Observational Study

Medicine

OPEN

Prognostic factors and survival prediction for patients with metastatic lung adenocarcinoma

A population-based study

Bo Wu, MD[a], Jianhui Chen, MD[a], Xiang Zhang, MD[a], Nan Feng, MD[a], Zhongtian Xiang, MD[a], Yiping Wei, PhD, MD[a], Junping Xie, MD[b], Wenxiong Zhang, MD[a,*]

Abstract

The prognosis of metastatic lung adenocarcinoma (MLUAD) varies greatly. At present, no studies have constructed a satisfactory prognostic model for MLUAD. We identified 44,878 patients with MLUAD. The patients were randomized into the training and validation cohorts. Cox regression models were performed to identify independent prognostic factors. Then, R software was employed to construct a new nomogram for predicting overall survival (OS) of patients with MLUAD. Accuracy was assessed by the concordance index (C-index), receiver operating characteristic curves and calibration plots. Finally, clinical practicability was examined via decision curve analysis. The OS time range for the included populations was 0 to 107 months, and the median OS was 7.00 months. Nineteen variables were significantly associated with the prognosis, and the top 5 prognostic factors were chemotherapy, grade, age, race and surgery. The nomogram has excellent predictive accuracy and clinical applicability compared to the TNM system (C-index: 0.723 vs 0.534). The C-index values were 0.723 (95% confidence interval: 0.719–0.726) and 0.723 (95% confidence interval: 0.718–0.729) in the training and validation cohorts, respectively. The area under the curve for 6-, 12-, and 18-month OS was 0.799, 0.764, and 0.750, respectively, in the training cohort and 0.799, 0.762, and 0.746, respectively, in the validation cohort. The calibration plots show good accuracy, and the decision curve analysis values indicate good clinical applicability and effectiveness. The nomogram model constructed with the above 19 prognostic factors is suitable for predicting the OS of MLUAD and has good predictive accuracy and clinical applicability.

Abbreviations: AJCC = American Joint Committee on Cancer, AUC = area under curve, CI = confidence interval, C-index = concordance index, DCA = decision curve analysis, HR = hazard ratio, LC = lung cancer, LN = lymph node, LUAD = lung adenocarcinoma, MLUAD = metastatic lung adenocarcinoma, NSCLC = non-small-cell lung cancer, OS = overall survival, SEER = the Surveillance Epidemiology and End Results database.

Keywords: lung adenocarcinoma, metastatic, nomogram, overall survival, SEER

1. Introduction

Lung cancer (LC) ranks as the second highest in incidence and the first in death among all cancers; there were approximately 2,206,771 new cases and 1,796,144 deaths in 2021 worldwide.[1] Non-small cell LC (NSCLC) is one of the major pathology types in LC, in which lung adenocarcinoma (LUAD) accounts for 48.2%.[2] The 5-year overall survival (OS) rate for LAUD patients is 19%. In a large cohort of patients diagnosed with distant metastasis of LAUD, the 5-year OS rate was only 6%.[1,3] Although targeted therapy, immunotherapy, chemotherapy and radiotherapy are used to treat metastatic lung adenocarcinoma (MLUAD), the outcome is still not ideal and the tumor response is quite variable.[4] Moreover, the liver, bone, lung and brain are the most frequent sites of metastasis in MLUAD patients. Interestingly, LUAD patients with different metastatic sites have different prognoses, with median OS ranging from 3 to 8 months, which suggests that MLUAD is highly heterogeneous.[5,6] Therefore, the prognosis of patients with MLUAD varies greatly.

BW, JC, JX, and WZ contributed equally to this work.

This study was supported by National Natural Science Foundation of China (NSFC, grant numbers: 81160294 and 81960425) and the Key Research and Development Program of Jiangxi Province (grant number: 20181BBH80005). The funding had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

The authors have no conflicts of interest to disclose.

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

This article does not contain any studies with human participants or animals performed by any of the authors. The data of this paper is extracted from SEER database, where it is publicly available and unrestricted re-use is permitted via an open license.

Supplemental Digital Content is available for this article.

[a] Department of Thoracic Surgery, The Second Affiliated Hospital of Nanchang University, Nanchang, China, [b] Department of Respiratory and Critical Care Medicine, The Second Affiliated Hospital of Nanchang University, Nanchang, China.

** Correspondence: Wenxiong Zhang, Department of Thoracic Surgery, The Second Affiliated Hospital of Nanchang University, 1 Minde Road, Nanchang 330006, China (e-mail: zwx123dr@126.com).*

How to cite this article: Wu B, Chen J, Zhang X, Feng N, Xiang Z, Wei Y, Xie J, Zhang W. Prognostic factors and survival prediction for patients with metastatic lung adenocarcinoma: A population-based study. Medicine 2022;101:49(e32217).

Received: 11 October 2022 / Received in final form: 16 November 2022 / Accepted: 16 November 2022

http://dx.doi.org/10.1097/MD.0000000000032217

At present, clinicians assess patient prognosis and develop optimal treatment plans based on the eighth edition TNM staging system of the American Joint Committee on Cancer (AJCC).[7] However, the TNM system only evaluates tumor-related pathological features, including the size of the primary tumor, lymph node (LN) status and metastasis status. It lacks important clinicopathological features of the patients, such as age, sex, race, year of diagnosis, living status and treatment modality. Therefore, the TNM system cannot be used as a gold standard for predicting outcomes and formulating treatment plans. We need to construct a more comprehensive and accurate prediction model.

A nomogram is a prediction model based on multivariate regression analysis.[8] The nomogram model's theory is systematic, and the accuracy of the prediction is more satisfactory than the conventional TNM system. Therefore, nomogram models have been widely used in cancer prognostication.[9] Various satisfactory nomograms have been constructed for predicting outcomes in breast cancer, colon cancer, and LC,[10–12] and there are some studies on LUAD.[2] However, no studies have constructed a satisfactory prognostic nomogram for MLUAD.

Therefore, in this study, we constructed a new nomogram by assessing related prognostic factors to predict the OS of MLUAD patients based on the Surveillance, Epidemiology and End Results (SEER) database. Our findings can help clinicians make optimal treatment decisions and assess their prognosis.

2. Methods

2.1. Data source and patient selection

Study data were accessed through a large database (SEER database) based on cancer statistics in the United States of America. We downloaded the relevant information of each patient by SEER*stat version 8.4.0 from 18 registries. Because SEER database patient data is publicly available worldwide, this study is

