≥60	50	37		
Smoking history	No	61	46	Ref 1.068(0.574–1.988)	0.835
	Yes	36	29		
N stages	N0	47	23	Ref 0.472(0.122–1.821) 2.316(1.15–4.663) 4.379(1.569–12.222)	0.003
	N1	13	3		0.275
	N2	30	34		0.019
	N3	7	15		0.005
EGFR mutation	Mutation	37	27	Ref 0.775(0.333–1.803) 0.93(0.41–2.11)	0.802
	Wild	17	16		0.555
	Unknown	40	35		0.862
WBC( $10^9/L$ )	≤6.6	67	41	Ref 1.852(0.99–3.464)	0.054
	>6.6	30	34		
RBC( $10^{12}/L$ )	≤4.4	53	39	Ref 1.112(0.608–2.034)	0.731
	>4.4	44	36		
HB(g/L)	≤130	55	32	Ref 1.76(0.957–3.235)	0.069
	>130	42	43		
PLT( $10^9/L$ )	≤242	41	28	Ref 1.229(0.663–2.279)	0.513
	>242	56	47		
Lymphocyte( $10^9/L$ )	≤1.6	47	49	Ref 0.499(0.268–0.927)	0.028
	>1.6	50	26		
Neutrophils( $10^9/L$ )	≤4.4	67	36	Ref 2.419(1.295–4.52)	0.006
	>4.4	30	39		
Monocyte( $10^9/L$ )	≤0.47	67	48	Ref 1.256(0.663–2.379)	0.484
	>0.47	30	27		
NLR	≤3.3	72	41	Ref 2.388(1.255–4.544)	0.008
	>3.3	25	34		
PLR	≤172.6	69	37	Ref 2.531(1.347–4.755)	0.004
	>172.6	28	38		
MLR	≤0.34	73	46	Ref 1.918(0.996–3.691)	0.051
	>0.34	24	29		
NMR	≤11.6	69	50	Ref 1.232(0.643–2.362)	0.529
	>11.6	28	25		
LDH (U/L)	≤250	81	30	Ref 7.594(3.742–15.411)	<0.001
	>250	16	45		
CEA (ng/mL)	≤6.5	72	25	Ref 5.76(2.973–11.16)	<0.001
	>6.5	25	50		

(Continued)



**Table 2** (Continued).

Characteristic		Training Cohort		OR (95% CI)	P value
		T1-2 Stages without Metastasis	T1-2 Stages with Metastasis		
CA-125 (U/mL)	≤35 >35	72 25	30 45	Ref 4.32(2.259–8.262)	<0.001
NSE (ng/mL)	≤16.3 >16.3	67 30	22 53	Ref 5.38(2.787–10.385)	<0.001
Cyfra211 (ng/mL)	≤3.3 >3.3	67 30	23 52	Ref 5.049(2.628–9.7)	<0.001

**Abbreviations:** EGFR, epidermal growth factor receptor; WBC, white blood cell; HB, hemoglobin; RBC, red blood cell; PLT, platelet; PLR, platelet-to-lymphocyte ratio; NLR, neutrophil-to-lymphocyte ratio; MLR, monocyte-to-lymphocyte ratio; NMR, neutrophil-to-monocyte ratio; LDH, lactate dehydrogenase; CEA, carcinoembryonic antigen; NSE, neural-specific enolase; CA-125, carbohydrate antigen 125; Cyfra211, cytokeratin 19 fragment; Ref, reference.

**Table 3** Multivariate Logistic Proportional Hazards Regression Analysis in the Training Cohort

Characteristic	Groups	OR (95% CI)	P value
LDH (U/L)	≤250 >250	Ref 3.494 (1.515–8.058)	0.003
PLR	≤172.6 >172.6	Ref 2.905 (1.27–6.643)	0.012
CEA (ng/mL)	≤6.5 >6.5	Ref 3.506 (1.53–8.033)	0.003
NSE (ng/mL)	≤16.3 >16.3	Ref 2.404 (1.064–5.433)	0.035
Cyfra211 (ng/mL)	≤3.3 >3.3	Ref 2.527 (1.131–5.643)	0.024

**Abbreviations:** EGFR, epidermal growth factor receptor; PLR, platelet-to-lymphocyte ratio; LDH, lactate dehydrogenase; CEA, carcinoembryonic antigen; NSE, neural-specific enolase; Cyfra211, cytokeratin 19 fragment; Ref, reference.

## Validation of the Nomogram

The C-index, calibration curve and DCA were used to evaluate the accuracy of the nomogram in the training cohort and validation cohort. The C-index of the established nomogram was 0.792, meaning that the prediction model had high accuracy in distinguishing patients with distant metastasis (0.71–0.90). Moreover, we continue to observe the calibration curve in [Figure 4](#) and find that the

regression fitting curve was very close to the standard curve. The results also showed that there was no significant difference between the two curves ( $P = 0.404$ ), meaning that the prediction model had a high degree of calibration, and the prediction results of the model were very close to the actual outcome ([Figure 4](#)).

