Heliyon 9 (2023) e15924

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Development and validation of a novel nomogram to predict the overall survival of patients with large cell lung cancer: A surveillance, epidemiology, and end results population-based study

Hongxia Zhou [a,1], Pengxiang Gao [b,1], Fangpeng Liu [b], Liangliang Shi [c], Longhua Sun [b], Wei Zhang [b,d,e], Xinping Xu [b,d,e,*], Xiujuan Liu [a,**]

[a] Department of Nephrology, The 908th Hospital of the People's Liberation Army Joint Logistics Support Force, The Great Wall Affiliated Hospital of Nanchang University, Nanchang City, Jiangxi 330006, China
[b] Department of Pulmonary and Critical Care Medicine, The First Affiliated Hospital of Nanchang University, Nanchang City, Jiangxi Province 330006, China
[c] Medical Center of Burn Plastic and Wound Repair, The First Affiliated Hospital of Nanchang University, Nanchang City, Jiangxi Province 330006, China
[d] Jiangxi Institute of Respiratory Diseases, The First Affiliated Hospital of Nanchang University, Nanchang City, Jiangxi Province 330006, China
[e] Jiangxi Clinical Research Center for Respiratory Diseases, The First Affiliated Hospital of Nanchang University, Nanchang City, Jiangxi Province 330006, China

ARTICLE INFO

Keywords:
Large cell lung cancer
Nomogram
Overall survival
Prognostic

ABSTRACT

Background: Large cell lung cancer (LCLC) is a rare subtype of non-small cell lung carcinoma (NSCLC), and little is known about its clinical and biological characteristics.
Methods: LCLC patient data were extracted from the Surveillance, Epidemiology, and End Results (SEER) database between 2004 and 2015. All patients were randomly divided into a training group and a validation group at a ratio of 7:3. The independent prognostic factors that were identified ($P < 0.01$) by stepwise multivariate Cox analysis were incorporated into an overall survival (OS) prediction nomogram, and risk-stratification systems, C-index, time-ROC, calibration curve, and decision curve analysis (DCA) were applied to evaluate the quality of the model.
Results: Nine factors were incorporated into the nomogram: age, sex, race, marital status, 6th AJCC stage, chemotherapy, radiation, surgery and tumor size. The C-index of the predicting OS model in the training dataset and in the test dataset was 0.757 ± 0.006 and 0.764 ± 0.009, respectively. The time-AUCs exceeded 0.8. The DCA curve showed that the nomogram has better clinical value than the TNM staging system.
Conclusions: Our study summarized the clinical characteristics and survival probability of LCLC patients, and a visual nomogram was developed to predict the 1-year, 3-year and 5-year OS of

Abbreviations: LCLC, large cell lung cancer; SEER, Surveillance, Epidemiology and End Results; OS, overall survival; AJCC, American Joint Committee on Cancer; C-index, Concordance index; ROC, receiver operating characteristic; DCA, decision curve analysis; AUC, area under the curve; NSCLC, non-small cell lung cancer; CSS, cancer-specific survival.

* Corresponding author. Department of Pulmonary and Critical Care Medicine, The First Affiliated Hospital of Nanchang University, Nanchang City, Jiangxi Province 330006, China.
** Corresponding author.
E-mail addresses: xinpingxu@ncu.edu.cn (X. Xu), liuxinfanjian@126.com (X. Liu).
[1] Co-first author, Hongxia Zhou and Pengxiang Gao contributed equally to this article.

https://doi.org/10.1016/j.heliyon.2023.e15924
Received 8 August 2022; Received in revised form 12 April 2023; Accepted 26 April 2023
Available online 6 May 2023

LCLC patients. This provides more accurate OS assessments for LCLC patients and helps clinicians make personal management decisions.

1. Introduction

Lung carcinoma has the second-highest incidence of malignancy and is the primary cause of tumor-related deaths at present according to global cancer statistics 2020 [1], resulting in a serious burden on global health. In the world's lung cancer population, large cell lung cancer (LCLC) is a rare subtype that accounts for only 9% of cases, as reported by the World Health Organization (WHO) [2]. LCLC is also a type of undifferentiated carcinoma that lacks the cellular differentiation and structural features of small-cell lung cancer, squamous cell carcinoma and adenocarcinoma [3]. According to the 2015 WHO Lung Tumors Classification, large cell neuroendocrine carcinoma (LCNEC) was excluded from traditional LCLC and was grouped with other neuroendocrine carcinomas at the same time [4]. As a predictive method, nomograms have been used widely to predict the outcome of related cancers, and visualization and ease of operation are two major advantages [5,6].

However, few prognostic models are available to predict LCLC. In particular, studies that have evaluated the treatment modalities, including radiation and chemotherapy, are rare [7,8]. LCLC grows quickly and spreads quickly [9]. It is very important to carry out correct diagnosis and treatment strategies within a short time. Because LCLC has a lower incidence than common types of lung cancer and because there is a lack of collation of relevant clinical information, and little understanding of their clinical and biological characteristics, the relevant clinical data from patients diagnosed with LCLC were collected and analyzed in our retrospective study, and we decided to develop a novel and high-efficiency nomogram to facilitate the clinician in making treatment decisions.

2. Methods

2.1. Statement of ethical principles

We have been granted access to the Surveillance, Epidemiology, and End Results (SEER) database, and the login account was 15885-Nov2020. The LCLC patients' informed consent was not required in our study because the data were public and delabeled at the same time. We conducted this study in accordance with the requirements from the Declaration of Helsinki.