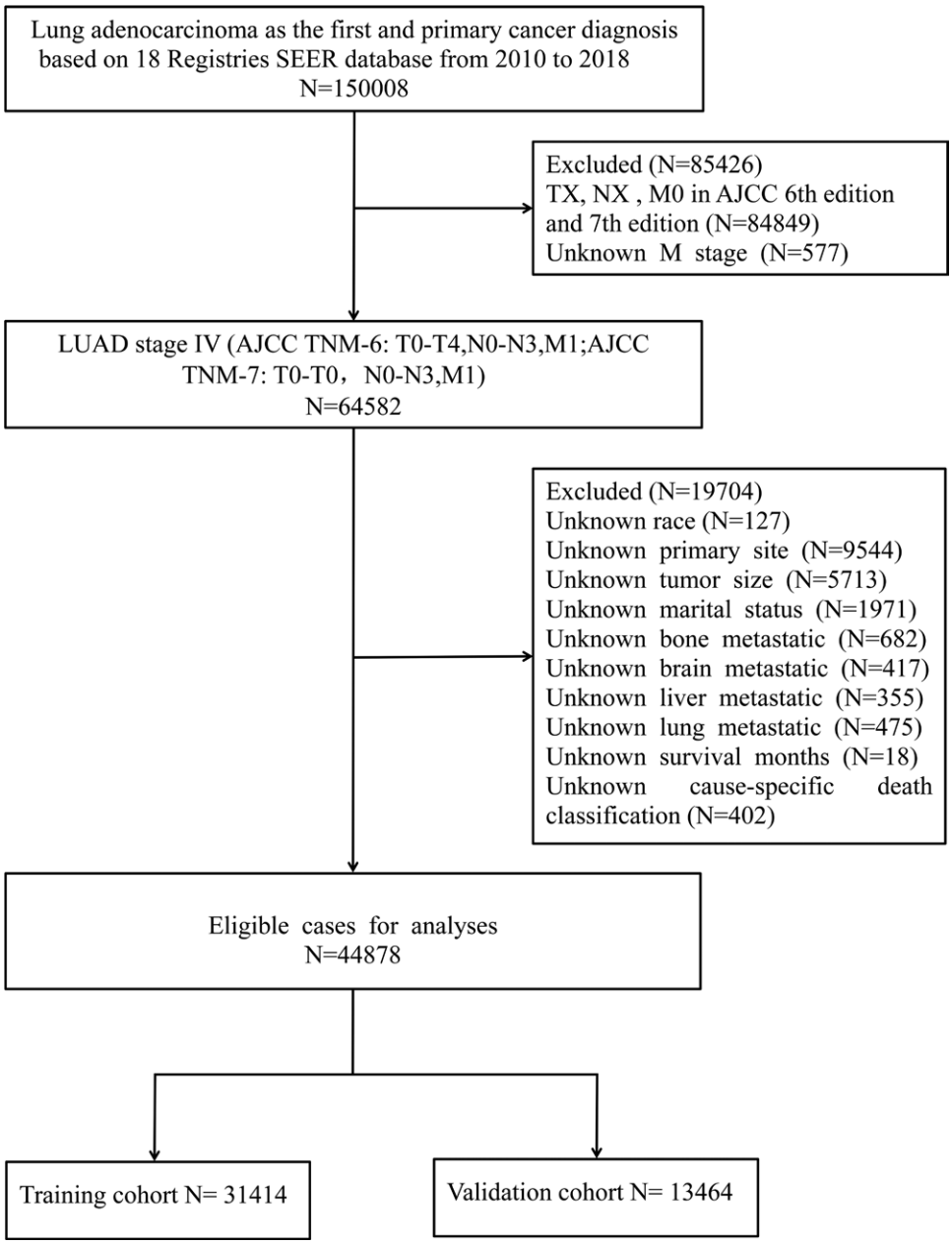

Figure 1. The flow chart of MLUAD patients. AJCC = American Joint Committee on Cancer, LUAD = lung adenocarcinoma, MLUAD = metastatic lung adenocarcinoma, SEER = the Surveillance Epidemiology and End Results database.

Table 1

Clinical and pathological characteristics of patients with metastatic lung adenocarcinoma.

	Total cohort	Training cohort	Validation cohort	
Characteristic	**44,878 (100%)**	**31,414 (70%)**	**13,464 (30%)**	***P* value**
Age				
<50	2571 (5.7%)	1776 (5.7%)	795 (5.9%)	.667
50–59	9119 (20.3%)	6380 (20.3%)	2739 (20.3%)	
60–69	14,554 (32.5%)	10,130 (32.3%)	4424 (32.9%)	
70–79	12,270 (27.3%)	8685 (27.6%)	3585 (26.6%)	
≥80	6364 (14.2%)	4443 (14.1%)	1921 (14.3%)	
Year of diagnosis				
2010–2014	23,419 (52.2%)	16,438 (52.3%)	6981 (51.8%)	.650
2015–2018	21,459 (47.8%)	14,976 (47.7%)	6483 (48.2%)	
Sex				
Female	22,464 (50.1%)	15,791 (50.3%)	6673 (49.6%)	.391
Male	22,414 (49.9%)	15,623 (49.7%)	6791 (50.4%)	
Race				
White	34,036 (75.8%)	23,886 (76.0%)	10,150 (75.4%)	.619
Black	5786 (12.9%)	3985 (12.7%)	1801 (13.4%)	
Asian	4852 (10.8%)	3396 (10.8%)	1456 (10.8%)	
American Indian	204 (0.5%)	147 (0.5%)	57 (0.4%)	
Living status				
With others*	24,464 (54.5%)	17,062 (54.3%)	7402 (55.0%)	.434
Alone†	20,414 (45.5%)	14,352 (45.7%)	6062 (45.0%)	
Primary site				
Upper lobe	27,526 (61%)	19,270 (61%)	8256 (61%)	.426
Middle lobe	2259 (5.0%)	1587 (5.1%)	672 (5.0%)	
Lower lobe	13,297 (30%)	9345 (30%)	3952 (29.7%)	
Main bronchus	1796 (4.0%)	1212 (3.9%)	584 (4.3%)	
Laterality				
Right	26,367 (58.8%)	18,389 (58.5%)	7978 (59.3%)	.608
Left	18,324 (40.8%)	12,896 (41.1%)	5428 (40.3%)	
Right+Left	115 (0.3%)	84 (0.3%)	31 (0.2%)	
Unknown	72 (0.2%)	45 (0.1%)	27 (0.2%)	
Tumor size (cm)				
≤3	14,718 (32.8%)	10,283 (32.7%)	4435 (32.9%)	.990
3–7	23,419 (52.2%)	16,419 (52.3%)	7000 (52.0%)	
≥7	6741 (15.0%)	4712 (15.0%)	2029 (15.1%)	
Grade				
I	1199 (2.7%)	856 (2.7%)	343 (2.5%)	.640
II	4663 (10.4%)	3290 (10.5%)	1373 (10.2%)	
III	9491 (21.1%)	6591 (21.0%)	2900 (21.5%)	
IV	349 (0.8%)	230 (0.7%)	119 (0.9%)	
Unknown	29,176 (65.0%)	20,447 (65.1%)	8729 (64.9%)	
T stage				
T0	37 (0.1%)	23 (0.1%)	14 (0.1%)	.958
T1	7296 (16.3%)	5111 (16.3%)	2185 (16.2%)	
T2	12,750 (28.4%)	8929 (28.4%)	3821 (28.4%)	
T3	10,431 (23.2%)	7256 (23.1%)	3175 (23.6%)	
T4	14,364 (32.0%)	10,095 (32.1%)	4269 (31.7%)	
N stage				
N0	11,175 (24.9%)	7807 (24.9%)	3368 (25.0%)	>.999
N1	3637 (8.1%)	2537 (8.1%)	1100 (8.2%)	
N2	19,675 (43.8%)	13,788 (43.8%)	5887 (43.7%)	
N3	10,391 (23.2%)	7282 (23.2%)	3109 (23.1%)	
Regional LN surgery				
≤3	608 (1.4%)	442 (1.4%)	166 (1.2%)	.409
>3	920 (2.1%)	625 (2.0%)	295 (2.2%)	
Unknown	43,350 (96.5%)	30,347 (96.6%)	13,003 (96.6%)	
Regional LN examined				
Yes	7986 (17.8%)	5649 (18.0%)	2337 (17.4%)	.284
No	36,892 (82.2%)	25,765 (82.0%)	11,127 (82.6%)	
Regional LN status				
Positive	6821 (15.2%)	4842 (15.4%)	1979 (14.7%)	.412
Negative	1156 (2.6%)	800 (2.5%)	356 (2.6%)	
Unknown	36,901 (82.2%)	25,772 (82.1%)	11,129 (82.7%)	
Bone metastases				
No	25,756 (57.4%)	18,004 (57.3%)	7752 (57.6%)	.875
Yes	19,122 (42.6%)	13,410 (42.7%)	5712 (42.4%)	
Brain metastases				
No	30,289 (67.5%)	21,178 (67.4%)	9111 (67.7%)	.871
Yes	14,589 (32.5%)	10,236 (32.6%)	4353 (32.3%)	

(*Continued*)

Table 1

(*Continued*)

Characteristic	Total cohort 44,878 (100%)	Training cohort 31,414 (70%)	Validation cohort 13,464 (30%)	*P* value
Liver metastases				
No	37,106 (82.7%)	26,020 (82.8%)	11,086 (82.3%)	.452
Yes	7772 (17.3%)	5394 (17.2%)	2378 (17.7%)	
Lung metastases				
No	31,089 (69.3%)	21,790 (69.4%)	9299 (69.1%)	.821
Yes	13,789 (30.7%)	9624 (30.6%)	4165 (30.9%)	
Distant LN metastases				
Yes	2696 (6.0%)	1864 (5.9%)	832 (6.2%)	.900
No	13,193 (29.4%)	9232 (29.4%)	3961 (29.4%)	
Unknown	28,989 (64.6%)	20,318 (64.7%)	8671 (64.4%)	
Surgery				
No surgery	43,127 (96.1%)	30,196 (96.1%)	12,931 (96.0%)	.923
Local‡	60 (0.1%)	40 (0.1%)	20 (0.1%)	
Wedge resection	556 (1.2%)	394 (1.3%)	162 (1.2%)	
Segmentectomy	66 (0.1%)	50 (0.2%)	16 (0.1%)	
Lobectomy	780 (1.7%)	545 (1.7%)	235 (1.7%)	
Surgery NOS	289 (0.6%)	189 (0.6%)	100 (0.7%)	
Radiation				
Yes	21,442 (47.8%)	15,022 (47.8%)	6420 (47.7%)	.965
No	23,436 (52.2%)	16,392 (52.2%)	7044 (52.3%)	
Chemotherapy				
Yes	26,528 (59.1%)	18,572 (59.1%)	7956 (59.1%)	.998
No	18,350 (40.9%)	12,842 (40.9%)	5508 (40.9%)	

LN = lymph node, NOS = not otherwise specified.