The DCA was conducted to further evaluate the clinicopathological data in the validation cohort ([Figure 5](#)). The DCA showed that the clinical value of the model presented more net benefits at 0%–81% threshold probability. Therefore, patients with high-risk metastasis possibility receiving treatment had a more net benefit than either the treat all patients or treat none patients. These results confirmed the prediction efficacy of the nomogram.

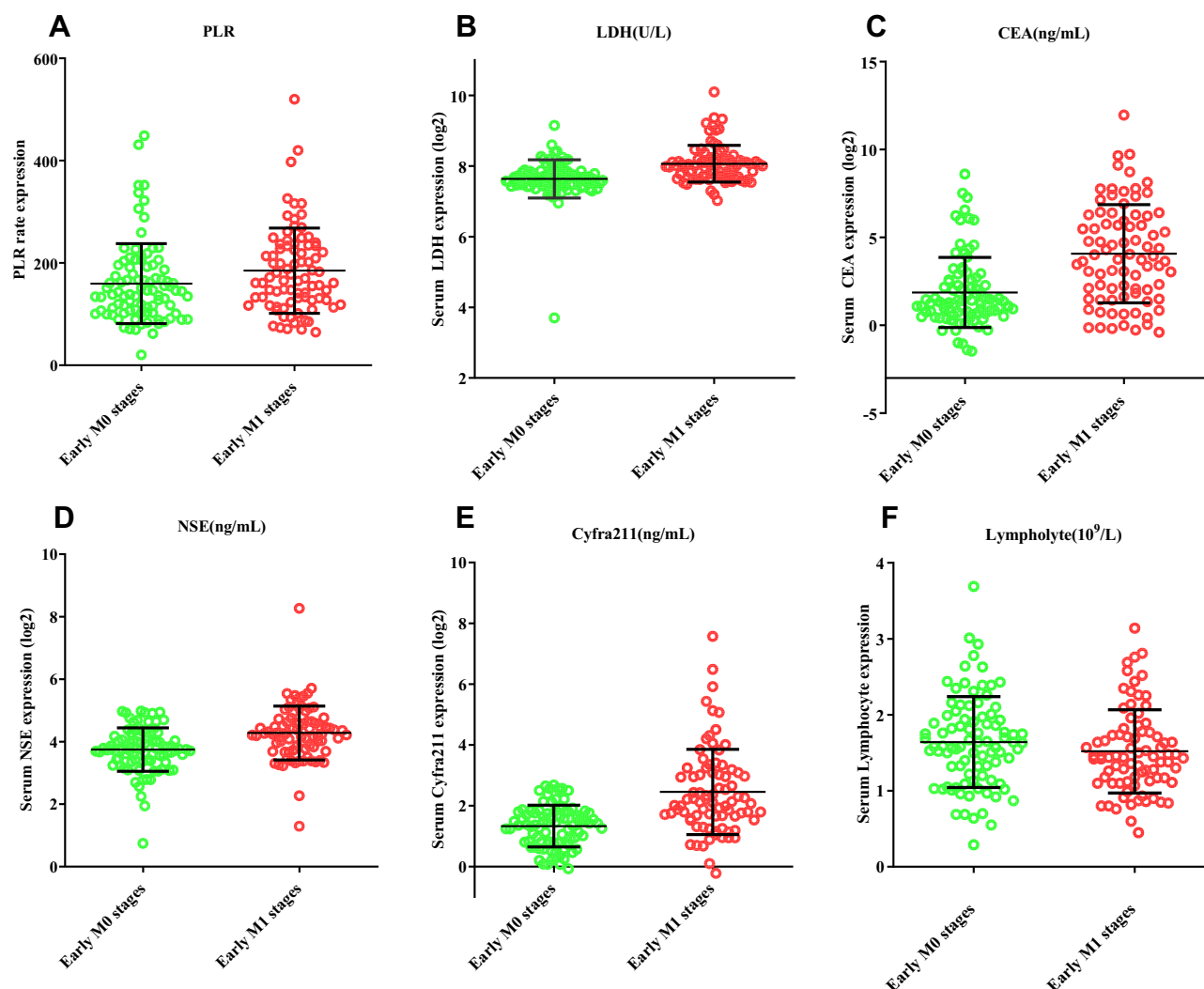
## Discussion

NSCLC is one of the malignant tumors with high incidence and poor prognosis in China. Even with active anti-tumor treatment, some patients have progressive disease and distant metastasis and require further treatment. TNM stage is an important prognostic biomarker of lung cancer,<sup>17</sup> and for early-stage NSCLC, the 5-year survival rate is more than 70%.<sup>18,19</sup> However, many patients have distant metastasis at the initial diagnosis of early T stages (diameter≤5cm). Moreover, we unable to find micrometastasis out early for

**Table 4** Spearman Correlation Analysis ET-LUAD and Clinicopathological Characteristics

Spearman		PLR	LDH	CEA	NSE	Cyfra211	Lymphocyte
Distant metastasis	r	0.222	0.451	0.409	0.394	0.381	–0.169
	P value	0.004	<0.001	<0.001	<0.001	<0.001	0.027

**Abbreviations:** PLR, platelet-to-lymphocyte ratio; LDH, lactate dehydrogenase; CEA, carcinoembryonic antigen; NSE, neural-specific enolase; Cyfra211, cytokeratin 19 fragment.



