Characteristics and nomogram for primary lung lepidic adenocarcinoma

Hui Tang
Yingyi Wang (waltwyy@163.com)
Chunmei Bai

Research Article

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

Posted Date: January 7th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-246772/v2

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.

Methods: Data from the Surveillance, Epidemiology, and End Results (SEER) database of 4087 LPA patients were retrospectively analyzed and compared with non-LPA pulmonary ADC to explore the clinicopathological and prognosis features of LPA. Univariate and multivariate Cox proportional hazard models were performed to identify independent survival predictors for further nomogram development. The nomograms were validated by using the concordance index, receiver operating characteristic curves, and calibration plots, as well as decision curve analysis, in both the training and validation cohorts.

Results: Compared with non-LPA pulmonary ADC patients, those with LPA exhibited unique clinicopathological features, including more elderly and female patients, smaller tumor size, less pleural invasion, and 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 compared with the actual observed results and demonstrated improved prognostic capacity compared with TNM stage.

Conclusions: Patients with LPA are more likely to be older and female. 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 one of the most commonly diagnosed cancers 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 the 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 relying only 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 covers approximately 28% of the population in the USA (7). Using the SEER database, we retrospectively analyzed the clinicopathological features and survival data of 4087 LPA patients to confirm their clinicopathological characteristics and prognostic factors. We then developed nomograms estimating the overall survival (OS) and cancer-specific survival (CSS) of patients with LPA. Furthermore, we performed nomogram validation in both the training and validation cohorts, as well as decision curve

analysis (DCA), to evaluate the accuracy of the nomograms. In addition, 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 SEER*stat software, version 8.3.8 (https://seer.cancer.gov/seerstat/). Lung adenocarcinoma was classified according to the 2015 WHO classification system. The 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 as follows: (1) primary lung cancer; (2) ICD-O-3 histology code 8250/3 (lepidic adenocarcinoma) or 8140/3 (adenocarcinoma-NOS); (3) positive histological confirmation; and (4) diagnosis between 2005 and 2014. The exclusion criteria were as follows: (1) patients who had multiple primary tumors in their lifetime; (2) unknown survival data or TNM stage; and (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 the LPA group and the ADC-NOS group. After propensity score matching (PSM), OS and CSS were compared between ADC-NOS and LPA patients. To create and validate the nomograms, patients with LPA diagnosed in 2009 and 2010 (n = 979) were assigned to the validation cohort, and those diagnosed between 2005 and 2014, except for 2009 and 2010 (n = 3108), were assigned to the training cohort.

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 races were recorded as “Other”, except for white and black races. “Married (including common law)” was recorded as “Married”, and other marital statuses were recorded as “Single”. The status of education and income were defined as “Low” or “High”, meaning that patients resided in counties with lower/higher education or income than the median level. Considering that the survival time in the SEER database was expressed in months, the survival time of 0 months was recorded as 0.5 months. OS was defined as the period from diagnosis to death caused by any cause or the last follow-up, while CSS was defined as the period from diagnosis to death caused by lung cancer.

2.3. Statistical analysis

For descriptive statistics, the absolute number and percentage of variables were described. Chi-square tests were used to compare the demographic and clinicopathological characteristics among different groups. PSM analysis was used to minimize the impact of confounding factors. The propensity score for each 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 levels, primary tumor location and size, separate tumor nodules, pleural invasion, histological grade, TNM stage, and treatment. Caliper matching within a caliper of 0.02 was performed among the two groups. After PSM analysis, 4060 pairs of patients were successfully matched among the patients included in our study. OS

and CSS were compared between matched patients with ADC-NOS and LPA by Kaplan–Meier curves and log-rank tests. Then, the LPA patient data 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), receiver operating characteristic (ROC) curves, 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 validated in both the training and validation cohorts.

The ages of patients were stratified by the X-tile program (Yale University, USA) (8). According to the desirable 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/). A two-tailed value of $P < 0.05$ was considered to be statistically significant.