*Including marital status: married or with partner.

†Including marital status: single, divorced/separated or widowed.

‡Local tumor destruction (includes laser ablation, cryosurgery, electrocautery and fulguration).

exempt from ethics committee approval by default. The flow chart of the MLUAD patient inclusion is shown in Figure 1.

Inclusion criteria for this study: pathological type is LUAD, diagnosed 2010 through 2018, stage IV (AJCC TNM-6 M1; AJCC TNM-7: M1), LUAD is the primary tumor type leading to distant metastasis. The exclusion criteria were as follows: unknown pathological type, AJCC TNM-6: TX, NX; AJCC TNM-7: TX, NX), incomplete records of essential clinical and pathological information, and unknown survival times and cause-specific death classification.

2.2. Study variables

We screened a total of 22 variables that may be related to the prognosis and changed the continuous variables into categorical variables. The specific stratification of the variables was as follows: age (<50, 50–59, 60–69, 70–79, and ≥80), sex (female and male), year of diagnosis (2010–2014 and 2015–2018), living status (with others and alone), race (white, black, Asian and American Indian), primary site (upper, middle, lower lobes, and main bronchus), grade (I, II, III, IV, and unknown), laterality (right side, left side, right+left and unknown), tumor size (≤3 cm, 3–7 cm, and ≥7 cm), T stage (T0, T1, T2, T3, and T4), regional LN surgery (≤3, >3, and unknown), regional LN examined (yes and no), regional LN status (positive, negative, and unknown), N stage (N0, N1, N2, and N3), bone metastases (yes and no), brain metastases (yes and no), liver metastases (yes and no), lung metastases (yes and no), distant LN metastases (yes, no and unknown), surgery (no surgery, local tumor destruction [includes laser ablation, cryosurgery, electrocautery and fulguration], wedge resection, segmentectomy, lobectomy and surgery not otherwise specified), radiotherapy (no and yes) and chemotherapy (no and yes). In addition to the above variables, we also needed information about the primary outcome, OS, in terms of survival months and survival status.

2.3. Statistical analysis

The included patients were randomized to the training cohort and validation cohorts at a ratio of 7:3. To determine whether there were significant differences in different variables between the 2 cohorts of data, Pearson's chi-square test was used. The independent prognostic factors were selected for the training cohorts via univariate and multivariate Cox regression models. Least absolute shrinkage and selection operator regression analysis was also used to select significant prognostic factors. Kaplan–Meier analysis was used to build survival curves for the above factors. Then, a new nomogram model was built utilizing the independent prognostic factors. The training cohort and validation cohorts were used separately for internal and external validation of the nomogram model. The concordance index (C-index) and receiver operating characteristic curve were applied to calculate the discrimination ability of the nomogram model. The area under the curve (AUC) was between 0.5–1, with = 0.5 as completely random, 0.50–0.70 low accuracy, 0.70–0.90 medium accuracy, and 0.90–1.00 high accuracy.[13] Calibration plots were applied to measure the agreement of the predicted value from the nomogram with the actual value. Decision curve analysis (DCA) was used to prove this nomogram model's practical clinical value.[14] *P* values <.05 were accepted as statistically significant using the R (version 4.1.3, Vienna, Austria) software package to process all statistical analyses.

3. Results

3.1. Patient characteristics

We included 44,878 patients with MLUAD in this study, and they were randomized to the training cohort and validation cohort at a ratio of 7:3. The training set included 31,414 patients, and the validation set included 13,464 patients. The clinical and pathological characteristics of the MLUAD patients

Table 2

Univariate and multivariate cox regression analysis based on all variables for overall survival (training cohort).

	Univariate analysis		Multivariate analysis	
Characteristic	**HR (95% CI)**	***P* value**	**HR (95% CI)**	***P* value**
Age				
<50	Reference		Reference	
50–59	1.26 (1.19, 1.34)	<.001	1.16 (1.10, 1.24)	<.001
60–69	1.41 (1.33, 1.49)	<.001	1.29 (1.21, 1.36)	<.001
70–79	1.65 (1.55, 1.75)	<.001	1.48 (1.40, 1.57)	<.001
≥80	2.04 (1.92, 2.18)	<.001	1.58 (1.48, 1.68)	<.001
Year of diagnosis				
2010–2014	Reference		Reference	
2015–2018	0.87 (0.84, 0.89)	<.001	0.92 (0.89, 0.96)	<.001
Sex				
Female	Reference		Reference	
Male	1.28 (1.25, 1.31)	<.001	1.25 (1.22, 1.29)	<.001
Race				
White	Reference		Reference	
Black	1.04 (1.00, 1.08)	.030	1.02 (0.99, 1.06)	.204
Asian	0.67 (0.64, 0.70)	.000	0.68 (0.66, 0.71)	<.001
American Indian	0.95 (0.79, 1.14)	.575	0.91 (0.76, 1.09)	.292
Living status				
With others*	Reference		Reference	
Alone†	1.21 (1.18, 1.24)	<.001	1.14 (1.11, 1.17)	<.001
Primary site				
Upper lobe	Reference		Reference	
Middle lobe	0.91 (0.86, 0.97)	.002	0.93 (0.88, 0.98)	.013
Lower lobe	0.96 (0.93, 0.98)	.002	0.97 (0.95, 1.00)	.047
Main bronchus	1.23 (1.15, 1.31)	.000	1.12 (1.05, 1.19)	<.001
Laterality				
Right	Reference		Reference	
Left	0.96 (0.94, 0.99)	.004	0.96 (0.93, 0.98)	.001
Right+Left	0.79 (0.62, 1.02)	.067	0.87 (0.68, 1.12)	.29
Unknown	1.25 (0.90, 1.74)	.189	1.30 (0.93, 1.81)	.128
Tumor size (cm)				
≤3	Reference		Reference	
3–7	1.17 (1.13, 1.2)	<.001	1.14 (1.11, 1.17)	<.001
≥7	1.45 (1.4, 1.51)	<.001	1.41 (1.35, 1.46)	<.001
Grade				
I	Reference		Reference	
II	1.26 (1.16, 1.37)	<.001	1.28 (1.17, 1.39)	<.001
III	1.74 (1.60, 1.88)	<.001	1.67 (1.54, 1.81)	<.001
IV	1.89 (1.62, 2.20)	<.001	1.81 (1.55, 2.11)	<.001
Unknown	1.63 (1.51, 1.76)	<.001	1.56 (1.44, 1.69)	<.001
T stage				
T0	Reference			
T1	0.94 (0.60, 1.45)	.765		
T2	1.07 (0.69, 1.65)	.776		
T3	1.11 (0.71, 1.72)	.651		
T4	1.12 (0.72, 1.74)	.608		
N stage				
N0	Reference		Reference	
N1	1.08 (1.03, 1.14)	.002	1.13 (1.07, 1.18)	<.001
N2	1.28 (1.24, 1.32)	<.001	1.33 (1.29, 1.37)	<.001
N3	1.25 (1.20, 1.29)	<.001	1.39 (1.34, 1.45)	<.001
Regional LN surgery				
≤3	Reference		Reference	
>3	0.59 (0.51, 0.68)	<.001	0.84 (0.72, 0.99)	.001
Unknown	1.64 (1.48, 1.83)	<.001	1.08 (0.96, 1.21)	.193
Regional LN examined				
Yes	Reference		Reference	
No	1.37 (1.32, 1.41)	<.001	0.72 (0.32, 1.61)	.426
Regional LN status				
Positive	Reference		Reference	
Negative	0.55 (0.50, 0.61)	<.001	0.83 (0.75, 0.92)	<.001
Unknown	1.25 (1.20, 1.29)	<.001	1.57 (0.70, 3.51)	.269
Bone metastases				
No	Reference		Reference	
Yes	1.31 (1.27, 1.34)	<.001	1.35 (1.32, 1.39)	<.001
Brain metastases				
No	Reference		Reference	