2.2. Patient data

All LCLC patients concrete information was obtained from the SEER database, and SEER*Stat software (version 8.35) was used for data filtration in this large-scale, population-based retrospective study. The SEER database converges data from the National Cancer Institute's 18 cancer registries, including data from approximately 30% of the US population [10]. The rational cutoff date was set at December 31, 2015. We conducted rigorous patient data screening by collating case data (Supplementary Fig. 1). The detailed criteria for inclusion and exclusion were as follows: (1) diagnosed with large cell lung cancer between 2004 and 2015, based on pathological confirmation; (2) first malignant primary indicator, originating in the main bronchus and lung (primary site code: C34.0-C34.9); (3) International Classification of Diseases code O-3 (ICD-O-3) morphology was 8012/3; and (4) incomplete records for variables such as American Joint Committee on Cancer (AJCC) TNM staging, therapeutic method, and other key clinical feature information were excluded.

All selected patients were followed up based on the SEER database protocols, and complete information was recorded, such as follow-up time, survival status and survival time. Therefore, we were able to investigate the duration of follow-up and overall survival (OS) of these LCLC patients. Of special note, missing data that could not be used to assess the survival status were eliminated before the statistical analysis. We gathered LCLC patients' demographic characteristics, including age at diagnosis, race, sex, and marital status, and solid tumor clinicopathological features, such as primary site, lateral location, grade classification, AJCC stage (6th edition), tumor size, chemotherapy, radiation, and surgery.

2.3. Follow-up

Overall survival, which was defined as the time interval from diagnosis to death from any cause, was regarded as the observational endpoint in this study. The survival analysis had a follow-up duration from 0 to 179.0 months and a median of 12.0 months (95% CI 10.977–13.023). The selected LCLC patients who participated in the survival analysis had a clear status of survival, death or alive.

2.4. Statistical analysis

Xtile (3.6.1 Yale University USA) [11] was applied to find the best cutoff value for the continuous variables, including age and tumor size. The Kaplan–Meier method was used for cumulative survival curve determining, and log-rank tests were applied to compare different curves. The Cox proportional hazards model was employed to conduct univariate and multivariate analyses. In the multivariate Cox analysis, only those variables that were significantly associated with survival in univariate Cox analysis Hazard ratios (HRs) and 95% confidence intervals (CIs) were selected for inclusion in the model. The novel nomogram was delineated on the basis of

the results ($P < 0.01$) of the multivariate Cox analysis. Calibration curves (Fig. 3) and time-ROC curves (Fig. 4A and B) were applied to assess the versatility and accuracy of the established nomogram. Decision curve analysis (DCA) [12,13]was used to compare clinical benefits with different prediction models. The prediction error was estimated with 1000 bootstrap samples. Both reported significance levels were two-sided, and the statistical significance level was 0.05.

Statistical analysis was performed with the software R version 4.2.0 http://www.r-project.org and SPSS version 25.0 (IBM, Chicago, IL). GraphPad Prism version 8.2.1 (GraphPad Software, San Diego, CA) was launched to depict the Kaplan–Meier curve. R packages such as rms, survival, nomogram Formula, timeROC, and ggDCA were used in statistical analyses.

Table 1
Baseline data and clinical features of 2817 patients with large cell lung cancer.

Variables	Total	Training dataset	Testing dataset	P value
Year of diagnosis	2004–2015			
Number of patients	2817	1972	845	
Age				
70≤	1722 (61.1%)	1198 (60.8%)	524 (62%)	0.668
71–77	627 (22.3%)	448 (22.7%)	179 (21.2%)	
≥78	468 (16.6%)	326 (16.5%)	142 (16.8%)	
Sex				
female	1182 (42%)	826 (41.9%)	356 (42.1%)	0.904
male	1635 (58%)	1146 (58.1%)	489 (57.9%)	
Race				
white	2271 (80.6%)	1597 (81%)	674 (79.7%)	0.608
black	402 (14.3%)	273 (13.8%)	129 (15.3%)	
other	144 (5.1%)	102 (5.2%)	42 (5%)	
Marital status				
married	1499 (53.2%)	1066 (54.1%)	433 (52.4%)	0.426
unmarried	1318 (46.8%)	906 (45.9%)	412 (47.6%)	
Primary Site labeled				
upper lobe	1724 (61.2%)	1224 (62.1%)	500 (59.2%)	0.387
middle lobe	135 (4.8%)	97 (4.9%)	38 (4.5%)	
Lower lobe	683 (24.2%)	466 (23.6%)	217 (25.7%)	
other	271 (9.6%)	185 (9.4%)	90 (10.6%)	
Laterality				
left	1159 (41.1%)	825 (41.8%)	334 (39.5%)	0.254
right	1658 (58.9%)	1147 (58.2%)	511 (60.5%)	
Histologic Grade				
I-II	53 (1.9%)	40 (2%)	13 (1.5%)	0.380
III-IV	2764 (98.1%)	1932 (98%)	832 (98.5%)	
Tumor size				
≤45 mm	1521 (54%)	1089 (55.2%)	432 (51.1%)	0.045
> 45	1246 (46%)	883 (44.8%)	413 (48.9%)	
T stage				
T1	586 (20.8%)	417 (21.1%)	169 (20.0%)	0.627
T2	1178 (41.8%)	824 (41.8%)	354 (41.9%)	
T3	255 (9.1%)	184 (9.3%)	71 (8.4%)	
T4	798 (28.3%)	547 (27.7%)	251 (29.7%)	
N stage				
N0	1349 (47.9%)	955 (48.4%)	394 (46.6%)	0.299
N1	297 (10.5%)	200 (10.1%)	97 (11.5%)	
N2	907 (32.2%)	623 (31.6%)	284 (33.6%)	
N3	264 (9.4%)	194 (9.8%)	70 (8.3%)	
M stage				
M0	1817 (64.5%)	1277 (64.8%)	540 (63.9%)	0.665
M1	1000 (35.5%)	695 (35.2%)	305 (36.1%)	
Chemotherapy				
No/unkown	1619 (57.5%)	1135 (57.6%)	484 (57.3%)	0.891
Yes	1198 (42.5%)	837 (42.4%)	361 (42.7%)	
Radiation				
No/unknown	1694 (60.1%)	1192 (60%)	502 (59.4%)	0.606
Yes	1123 (39.9%)	780 (40%)	343 (40.6%)	
Surgery				
No	1546 (54.9%)	1079 (54.7%)	467 (55.3%)	0.788
Yes	1271 (45.1%)	893 (45.3%)	378 (44.7%)	
Stage				
I	859 (30.5%)	602 (30.5%)	257 (30.4%)	0.943
II	249 (8.8%)	178 (9.0%)	71 (8.4%)	
III	709 (25.2%)	497 (25.2%)	212 (25.1%)	
IV	1000 (35.5%)	695 (35.2%)	305 (36.1%)	

