



Development and validation of nomograms for predicting overall and cancer-specific survival in young patients with non-small cell lung cancer

Yizhou Peng^{1,2}, Yihua Sun^{1,2}

¹Department of Thoracic Surgery, Fudan University Shanghai Cancer Center, Shanghai 200032, China; ²Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200032, China

Contributions: (I) Conception and design: Y Sun; (II) Administrative support: Y Sun; (III) Provision of study materials or patients: Y Peng; (IV) Collection and assembly of data: Y Peng; (V) Data analysis and interpretation: Y Peng; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Yihua Sun. Department of Thoracic Surgery, Fudan University Shanghai Cancer Center, 270 Dong-An Road, Shanghai 200032, China. Email: sun_yihua76@hotmail.com.

Background: Young patients with non-small cell lung cancer (NSCLC) represent a distinct subgroup of patients with this disease. This study aimed to construct nomograms to predict the overall survival (OS) and cancer-specific survival (CSS) of young patients with NSCLC.

Methods: NSCLC patients under 50 years old diagnosed between 2010 and 2016 were selected from the Surveillance, Epidemiology, and End Results (SEER) database and randomly divided into training (n=1,357) and validation (n=678) cohorts at a ratio of 2:1. Independent prognostic factors for OS or CSS were identified through the log-rank test, Cox proportional hazards models or competing risk model and further integrated to construct nomograms. The predictive capability of the nomogram was assessed by Harrell's concordance index (C-index), the calibration curve and risk group stratification.

Results: A total of 2,035 patients were enrolled. In the training cohort, insurance, marital status, histological type, grade, T stage, N stage and surgery were identified as independent prognostic for OS and CSS. The C-index value were 0.759 [95% confidence interval (CI): 0.731–0.787] for OS and 0.810 (95% CI: 0.803–0.818) for BCSS in the training cohort and 0.751 (95% CI: 0.711–0.790) for OS and 0.807 (95% CI: 0.795–0.819) for CSS in the validation cohort. The calibration curves showed optimal agreement between the predicted and actual survival both in internal and external validation. In addition, patients in the validation cohort within different risk groups exhibited significantly different survival even in each TNM stage.

Conclusions: Nomograms were developed and validated to predict OS and CSS of young patients with NSCLC in our study. A prospective study with more potential prognostic factors and the latest TNM classification is required to ameliorate this model.

Keywords: Non-small cell lung cancer (NSCLC); young patients; nomogram; survival; Surveillance, Epidemiology, and End Results (SEER)

Submitted Nov 05, 2019. Accepted for publication Feb 14, 2020.

doi: 10.21037/jtd.2020.03.03

View this article at: <http://dx.doi.org/10.21037/jtd.2020.03.03>

Introduction

Accounting for 5% to 10% of non-small cell lung cancer (NSCLC), young patients with NSCLC represent a rare but distinct subgroup of patients with this disease (1). Previous

studies have reported that young patients with NSCLC have different clinicopathological features. It is more likely to be non-white, female, non-smokers, adenocarcinoma (ADC), lymph node positivity and invasive disease in young patients

(2-4). With the popular application of low-dose computed tomography (LDCT), there are an increasing number of young patients with lung cancer (5). It is reasonable that the expecting life span of young patients will be longer than the elders when receiving curative treatments. Adjuvant therapy could improve the overall survival (OS) and cancer-specific survival (CSS). Current Comprehensive Cancer Network (NCCN) guidelines of NSCLC recommends adjuvant therapy based on the TNM classification (6). Therefore, it is crucial to identify patients with high risk of recurrence and cancer-specific death among young patients, who may benefit from adjuvant therapies. However, current managements for young NSCLC patients are not distinguishable. Nomogram, that estimates a probability about recurrence or death of each patient by integrating independent prognostic factors, has been widely used in many subtypes of diseases (7,8). To our knowledge, there is no nomogram for predicting OS and CSS of young patients with NSCLC yet.

In this study, we aimed to develop and validate a nomogram for young patients with NSCLC based on the cases diagnosed between 2010 and 2016 from the Surveillance, Epidemiology, and End Results (SEER) database. Subsequently, we compared nomograms with those based on the traditional TNM system.

Methods

Study cohort

SEER is a public database collecting data from 18 population-based cancer registries and accounts for approximately 30% of the United States population. Patients aged 18–50 diagnosed from 2010 to 2016 were extracted (released in April 2019, based on the November 2018 submission). T and N stage were suggested by the American Joint Committee on Cancer (AJCC) 7th edition of the TNM staging system. The inclusion and exclusion criteria are listed as follows:

Inclusion criteria:

- (I) Malignant Lung cancer (site record: C34.0-C34.9) with the histological types of ADC (8140–8147, 8255, 8260, 8310, 8323, 8480, 8481, 8490, 8550, 8572), squamous cell carcinoma (8050, 8051, 8052, 8070–8078) and other types of NSCLC (8010, 8012, 8014, 8015, 8020, 8021, 8022, 8030, 8036) based on the International Classification of Diseases for Oncology, 3rd Edition (ICD-O-3);
- (II) Lung cancer was the only or first primary cancer

diagnosis;

- (III) Unilateral NSCLC;
- (IV) Survival time was 1 month at least.