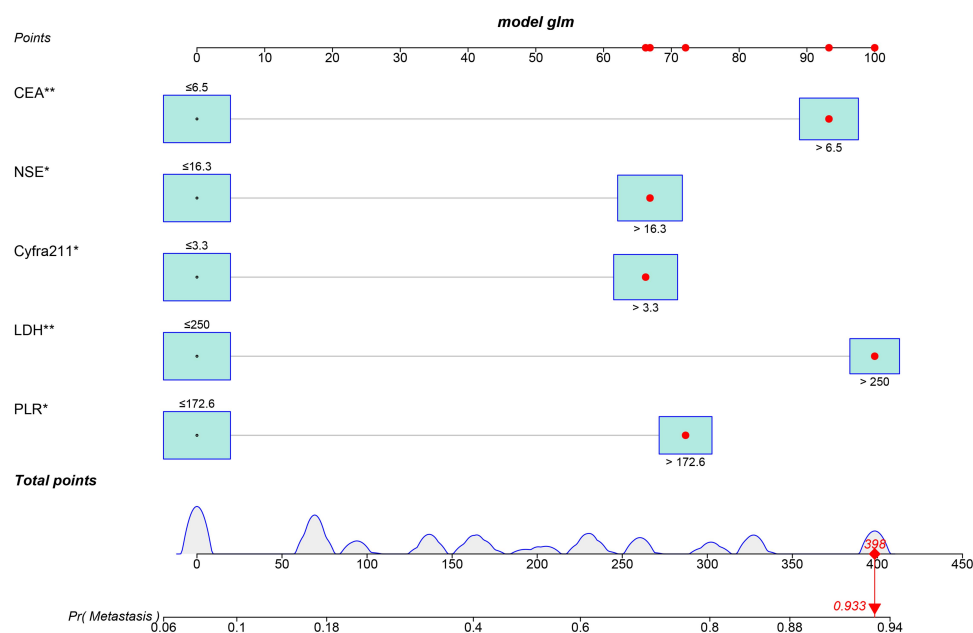
**Figure 2** Correlation analysis between ET-LUAD with distant metastases and clinicopathological characteristics in the training cohort. The early M1 stages in (A–E) expression is higher than early M0 stages, but lower than the lymphocyte count (F). The serum LDH, CEA, NSE and Cyfra211 expression were calculated by log2. **Abbreviations:** PLR, platelet-to-lymphocyte ratio; LDH, lactate dehydrogenase; CEA, carcinoembryonic antigen; NSE, neural-specific enolase; M0, T1-2 stages without distant metastasis; M1, T1-2 stages with distant metastasis.

the limitations of imaging technology, and the patients lost the best opportunity for treatment. In advanced NSCLC, the 5-year survival rate is less than 4%.<sup>6,7</sup> Therefore, it is important to find new detection techniques for diagnosis of early T stages in patients likely to develop into distant metastasis and provide the best support treatment.

The study by Wu et al<sup>17</sup> showed that early T1 stages of NSCLC preoperative related clinical indicators (gender, ages, serum CEA and tumor size) included in nomogram analysis, can be used as important biomarkers for the prediction of regional lymph node metastasis in early-stage lung cancer. Liao et al<sup>20</sup> enrolled 6245 T1-4N1-3M0 NSCLC patients from SEER database and analysis the correlation between clinical-pathological features and prognosis of survival, basic the positive lymph node ratio