3. Results

3.1. Clinicopathologic characteristics of patients with LCLC

A total of 2817 patients were diagnosed with LCLC from 2004 to 2015. The LCLC patients' average age was 66.5 ± 10.8 years, and the males (58%) slightly outnumbered the females (42%) regarding the percentage of these patients. The majority (53.2%) of the patients were married, and 58.9% of the solid tumors were located in the right lung more than in the left lung (41.1%). Meanwhile, the most common LCLC lesions were in the upper lobe (61.2%), and the lower lobe was a close second (24.2%). In addition, the overwhelming majority (98.1%) of the LCLCs were poorly differentiated or nondifferentiated. The mean tumor size of the LCLCs in the patients was 49.6 ± 31 mm.

Accordingly, tumors at stage III/IV accounted for 60.7%, while stage I and stage II tumors accounted for 30.5% and 8.8%, respectively, and all characteristics of the traindataset were similar compared to the testdataset.

3.2. One-, three- and five-year OS prediction nomogram model

In total, 2817 LCLC patients were randomly divided into traindataset (1972) and testdataset (845) by a ratio of 7:3; the 7:3 ratio is a mainstream training/testing dataset classification in clinical prognostic model study when the sample size is not huge [14,15] (Table 1). Simultaneously, these patients were classified into 3 groups in terms of age (≤70 years,71–77years, ≥78 years) and 2 groups in terms of tumor size (≤45 mm, >45 mm) in the traindataset. In a single-factor analysis, we found that variables were significantly correlated with OS in the training cohort (Table 2). The potential redundancy was eliminated on the basis of the AIC-based backward

Table 2
Univariate and multivariate Cox analyses of OS for patients with large cell lung cancer HR, hazard ratio; 95% CI, 95% confidence interval.

Variables	Univariate analysis		Multivariate analysis	
	HR (95% CI)	P value	HR (95% CI)	P value
Age				
70≤	1.0 (reference)		1.0 (reference)	
71–77	1.156 (1.029–1.298)	0.015	1.273 (1.130–1.434)	< 0.001
≥78	1.805 (1.588–2.051)	< 0.001	1.563 (1.362–1.793)	< 0.001
Sex				
male	1.0 (reference)		1.0 (reference)	
female	0.828 (0.752–0.913)	< 0.001	0.837 (0.755–0.927)	0.001
Race				
white	1.0 (reference)		1.0 (reference)	
black	1.044 (0.909–1.2)	0.539	0.903 (0.783–1.041)	0.160
other	0.836 (0.670–1.042)	0.110	0.693 (0.555–0.866)	0.001
Marital status				
married	1.0 (reference)		1.0 (reference)	
unmarried	1.153 (1.048–1.268)	0.004	1.203 (1.087–1.331)	< 0.001
PrimarySitelabeled				
upper lobe	1.0 (reference)			
middle lobe	0.951 (0.760–1.191)	0.662		
Lower lobe	1.133 (1.012–1.270)	0.031		
other	1.518 (1.290–1.786)	< 0.001		
Laterality				
left	1.0 (reference)		1.0 (reference)	
right	0.953 (0.865–1.049)	0.323	0.893 (0.810–0.984)	0.022
Histologic Grade				
I-II	1.0 (reference)			
III-IV	1.108 (0.789–1.556)	0.553		
Tumor size				
≤45 mm	1.0 (reference)		1.0 (reference)	
> 45	1.62 (1.472–1.784)	< 0.001	1.272 (1.150–1.408)	< 0.001
Chemotherapy				
No/unkown	1.0 (reference)		1.0 (reference)	
YES	0.918 (0.834–1.011)	0.083	0.550 (0.491–0.617)	< 0.001
Radiation				
No/unknown	1.0 (reference)		1.0 (reference)	
YES	1.382 (1.254–1.524)	< 0.001	0.806 (0.719–0.902)	< 0.001
Surgery				
No	1.0 (reference)		1.0 (reference)	
Yes	0.294 (0.265–0.326)	< 0.001	0.390 (0.338–0.450)	< 0.001
Stage				
I	1.0 (reference)		1.0 (reference)	
II	1.473 (1.220–1.778)	< 0.001	1.687 (1.389–2.049)	< 0.001
III	2.165 (1.893–2.476)	< 0.001	2.041 (1.736–2.401)	< 0.001
IV	4.551 (4.008–5.168)	< 0.001	3.565 (3.035–4.188)	< 0.001

selection procedure in the multivariate Cox proportional hazards regression analysis. Ultimately, multivariate factors ($P < 0.01$), including age, sex, race, marital status, laterality, stage I-IV, tumor size, chemotherapy, radiation and surgery, were recruited to exploit the nomogram (Fig. 1). Supplementary Table 1 shows the hazard ratios for the nomogram parameters. Subgroup analysis of the nomogram: the total nomogram scores were divided into three levels: high risk (total nomogram scores > 250), medium risk (150 < total nomogram scores ≤250) and low risk (total nomogram scores≤150). We plotted the Kaplan–Meier curves of each subset (Fig. 2).