(*Continued*)

Table 2

(*Continued*)

Characteristic	Univariate analysis		Multivariate analysis	
	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Yes	1.11 (1.08, 1.14)	<.001	1.30 (1.26, 1.34)	<.001
Liver metastases				
No	Reference		Reference	
Yes	1.59 (1.54, 1.65)	<.001	1.48 (1.44, 1.53)	<.001
Lung metastases				
No	Reference			
Yes	1.00 (0.97, 1.03)	.968		
Distant LN metastases				
Yes	Reference		Reference	
No	0.89 (0.84, 0.95)	.000	0.94 (0.88, 1.00)	.041
Unknown	1.07 (1.01, 1.13)	.025	1.16 (1.08, 1.24)	<.001
Surgery				
No surgery	Reference		Reference	
Local‡	0.83 (0.59, 1.19)	.311	0.91 (0.64, 1.29)	.599
Wedge resection	0.50 (0.45, 0.57)	<.001	0.68 (0.60, 0.77)	<.001
Segmentectomy	0.43 (0.30, 0.60)	<.001	0.48 (0.34, 0.68)	<.001
Lobectomy	0.31 (0.28, 0.35)	<.001	0.51 (0.44, 0.60)	<.001
Surgery NOS	0.74 (0.63, 0.87)	<.001	0.87 (0.74, 1.02)	.087
Radiation				
Yes	Reference		Reference	
No	1.07 (1.04, 1.09)	<.001	1.14 (1.10, 1.17)	<.001
Chemotherapy				
Yes	Reference		Reference	
No	2.47 (2.41, 2.53)	<.001	2.69 (2.62, 2.76)	<.001

CI = confidence interval, HR = hazard ratio, LN = lymph node, NOS = not otherwise specified.

*Including marital status: married or with partner.

†Including marital status: single, divorced/separated or widowed.

‡Local tumor destruction (includes laser ablation, cryosurgery, electrocautery and fulguration).

are shown in Table 1. In the total cohort, the OS time range for the included population was 0 to 107 months, median OS of 7 months. MLUAD patients were concentrated within 60 to 69 years (32.5%), white (75.8%), upper lobe primary site (61.0%), unilateral laterality (99.6%), tumor size 3 to 7 cm (52.2%), and N2 (43.8%). The most common metastatic sites in MLUAD were bone (42.6%), brain (32.5%), lung (30.7%), and liver (17.3%). Moreover, 3.9% of patients received surgery, 47.8% of patients received radiotherapy, and 59.1% of patients were treated with chemotherapy. The basic characteristics of the patients were basically the same in the training and validation cohorts (all $P > .284$).

3.2. Prognostic factors of OS for MLUAD

To screen the independent prognostic factors of MLUAD patients, first, 22 variables were analyzed via univariate Cox regression. Lung metastases and T stage were excluded ($P > .05$), and the other variables had significant statistical significance (all $P < .05$) (Table 2, see Figure S1A, Supplemental Digital Content, http://links.lww.com/MD/I77, which illustrates the univariate (A) and multivariate (B) cox regression analysis for MLUAD patients in training cohort). Second, the significant variables were further subjected to multivariate Cox regression. The final multivariate Cox regression statistical results showed that age, primary site, sex, race, living status, year of diagnosis, laterality, tumor size, grade, N stage, regional LN status, regional LN surgery, distant LN metastases, bone metastases, brain metastases, liver metastases, surgery, chemotherapy and radiation were independent prognostic factors for OS (Table 2, see Figure S1B, Supplemental Digital Content, http://links.lww.com/MD/I77, which illustrates the univariate (A) and multivariate (B) cox regression analysis for MLUAD patients in training cohort). The same 19 independent prognostic factors were identified via least absolute shrinkage and selection operator regression (Fig. 2). These variables were further used for Kaplan–Meier curve analysis. Kaplan–Meier curves confirmed these independent prognostic factors (log-rank tests $P < .05$), in which no chemotherapy, grade IV, age ≥ 80 years, with no surgery (Fig. 3), black race, liver metastases, unknown laterality, unknown regional LN surgery, unknown regional node status, tumor size ≥ 7 cm, N2, bone metastases, brain metastases, male sex, unknown distant LN metastases, main bronchus site, living alone, no radiation and diagnosed in 2010 to 2014 were predictive risk factors for MLUAD survival (see Figure S2, Supplemental Digital Content, http://links.lww.com/MD/I78, which illustrates the Kaplan–Meier curves of OS for MLUAD patients in race (A), liver metastases (B), laterality (C), regional LN surgery (D), regional LN status (E), tumor size (F), N stage (G), bone metastases (H), brain metastases (I), sex (J), distant LN metastases (K), primary site (L), living status (M), radiation (N) and year of diagnosis (O)).

3.3. Construction of the MLUAD nomogram

The above 19 independent predictors were used to construct the prognostic nomogram (Fig. 4) for MLUAD in the training cohort. Different predictors have their corresponding points, and the total points are equivalent to the total of all scores of predictors. According to the probability of different OS corresponding to the total points, we estimated the probability of 6-, 12- and 18-month OS for each patient. The top 5 contributing factors were chemotherapy, grade, age, race and surgery, and the others were all important predictors. Patients who refused chemotherapy or surgery or age ≥ 80 had the worst OS.

3.4. Validation of the nomogram

The training cohort and validation cohorts were used separately in internal and external validation of the nomogram

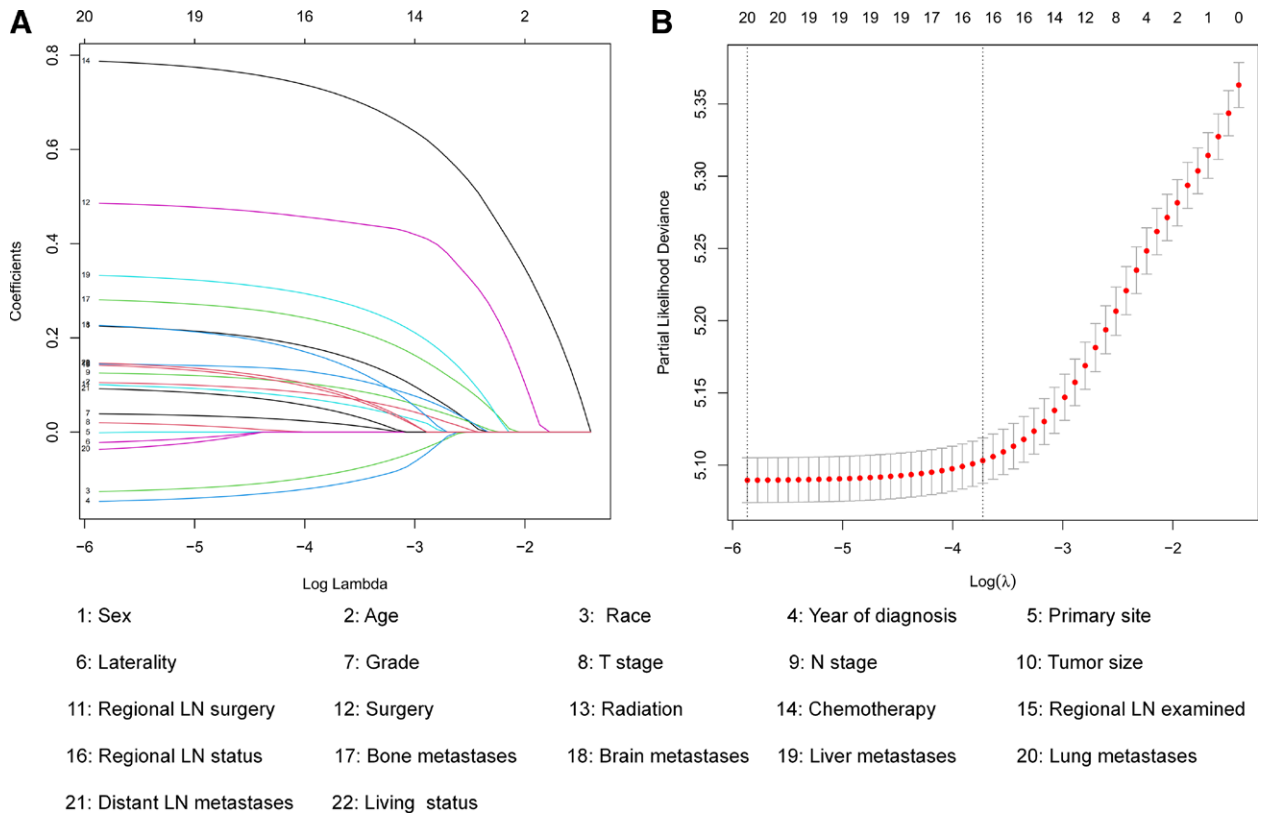

Figure 2. Feature selection using the least absolute shrinkage and selection operator (LASSO) regression in the training cohort. (A) LASSO coefficient profiles of 22 clinicopathologic characteristics. (B) Selecting the tuning parameters (lambda) in lasso regression using 5-fold cross-validation.

