Ann Surg Oncol (2019) 26:2053–2062
https://doi.org/10.1245/s10434-019-07318-7

Annals of
SURGICAL ONCOLOGY
OFFICIAL JOURNAL OF THE SOCIETY OF SURGICAL ONCOLOGY

ORIGINAL ARTICLE – THORACIC ONCOLOGY

A Nomogram for Predicting Cancer-Specific Survival of TNM 8th Edition Stage I Non-small-cell Lung Cancer

Yuan Zeng, MD[1,2], Nicholas Mayne, MD[3], Chi-Fu Jeffrey Yang, MD[3], Thomas A. D'Amico, MD[3], Calvin S. H. Ng, MD[4], Chia-chuan Liu, MD[5], René Horsleben Petersen, MD[6], Gaetano Rocco, MD[7], Alessandro Brunelli, MD[8], Jun Liu, MD[1,2], Yang Liu, MD[1,2], Weizhe Huang, MD[1,2], Jiaxi He, MD[1,2], Wei Wang, MD[1,2], Long Jiang, MD[1,2], Fei Cui, MD[1,2], Wenjun Wang, MD[1,2], Wenhua Liang, MD[1,2], and Jianxing He, PhD, MD[1,2] on behalf of the AME Thoracic Surgery Collaborative Group

[1]Department of Thoracic Surgery, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou, Guangdong Province, People's Republic of China; [2]Guangzhou Institute of Respiratory Disease and China State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou, Guangdong Province, People's Republic of China; [3]Section of General Thoracic Surgery, Duke University Medical Center, Durham, NC; [4]Department of Surgery, Prince of Wales Hospital, The Chinese University of Hong Kong, Sha Tin, Hong Kong; [5]Division of Thoracic Surgery, Sun Yat-Sen Cancer Center, Taipei, Taiwan; [6]Department of Cardiothoracic Surgery, University Hospital of Copenhagen, Rigshospitalet, Copenhagen, Denmark; [7]Division of Thoracic Surgical Oncology, Department of Thoracic Surgery and Oncology, Istituto Nazionale Tumori, IRCCS, Fondazione Pascale, Naples, Italy; [8]Department of Thoracic Surgery, St James's University Hospital, Leeds, UK

ABSTRACT

Background. Models for predicting the survival outcomes of stage I non-small-cell lung cancer (NSCLC) defined by the newly released 8th edition TNM staging system are scarce. This study aimed to develop a nomogram for predicting the cancer-specific survival (CSS) of these patients and identifying individuals with a higher risk for CSS.

Methods. A total of 30,475 NSCLC cases were extracted from the Surveillance, Epidemiology, and End Results (SEER) database. We identified and integrated the risk factors to build a nomogram. The model was subjected to bootstrap internal validation with the SEER database, and external validation with a multicenter cohort of 1133 patients from China. The difference in the impact of adjuvant chemotherapy on model-defined high- and low-risk patients was examined using the National Cancer Database (NCDB).

Results. Eight independent prognostic factors were identified and integrated into the model. The calibration curves showed good agreement. The concordance index (C-index) of the nomogram was higher than that of the staging system (IA1, IA2, IA3, and IB) (internal validation set 0.63 vs. 0.56; external validation set 0.66 vs. 0.55; both $p < 0.01$). Specifically, 21.7% of stage IB patients (7.5% of all stage I) were categorized into the high-risk group (score > 30). There was a significant interaction effect between the adjuvant chemotherapy and risk groups in the NCDB cohort ($p = 0.003$).

Conclusions. We established a practical nomogram to predict CSS for 8th edition stage I NSCLC. A prospective study is warranted to determine its role in identifying adjuvant chemotherapy candidates.

Yuan Zeng, Nicholas Mayne, and Chi-Fu Jeffrey Yang have contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1245/s10434-019-07318-7) contains supplementary material, which is available to authorized users.

First Received: 21 July 2018;
Published Online: 21 March 2019

W. Liang, MD
e-mail: liangwh1987@163.com

J. He, PhD, MD
e-mail: drhe_jianxing@163.com

Lung cancer is the most common and deadly cancer worldwide, accounting for approximately 1.8 million new cases and 1.6 million deaths.[1] Non-small-cell lung cancer

(NSCLC) represents approximately 85% of lung cancer cases and has a 5-year overall survival (OS) rate of approximately 15%.[2] While surgery remains the most beneficial therapy for stage I NSCLC,[3] recurrence rates after surgical resection are still as high as 18–29%.[4] Even for patients with 8th edition TNM stage I NSCLC, which does not include tumors > 4 cm in size, the 5-year OS is only approximately 80%.[5] Postoperative chemotherapy may effectively treat micrometastatic disease, and, in a number of large, randomized, phase III trials, has been shown to have survival benefits in resected stage II and IIIA NSCLC.[6–8] However, the role of adjuvant therapy is controversial in stage I patients.[9,10] In the 7th edition of the TNM classification, adjuvant chemotherapy is recommended for stage IB patients with high risk factors, such as a tumor ≥ 4 cm, according to the National Comprehensive Cancer Network (NCCN) guidelines.[10] Therefore, the International Association for the Study of Lung Cancer (IASLC) has changed the classification of tumors ≥ 4 cm from stage IB in the 7th edition to stage IIA in the 8th edition, for NSCLC.[5] Theoretically, most of the current 8th edition stage I patients are not candidates for adjuvant chemotherapy, while the current stage IIA patients are usually offered adjuvant chemotherapy. Therefore, stratifying stage I NSCLC into different prognostic categories is important to identify patients who may benefit from postoperative adjuvant therapy. Studies have proposed that other independent prognostic factors, aside from TNM staging, such as examined lymph node (ELN) count, extent of surgery, differentiation grade, histology, and visceral pleural invasion (VPI) could significantly contribute to individualized prediction of survival.[11–14]

A nomogram is a simple graphical representation of a statistical predictive model that calculates a numerical probability of a clinical event.[15] Multivariate nomograms have been developed for prediction modeling in many types of cancer and have shown better performance than traditional TNM staging systems;[16–18] however, nomograms for predicting survival outcomes in stage I NSCLC are scarce. In this study, we aimed to develop a clinical nomogram for predicting cancer-specific survival (CSS) of the current stage I NSCLC patients after resection to identify individuals with a higher risk of cancer-related death.

