

The characteristics and nomogram for primary lung lepidic adenocarcinoma

Hui Tang

Peking Union Medical College Hospital <https://orcid.org/0000-0001-8870-6726>

Yingyi Wang (✉ waltwyy@163.com)

Department of Medical Oncology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing 100730, China

Chunmei Bai

Department of Medical Oncology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing 100730, China

Research

Keywords: lung cancer, lepidic adenocarcinoma, clinicopathological feature, prognostic evaluation, nomogram, SEER database

Posted Date: February 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-246772/v1>

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

Abstract

Background Lepidic adenocarcinoma (LPA) is an infrequent subtype of invasive pulmonary adenocarcinoma (ADC). However, the clinicopathological features and prognostic factors of LPA have not been elucidated. We undertook a retrospective population-based analysis to examine the clinicopathological features of LPA, and construct nomograms predicting long-term survival of LPA patients. Methods Data from the Surveillance, Epidemiology, and End Results (SEER) database of 4087 LPA patients diagnosed between 2005 and 2014 were retrospectively analyzed and compared with non-LPA pulmonary ADC to explore the clinicopathological and prognosis features of LPA. All patients included in our study were histologically confirmed. Those with multiple primary tumors in their lifetime, unknown survival data, unknown tumor-node-metastasis (TNM) stage, and unknown information including age, race, and marital status were excluded. Univariate and multivariate Cox proportional hazard models were performed to identify independent survival predictors for further nomogram development. The nomograms were internally and externally validated for concordance index and calibration plots as well as decision curve analysis. Results Compared with non-LPA pulmonary ADC patients, those with LPA exhibited unique clinicopathological features, including more elderly and female patients, smaller tumor size, lesser pleural invasion, lower histological grade and stage. Multivariate analyses showed that age, sex, marital status, primary tumor size, pleural invasion, histological grade, stage, primary tumor surgery, and chemotherapy were independently associated with overall survival (OS) and cancer-specific survival (CSS) in patients with LPA, while race was the only independent prognostic factor for OS, not for CSS. The nomograms showed good accuracy comparing with actual observed results and demonstrated improved prognostic capacity than TNM stage. Conclusions Patients with LPA are more common in older age and female. The smaller tumor size, lower histological grade and stage are the clinicopathological features of LPA which may indicate a good prognosis. The constructed nomograms accurately predict the long-term survival of LPA patients.

Introduction

Lung cancer is the leading cause of cancer death and the most commonly diagnosed cancer worldwide^[1]. Lepidic adenocarcinoma (LPA), also known as lepidic predominant adenocarcinoma or nonmucinous bronchioloalveolar carcinoma^[2], is an infrequent subtype of lung adenocarcinoma (ADC) without precise incidence data. LPA is defined as an ADC of > 3 cm in tumor size and/or has > 5 mm lymphatic, vascular, or pleural invasion with a nonmucinous lepidic predominant growth pattern^[3]. The definition was proposed by the International Association for the Study of Lung Cancer in 2011 and subsequently accepted by the World Health Organization (WHO) in 2015^[4]. LPA exhibits unique clinicopathological features, specific gene mutation profiles, and desirable survival outcomes compared with lung adenocarcinoma, not otherwise specified (NOS)^[4-6]. However, very few population-based studies have been completed on analysis of the demographic and clinicopathological features as well as the factors influencing the prognosis of LPA. Meanwhile, it is quite challenging for clinicians to accurately predict the prognosis of patients only relying on tumor-node-metastasis (TNM) stage. Therefore, it is necessary to develop tools for estimating the probability of long-term survival in patients with LPA.

The Surveillance, Epidemiology, and End Results (SEER) database provides a wide range of demographic, clinical and follow-up information of cancer patients, which was established in 1973 and covered about 28% of the population in USA^[7]. Using the SEER database, we retrospectively analyzed the clinicopathological features and survival of 4087 LPA patients to confirm their clinicopathological characteristics and prognostic factors. We then developed the nomograms estimating overall survival (OS) and cancer-specific survival (CSS) for patients with LPA. Furthermore, we performed the internal and external validation as well as decision curve analysis (DCA) for nomograms to evaluate the

accuracy of nomograms. Besides, we estimated the incidence of LPA and explored the risk factors associated with distant and lymph node metastases of LPA.