3.3. External validation

The concordance index (C-index) of the nomogram for predicting OS in the training dataset and test dataset was 0.757 ± 0.006 and 0.764 ± 0.009, respectively. The calibration curve for one-year, three-year, and five-year OS showed little gap between the predictions and actual outcomes in the validation cohort (Fig. 3). The time-dependent AUCs [16] of the validation cohort for one-year, three-year, and five-year OS were 0.85, 0.84, and 0.85, respectively. The time-dependent AUCs of the training data cohort for one-year, three-year, and five-year OS were 0.84, 0.85, and 0.84, respectively (Fig. 4A and B). DCA [17] was used to compare the clinical benefits of the nomogram with the 6th TNM staging system (Fig. 4C); the nomogram achieved greater net benefit than the 6th TNM staging system in terms of a wide range of threshold probabilities in traindataset. At one year, three years and five years and with the same risk threshold, it can always obtain a higher net benefit.

4. Discussion

The diagnosis and treatment of LCLC has always been a difficult problem. We reviewed the proportion of LCLC cases within the cases of lung cancer rom 2004 to 2015 (Supplementary Fig. 2) [18], and the number of LCLC cases have decreased year by year distinctly. The median survival of LCLC was approximately 1 year regardless of whether OS or CSS was the event endpoint (Supplementary Fig. 3). According to a recent IHC analysis, a portion of LCLC cases are now classified as lung squamous cell carcinoma or lung adenocarcinoma. Consequently, LCLC has now become one of the rarest subtypes of NSCLC [19]. Rare tumors have received increasing attention and research in recent years [20,21]. In a single-factor analysis, we found that radiation therapy was a risk factor HR = 1.382 95% CI (1.254–1.524), but among multiple factors, radiation therapy was protective HR = 0.86 95% CI (0.719–0.902).

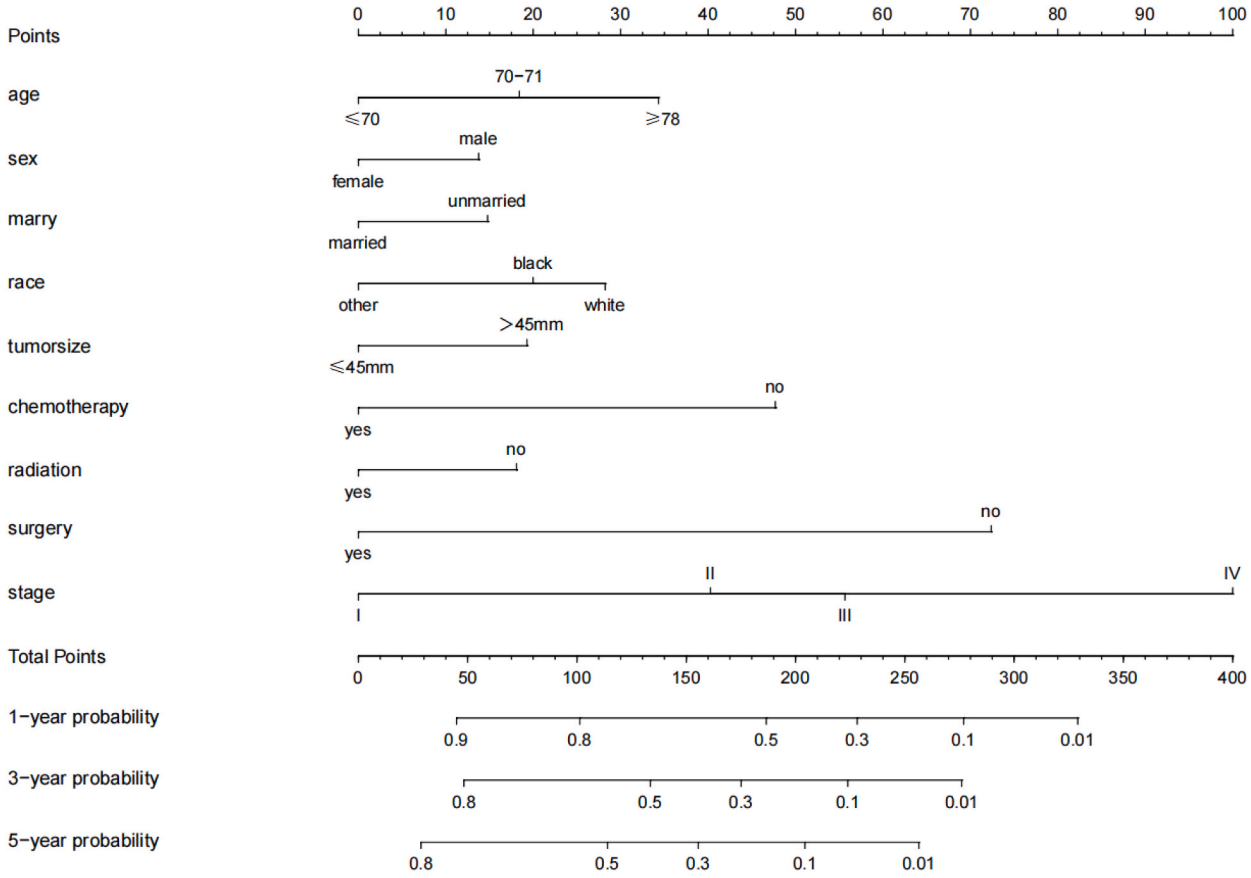

Fig. 1. The nomogram to predict 1-, 3-, 5-year overall survival of large cell lung cancer patients.

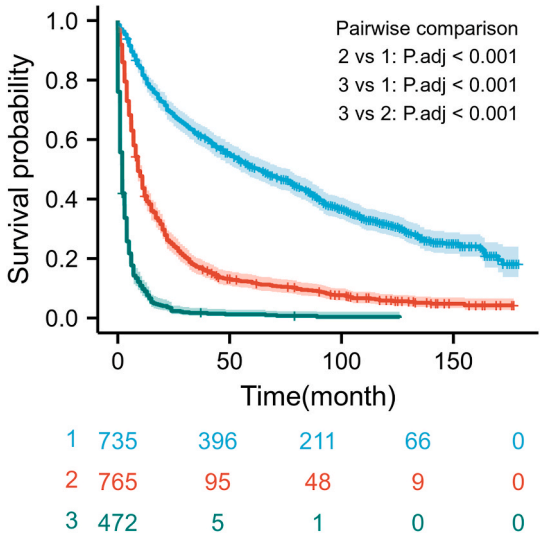

Fig. 2. The Kaplan Meier curve for re-grouped patients with large cell lung cancer according to the nomogram scores. OS, cancer-specific survival. Risk score level:3:high risk (total nomo scores > 250) 2:medium risk (150 < total nomo scores ≤250) and 1: low risk (total nomo scores≤150).