METHODS

Patients and Selection Criteria

The Surveillance, Epidemiology, and End Results (SEER) database consists of 18 population-based cancer registries including nearly 28% of the US population (http://seer.cancer.gov/). Information for patients with stage I NSCLC (defined according to the 8th edition) between 1998 and 2013 was extracted from the SEER database. The inclusion criteria consisted of pathologically confirmed primary T1/T2N0M0 NSCLC (size ≤ 5 cm) between January 1998 and December 2013; history of lobectomy, pneumonectomy, bronchial sleeve resection, segmentectomy, or wedge resection, but no radiotherapy; and the presence of one malignant primary lesion. The exclusion criteria were tumor involvement of the main bronchus; history of atelectasis or obstructive pneumonitis; and no information on extracted data. Therefore, tumor size and pleural infiltration are factors that determine T staging.

An external validation cohort was provided to examine the generalizability of the model. The cohort composed 1133 patients with stage I NSCLC diagnosed between 2009 and 2014 in the First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China.

The patient baseline demographics (age, sex, and race/ethnicity), characteristics of tumors (size, location, differentiation grade, and histologic type), ELN, VPI, and treatment details (surgical type) were collected from the SEER database. In this study, histologic subtypes were classified as bronchioloalveolar cell carcinoma (BAC), squamous cell carcinoma, adenocarcinoma, and others.

Construction of the Nomogram

Eligible stage I patients from the SEER database were randomly divided into the training (75%, $n = 22{,}848$) and validation (25%, $n = 7627$) cohorts to establish and validate the nomogram. CSS was defined as the time from medical diagnosis to cancer-related death. In the training set, CSS was estimated using the Kaplan–Meier method and compared using the log-rank test. Multivariable Cox proportional hazard regression was used to determine independent prognostic factors. On the basis of the results of the multivariable analysis, a nomogram was formulated using R 2.14.1 (http://www.r-project.org) with the survival and rms package.[19]

Validation and Calibration of the Nomogram

The model was subjected to bootstrap internal validation in the training cohort, independent validation in the validation cohort, and external validation with the cohort from the First Affiliated Hospital of Guangzhou Medical University. Discrimination ability was determined using the concordance index (C-index). The values of the C-index ranged from 0.5 to 1.0, with 0.5 indicating random chance and 1.0 indicating a perfect ability to correctly discriminate the outcome using the nomogram.[15] Comparison of the C-index of two different models was based on previously described methods.[20] Statistical analyses were performed using MedCalc 18.2.1, and calibration for

1-, 3-, and 5-year CSS, which compares the predicted survival with the observed survival, was evaluated using a calibration curve.

Risk Group Stratification Based on a Modified Nomogram

The significant independent factors related to adjuvant chemotherapy selection were incorporated to establish a modified nomogram. In addition to comparing the discrimination ability by C-index, we sought to illustrate the independent discrimination ability of the nomogram for identifying stage I NSCLC patients with high- and low-risk of cancer-related death. Thus, patients with stage IIA NSCLC were also selected from the SEER database between 1998 and 2013. The 5-year CSS in patients with stage IIA was calculated. Postoperative chemotherapy was recommended for patients with stage II and III tumors. Furthermore, the 8th edition staging system has reclassified tumors > 4 cm as stage IIa,[5], thus, theoretically, most of the current stage I patients are not candidates for adjuvant chemotherapy. Patients were stratified into high- and low-risk groups using the cut-off value that was determined by 5-year survival rates of current stage IIA patients. The value was then applied to the validation cohort, and the respective Kaplan–Meier survival curves were delineated. Moreover, patients with stage IA and IB disease were combined to calculate the proportion of patients who were categorized into the high-risk group.

Examination of the Predictive Value of the Nomogram on Postoperative Chemotherapy

Patients entered in the National Cancer Database (NCDB) between 2004 and 2013 were included in this study according to our inclusion and exclusion criteria described above. The extracted information included age, sex, race, size, differentiation grade, histologic type, ELN, VPI, surgical type, adjuvant chemotherapy, and Charlson–Deyo combined comorbidity score (CDCC). Patients were divided into those who did and those who did not receive postoperative adjuvant chemotherapy. Multivariable Cox proportional hazard regression was used to determine independent prognostic factors and estimate the interaction effect, and the hazard ratio (HR) and corresponding 95% confidence interval (CI) were calculated. The interaction effect was illustrated using Kaplan–Meier curves, which were delineated by adjuvant chemotherapy history with risk group stratification, and statistically tested by introducing an interaction item of (risk group) * (history of adjuvant chemotherapy) in the Cox regression model. Statistical analyses were performed using SPSS version 22.0 (IBM Corporation, Armonk, NY, USA); a p value < 0.05 was considered statistically significant.

RESULTS

Study Cohorts

A total of 30,475 patients with stage I NSCLC and 2695 patients with stage IIa NSCLC from the SEER database (electronic supplementary Fig. S1), as well as 1133 patients with stage I NSCLC from the First Affiliated Hospital of Guangzhou Medical University, met the inclusion criteria for this study. Eligible stage I patients from the SEER database were randomly divided into the training ($n = 22{,}848$) and internal validation ($n = 7627$) cohorts to establish and validate the nomogram. The external validation cohort consisted of 1133 patients with stage I NSCLC between 2009 and 2014 from the First Affiliated Hospital of Guangzhou Medical University. The median [interquartile range (IQR)] follow-up times on CSS were 45 months (20–81), 45 months (21–81), and 38.3 months (25.5–54.9) for the training, internal validation, and external validation cohorts, respectively. The demographics and clinicopathological characteristics assessed in the three cohorts are presented in Table 1.

