

A CT-Based and Serum-Derived Nomogram for Preoperative Prediction of Lymphovascular Invasion in Stage I Lung Adenocarcinoma

Hanbing Zhang

Hebei Medical University

Peng Qie

Hebei General Hospital

Junpeng Feng

Hebei General Hospital

Yang Liu

Hebei General Hospital

Dehao Liu

Hebei General Hospital

Huien Wang

wanghe829@126.com

Hebei General Hospital

Research Article

Keywords: Lung adenocarcinoma, Lymphovascular invasion, Nomogram, CT imaging, NSE, Preoperative prediction

Posted Date: October 5th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10816605/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Lymphovascular invasion (LVI) is a crucial prognostic factor in stage I lung adenocarcinoma (LUAD), yet it can only be confirmed postoperatively. This study aimed to develop a preoperative predictive model for LVI by integrating hematological, clinical, and CT-based radiological features.

Methods

We retrospectively analyzed 800 patients with stage I LUAD who underwent curative resection. Least Absolute Shrinkage and Selection Operator (LASSO) regression and multivariate logistic regression were employed to identify independent predictors. A nomogram was constructed and internally validated via bootstrap resampling. Model performance was assessed using the area under the receiver operating characteristic curve (AUC), calibration curves, and decision curve analysis (DCA).

Results

Six independent predictors were identified: neuron-specific enolase (NSE), tumor diameter, consolidation tumor ratio (CTR) > 0.5, spiculation, vacuolar sign, and vessel convergence sign. The nomogram demonstrated good discrimination, with AUCs of 0.841 (95% CI: 0.807–0.875) in the training cohort and 0.786 (95% CI: 0.704–0.868) in the validation cohort. Calibration curves showed excellent agreement between predicted and observed probabilities, and DCA confirmed clinical net benefit across a wide threshold range (5%–80%).

Conclusion

This simple nomogram incorporating routine CT features and serum NSE offers a robust, noninvasive tool for preoperative individualized prediction of LVI in stage I LUAD, potentially aiding in risk stratification and treatment planning.

Introduction

Lung cancer is the most frequently diagnosed malignant tumor globally and a major cause of cancer-related mortality^[1]. Non-small cell lung cancer (NSCLC) accounts for roughly 89.2% of all lung cancer cases. Lung adenocarcinoma (LUAD), the most common histological subtype of NSCLC, continues to show rising incidence rates worldwide^[2]. Surgical resection is the mainstay of treatment for early-stage lung adenocarcinoma. However, a substantial proportion of patients still experience postoperative recurrence and metastasis, even among those with completely resected stage I NSCLC^[3–4].

Lymphovascular invasion (LVI) refers to tumor cell infiltration into arteries, veins, or lymphatic vessels. It is considered an initial step in local recurrence and distant metastasis and represents an important independent adverse prognostic factor for lung adenocarcinoma^[5]. Extensive studies have consistently demonstrated that LVI-positive patients have a substantially higher risk of postoperative recurrence and significantly shorter survival^[6–7]. Given the well-established prognostic value of LVI, the 2026 NCCN Clinical Practice Guidelines for Non-Small Cell Lung Cancer have incorporated it as a key reference factor in postoperative adjuvant therapy decisions^[8]. However, as a pathological indicator, LVI can

only be confirmed through postoperative pathological examination. This precludes clinicians from knowing a patient's LVI status before surgery, making it difficult to perform accurate risk stratification or tailor individualized treatment plans during the preoperative phase. Therefore, there is considerable clinical value in developing effective noninvasive methods to predict LVI prior to surgery.

In recent years, hematological markers have been widely used in cancer screening, prognostic assessment, and treatment monitoring, owing to their noninvasive nature, simplicity, and low cost. For instance, composite indicators such as carcinoembryonic antigen (CEA) and the systemic immune-inflammation index (SII) have been demonstrated to be closely associated with the development and progression of LUAD^[9–10]. However, few studies have explored the value of hematological markers in predicting LVI. In other tumor types, considerable research has investigated the ability of routine hematological indicators to predict LVI in gastric cancer and hepatocellular carcinoma^[11–12]. Furthermore, other studies have confirmed the prognostic value of composite markers such as the prognostic nutritional index (PNI) and SII in prostate cancer and breast cancer, respectively^[13–14]. Therefore, there is theoretical support for investigating the value of hematological markers in predicting LVI in LUAD patients.

In terms of imaging features, the rapid development of preoperative radiomics and deep learning techniques has opened new avenues for noninvasive prediction of LVI. Multiple studies have demonstrated that computed tomography (CT) and magnetic resonance imaging (MRI) can effectively predict LVI in rectal cancer, breast cancer, and hepatocellular carcinoma, showing favorable diagnostic performance^[15–17].

This study aims to develop a predictive model for LVI in stage I lung adenocarcinoma by integrating hematological markers, clinical characteristics, and imaging features. Through this multidimensional integrative analysis, we expect to enhance predictive accuracy and offer clinicians a more effective reference for guiding early intervention and treatment selection in lung adenocarcinoma.

Methods

- This study retrospectively enrolled 1,373 patients who underwent surgical resection for lung cancer at Hebei General Hospital between January 2024 and July 2026. Ultimately, 800 patients were included in the final analysis. (Fig. 1: Flow diagram shows the patients who were included and excluded in this study).
- Inclusion criteria for this study subjects: (1) postoperative pathological diagnosis of lung adenocarcinoma with clear documentation of LVI status in the pathology report; (2) R0 resection; (3) chest CT scan performed within 1 month before surgery; (4) complete clinical records with detailed patient information available. Exclusion criteria: (1) prior chemotherapy, radiotherapy, immunotherapy, or any other neoadjuvant therapy before surgery; (2) history of other malignancies; (3) presence of infection-related symptoms (e.g., fever or chills) within 1 month before surgery; (4) pre-existing severe liver disease, hematological disorders, or organ failure. The present study was approved by the Ethics Committee of Hebei General Hospital (2026-LW-291). Written informed consent was obtained from all surviving participants, as well as from the relatives who provided information on behalf of deceased participants.

Data Collection

- Clinical and pathological data of all enrolled patients were collected and reviewed through the electronic medical record system of our hospital. The study was conducted in strict compliance with the Declaration of Helsinki, ensuring data privacy and security. The specific contents collected are detailed below.

(1) Basic information: age, sex, smoking history, smoking index, body mass index (BMI), tumor size, and tumor location

(2) Preoperative hematological examination data (within 1 week before surgery): White blood cell count (WBC), neutrophil count (NEUT), lymphocyte count (LYMPH), monocyte count (MONO), eosinophil count (EO), hemoglobin (HGB), platelet count (PLT), total protein (TP), albumin (ALB), lactate dehydrogenase (LDH), carcinoembryonic antigen (CEA), cytokeratin 19 fragment (CYFRA21-1), neuron-specific enolase (NSE), squamous cell carcinoma antigen (SCC), prothrombin time (PT), activated partial thromboplastin time (APTT), thrombin time (TT), fibrinogen (FIB), and D-dimer (DD). All laboratory tests were performed by the clinical laboratory at our hospital.

(3) Definition and calculation of composite hematological indices: The main composite indices analyzed in this study included the prognostic nutritional index (PNI), systemic immune-inflammation index (SII), neutrophil-to-lymphocyte ratio (NLR), derived neutrophil-to-lymphocyte ratio (d-NLR), platelet-to-lymphocyte ratio (PLR), lymphocyte-to-monocyte ratio (LMR), and systemic inflammation response index (SIRI). According to the relevant literature, these indices were calculated using the following formulas:

☒ $PNI = \text{serum albumin (g/dL)} + 5 \times \text{total lymphocyte count } (\times 10^9/L)$

☒ $SII = \text{platelet count} \times \text{neutrophil count} / \text{lymphocyte count}$

☒ $NLR = \text{neutrophil count} / \text{lymphocyte count}$

☒ $d\text{-NLR} = \text{neutrophil count} / (\text{white blood cell count} - \text{neutrophil count})$

☒ $PLR = \text{platelet count} / \text{lymphocyte count}$

☒ $LMR = \text{lymphocyte count} / \text{monocyte count}$

☒ $SIRI = \text{neutrophil count} \times \text{monocyte count} / \text{lymphocyte count}$

(4) Imaging data: Chest CT images were retrospectively reviewed independently by two thoracic surgeons with more than 5 years of experience. Disagreements were resolved through discussion and mutual validation until consensus was reached. The evaluated CT features included: maximum nodule diameter, proportion of solid component, Vacuolar sign (Vac), Lobulation sign (Lob), Spiculation sign (Spic), Vascular sign (VasC), and Pleural Indentation (PI).