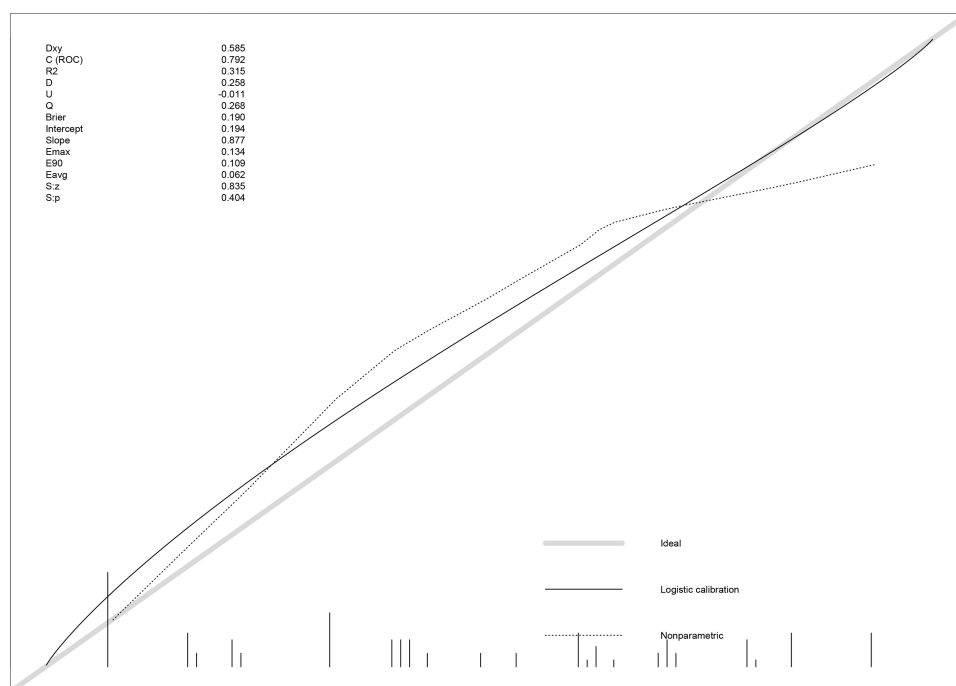
(pLNR) of variables in nomogram model can be better predict survival in patients with T1-4N1-3M0 NSCLC. Therefore, a nomogram focusing on clinicopathological features can be used to better predict the survival and prognosis of early-stage NSCLC.<sup>21,22</sup> However, only a few studies have reported on why some early T stages of LUAD quickly develop into distant metastasis, and clinical factors related to distant metastasis remain largely unknown. In this study, patients were randomly divided in a training cohort and validation cohort in a 1:2 ratio. The nomogram prediction model was established using relevant clinicopathological information of patients before receiving any treatment. Based on the univariate and multivariate logistic proportional hazards regression analysis, five factors were selected into the nomogram analysis in





**Figure 3** The nomogram predicting metastasis possibility for patients with non-metastatic ET-LUAD in the training cohort. The univariate and multivariate logistic hazards regression analysis identified independent risk factors and incorporated those variables to established nomogram model, including PLR, serum CEA, LDH, NSE, and Cyfra211. Then mark the data of those five variables on the interactive nomogram. Each variable is represented by a horizontal line, and the patient's information is marked on the coordinates. The regression coefficient of each predictive variable corresponds to the score value in the range of 0–100. The total score of the five variables was obtained by adding the corresponding scores, and the metastasis possibility obtained for ET-LUAD. The red dot on the scale represents the corresponding score of the variable. Pr (Metastasis), ET-LUAD with metastasis. (\* $P < 0.05$ , \*\* $P < 0.01$ ).

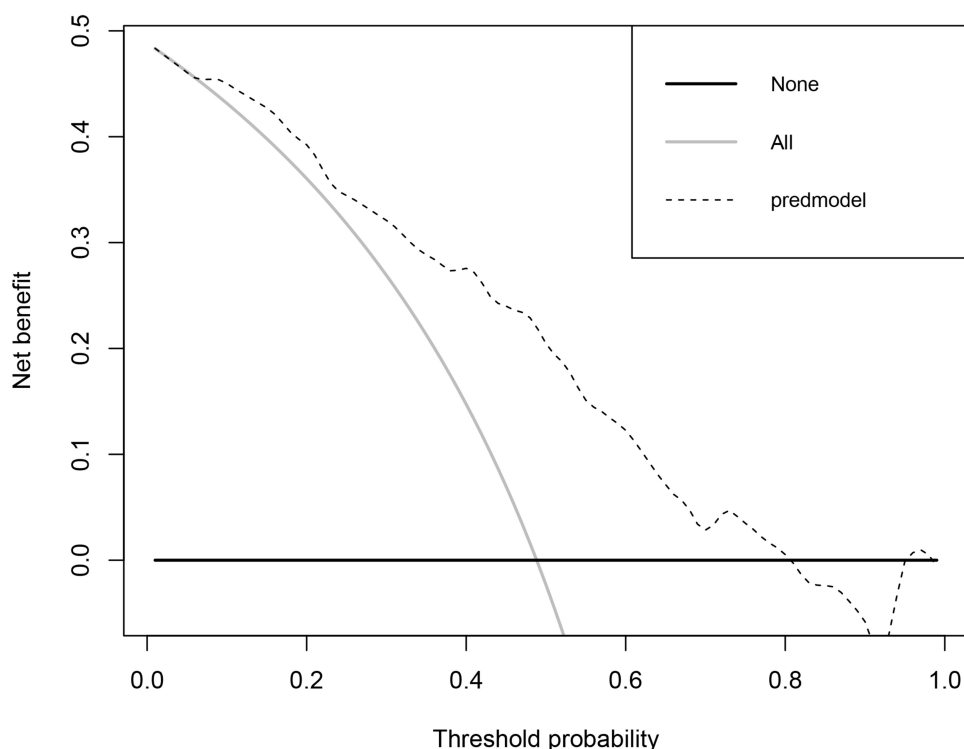
**Abbreviations:** PLR, platelet-to-lymphocyte ratio; LDH, lactate dehydrogenase; CEA, carcinoembryonic antigen; NSE, neural-specific enolase.