Exclusion criteria:

- (I) Diagnosis was obtained through death certificate or autopsy;
- (II) Patients with unknown race information, unknown insurance status, unknown marital status, unknown histological grade, unknown T stage or N stage and stage IV lung cancer.

The eligible patients were randomly divided into the training cohort and validation cohort at a ratio of 2:1.

Construction of nomograms

OS was defined as the time from the date of diagnosis to the date of death or date of last follow-up. After the Kaplan-Meier estimate and log-rank test in the training cohort, univariate prognostic factors with $P < 0.1$ were identified and further included in the multivariable analysis (Cox proportional hazards models). A nomogram was constructed by integrated independent prognostic factors with $P < 0.05$.

CSS was calculated from the data of diagnosis to the data of death due to NSCLC or date of the last follow-up. Deaths attributable to causes other than this cancer were considered competing risk. Cumulative incidence function (CIF) was used to assess the probability of death. Gray's test was used to estimate the difference of CIF among groups. The competing risk model was performed through a subdistribution analysis of competing risks. The independent prognostic factors determined with $P < 0.05$ identified by the multivariate analysis were used to develop a nomogram for CSS.

Validation of nomograms

Internal (1,000 bootstrap resamples in the training cohort) and external (in the validation cohort) validation were utilized to assess the predictive capability of nomograms. Harrell's concordance index (C-index) with a 95% confidence interval (CI) was calculated to assess the discriminative ability of the model. The calibration curve was plotted with predicted and observed survival to evaluate the calibration of the model.

Comparison of nomograms

In the training and validation cohort, the capability of

nomograms based on the TNM stage and those developed in our study were compared based on the C-index.

Risk group stratification

In addition to comparing the discriminative ability by C-index, we sought to illustrate the independent discrimination ability of the nomogram of OS for identifying patients with different risk of death. Patients in the validation cohort were stratified into high-, intermediate- and low-risk groups using the cut-off value which was determined by evenly dividing patients in the training cohort into different risk groups according to the total risk scores of the nomogram for predicting OS. The respective Kaplan-Meier survival curves were delineated.

Statistical analysis

Data were collected using SEER*Stat Software (version 8.3.6, <https://seer.cancer.gov/seerstat/>). All statistical analyses were performed in R software (version 3.6.1, <http://www.r-project.org>) with the R package as the following: survival (<http://CRAN.R-project.org/package=survival>),

plyr (<http://CRAN.R-project.org/package=plyr>), rms (<http://CRAN.R-project.org/package=rms>), cmprsk (<http://CRAN.R-project.org/package=cmprsk>), mstate (<http://CRAN.R-project.org/package=mstate>), ggplot2 (<http://CRAN.R-project.org/package=ggplot2>), nomogramEx (<http://CRAN.R-project.org/package=nomogramEx>).

Results

Clinicopathological characteristics

From January 1, 2010, to December 31, 2016, a total of 2,035 eligible young patients with NSCLC in the SEER database were identified and randomly divided into the training cohort (n=1,357) and validation cohort (n=678) at a ratio of 2:1. The demographic and clinical characteristics were shown in *Table 1*. Specifically, most of the patients in our study were female (54.5%) and lung ADC (72.8%). Median follow-up time was 26 months [interquartile range (IQR), 11–49]. The median survival time was 71 months (*Figure 1*). By the date of the last follow-up, 724 (35.6%) patients had died [642 (31.5%) patients and 82 (4.0%) died from NSCLC and other causes, respectively].

Table 1 Patients' demographics and clinicopathological characteristics

Variables	All patients, N=2,035 (%)	Training cohort, N=1,357 (%)	Validation cohort, N=678 (%)
Sex			
Female	1,109 (54.50)	741 (54.61)	368 (54.28)
Male	926 (45.50)	616 (45.39)	310 (45.72)
Race			
White	1,546 (75.97)	1,025 (75.53)	521 (76.84)
Black	282 (13.86)	196 (14.44)	86 (12.68)
Others ^a	207 (10.17)	136 (10.02)	71 (10.47)
Insurance			
Insured	1,389 (68.26)	915 (67.43)	474 (69.91)
Medicaid	480 (23.59)	331 (24.39)	149 (21.98)
Uninsured	166 (8.16)	111 (8.18)	55 (8.11)
Marital status			
Married	1,098 (53.96)	715 (52.69)	383 (56.49)
Unmarried	937 (46.04)	642 (47.31)	295 (43.51)

Table 1 (continued)

Table 1 (continued)

Variables	All patients, N=2,035 (%)	Training cohort, N=1,357 (%)	Validation cohort, N=678 (%)
Laterality			
Right	1,231 (60.49)	818 (60.28)	413 (60.91)
Left	804 (39.51)	539 (39.72)	265 (39.09)
Histological type			
Adenocarcinoma	1,478 (72.63)	988 (72.81)	490 (72.27)
Squamous carcinoma	495 (24.32)	331 (24.39)	164 (24.19)
Others	62 (3.05)	38 (2.80)	24 (3.54)
Grade			
I	279 (13.71)	191 (14.08)	88 (12.98)
II	853 (41.92)	568 (41.86)	285 (42.04)
III-IV	903 (44.37)	598 (44.07)	305 (44.99)
T stage			
T1	674 (33.12)	450 (33.16)	224 (33.04)
T2	689 (33.86)	468 (34.49)	221 (32.60)
T3	369 (18.13)	244 (17.98)	125 (18.44)
T4	303 (14.89)	195 (14.37)	108 (15.93)
N stage			
N0	1,094 (53.76)	728 (53.65)	366 (53.98)
N1-3	941 (46.24)	629 (46.35)	312 (46.02)
Surgery			
No	577 (28.35)	378 (27.86)	199 (29.35)
Yes	1,458 (71.65)	979 (72.14)	479 (70.65)

^a, others: American Indian/Alaskan native and Asian/Pacific Islander.