model to evaluate the accuracy of predicting different time survival rates. The C-index values were 0.723 (95% confidence interval [CI]: 0.719–0.726) and 0.723 (95% CI: 0.718–0.729) in the training and validation set, respectively. The C-index of the TNM system in MLUAD patients was 0.534 (95% CI: 0.530–0.538). The AUC results for 6-, 12-, and 18-month OS were 0.799, 0.764, and 0.750 for the training cohort and 0.799, 0.762, and 0.746 for the validation cohort, respectively (Fig. 5). These C-index and AUC values strongly indicated that the nomogram had high predictive ability and discrimination. Moreover, the calibration curves showed good agreement in the 2 cohorts between the predicted values for this nomogram and the actual values for 6-, 12- and 18-month OS (see Figure S3, Supplemental Digital Content, http://links.lww.com/MD/I79, which illustrates the calibration plots for 6-, 12-, 18-month OS in the training (A–C) and validation (D–F) cohorts), and DCA of 6-, 12-, and 18-month OS prediction demonstrated the practical clinical value of the nomogram (Fig. 6).

We calculated all patient risk scores according to the nomogram model and calculated the cutoff with X-tile software. The results show that the cutoff values of the risk scores (training cohort: 2.220, validation cohort: 1.995), according to the cutoff values, were assigned to the high-risk group and low-risk group in the 2 cohorts. Kaplan–Meier curve results indicated that patients in the low-risk group had a better prognosis (Fig. 7). Finally, we predicted the 6-, 12-, and 18-month OS for a randomly selected patient in the validation cohort (see Figure S4, Supplemental Digital Content, http://links.lww.com/MD/I80, which illustrates the nomograms for predicting 6-, 12-, 18-month OS for a randomly selected patient in validation cohort (*P < .05, **P < .01, ***P < .001).

4. Discussion

LUAD is one of the major pathological types of NSCLC and has high heterogeneity and diversity.[15] Most patients have distant metastases when first diagnosed. Although many newly gene-mutation targeted therapy and immunotherapy options have emerged in the last decade,[16] the prognosis of MLUAD patients is still unsatisfactory and varies greatly.[17] The TNM system is widely used in clinical practice to evaluate the prognosis of patients.[7,18] However, due to the lack of inclusion of many potential prognostic factors, its prognosis is not accurate. Therefore, there is currently no satisfactory prognostic model for MLUAD patients. Studies have shown that nomograms can accurately predict patient outcomes.[19] Therefore, we need to construct a new nomogram by assessing relevant prognostic factors to predict the OS of MLUAD patients, which can help clinicians make the best treatment decisions.

The nomogram model constructed with the above 19 prognostic factors is suitable to predict the OS of MLUAD, having excellent predictive accuracy and clinical applicability. Compared to the TNM system, the C-index of our model was obviously higher (0.723 vs 0.534). In both the training and validation cohorts, the AUC values indicated accurate prediction ability, showing that it provides high discrimination. The calibration curves correspond well to the actual reference line, demonstrating the reproducibility and reliability of this nomogram. The DCA values indicate that our nomogram has good clinical practicability and effectiveness.

This study shows that with increasing age, the prognosis is worse, and the study by Chen et al et al[6] also confirmed this view. With the advances in LC screening technology and treatment drugs,[20] the survival rate of MLUAD patients has improved. Patients diagnosed in 2015 to 2018 had better survival than patients diagnosed in 2010 to 2014. Our study also shows that men have a better prognosis, which may be associated with the

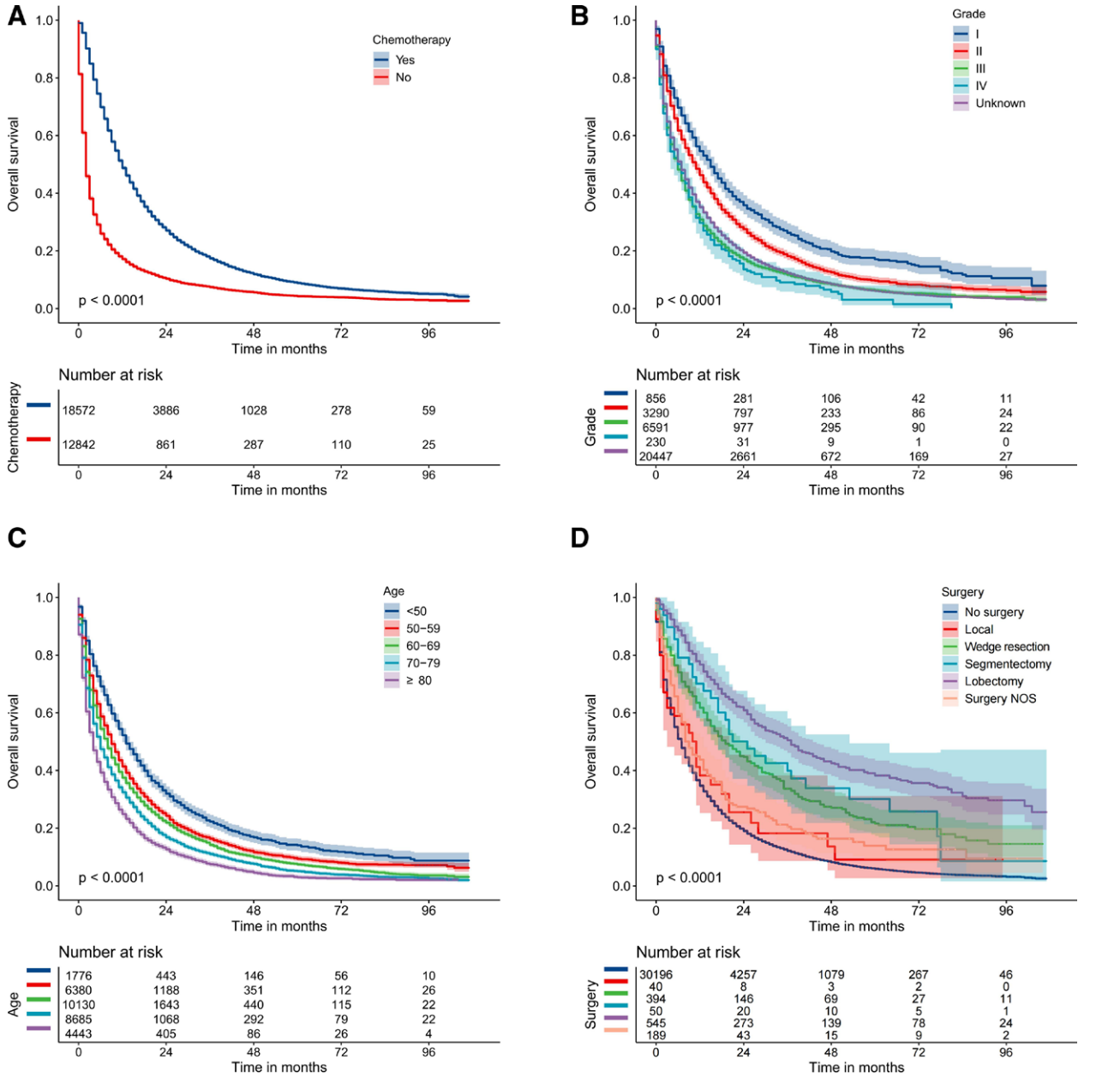

Figure 3. Kaplan–Meier curves of OS for MLUAD patients in chemotherapy (A), grade (B), age (C) and surgery (D). MLUAD = metastatic lung adenocarcinoma, NOS = not otherwise specified, OS = overall survival.