Independent Prognostic Factors in the Training Cohort

The results of the univariable analysis are listed in Table 2. All significant factors in the univariable analysis were entered into the multivariable analysis based on the Cox regression. The multivariate analysis indicated that age ($p < 0.001$), sex ($p < 0.001$), race ($p < 0.05$), ELN ($p < 0.001$), tumor size ($p < 0.001$), extent of surgery ($p < 0.001$), differentiation grade ($p < 0.001$), histology ($p < 0.001$), and VPI ($p < 0.001$) were independent prognostic factors for CSS (Table 2).

Prognostic Nomogram for Cancer-Specific Survival

Significant independent factors, including age, sex, ELN, tumor size, extent of surgery, differentiation grade, histology, and VPI, were incorporated to establish the nomogram (Fig. 1). The nomogram showed tumor size and differentiation grade as sharing the largest contribution to prognosis, followed by age and histology. The number of ELNs and the extent of surgery showed a moderate impact on survival. Each factor within these variables was assigned a score on the point scale. By accumulating the total score and locating it on the total point scale, it was easy to draw a straight line down to determine the estimated probability of survival at each score point.

TABLE 1 Demographics and clinicopathologic characteristics of the training, internal validation, and external validation cohort

Characteristic	Training cohort (n = 22,848)		Internal validation cohort (n = 7627)		External validation cohort (n = 1133)	
	N	%	*N*	%	*N*	%
Age (years)						
< 60	4725	20.7	1571	20.6	503	44.4
60–70	7999	35	2690	35.3	387	34.2
≥ 70	10,124	44.3	3366	44.1	243	21.4
Sex						
Male	10,344	45.3	3464	45.4	572	50.5
Female	12,504	54.7	4163	54.6	561	49.5
Race						
White	19,545	85.5	6503	85.3	–	–
Black	1779	7.8	605	7.9	–	–
Other	1524	6.7	519	6.8	–	–
Location						
Upper	14,543	63.7	4761	62.4	648	57.2
Middle	1217	5.3	425	5.6	87	7.7
Lower	6899	30.2	2381	31.2	361	31.9
Other	189	0.8	60	0.8	37	3.3
ELN						
< 16	20,145	88.2	6691	87.7	671	59.2
≥ 16	2682	11.8	936	12.3	462	40.8
Tumor size (cm)						
≤ 1	1526	6.7	562	7.4	246	21.7
> 1–2	9112	39.9	3041	39.9	471	41.6
> 2–3	7992	35	2634	34.5	329	29
> 3–4	4218	18.5	1390	18.2	87	7.7
Extent of surgery						
Lobectomy[a]	20,032	87.7	6693	87.8	889	78.5
Segmentectomy	782	3.4	259	3.4	91	8
Wedge resection	2034	8.9	675	8.9	153	13.5
Differentiation grade						
I	4388	19.2	1452	19	13	1.1
II	10,631	46.5	3535	46.3	1028	90.7
III or IV	7829	34.3	2640	33.6	92	8.1
Histology						
BAC	2244	9.8	746	9.8	81	7.1
SC	5629	24.6	1878	24.6	114	10.1
Adenocarcinoma	12,501	54.7	4187	54.9	884	78
Others	2474	10.8	816	10.7	54	4.8
VPI						
No	19,431	85	6519	85.5	643	56.8
Yes	3417	15	1108	14.5	490	43.2

ELN examined lymph node, *BAC* bronchioloalveolar carcinoma, *SC* squamous carcinoma, *VPI* visceral pleural invasion

[a]Includes lobectomy, pneumonectomy, and bronchial sleeve resection

TABLE 2 Univariable analysis and Cox proportional hazards regression analysis

Characteristic	Univariable analysis *p* value	Multivariable analysis			Selected factors for building the model		
		HR	95% CI	*p* value	HR	95% CI	*p* value
Age (years)	< 0.001						
< 60		1 (reference)			1 (reference)		
60–70		1.258	1.155–1.370	< 0.001	1.251	1.149–1.362	< 0.001
≥ 70		1.692	1.560–1.836	< 0.001	1.679	1.548–1.821	< 0.001
Sex	< 0.001						
Male		1.262	1.192–1.337	< 0.001	1.258	1.188–1.332	< 0.001
Female		1 (reference)			1 (reference)		
Race	< 0.001						
White		1 (reference)					
Black		1.131	1.018–1.256	0.021			
Other		0.849	0.749–0.962	0.01			
Location	< 0.001						
Upper		1 (reference)					
Middle		1.037	0.908–1.184	0.592			
Lower		1.070	1.005–1.139	0.034			
Other		1.507	1.177–1.928	0.001			
ELN	< 0.001						
0–15		1.312	1.186–1.452	< 0.001	1.314	1.187–1.453	< 0.001
≥ 16		1 (reference)			1 (reference)		
Tumor size (cm)	< 0.001						
≤ 1		1 (reference)			1 (reference)		
> 1–2		1.121	0.979–1.283	0.098	1.198	1.039–1.381	0.013
> 2–3		1.403	1.225–1.606	< 0.001	1.518	1.317–1.750	< 0.001
> 3–4		1.579	1.371–1.819	< 0.001	1.756	1.515–2.035	< 0.001
Extent of surgery	< 0.001						
Lobectomy[a]		1 (reference)			1 (reference)		
Segmentectomy		1.255	1.079–1.458	< 0.001	1.263	1.087–1.468	0.002
Wedge resection		1.399	1.269–1.542	< 0.001	1.411	1.280–1.556	< 0.001
Differentiation grade	< 0.001						
I		1 (reference)			1 (reference)		
II		1.593	1.443–1.760	< 0.001	1.588	1.438–1.754	< 0.001
III or IV		1.833	1.651–2.035	< 0.001	1.822	1.641–2.023	< 0.001
Histology	< 0.001	1 (reference)					
BAC		1 (reference)					
SC		1.439	1.268–1.632	< 0.001	1.464	1.291–1.660	< 0.001
Adenocarcinoma		1.288	1.146–1.447	< 0.001	1.292	1.150–1.451	< 0.001
Others		1.416	1.231–1.629	< 0.001	1.436	1.249–1.652	< 0.001
VPI	< 0.001						
No		1 (reference)					
Yes		1.241	1.144–1.345	< 0.001	1.235	1.139–1.339	< 0.001