Methods

2.1. Data source and selection

Patient data were obtained from the SEER database using the SEER*stat software, version 8.3.8 (<https://seer.cancer.gov/seerstat/>). Lung adenocarcinoma was classified according to the 2015 WHO classification system. International Classification of Diseases for Oncology, third edition (ICD-O-3) histology code was used in this study to identify patients. The inclusion criteria were: (1) primary lung cancer; (2) the ICD-O-3 histology code was 8250/3 (lepidic adenocarcinoma) or 8140/3 (adenocarcinoma-NOS); (3) positive histological confirmation; (4) diagnosed between 2005 and 2014. The exclusion criteria were: (1) patients who had multiple primary tumors in their lifetime; (2) unknown survival data and TNM stage; (3) unknown important and easily accessible information in clinical practice including age at diagnosis, race, and marital status. The unique demographic and clinicopathological features were explored and compared between LPA group and ADC-NOS group. After propensity score matching (PSM), OS and CSS were compared between ADC-NOS and LPA patients. For creating and validating the nomograms, patients with LPA diagnosed in 2009 and 2010 (n = 979) were assigned into validation cohorts, and those diagnosed between 2005 and 2014, except for 2009 and 2010 (n = 3108) were assigned into training cohorts.

2.2. Study variables

Demographic and clinicopathological variables of the included patients were extracted, including age, sex, race, marital status, tumor location, primary tumor size, separate tumor nodules, pleural invasion, histological grade, 6th edition TNM stage, treatment, vital status, survival time, corresponding death causes, and the status of education and income in the county where patients resided in. In the present study, other race was recorded as "Other", except for the white and the black. "Married (including common law)" was recorded as "Married", other marital statuses were recorded as "Single". The status of education and income were defined either as "Low" or "High", meaning that patients resided in counties with lower/higher education or income than median level. Considering there was no precise survival day data, the survival time of 0 months was recorded as 0.5 months. OS was defined as the time period from the diagnosis to the death caused by any cause or last follow-up, while CSS was defined as the time period from the diagnosis to the death caused by lung cancer.

2.3. Statistical analysis

For descriptive statistics, absolute numbers with proportion for variables was used. Chi-square test was used to compare the demographic and clinicopathological characteristics among different groups. A PSM method was used to minimize the impact of confounding factors. The propensity score for every patient with ADC-NOS or LPA was calculated with a logistic regression model, which included the following variables: age, sex, race, marital status, income and education level, primary tumor location and size, separate tumor nodules, pleural invasion, histological grade, TNM stage, and treatment. The caliper matching within a caliper of 0.02 was performed among two groups. With PSM analysis to exclude certain confounding factors, 4060 pairs of patients were successfully matched among patients included in our study. After PSM, OS and CSS were compared between patients with ADC-NOS and LPA by Kaplan-Meier curves and log-rank test. Then, the data of LPA patients were used for further analyses. Multivariate binary logistic regression analyses were performed to identify risk factors for distant and lymph node metastases in all LPA patients. Univariate and multivariate Cox proportional hazard models were performed to calculate the hazard ratios (HR) with 95% confidence intervals (CI) of variables associated with OS and CSS in the training cohort of LPA

patients. Based on multivariate Cox analyses, nomograms were constructed and evaluated by the concordance index (C-index) and calibration curves, which were used for the comparison between the observed and nomogram-predicted survival outcomes. Ultimately, decision curve analysis (DCA) was performed to compare the prognostic capacity of the nomogram model and TNM stage. To verify the applicability of the nomogram model, nomograms were both internally and externally validated.

The ages of patients were stratified by using the X-tile program (Yale University, USA) [8]. According to the cutoff value of age, in terms of OS, determined by X-tile analysis (Supplementary Figure S1A-C), the patients were divided into 3 groups (0–69, 70–79, and 80+ years old). All statistical analyses were performed using SPSS software version 21.0 (IBM Inc) or R version 3.6.1 (<http://www.r-project.org/>). Two-tailed value of $P < 0.05$ was considered to be statistically significant.