Radiotherapy is known to benefit non-small cell lung cancer in many cases, but the indications for patient treatment and the patient's willingness also need to be taken into account [22,23].

Nomograms are widely used as prediction tools in oncology. Rapid computation through user-friendly digital interfaces, together with increased accuracy, and more easily understood prognoses compared with conventional staging, allow for seamless incorporation of nomogram-derived prognosis to aid clinical decision making [24]. Striking a balance between the ease of using the nomogram and the accuracy of prediction, we included the independent risk Factor $P < 0.01$ and excluded the variable laterality ($P = 0.022$). Through the screening of variables and consideration of clinical significance, we finally selected 9 variables: age, sex, race, marital status, stage, tumor size, chemotherapy, radiation, and surgery for the drawing of the nomogram. When the patient is diagnosed, the information of the first six variables will be determined. We can temporarily choose no treatment to predict the survival rate of the patient. Then, the treatment methods were included separately to observe the change in the predicted survival rate; that is, the patients knew the improvement in the predicted survival rate after treatment. Supplementary Table 1 can help clinicians determine nom scores very quickly. According to the total score, we can obtain the 1-year, 3-year and 5-year survival rates by using the nomogram (Fig. 1). In the meantime, the patients with LCLC were segmented into three levels on the basis of total Nom score strata: high risk, medium risk, and low risk. This model is easy to operate and can better hierarchically indicate the outcome of patients. A series of methods, including the C-index, calibration plots, time-ROC and decision curve analysis, were used to strictly evaluate the LCLC prediction model [25]. Our evaluation system is more advantageous in the evaluation of patient prognosis than the traditional TNM staging by DCA.

There are few studies of LCLC in the SEER database, and clinical data are not fully utilized. Shi et al. performed a study on LCLC that focused only on treatment with surgery and not radiation or chemotherapy [7]. Tai and his colleagues ignored the importance of radiotherapy in their study of LCLC [26]. Radiation therapy, as a common treatment, will be routinely administered, and patients will certainly benefit [8,27], which is consistent with our research conclusions. Common therapies such as surgery, chemotherapy and radiotherapy have a significant impact on the prognosis of LCLC, and there have been studies that consider combinations of these therapies [26]. The forest plot reveals huge benefits of surgery for LCLC patients (Supplementary Fig. 4). The goal of treatment of early lung cancer patients is cure. Multidisciplinary discussions on surgical resectability and medical feasibility have determined local (surgical or radiotherapy) and associated systemic approaches to further improve the possibility of cure [28,29]. At the same time, we used all-cause death as our primary endpoint because we hold the opinion that therapeutic factors are greatly influential in the LCLC prediction model. Since 2015, large cell neuroendocrine carcinoma and combined large cell neuroendocrine carcinoma (8013/3) have been classified as lung neuroendocrine neoplasms [30]. To ensure that the nomogram model can continue to be suitable for future clinical application, we only selected patients with large cell carcinoma (8012/3). Our study is subject to several limitations. On the one hand, even though our nomogram included key variables wherever possible, smoking status, radiation intensity, and options of chemotherapeutics were not recorded in the database, and consequently, these were not included in the nomogram. Because our patient data were selected from 2004 to 2015, we first used version 6th TNM staging to compare with our prognostic model via DCA, and then, we completed the transformation of TNM staging and evaluated the patient data according to the 8th edition. There is no essential difference in staging between the two editions, while the staging of the eighth edition TNM is more detailed [31]. This model is superior to TNM staging, whether it is the 6th (Fig. 4C) or the 8th edition (Supplementary Fig. 5). On the other hand, selective bias is unavoidable in retrospective studies. Our conclusions need to be further confirmed and improved with prospective studies. The prognostic nomogram that we created will help doctors better address LCLC patients' practical problems and provide a guide for treatment decision-making to improve long-term survival and quality of life.

In summary, the novel prediction model developed in our study precisely predicted the prognosis of LCLC patients and performed