HR hazard ratio, *CI* confidence interval, *ELN* examined lymph node, *BAC* bronchioloalveolar carcinoma, *SC* squamous carcinoma, *VPI* visceral pleural invasion

[a]Includes lobectomy, pneumonectomy, and bronchial sleeve resection

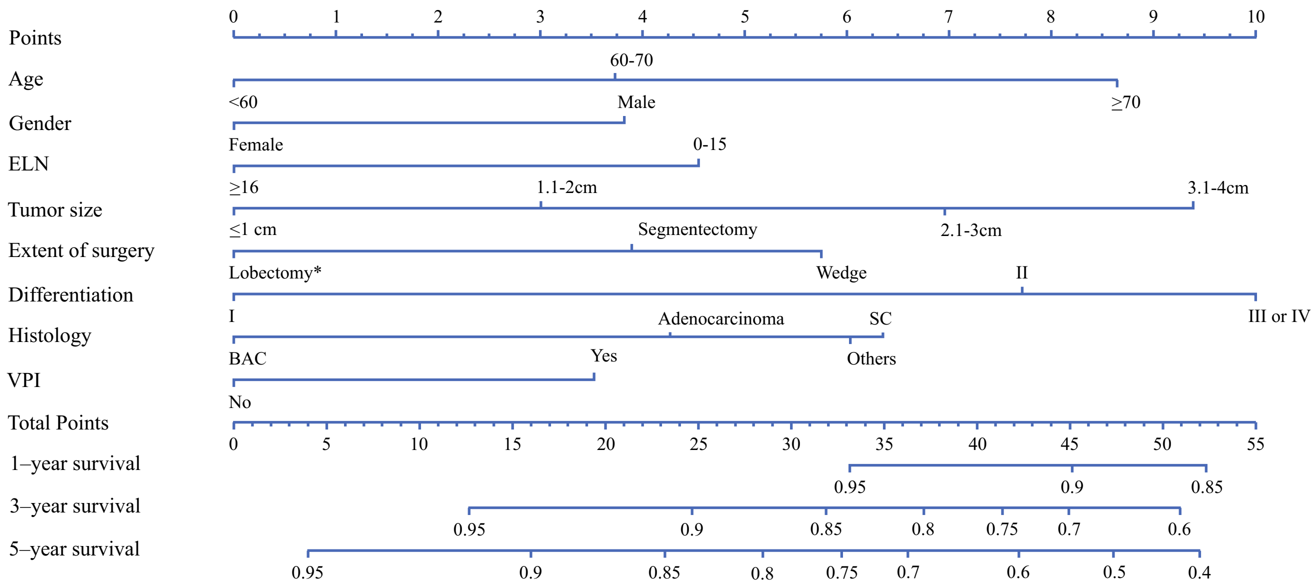

FIG. 1 Postoperative prognostic nomogram for patients with stage I disease. *Lobectomy includes lobectomy, pneumonectomy, and bronchial sleeve resection. *ELN* examined lymph node, *BAC* bronchioloalveolar carcinoma, *SC* squamous carcinoma, *VPI* visceral pleural invasion

Calibration and Validation of the Nomogram

The calibration plots presented an acceptable agreement in the training, internal validation, and external validation cohorts between the nomogram predictions and actual observations for 1-, 3-, and 5-year CSS (Fig. 2). In the training cohort, Harrell's C-index for the established nomogram to predict CSS (0.64; 95% CI 0.63–0.65) was significantly greater than that of the IASLC 8th edition staging system (stages IA1, IA2, IA3, and IB) (0.57; 95% CI 0.56–0.57; $p < 0.01$). In the internal validation cohort, the C-index was higher for the nomogram prediction (0.63; 95% CI 0.62–0.65) than for the TNM category prediction (0.56; 95% CI 0.54–0.57; $p < 0.01$). In the external validation cohort, the C-index was also higher for the nomogram prediction (0.66; 95% CI 0.60–0.73) than for the TNM category prediction (0.55; 95% CI 0.49–0.61; $p < 0.01$). The C-indexes for tumor size, ELN, extent of surgery, differentiation grade, histology, and VPI prediction were also greater for the nomogram prediction (electronic supplementary Table S1).

Impact of Adjuvant Chemotherapy on High- and Low-Risk Patients

To perform an exploratory examination of the predictive value of the nomogram on adjuvant chemotherapy, a modified nomogram, including ELN, tumor size, extent of surgery, differentiation grade, histology, and VPI was established. Age and sex were not included due to the lack of direct association with the decision for adjuvant chemotherapy. Details are summarized in the electronic supplementary material. The 5-year CSS of patients with stage IIa disease was 66.8%, which corresponded to a score of 30. After applying this cut-off value to all patients, stratification into high-risk (score > 30) and low-risk (score 0–30) subgroups allowed significant distinction between Kaplan–Meier curves for survival outcomes within each T category (Fig. 3a). All high-risk patients (score > 30) shared similar prognoses despite the T category. Specifically, 7.5% of all stage Ia patients and 21.7% of stage IB patients were categorized into the high-risk group (score > 30) and had inferior CSS to that of stage IIA patients. Kaplan–Meier curves for the external validation cohort are also shown in Fig. 3b.