Results

3.1. Patients and tumor characteristics

Of the 1,244,493 patients diagnosed with a primary lung or bronchus malignancy in the SEER database between 1975 and 2016, a total of 27,142 patients were diagnosed with LPA, which were accounted for 2.18% of all lung cancer. After applying the criteria of inclusion and exclusion, the number of patients with lung ADC-NOS and LPA enrolled in our study were 84267 and 4087, respectively. The demographic and clinicopathological characteristics of the eligible patients were shown in Table 1. Among the eligible patients, those with LPA were more common in older age, female, and yellow race. Besides, patients with LPA were more inclined to smaller tumor size, fewer separate tumor nodules, lesser pleural invasion, lower histological grade and stage.

After PSM, Kaplan-Meier curves were performed and showed that patients with LPA had better survival outcomes than those with ADC-NOS (Supplementary Figure S2A-B). Moreover, the median OS of the total 4060 LPA patients who were successfully matched by PSM was 49 months (Supplementary Figure S2A).

3.2. Factors associated with distant and lymph node metastases

As shown in Supplementary Table S1, the factors significantly associated with distant metastasis were identified by chi-square test and further examined by multivariate analysis, which showed that yellow race, the large size of the tumor, positive separate tumor nodule, and higher histological grade were the independent risk factors for distant metastasis. However, age, sex, race, tumor size, separate tumor nodule, pleural invasion, and histological grade were significantly associated with lymph node metastasis in the multivariate analysis (Supplementary Table S2).

3.3. Establishment of the nomograms predicting OS and CSS of LPA patients

In the training cohort, univariate analysis showed that age, sex, race, marital status, education, income, tumor location, primary tumor size, separate tumor nodule, pleural invasion, histological grade, TNM stage, primary tumor surgery, radiotherapy, and chemotherapy were significantly associated with OS (Table 2). Further multivariate analysis showed that age, sex, race, marital status, primary tumor size, pleural invasion, histological grade, TNM stage, primary tumor surgery, and chemotherapy were significantly associated with OS. While multivariate analysis identified that age, sex, marital status, primary tumor size, pleural invasion, histological grade, TNM stage, primary tumor surgery, and chemotherapy were significantly associated with CSS (Table 2). According to the multivariate results, two nomograms predicting the survival probability of 1- and 5- year OS (Figure 1) and CSS (Figure 2) were constructed with these independent variables.

To use the nomograms, each variable was first assigned to a specific score by the point scale at the top of the nomograms. Based on the sum of those scores, the point scale at the bottom of nomograms was used to estimate the survival probability of one individual patient.

3.4. Nomogram validation

The internal and external validation for both OS and CSS nomograms were performed. The internal validation was performed via the training cohort, and the C-index values of the nomogram predicting OS and CSS were 0.786 (95% CI, 0.776-0.796) and 0.812 (95% CI, 0.802-0.822), respectively. The external validation was performed via the validation cohort, and the C-index values of the nomogram predicting OS and CSS were 0.781 (95% CI, 0.762-0.800) and 0.812 (95% CI, 0.793-0.831), respectively. Furthermore, internal and external calibration plots both indicated that the OS and CSS nomograms demonstrated excellent agreement between the predicted and actual survival outcome (Figure 3A-H). Besides, the DCA results demonstrated that the nomograms showed better prognostic capacity than TNM stage (Supplementary Figure S3A-D).

Discussion

Concise and accurate prognostic prediction models for patients with malignancy are essential for clinical decision-making and scientific research. Clinically, TNM stage is the most widely used survival predictor for cancer patients. Therefore, identifying more prognostic factors and a more individualized model will certainly improve the accuracy of clinical outcome prediction. In this study, we used the SEER database, a large-scale population-based cancer registry program, to explore the clinical characteristics of 4087 patients with LPA and identified the factors associated with distant and lymph node metastases in LPA patients. After that, we developed and validated accurate and personalized prognostic nomograms predicting 1- and 5-year OS and CSS of patients with LPA. Moreover, the nomograms demonstrated better prognostic capacity than TNM stage.

Survival outcomes of LPA patients with poor prognostic factors were undesirable, the median OS of advanced LPA patients was 20.1 months^[9]. However, the prognosis of advanced LPA patients could be improved by appropriate treatments, including chemotherapy and EGFR tyrosine kinase inhibitors (TKIs)^[9]. The 5-year disease-free survival of LPA patients after complete surgical resection was about 90%^[10]. With the evaluation of the nomograms that generated in our study, more aggressive treatments are recommended for high-risk patients with LPA, and appropriate shortening of the follow-up interval is encouraged to detect the occurrence of endpoint events as early as possible. For example, older, unmarried, black men with sizeable tumor and advanced TNM stage are recommended for frequent follow-up and more aggressive treatments, including primary tumor resection when they meet the operational criteria.