Statistical Analysis

Statistical analyses were performed using SPSS version 27.0 and R software (version 4.6.1). Continuous variables with a normal distribution were presented as mean $\pm$ standard deviation, and intergroup comparisons were conducted using the independent samples t-test. Non-normally distributed continuous variables were expressed as median (interquartile range), with group comparisons performed using non-parametric tests. Categorical variables were presented as frequencies (percentages) and compared using the chi-square test. Predictive model development and validation were carried out in the R environment. Model discrimination performance was assessed using the area under the receiver operating characteristic curve (AUC), sensitivity, and specificity.

Result

Univariate Analysis of the Study Population

- A total of 800 patients were enrolled in this study and divided into LVI-negative and LVI-positive groups according to LVI status. Univariate analysis revealed that **[Insert Table 1 here]**, compared with the LVI-negative group, the LVI-positive group exhibited larger maximum tumor diameter, a higher proportion of solid component, and more frequent presence of Spic, Vac, VasC, and PI on imaging. In addition, levels of hematological markers such as HGB and TP were lower in the LVI-positive group, whereas tumor marker levels including CEA, SCC, NSE, and CYFRA21-1 were higher. All these differences were statistically significant ($P < 0.05$).

Table 1
Univariate analysis comparing the LVI-positive and LVI-negative groups.

Factors	Total(n = 800)	LVI+(n = 158)	LVI-(n = 642)	p
SEX,n(%)				0.682
male	305(38.12)	58(36.7)	247(38.5)	
female	495(61.88)	100(63.3)	395(61.5)	
AGE,Median(Q1,Q3)	64.00(56.25,70.00)	66.00(58.75,70.00)	63.00(56.00,70.00)	0.202
BMI,Median(Q1,Q3)	25.27(23.03,27.83)	24.99(23.02,27.56)	25.34(23.02,27.90)	0.455
Smoking history,n(%)				0.440
yes	184(23.00)	40(25.3)	144(22.4)	
no	616(77.00)	118(74.7)	498(77.6)	
Smoking index,Median(Q1,Q3)	0(0,0)	0(0,300)	0(0,0)	0.356
Laterality,n(%)				0.995
Left	334(41.75)	66(41.8)	268(41.7)	
Right	466(58.25)	92(58.2)	374(58.3)	
Location,n(%)				0.797
Upper	468(58.50)	91(57.6)	377(58.7)	
Middle/Lower	332(41.50)	67(42.4)	265(41.3)	
CTR > 0.5,n(%)				< 0.001
yes	353(44.13)	105(66.5)	248(38.6)	
no	447(55.87)	53(33.5)	394(61.4)	
tumor size,Median(Q1,Q3)	1.60(1.20,2.30)	2.30(1.60,3.00)	1.50(1.10,2.10)	< 0.001
Spic,n(%)				< 0.001
yes	239(29.88)	73(46.2)	166(25.9)	
no	561(70.12)	85(53.8)	476(74.1)	
Lob,n(%)				0.412
yes	178(22.25)	39(24.7)	139(21.7)	
no	622(77.75)	119(75.3)	503(78.3)	
Vac,n(%)				0.001
yes	125(15.63)	38(24.1)	87(13.6)	
no	675(84.37)	120(75.9)	555(86.4)	

Factors	Total(n = 800)	LVI+(n = 158)	LVI-(n = 642)	p
VasC,n(%)				< 0.001
yes	218(27.25)	100(63.3)	118(18.4)	
no	582(72.75)	58(36.7)	524(81.6)	
PI,n(%)				0.002
yes	300(37.50)	76(48.1)	224(34.9)	
no	500(62.50)	82(51.9)	418(65.1)	
WBC,Median(Q1,Q3)	5.64(4.73,6.81)	5.51(4.62,6.72)	5.67(4.74,6.85)	0.286
RBC,Mean ± SD	4.46 ± 0.48	4.40 ± 0.49	4.47 ± 0.48	0.095
HGB,Mean ± SD	135.03 ± 14.95	132.59 ± 14.19	135.62 ± 15.08	0.022
PLT,Median(Q1,Q3)	231.00(195.25,276.00)	231.50(201.00,275.50)	231.00(195.00,276.00)	0.553
NEUT,Median(Q1,Q3)	3.55(2.85,4.58)	3.34(2.81,4.34)	3.58(2.85,4.68)	0.118
LY,Median(Q1,Q3)	1.58(1.24,1.97)	1.56(1.24,1.98)	1.58(1.25,1.97)	0.770
MONO,Median(Q1,Q3)	0.26(0.21,0.33)	0.27(0.22,0.33)	0.26(0.21,0.34)	0.425
EO,Median(Q1,Q3)	0.08(0.05,0.14)	0.09(0.05,0.14)	0.08(0.04,0.15)	0.902

Table 1
(continued)

Factors	Total(n = 800)	LVI+(n = 158)	LVI-(n = 642)	p
PT,Median(Q1,Q3)	10.90(10.50,11.40)	11.00(10.50,11.40)	10.90(10.50,11.40)	0.864
APTT,Median(Q1,Q3)	25.90(24.70,27.30)	25.55(24.40,27.30)	26.00(24.80,27.40)	0.072
TT,Median(Q1,Q3)	17.90(17.20,18.50)	18.00(17.30,18.50)	17.90(17.20,18.50)	0.398
FIB,Median(Q1,Q3)	2.69(2.36,3.16)	2.73(2.40,3.12)	2.67(2.33,3.16)	0.705
D-D,Median(Q1,Q3)	0.28(0.18,0.46)	0.30(0.19,0.48)	0.28(0.18,0.46)	0.440
TP,Median(Q1,Q3)	72.20(67.53,75.58)	70.70(65.70,75.60)	72.35(68.10,75.53)	0.034
ALB,Median(Q1,Q3)	43.70(40.90,45.90)	43.30(39.78,45.70)	43.80(41.10,45.90)	0.141
LDH,Median(Q1,Q3)	174.20(153.53,196.50)	176.75(155.88,200.83)	173.95(153.05,195.93)	0.241
CEA,Median(Q1,Q3)	2.20(1.30,3.87)	3.44(1.76,6.30)	2.02(1.20,3.19)	< 0.001
SCC,Median(Q1,Q3)	0.40(0.28,0.58)	0.43(0.29,0.69)	0.39(0.27,0.57)	0.035
NSE,Median(Q1,Q3)	10.35(8.82,12.03)	11.02(9.24,12.50)	10.13(8.72,11.91)	0.005
CYFRA21-1,Median(Q1,Q3)	1.87(1.33,2.74)	1.98(1.46,3.02)	1.84(1.28,2.70)	0.028
PNI,Median(Q1,Q3)	51.65(48.70,54.34)	51.10(48.15,54.13)	51.83(48.79,54.35)	0.137
SII,Median(Q1,Q3)	529.18(379.06,729.04)	516.98(383.56,719.27)	532.33(378.66,731.10)	0.837
NLR,Median(Q1,Q3)	2.24(1.71,3.07)	2.21(1.68,2.99)	2.24(1.72,3.15)	0.366
d-NLR,Median(Q1,Q3)	1.79(1.37,2.41)	1.74(1.33,2.33)	1.79(1.38,2.44)	0.336
PLR,Median(Q1,Q3)	147.75(115.70,188.11)	153.39(117.12,195.46)	146.62(115.21,187.20)	0.285
LMR,Median(Q1,Q3)	5.98(4.62,7.57)	5.82(4.68,7.36)	6.04(4.59,7.61)	0.348
SIRI,Median(Q1,Q3)	0.60(0.41,0.89)	0.58(0.42,0.82)	0.60(0.40,0.91)	0.735

Variable Selection Using LASSO Regression

- To address collinearity among independent variables and avoid model overfitting, we applied a dimensionality reduction approach using least absolute shrinkage and selection operator (LASSO) regression to identify risk factors closely associated with LVI in lung adenocarcinoma Fig. 2(LASSO regression variable selection results). As shown in Fig. 2, the coefficients of some variables were progressively compressed to zero as more predictors were incorporated into the model. To reduce dimensionality and select the most robust risk factors, we performed LASSO regression with 10-fold cross-validation. The optimal penalty parameter λ was determined based on the minimum deviance criterion ($\lambda_{\min}$). As illustrated in Fig. 3, cross-validation error decreased substantially as $\log(\lambda)$ diminished. Ultimately, under the $\lambda_{\min}$ criterion, six candidate predictors with non-zero coefficients were selected: NSE, maximum tumor diameter (cm), CTR > 0.5, Spic, Vac, and VasC.

Multivariable Logistic Regression Results

- Multivariable logistic regression analysis was performed on the six variables selected by LASSO regression. A backward stepwise regression method was then applied for further variable selection[Insert Table 2 here]. Ultimately, six independent predictors were identified: NSE, maximum tumor diameter (cm), CTR > 0.5, Spic, Vac, and VasC.