A total of 38,927 patients with stage I NSCLC in NCBD were included in the study. The median (IQR) follow-up time was 32.1 months (20.6–46.2). In the low- and high-risk groups, 3.5% and 10.9% of patients, respectively, received postoperative chemotherapy. In the NCDB cohort, there was a significant interaction effect between adjuvant chemotherapy and the risk group (HR 0.7, 95% CI 0.55–0.89; $p = 0.003$), indicating a greater OS difference from adjuvant chemotherapy in high-risk patients than in low-risk patients (electronic supplementary Table S2). The Kaplan–Meier curves for OS stratified by risk and adjuvant chemotherapy are shown in electronic supplementary Fig. S2.

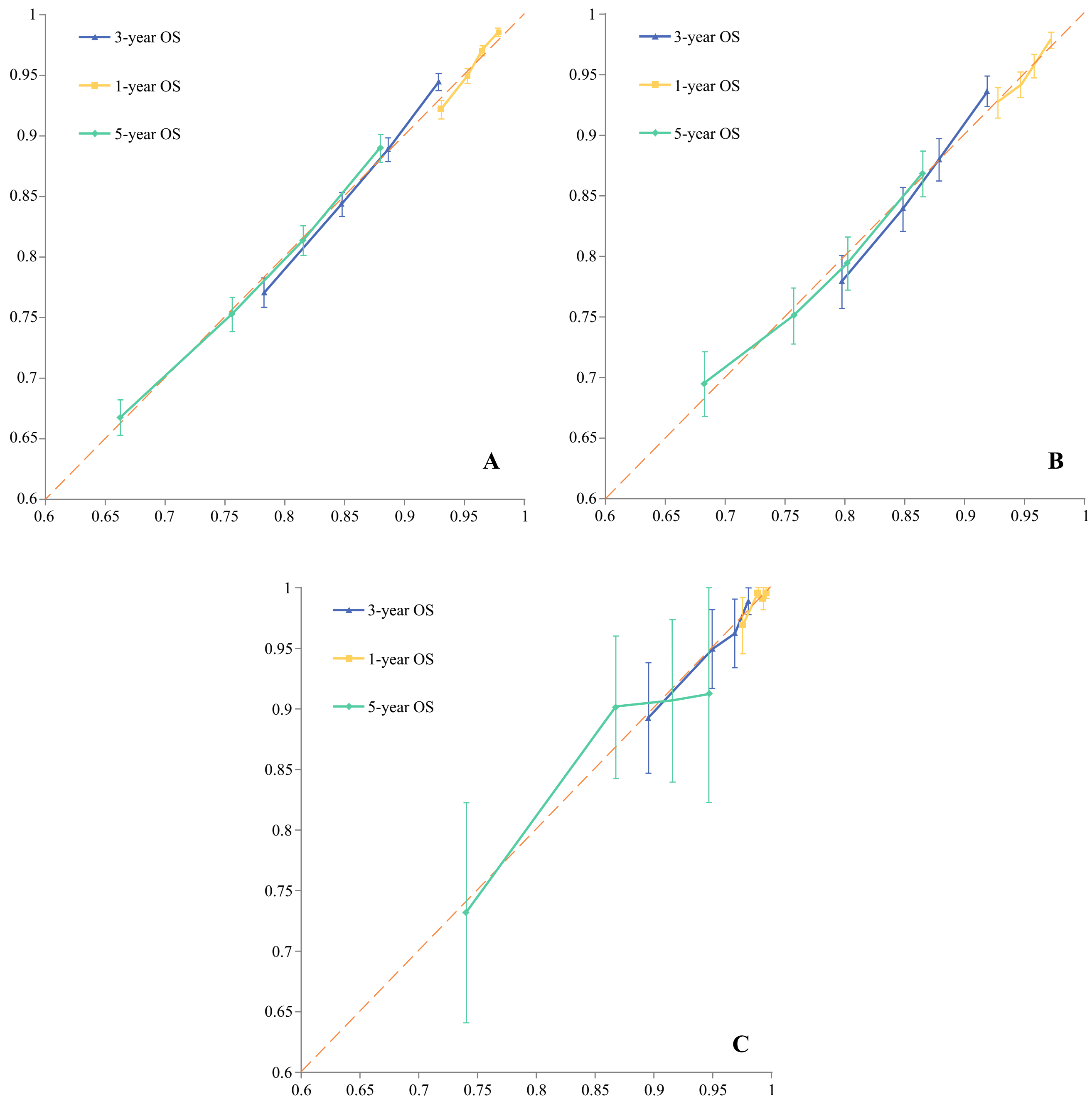

FIG. 2 Calibration curves for predicting patient survival at each time point in the **a** training cohort, **b** internal validation cohort, and **c** external validation cohort. Nomogram-predicted CSS is plotted on the *x*-axis, and actual CSS is plotted on the y-axis. A plot along the 45-degree line (*dashed line*) would indicate a perfect calibration model in which the predicted probabilities are identical to the actual outcomes. *CSS* cancer-specific survival, *OS* overall survival

DISCUSSION

Previous studies have suggested that patients with 7th edition stage IB tumors > 4 cm might benefit from adjuvant chemotherapy.[9,21] Thus, the 8th edition staging system has reclassified tumors > 4 cm as stage IIA.[5] Most of the current stage I patients are, theoretically, not candidates for adjuvant chemotherapy; however, some of the current edition stage I patients still presented high-risk factors. In this study, we identified age, sex, ELN, tumor size, extent of surgery, differentiation grade, histology, and VPI as independent factors for CSS.

