ORIGINAL RESEARCH

Establishment and Validation of a Predictive Nomogram for Postoperative Survival of Stage I Non-Small Cell Lung Cancer

Zhi-Hui Wang[1], Lili Deng[2]

[1]Department of Medical Oncology, The Fifth People's Hospital of Shenyang, Shenyang, People's Republic of China; [2]Department of Medical Oncology, The Second Affiliated Hospital of Harbin Medical University, Harbin, People's Republic of China

Correspondence: Lili Deng, Department of Medical Oncology, The Second Affiliated Hospital of Harbin Medical University, Harbin, People's Republic of China, Email dengdoctor@126.com

Background: Surgical procedure is the preferred option for people with early-stage non-small cell lung cancer (NSCLC), while nearly 30% of patients experienced metastatic or recurrent tumor after operation. The primary intention of this context is to summarize high-risk prognostic factors and set up a novel nomogram to predict the overall survival of individuals with stage I NSCLC after resection.

Methods: Research objects, 10,218 patients with stage I NSCLC after operation from 2010 to 2015, were extracted from the Surveillance, Epidemiology, and End Results (SEER) database. Independent prognostic factors, confirmed by Cox regression analyses, were integrated into a nomogram, to predict the 3-and 5-year overall survival of these individuals. The model experienced internal validation of testing cohorts above and external validation crewed by 160 patients from China. Finally, the nomogram was evaluated through several verification methods such as concordance index (C-index), calibration plots and receiver operating characteristic curve (ROC).

Results: Multivariate analysis identified that age, gender, histologic type, differentiation class, type of operation, T stage and treatment were significant predictive factors for the survival of stage I NSCLC. Based on these factors, a nomogram was constructed to predict the 3- and 5-year overall survival of these individuals. Meanwhile, in the training set, this nomogram displayed excellent superiority over the TNM staging system with abroad application, especially in C-index (0.669 vs 0.580) and the AUC (the Area Under ROC Curve) for the 3- and 5-year survival (0.678 vs 0.582; 0.650 vs 0.576). In the calibration curve, the curve representing predicted survival tended to align with the line representing actual survival as well.

Conclusion: A nomogram was successfully created and verified to achieve the goal that made a rounded accurate prediction on the survival of postoperative I NSCLC patients in terms of the SEER database.

Keywords: NSCLC, nomogram, prognosis, SEER, survival

Introduction

It is reported that the 5-year overall survival rate for persons with stage I NSCLC is up to 80%.[1] Surgery remains the accepted standard for stage I NSCLC. Unfortunately, people with stage I NSCLC may take a risk for recurrence or experience localized lesions die from NSCLC even if the tumor is completely resected via surgery. Locoregional recurrence rate is approximately 5–19% for stage I NSCLC.[2] Besides, long-term survival after recurrence remains not optimistic.[3] Thus, it often makes sense to accurately assess the risk associated with survival prognosis of patients, and monitor patients with high risks at regular intervals.

The main prognosis elements for stage I NSCLC, as reported in previous studies, included age, sex, smoking status, visceral pleural invasion, type of resection, histology subtype, and differentiation.[4–6] And there is no consensus on whether postoperative adjuvant therapy is beneficial for stage I NSCLC.

The traditional TNM staging has always been an important basis for prognosis. Nevertheless, there is some evidence supporting that patients, classified as having the same pathological or clinical staging of NSCLC, are likely to still have

diverse prognosis and ultimate ending. This result was largely due to the failure to in consideration of biological heterogeneity in patients with NSCLC.[7] Therefore, it is necessary to establish a comprehensive prognostic prediction model for multiple biomarkers to guide individualized treatment.

Recently, a nomogram, which has been regarded as a powerful tool for visual representation of prediction, could incorporate more comprehensive prognostic elements and give a more accurate prognostic evaluation to help guide follow-up.[8] Nomograms have been gradually developing in NSCLC in previous studies.[9–11] Yet, very few researchers poured concentration on predicting the prognostic nomogram of stage I NSCLC after resection. Thus, this paper managed to explore high-risk factors affecting survival and establish a nomogram to predict the outcome of stage I NSCLC patients after resection to guide the individualized survival prediction sufficiently.

Materials and Methods
Data Source
The Surveillance, Epidemiology, and End Results (SEER) database, composed mainly of 18 registries in U.S.A, is a large-scale data set that a variety of tumor-related information could be obtained publicly. Data enrolled in the study were filtered and downloaded from SEER*Stat Software (version 8.3.6.1). Detailed clinical and pathological information and prognosis of people with stage I NSCLC from 2010 to 2015 were accessible by logging in SEER*Stat with account 12,065-Nov2019. Referring to the above criteria, the external validation was made up of 160 patients diagnosed from 2010 to 2015 in the Second Affiliated Hospital of Harbin Medical University, Heilongjiang, China, which were approved by the ethics committee of the hospital. Corresponding to the data of the training set, more than 200 patients with I NSCLC who received surgery from 2010 to 2015 were selected from medical record system. Due to the small scale of the hospital, mainly rural residents and weak awareness of medical treatment, there is no large amount of data available for research, and only 160 patients were obtained in the end. Only patients who met the following requirement were included: 1) confirmed pulmonary malignant lung tumor and no history of other malignancies; 2) confirmed primary I NSCLC (tumor size ≤4cm) based on 8th TNM staging[12]; 3) no history of preoperative antitumor therapy; 4) history of pneumonectomy/extended pneumonectomy, lobectomy, wedge/sleeve lobectomy, segmental resection. Exclusion criteria were as follows: lack of complete clinical information; other pathological types except non-small cell lung cancer; preoperative radiotherapy and chemotherapy and other treatments; patients who died of causes unrelated to lung cancer.