Results

3.1. Patients and tumor characteristics

Among 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 accounted for 2.18% of all lung cancer patients. After applying the inclusion and exclusion criteria, the numbers 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 are shown in Table 1. Among the eligible patients, those with LPA were more common in older age, female, and yellow race. In addition, patients with LPA were inclined to have smaller tumor sizes, fewer separate tumor nodules, less pleural invasion, and lower histological grades and stages. After PSM analysis, Kaplan–Meier curves and log-rank tests were performed and showed that patients with LPA had better survival outcomes than those with ADC-NOS (Supplementary Figure S2A-B).

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 tests and further examined by multivariate analysis, which showed that yellow race, large tumor size, positive separate tumor nodules, and higher histological grade were independent risk factors for distant metastasis. Moreover, age, sex, race, tumor size, separate tumor nodules, 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. 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 the nomograms was used to estimate the survival probability of one individual patient.

3.4. Nomogram validation

Validation of the OS and CSS nomograms was performed in both the training and validation cohorts. 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) in the training cohort, respectively. 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) in the validation cohort, respectively. The sensitivity and specificity of predicting the prognosis of LPA were identified by ROC curves. As shown in Figure 3, the area under the curve (AUC) values of the nomogram predicting 1- and 5-year OS were 0.839 and 0.859, respectively, in the training cohort (Figure 3A); the AUC values of the nomogram predicting 1- and 5-year OS were 0.850 and 0.873, respectively, in the validation cohort (Figure 3B). While the AUC values of the nomogram predicting 1- and 5-year CSS were 0.858 and 0.889, respectively, in the training cohort (Figure 3C), the AUC values of the nomogram predicting 1- and 5-year CSS were 0.872 and 0.907, respectively, in the validation cohort (Figure 3D). Furthermore, calibration plots conducted using the training and validation cohorts both indicated that the OS and CSS nomograms demonstrated excellent agreement between the predicted and actual survival outcomes (Figure 4A-H). In addition, the DCA results demonstrated that the nomograms showed better prognostic capacity than TNM stage (Figure 5A-D).

Furthermore, LPA patients were divided into two groups (“low risk” or “high risk”) based on the median total scores calculated by the nomograms. As shown in Figure 6A, the Kaplan–Meier curves and log-rank tests suggested that the median OS of LPA patients in the high-risk group (17.0 months; 95% CI, 16.0-18.0 months) was significantly shorter than that in the low-risk group (not reached) ($P < 0.0001$). Likewise, as shown in Figure 6B, the median CSS of LPA patients in the high-risk group (19.0 months; 95% CI, 18.0-21.0 months) was significantly shorter than that in the low-risk group (not reached) ($P < 0.0001$). Moreover, the median OS of all LPA patients was 50 months (95% CI, 47.0-50.0 months), and the median CSS of all LPA patients was 75 months (95% CI, 69.0-88.0 months) (Figure 6A-B).

Discussion

Concise and accurate prognostic prediction models for patients with malignancy are essential for clinical decision-making and scientific research. Indeed, TNM stage is the most widely used survival predictor for cancer patients. However, 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 the 1- and 5-year OS and CSS of patients with LPA.