The independent prognostic factors in our results were similar to those identified in several previous studies[11,22–24] for stage I NSCLC. Notably, ELN count was an important prognostic factor for node-negative NSCLC; a cut-off point

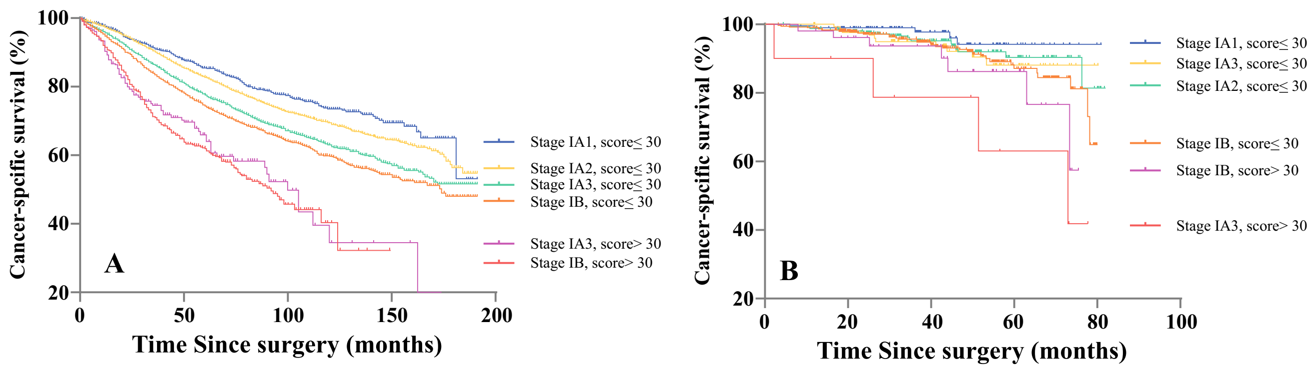

FIG. 3 Risk group stratification within each staging system (IA1, IA2, IA3, IB) in the **a** SEER database and **b** external validation cohort. *SEER* Surveillance, Epidemiology, and End Result

of 16 ELNs was recommended in our previous study.[12] Recently, lobectomy, segmentectomy, and wedge resection showed significantly different CSS for NSCLC $\leq$ 2 cm.[11] Several studies[25–27] and a meta-analysis[13] demonstrated that VPI is associated with worse prognosis in early-stage NSCLC. Since 2011, the use of the term BAC in the IASLC has been revised: the terms adenocarcinoma in situ (AIS) and minimally invasive adenocarcinoma (MIA) were introduced to replace some of the BAC category.[28] As this new coding system has not yet been integrated into the SEER database, many BAC cases in this study should be classified as AIS/MIA.[29] Adjuvant chemotherapy is recommended for 7th edition stage IB patients with high-risk factors, including poorly differentiated tumors, vascular invasion, wedge resection, tumors $>$ 4 cm, and visceral pleural involvement, in the NCCN guidelines index.[10] Therefore, age, sex, ELN count, tumor size, extent of surgery, differentiation grade, histology, and VPI were the final variables used in the nomogram. In addition, age, sex, race, and location were not included in the risk group stratification and examination of the predictive value of the nomogram on adjuvant chemotherapy because these variables were not directly associated with the decision for adjuvant chemotherapy.

By incorporating these factors, the C-index of the nomogram was higher than that of the staging system (IA1, IA2, IA3, IB). In the independent validation cohort and the external validation cohort, the discriminative abilities were similar. Moreover, our nomogram (0.63 and 0.66 for the primary and external validation cohorts, respectively) performed significantly better than TNM staging (0.56 and 0.55 for the training and external validation cohorts, respectively) for the prediction of individual CSS. Therefore, by incorporating multiple risk factors, this model had greater discrimination ability than the TNM staging system. Despite several previously reported NSCLC prognostic models,[14,18,30–32] a nomogram has not been developed for 8th edition stage I NSCLC. Kratz et al.[31] reported a practical molecular assay to predict survival in resected NSCLC; however, this system was based on data from a 14-gene expression assay, and its applicability to clinical practice is still difficult to comprehend. We previously developed a nomogram model for predicting OS in patients with stage I–IIIA resected NSCLC, but it was a prognosis estimation model and was not directly associated with tailoring for adjuvant chemotherapy.[18] Some studies[14,33] reported models for postoperative recurrence in stage IA NSCLC, however their sample sizes were relatively small. In this study, we used a large multicenter population to establish a risk prediction nomogram incorporating most recognized risk factors, which targeted very early-stage patients who were considered very low risk. Although the C-index increased by only 0.07 (0.63 vs. 0.56), we did not need to detect additional indicators, such as molecular tests, because all factors of the nomogram were the existing clinical data. In clinical practice, it is easy to perform an individualized survival prediction after surgery for patients with stage I NSCLC using the nomogram scoring system.

Adjuvant chemotherapy is routinely recommended in patients with stage II and III tumors,[6–8] but not in patients with stage I disease; however, some studies have demonstrated that adjuvant chemotherapy improved survival, even in patients with stage IA NSCLC compared with surgery alone.[34,35] This effect may be attributable to patients with high-risk of cancer-related death in stage IA NSCLC. In addition, stage IIA patients are considered to have the lowest risk for death among those who could benefit from adjuvant chemotherapy. Thus, by stratifying patients into high- and low-risk groups using the cut-off value, which was determined by the 5-year survival rates of current stage IIA patients, we separated patients with distinct survival outcomes (Fig. 3). Patients in each high-risk group had inferior CSS to that of stage IIA patients, and, interestingly, all high-risk patients ($>$ 30) shared similar prognosis despite their tumor stage.