Variables
The reliable and thorough information was provided with SEER as follows: age at diagnosis, sex, race, histologic type, differentiation grade, location, surgical type, TNM stage (AJCC, 7th ed.), tumor size, pleural/elastic layer invasion, radiation, chemotherapy, the reason for death, survival months, and vital status. Significantly, the data were rearranged with reference to the 8th edition TNM staging by integrating TNM data, tumor size and visceral pleural invasion (VPI) in the new research; histopathological type was classified into adenocarcinoma (ADC), squamous cell carcinoma (SC), and others (large cell carcinoma, adeno squamous carcinoma and so on). The variables available for external validation are age, gender, histologic type, differentiation grade, location, surgical type, TNM stage (8th ed.), VPI, radiation, chemotherapy, survival months, and current status. In addition, we eliminated individuals with a shortage of specific information about these characteristics. Overall survival (OS), namely the endpoint of current research, was started with the date of making a definite diagnosis and ended with death or last contact.

Construction and Validation of Nomogram Model
We collected a total of more than 1000,000 lung cancer data retrospectively. After deleting ineligible objects, we selected a total of more than 10,000 stage I non-small cell lung cancer data with complete variables. Finally, these data were taken as the research object. The eligible people were randomized in a training set and (70%, n = 7183) a validation set (30%, n = 3035) by a ratio of 7:3. Detailed sampling method and process are shown in Figure 1. A training set was conducted to build a nomogram. Categorical variables were described as count and percentage, while continuous variables were transformed into categorical variables according to the median. The Kaplan–Meier method and Log rank test were used for survival analysis,
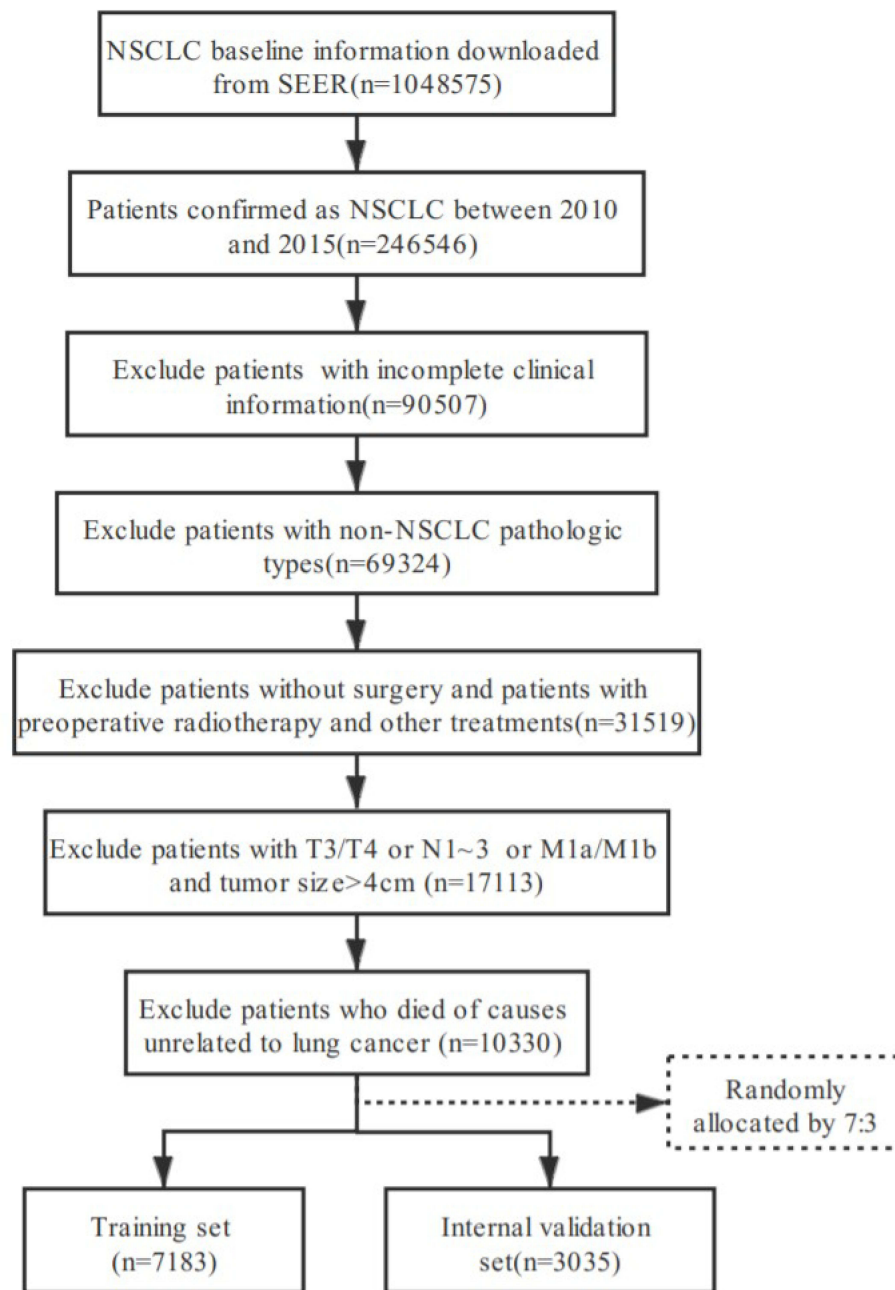
Figure 1 Flow diagram of the 1 NSCLC patients with training and validation sets.

while chi-square test was used for the comparison of categorical variables. Independent prognostic elements, affecting prognosis to different degrees, were confirmed by univariate and multivariate Cox regression analyses. Variables that were statistically significant in univariate analysis (p-value <0.05) were carried into a multivariate analysis subsequently. Ultimately, factors with P values <0.05 in both analyses became the key to establishing the model. The level of risk factors was measured by the hazard ratio (HR) and the corresponding 95% confidence interval (CI), both of which were figured out by the Cox proportional hazard regression model. All the processes of calculating and modeling were completed by R version 4.0.2 (http://www.r-progect.org).

Validation indicators generally included the Harrell's C-index, the calibration curve, and the receiver operating characteristic curve (ROC). The Harrell's C-index, range from 0.5 to 1, reflected various degrees of discrimination abilities of the nomogram. The larger the C-index, the better is its prognostic accuracy. Similarly, the predictive accuracy of the nomogram model was demonstrated by the value of area under curve (AUC). The value of AUC denoted the size of the area under the ROC curve. And a larger AUC represents better performance in general. Additionally, calibration plots, regarding a 45-degree line as an optimal model, served to value the model by a comparison of predicted 3- and 5-year survival and actual 3- and 5-year survival. Furthermore, the nomogram was subjected to 1000 bootstrap resamples for the verification of the training cohort. R software (version 4.0.2), as well as its installed packages of survival and rms, was used for analysis.