The survival outcomes of LPA patients with poor prognostic factors were undesirable, and 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 approximately 90% (10). With the evaluation of the nomograms 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 tumors and advanced TNM stages 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 that the incidence of LPA was much higher. Our results also indicated that LPA patients were more common in older age and females, which is consistent with previous studies (13, 14). In addition, some clinicopathological features of LPA patients indicated a good prognosis, including smaller tumor size, fewer separate tumor nodules, less pleural invasion, and 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 possessed some characteristics differing from other histologic subtypes of invasive pulmonary ADC, such as being more common in nonsmokers or light smokers, a preference for pulmonary peripheral location and being false-negative in positron-emission tomographic scans (13, 16). Clinically, asymptomatic at presentation or excessive airway secretion were more common in patients with LPA (17). In the genetic alteration profiles, EGFR mutations occurred in approximately 50% of patients with LPA, which was significantly higher than other subtypes (5), especially mutations in exon 21 (17, 18). However, KRAS mutations are much less common and account for approximately 10% of the LPA population (5). Compared 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. 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 (21, 22). In patients with advanced LPA, studies have suggested that taxane-based chemotherapy and pemetrexed might be effective and well tolerated (23, 24). With higher frequencies of EGFR mutations, EGFR-TKI therapy 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 was associated with a lower risk of recurrence, and invasive component size was a better predictor for survival than overall tumor diameter (16, 17, 28, 29). Furthermore, no recurrence was observed in any of the 18 LPA patients with a 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 factor of LPA patients following histological grade and stage. In 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 or 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. Notably, 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. Moreover, our nomograms were examined by C-index, ROC curves, calibration plots, and DCA curves, which demonstrated that the nomograms showed excellent agreement between the nomogram-predicted and actual survival outcomes of patients with LPA, as well as better prognostic capacity than TNM stage.

To date, this is the first time that the demographic and clinicopathological features, as well as the incidence of LPA, have been elucidated based on a large-scale population-based database. Meanwhile, this is the first nomogram predicting the survival outcomes of LPA patients, which could aid in personalized prognostic evaluation and clinical decision-making. However, there were still some limitations in our study, although the nomograms demonstrated good accuracy and applicability. First, nomograms were constructed based on retrospective data, and prospective external validation is needed. Second, some critical information, such as the diameter of the invasive area in LPA, tumor biomarkers, chemotherapy regimen, targeted therapy, molecular pathology, and genetic tests, was absent in the database. Moreover, the TNM staging information provided by the database is the result of the 6th edition staging system, instead of the latest edition staging system. Therefore, we could not analyze those variables or improve the 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 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 the OS and CSS of LPA patients individually. The nomograms showed good accuracy and applicability, which may aid in individualized prognostic prediction for LPA patients and clinical decision-making.

Declarations

Funding

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

Competing interests

The authors declare that they have no conflict of interest.

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 and/or analyzed 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, Y W, and C B conceived of and designed the study. H T and Y W performed literature searches. H T generated the figures and Tables. H T and Y W analyzed the data. H T wrote the manuscript, and Y W and C B critically reviewed the manuscript. C 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 are publicly available and deidentified.

Consent for publication

Not applicable.

Tables

Table 1. Characteristics of patients with LPA compared with lung ADC-NOS.