In this study, the most important point is that 21.7% of stage IB patients (7.5% of all stage I) were categorized into the high-risk group. Therefore, we examined the predictive value of this nomogram for the impact of adjuvant chemotherapy. In the NCDB ($p = 0.003$), there was a significant interaction effect between the chemotherapy and risk groups, which strongly proved a greater survival difference from adjuvant chemotherapy in high-risk patients than in low-risk patients. This finding indicated that a prospective study is warranted to determine its role in identifying adjuvant chemotherapy candidates.

We established and validated this nomogram with a large population of patients from the SEER database and external data, and the calibration curves showed good agreement between nomogram prediction and actual observation; thus, the repeatability and reliability of the nomogram can be guaranteed. In addition, this model includes only the important clinically available variables and is cheaper than molecular tests, making it a more economical and practical option worldwide.

There are still some limitations of this study. First, this study was limited by the retrospective nature of data collection, which represents an inevitable bias. This indication bias also included more patients with a high risk of metastasis who chose chemotherapy. Therefore, we could only use interaction to test and demonstrate the predictive role of the model. Even if the interaction was significant, the confounding by indication bias was still inevitable. Second, the SEER program lacks data on lymphovascular invasion, the latest classification of adenocarcinoma, some relevant molecular factors, spread through air space, and a proportion of pathological subtypes. Furthermore, there are also limitations associated with the NCDB. Although we used multivariable analysis to reduce the impact of confounding and investigate specific covariates, there are certain covariates that are not available in the NCDB, such as surgeon experience and detailed comorbidity information. In addition, the NCDB does not have information on CSS, which limited our NCDB analysis to examine OS. Although these factors are important, they are not widely available in the current stage, which restricts their clinical use. Despite these limitations, this nomogram was developed using a large population from the SEER database providing unique opportunities to predict CSS for patients with 8th edition stage I NSCLC. Future studies using prospective data collection that includes data on chemotherapy or targeted therapy are encouraged to further determine the impact of adjuvant chemotherapy on these high-risk patients.

CONCLUSION

We established a practical nomogram that can provide individual predictions of CSS for patients with 8th edition stage I NSCLC. This nomogram may help clinicians identify high-risk stage I patients after surgical resection. Further prospective studies to determine the impact of adjuvant chemotherapy on these high-risk patients are warranted.

AUTHOR CONTRIBUTIONS Conception and design: Jianxing He, Yuan Zeng, Wenhua Liang. Collection and assembly of data: Jianxing He, Yuan Zeng, Wenhua Liang, Nicholas Mayne, Chi-Fu Jeffrey Yang. Data analysis and interpretation: All authors. Manuscript preparation: All authors. Manuscript proofing: All authors. Final approval of the manuscript. All authors.

FUNDING This study was supported by the China State Key Laboratory of Respiratory Disease Independent Subject (Grant No. SKLRD-QN-201925); China National Science Foundation (Grant Nos. 81871893 and 81501996); Key Project of Guangzhou Scientific Research Project (Grant Nos. 201804020030); high-level university construction project of Guangzhou Medical University (Grant Nos. 20182737, 201721007, 201715907, and 2017160107); and the National Key R & D Program (Grant Nos. 2017YFC0907903 and 2017YFC0112704).

DISCLOSURE We have no conflicts of interest to declare.

REFERENCES

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin.* 2015;65:87–108.
2. Goldstraw P, Ball D, Jett JR, Le CT, Lim E, Nicholson AG, et al. Non-small-cell lung cancer. *Lancet.* 2011;378:1727–40.
3. Datta D, Lahiri B. Preoperative evaluation of patients undergoing lung resection surgery. *Chest.* 2003;123:2096–103.
4. Varlotto JM, Recht A, Flickinger JC, Medford-Davis LN, Dyer AM, DeCamp MM. Varying recurrence rates and risk factors associated with different definitions of local recurrence in patients with surgically resected, stage I nonsmall cell lung cancer. *Cancer.* 2010;116:2390–400.
5. Goldstraw P, Chansky K, Crowley J, Rami-Porta R, Asamura H, Eberhardt WE, 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.
6. Arriagada R, Bergman B, Dunant A, Le CT, Pignon JP, Vansteenkiste J. Cisplatin-based adjuvant chemotherapy in patients with completely resected non-small-cell lung cancer. *N Engl J Med.* 2004;350:351–60.
7. Arriagada R, Dunant A, Pignon JP, Bergman B, Chabowski M, Grunenwald D, et al. Long-term results of the international adjuvant lung cancer trial evaluating adjuvant Cisplatin-based chemotherapy in resected lung cancer. *J Clin Oncol.* 2010;28:35–42.
8. Douillard JY, Rosell R, De Lena M, Carpagnano F, Ramlau R, Gonzáles-Larriba JL, et al. Adjuvant vinorelbine plus cisplatin versus observation in patients with completely resected stage IB-IIIA non-small-cell lung cancer (Adjuvant Navelbine International Trialist Association [ANITA]): a randomised controlled trial. *Lancet Oncol.* 2006;7:719–27.
9. Strauss GM, Herndon JE, Maddaus MA, Johnstone DW, Johnson EA, Harpole DH, et al. Adjuvant paclitaxel plus carboplatin compared with observation in stage IB non-small-cell lung cancer: CALGB 9633 with the Cancer and Leukemia Group B,