Results

Patients Characteristics

Training (n = 7183) cohorts, consisted of 70% of the selected persons from the SEER database, were mainly carried out to create the nomogram. Meanwhile, the remaining 30% as internal validation (n = 3035) cohorts, were responsible for validating the model. An external validation was composed of 160 sufferers with stage I NSCLC diagnosed in the Second Affiliated Hospital of Harbin Medical University from 2010 to 2015. The clinicopathological characteristics of all cases in the training cohorts and the validation cohorts are clearly presented in Table 1. In the training group, the 3- and 5-year OS rates were 84.3% and 76.3%, respectively. In the internal validation cohort, the 3- and 5-year OS rates were 85.1% and 75.7%, respectively. In addition, the 3- and 5-year OS rates of the external cohort were 66.1% and 60.4%, respectively.

Cox Regression Analysis

In the univariable analysis, age, gender, histopathological type, differentiation grade, location, type of surgery, T stage and treatment were highly correlated with overall survival. There seemed to be no difference in the race of patients and visceral pleural invasion (VPI) in prognosis in the univariable analysis. Variables with $p < 0.005$ in univariate analysis were involved in the multivariate analysis. Multivariable analysis suggested that age, gender, histopathological type, differentiation grade, type of surgery, T stage and treatment were also verified to be independent prognostic factors for OS. The specific results of Cox regression analysis are exhibited in Table 2.

According to the analysis results, male patients who were >70 years old, diagnosed with squamous cell carcinoma, were associated with poor prognosis. The lower the tumor grade and the later the pathological stage, the worse the prognosis. Lobectomy had the best prognosis compared to other procedures in stage I NSCLC. Patients without postoperative adjuvant therapy had a relatively good prognosis.

Construction and Validation of Nomogram

Significant independent factors, including age, gender, histopathological type, differentiation class, type of surgery, T stage and treatment, were selected to establish a nomogram predicting 3-year or 5-year overall survival rates. To determine the survival rates that the total score represented, it was necessary to figure out the scores of each of these variables and add them up to get the total score in the nomogram. A vertical line was drawn from the calculated sum value, to obtain the corresponding estimated probability of survival (Figure 2).

The C-index calculated from the training group was 0.669 (95% CI: 0.651–0.687), which indicated discrimination ability, relatively higher than it in IASLC 8th edition TNM staging system (0.580 95% CI: 0.562–0.598). Similarly, the C-indices (internal set 0.676,95% CI: 0.649–0.703; external set 0.733, 95% CI: 0.674–0.792) were considerably greater than the outcomes of TNM category (internal set 0.582,95% CI: 0.555–0.609; external set 0.576, 95% CI: 0.515–0.637) in both internal and external validation groups.

As shown in the training cohort, the area under curves (AUCs) of the nomogram for predicting the 3- and 5-year overall survival rates were 0.678 and 0.650, respectively, while the AUCs of the TNM staging for predicting the 3- and 5-year overall survival rates were 0.582 and 0.576, respectively. In the internal validation cohort, the area under curves (AUCs) of the nomogram for predicting the 3- and 5-year overall survival rates were 0.689 and 0.647, respectively, while

Table 1 Demographics and Clinicopathologic Characteristics in Different Sets

Characteristics	Training Set (N, %)	Internal Validation Set (N, %)	External Validation Set (N, %)
Age			
<60	1383(19.2)	553(18.2)	71(44.3)
60–69	2578(35.9)	1141(37.6)	62(38.8)
≥70	3222(44.9)	1341(44.2)	27(16.9)
Gender			
Female	3968(55.2)	1659(54.7)	74(46.2)
Male	3215(44.7)	1376(45.3)	86(53.8)
Race			
Black	5930(82.6)	2531(83.4)	
White	614(8.5)	244(8.0)	
Others	639(8.9)	260(8.6)	
Histologic type			
ADC	4883(68.0)	2050(67.5)	103(64.4)
SC	2028(28.2)	869(28.6)	47(29.4)
Others	272(3.8)	116(3.8)	10(6.2)
Differentiation grade			
Grade I	1246(17.3)	524(17.2)	21(13.1)
Grade II	3640(50.7)	1522(50.1)	86(53.8)
Grade III+IV	2297(32.0)	989(32.7)	53(33.1)
Location			
Left lower	946(13.2)	391(12.9)	21(13.1)
Left upper	1882(26.2)	851(28.4)	50(31.2)
Right lower	1137(15.8)	503(16.6)	22(13.8)
Right middle	369(5.1)	150(4.9)	19(11.9)
Right upper	2708(37.7)	1088(35.8)	47(29.4)
Other	141(2.0)	52(1.7)	1(0.6)
Surgery			
Lobectomy	5636(78.5)	2388(78.7)	130(81.2)
Pneumonectomy	59(0.8)	27(0.9)	4(2.5)
Wedge resection	1471(20.5)	614(20.2)	23(14.4)
Segmental Resection	17(0.2)	6(0.2)	3(1.9)
T			
T1a	472(6.6)	209(6.9)	13(8.1)
T1b	3027(42.1)	1258(41.4)	65(40.6)
T1c	2454(34.2)	1034(34.1)	54(33.8)
T2a	1230(17.1)	534(17.6)	28(17.5)
VPI			
No	6694(93.2)	2824(93.0)	147(91.9)
Yes	489(6.8)	211(6.9)	13(8.1)
Treatment			
No/unknown	6644(92.5)	2822(92.9)	151(94.4)
Chemo only	445(6.2)	182(5.9)	7(4.4)
Chemo+Rad	94(1.3)	31(1.0)	2(1.2)

Abbreviations: ADC, adenocarcinoma; SC, squamous carcinoma; Others large cell carcinoma, adeno squamous carcinoma and so on; Grade I, well differentiated; Grade II, moderately differentiated; Grade III, poorly differentiated; Grade IV, undifferentiated; VPI, visceral pleural invasion; Chemo, chemotherapy; Rad, radiotherapy.