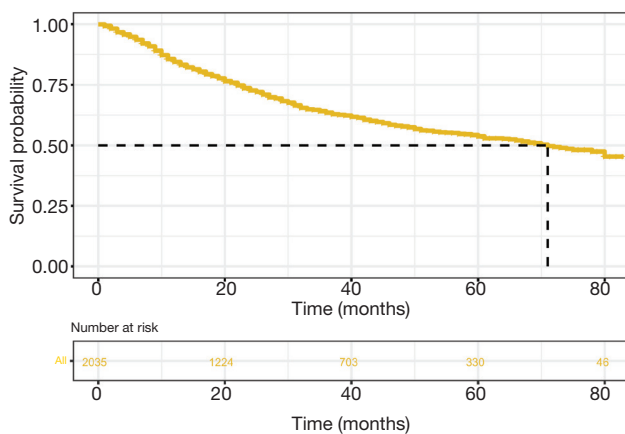


Figure 1 Kaplan-Meier curves for overall survival of all enrolled patients.

Factors associated with OS

The 1-, 2- and 5-year OS of young patients with NSCLC were 84.5%, 72.7% and 52.7%, respectively. After univariate analysis (log-rank test) of OS, all factors excepted race and had strong correlations with OS and subsequently subjected to the multivariate analysis (Cox proportional-hazards regression model). Insurance, marital status, histological type, T stage, N stage and surgery were confirmed to be independent prognostic factors associated with OS ($P < 0.05$) (Table 2).

Factors associated with CSS

Detailed estimates of the cumulative incidence of death

Table 2 Univariate and multivariate analyses of OS in the training cohort

Variables	Univariate analysis		Multivariate analysis	
	P value	HR	95% CI	P value
Sex	<0.001			
Female			Reference	
Male		1.13	0.94–1.37	0.192
Race	0.231	–	–	–
White		–	–	–
Black		–	–	–
Others ^a		–	–	–
Insurance	<0.001			
Insured			Reference	
Medicaid		1.36	1.09–1.69	0.006
Uninsured		1.36	1.01–1.83	0.04
Marital status	<0.001			
Married			Reference	
Unmarried		1.28	1.06–1.55	0.012
Laterality	0.948	–	–	–
Right		–	–	–
Left		–	–	–
Histological type	<0.001			
Adenocarcinoma			Reference	
Squamous carcinoma		1.46	1.19–1.80	<0.001
Others		1.28	0.83–1.98	0.267
Grade	<0.001			
I			Reference	
II		1.27	0.87–1.86	0.212
III–IV		1.57	1.08–2.29	0.017
T stage	<0.001			
T1			Reference	
T2		1.42	1.08–1.87	0.012
T3		1.65	1.22–2.23	0.001
T4		1.98	1.44–2.72	<0.001
N stage	<0.001			
N0			Reference	
N1–3		1.74	1.42–2.14	<0.001
Surgery	<0.001			
No			Reference	
Yes		0.40	0.32–0.50	<0.001

^a, others: American Indian/Alaskan native and Asian/Pacific Islander.

resulting from NSCLC and other causes according to clinical characteristics were listed in *Table 3*. Specifically, the 1-, 2- and 5-year cumulative incidences of death from NSCLC were 13.4%, 24.9% and 40.9%, respectively, while probabilities of death from other causes were 1.3%, 2.2% and 5.2%, respectively. After univariate (Gray's test) and multivariate analysis (subdistribution analysis of competing risks), patients with uninsured status, unmarried status, non-ADC, grade II-IV, T2-4 or N1-3 exhibited a higher cumulative incidence of death of NSCLC.

Nomogram construction and validation

Nomograms were constructed based on the variables which were strongly prognostic factors of OS and CSS after multivariate analysis (*Figure 2*). Surgery made the largest contribution to the prognosis of young patients with NSCLC, followed by T stage, N stage and grade. For the training cohort, the C-index values of nomograms were 0.759 (95% CI: 0.731–0.787) for OS and 0.810 (95% CI: 0.803–0.818) for CSS. Meanwhile the C-index values were 0.751 (95% CI: 0.711–0.790) for OS and 0.807 (95% CI:

Table 3 The 1-, 2-, 5-year cumulative incidence of death among patients in the training cohort