Radiation Therapy Oncology Group, and North Central Cancer Treatment Group Study Groups. *J Clin Oncol.* 2008;26:5043–51.
10. National Comprehensive Cancer Network: NCCN clinical practice guidelines in oncology: nonsmall cell lung cancer. V. 1.2017. www.nccn.org/professionals/physician_gls/PDF/nscl.pdf.
11. Dai C, Shen J, Ren Y, Zhong S, Zheng H, He J, et al. Choice of surgical procedure for patients with non-small-cell lung cancer ≤ 1 cm or > 1 to 2 cm among lobectomy, segmentectomy, and wedge resection: a population-based study. *J Clin Oncol.* 2016;34:3175–82.
12. Liang W, He J, Shen Y, Shen J, He Q, Zhang J, et al. Impact of examined lymph node count on precise staging and long-term survival of resected non-small-cell lung cancer: a population study of the US SEER database and a chinese multi-institutional registry. *J Clin Oncol.* 2017;35(11):1162–1170.
13. Jiang L, Liang W, Shen J, Chen X, Shi X, He J, et al. The impact of visceral pleural invasion in node-negative non-small cell lung cancer: a systematic review and meta-analysis. *Chest.* 2015;148:903–11.
14. Zhang Y, Sun Y, Xiang J, Zhang Y, Hu H, Chen H. A clinicopathologic prediction model for postoperative recurrence in stage Ia non-small cell lung cancer. *J Thorac Cardiovasc Surg.* 2014;148:1193–9.
15. Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol.* 2008;26:1364–70.
16. Wang Y, Li J, Xia Y, Gong R, Wang K, Yan Z, et al. Prognostic nomogram for intrahepatic cholangiocarcinoma after partial hepatectomy. *J Clin Oncol.* 2013;31:1188–95.
17. Zaak D, Burger M, Otto W, Bastian PJ, Denzinger S, Stief CG, et al. Predicting individual outcomes after radical cystectomy: an external validation of current nomograms. *BJU Int.* 2010;106:342–8.
18. Liang W, Zhang L, Jiang G, Wang Q, Liu L, Liu D, et al. Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer. *J Clin Oncol.* 2015;33:861–9.
19. Frank EHJR. Regression Modeling Strategies. R Package version 3.4-0. http://www.rproject.org/.
20. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology.* 1983;148:839–43.
21. Ou SH, Zell JA, Ziogas A, Anton-Culver H. Prognostic factors for survival of stage I nonsmall cell lung cancer patients: a population-based analysis of 19,702 stage I patients in the California Cancer Registry from 1989 to 2003. *Cancer.* 2007;110:1532–41.
22. Kato H, Ichinose Y, Ohta M, Hata E, Tsubota N, Tada H, et al. A randomized trial of adjuvant chemotherapy with uracil-tegafur for adenocarcinoma of the lung. *N Engl J Med.* 2004;350:1713–21.
23. Woo T, Okudela K, Mitsui H, Tajiri M, Yamamoto T, Rino Y, et al. Prognostic value of the IASLC/ATS/ERS classification of lung adenocarcinoma in stage I disease of Japanese cases. *Pathol Int.* 2012;62:785–91.
24. Whitson BA, Groth SS, Andrade RS, Maddaus MA, Habermann EB, D'Cunha J. Survival after lobectomy versus segmentectomy for stage I non-small cell lung cancer: a population-based analysis. *Ann Thorac Surg.* 2011;92:1943–50.
25. Shimizu K, Yoshida J, Nagai K, Nishimura M, Ishii G, Morishita Y, et al. Visceral pleural invasion is an invasive and aggressive indicator of non-small cell lung cancer. *J Thorac Cardiovasc Surg.* 2005;130:160–5.
26. Lakha S, Gomez JE, Flores RM, Wisnivesky JP. Prognostic significance of visceral pleural involvement in early-stage lung cancer. *Chest.* 2014;146:1619–26.
27. Schuchert MJ, Schumacher L, Kilic A, Close J, Landreneau JR, Pennathur A, et al. Impact of angiolymphatic and pleural invasion on surgical outcomes for stage I non-small cell lung cancer. *Ann Thorac Surg.* 2011;91:1059–65; discussion 1065.
28. Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger KR, Yatabe Y, et al. International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society international multidisciplinary classification of lung adenocarcinoma. *J Thorac Oncol.* 2011;6:244–85.
29. Veluswamy RR, Ezer N, Mhango G, Goodman E, Bonomi M, Neugut AI, et al. Limited resection versus lobectomy for older patients with early-stage lung cancer: impact of histology. *J Clin Oncol.* 2015;33:3447–53.
30. Kent MS, Mandrekar SJ, Landreneau R, Nichols F, Foster NR, DiPetrillo TA, et al. A Nomogram to Predict Recurrence and Survival of High-Risk Patients Undergoing Sublobar Resection for Lung Cancer: An Analysis of a Multicenter Prospective Study (ACOSOG Z4032). *Ann Thorac Surg.* 2016;102:239–46.
31. Kratz JR, He J, Van Den Eeden SK, Zhu ZH, Gao W, Pham PT, et al. A practical molecular assay to predict survival in resected non-squamous, non-small-cell lung cancer: development and international validation studies. *Lancet.* 2012;379:823–32.
32. Birim O, Kappetein AP, Waleboer M, Puvimanasinghe JP, Eijkemans MJ, Steyerberg EW, et al. Long-term survival after non-small cell lung cancer surgery: development and validation of a prognostic model with a preoperative and postoperative mode. *J Thorac Cardiovasc Surg.* 2006;132:491–8.
33. Chen H, Sui X, Yang F, Liu J, Wang J. Nomograms for predicting recurrence and survival of invasive pathological stage IA non-small cell lung cancer treated by video assisted thoracoscopic surgery lobectomy. *J Thorac Dis.* 2017;9:1046–53.
34. Hamada C, Tsuboi M, Ohta M, Fujimura S, Kodama K, Imaizumi M, et al. Effect of postoperative adjuvant chemotherapy with tegafur-uracil on survival in patients with stage IA non-small cell lung cancer: an exploratory analysis from a meta-analysis of six randomized controlled trials. *J Thorac Oncol.* 2009;4:1511–6.
35. Tsuchiya T, Akamine S, Muraoka M, Kamohara R, Tsuji K, Urabe S, et al. Stage IA non-small cell lung cancer: vessel invasion is a poor prognostic factor and a new target of adjuvant chemotherapy. *Lung Cancer.* 2007;56:341–8.