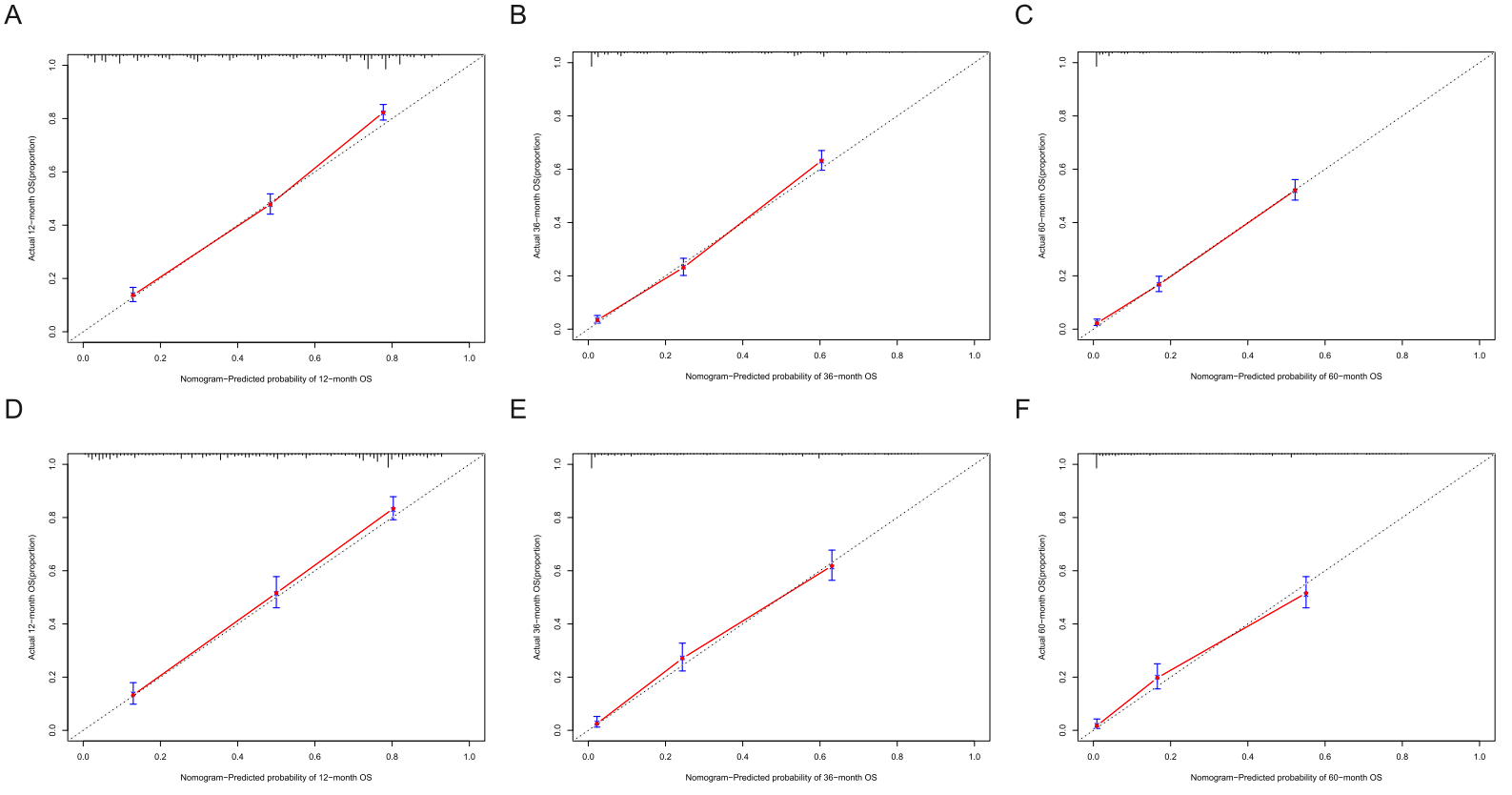

Fig. 3. Calibration curves for OS by using traindata (A, B, C) and testdata (D, E, F).

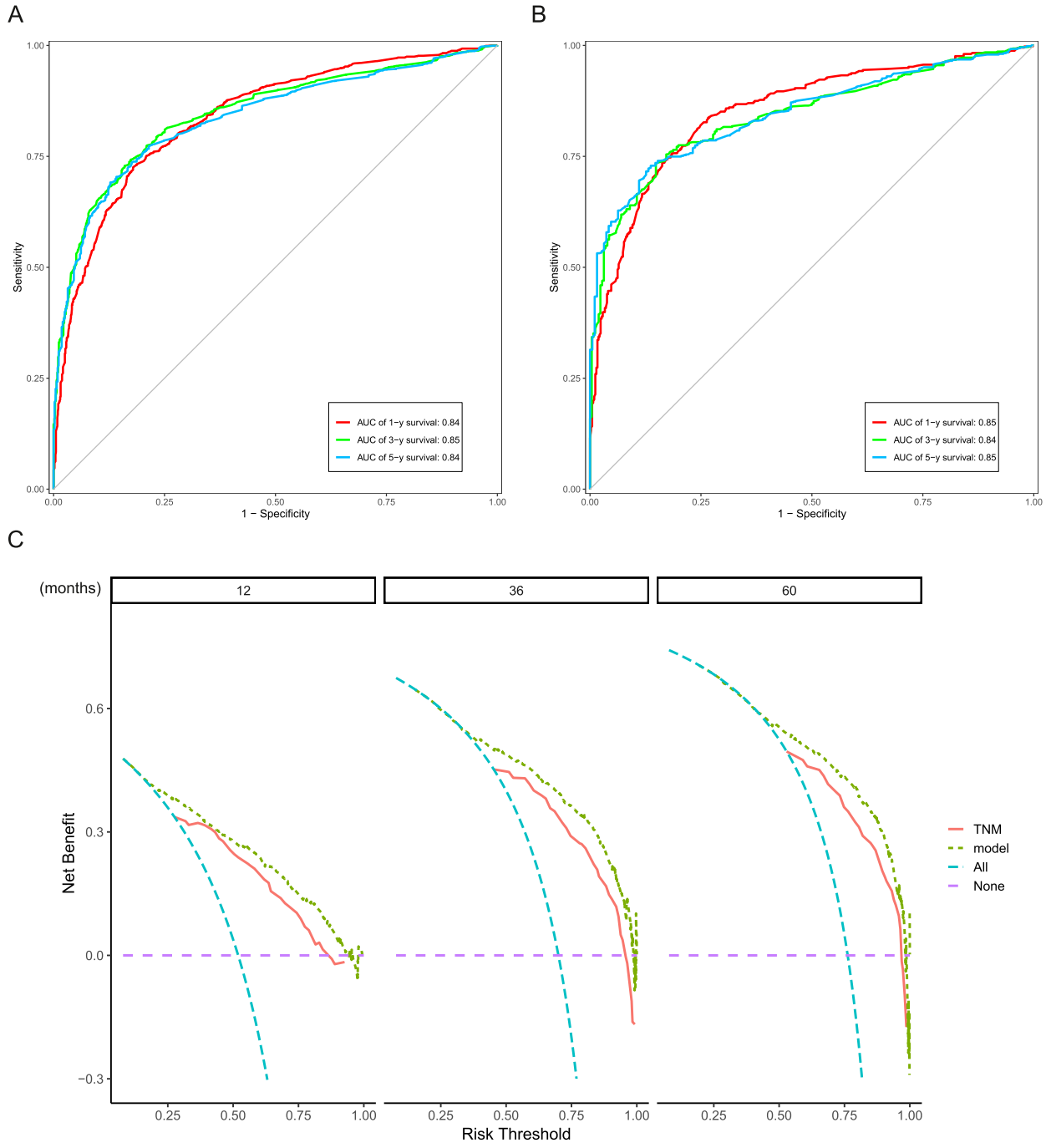

Fig. 4. Time-ROC of 1-, 3-, 5-year AUC for OS by using traindata(A) and testdata(B), Decision curve analysis of the current nomogram model and the 6th edition AJCC TNM staging system for predicting 12,36 and 60 months overall survival probability(C).