Compared with other rare histologic subtypes of lung cancer, such as papillary adenocarcinoma^[11] and carcinosarcoma^[12], our results suggested the incidence of LPA was much higher. Our results indicated that female and older age were highly associated with LPA rather than ADC-NOS, which is consistent with previous studies^[13,14]. In addition, some clinicopathological features of LPA patients indicate a good prognosis, including smaller tumor size, fewer separate tumor nodules, lesser pleural invasion, lower histological grade and stage. This is consistent with previous studies^[15] and in line with the good prognosis of LPA^[3,13,15]. Moreover, LPA have some characteristics differing from other histologic subtypes of invasive pulmonary ADC, such as being more common in non-smokers or light smokers, a preference for pulmonary peripheral location and be false-negative in positron-emission tomographic scan^[13,16]. Clinically, asymptomatic at presentation and excessive airway secretion were more common in patients with LPA^[17]. In the genetic alteration profiles, EGFR mutations were occurred in about 50% patients with LPA, which is significantly higher than other subtypes^[5], especially the mutations in exon 21^[17,18]. Whereas KRAS mutations were

much less common that account for about 10% of LPA population [5]. Comparing with other histologic subtypes, a lower rate of ALK rearrangement and a higher rate of RET rearrangement were reported [6,19,20].

Most studies supported that patients with LPA had desirable survival outcomes compared with other subtypes of invasive pulmonary ADC. In the treatment strategies, surgery is still the superior option for LPA patients whereas adjuvant chemotherapy including oral fluoropyrimidines and platinum-based regimens conferred no survival benefit on patients with LPA, regardless of the tumor stage at presentation [21,22]. In patients with advanced LPA, studies suggested that taxane-based chemotherapy and pemetrexed might be effective and well-tolerated in LPA [23,24]. With higher frequencies of EGFR mutations, EGFR-TKIs as first- or second-line treatment for advanced LPA demonstrated encouraging efficacy [9]. Nevertheless, due to the lower expression level of programmed cell death-ligand 1, the efficacy of immune checkpoint inhibitors in patients with LPA may be poor [25-27]. Moreover, multiple studies suggested that a higher percentage of lepidic growth patterns were associated with a lower risk of recurrence, and the invasive component size were better predictors of survival than overall tumor diameter [16,17,28,29]. Furthermore, no recurrence was observed in any of the 18 LPA patients with maximum tumor diameter > 3 cm but the maximum diameter of the invasive area < 5 mm [30]. Therefore, Suzuki et al. [30] proposed that LPA with an invasion of 5 mm or less can be regarded as minimally invasive ADC even if the tumor is larger than 3 cm in diameter. Unsurprisingly, our results suggested that primary tumor surgery was a major prognostic factors of LPA patients following the histological grade and stage. By contrast, chemotherapy was far less important to the prognosis of LPA patients. Furthermore, our results suggested that radiotherapy had no significant effect on the survival outcomes of LPA patients. Regrettably, we could not explore the prognostic significance of chemotherapy regimens, targeted therapy, immunotherapy and the diameter of the invasive area.

In the current study, we identified that age, sex, marital status, primary tumor size, pleural invasion, histological grade, TNM stage, primary tumor surgery, and chemotherapy were independently associated with OS and CSS in patients with LPA. It is of note that few patients with histological grade IV LPA were included in this study. Therefore, the nomograms we constructed to predict the survival outcomes were not suitable for patients with histological grade IV LPA. Similar to previous studies, our results suggested that treatment, tumor size and some demographic characteristics also had an impact on the prognosis of LPA patients, and we provided a statistical prediction tool that can incorporate and quantify the selected prognostic factors to estimate the survival outcome for an individual patient.