the AUCs of the TNM staging for predicting the 3- and 5-year overall survival rates were 0.575 and 0.552, respectively. In the external validation cohort, the area under curves (AUCs) of the nomogram for predicting the 3- and 5-year overall survival rates were 0.782 and 0.775, respectively, while the AUCs of the TNM staging for predicting the 3- and 5-year overall survival rates were 0.588 and 0.594, respectively. Higher area under receiver operating characteristic curve

Table 2 Univariate and Multivariate Cox Analyses for I NSCLC

Variable	Univariate Analyses			Multivariate Analyses		
	HR	95% CI	P	HR	95% CI	P
Age						
<60	Reference			Reference		
60–69	1.16	0.96–1.41	0.125	1.11	0.91–1.35	0.283
≥70	1.63	1.36–1.96	<0.001	1.50	1.25–1.82	<0.001
Sex						
Male	Reference			Reference		
Female	0.70	0.61–0.79	<0.001	0.75	0.66–0.85	<0.001
Race						
Black	Reference					
White	1.01	0.81–1.26	0.925			
Others	0.73	0.53–1.01	0.059			
Histology						
SC	Reference			Reference		
ADC	0.63	0.55–0.72	<0.001	0.85	0.74–0.98	0.021
Others	1.04	0.78–1.37	0.793	1.03	0.78–1.37	0.815
Grade						
I	Reference			Reference		
II	2.63	2.03–3.40	<0.001	2.27	1.75–2.96	<0.001
III+IV	3.49	2.67–4.53		2.80	2.13–3.66	<0.001
Location						
Left upper	Reference			Reference		
Left lower	0.91	0.73–1.13	0.387	0.94	0.76–1.17	0.600
Right upper	0.86	0.73–1.01	0.067	0.92	0.78–1.08	0.289
Right middle	1.11	0.83–1.48	0.474	1.29	0.97–1.72	0.082
Right lower	1.09	0.90–1.32	0.371	1.19	0.98–1.44	0.079
Others	1.49	1.02–2.17	0.040	1.44	0.98–2.12	0.067
Surgery						
Lobectomy	Reference			Reference		
Pneumonectomy	2.14	1.28–3.57	0.003	1.55	0.91–2.64	0.105
Wedge resection	1.75	1.52–2.01	<0.001	1.92	1.66–2.22	<0.001
Segmental resection	3.02	1.35–6.75	0.007	2.98	1.13–6.68	0.008
T						
T1a	Reference			Reference		
T1b	1.60	1.13–2.27	0.007	1.65	1.16–2.34	0.005
T1c	2.15	1.52–3.04	<0.001	2.21	1.55–3.14	<0.001
T2a	2.91	2.04–4.14	<0.001	2.86	1.99–4.12	<0.001
VPI						
Yes	Reference					
No	0.99	0.77–1.30	0.980			
Treatment						
No/unknown	Reference			Reference		
Chemo only	1.26	0.99–1.59	0.057	1.06	0.83–1.34	0.640
Chemo+Rad	2.35	1.62–3.40	<0.001	1.71	1.17–2.49	0.005

Abbreviations: HR, hazard ratio; CI, confidence interval.

(ROC) curve, a better discriminative ability. Similar results were observed for both internal and external validation cohorts, respectively (Figure 3).

In addition, the calibration curve did not deviate significantly by 1000 bootstrap resamples. Calibration plots for the survival rate of 3- or 5-year after surgery performed good consistency between the prognostication by our model and
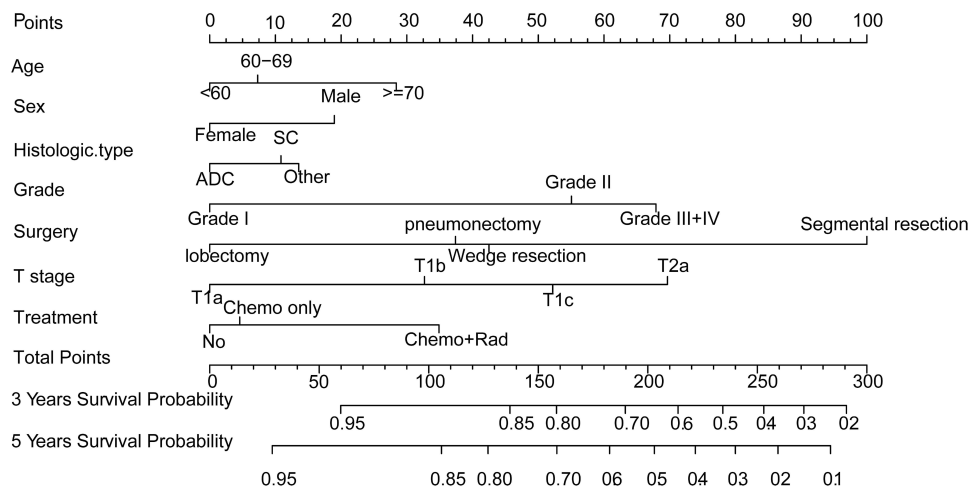
Figure 2 Nomogram for predicting 3-year and 5-year overall survival rates. The nomogram adopts the principle of score accumulation, and each factor corresponds to the top "Points" to get a score. The total score "Total Points" obtained by adding the scores of each factor corresponds to "3-years Survival Probability" and "5-years Survival Probability" to get the patient's 3-year or 5-year survival rate. ADC adenocarcinoma; SC squamous carcinoma; Other large cell carcinoma, adeno squamous carcinoma and so on; Chemo chemotherapy; Rad radiotherapy.

practical observation, indicating that the predicted survival rate was in high agreement with the actual survival rate of patients in both training sets and validation sets (Figure 4).

Discussion

In this study, we confirmed that age, gender, histologic type, differentiation class, surgical type, T stage and treatment were strongly correlated with prognostic in the OS rates of sufferers with postoperative stage I NSCLC. On this basis, a nomogram, for patients with postoperative stage I NSCLC covered those prognostic factors above, was constructed and verified, which proved to make a better assessment of the prognosis than traditional TNM staging. This nomogram could estimate the prognosis of the patients clearly, screen out patients with short survival and make personalized decision with respect to the surveillance and therapy effectively.