Variables	Cumulative incidence of death resulting from NSCLC							Cumulative incidence of death resulting from other causes			
	Univariate analysis				Multivariate analysis			Univariate analysis			
	1 year	2 years	5 years	P value	HR	95% CI	P value	1 year	2 years	5 years	P value
All patients	0.134	0.249	0.409					0.013	0.022	0.052	
Sex				<0.001							0.406
Female	0.106	0.197	0.364		Reference			0.017	0.020	0.049	
Male	0.167	0.309	0.463		1.16	0.945–1.420	0.160	0.007	0.023	0.055	
Race				0.601							0.860
Black	0.150	0.302	0.429					0.016	0.023	0.047	
Others ^a	0.126	0.199	0.363					0.000	0.032	0.061	
White	0.132	0.244	0.411					0.013	0.020	0.052	
Insurance				<0.001							0.033
Insured	0.116	0.214	0.368		Reference			0.007	0.017	0.047	
Medicaid	0.176	0.312	0.475		1.21	0.947–1.550	0.130	0.025	0.036	0.076	
Uninsured	0.150	0.330	0.533		1.41	1.049–1.900	0.023	0.019	0.019	0.030	
Marital status				<0.001							0.265
Married	0.116	0.205	0.355		Reference			0.006	0.017	0.050	
Unmarried	0.153	0.295	0.467		1.29	1.037–1.600	0.022	0.020	0.027	0.054	
Laterality				0.405							0.066
Left	0.138	0.248	0.435					0.008	0.010	0.037	
Right	0.130	0.249	0.392					0.016	0.030	0.062	
Histological type				<0.001							0.449
Adenocarcinoma	0.094	0.174	0.340		Reference			0.008	0.019	0.048	
Squamous carcinoma	0.241	0.445	0.574		1.39	1.111–1.740	0.004	0.025	0.029	0.068	
Others	0.211	0.400	0.647		1.32	0.829–2.120	0.240	0.026	0.026	0.026	

Table 3 (continued)

Table 3 (continued)

Variables	Cumulative incidence of death resulting from NSCLC							Cumulative incidence of death resulting from other causes			
	Univariate analysis				Multivariate analysis			Univariate analysis			
	1 year	2 years	5 years	P value	HR	95% CI	P value	1 year	2 years	5 years	P value
Grade				<0.001							0.604
I	0.036	0.099	0.216		Reference			0.005	0.013	0.036	
II	0.130	0.233	0.344		1.30	0.879–1.920	0.190	0.009	0.020	0.048	
III–IV	0.167	0.308	0.524		1.55	1.055–2.270	0.025	0.018	0.026	0.059	
T				<0.001							0.406
T1	0.040	0.098	0.183		Reference			0.022	0.031	0.063	
T2	0.108	0.240	0.436		1.78	1.305–2.430	<0.001	0.005	0.015	0.044	
T3	0.187	0.307	0.512		2.10	1.487–2.960	<0.001	0.004	0.014	0.042	
T4	0.334	0.522	0.707		2.33	1.602–3.400	<0.001	0.021	0.028	0.059	
N				<0.001							0.544
N0	0.049	0.122	0.267		Reference			0.012	0.017	0.049	
N1–3	0.231	0.393	0.574		1.71	1.366–2.140	<0.001	0.014	0.027	0.054	
Surgery				<0.001							0.678
No	0.341	0.548	0.745		Reference			0.014	0.021	0.056	
Yes	0.052	0.132	0.281		2.55	2.006–3.250	<0.001	0.012	0.022	0.050	

^a, others: American Indian/Alaskan native and Asian/Pacific Islander. NSCLC, non-small cell lung cancer.

0.795–0.819) for CSS in the validation cohort. Calibration plots showed good consistency between the actual and estimated outcomes in both internal and external validation (Figures 3,4).

Comparison of nomograms

The C-index values of nomograms based TNM-staging system were 0.702 (95% CI: 0.729–0.674) for OS and 0.750 (95% CI: 0.743–0.758) for CSS in the training cohort, which were significantly lower than those of nomogram that integrate all independent prognostic factors. In the validation cohort, C-index in our study were also significantly greater than those of TNM-system nomograms with C-index at 0.691 (95% CI: 0.730–0.652) and 0.746 (95% CI: 0.733–0.758) for OS and CSS, respectively.

Performance of the nomograms

The total risk score for each patient was calculated

based on nomogram of OS, cut-off values for risk group stratification were determined by evenly dividing patients in the training cohort into three different risk groups. Significant differences of survival were observed in patients within different risk groups even in each TNM stage in the validation cohort (Figure 5).

Discussion

Young patients with NSCLC is a rare but distinctive subset. Although early studies found young patients with NSCLC have a similar outcome with other groups (9,10), recent studies have confirmed this subgroup has improved survival. The AJCC staging system could be used to predict prognosis for population, but patients with the same stage of TNM may have distinct outcomes mainly because various factors affect the prognosis of patients. Therefore, identifying prognostic factors for young patients with NSCLC is crucial. Therefore, how to identify and integrate all prognostic factors is important to predict the outcome