Characteristics	LPA n = 4087 Number (%)	ADC n = 84267 Number (%)	Pearson χ^2 P value
Age, years old			<0.001
0-69	1952 (47.8)	49110 (58.3)	
70-79	1400 (34.3)	23561 (28.0)	
>=80	735 (18.0)	11596 (13.8)	
Sex			<0.001
Male	1546 (37.8)	41672 (49.5)	
Female	2541 (62.2)	42595 (50.5)	
Race			<0.001
White	3281 (80.3)	66376 (78.8)	
Black	350 (8.6)	10172 (12.1)	
Other	456 (11.2)	7719 (9.2)	
Marital status			<0.001
Married	2397 (58.6)	46353 (55.0)	
Single	1690 (41.4)	37914 (45.0)	
Education			0.174
Low	2071 (50.7)	41784 (49.6)	
High	2016 (49.3)	42483 (50.4)	
Income			0.01
Low	2072 (50.7)	40983 (48.6)	
High	2015 (49.3)	43284 (51.4)	
Laterality			<0.001
Left	1574 (38.5)	32024 (38.0)	
Right	2389 (58.5)	48216 (57.2)	
Unknown	124 (3.0)	4027 (4.8)	
Lobe			<0.001
Upper	2145 (52.5)	45369 (53.8)	
Middle	196 (4.8)	3850 (4.6)	
Lower	1400 (34.3)	21229 (25.2)	
Unknown	346 (8.5)	13819 (16.4)	
Tumor Size			<0.001
<3 cm	2173 (53.2)	27619 (32.8)	
>=3 & <5 cm	894 (21.9)	21676 (25.7)	
>=5 cm	514 (12.6)	19719 (23.4)	
Unknown	506 (12.4)	15253 (18.1)	
Separate tumor nodules			<0.001
Yes	332 (8.1)	12568 (14.9)	
No	1121 (27.4)	33194 (39.4)	
Unknown	2634 (64.4)	38505 (45.7)	
Pleural invasion			<0.001
Yes	110 (2.7)	4190 (5.0)	
No	677 (16.6)	9439 (11.2)	
Unknown	3300 (80.7)	70638 (83.8)	
Grade			<0.001
I	1592 (39.0)	5433 (6.4)	
II	1069 (26.2)	18796 (22.3)	
III	244 (6.0)	25228 (29.9)	
IV	7 (0.2)	548 (0.7)	
Unknown	1175 (28.7)	34262 (40.7)	
Stage			<0.001
I	2207 (54.0)	17011 (20.2)	
II	167 (4.1)	3646 (4.3)	
III	585 (14.3)	18059 (21.4)	
IV	1128 (27.6)	45551 (54.1)	
Primary tumor surgery			<0.001
Yes	2597 (63.5)	22366 (26.5)	
No/unknown	1490 (36.5)	61901 (73.5)	
Metastatic tumor surgery			<0.001

Yes	74 (1.8)	4887 (5.8)	
No/unknown	4013 (98.2)	79380 (94.2)	
Radiotherapy			<0.001
Yes	612 (15.0)	34543 (41.0)	
No/unknown	3475 (85.0)	49724 (59.0)	
Chemotherapy			<0.001
Yes	1194 (29.2)	41016 (48.7)	
No/unknown	2893 (70.8)	43251 (51.3)	

Abbreviations: ADC, adenocarcinoma; LPA, lepidic adenocarcinoma; NOS, not otherwise specified.

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			
<3 cm		1.000				1.000		
>=3 & <5 cm		1.277	(1.131,1.443)	<0.001		1.361	(1.182,1.569)	<0.001
>=5 cm		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.

References

1. Fitzmaurice C, Abate D, Abbasi N, Abbastabar H, Abd-Allah F, Abdel-Rahman O, et al. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2017: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol* (2019) 5:1749-68. doi: 10.1001/jamaoncol.2019.2996

2. 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. doi: 10.1097/JTO.0b013e318206a221

3. Lambe G, Durand M, Buckley A, Nicholson S, McDermott R. Adenocarcinoma of the lung: from BAC to the future. *Insights Imaging* (2020) 11:69. doi: 10.1186/s13244-020-00875-6

4. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB, et al. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. *J Thorac Oncol* (2015) 10:1243-60. doi: 10.1097/jto.0000000000000630

5. Jiang L, Mino-Kenudson M, Roden AC, Rosell R, Molina M, Flores RM, et al. Association between the novel classification of lung adenocarcinoma subtypes and EGFR/KRAS mutation status: A systematic literature review and pooled-data analysis. *Eur J Surg Oncol* (2019) 45:870-6. doi: 10.1016/j.ejso.2019.02.006

6. Kim J, Jang SJ, Choi CM, Ro JY. Correlation of Histologic Subtypes and Molecular Alterations in Pulmonary Adenocarcinoma: Therapeutic and Prognostic Implications. *Adv Anat Pathol* (2016) 23:330-8. doi: 10.1097/pap.0000000000000121

7. Doll KM, Rademaker A, Sosa JA. Practical Guide to Surgical Data Sets: Surveillance, Epidemiology, and End Results (SEER) Database. *JAMA Surg* (2018) 153:588-9. doi: 10.1001/jamasurg.2018.0501

8. Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res* (2004) 10:7252-9. doi: 10.1158/1078-0432.Ccr-04-0713