ability of androgens to stimulate LUAD growth.[21] Asians have better outcomes among the different races, and a large survey showed that tumor histology is an important factor leading to the OS differences by race.[22]

Because our subjects were patients with MLUAD, the M stage was M1. The higher the N stage, the worse the prognosis.[23] Surprisingly, T stage was not an independent prognostic factor of MLUAD, but we found that the prognosis of patients with larger tumors was worse, which may be related to the fact that other factors defining T stage mask the impact of tumor size on prognosis[24]; the specific reasons need to be further studied. In terms of tumor grade, patients with higher tumor grade have shorter survival, previous studies showed that poorly differentiated related proteins are associated with higher proliferation.[25] In addition, tumor site was also incorporated into this nomogram model.[26] The prognosis of major bronchial LC is poor, which is in agreement with previous research findings.[27] In regional nodes, regional LN surgery > 3 and regional LN negative were considered protective factors for MLUAD. This is associated with the ability to determine regional LN status after removing more LNs, and patients with regional LN metastases have a high disease recurrence risk and a worse prognosis.[28] Our study found that patients who lived with others had a better prognosis than those who lived alone, possibly because patients living alone lack family support and face greater pressure.[29]

Moreover, Campos-Balea et al[3] reported that the probabilities of LUAD metastasis to the bone, lung, brain and liver were 41.3%, 31.8%, 28.9% and 17.1%, respectively. Similar results were observed in our study. Among 44,848 patients with MLUAD, 19,122 (42.6%), 14,589 (32.5%), 13,789 (30.7%), and 7772 (17.3%) had bone, brain, lung and liver metastases, respectively (Table 1). Multivariate Cox regression analyses indicated that bone, brain and liver metastasis are adverse prognostic factors of MLUAD, but lung metastasis is not an independent prognostic factor for MLUAD ($P > .05$) (Table 2). Liver metastasis has the worst prognosis, and lung metastasis

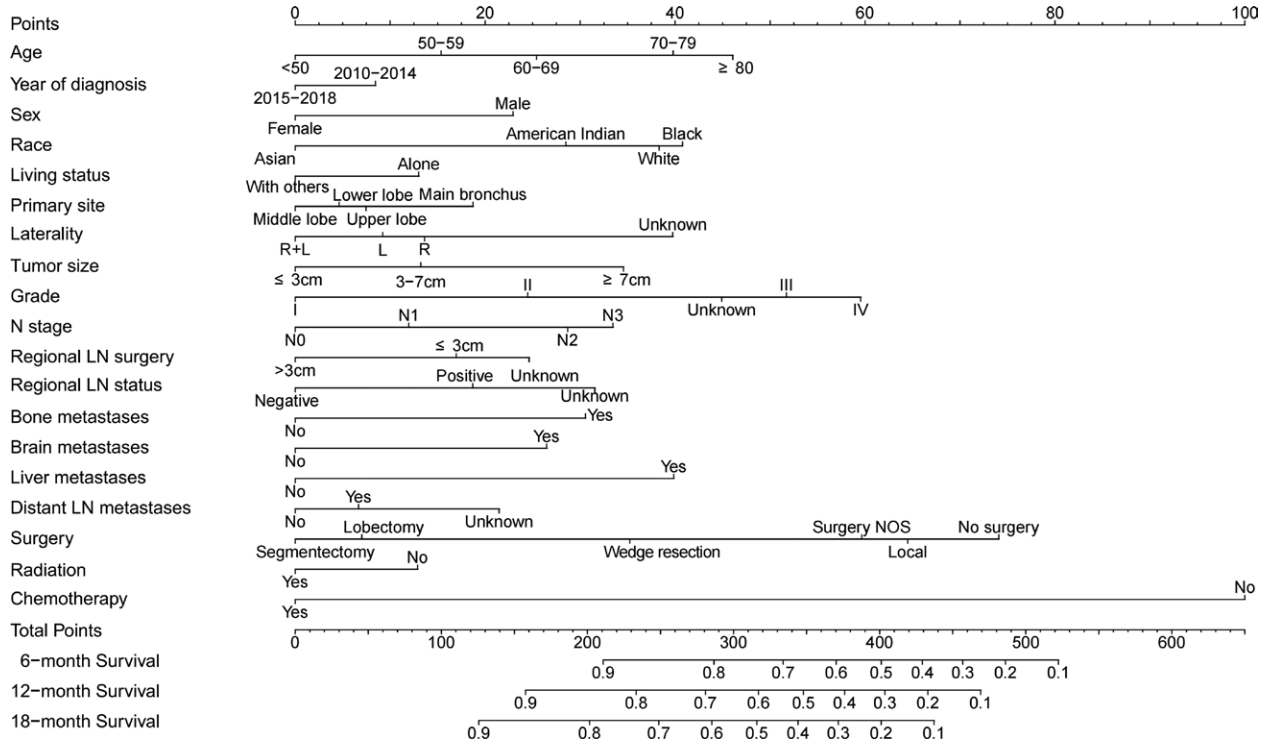

Figure 4. Nomograms for predicting 6-, 12-, 18-mo OS for MLUAD patients. MLUAD = metastatic lung adenocarcinoma, OS = overall survival.

has the best prognosis among the common metastatic sites.[30] Additionally, compared with distant LN metastasis, no distant LN metastasis was associated with better OS.

Systemic comprehensive therapy is the standard treatment for metastatic NSCLC, including traditional chemotherapy and radiotherapy, emerging immunotherapy and targeted therapy.[31] It is worth mentioning that EGFR mutations were found in some tumors of NSCLC patients in 2004, and the survival of EGFR-positive patients was better.[32] Due to some conceptual changes, surgery has been included in the treatment considered. A study of the California Cancer Registry reported longer survival in IV NSCLC patients treated with surgery compared to those who refused surgery (9.4–28 vs 2–10) months in various treatment options,[33] our study results also confirm this idea, It may be related to patients with serious comorbidities that cannot tolerate surgery, such as heart disease.[34] The prognosis is also related to the surgical methods, with local tumor destruction (hazard ratio [HR] 0.91, $P = .599$), wedge resection (HR 0.68, $P < .001$), lobectomy (HR 0.51, $P < .001$), and segmentectomy (HR 0.48, $P < .001$) were all associated with Increased survival when compared with no surgery. The study of Yang et al[35] reported that better survival with lobectomy or segmentectomy than wedge resection. In our nomogram model, surgery and chemotherapy had a greater contribution to the prognosis of MLUAD patients. Although radiotherapy is an independent prognostic factor, it has little contribution to the prognosis of MLUAD patients, suggesting that radiotherapy may be used as a palliative treatment to reduce pain.[36]

This study's strengths include the following 2 points. First, a large sample of patients was included, which could only be obtained in a large multicenter clinical database. Second, more variables were incorporated, significantly improving the predictive ability of the nomogram. However, our research has some limitations. First, the SEER database was missing several significant factors, including targeted therapy, immunotherapy information and some gene mutations, such as EGFR, L858R, ALK and ROS1-related mutations, that have been shown to be associated with a worse prognosis.[37] The inclusion of these potential prognostic factors may enhance the accuracy of the predictive models. Second, we based the nomogram on MLUAD patients in the USA, which may not be representative of MLUAD patients from different countries or other ethnicities. Last, this is a retrospective study, and inevitably, there is some selection bias. Thus, further prospective research is needed to verify and improve the accuracy of this nomogram.

5. Conclusion

In summary, the nomogram model constructed with the above 19 prognostic factors has good predictive accuracy for the OS of MLUAD and good clinical applicability. The top 5 prognostic factors are chemotherapy, grade, age, race and surgery. The model can quantify the independent risk factors into an assessment scale to predict the outcomes of MLUAD patients. We anticipate that more factors will be incorporated into this model in future clinical trials to improve the prognostication for MLUAD patients.