Table 2
Multivariable logistic regression results

Intercept	Coef	S.E.	Wald Z	Pr(> Z)
	-4.4780	0.4644	-9.64	< 0.0001
NSE	0.0893	0.0327	2.73	0.0063
tumor diameter(cm)	0.3327	0.1236	2.69	0.0071
CTR > 0.5	0.6628	0.2419	2.74	0.0062
Spic	0.5376	0.2290	2.35	0.0189
Vac	0.6094	0.2532	2.41	0.0161
VasC	1.9577	0.2103	9.31	< 0.0001

Construction of the nomogram prediction model

- Based on the results of LASSO regression and multivariable logistic regression analysis, six independent factors for predicting LVI in lung adenocarcinoma were identified, and a predictive model was subsequently established. The predicted probability of LVI was calculated as $P = 1/(1 + e^{-X})$, $X = -4.4780 + 0.0893 \times \text{NSE} + 0.3327 \times \text{tumor diameter} + 0.6628 \times \text{CTR} > 0.5 + 0.5376 \times \text{Spic} + 0.6094 \times \text{Vac} + 1.9577 \times \text{VasC}$. Here, e denotes the base of the natural logarithm, and X represents the linear combination of regression coefficients from the logistic model. Tumor diameter is measured in centimeters.
- Based on the constructed model, a nomogram was developed, as shown in Fig. 4. In this nomogram, each predictor is assigned a corresponding point value. The points for all predictors are summed to obtain a total score, which is then located on the total score axis. A vertical line is drawn downward from that point to determine the individual probability of LVI. This model offers the advantage of not only quantifying LVI risk but also providing clinicians with a simple and intuitive tool to better guide treatment decisions.

Model discrimination

- Receiver operating characteristic (ROC) curves were used to evaluate the discriminative ability of the nomogram and to quantitatively assess its predictive performance for LVI in lung adenocarcinoma patients. The AUC value of the model in the training cohort was 0.8409 (95% CI: 0.8074–0.8745), with sensitivity and specificity of 85.94% and 66.80%, respectively, at the optimal cutoff point ($T = -1.9848$). In the test cohort, the AUC was 0.7859 (95% CI: 0.7042–0.8676), with sensitivity and specificity of 80.00% and 58.46% at the same cutoff point. The model demonstrated good discriminative ability, effectively distinguishing lung cancer patients with and without LVI. Figure 5 presents the ROC curves for the training set (left panel) and validation set (right panel).

Model calibration

- To evaluate the agreement between the nomogram-predicted probabilities and actual observed outcomes, calibration curves were plotted, and internal validation was performed using the bootstrap method with 1000 resamples. As shown in the Fig. 6, in the training set (left panel), both the apparent curve (unadjusted) and the bias-corrected curve closely aligned with the ideal 45° reference line, indicating good calibration performance of the model on the modeling data. In the validation set (right panel), the apparent curve also closely followed the ideal diagonal, with no evident deviation. Notably, the two calibration curves exhibited highly consistent shapes across different datasets, without any clear trend of systematic overestimation or underestimation. These findings further confirm that the nomogram has good generalizability and stable calibration performance. Collectively, these results demonstrate that the model can provide reasonably accurate individualized risk prediction of LVI for patients with stage I lung adenocarcinoma.

Clinical utility of the model

- To further evaluate the clinical applicability of the nomogram, decision curve analysis (DCA) was performed to assess the net benefit of the model across different risk thresholds. In the DCA curve, the horizontal axis represents the risk threshold and the vertical axis represents the net benefit. The "All" curve corresponds to the strategy of treating all patients as positive, with net benefit decreasing as the threshold increases. The "None" curve corresponds to the strategy of treating no patients, with net benefit constantly at zero. When the model curve lies above both the "All" and "None" curves, it indicates a clinical net benefit within that threshold range.
- As shown in Fig. 7, in the training set, the model's decision curve remained above both the "treat-all" and "treat-none" strategies across a threshold probability range of 5% to 80%, confirming its clinical utility, suggesting that clinical decisions guided by the model could provide positive net benefit for patients within this interval. Results in the validation set were highly consistent with those in the training set, with the model curve also maintaining a net benefit advantage over a wide threshold range. These findings indicate that the nomogram has stable clinical utility and can serve as an effective auxiliary tool for preoperative assessment of LVI risk.

Discussion

- Cancer has become the leading cause of death worldwide. Globally, approximately 4,986 people die from lung cancer each day, accounting for 18.7% of all cancer-related deaths^[18]. In 2022, the age-standardized incidence rate (ASIR) for all cancers in China was 201.6 per 100,000 population, with lung cancer having the highest ASIR at 40.8 per 100,000. The age-standardized mortality rate (ASMR) for all cancers was 96.5 per 100,000, and lung cancer was the leading cause of cancer death, with an ASMR of 26.7 per 100,000^[19]. Lung cancer remains the most common and deadliest cancer type both in China and worldwide, with LUAD accounting for approximately 45.6% of all lung cancer cases^[2]. LVI is a key prognostic factor affecting outcomes in LUAD patients. Studies have shown that LUAD patients with LVI have significantly inferior overall survival (OS) and recurrence-free survival (RFS) compared with those without LVI (72.0 months vs. 81.8 months; 49.0 months vs. 66.6 months). Moreover, the prognostic impact of LVI appears to be greater in stage I patients than in those with stage II–III disease^[20].
- Therefore, accurate assessment of LVI in stage I LUAD patients offers valuable clinical reference, helping clinicians develop individualized treatment strategies—including the timing of surgery, choice of preoperative adjuvant therapy, and extent of resection. This, in turn, is expected to maximize patient prognosis and improve quality of life.
- In our univariate analysis, significant differences were observed between the LVI-positive and LVI-negative groups in several hematological markers, including HGB, TP, CEA, SCC, NSE, and CYFRA21-1. However, after selection by LASSO regression and multivariable logistic regression, only NSE was retained in the final prediction model. This

finding suggests that the independent contribution of conventional hematological markers to LVI prediction is limited, and their significance in univariate analysis may partly arise from confounding factors or collinearity among the markers.

- Serum tumor markers, as noninvasive laboratory indicators, have been demonstrated to reflect the degree of differentiation and malignancy of lung cancer. They are closely associated with tumor metastasis and hold certain predictive value in diagnosis, staging, prognostic assessment, and treatment monitoring^[21]. NSE is a neuron-specific isoenzyme of enolase, also known as enolase- γ , and has long been regarded as a marker of neuroendocrine tumors. Some studies have proposed that NSE is a key enzyme in glycolysis and plays an important role in aerobic glycolysis. Cells with high NSE expression show faster proliferation and are associated with poorer prognosis^[22–23]. The biological mechanism underlying elevated NSE in LVI of lung adenocarcinoma remains unclear. Some studies have investigated the mechanism of NSE in small cell lung cancer, and the results demonstrated that NSE activated the BMP2/Smad/ID1 signaling pathway by interacting with NBL1 and downregulating it, promoting the stem cell-like characteristics of SCLC cells, thereby promoting tumor proliferation, invasion, metastasis, tissue fibrosis, angiogenesis, and immune function^[24]. This may partly explain the potential involvement of NSE in LVI development in lung adenocarcinoma. In our study, NSE levels were significantly elevated in the LVI-positive group (median 11.02 vs. 10.13, $P = 0.005$) and retained independent predictive value in multivariable analysis ($OR = 1.093$, $P = 0.0063$), suggesting that NSE may serve as an indirect indicator of aggressive biological behavior in lung adenocarcinoma. However, it should be noted that NSE has relatively low clinical specificity and is susceptible to interference from hemolysis and neurological disorders. Moreover, the precise biological mechanism underlying its role in LVI prediction warrants further investigation.
- CEA, as the most classical serum tumor marker for lung adenocarcinoma, has been shown in previous studies to serve as a predictor of LVI^[25]. In our study, CEA showed a highly significant difference in univariate analysis ($P < 0.001$), but it was not retained in the final model. This may be attributed to strong collinearity between CEA and imaging variables such as maximum tumor diameter and the proportion of solid component. Larger tumor volume and greater solid component are associated with increased opportunities for CEA release into the bloodstream, and this overlap in information likely attenuated the independent predictive contribution of CEA^[26].
- Notably, the systemic inflammation-related composite indices that have received considerable attention in previous studies—such as SII, NLR, PLR, and PNI, showed no significant association with LVI in our study (all $P > 0.05$). This negative finding may reflect a relatively weaker role of inflammation-driven mechanisms in LVI formation in lung adenocarcinoma. Unlike certain gastrointestinal tumors, vascular invasion in early-stage lung adenocarcinoma may depend more on the intrinsic invasive properties of tumor cells, such as epithelial–mesenchymal transition capacity and matrix metalloproteinase secretion levels, rather than on host systemic inflammatory responses^[27–28]. Furthermore, all patients enrolled in this study had stage I disease with relatively low tumor burden, and the systemic inflammatory response may not have been fully activated. This may also explain why the composite inflammatory indices showed poor predictive performance.
- Notably, imaging features played a dominant role in our prediction model, with five of the six final predictors derived from CT assessment. This finding highlights the central value of preoperative CT in noninvasive LVI prediction.
- Among these, maximum tumor diameter is one of the most stable imaging predictors of LVI. In our multivariable analysis, each 1-cm increase in tumor diameter was associated with an approximately 39.5% increase in LVI risk ($OR = 1.395$). From a biological perspective, larger tumors tend to have longer growth cycles and more abundant angiogenesis, which may increase the opportunity for tumor cells to invade vessel walls and enter the circulatory system^[29]. Shimada et al. investigated the relationship between tumor diameter and LVI in patients with pT1-4N0-2M0 NSCLC and found that the probability of LVI varied with tumor size^[30]. Among patients with a maximum