9. Cadranel J, Gervais R, Merle P, Moro-Sibilot D, Westeel V, Bigay-Game L, et al. Erlotinib versus carboplatin and paclitaxel in advanced lepidic adenocarcinoma: IFCT-0504. *Eur Respir J* (2015) 46:1440-50. doi: 10.1183/13993003.02358-2014

10. Glanville AR, Wilson BE. Lung transplantation for non-small cell lung cancer and multifocal bronchioalveolar cell carcinoma. *Lancet Oncol* (2018) 19:e351-e8. doi: 10.1016/s1470-2045(18)30297-3

11. Zhang Y, Xie H, Zhang Z, Zhang P, Chen P, Wang X. The Characteristics and Nomogram for Primary Lung Papillary Adenocarcinoma. *Open Med* (2020) 15:92-102. doi: 10.1515/med-2020-0014

12. Ersek JL, Symanowski JT, Han Y, Howard A, Dumas K, Ahrens W, et al. Pulmonary Carcinosarcoma: A Surveillance, Epidemiology, and End Results (SEER) Analysis. *Clin Lung Cancer* (2020) 21:160-70. doi: 10.1016/j.cllc.2019.07.001

13. Arenberg D. Bronchioloalveolar carcinoma. *Semin Respir Crit Care Med* (2011) 32:52-61. doi: 10.1055/s-0031-1272869

14. Forest F, Patoir A, Dal-Col P, Da Cruz V, Camy F, Stachowicz ML, et al. Lepidic, Papillary Components and EGFR Mutations are Frequent in Patients With Lung Adenocarcinoma Who are Over 75 Years Old. *Appl Immunohistochem Mol Morphol* (2019) 27:667-71. doi: 10.1097/pai.0000000000000678

15. Borczuk AC. Prognostic considerations of the new World Health Organization classification of lung adenocarcinoma. *Eur Respir Rev* (2016) 25:364-71. doi: 10.1183/16000617.0089-2016

16. Moon Y, Lee KY, Sung SW, Park JK. Differing histopathology and prognosis in pulmonary adenocarcinoma at central and peripheral locations. *J Thorac Dis* (2016) 8:169-77. doi: 10.3978/j.issn.2072-1439.2016.01.15

17. Chakraborty RK, Sharma S. Bronchoalveolar Cancer. StatPearls Publishing. 2020.

18. Zhao FN, Zhao YQ, Han LZ, Xie YS, Liu Y, Ye ZX. Clinicoradiological features associated with epidermal growth factor receptor exon 19 and 21 mutation in lung adenocarcinoma. *Clin Radiol* (2019) 74:7-17. doi: 10.1016/j.crad.2018.10.001

19. Zheng D, Wang R, Zhang Y, Pan Y, Cheng X, Cheng C, et al. Prevalence and clinicopathological characteristics of ALK fusion subtypes in lung adenocarcinomas from Chinese populations. *J Cancer Res Clin Oncol* (2016) 142:833-43. doi: 10.1007/s00432-015-2081-4

20. Song Z, Yu X, Zhang Y. Clinicopathologic characteristics, genetic variability and therapeutic options of RET rearrangements patients in lung adenocarcinoma. *Lung Cancer* (2016) 101:16-21. doi: 10.1016/j.lungcan.2016.09.002

21. Sasada S, Miyata Y, Mimae T, Tsutani Y, Mimura T, Okada M. Application of Lepidic Component Predominance to Adjuvant Chemotherapy with Oral Fluoropyrimidines for Stage I Lung Adenocarcinoma. *Clin Lung Cancer* (2016) 17:433-40.e1. doi: 10.1016/j.cllc.2015.11.015

22. Sasada S, Miyata Y, Mimae T, Mimura T, Okada M. Impact of Lepidic Component Occupancy on Effects of Adjuvant Chemotherapy for Lung Adenocarcinoma. *Ann Thorac Surg* (2015) 100:2079-86. doi: 10.1016/j.athoracsur.2015.05.102