Acknowledgments

The authors would like to thank all public health workers of cancer statistics, all patients and doctors relevant to this research.

Author contributions

JX and WZ had full access to all of the data in the manuscript and takes responsibility for the integrity of the data and the accuracy of the data analysis. All authors read and approved the final manuscript.

Concept and design: All authors.

Acquisition, analysis, or interpretation of data: All authors.

Drafting of the manuscript: Bo Wu, Jianhui Chen, Nan Feng Zhongtian Xiang, Xiang Zhang.

Critical revision of the manuscript for important intellectual content: Bo Wu, Junping Xie, Wenxiong Zhang.

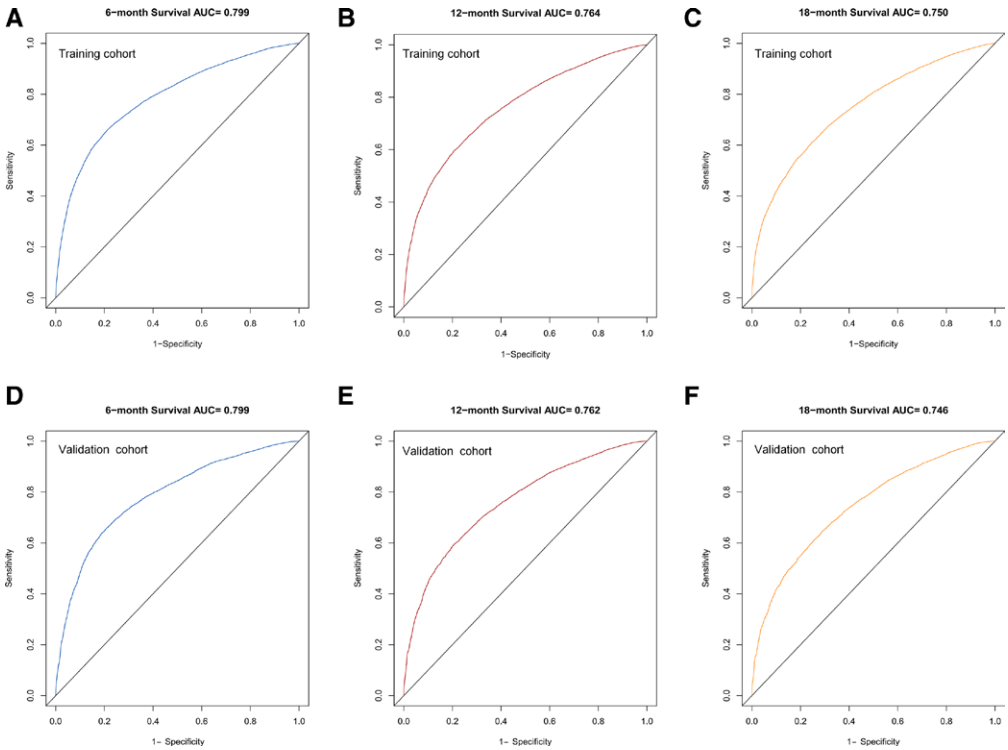

Figure 5. The ROC curve and AUC for 6-, 12-, 18-mo OS in the training (A–C) and validation (D–F) cohorts. AUC = area under the curve, OS = overall survival, ROC = receiver operating characteristics.

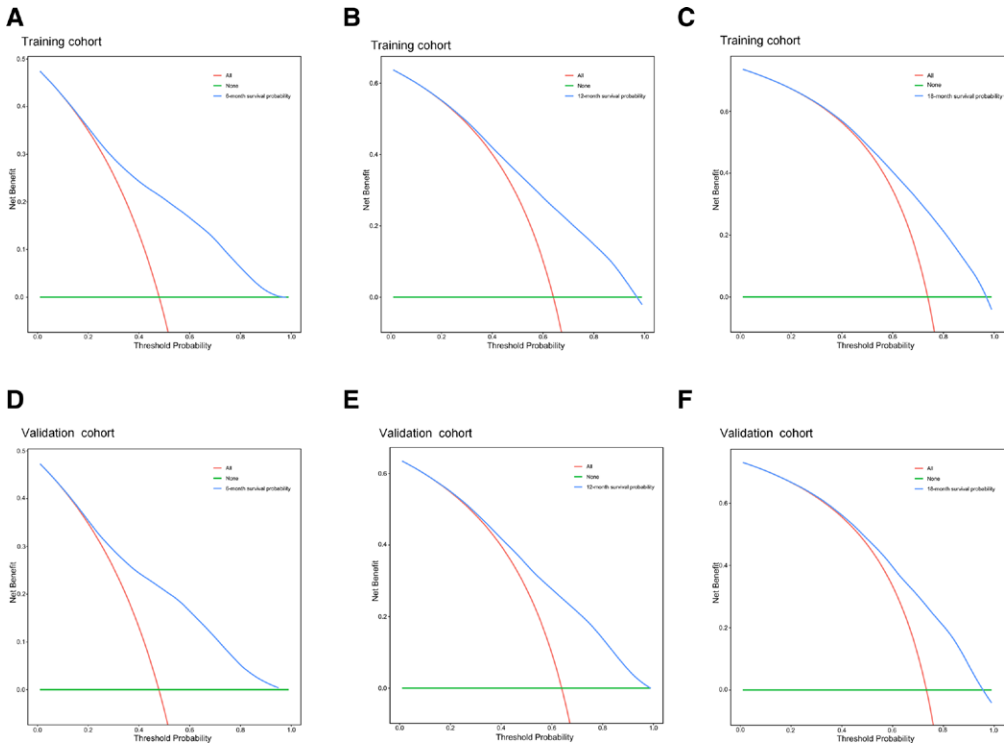

Figure 6. DCA of the nomogram for 6-, 12-, 18-mo OS in the training (A–C) and validation (D–F) cohorts. DCA = decision curve analyses, OS = overall survival.

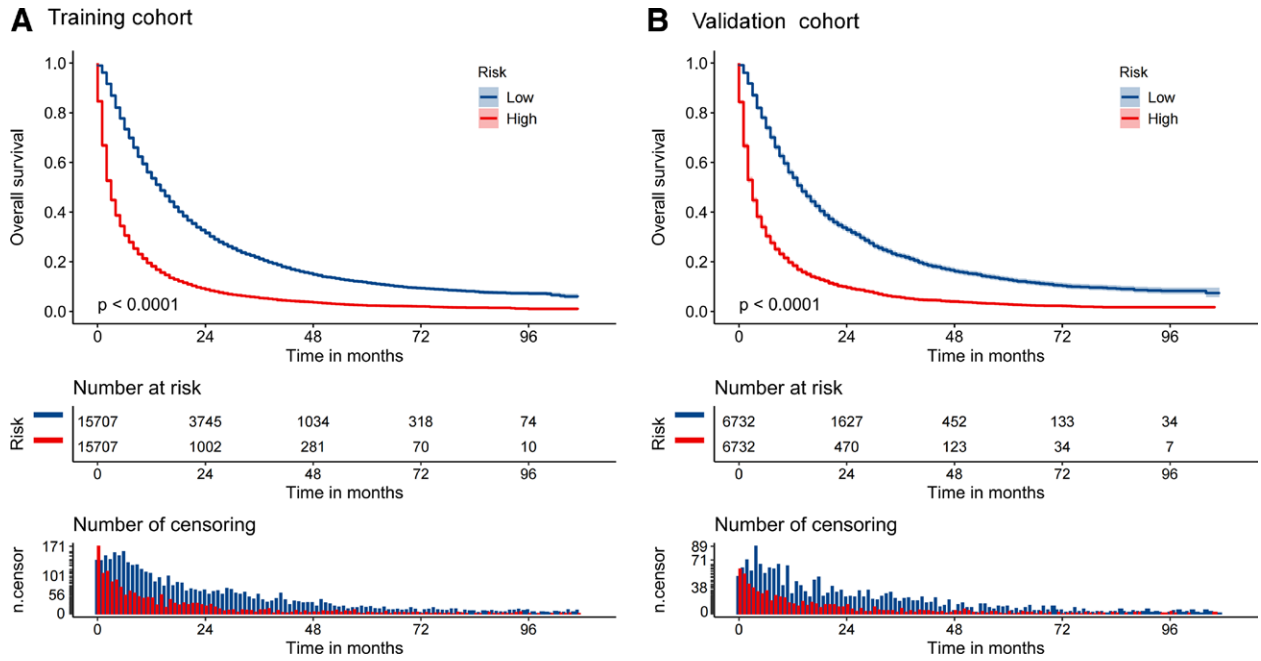

Figure 7. Kaplan–Meier curves of the OS for patients the low-risk group and the high-risk group both the training (A) and validation cohorts (B). OS = overall survival.