diameter ≤ 5 cm, larger tumor diameter was associated with a higher likelihood of LVI. In addition, other studies have suggested that larger tumors are more prone to internal hypoxia, and upregulation of the hypoxia-inducible factor-1 α (HIF-1 α) pathway can promote vascular endothelial growth factor expression and alterations in vascular permeability, thereby creating favorable conditions for tumor cell invasion into blood and lymphatic vessels^[31].

- In our study, the presence of a solid component ratio > 0.5 showed a statistically significant difference between the LVI-positive and LVI-negative groups, which is consistent with the findings reported by Choe et al^[32]. Furthermore, it emerged as a strong independent predictor (OR = 1.940, P = 0.0062). In a study of 1,004 patients, Li et al. found that ground-glass opacity-predominant lung adenocarcinoma typically corresponds to the lepidic-predominant pathological subtype, which exhibits low invasiveness and rare vascular invasion. In contrast, an increased proportion of solid component often indicates the presence of acinar, papillary, or micropapillary growth patterns^[33]. The latter growth patterns have been confirmed by multiple studies to be closely associated with aggressive biological behavior and poor prognosis^[34]. In our study, 66.5% of LVI-positive patients had a CTR > 0.5 , compared with only 38.6% in the LVI-negative group, with a highly statistically significant difference. This further supports the clinical value of solid component ratio as a predictive marker for LVI.
- Spic, Vac, VasC, and PI were all significantly associated with LVI in univariate analysis (P < 0.05). However, after adjustment in multivariable analysis, PI did not remain in the final model. In a study by Onoda et al. involving 221 patients^[35], in a study by Onoda et al. involving 221 patients, it was found that pleural indentation (pleural tags) may be caused by a variety of factors, including local edema, tumor invasion, inflammation, and fibrosis, and its imaging features may overlap with those of Spic, leading to an insignificant difference in multivariable analysis. VasC refers to vascular structures being pulled toward the lesion, converging in the direction of the lesion, or being interrupted within the lesion or at its margin. This phenomenon is mainly attributable to the fact that malignant nodules promote microvascular growth in tumor tissue through vascular endothelial growth factor (VEGF) secreted by cancer cells, while also disrupting parts of the vessel wall^[36]. In our study, VasC (OR = 7.084) was the predictor with the greatest weight, with an odds ratio far exceeding those of the other variables. Long et al. in a clinical study involving 366 patients also reached the conclusion that VasC is an effective predictor of LVI^[36]. In clinical practice, the assessment of VasC is relatively straightforward and shows good reproducibility, making it a useful imaging clue for the early identification of patients at high risk of LVI.
- In our study, Spic and Vac also showed independent predictive value for LVI, which is consistent with previous findings^[25]. Some studies suggest that the pathological basis of Spic in malignant tumors lies in tumor cell infiltration and proliferation along blood vessels and lymphatic vessels into the adjacent stroma, which manifests on CT images as short and dense spiculated changes^[37].
- Therefore, for lung cancer patients with positive Spic, attention should be paid to the possible presence of LVI. In a study involving 1,683 patients and 2,125 tumor nodules, Liu et al. compared preinvasive lesions, minimally invasive adenocarcinoma (MIA), and invasive adenocarcinoma (IAC) and found that the incidence of bubble-like lucency increased significantly with increasing invasiveness (OR = 1.81). This finding also helps explain why the proportion of Vac was higher in the LVI-positive group than in the LVI-negative group in our study (24.1% vs. 13.6%)^[38].
- The nomogram developed in this study innovatively integrated NSE with five CT-based imaging features into an intuitive visualization tool. The model achieved an AUC of 0.841 in the training set and 0.786 in the validation set. The calibration curve closely approximated the ideal 45° reference line, and DCA demonstrated clinical net benefit across a threshold range of 5% to 80%. Collectively, these findings indicate that the model possesses good discriminative ability, satisfactory calibration, and promising clinical applicability. Compared with previous studies, a notable advantage of our model is that it relies solely on five morphological features clearly described in routine CT reports and a single serum marker, without the need for complex radiomic feature extraction or machine-learning modeling. This translates to lower implementation costs and stronger generalizability in clinical practice.

- Nevertheless, several limitations of this model should be acknowledged. First, this was a single-center retrospective study. Although the sample was representative to some extent, there remains a risk of selection bias, as randomly sampled cohorts may not fully reflect the target population. In addition, the model lacks external validation and the sample size is relatively limited. Future studies with a multicenter design and larger cohorts are warranted. Second, interpretation of imaging features is subject to the reader's experience, and inter-observer agreement for qualitative indicators such as VasC and Spic may vary. Future research could incorporate deep-learning techniques for automated feature extraction and quantification to further improve reproducibility. Finally, although the model showed net clinical benefit across a threshold range of approximately 5% to 80% on DCA, defining an acceptable risk threshold in clinical practice requires careful consideration of the trade-offs between benefits and harms of intervention.
- In summary, this nomogram, built on readily available preoperative indicators, provides a practical tool for treatment planning and prognostic evaluation in lung cancer patients. It not only has the potential to improve patient outcomes but also offers valuable reference for future clinical practice.

Abbreviations

NSCLC

non-small cell lung cancer

LUAD

lung adenocarcinoma,

LVI

lymphovascular invasion,

CEA

Carcinoembryonic Antigen

SII

Systemic Immune-Inflammation Index

PNI

Prognostic Nutritional Index

CT

Computed Tomography

MRI

Magnetic Resonance Imaging

NSE

Neuron-Specific Enolase

NLR

Neutrophil-to-Lymphocyte Ratio

d-NLR

DerivedNeutrophil-to-Lymphocyte Ratio

PLR

Platelet-to-Lymphocyte Ratio

LMR

Lymphocyte-to-MonocyteRatio

SIRI

Systemic Inflammation Response Index

CTR

Consolidation Tumor Ratio

Vac
Vacuolar sign
Lob
Lobulation sign
Spic
Spiculation sign
VasC
Vascular sign
PI
Pleural indentation
AUC
Area Under the Curve
LASSO Regression
Least Absolute Shrinkage and Selection Operator Regression,
ROC
Receiver Operating Characteristic
DCA
Decision Curve Analysis
ASIR
All age-standardized incidence ity rates
ASMR
All age-standardized mortal ity rates
OS
Overall Survival
RFS
Recurrence-Free Survival
HIF-1 α
Hypoxia-inducible factor-1 α
MIA
minimal invasive adenocarcinoma
IAC
invasive adenocarcinoma

Declarations

Ethics approval and consent to participate

The present study was approved by the Ethics Committee of Hebei General Hospital(2026-LW-291). Written informed consent was obtained from all surviving participants, as well as from the relatives who provided information on behalf of deceased participants.

Disclosures

The authors have no conflicts of interest to disclose.

Funding sources

The authors have no funding sources to disclose.

Author Contribution

Z and Q conceived the study and designed the methodology. Z, F, Y.L and D.L collected and verified the clinical data. Z. performed the statistical analysis and wrote the original manuscript. Q and F assisted with data processing and prepared Figures and Tables . W supervised the entire project, revised the manuscript critically, and is the corresponding author. All authors reviewed and approved the final manuscript.

Acknowledgement

We thank all the participating staff for their involvement with this project. ChatGPT and DeepSeek were used for language refinement, with the authors ensuring that the intended meaning was maintained.

Availability of data and materials

The dataset(s) supporting the conclusions of this article is(are) included within the article (and its additional file(s)).