**Figure 4** The discrimination and calibration curves of prediction model.

the training cohort, the results showed that the nomogram can better predict the possibility of distant metastasis for the high-risk population in ET-LUAD. Moreover, the

prediction model was validated using a validation cohort group, and the result also achieved a high consistency and clinical application value. The nomogram model



**Figure 5** The decision curve analyses (DCA) for the clinical values of this model. The Y-axis represents the net benefit. The dotted line represents the clinicopathologic nomogram. The gray line represents the hypothesis that all patients are involved in distant metastases. The black solid line represents the hypothesis that no patients are involved in distant metastases. The X-axis represents the metastasis possibility. The metastasis possibility is where the expected benefit of treatment is equal to the expected benefit of avoiding treatment.

established in this study was similar to that by Zhang et al<sup>23</sup> which showed that tumor size is an independent risk factor for survival prognosis and regional lymph node metastasis in there was established model through SEER database.

Tumor biomarkers play an important role in diagnosis, recurrence, metastasis, and prediction of the curative effect after the treatment of lung cancer. To study the influence of tumor biomarkers on the prognosis of lung cancer, Yang et al<sup>9</sup> suggested that serum CEA, CA125, Cyfra-211 and NSE were important biomarkers for the prognosis of lung cancer, and the combination of the two biomarkers is even better for preliminary diagnosis of lung cancer. Moreover, Cacho-Díaz et al<sup>24</sup> studied the correlation between serum CEA and metastasis and showed that the presence of certain clinical neurologic symptoms with high serum CEA levels was associated with a higher risk of brain metastasis in lung cancer patients. Ayan et al<sup>25</sup> suggested that high levels of serum osteopontin and CEA were associated with a higher risk of bone metastasis in lung cancer. Another study also showed that CA-125, CA-153, and Cyfra21-1 is an important clinical biomarker in male lung cancer with ocular metastasis.<sup>26</sup> In this study, the

univariate and multivariate logistic analysis showed that high levels of CEA, NSE, and Cyfra21-1 were independent risk factors for ET-LUAD and may more easily become distant metastasis. Besides, the established nomogram had a high consistency (C-index=0.792) and good calibration, suggesting that nonmetastatic patients with high serum tumor biomarkers may be provided an important clinical application value to predict the ET-LUAD micrometastasis.

LDH is an enzyme that plays an important role in anaerobic glycolysis and induces cell proliferation, high levels of LDH are related to the promotion of tumor invasion, metastasis, and indicate short survival of NSCLC.<sup>27–29</sup> Zhang et al<sup>27</sup> conducted a systematic meta-analysis to analyze the relationship between LDH and small cell lung cancer (SCLC). The results showed that serum LDH levels were significantly related to the prognosis of SCLC, and LDH was considered as prognosis and potential tumor-promoting factor for lung cancer. Moreover, de Jong et al<sup>30</sup> studied the association between serum biomarkers CEA and LDH and response in advanced non-small-cell lung cancer patients treated with platinum-based chemotherapy. The results showed

that a decrease in CEA and LDH ( $\geq 20\%$ ), particularly in early treatment, were significantly associated with better radiological response. Dong et al<sup>31</sup> studied the correlation between LDH and distant metastasis and showed that immunohistochemistry of LDH-A expression and serum LDH status was closely associated with triple-negative breast cancer brain metastasis and brain metastasis-free survival. Besides, this study indicated that tumor LDH and serum LDH status are two predictors for triple-negative breast cancer brain metastasis. However, few clinical studies analyze the relationship between serum LDH and ET-LUAD, and whether serum LDH can be a potential biomarker for predicting distant metastasis was unknown. In this study, serum LDH was positively correlated to distant metastasis group, high LDH expression may more easily become distant metastasis. Moreover, the nomogram model of serum LDH combined with tumors biomarkers can better predict metastasis possibility in the training cohort group, and the validation cohort also shows good consistency.

Numerous studies also show that systemic inflammatory reaction (SIR) plays an important role in promoting tumor microvascular formation, tumor cell differentiation, proliferation, and inhibiting host immune cell activity.<sup>13,32,33</sup> Some inflammatory factors show changes in the occurrence and development of cancer in peripheral blood, including lymphocytes, neutrophils, monocytes, and platelets.<sup>14,34</sup> Amaral et al<sup>35</sup> predicted the effect of immunotherapy on advanced NSCLC and the results showed that an increase in pre-treatment NLR and PLR was associated with shorter OS and PFS, and it was suggested that these biomarkers may be predictive factors of poor prognosis in NSCLC patients. For NSCLC patients with distant metastasis, an increase in serum platelet count and a decrease in the average platelet volume is related to NSCLC patients with brain metastasis,<sup>14</sup> and the ratio of NLR can be used as a risk predictor.<sup>36</sup> Li et al<sup>15</sup> analyzed the predictive value of peripheral serum inflammation index on 310 cases of LUAD with brain metastasis and EGFR gene mutation. The results showed that systemic immune inflammation index (SII) composed of lymphocytes, neutrophils, and platelets was an independent prognostic factor. Furthermore, Kudo et al<sup>10</sup> showed that the maturation of dendritic cells, Th1, and the signal pathway of leukocyte exosmosis were inhibited in NSCLC with brain metastasis ( $P < 0.01$ ) and 64% of brain metastasis had T cell clone expansion. Besides, when T-cell receptor

sequences are shared between brain metastasis and primary lung tumors, the density of T cells in metastatic tumors is very low, hence it is considered that human immunity is inhibited in NSCLC with brain metastasis.