Statistical analysis: Bo Wu, Jianhui Chen, Junping Xie, Wenxiong Zhang.
Supervision: Bo Wu, Jianhui Chen, Junping Xie, Wenxiong Zhang.

References

[1] Siegel RL, Miller KD, Fuchs HE, et al. Cancer statistics, 2022. CA Cancer J Clin. 2022;72:7–33.
[2] You H, Teng M, Gao CX, et al. Construction of a nomogram for predicting survival in elderly patients with lung adenocarcinoma: a retrospective cohort study. Front Med (Lausanne). 2021;8:680679.
[3] Campos-Balea B, de Castro Carpeño J, Massutí B, et al. Prognostic factors for survival in patients with metastatic lung adenocarcinoma: an analysis of the SEER database. Thorac Cancer. 2020;11:3357–64.
[4] Sugawara S, Lee JS, Kang JH, et al. Nivolumab with carboplatin, paclitaxel, and bevacizumab for first-line treatment of advanced nonsquamous non-small-cell lung cancer. Ann Oncol. 2021;32:1137–47.
[5] Riihimäki M, Hemminki A, Fallah M, et al. Metastatic sites and survival in lung cancer. Lung Cancer. 2014;86:78–84.
[6] Chen H, Huang C, Ge H, et al. A novel LASSO-derived prognostic model predicting survival for non-small cell lung cancer patients with M1a diseases. Cancer Med. 2022;11:1561–72.
[7] Wankhede D. Evaluation of Eighth AJCC TNM sage for lung cancer NSCLC: a meta-analysis. Ann Surg Oncol. 2021;28:142–7.
[8] Park SY. Nomogram: an analogue tool to deliver digital knowledge. J Thorac Cardiovasc Surg. 2018;155:1793.
[9] Diao JD, Ma LX, Wu CJ, et al. Construction and validation a nomogram to predict overall survival for colorectal signet ring cell carcinoma. Sci Rep. 2021;11:3382.
[10] Hua X, Huang X, Huang ZZ, et al. Establishment of prognostic nomograms based on skeletal muscle index and serum biomarker in breast cancer patients receiving radiotherapy. Clin Transl Med. 2020;10:e115.
[11] Sjoquist KM, Renfro LA, Simes RJ, et al.; Fondation Aide et Recherche en Cancerologie Digestive Group (ARCAD). Personalizing survival predictions in advanced colorectal cancer: the ARCAD nomogram project. J Natl Cancer Inst. 2018;110:638–48.
[12] Liang W, Zhang L, Jiang G, et al. Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer. J Clin Oncol. 2015;33:861–9.
[13] Kagiyama N, Piccirilli M, Yanamala N, et al. Machine learning assessment of left ventricular diastolic function based on electrocardiographic features. J Am Coll Cardiol. 2020;76:930–41.
[14] Zhang XS, Gale RP, Zhang MJ, et al. A predictive scoring system for therapy-failure in persons with chronic myeloid leukemia receiving initial imatinib therapy. Leukemia. 2022;36:1336–42.
[15] Succony L, Rassl DM, Barker AP, et al. Adenocarcinoma spectrum lesions of the lung: detection, pathology and treatment strategies. Cancer Treat Rev. 2021;99:102237.
[16] Bai W, Zhang J, Wang Y, et al. Comparative analysis of the long-term outcomes of segmentectomy and lobectomy for stage IA1 lung adenocarcinoma in patients with or without previous malignancy of other organs: a population-based study. Expert Rev Anticancer Ther. 2022;22:215–28.
[17] Wang M, Herbst RS, Boshoff C. Toward personalized treatment approaches for non-small-cell lung cancer. Nat Med. 2021;27:1345–56.
[18] Hwang JK, Page BJ, Flynn D, et al. Validation of the Eighth Edition TNM lung cancer staging system. J Thorac Oncol. 2020;15:649–54.
[19] Gittleman H, Sloan AE, Barnholtz-Sloan JS. An independently validated survival nomogram for lower-grade glioma. Neuro Oncol. 2020;22:665–74.
[20] Oudkerk M, Liu S, Heuvelmans MA, et al. Lung cancer LDCT screening and mortality reduction – evidence, pitfalls and future perspectives. Nat Rev Clin Oncol. 2021;18:135–51.
[21] Cheng TD, Darke AK, Redman MW, et al. Smoking, sex, and non-small cell lung cancer: steroid hormone receptors in tumor tissue (S0424). J Natl Cancer Inst. 2018;110:734–42.
[22] Ellis L, Canchola AJ, Spiegel D, et al. Racial and ethnic disparities in cancer survival: the contribution of tumor, sociodemographic, institutional, and neighborhood characteristics. J Clin Oncol. 2018;36:25–33.
[23] Song H, Yoon SH, Kim J, et al. Application of N descriptors proposed by the international association for the study of lung cancer in clinical staging. Radiology. 2021;300:450–7.
[24] Hoeijmakers F, Heineman DJ, Daniels JM, et al. Variation between multidisciplinary tumor boards in clinical staging and treatment recommendations for patients with locally advanced non-small cell lung cancer. Chest. 2020;158:2675–87.
[25] Wang T, Zhou J, Wang Y, et al. Clinicopathological characteristics and prognosis of resectable lung adenosquamous carcinoma: a population-based study of the SEER database. Jpn J Clin Oncol. 2022;52:1191–200.
[26] Yin Y, Qin J. Multi-line treatment with tyrosine kinase inhibitors enabled in 4-year survival for patient with stage IV lung adenocarcinoma: a case report. Ann Palliat Med. 2021;10:6990–6.
[27] Liao Y, Fan X, Wang X. Effects of different metastasis patterns, surgery and other factors on the prognosis of patients with stage IV non-small cell lung cancer: a surveillance, epidemiology, and end results (SEER) linked database analysis. Oncol Lett. 2019;18:581–92.

[28] Liang W, He J, Shen Y, et al. Impact of examined lymph node count on precise staging and long-term survival of resected non-small-cell lung cancer: a population study of the US SEER database and a Chinese multi-institutional registry. J Clin Oncol. 2017;35:1162–70.
[29] Kehl KL, Zahrieh D, Yang P, et al. Rates of guideline-concordant surgery and adjuvant chemotherapy among patients with early-stage lung cancer in the US ALCHEMIST study (Alliance A151216). JAMA Oncol. 2022;8:717–28.
[30] Corrao G, Marvaso G, Ferrara R, et al. Stereotatic radiotherapy in metastatic non-small cell lung cancer: combining immunotherapy and radiotherapy with a focus on liver metastases. Lung Cancer. 2020;142:70–9.
[31] Ettinger DS, Wood DE, Aisner DL, et al. NCCN guidelines insights: non-small cell lung cancer, version 2.2021. J Natl Compr Canc Netw. 2021;19:254–66.
[32] Hanna NH, Schneider BJ, Temin S, et al. Therapy for stage IV non-small-cell lung cancer without driver alterations: ASCO and OH (CCO) joint guideline update. J Clin Oncol. 2020;38:1608–32.
[33] David EA, Clark JM, Cooke DT, et al. The role of thoracic surgery in the therapeutic management of metastatic non-small cell lung cancer. J Thorac Oncol. 2017;12:1636–45.
[34] Abdel-Rahman O. Risk of cardiac death among cancer survivors in the United States: a SEER database analysis. Expert Rev Anticancer Ther. 2017;17:873–8.
[35] Yang CJ, Gu L, Shah SA, et al. Long-term outcomes of surgical resection for stage IV non-small-cell lung cancer: a national analysis. Lung Cancer. 2018;115:75–83.
[36] Vinod SK, Hau E. Radiotherapy treatment for lung cancer: current status and future directions. Respirology. 2020;25(Suppl 2):61–71.
[37] Tan AC, Tan DSW. Targeted therapies for lung cancer patients with oncogenic driver molecular alterations. J Clin Oncol. 2022;40:611–25.