23. Duruisseaux M, Baudrin L, Quoix E, Wislez M, Moro-Sibilot D, Coëtmeur D, et al. Chemotherapy effectiveness after first-line gefitinib treatment for advanced lepidic predominant adenocarcinoma (formerly advanced bronchioloalveolar carcinoma): exploratory analysis of the IFCT-0401 trial. *J Thorac Oncol* (2012) 7:1423-31. doi: 10.1097/JTO.0b013e31825ab897

24. Lau DH, Moon J, Davies AM, Sanborn RE, Hirsch FR, Franklin WA, et al. Southwestern oncology group phase II trial (S0526) of pemetrexed in bronchioloalveolar carcinoma subtypes of advanced adenocarcinoma. *Clin Lung Cancer* (2013) 14:351-5. doi: 10.1016/j.cllc.2012.12.004

25. Takada K, Toyokawa G, Kinoshita F, Jogo T, Kohashi K, Wakasu S, et al. Expression of PD-L1, PD-L2, and IDO1 on tumor cells and density of CD8-positive tumor-infiltrating lymphocytes in early-stage lung adenocarcinoma according to histological subtype. *J Cancer Res Clin Oncol* (2020) 146:2639-50. doi: 10.1007/s00432-020-03250-6

26. Naso JR, Wang G, Pender A, Wong SK, Zhu J, Ho C, et al. Intratumoral heterogeneity in programmed death-ligand 1 immunoreactivity is associated with variation in non-small cell lung carcinoma histotype. *Histopathology* (2020) 76:394-403. doi: 10.1111/his.13983

27. Toyokawa G, Takada K, Okamoto T, Kawanami S, Kozuma Y, Matsubara T, et al. Relevance Between Programmed Death Ligand 1 and Radiologic Invasiveness in Pathologic Stage I Lung Adenocarcinoma. *Ann Thorac Surg* (2017) 103:1750-7. doi: 10.1016/j.athoracsur.2016.12.025

28. Xu Y, Zhu C, Qian W, Zheng M. Comprehensive study of mutational and clinicopathologic characteristics of adenocarcinoma with lepidic pattern in surgical resected lung adenocarcinoma. *J Cancer Res Clin Oncol* (2017) 143:181-6. doi: 10.1007/s00432-016-2255-8

29. Aherne EA, Plodkowski AJ, Montecalvo J, Hayan S, Zheng J, Capanu M, et al. What CT characteristics of lepidic predominant pattern lung adenocarcinomas correlate with invasiveness on pathology? *Lung Cancer* (2018) 118:83-9. doi: 10.1016/j.lungcan.2018.01.013

30. Suzuki S, Sakurai H, Masai K, Asakura K, Nakagawa K, Motoi N, et al. A Proposal for Definition of Minimally Invasive Adenocarcinoma of the Lung Regardless of Tumor Size. *Ann Thorac Surg* (2017) 104:1027-32. doi: 10.1016/j.athoracsur.2017.02.067

Figures

Figure 1

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

Figure 2

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

Figure 3

Receiver operating characteristic curves of the nomograms predicting OS and CSS in the training and validation cohorts. Receiver operating characteristic curves of 1- and 5-year OS in the training cohort (A) and the validation cohort (B); receiver operating characteristic curves of 1- and 5-year CSS in the training cohort (C) and the validation cohort (D).