In this study, the indexes of peripheral serum inflammation indicators were included in univariate logistic regression analysis and the results showed that PLR, NLR, neutrophils count, and lymphocytes count were correlated with distant metastasis in ET-LUAD. Moreover, the multivariate logistic indicated that PLR ratio was an independent risk factor for distant metastasis. Besides, lymphocytes count was taken into Spearman correlation analysis and the results showed that it was negatively correlated to metastasis in ET-LUAD, although the multivariate logistic analysis was not significant. Furthermore, the nomogram, C-index and DCA showed that PLR was also a better predictive risk factor of early T stages in nonmetastatic patients. Therefore, it was suggested that the low lymphocyte count may lead to the low ability of cellular immunity and tumor suppression, and may also be an important factor for predicted distant metastasis in ET-LUAD. These results are similar to those reported in many studies showing that tumor-infiltrating lymphocytes (TIL) play an important role in tumor recurrence, immune response, and patient survival.<sup>37–39</sup> Besides, tumor-infiltrating B lymphocytes (TIB) in lung cancer tissues produce IgG that recognizes common tumor-specific antigen and relative to the metastasis,<sup>40</sup> and the number of tumor-infiltrating lymphocytes cells (CD3+, CD4+, CD8+, and CD20+) is decreased in early recurrence patients. The increase in the quantity of tumor-infiltrating T and B cells may contribute to the prevention of early-stage lung adenocarcinoma recurrence after surgical resection.<sup>41</sup>

There are some limitations in this study that should be addressed. Firstly, we only observed the risk and correlation between clinicopathological characteristics and ET-LUAD, but the influence of these factors on the prognosis was not analyzed and follow-up, such as time to disease progressive or overall survival, treatment regimens, short-term effect with Response Evaluation Criteria in Solid Tumors 1.1 (RECIST1.1) etc. Moreover, there are 112 patients with *EGFR* mutation status were unknown, and the small samples may lead to analyze the relationship between *EGFR* gene status and ET-LUAD were not accuracy, and need more samples to analysis. Thirdly, we only use the external and internal data of single medical center to verify the accuracy of nomogram model, but those clinical variables were significant correlation between

other countries and races also unknown. Finally, this study is a small sample size, retrospective study, and the enrollment of patients may produce selective bias. A larger sample size, multi-centered clinical study should be further analyzed these variables and confirm its clinical application value for ET-LUAD patients in the future.

## Conclusion

The prediction model of variables including PLR, CEA, LDH, NSE and Cyfra211 can be well used to predict the high-risk population of ET-LUAD without distant metastasis, and serum lymphocytes count may be an important factor for human immunity to overcome tumor metastasis. Besides, the nomogram model of our establishment has demonstrated its clinical accuracy and can be used as a tool for doctors to guide preventive and individualized treatment for ET-LUAD patients.

## Date Availability Statement

The data generated and analyzed during the current study are available from the corresponding author on reasonable request.

## Ethical Approval

This study was a retrospective observational study and without involved any human participant drug experiments, the present study has been reviewed and approved by the Ethics Committee of First Affiliated Hospital of Nanchang University, Nanchang, China. Ethics number is (2020) medicine research (78), and all patients provided written informed consent. According to the declaration of Helsinki, all the patients' data were kept confidential.

## Acknowledgments

We appreciate the effort of the physicians for enrolling patients and thank all the patients involved for allowing us to analyze their clinical data.

## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

This work was supported by the National Natural Science Foundation of China (grant number 81560410 to JH Mei). The funding sources had no role in the data collection, analysis or interpretation.

## Disclosure

The authors declare that they have no competing interests.