To date, it is the first time that the demographic and clinicopathological features, as well as the incidence of LPA, were elucidated based on a large-scale population-based database. Meanwhile, this is the first nomograms predicting the survival outcomes in LPA patients, which could aid in the personalized prognostic evaluation and clinical decision-making. However, there were still some limitations in our study though the nomograms demonstrated good accuracy and applicability. First, the nomograms were constructed based on the retrospective data, and the prospective validation is needed. Second, some critical information such as the diameter of the invasive area in LPA, tumor biomarkers, chemotherapy regimens, targeted therapy, molecular pathology, and genetic tests were absent in the database. Therefore, we could not analyze those variables and improve prognostic nomograms in our study. Third, the patients were almost all Americans, and the results might be different in other races. Such drawbacks are inherent to almost all retrospective population-based studies. However, the large size and the long follow-up duration of the present study compensate to a great extent and provide a comprehensive knowledge of LPA. Further prospective studies with more important information are needed for model improvement and independent validation.

Conclusion

In summary, we explored the clinical characteristics of LPA patients and developed nomograms predicting OS and CSS of LPA patients individually. The nomograms showed good accuracy and applicability which may provide the function of individualized prognostic prediction for LPA patients and aid in clinical decision-making.

Declarations

Funding

This work was supported by grants from CAMS Innovation Fund for Medical Sciences (No. 2016-I2M-1-001).

Competing interests

No potential conflict of interest was reported by the authors.

Acknowledgment

The authors would like to thank the Surveillance, Epidemiology, and End Results (SEER) database for the support.

Availability of data and materials

The datasets generated during and/or analysed during the current study are available in the Surveillance, Epidemiology, and End Results (SEER) database (<https://seer.cancer.gov/>). R code is available upon request.

Contributions

H T, YY W, and CM B conceived of and designed the study. H T and YY W performed literature search. H T generated the figures and Tables. H T and YY W analyzed the data. H T wrote the manuscript, YY W and CM B critically reviewed the manuscript. CM B supervised the research. All authors have read and approved the manuscript.

Ethics approval and consent to participate

We received permission to access the research data file in the SEER program from the National Cancer Institute, US (reference number 12317-Nov2019). Approval was waived by the local ethics committee, as SEER data is publicly available and de-identified.

Consent for publication

Not applicable.

Tables

Table 2. Univariate and multivariate analyses of OS and CSS in the training cohort of LPA patients (n = 3108).

Characteristics	Univariate analysis P-value for OS	Multivariate analysis for OS			Univariate analysis P-value for CSS	Multivariate analysis for CSS		
		HR	95% CI	P-value		HR	95% CI	P-value
Age, years old	<0.001				<0.001			
0-69		1.000				1.000		
70-79		1.298	(1.169,1.442)	<0.001		1.189	(1.058,1.338)	0.004
>=80		1.557	(1.374,1.765)	<0.001		1.261	(1.092,1.456)	0.002
Sex	<0.001				<0.001			
Male		1.000				1.000		
Female		0.626	(0.570,0.688)	<0.001		0.655	(0.590,0.728)	<0.001
Race	0.032				0.010			
White		1.000				1.000		
Black		1.030	(0.874,1.214)	0.724		1.043	(0.872,1.248)	0.646
Other		0.853	(0.734,0.992)	0.039		0.859	(0.728,1.015)	0.074
Marital status	<0.001				0.004			
Married		1.000				1.000		
Single		1.175	(1.070,1.290)	0.001		1.125	(1.013,1.251)	0.028
Education	0.003				0.006			
Low		1.000				1.000		
High		0.985	(0.888,1.091)	0.768		0.990	(0.892,1.100)	0.853
Income	0.049				0.064			
Low		1.000				NI		
High		0.986	(0.890,1.091)	0.781				
Laterality	<0.001				<0.001			
Left		1.000				1.000		
Right		1.043	(0.947,1.148)	0.398		1.012	(0.906,1.129)	0.836
Unknown		1.347	(1.035,1.753)	0.027		1.458	(1.111,1.914)	0.007
Lobe	<0.001				<0.001			
Upper		1.000				1.000		
Middle		0.982	(0.790,1.220)	0.867		1.075	(0.959,1.205)	0.217
Lower		1.092	(0.987,1.207)	0.088		1.027	(0.808,1.305)	0.829
Unknown		0.935	(0.785,1.114)	0.454		0.878	(0.725,1.063)	0.182
Tumor Size	<0.001				<0.001			
<3cm		1.000				1.000		
>=3 & <5cm		1.277	(1.131,1.443)	<0.001		1.361	(1.182,1.569)	<0.001
>=5cm		1.915	(1.669,2.197)	<0.001		2.159	(1.853,2.514)	<0.001
Unknown		1.956	(1.674,2.285)	<0.001		2.058	(1.736,2.439)	<0.001
Separate tumor nodules	<0.001				<0.001			
Yes		1.000				1.000		
No		1.192	(0.961,1.478)	0.109		1.142	(0.908,1.437)	0.256
Unknown		1.347	(1.117,1.623)	0.002		1.277	(1.051,1.552)	0.014
Pleural invasion	<0.001				<0.001			
Yes		1.000				1.000		
No		0.499	(0.332,0.751)	0.001		0.433	(0.277,0.676)	<0.001
Unknown		0.737	(0.510,1.064)	0.103		0.671	(0.455,0.990)	0.044
Grade	<0.001				<0.001			
I		1.000				1.000		
II		1.206	(1.065,1.366)	0.003		1.233	(1.067,1.425)	0.004
III		1.637	(1.364,1.964)	<0.001		1.776	(1.453,2.172)	<0.001
IV		0.451	(0.063,3.225)	0.428		0.593	(0.083,4.250)	0.603
Unknown		1.158	(1.029,1.303)	0.015		1.165	(1.020,1.331)	0.024
Stage	<0.001				<0.001			
I		1.000				1.000		
II		2.074	(1.666,2.583)	<0.001		2.649	(2.068,3.393)	<0.001
III		2.268	(1.958,2.626)	<0.001		2.923	(2.467,3.464)	<0.001
IV		3.079	(2.656,3.570)	<0.001		4.075	(3.435,4.834)	<0.001
Primary tumor surgery	<0.001				<0.001			

