Nomogram for predicting the survival rate of primary pulmonary mucoepidermoid carcinoma patients: a retrospective study based on SEER database

Guo Lin[1,2#], Hengrui Liang[1#], Wei Wang[1], Jun Liu[1], Jianfu Li[1], Wenhua Liang[1], Jianxing He[1]

[1]Department of Thoracic Surgery and Oncology, the First Affiliated Hospital of Guangzhou Medical University, State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of Respiratory Health, Guangzhou, China; [2]Guangzhou Medical University, Guangzhou, China

Contributions: (I) Conception and design: H Liang, W Liang, J He; (II) Administrative support: J He, W Wang, W Liang; (III) Provision of study materials or patients: G Lin; (IV) Collection and assembly of data: G Lin; (V) Data analysis and interpretation: G Lin; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Jianxing He; Wenhua Liang. Department of Thoracic Surgery and Oncology, the First Affiliated Hospital of Guangzhou Medical University, State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou 510120, China. Email: drjianxing.he@gmail.com; liangwh1987@163.com.

Background: Primary pulmonary mucoepidermoid carcinoma (PMEC) is a rare malignant tumor, and the clinical manifestations lack specificity. The study evaluates the prognostic factors and constructs a practicable nomogram to estimate the individualized survival status for PMEC patients.

Methods: Surveillance, Epidemiology, and End Results (SEER) database was used to selected eligible patients between 1975 and 2016. The baseline characteristics including age, sex, race, marital status, tumor stage, differentiated degree, tumor laterality, primary tumor site, tumor size, lymph node metastases status, distant metastases status, surgery, chemotherapy, and radiation. We identified independent variables to build 3-, 5-, 10-year overall survival (OS) and cancer-specific survival (CSS) nomograms by univariate and multivariate analyses.

Results: A total of 438 PMEC patients met our selection criteria. In multivariate analysis, age, tumor stage, differentiated grade, tumor size, lymph node metastases status, distant metastases status, surgery and radiation were involved in the nomogram. The C-index (0.887 (95% CI: 0.863–0.911), calibrate plots and ROC curves (AUC =0.941, 0.951, 0.935 for 3-, 5-, 10-year OS, respectively) indicated the satisfied accuracy and practicability of our nomograms. Compared to TNM system, our model also showed a superior prediction (IDI =0.167, 0.171, 0.172, P<0.001).

Conclusions: We built OS (CSS) nomograms that can accurately estimate individualized survival time and identify the risk classification of PMEC.

Keywords: Pulmonary mucoepidermoid carcinoma (PMEC); Overall survival; Cancer-specific survival (CSS); Nomogram; Surveillance, Epidemiology, and End Results (SEER)

Introduction

Mucoepidermoid carcinoma is the most common malignant salivary gland tumor in children and adults (1). Most mucoepidermoid carcinoma identified in the head and neck can also occur in other sites, such as lung, skin, breast, liver, and so forth (1-3). Pulmonary mucoepidermoid carcinoma (PMEC) is rare, which originates from the submucosal glands of the tracheobronchial tree (4,5). It was first defined in 1952 by Smetana *et al.* (6) and accounted for only 0.1–0.2% of all primary pulmonary malignant tumors (7).

PMEC is composed of mucinous cells, intermediate cells, and squamous cells arranged in solid, cystic, or glandular shape (8). According to the degree of cell differentiation, atypia, and mitotic figures, PMEC was classified as low and high-grade (9). The majority of high-grade PMEC is the solid squamous cell with strong invasiveness, presenting with regional lymph node metastasis, and distant metastasis (10). Low-grade PMEC is mostly cystic components, with a favorable prognosis. However, due to the lack of specificity in PMEC patients' clinical manifestations, the grade classification and diagnosis of PMEC depends on bronchoscopy and lung biopsy (11,12). In addition, the low incidence of PMEC and limited data makes it challenging to characterize epidemiologically. Furthermore, the traditional AJCC TNM staging system is backward owing to its timeliness and lack of individuality.

Nomogram based on cox regression is a graphical calculating tool that involved patients' characteristics in estimating the personalized survival time and providing individualized risk prediction (13). Compared to other estimation models, nomogram has superior accuracy, utility, and risk classification (14). Based on specific samples, it can more effectively estimate patients' survival rate with individualized baseline characteristics, applied in the prognosis estimation of malignant tumors.

Whereas there is no study to construct a prediction model to forecast the individual survival time of primary PMEC patients. Our study addressed this gap, attempting to construct and validate authentic nomograms to estimate the survival outcome of PMEC patients. We present the following article in accordance with the TRIPOD reporting checklist (available at http://dx.doi.org/10.21037/atm-20-6555).

Methods

Data source

(I) This retrospective study was performed based on the SEER database (http://www.seer.cancer.gov), provided by the National Cancer Institute, which summarized 18 registries of the US from 1975–2016 and covered more than 30% of the US population. SEER*Stat 8.3.6 software was used to access information from the database. We identified the study population from the SEER cancer registry followed: (I) Diagnosis of MEC (ICD-O-3, Hist/behave: 8430/3 and primary sites: C34.1, C34.2, C34.3, C34.8, C34.9);

(II)   Only primary cancer;

(III)   Age of diagnosis ≥18 years;

(IV)   Active follow-up.

Those who had incomplete characteristics or only with a death certificate or autopsy were excluded.