## References

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin.* **2018**;68(6):394–424.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics 2019. *CA Cancer J Clin.* **2019**;69(1):7–34.
3. Nakazawa K, Kurishima K, Tamura T, et al. Specific organ metastasis and survival in small cell lung cancer. *Oncol Lett.* **2012**;4:617–620. doi:10.3892/ol.2012.792
4. Riihimäki M, Hemminki A, Fallah M, et al. Metastatic sites and survival in lung cancer. *Lung Cancer.* **2014**;86:78–84. doi:10.1016/j.lungcan.2014.07.020
5. Tamura T, Kurishima K, Nakazawa K, et al. Specific organ metastases and survival in metastatic non-small-cell lung cancer. *Mol Clin Oncol.* **2015**;3(1):217–221. doi:10.3892/mco.2014.410
6. Ashour Badawy A, Khedr G, Omar A, et al. Site of metastases as prognostic factors in unselected population of stage IV non-small cell lung cancer. *Asian Pac J Cancer Prev.* **2018**;19(7):1907–1910.
7. Goldstraw P, Chansky K, Crowley J, et al. The IASLC lung cancer staging project: proposals for revision of the TNM stage groupings in the forthcoming (Eighth) edition of the TNM classification for lung cancer. *J Thorac Oncol.* **2016**;11:39–51.
8. Chen Y-H, Chen Y-F, Chen C-Y, et al. Clinical factors associated with treatment outcomes in EGFR mutant non-small cell lung cancer patients with brain metastases: a case-control observational study. *BMC Cancer.* **2019**;19(1):1006. doi:10.1186/s12885-019-6140-0
9. Yang Q, Zhang P, Wu R, et al. Identifying the best biomarker combination in CEA, CA125, CY211, NSE, and SCC for lung cancer screening by combining ROC curve and logistic regression analyses: is it feasible? *Dis Biomarkers.* **2018**;2018:1–12.
10. Kudo Y, Haymaker C, Zhang J, et al. Suppressed immune micro-environment and repertoire in brain metastasis from patients with resected non-small cell lung cancer. *Ann Oncol.* **2019**;30(9):1521–1530. doi:10.1093/annonc/mdz207
11. Darlix A, Lamy PJ, Lopez-Crapez E, et al. Serum NSE, MMP-9 and HER2 extracellular domain are associated with brain metastasis in metastatic breast cancer patients: predictive biomarkers for brain metastasis? *Int J Cancer.* **2016**;139(10):2299–2311.
12. Wang J, Chu Y, Li J, et al. The clinical value of carcinoembryonic antigen for tumor metastasis assessment in lung cancer. *PeerJ.* **2019**;7:e7433. doi:10.7717/peerj.7433
13. De Raaf PJ, Sleijfer S, Lamers CH, et al. Inflammation and fatigue dimensions in advanced cancer patients and cancer survivors: an explorative study. *Cancer.* **2012**;118(23):6005–6011.
14. Li MM, Wang X, Yun ZY, et al. Platelet indices in non-small cell lung cancer patients with brain metastasis [J]. *Cancer Biobiomarker.* **2019**;24(4):515–519. doi:10.3233/CBM-192393
15. Li H, Wang G, Zhang H, et al. Prognostic role of the systemic immune-inflammation index in brain metastases from lung adenocarcinoma with different EGFR mutations. *Genes Immun.* **2019**;20(6):455–461. doi:10.1038/s41435-018-0050-z