better than the TNM clinical stage model. It is hoped that this model will provide support to oncologists and pathologists in formulating clinical strategies.

5. Conclusions

Based on an analysis of the United States population-based cohort in the SEER database, clinical characteristics and prognosis in LCLC patients were summarized. Moreover, a visual nomogram was developed for predicting the probability of outcome by taking

multivariate factors ($P < 0.01$) into account and into consideration. The differences in the nomogram were very small and completely acceptable between the training dataset and test dataset. This nomogram provides more accurate OS assessments for LCLC patients and gives assistance to clinicians to make more beneficial personal management decisions.

Author contribution statement

Hongxia Zhou, Pengxiang Gao:Performed the experiments; Analyzed and interpreted the data; Wrote the paper.
Fangpeng Liu, Liangliang Shi, Longhua Sun:Contributed analysis tools or data; Analyzed and interpreted the data.
Wei Zhang, Xinping Xu, Xiujuan Liu:Conceived and designed the experiments.

Availability of data and materials

Study data was publicly available in the SEER database (https://seer.cancer.gov/).

Funding

This work was supported by Science and Technology Plan of Jiangxi Provincial Health Commission (SKJP220203329), the National Natural Science Foundation Of Jiangxi Province (20202BAB206043), National Natural Science Foundation of China (81860140).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e15924.

References

[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA A Cancer J. Clin. 71 (3) (2021) 209–249, https://doi.org/10.3322/caac.21660.
[2] R. Liang, T.-X. Chen, Z.-Q. Wang, K.-W. Jin, L.-Y. Zhang, Q.-N. Yan, .W.-P. Wang, A retrospective analysis of the clinicopathological characteristics of large cell carcinoma of the lung, Exp. Ther. Med. 9 (1) (2015) 197–202.
[3] G. Pelosi, M. Barbareschi, A. Cavazza, P. Graziano, G. Rossi, M. Papotti, Large cell carcinoma of the lung: a tumor in search of an author. A clinically oriented critical reappraisal, Lung Cancer 87 (3) (2015) 226–231, https://doi.org/10.1016/j.lungcan.2015.01.008.
[4] W.D. Travis, E. Brambilla, A.G. Nicholson, Y. Yatabe, J.H.M. Austin, M.B. Beasley, W.H.O. Panel, The 2015 world Health organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification, J. Thorac. Oncol. 10 (9) (2015) 1243–1260, https://doi.org/10.1097/JTO.0000000000000630.
[5] Z. Zuo, G. Zhang, P. Song, J. Yang, S. Li, Z. Zhong, J. He, Survival nomogram for stage IB non-small-cell lung cancer patients, based on the SEER database and an external validation cohort, Ann. Surg Oncol. 28 (7) (2021) 3941–3950, https://doi.org/10.1245/s10434-020-09362-0.
[6] B. Jia, Q. Zheng, J. Wang, H. Sun, J. Zhao, M. Wu, Z. Wang, A nomogram model to predict death rate among non-small cell lung cancer (NSCLC) patients with surgery in surveillance, epidemiology, and end results (SEER) database, BMC Cancer 20 (1) (2020) 666, https://doi.org/10.1186/s12885-020-07147-y.
[7] Y. Shi, W. Chen, C. Li, S. Qi, X. Zhou, Y. Zhang, G. Li, Clinicopathological characteristics and prediction of cancer-specific survival in large cell lung cancer: a population-based study, J. Thorac. Dis. 12 (5) (2020) 2261–2269, https://doi.org/10.21037/jtd.2020.04.24.
[8] G. Lin, K. Qi, B. Liu, H. Liu, J. Li, A nomogram prognostic model for large cell lung cancer: analysis from the Surveillance, Epidemiology and End Results Database, Transl. Lung Cancer Res. 10 (2) (2021) 622–635, https://doi.org/10.21037/tlcr-19-517b.
[9] L. Zhang, M.-C. Hsieh, L. Rennert, P. Neroda, X.-C. Wu, C. Hicks, R. Gimbel, Diagnosis-to-surgery interval and survival for different histologies of stage I-IIA lung cancer, Transl. Lung Cancer Res. 10 (7) (2021) 3043–3058, https://doi.org/10.21037/tlcr-21-168.
[10] K.M. Doll, A. Rademaker, J.A. Sosa, Practical guide to surgical data sets: surveillance, epidemiology, and end results (SEER) database, JAMA Surg. 153 (6) (2018) 588–589, https://doi.org/10.1001/jamasurg.2018.0501.
[11] R.L. Camp, M. Dolled-Filhart, D.L. Rimm, X-tile, A new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization, Clin. Cancer Res. : Off. J. Am. Assoc. Canc. Res. 10 (21) (2004) 7252–7259.
[12] A.J. Vickers, E.B. Elkin, Decision curve analysis: a novel method for evaluating prediction models, Med. Decis. Making : Int. J. Soci. Med. Dec. Making 26 (6) (2006) 565–574.
[13] B. Van Calster, L. Wynants, J.F.M. Verbeek, J.Y. Verbakel, E. Christodoulou, A.J. Vickers, E.W. Steyerberg, Reporting and interpreting decision curve analysis: a guide for investigators, Eur. Urol. 74 (6) (2018) 796–804, https://doi.org/10.1016/j.eururo.2018.08.038.
[14] L. Yang, J. Yu, S. Zhang, Y. Shan, Y. Li, L. Xu, J. Zhang, A prognostic model of patients with ovarian mucinous adenocarcinoma: a population-based analysis, J. Ovarian Res. 15 (1) (2022) 26, https://doi.org/10.1186/s13048-022-00958-6.
[15] C. Liu, C. Hu, J. Huang, K. Xiang, Z. Li, J. Qu, T. Wen, A prognostic nomogram of colon cancer with liver metastasis: a study of the US SEER database and a Chinese cohort, Front. Oncol. 11 (2021), 591009, https://doi.org/10.3389/fonc.2021.591009.