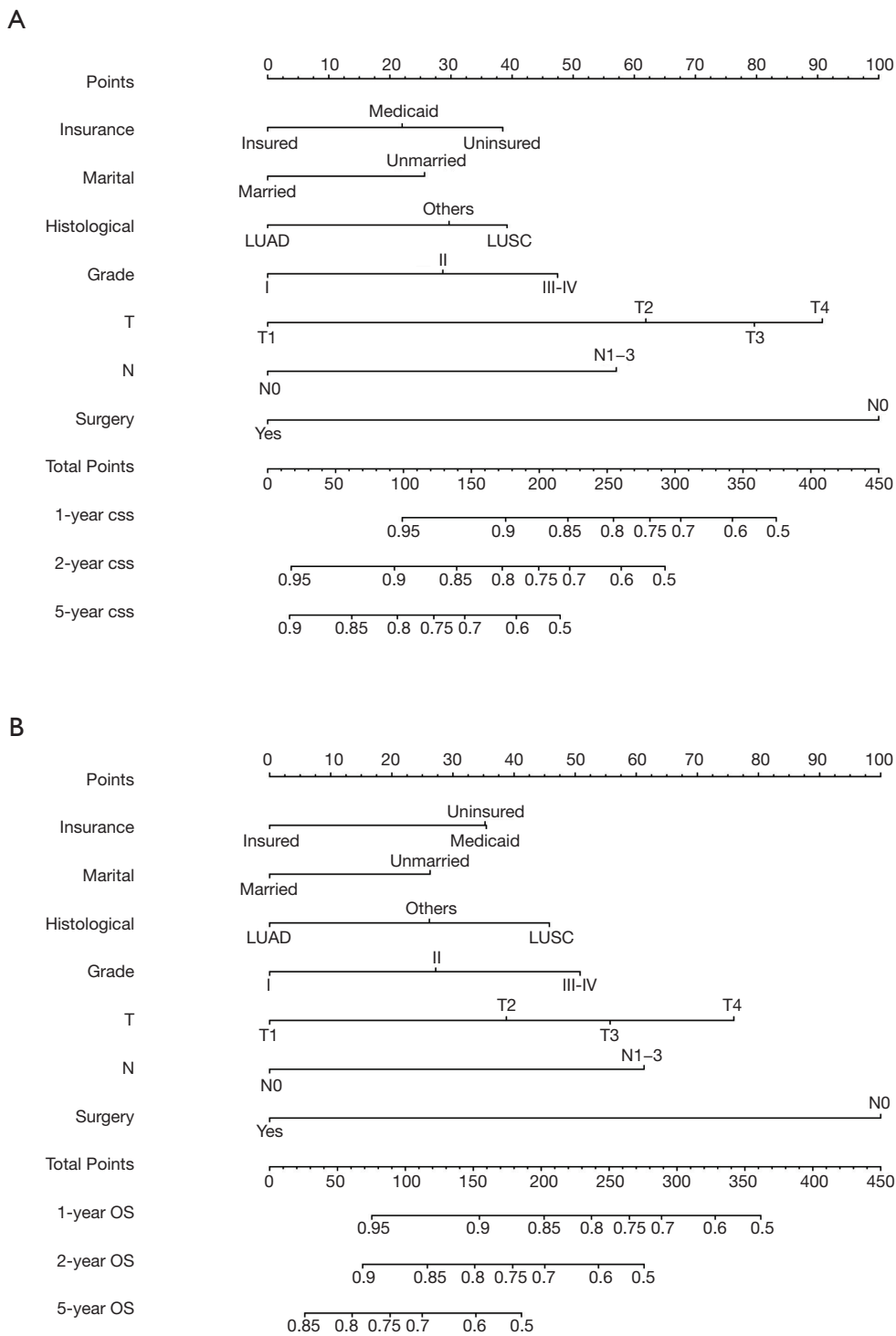


Figure 2 Nomograms for predicting the 1-, 2-, and 5-year (A) OS and (B) CSS of young patients with NSCLC. OS, overall survival; CSS, cancer-specific survival; NSCLC, non-small cell lung cancer.