A previous study shown that age did not determine survival of I NSCLC.[13] According to our study, the risk of an unfavorable prognosis increased with age, which was more evident in males than in females. Male patients with advanced age were revealed a poor prognosis in I NSCLC, which was consistent with previous studies.[14,15] Thus, bad prognosis may be due to poor tolerance of elderly patients and weak willingness to treat.[16] Male patients generally had a poor prognosis probably because the peak time of recurrence of men appeared earlier than that of women.[17]

Currently, there is still a debate on which pathologic type has better survival. A retrospective analysis demonstrated stage I squamous patients had a high risk of local disease recurrence.[18] Similarly, based on our results, persons with squamous cell carcinoma performed a higher risk of cancer recurrence and a poorer prognosis than adenocarcinoma. However, other studies reported that adenocarcinoma, as a poor prognostic determinant in resected lung cancer, appeared to have a higher recurrence rate than non-adenocarcinoma.[19,20] It might be explained that for postoperative patients of stage I, adenocarcinoma was more likely to have distant metastasis, while squamous cell carcinoma was more likely to have local recurrence.[21]

The T stage has always been an independent risk influence affecting prognosis after stage I NSCLC.[1,11] According to our nomogram, the greater the tumor, the worse the prognosis in I NSCLC. The 5-year survival probability of patients with tumor size <or= 2.0cm was higher than that of patients with tumors >2cm by CT screening in stage IA NSCLC.[22] In the 8th ed TNM staging, VPI was suggested as one of the important factors influencing upgrading of T1 to T2 stage.[23] However, several studies had identified that VPI was an independent risk factor with unfavorable prognosis in NSCLC patients.[24,25] Most of these researches covered each stage of NSCLC; however, the research object of this article limited to I NSCLC. In this study, the factor of VPI did not reach statistical significance, which manifested it did not affect the prognosis.
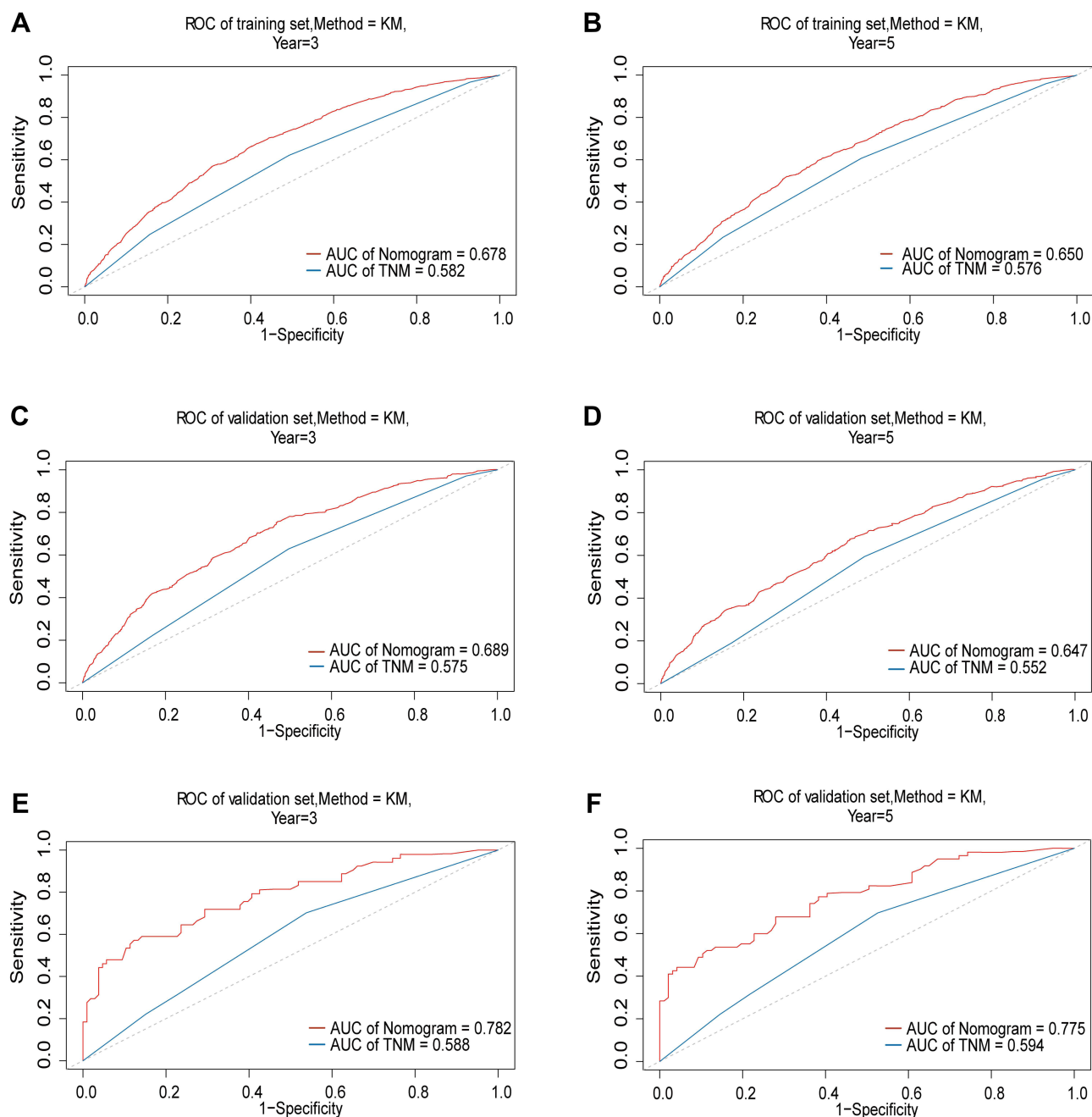
Figure 3 ROC curves of the Nomogram and the TNM stage system for predicting 3-, and 5-year overall survival (OS) (**A** and **B**) ROC curves in the training cohort; (**C** and **D**) ROC curves in the internal validation cohort; (**E** and **F**) ROC curves in the external validation cohort.
Abbreviations: AUC, Area under curve; OS, Overall survival; ROC, Receiver operating characteristic curve.