[16] X. Zhang, X. Shi, H. Zhao, X. Jia, Y. Yang, Identification and validation of a tumor microenvironment-related gene signature for prognostic prediction in advanced-stage non-small-cell lung cancer, BioMed Res. Int. 2021 (2021), 8864436, https://doi.org/10.1155/2021/8864436.
[17] J. Yu, W. Hu, N. Yao, M. Sun, X. Li, L. Wang, B. Li, Development and validation of a nomogram to predict overall survival of T1 esophageal squamous cell carcinoma patients with lymph node metastasis, Trans. Oncol. 14 (8) (2021), 101127, https://doi.org/10.1016/j.tranon.2021.101127.
[18] L.L. Wu, C.W. Li, W.K. Lin, L.H. Qiu, D. Xie, Incidence and survival analyses for occult lung cancer between 2004 and 2015: a population-based study, BMC Cancer 21 (1) (2021) 1009, https://doi.org/10.1186/s12885-021-08741-4.
[19] N. Rekhtman, W.D. Travis, Large No more: the journey of pulmonary large cell carcinoma from common to rare entity, J. Thorac. Oncol. 14 (7) (2019) 1125–1127, https://doi.org/10.1016/j.jtho.2019.04.014.
[20] P. Horak, C. Heining, S. Kreutzfeldt, B. Hutter, A. Mock, J. Hüllein, S. Fröhling, Comprehensive genomic and transcriptomic analysis for guiding therapeutic decisions in patients with rare cancers, Cancer Discov. 11 (11) (2021) 2780–2795, https://doi.org/10.1158/2159-8290.CD-21-0126.
[21] D.M. Finkelstein, N.K. Horick, R. Ramchandani, K.L. Boyd, H.Q. Rana, B.L. Bychkovsky, Are rare cancer survivors at elevated risk of subsequent new cancers? BMC Cancer 19 (1) (2019) 166, https://doi.org/10.1186/s12885-019-5358-1.
[22] Z. Hui, Y. Men, C. Hu, J. Kang, X. Sun, N. Bi, J. He, Effect of postoperative radiotherapy for patients with pIIIA-N2 non-small cell lung cancer after complete resection and adjuvant chemotherapy: the phase 3 PORT-C randomized clinical trial, JAMA Oncol. 7 (8) (2021) 1178–1185, https://doi.org/10.1001/jamaoncol.2021.1910.
[23] D.J. Tandberg, B.C. Tong, B.G. Ackerson, C.R. Kelsey, Surgery versus stereotactic body radiation therapy for stage I non-small cell lung cancer: a comprehensive review, Cancer 124 (4) (2018) 667–678, https://doi.org/10.1002/cncr.31196.
[24] V.P. Balachandran, M. Gonen, J.J. Smith, R.P. DeMatteo, Nomograms in oncology: more than meets the eye, Lancet Oncol. 16 (4) (2015) e173–e180, https://doi.org/10.1016/S1470-2045(14)71116-7.
[25] A.J. Vickers, B. van Calster, E.W. Steyerberg, A simple, step-by-step guide to interpreting decision curve analysis, Diagn. Prog. Res. 3 (2019) 18, https://doi.org/10.1186/s41512-019-0064-7.
[26] L. Xiaochuan, Y. Jiangyong, Z. Ping, W. Xiaonan, L. Lin, Clinical characteristics and prognosis of pulmonary large cell carcinoma: a population-based retrospective study using SEER data, Thor. Canc. 11 (6) (2020) 1522–1532, https://doi.org/10.1111/1759-7714.13420.
[27] Q. Tai, L. Zhang, X. Hu, Clinical characteristics and treatments of large cell lung carcinoma: a retrospective study using SEER data, Transl. Cancer Res. 9 (3) (2020) 1455–1464, https://doi.org/10.21037/tcr.2020.01.40.
[28] J.E. Chaft, A. Rimner, W. Weder, C.G. Azzoli, M.G. Kris, T. Cascone, Evolution of systemic therapy for stages I-III non-metastatic non-small-cell lung cancer, Nat. Rev. Clin. Oncol. 18 (9) (2021) 547–557, https://doi.org/10.1038/s41571-021-00501-4.
[29] A.K. Ganti, A.B. Klein, I. Cotarla, B. Seal, E. Chou, Update of incidence, prevalence, survival, and initial treatment in patients with non-small cell lung cancer in the US, JAMA Oncol. 7 (12) (2021) 1824–1832, https://doi.org/10.1001/jamaoncol.2021.4932.
[30] A.G. Nicholson, M.S. Tsao, M.B. Beasley, A.C. Borczuk, E. Brambilla, W.A. Cooper, W.D. Travis, The 2021 WHO classification of lung tumors: impact of advances since 2015, J. Thorac. Oncol. 17 (3) (2022) 362–387, https://doi.org/10.1016/j.jtho.2021.11.003.
[31] R. Rami-Porta, Future perspectives on the TNM staging for lung cancer, Cancers 13 (8) (2021), https://doi.org/10.3390/cancers13081940.