Figure 4

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

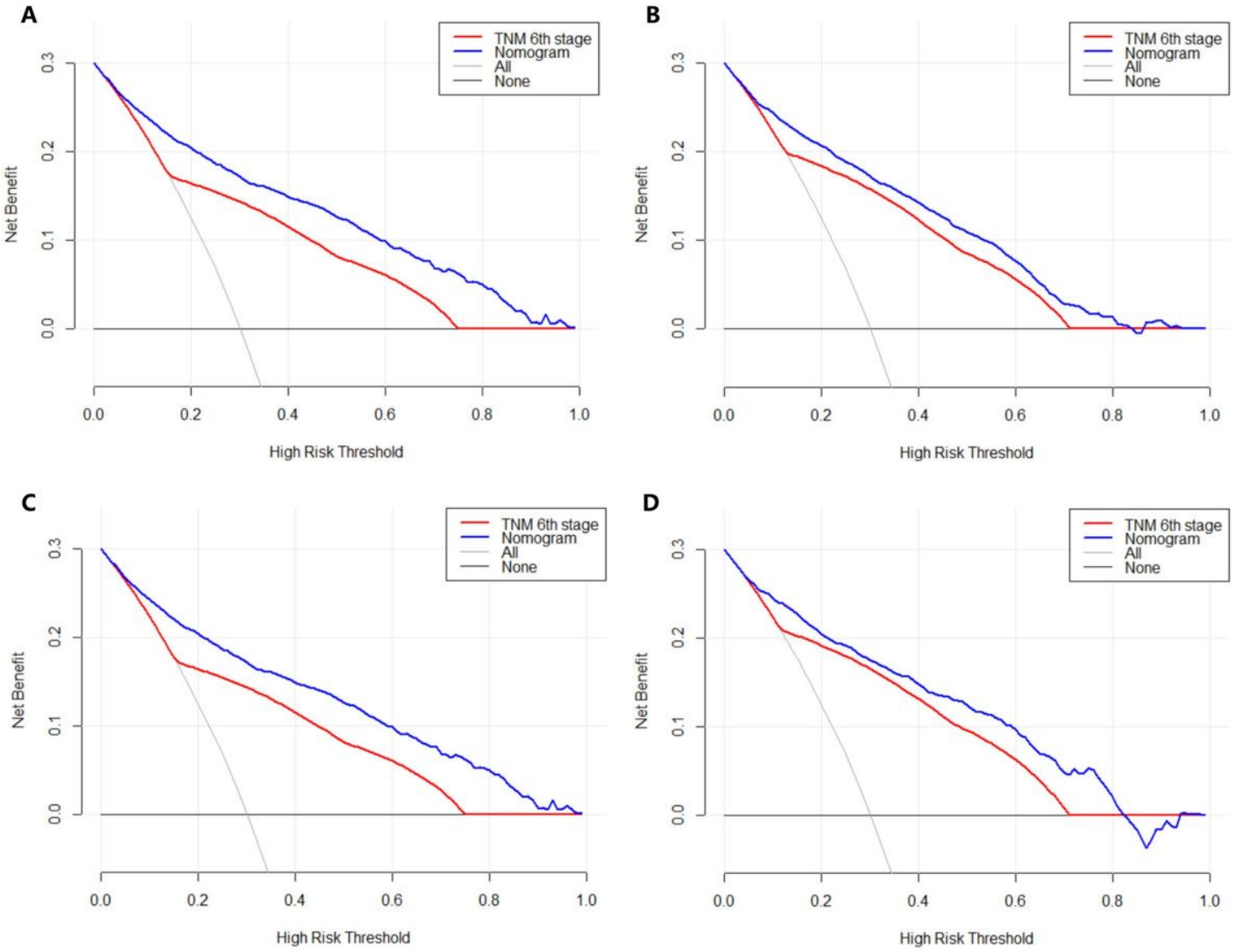

Figure 5

Decision curve analysis for the nomograms predicting OS and CSS. (A, B) Decision curve analysis of the nomogram for OS (A) and CSS (B) in the training cohort; (C, D) decision curve analysis of the nomogram for OS (C) and CSS (D) in the validation cohort.

Figure 6

Kaplan–Meier curves of overall survival (A) and cancer-specific survival (B) for all patients with lepidic adenocarcinoma divided into two risk stratifications based on the scores calculated by the nomograms.

Supplementary Files

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

- Supplementarymaterials.docx