References

1. Sung H, Filho AM, Laversanne M, et al. Global cancer statistics 2024: GLOBOCAN estimates of incidence and mortality worldwide for 34 cancers in 186 countries. *CA Cancer J Clin.* 2026;76(4):e70090.
2. Luo G, Zhang Y, Rumgay H, et al. Estimated worldwide variation and trends in incidence of lung cancer by histological subtype in 2022 and over time: a population-based study. *Lancet Respir Med.* 2025;13(4):348–63.
3. Schuchert MJ, Normolle DP, Awais O, et al. Factors influencing recurrence following anatomic lung resection for clinical stage I non-small cell lung cancer. *Lung Cancer.* 2019;128:145–51.
4. Potter AL, Costantino CL, Suliman RA, et al. Recurrence After Complete Resection for Non-Small Cell Lung Cancer in the National Lung Screening Trial. *Ann Thorac Surg.* 2023;116(4):684–92.
5. Nie P, Yang G, Wang N, et al. Additional value of metabolic parameters to PET/CT-based radiomics nomogram in predicting lymphovascular invasion and outcome in lung adenocarcinoma. *Eur J Nucl Med Mol Imaging.* 2021;48(1):217–30.
6. Locke M, Aung WY, Esposito M, Seetharamu N. Lymphovascular Invasion is a Predictor of Postoperative Recurrence or Death in Stage I Non-small-cell Lung Cancer (NSCLC). *Clin Lung Cancer.* 2025;26(7):e560–5.
7. Mollberg NM, Bennette C, Howell E, Backhus L, Devine B, Ferguson MK. Lymphovascular invasion as a prognostic indicator in stage I non-small cell lung cancer: a systematic review and meta-analysis. *Ann Thorac Surg.* 2014;97(3):965–71.
8. Riely GJ, Wood DE, Aisner DL, et al. Non-Small Cell Lung Cancer, Version 4.2026, NCCN Clinical Practice Guidelines In Oncology. *J Natl Compr Canc Netw.* 2026;24(4):e260017.
9. Kuo YS, Zheng MY, Shieh YS, Huang TW, Chou YT. Prognostic and Monitoring Utility of Serum CEA in Lung Adenocarcinoma: Differential Roles in EGFR-TKI and Chemotherapy Treatments. *Cancer Med.* 2025;14(17):e71170.

10. Zhang C, Yang C. Relationship between systemic immune-inflammation index and all-cause mortality in stages IIIB-IV epidermal growth factor receptor-mutated lung adenocarcinoma. *Front Endocrinol (Lausanne)*. 2025;16:1698317. Published 2025 Nov 5.
11. Li C, Wang Y, Liang S, et al. The Value of the Preoperative Hematological Inflammatory Markers for Predicting Lymphovascular Invasion in Gastric Cancer. *Immunotargets Ther*. 2026;15:564201. Published 2026 Jan 9.
12. Wei Y, Huang X, Pei W, Zhao Y, Liao H. MRI Features and Neutrophil-to-Lymphocyte Ratio (NLR)-Based Nomogram to Predict Prognosis of Microvascular Invasion-Negative Hepatocellular Carcinoma. *J Hepatocell Carcinoma*. 2025;12:275–87. Published 2025 Feb 15.
13. Tobing E, Tansol C, Tania C, Sihombing AT. Prognostic Nutritional Index (PNI) as Independent Predictor of Poor Survival in Prostate Cancer: A Systematic Review and Meta-Analysis. *Clin Genitourin Cancer*. 2024;22(5):102142.
14. Zhang Y, Sun Y, Zhang Q. Prognostic value of the systemic immune-inflammation index in patients with breast cancer: a meta-analysis. *Cancer Cell Int*. 2020;20:224. Published 2020 Jun 9.
15. Li M, Jin Y, Rui J, et al. Computed tomography-based radiomics for predicting lymphovascular invasion in rectal cancer. *Eur J Radiol*. 2022;146:110065.
16. Xu Z, Xie Y, Wu L, et al. Using Machine Learning Methods to Assess Lymphovascular Invasion and Survival in Breast Cancer: Performance of Combining Preoperative Clinical and MRI Characteristics. *J Magn Reson Imaging*. 2023;58(5):1580–9.
17. Li YM, Zhu YM, Gao LM, et al. Radiomic analysis based on multi-phase magnetic resonance imaging to predict preoperatively microvascular invasion in hepatocellular carcinoma. *World J Gastroenterol*. 2022;28(24):2733–47.
18. Ji Y, Zhang Y, Liu S, et al. The epidemiological landscape of lung cancer: current status, temporal trend and future projections based on the latest estimates from GLOBOCAN 2022. *J Natl Cancer Cent*. 2025;5(3):278–86. Published 2025 Feb 17.
19. Diao X, Guo C, Jin Y, et al. Cancer situation in China: an analysis based on the global epidemiological data released in 2024. *Cancer Commun (Lond)*. 2025;45(2):178–97.
20. Abolfathi H, Maranda-Robitaille M, Lamaze FC et al. Integrating Histologic Descriptors Into the Ninth Edition TNM Staging Improves Prognostic Stratification of Lung Adenocarcinoma. *J Thorac Oncol* Published online April 27, 2026.
21. Yang X, Xiao Y, Zhou Y, et al. Dynamic monitoring of serum tumor markers as prognostic factors in patients with advanced non-small-cell lung cancer treated with first-line immunotherapy: a multicenter retrospective study. *Ther Adv Med Oncol*. 2023;15:17588359231206282. Published 2023 Oct 31.
22. Ma Y, Yu C, Mohamed EM, et al. A causal link from ALK to hexokinase II overexpression and hyperactive glycolysis in EML4-ALK-positive lung cancer. *Oncogene*. 2016;35(47):6132–42.
23. Li S, Cao L, Wang X, Wang F, Wang L, Jiang R. Neuron-Specific Enolase Is an Independent Prognostic Factor in Resected Lung Adenocarcinoma Patients with Anaplastic Lymphoma Kinase Gene Rearrangements. *Med Sci Monit*. 2019;25:675–90. Published 2019 Jan 23.
24. Lu L, Zha Z, Zhang P, et al. Neuron-specific enolase promotes stem cell-like characteristics of small-cell lung cancer by downregulating NBL1 and activating the BMP2/Smad/ID1 pathway. *Oncogenesis*. 2022;11(1):21. Published 2022 Apr 29.
25. Lin M, Zhao X, Huang H, Lin H, Li K. A nomogram for predicting lymphovascular invasion in lung adenocarcinoma: a retrospective study. *BMC Pulm Med*. 2024;24(1):588. Published 2024 Nov 28.
26. Li Z, Wu W, Pan X, et al. Serum tumor markers level and their predictive values for solid and micropapillary components in lung adenocarcinoma. *Cancer Med*. 2022;11(14):2855–64.

27. Szalmás F, Nagy L, Bai P, Boratkó A. Loss of protein phosphatase 1 regulator TIMAP protein triggers EMT in A549 cells. *Cell Commun Signal*. 2026;24(1):321. Published 2026 Apr 13.
28. Vandooren J, Van den Steen PE, Opdenakker G. Biochemistry and molecular biology of gelatinase B or matrix metalloproteinase-9 (MMP-9): the next decade. *Crit Rev Biochem Mol Biol*. 2013;48(3):222–72.
29. Esteves M, Monteiro MP, Duarte JA. The effects of vascularization on tumor development: A systematic review and meta-analysis of pre-clinical studies. *Crit Rev Oncol Hematol*. 2021;159:103245.
30. Shimada Y, Saji H, Kato Y, et al. The Frequency and Prognostic Impact of Pathological Microscopic Vascular Invasion According to Tumor Size in Non-Small Cell Lung Cancer. *Chest*. 2016;149(3):775–85.
31. Finger EC, Giaccia AJ. Hypoxia, inflammation, and the tumor microenvironment in metastatic disease. *Cancer Metastasis Rev*. 2010;29(2):285–93.
32. Choe J, Lee SM, Park S, Choi S, Do KH, Seo JB. The prognostic value of lymphovascular invasion for stage I lung adenocarcinoma based on the presence of ground-glass opacity. *Eur Radiol*. 2025;35(4):2256–64.
33. Li R, Li Z, Yang Z, et al. The presence of micropapillary and/or solid subtypes is an independent prognostic factor for patients undergoing curative resection for stage I lung adenocarcinoma with ground-glass opacity. *Transl Lung Cancer Res*. 2024;13(2):256–68.
34. Kudo Y, Matsubayashi J, Saji H, et al. Association between high-resolution computed tomography findings and the IASLC/ATS/ERS classification of small lung adenocarcinomas in Japanese patients. *Lung Cancer*. 2015;90(1):47–54.
35. Onoda H, Higashi M, Murakami T, et al. Correlation between pleural tags on CT and visceral pleural invasion of peripheral lung cancer that does not appear touching the pleural surface. *Eur Radiol*. 2021;31(12):9022–9.
36. Long Y, Zhang H, Guo Y, et al. CT based quantification of intratumoral and peritumoral heterogeneity for diagnosing lymphovascular invasion for early stage non-small cell lung cancer. *BMC Med Imaging*. 2025;25(1):505. Published 2025 Dec 19.
37. Lin M, Zhao C, Huang H, et al. The clinical value of predicting lymphovascular invasion in patients with invasive lung adenocarcinoma based on the intratumoral and peritumoral CT radiomics models. *BMC Cancer*. 2025;25(1):1752. Published 2025 Nov 12.
38. Wan Q, Zou Q, Sun C, et al. CT-based Radiologic Ternary Classification Model in Predicting Pathologic Invasiveness of Pulmonary Nonsolid Nodules. *Radiology*. 2025;317(3):e251524.