Our study showed that lobectomy had a better effect on the long-term prognosis of stage I NSCLC than wedge resection and segmental resection, which was notably in line with previous findings. A randomized controlled trial initiated by The Lung Cancer Study Group proved that compared with lobectomy, total mortality and locoregional recurrence rate of limited pulmonary resection (segment or wedge) were significantly increased.[26] Segmental resection, featured fewer complications and a shorter treatment period in hospital, was more likely to occur regional recurrences in patients with stage I NSCLC.[27] Similar studies had confirmed that wedge resection was inferior to lobectomy with mediastinal lymph node dissection in stage I adenocarcinoma.[28] Somewhat differently, a context had found that
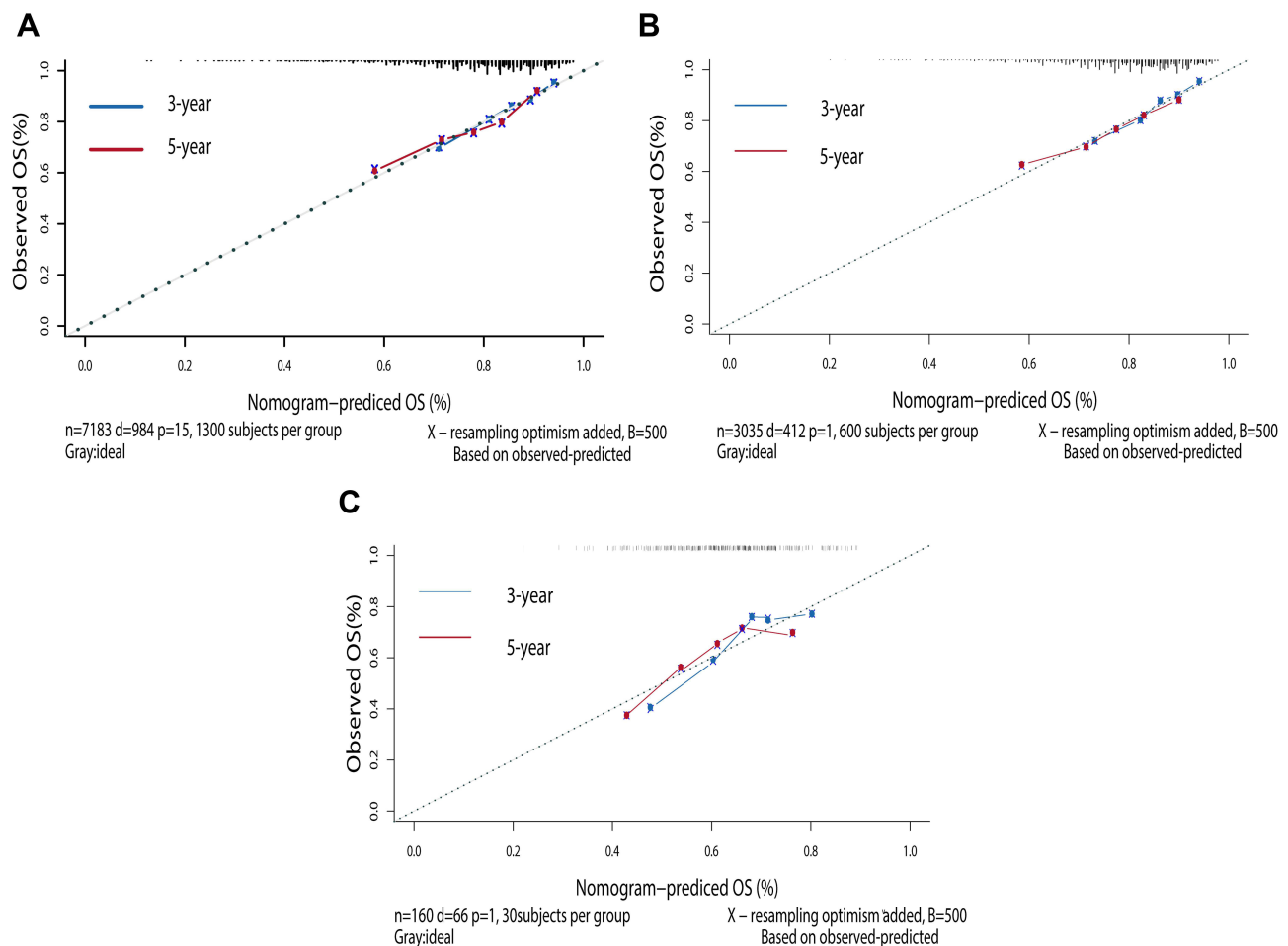
Figure 4 Calibration plots of the nomogram for predicting 3-and 5-year OS rates in stage I NSCLC patients. Calibration plots show the relationship between the predicted survival by the model and actual outcomes of training cohort (**A**), internal validation cohort (**B**) and external validation cohort (**C**).

lobectomy and pulmonary segmentectomy were nearly identical in prognosis for patients with tumor size ≤ 3cm, while lobectomy had the prognostic superiority in patients with >3 and ≤4cm.[28] Therefore, which type of surgery should be adopted for stage I NSCLC required a further concrete analysis.

Whether patients with stage I NSCLC need postoperative adjuvant chemotherapy or not has not a certain conclusion. We identified that adjuvant therapy did not benefit stage I NSCLC. Previous studies also have reported postoperative chemotherapy did not significantly improve the OS of stage IA NSCLC, which suggested that adjuvant chemotherapy may not be necessary after stage IA NSCLC surgery.[29] Suffers with stage IB NSCLC have benefited little from adjuvant chemotherapy.[30] It was not recommended to give adjuvant chemotherapy for stage IB NSCLC.[31] Given the potential of heart and lung damage, the national comprehensive cancer network (NCCN) guidelines did not recommend I, II NSCLC patients received postoperative radiotherapy.[32] This was consistent with our point of view.