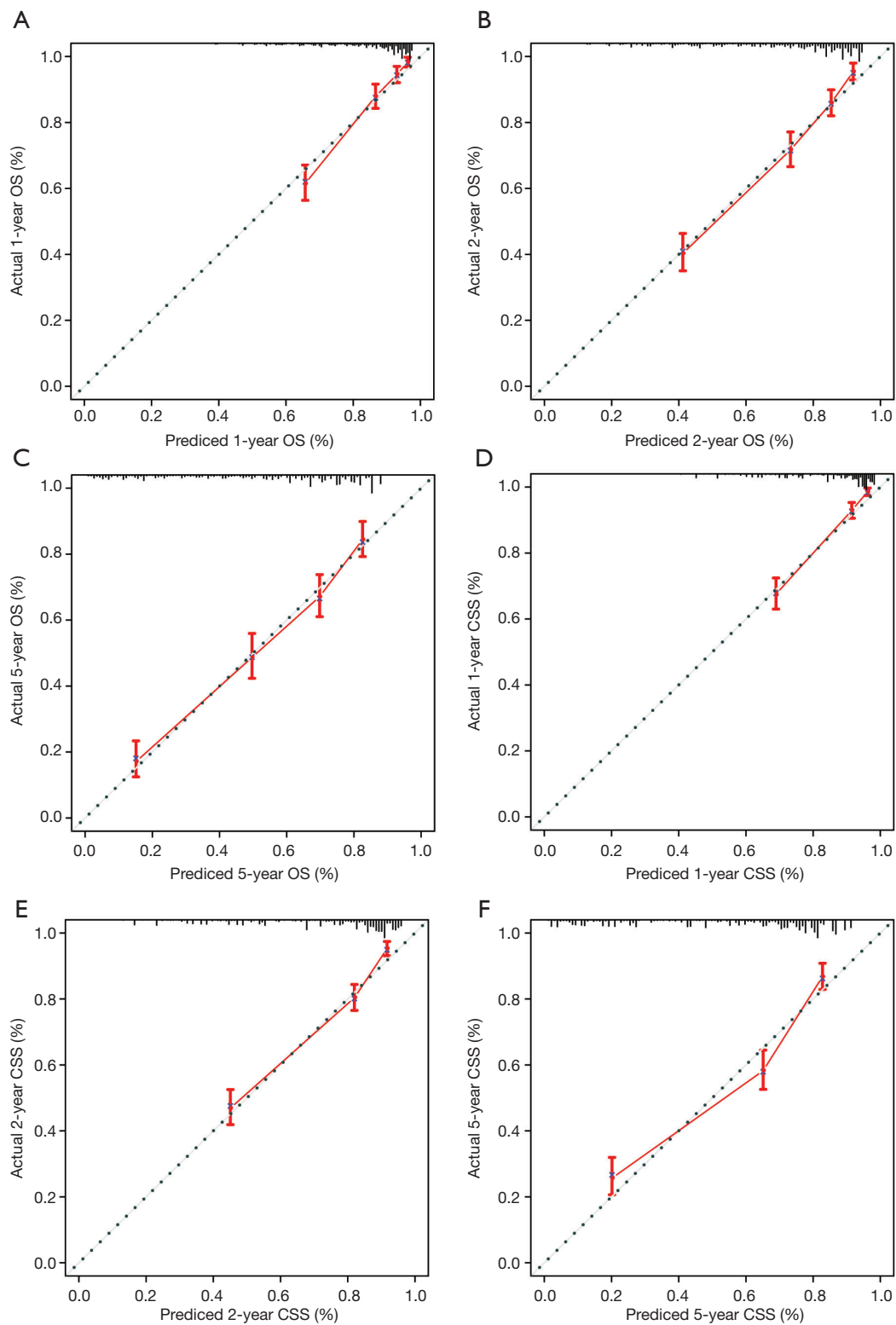


Figure 3 Internal calibrate curves for 1-, 2-, and 5-year OS (A-C) and CSS (D-F). The 45° line represents a perfect match between the nomogram-predicted survival (X-axis) and the actual survival (Y-axis). Vertical line at the top of plot indicates the 95% CI. OS, overall survival; CSS, cancer-specific survival.