Figures

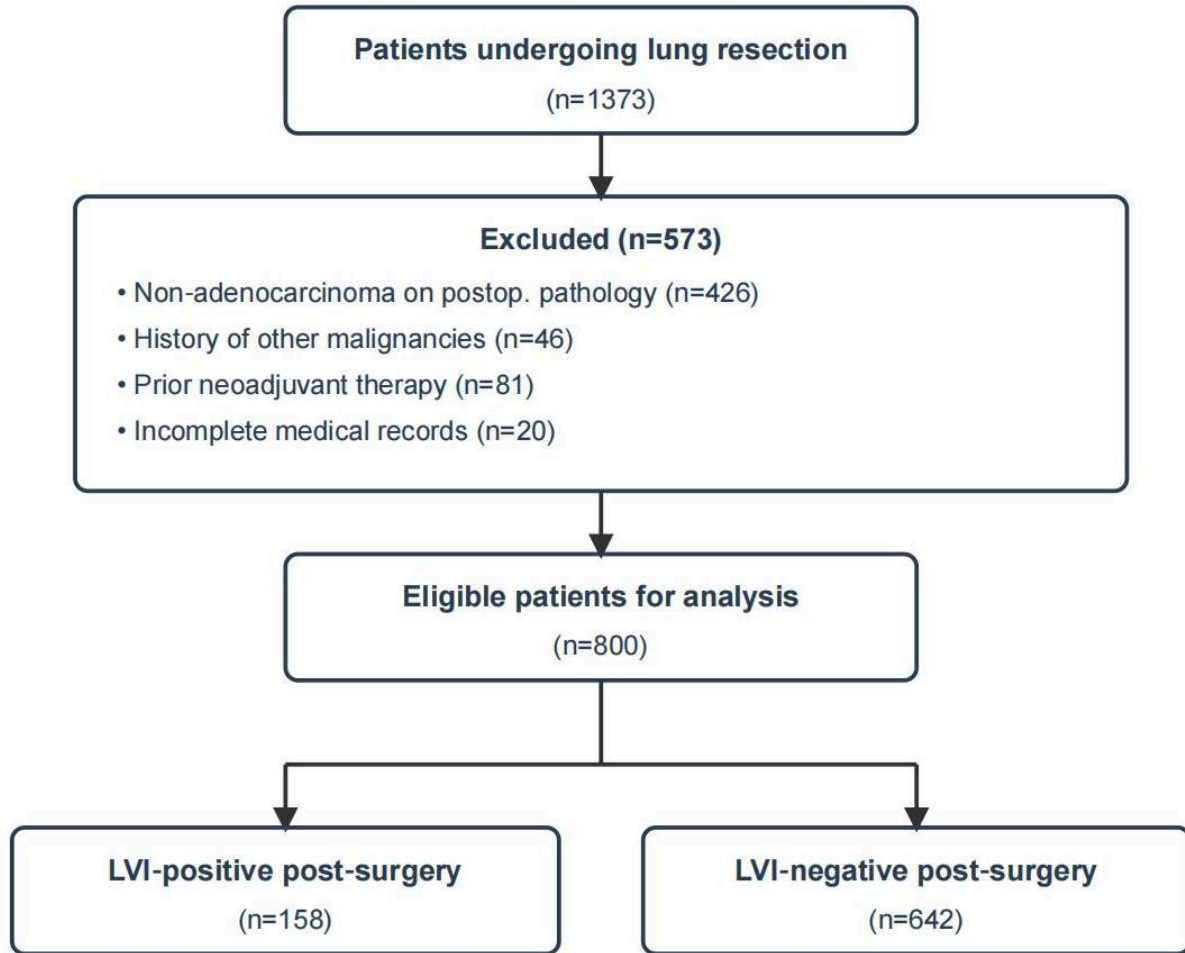


Figure 1:Flow diagram shows the patients who were included and excluded in this study

Figure 1

See image above for figure legend

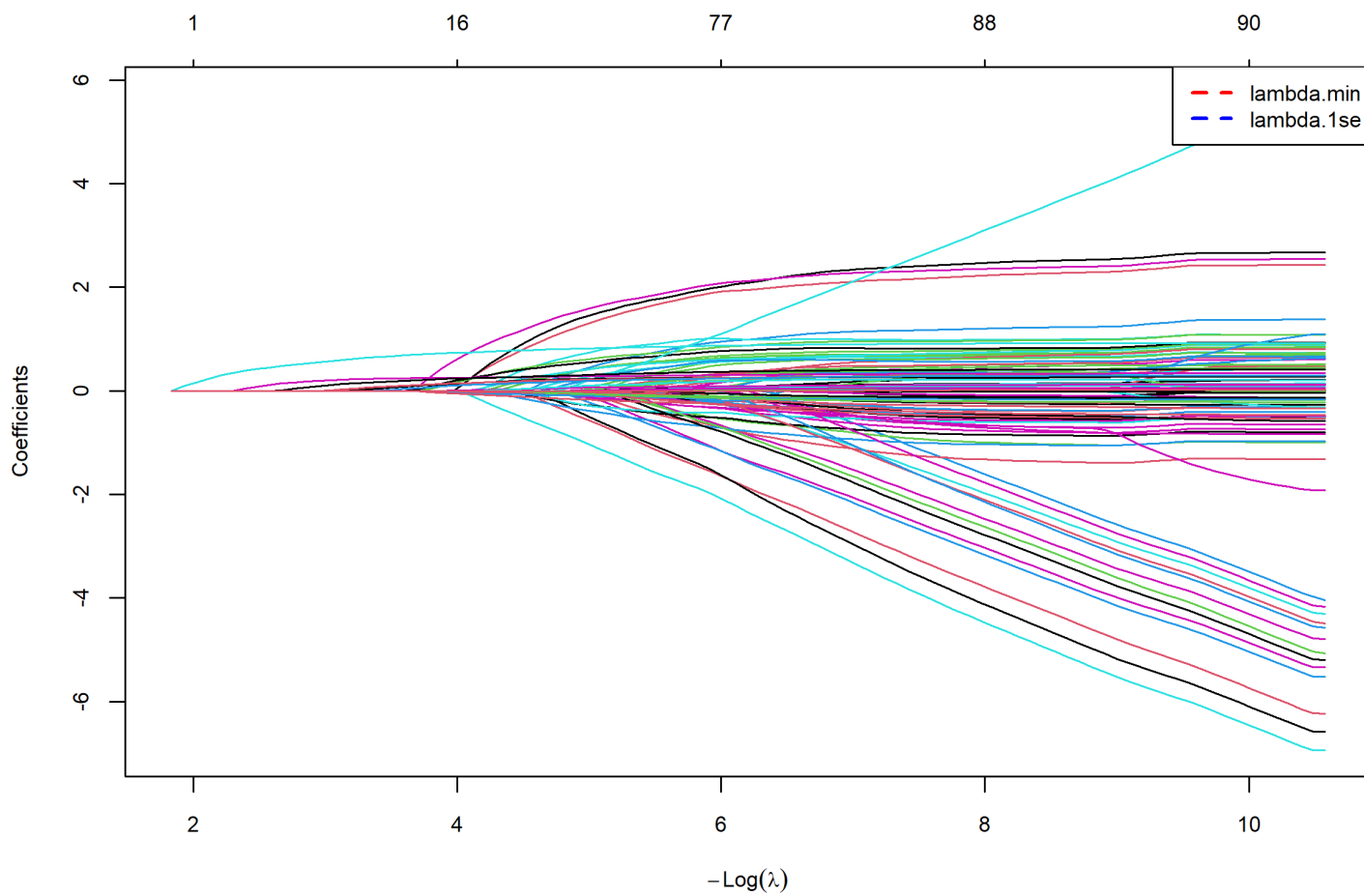


Figure.2 LASSO regression variable selection results.