Up to now, nomograms have been established for some of the cancers. However, far too little attention has been paid to predicting the prognostic nomogram exclusively target at stage I NSCLC persons underwent the surgical procedure. A previous study, with the goal of establishing a nomogram for predicting survival in individuals with resected NSCLC, covered every stages of NSCLC.[33] It was obviously inaccurate enough if applied to the prognosis of postoperative I NSCLC. Zhang et al[34] constructed a relatively incomprehensive nomogram with only four prognostic predictors for early NSCLC. A nomogram was created, as previously described, to predict Cancer-Specific Survival (CSS) for 8th edition stage I NSCLC. Unfortunately, they lacked of postoperative information.[11]

As the results show, our proposed nomogram is more robust and accurate than the existing TNM staging system in both training cohort and validation cohort. In the training set and the verification set, the C-indices of the nomogram prediction model were 0.669 (95% CI: 0.651–0.687), 0.676 (95% CI: 0.649–0.703) and 0.733 (95% CI: 0.674–0.792), respectively. All of them were higher than the C indices of TNM staging, which is 0.580 (95% CI: 0.562–0.598), 0.582 (95% CI: 0.555–0.609) and 0.576 (95% CI: 0.515–0.637), respectively. The 3-year and 5-year AUC of the prediction model performed more excellent, which also confirms that the newly established prediction model has higher accuracy than the traditional TNM staging model (training set: 0.678 vs 0.582; 0.650 vs 0.576, internal set: 0.689 vs 0.575; 0.647 vs 0.552, external set: 0.782 vs 0.588; 0.775 vs 0.594). The external validation revealed that the model was suitable to not merely western patients but also Asian patients with stage I NSCLC. In addition, there was a good consistency between the predictions and actual observations in the calibration plots in three cohorts, which guaranteed the validity and dependability of the instituted nomogram. Stability and reliability of the freshly prognostic model were revealed adequately by these observations.

Meanwhile, there remain several restrictions that ought to take into consideration despite the nomogram in this work demonstrated the abilities of predicting OS accurately. Firstly, the SEER database failed to cover more information such as specific chemotherapy regimens, recurrence-related data and other factors that might affect the prognosis of stage I NSCLC. Another disadvantage was small sample size for external validation. Finally, this nomogram followed a retrospective study design; as a result, larger sample size prospective clinical experiments would be demanded to further independent validation and prognostic stratification analysis in the future. In spite of these deficiencies, the newly created model is convenient for clinical manipulation to judge the predicted survival in stage I NSCLC cases.

Conclusion

To sum up, the present study identified age, sex, histologic type, differentiation grade, type of surgery, T stage and treatment as independent components determined medical prognosis for the OS of persons who suffered stage I NSCLC after surgery. Refer to our calculated proportion, these factors structured the nomogram for prediction, intended for forecasting the 3- and 5-year survival rates of persons with stage I NSCLC after resection and guiding individualized treatment accurately and stably.

Ethics Statement

The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Ethics Committee of the Second Affiliated Hospital of Harbin Medical University (Approval YY2020-016). Given the retrospective nature and anonymity of the present study, no informed consent was required. All patients' identifiable information was hidden, and anyone's identity cannot be deduced from the context.

Acknowledgments

The authors would like to thank SEER for open access to the database.

Funding

Disclosure

Zhi-Hui Wang is now affiliated with Hegang Hekuang Hospital. The authors report no conflicts of interest in this work.

References

1. Chansky K, Detterbeck FC, Nicholson AG, et al. The IASLC lung cancer staging project: external validation of the revision of the TNM stage groupings in the eighth edition of the TNM classification of lung cancer. *J Thorac Oncol*. 2017;7:1109–1121. doi:10.1016/j.jtho.2017.04.011
2. Mahvi DA, Liu R, Grinstaff MW, Colson YL, Raut CP. Local cancer recurrence: the realities, challenges, and opportunities for new therapies. *CA Cancer J Clin*. 2018;6:488–505. doi:10.3322/caac.21498
3. Sonoda D, Matsuura Y, Kondo Y, et al. Characteristics of surgically resected non-small cell lung cancer patients with post-recurrence cure. *Thorac Cancer*. 2020;11:3280–3288. doi:10.1111/1759-7714.13669

4. Kosaka T, Shimizu K, Nakazawa S, et al. Clinicopathological features of small-sized peripheral squamous cell lung cancer. *Mol Clin Oncol*. 2020;1:69–74. doi:10.3892/mco.2019.1951

5. Wang C, Wu Y, Shao J, Liu D, Li W. Clinicopathological variables influencing overall survival, recurrence and post-recurrence survival in resected stage I non-small-cell lung cancer. *BMC Cancer*. 2020;1:150. doi:10.1186/s12885-020-6621-1

6. Wang Y, Zheng D, Zheng J, et al. Predictors of recurrence and survival of pathological T1N0M0 invasive adenocarcinoma following lobectomy. *J Cancer Res Clin Oncol*. 2018;6:1015–1023. doi:10.1007/s00432-018-2622-8

7. Ettinger DS, Wood DE, Aggarwal C, et al. NCCN guidelines insights: non-small cell lung cancer, version 1.2020. *J Natl Compr Canc Netw*. 2019;12:1464–1472. doi:10.6004/jnccn.2019.0059

8. Li G, Lian L, Huang S, et al. Nomograms to predict 2-year overall survival and advanced schistosomiasis-specific survival after discharge: a competing risk analysis. *J Transl Med*. 2020;1:187. doi:10.1186/s12967-020-02353-5

9. Li H, Wang Z, Yang F, Wang J. Development and validation of a nomogram for predicting cancer-specific survival of surgical resected stage I-II adenosquamous carcinoma of the lung. *J Surg Oncol*. 2020;6:1027–1035. doi:10.1002/jso.25858

10. Wo Y, Yang H, Zhang Y, Wo J. Development and external validation of a nomogram for predicting survival in patients with stage IA non-small cell lung cancer ≤2 cm undergoing sublobectomy. *Front Oncol*. 2019;1385. doi:10.3389/fonc.2019.01385

11. Zeng Y, Mayne N, Yang CJ, et al. A nomogram for predicting cancer-specific survival of TNM 8th edition stage I non-small-cell lung cancer. *Ann Surg Oncol*. 2019;7:2053–2062. doi:10.1245/s10434-019-07318-7