16. Weichselbaum RR, Hellman S. Oligometastases revisited. *Nat Rev Clin Oncol*. 2011;8(6):378–382. doi:10.1038/nrclinonc.2011.44
17. Wu Y, Liu J, Han C, et al. Preoperative prediction of lymph node metastasis in patients with early-T-stage non-small cell lung cancer by machine learning algorithms. *Front Oncol*. 2020;10:743.
18. Rami-Porta R, Bolejack V, Crowley J, et al. The IASLC lung cancer staging project: proposals for the revisions of the T descriptors in the forthcoming eighth edition of the TNM classification for lung cancer. *J Thorac Oncol*. 2015;10(7):990–1003.
19. Takenaka T, Takenoyama M, Yamaguchi M, et al. Impact of the epidermal growth factor receptor mutation status on the post-recurrence survival of patients with surgically resected non-small-cell lung cancer. *Eur J Cardiothorac Surg*. 2015;47(3):550–555. doi:10.1093/ejcts/ezu227
20. Liao Y, Yin G, Fan X. The positive lymph node ratio predicts survival in T1-4N1-3M0 non-small cell lung cancer: a nomogram using the SEER database. *Front Oncol*. 2020;10:1356. doi:10.3389/fonc.2020.01356
21. Wo Y, Yang H, Zhang Y, et al. Development and external validation of a nomogram for predicting survival in patients with stage IA non-small cell lung cancer  $\leq 2$  cm undergoing sublobectomy. *Front Oncol*. 2019;9:1385. doi:10.3389/fonc.2019.01385
22. Kang J, Ning MS, Feng H, et al. Predicting 5-year progression and survival outcomes for early stage non-small cell lung cancer treated with stereotactic ablative radiation therapy: development and validation of robust prognostic nomograms. *Int J Radiat Oncol Biol Phys*. 2020;106(1):90–99. doi:10.1016/j.ijrobp.2019.09.037
23. Zhang J, Gold KA, Lin HY, et al. Relationship between tumor size and survival in non-small-cell lung cancer (NSCLC): an analysis of the surveillance, epidemiology, and end results (SEER) registry. *J Thorac Oncol*. 2015;10(4):682–690. doi:10.1097/JTO.0000000000000456
24. Cacho-Díaz B, Spinola-Maróño H, Mendoza-Olivas LG, et al. Association of neurologic manifestations and CEA levels with the diagnosis of brain metastasis in lung cancer patients. *Clin Transl Oncol*. 2019;21(11):1538–1542. doi:10.1007/s12094-019-02086-y
25. Ayan AK, Erdemci B, Orsal E, et al. Is there any correlation between levels of serum osteopontin, CEA, and FDG uptake in lung cancer patients with bone metastasis? *Rev Esp Med Nucl Imagen Mol*. 2016;35(2):102–106.
26. Li B, Yuan Q, Zou YT, et al. CA-125, CA-153, and CYFRA21-1 as clinical indicators in male lung cancer with ocular metastasis. *J Cancer*. 2020;11(10):2730–2736.
27. Zhang X, Guo M, Fan J, et al. Prognostic significance of serum LDH in small cell lung cancer: a systematic review with meta-analysis. *Cancer Biomark*. 2016;16(3):415–423. doi:10.3233/CBM-160580
28. Deng T, Zhang J, Meng Y, et al. Higher pretreatment lactate dehydrogenase concentration predicts worse overall survival in patients with lung cancer. *Medicine (Baltimore)*. 2018;97(38):e12524.
29. Lee DS, Park KR, Kim SJ, et al. Serum lactate dehydrogenase levels at presentation in stage IV non-small cell lung cancer: predictive value of metastasis and relation to survival outcomes. *Tumour Biol*. 2016;37(1):619–625. doi:10.1007/s13277-015-3776-5
30. de Jong C, Deneer VHM, Kelder JC, et al. Association between serum biomarkers CEA and LDH and response in advanced non-small cell lung cancer patients treated with platinum-based chemotherapy. *Thorac Cancer*. 2020;11(7):1790–1800. doi:10.1111/1759-7714.13449
31. Dong T, Liu Z, Xuan Q, et al. Tumor LDH-A expression and serum LDH status are two metabolic predictors for triple negative breast cancer brain metastasis. *Sci Rep*. 2017;7(1):6069. doi:10.1038/s41598-017-06378-7
32. Baniyash M. Myeloid-derived suppressor cells as intruders and targets: clinical implications in cancer therapy. *Cancer Immunol Immunother*. 2016;65(7):857–867. doi:10.1007/s00262-016-1849-y
33. Yang LY, Wu XJ, Ye SB, et al. Tumor-induced myeloid-derived suppressor cells promote tumor progression through oxidative metabolism in human colorectal cancer. *J Transl Med*. 2015;13(47):1–12.
34. Kedzierska M, Czernek U, Szydzowska-Pazera K, et al. The changes of blood platelet activation in breast cancer patients before surgery, after surgery, and in various phases of the chemotherapy. *Platelets*. 2013;24(6):462–468.
35. Amaral SR, Casal Moura M, Carvalho J, et al. Prognostic significance of neutrophil-to-lymphocyte ratio (NLR) and platelet-to-lymphocyte ratio (PLR) in non-small cell lung cancer (NSCLC) treated with immune checkpoint inhibitors. *Ann Oncol*. 2019;30(1):i3.
36. Koh YW, Choi JH, Ahn MS, et al. Base line neutrophil-lymphocyte ratio is associated with baseline and subsequent presence of brain metastasis in advanced non-small-cell lung cancer. *Sci Rep*. 2016;6(1):38585. doi:10.1038/srep38585
37. Montfort A, Pearce O, Maniati E, et al. A strong B-cell response is part of the immune landscape in human high-grade serous ovarian metastases. *Clin Cancer Res*. 2017;23(1):250–262. doi:10.1158/1078-0432
38. DiLillo DJ, Yanaba K, Tedder TF. B cells are required for optimal CD4+ and CD8+ T cell tumor immunity: therapeutic B cell depletion enhances B16 melanoma growth in mice. *J Immunol*. 2010;184(7):4006–4016. doi:10.4049/jimmunol.0903009
39. Brambilla E, Le Teuff G, Marguet S, et al. Prognostic effect of tumor lymphocytic infiltration in resectable non-small-cell lung cancer. *J Clin Oncol*. 2016;34(11):1223–1230. doi:10.1200/JCO.2015.63.0970
40. Imahayashi S, Ichiyoshi Y, Yoshino I, et al. Tumor-infiltrating B-cell-derived IgG recognizes tumor components in human lung cancer. *Cancer Invest*. 2000;18(6):530–536.
41. Liu J, Yang X, Lu X, et al. Impact of T-cell receptor and B-cell receptor repertoire on the recurrence of early stage lung adenocarcinoma. *Exp Cell Res*. 2020;112134.

## Therapeutics and Clinical Risk Management

### Publish your work in this journal

Therapeutics and Clinical Risk Management is an international, peer-reviewed journal of clinical therapeutics and risk management, focusing on concise rapid reporting of clinical studies in all therapeutic areas, outcomes, safety, and programs for the effective, safe, and sustained use of medicines. This journal is indexed on PubMed Central, CAS,

EMBASE, Scopus and the Elsevier Bibliographic databases. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/therapeutics-and-clinical-risk-management-journal>

Dovepress