Yes		1.000			1.000		
No/unknown		2.383	(2.068,2.748)	<0.001	2.533	(2.162,2.968)	<0.001
Metastatic tumor surgery	0.511				0.053		
Yes		NI			NI		
No/unknown							
Radiotherapy	<0.001				<0.001		
Yes		1.000			1.000		
No/unknown		0.921	(0.812,1.043)	0.195	0.960	(0.838,1.099)	0.556
Chemotherapy	<0.001				<0.001		
Yes		1.000			1.000		
No/unknown		1.329	(1.191,1.482)	<0.001	1.260	(1.119,1.418)	<0.001

Abbreviations: CI, confidence interval; CSS, cancer-specific survival; HR hazard ratio; LPA, lepidic adenocarcinoma; NI, not included; OS, overall survival.

Figures

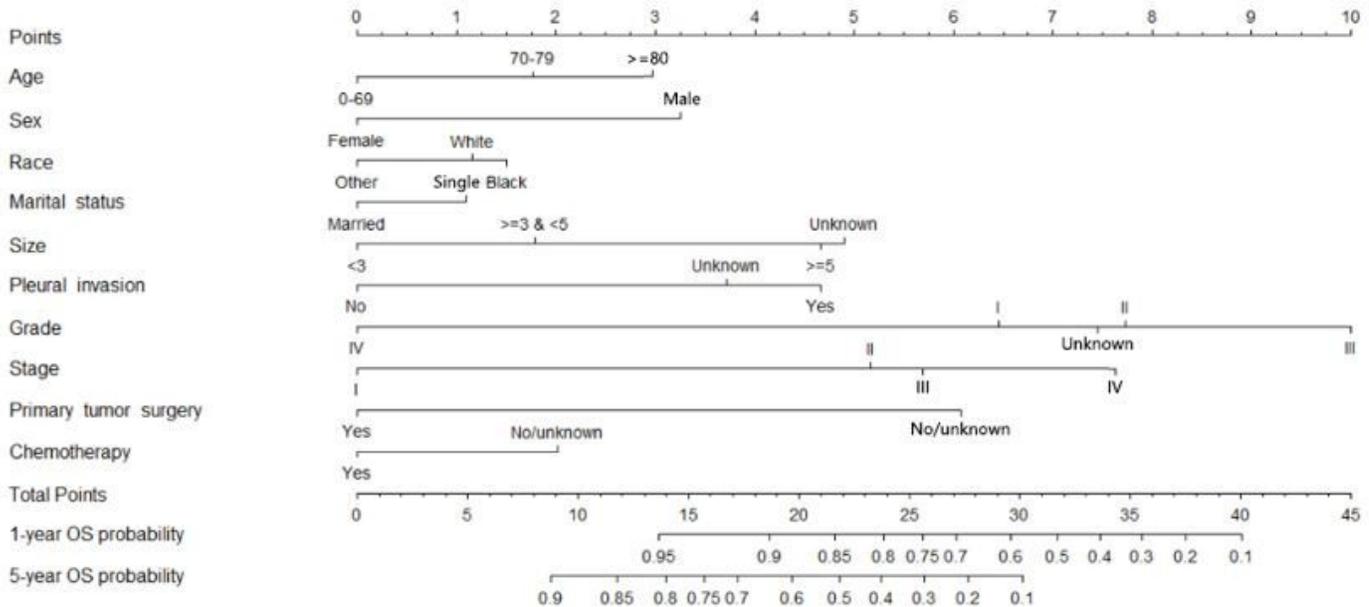


Figure 1

Nomograms predicting the survival probability of 1- and 5- year overall survival in patients with lepidic adenocarcinoma.

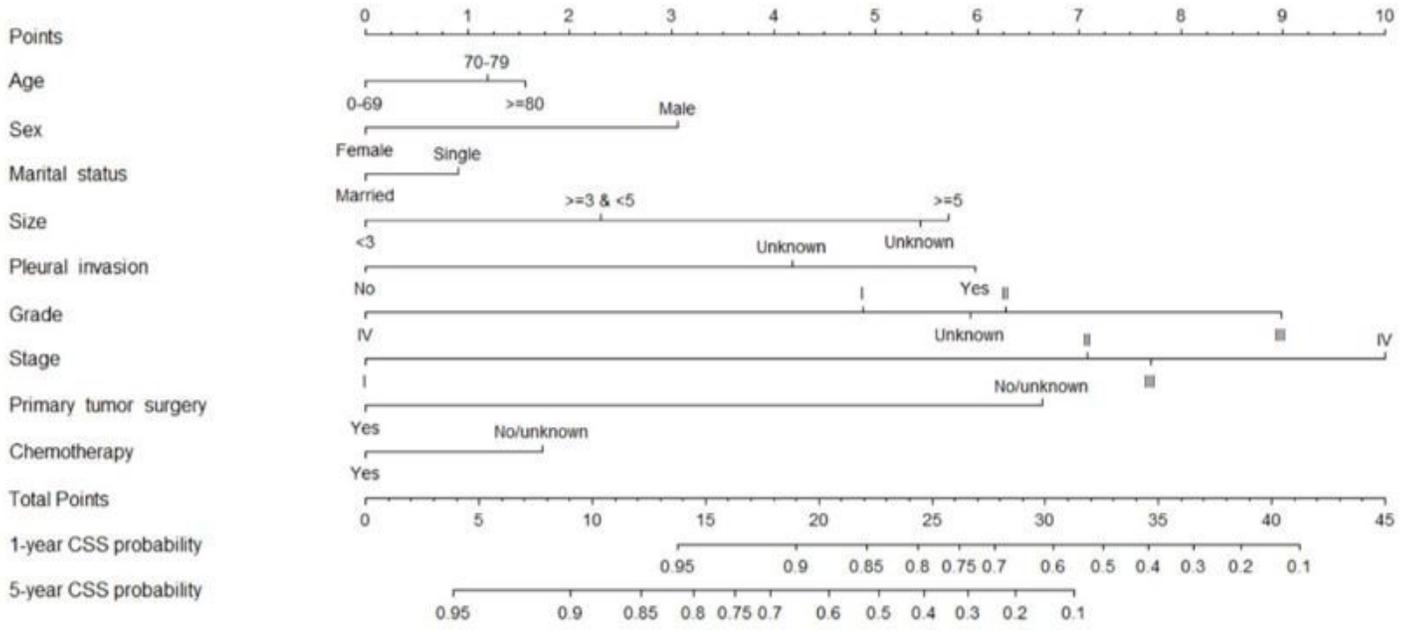


Figure 2

Nomograms predicting the survival probability of 1- and 5- year cancer-specific survival in patients with lepidic adenocarcinoma.