12. 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;1:39–51. doi:10.1016/j.jtho.2015.09.009

13. Palma DA, Tyldesley S, Sheehan F, et al. Stage I non-small cell lung cancer (NSCLC) in patients aged 75 years and older: does age determine survival after radical treatment? *J Thorac Oncol*. 2010;6:818–824. doi:10.1097/jto.0b013e3181d6e052

14. Fan X, Liang Y, Bai Y, Yang C, Xu S. Conditional survival rate estimates of lobectomy, segmentectomy and wedge resection for stage IA1 non-small cell lung cancer: a population-based study. *Oncol Lett*. 2020;2:1607–1618. doi:10.3892/ol.2020.11713

15. Nakada T, Noda Y, Kato D, et al. Risk factors and cancer recurrence associated with postoperative complications after thoracoscopic lobectomy for clinical stage I non-small cell lung cancer. *Thorac Cancer*. 2019;10:1945–1952. doi:10.1111/1759-7714.13173

16. Schulkes KJG, Hamaker ME, Lammers JJ, van Rens MTM, Geerts M, van Elden LJR. Multidisciplinary decision-making regarding chemotherapy for lung cancer patients – an age-based comparison. *Eur J Cancer Care*. 2018;1. doi:10.1111/ecc.12768

17. Watanabe K, Sakamaki K, Nishii T, et al. Gender differences in the recurrence timing of patients undergoing resection for non-small cell lung cancer. *Asian Pac J Cancer Prev*. 2018;3:719–724. doi:10.22034/apjcp.2018.19.3.719

18. Kelsey CR, Marks LB, Hollis D, et al. Local recurrence after surgery for early stage lung cancer: an 11-year experience with 975 patients. *Cancer*. 2009;22:5218–5227. doi:10.1002/cncr.24625

19. Murakami S, Ito H, Tsubokawa N, et al. Prognostic value of the new IASLC/ATS/ERS classification of clinical stage IA lung adenocarcinoma. *Lung Cancer*. 2015;2:199–204. doi:10.1016/j.lungcan.2015.06.022

20. Yaldız D, Örs Kaya Ş, Ceylan KC, et al. Prognostic effects of predominant histologic subtypes in resected pulmonary adenocarcinomas. *Balkan Med J*. 2019;6:347–353. doi:10.4274/balkanmedj.galenos.2019.2019.1.130

21. Zhang Y, Zheng D, Xie J, et al. Development and validation of web-based nomograms to precisely predict conditional risk of site-specific recurrence for patients with completely resected non-small cell lung cancer: a multiinstitutional study. *Chest*. 2018;3:501–511. doi:10.1016/j.chest.2018.04.040

22. Port JL, Kent MS, Korst RJ, Libby D, Pasmantier M, Altorki NK. Tumor size predicts survival within stage IA non-small cell lung cancer. *Chest*. 2003;5:1828–1833. doi:10.1378/chest.124.5.1828

23. Detterbeck FC, Franklin WA, Nicholson AG, et al. The IASLC lung cancer staging project: background data and proposed criteria to distinguish separate primary lung cancers from metastatic foci in patients with two lung tumors in the forthcoming eighth edition of the TNM classification for lung cancer. *J Thorac Oncol*. 2016;5:651–665. doi:10.1016/j.jtho.2016.01.025

24. Iizuka S, Kawase A, Oiwa H, Ema T, Shiiya N, Funai K. A risk scoring system for predicting visceral pleural invasion in non-small lung cancer patients. *Gen Thorac Cardiovasc Surg*. 2019;10:876–879. doi:10.1007/s11748-019-01101-x

25. Zhang T, Zhang JT, Li WF, et al. Visceral pleural invasion in T1 tumors (≤3 cm), particularly T1a, in the eighth tumor-node-metastasis classification system for non-small cell lung cancer: a population-based study. *J Thorac Dis*. 2019;7:2754–2762. doi:10.21037/jtd.2019.06.32

26. Ginsberg RJ, Rubinstein LV. Randomized trial of lobectomy versus limited resection for T1 N0 non-small cell lung cancer. Lung Cancer Study Group. *Ann Thorac Surg*. 1995;3:615–622. doi:10.1016/0003-4975(95)00537-u

27. De Zoysa MK, Hamed D, Routledge T, Scarci M. Is limited pulmonary resection equivalent to lobectomy for surgical management of stage I non-small-cell lung cancer? *Interact Cardiovasc Thorac Surg*. 2012;6:816–820. doi:10.1093/icvts/ivs031

28. Li F, Zhao Y, Yuan L, Wang S, Mao Y. Oncologic outcomes of segmentectomy vs lobectomy in pathologic stage IA (≤2 cm) invasive lung adenocarcinoma: a population-based study. *J Surg Oncol*. 2020;7:1132–1139. doi:10.1002/jso.25880

29. Pignon JP, Tribodet H, Scagliotti GV, et al. Lung adjuvant cisplatin evaluation: a pooled analysis by the LACE Collaborative Group. *J Clin Oncol*. 2008;21:3552–3559. doi:10.1200/jco.2007.13.9030

30. Heon S, Johnson BE. Adjuvant chemotherapy for surgically resected non-small cell lung cancer. *J Thorac Cardiovasc Surg*. 2012;3:S39–42. doi:10.1016/j.jtcvs.2012.03.039

31. Li R, Yang G, Tian Y, Tian D. Comparing the benefits of postoperative adjuvant chemotherapy vs. observation for stage IB non-small cell lung cancer: a meta-analysis. *J Thorac Dis*. 2019;7:3047–3054. doi:10.21037/jtd.2019.07.47

32. Ettinger DS, Wood DE, Aisner DL, et al. Non-small cell lung cancer, version 5.2017, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw*. 2017;4:504–535. doi:10.6004/jnccn.2017.0050

33. Liang W, Zhang L, Jiang G, et al. Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer. *J Clin Oncol*. 2015;8:861–869. doi:10.1200/jco.2014.56.6661

34. Zhang J, Fan J, Yin R, et al. A nomogram to predict overall survival of patients with early stage non-small cell lung cancer. *J Thorac Dis*. 2019;12:5407–5416. doi:10.21037/jtd.2019.11.53