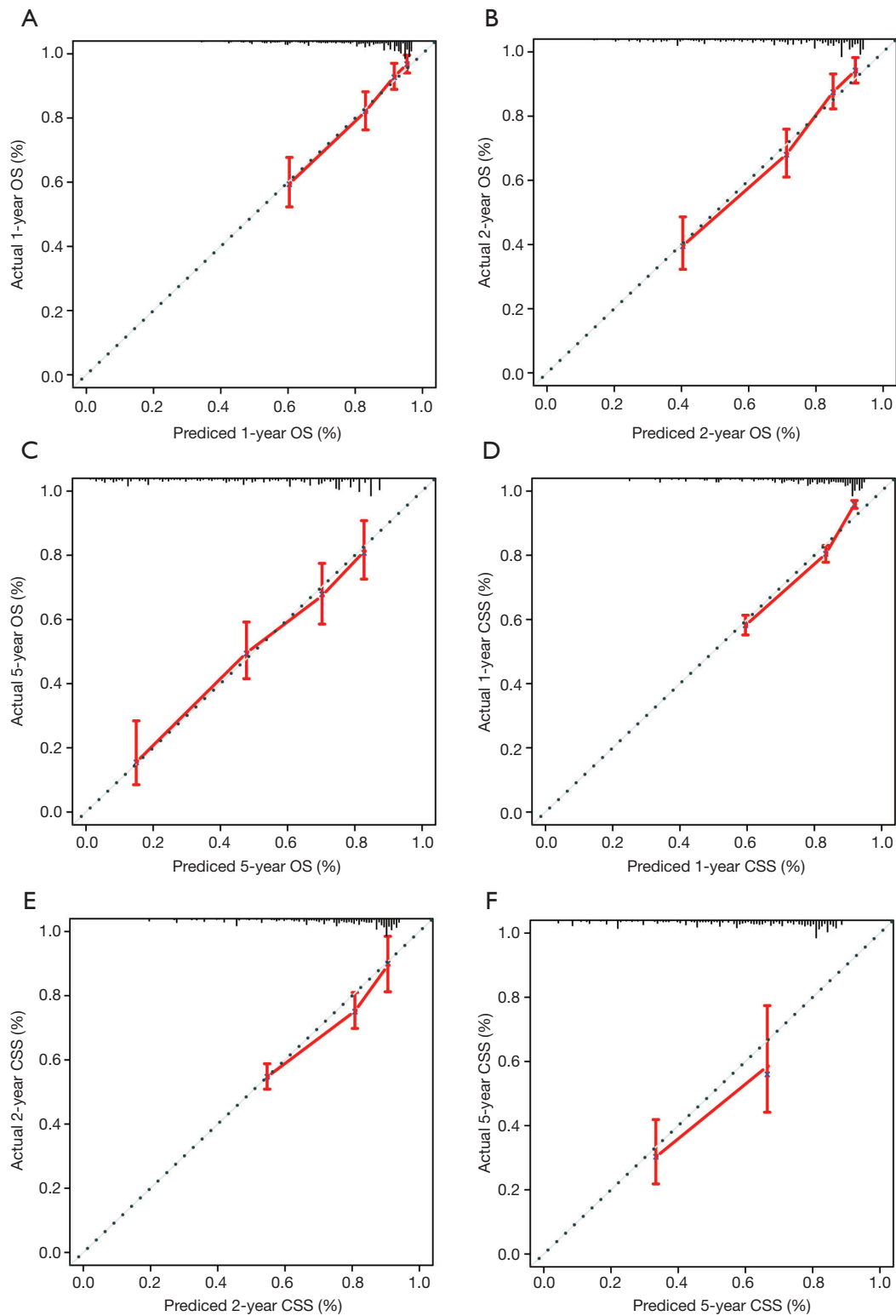


Figure 4 External calibrate curves for 1-, 2-, and 5-year OS (A-C) and CSS (D-F). The 45° line represents a perfect match between the nomogram-predicted survival (X-axis) and the actual survival (Y-axis). Vertical line at the top of plot indicates the 95% CI. OS, overall survival; CSS, cancer-specific survival.

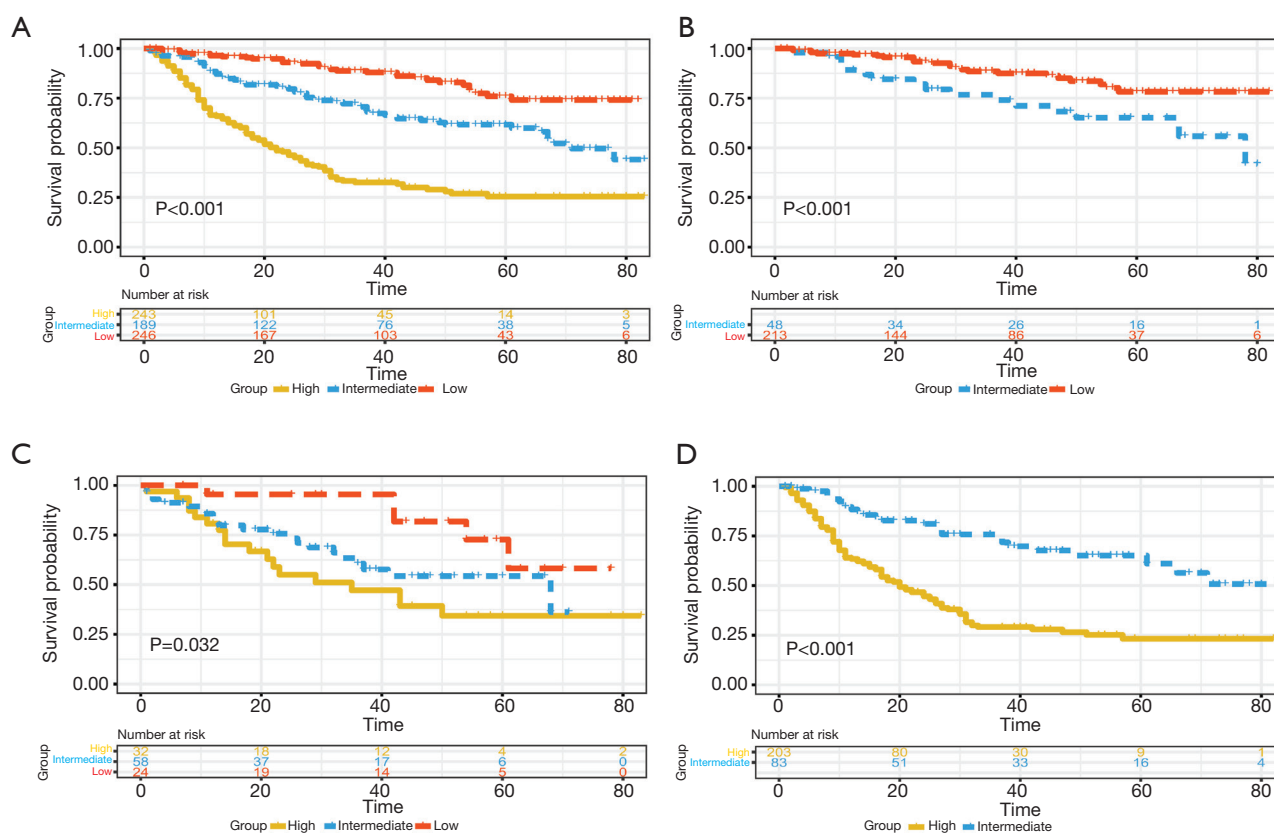


Figure 5 Kaplan-Meier curves for OS of patients (A, all patients; B-D, stage I-III) stratified by risk score within each TNM stage in the validation cohort. Subgroups less than 10 patients were not shown in graphs.

accurately. Nomogram, a visualized calculating score tool, is effectively used to estimate the probability of survival at certain time point by calculating the accumulative effect of all prognostic factors. Nevertheless, no prognostic nomogram for young patients with NSCLC had been constructed ascribed to the limited number of cases in a single institution.