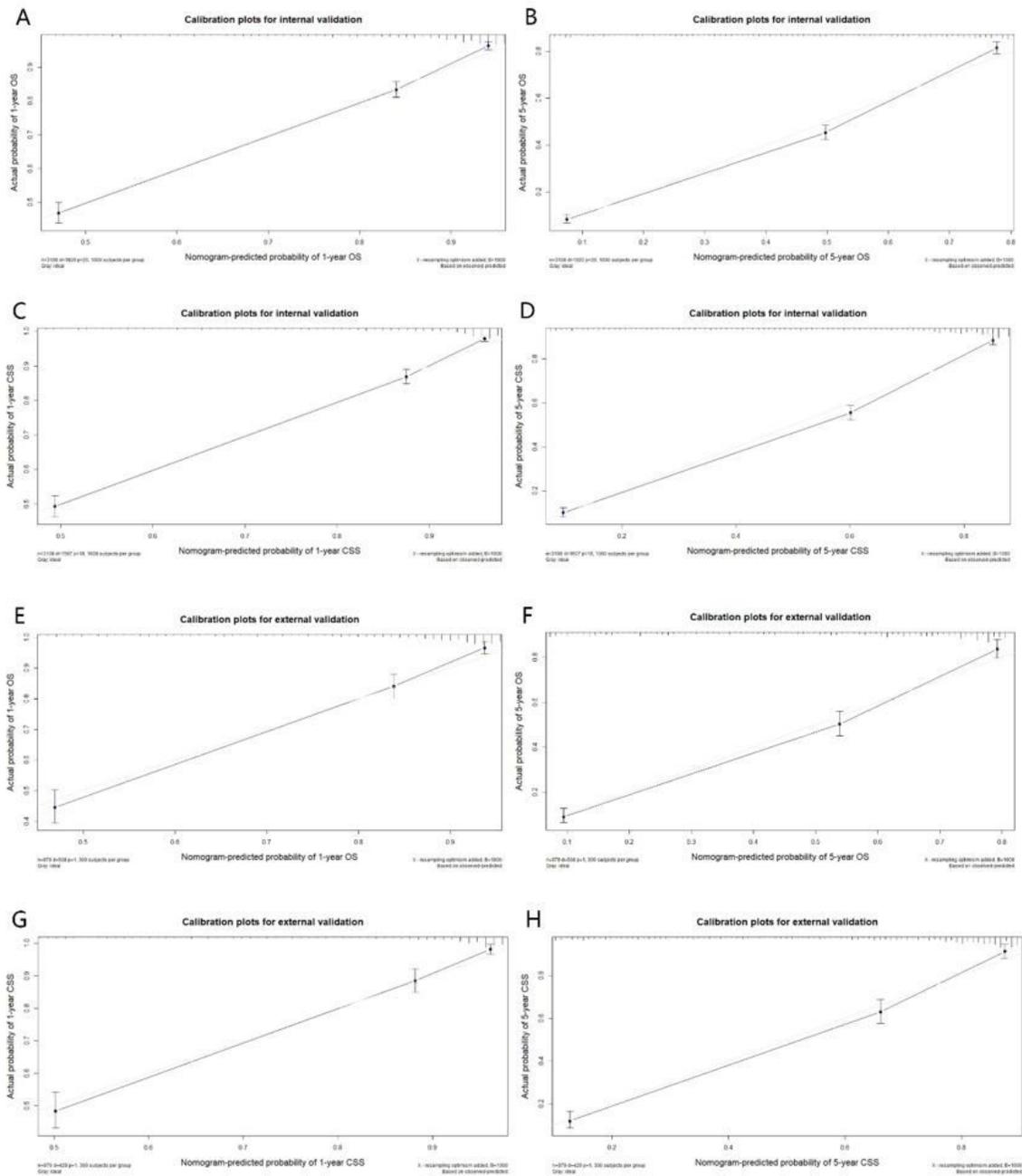


Figure 3

Calibration plots of the nomograms predicting OS and CSS in internal and external validation cohorts. (A, B) Calibration plots of 1- and 5-year OS in internal validation; (C, D) calibration plots of 1- and 5-year CSS in internal validation; (E, F) calibration plots of 1- and 5-year OS in external validation; (G, H) calibration plots of 1- and 5-year CSS in external validation.

Supplementary Files

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

- [Supplementarymaterials.docx](#)