Figure 2

See image above for figure legend

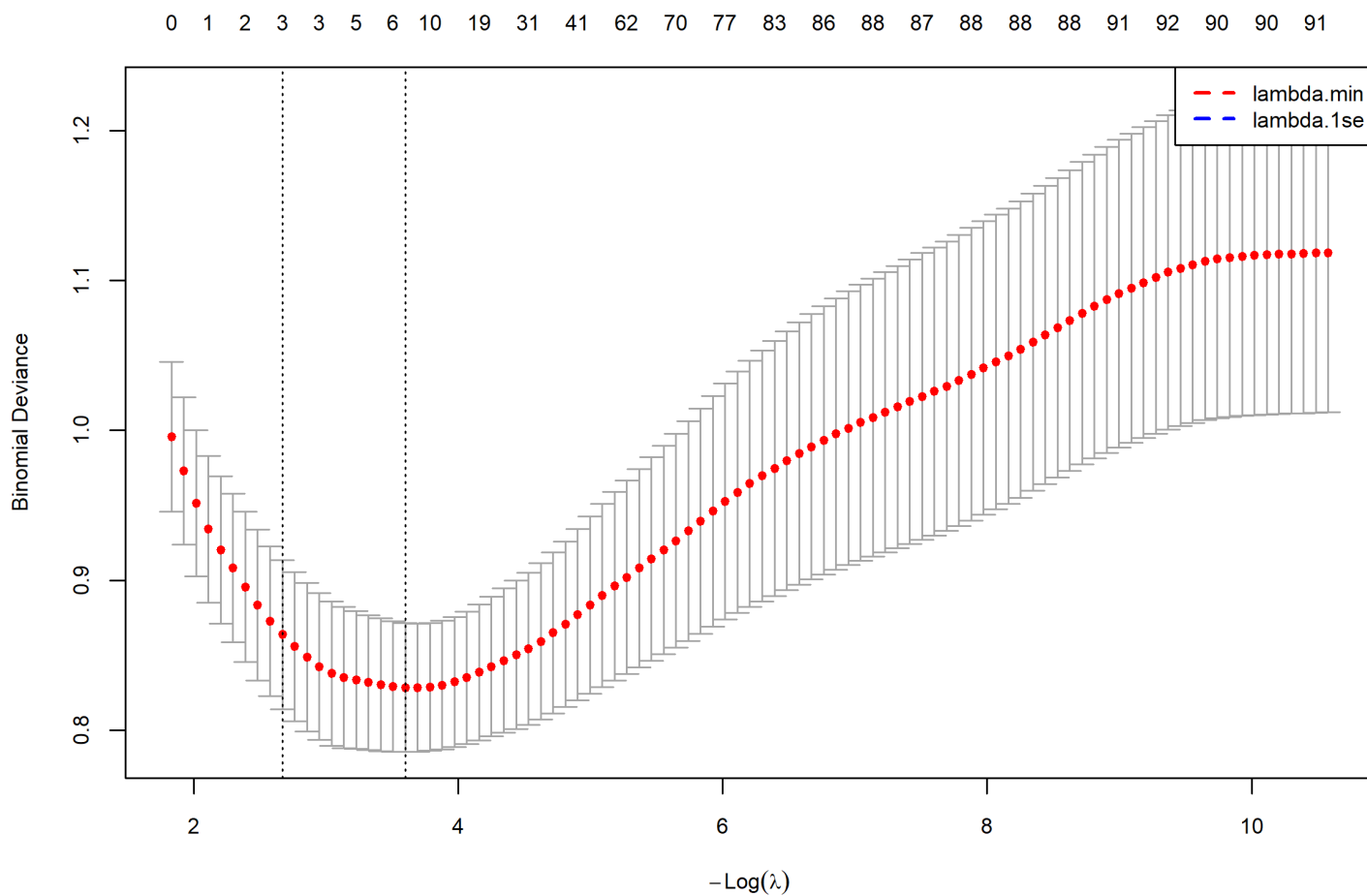


Figure.3 The process of selecting the most appropriate predictor variables in a LASSO model using ten-fold cross-validation

Figure 3

See image above for figure legend

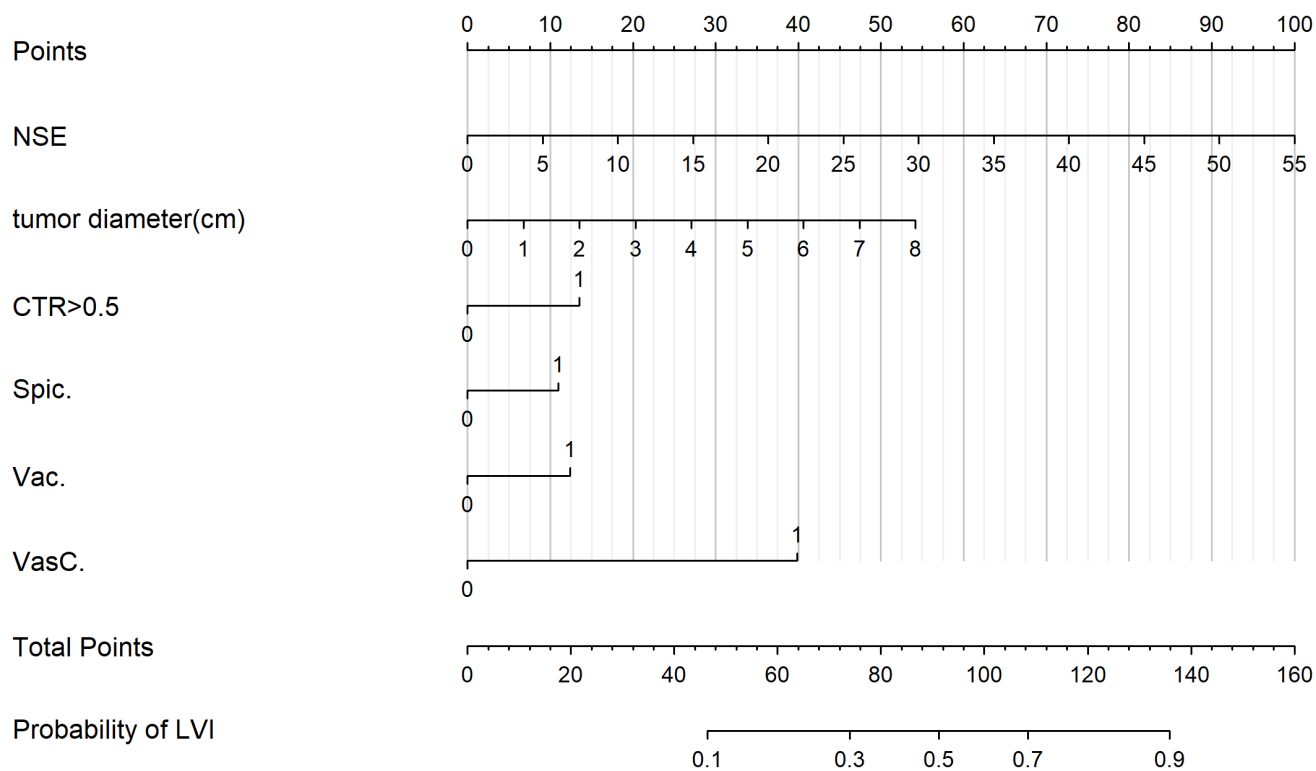


Figure.4 Nomogram Model for the Risk of Lymphovascular Invasion in stage I lung adenocarcinoma.