Our study enrolled 2,035 eligible patients between 2010 and 2016 from the SEER database which is a population-based database. Patients who were diagnosed between 2010 and 2016 were included because the long-time span may have a direct impact on the results of the study. On the one hand, the incidence of the smoking population has fallen significantly and the structure of sexual distinction has changed (11). On the other hand, therapeutic strategies have been also improved over time, especially the application of targeted therapy (12).

The log-rank test and Cox proportional hazards regression were used to identify the independent prognostic factors for OS in our study. But traditional Cox proportional

hazard model is not adequate for the analysis of CSS, because the cause-specific Cox model treats the competing risks of the event of interest as censored observations and does not directly translate to the survival probability (cumulative incidence) (13). Therefore the subdistribution hazard regression model was applied to the analysis of CSS. Insurance status, marital status, histological type, grade, T stage, N stage and surgery were closely associated with the OS and CSS of the young patients with NSCLC, except for sex, race and laterality. Consistent with previous studies, non-ADC histology, advanced stage and non-surgery were independent negative prognostic factors. All independent prognosis factors were integrated to construct nomograms. The nomograms showed great discrimination and reliability. Specifically, the C-index of our nomograms for OS and CSS were significantly higher than those based on the TNM staging system internally and externally. Calibration plots showed optimal agreement between predicted and actual probabilities of 1-, 2- and 5-year OS and CSS.

There are some limitations to our study. First, the follow-

up period was relatively short, as the data for re-staging based on the 7th TNM staging system was incomplete until 2010. Second, some important information was not captured in the SEER database including smoking status, histological subtype, etc. These factors might affect the effectiveness of the nomograms. Although information on radiation therapy and chemotherapy can be accessed from the SEER database, chemotherapy and radiotherapy were not selected as candidate factors due to the “no treatment” and “unknown if patients received treatment” were mixed up. Third, Patients were staged according to the 7th edition of the TNM classification, because the information for re-staging based on 7th TNM staging system in SEER were insufficient. Finally, although the eligible patients were randomly divided into the training and validation cohorts to evaluate the nomograms internally and externally, another independent population-based database or prospectively cohort is needed to validate our nomograms before clinical application.

Conclusions

Seven independent prognostic factors (including insurance, marital status, histological type, grade, T stage, N stage and surgery) for OS and CSS in young patients with NSCLC were identified and further utilized to develop nomograms based the large population database. The nomograms performed excellently internally and externally, which may help clinicians predict the prognosis of patients and make a decision for treatment.

Acknowledgments

Funding: This work was supported by Ministry of Science and Technology of the People’s Republic of China (2017YFA0505500; 2016YFA0501800), Science and Technology Commission of Shanghai Municipality (19XD1401300).

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/jtd.2020.03.03>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related

to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was approved by Shanghai Cancer Center Ethical Committee. Informed patient consent is not required for the data released by the SEER database because cancer is a reportable disease in every state in the United States.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Dell'Amore A, Monteverde M, Martucci N, et al. Surgery for non-small cell lung cancer in younger patients: what are the differences? *Heart Lung Circ* 2015;24:62-8.
2. Subramanian J, Morgensztern D, Goodgame B, et al. Distinctive characteristics of non-small cell lung cancer (NSCLC) in the young: a surveillance, epidemiology, and end results (SEER) analysis. *J Thorac Oncol* 2010;5:23-8.
3. Thomas A, Chen Y, Yu T, Trends and characteristics of young non-small cell lung cancer patients in the United States. *Front Oncol* 2015;5:113.
4. Xia W, Wang A, Jin M, et al. Young age increases risk for lymph node positivity but decreases risk for non-small cell lung cancer death. *Cancer Manag Res* 2018;10:41-8.
5. Zhang Y, Jheon S, Li H, et al. Results of low-dose computed tomography as a regular health examination among Chinese hospital employees. *J Thorac Cardiovasc Surg* 2019. [Epub ahead of print].
6. 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;17:1464-72.
7. Fakhry C, Zhang Q, Nguyen-Tân PF, et al. Development and Validation of Nomograms Predictive of Overall and Progression-Free Survival in Patients With Oropharyngeal Cancer. *J Clin Oncol* 2017;35:4057-65.
8. Callegaro D, Miceli R, Bonvalot S, et al. Development and

- external validation of two nomograms to predict overall survival and occurrence of distant metastases in adults after surgical resection of localised soft-tissue sarcomas of the extremities: a retrospective analysis. *Lancet Oncol* 2016;17:671-80.
9. Gadgeel SM, Ramalingam S, Cummings G, et al. Lung cancer in patients < 50 years of age: the experience of an academic multidisciplinary program. *Chest* 1999;115:1232-6.
 10. Maruyama R, Yoshino I, Yohena T, et al. Lung cancer in patients younger than 40 years of age. *J Surg Oncol* 2001;77:208-12.
 11. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin* 2018;68:7-30.
 12. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *Nature* 2018;553:446-54.
 13. Zhang Z. Survival analysis in the presence of competing risks. *Ann Transl Med* 2017;5:47.

Cite this article as: Peng Y, Sun Y. Development and validation of nomograms for predicting overall and cancer-specific survival in young patients with non-small cell lung cancer. *J Thorac Dis* 2020;12(4):1404-1416. doi: 10.21037/jtd.2020.03.03