Figure 4

See image above for figure legend

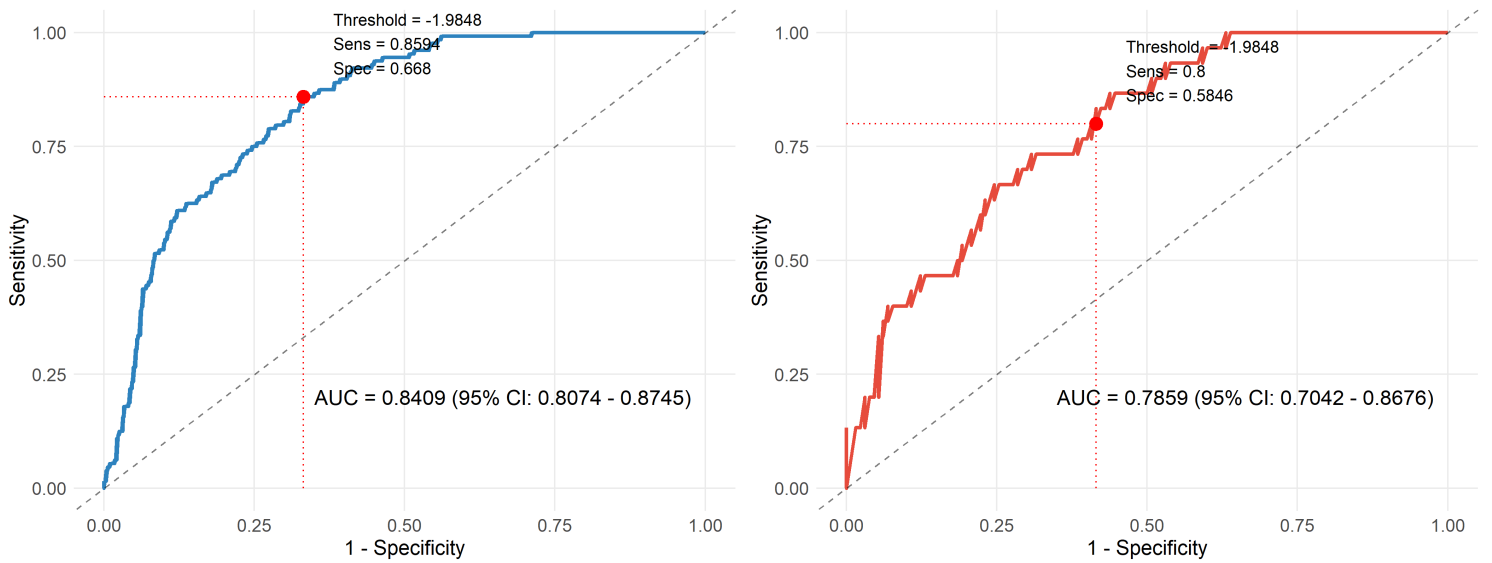


Figure.5 The ROC curves for the training set (left panel) and validation set (right panel).

Figure 5

See image above for figure legend

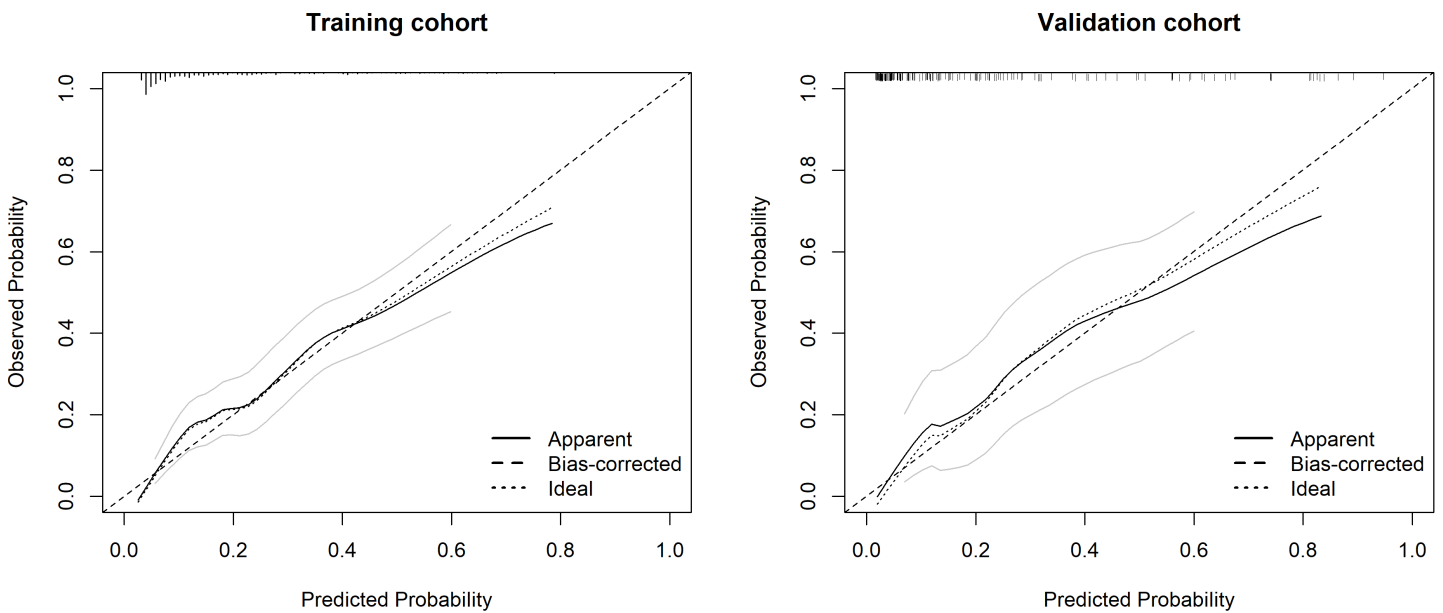


Figure.6 The calibration curve of the prediction model (left: training cohort; right: validation cohort)

Figure 6

See image above for figure legend

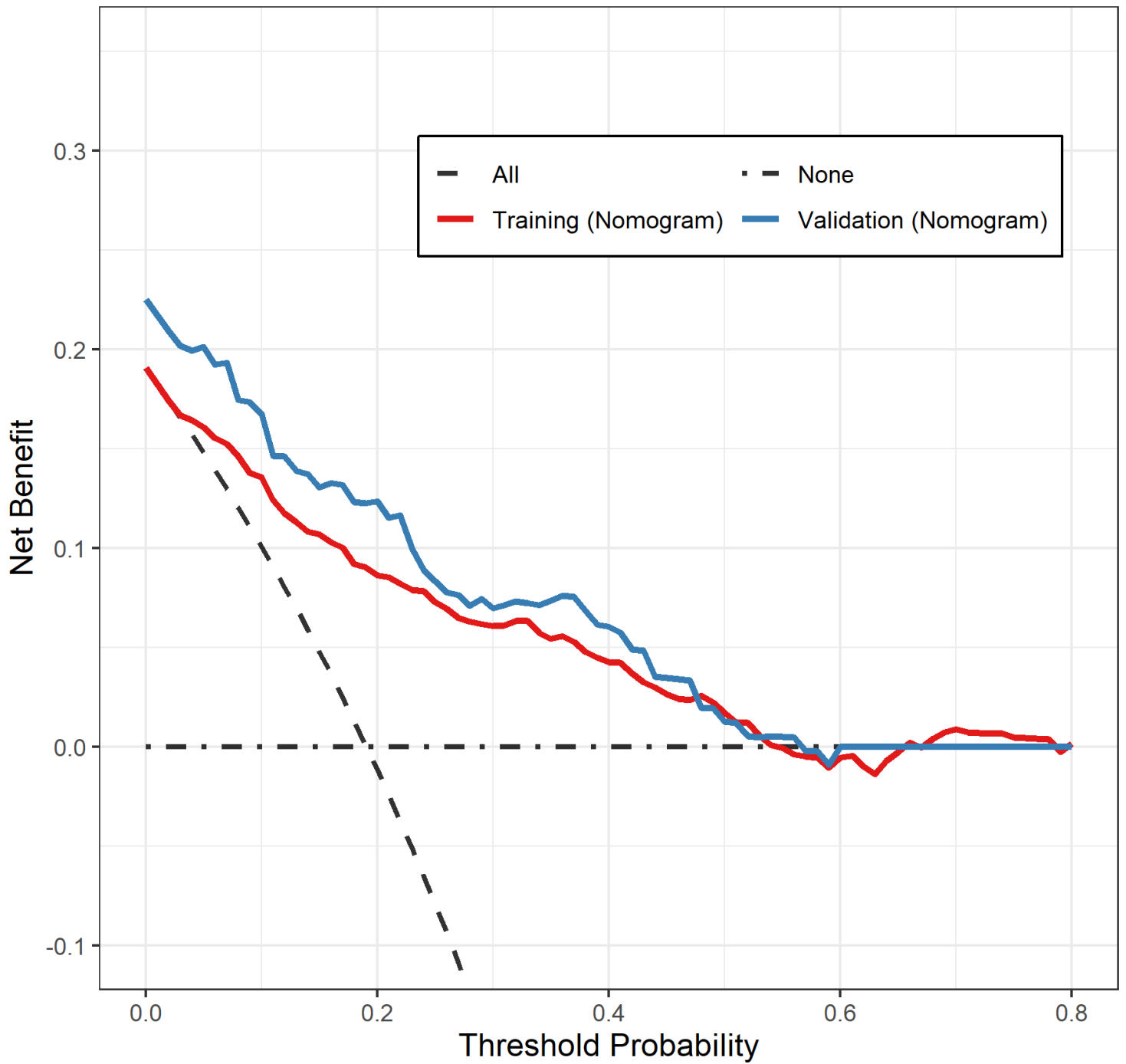


Figure.7 The clinical decision curve of the prediction model

Figure 7

See image above for figure legend

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [data.xlsx](#)