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Data extraction

We recorded eligible patients' demographics, clinicopathological features, and follow-up information, covering age at diagnosis, sex, race, marital status, tumor stage, differentiated grade, tumor laterality, primary tumor site, tumor size (T, the largest diameter of the primary tumor), lymph node metastases status (N), tumor distant metastases status (M), surgery, radiation and chemotherapy. These continuous variables, including age and tumor size were translated into categorical variables using the X-tile 3.6.1 program (Yale University, New Haven, CT), according to the minimum P value of the survival time. The age was divided into ≤44 years old, 44–69 years old, and ≥69 years old.

Owing to the traditional AJCC TNM system had multiple versions in the SEER database and did not apply to all patients, we created a customized TNM system. Collecting tumor size, lymph metastases status, and tumor distant metastases status from different periods, and divided into N0/N+, M0/M1. Tumor size was divided into two groups: ≤2.7 centimeters, 2.7–4.8 centimeters, and ≥4.8 centimeters by X-tile. Survival analysis indicated no survival difference was observed among N1, N2, and N3 (P=0.067). Hence, we redefined the positive lymph node metastasis (N1, N2, or N3) as one group named N+.

Statistical analysis

OS was the primary outcome and cancer-specific survival (CSS) as the secondary outcome in this study. CSS was

defined as the time from diagnosis to death caused by MEC; alive or death from other diseases was considered as censored cases. We identified 438 eligible cases in this study. In order to construct and validate the nomograms, we divided these eligible cases into training cohort (n=308, 70%) and validation cohort (n=130, 30%) randomly by 7:3. All categorical variables of baseline characteristics (age, sex, race, marital status, tumor stage, differentiated grade, tumor laterality, primary tumor site, T, N, M, surgery, radiation, and chemotherapy) were compared using the χ^2 test. In the training cohort, univariate and multivariate analyses were used to identify the independent prognostic variables of nomogram construction used to build a risk classification system-Only variables with significant (P<0.05) in univariate analyses can be retained into multivariate analyses. Considering the small sample size of CSS, we used OS independent factors to build a nomogram.

Validation cohorts were devoted to examine the accuracy of the nomograms. Internal validation was evaluated under bootstrapping validation with 1,000 resamples. The calibration curve and concordance index (C-index) were applied to reflect the conformance between predicting survival and reality survival. All patients in the training cohort were divided into two groups according to their risk scores for examining the nomograms' discriminability. Scores above the average are considered a high-risk group, while below-average account for the low-risk group. The survival curves of risk groups were evaluated by the log-rank test and shown by the Kaplan-Meier plot. IDI was used to compare our nomogram with other models.

All statistical analyses were conducted using SPSS version 26.0 (SPSS Chicago, IL) and rms, foreign, survival, and survival ROC packages in R version 3.6.1 software (R Foundation for Statistical Computing, Vienna, Austria, http://www.r-project.org/). All tests with two sides P value less than 0.05 was considered a significant difference.

Results

Characteristics

A total of 438 patients with primary PMEC were reviewed in this study. According to age, all patients were divided to ≤44 years old (n=185, 42.2%), 44–69 years old (n=177, 40.4%) and ≥69 years old (n=76, 17.4%). The whole sample had the similar distribution of gender and laterality: males (n=234, 53.4%) and females (n=204, 46.6%), left (n=211,

48.2%) and right (n=227, 51.8%). The majority of race was white (n=338, 77.2%).

For differentiation degree, tumors with well, moderately, poorly, and undifferentiated accounted for 23.5%, 34.7%, 9.6%, and 8.7% of all patients, respectively. Of all cases, more than half were negative lymph node metastases (n=251, 57.3%) and negative distant metastases (n=289, 66.0%). Moreover, most of patients received surgery (n=339, 77.4%) and did not accept chemotherapy (n=377, 86.1%). Only 20.1% of patients underwent radiation. There is no significant discrepancy between the training cohort and validation cohort. The detailed baseline characteristics were shown in *Table 1*.

Independent prognostic factors of nomogram construction

Prognostic factors in predicting OS
Univariate and multivariate Cox proportional hazards regression models were conducted to identify prognostic factors. As shown in *Table 2*, in univariate analyses, we noticed except for sex (P=0.364), race (P=0.609), marital status (P=0.385), and tumor laterality (P=0.715), the other variables were shown a significant association with OS. Multivariate analyses suggested that age, tumor stage, differentiated grade, T, N, M, surgery, and radiation as the independent prognostic factors for OS.

Prognostic factors in predicting CSS
Of 175 cases. 136 patients died of MEC, while the other 39 patients were ascribed to non-MEC causes. As shown in Table S1, tumor stage, T, M, surgery, and radiation were analyzed as the prognostic factors predicting CSS of MEC patients.

Construction of nomograms

We constructed two nomograms based on independent prognostic factors identified by Cox proportional hazards analyses, including age, tumor stage, differentiated grade, T, N, M, surgery, and radiation. The eight variables were enrolled in nomograms building for 3-, 5- and 10-year OS and CSS of MEC patients in this study. Each variable had a corresponding score on the point scale, locating and summing up each score according to the baseline characteristic; the total scores were obtained. The total scores represent the probabilities of 3-, 5- and 10-year OS and CSS. Particulars score of each variable in the nomograms were shown in *Figures 1,2*.

Table 1 Demographic/clinical characteristics of PMEC patients

Demographic/clinical characteristic	N (%)	Training cohort (N)	Validation cohort (N)	P value
All	438 (100.0)	308	130	
Age				0.092
≤44 years	185 (42.2)	126	59	
44–69 years	177 (40.4)	134	43	
≥69 years	76 (17.4)	48	28	
Sex				0.351
Male	234 (53.4)	169	65	
Female	204 (46.6)	139	65	
Race				0.068
White	338 (77.2)	245	93	
Other	100 (22.8)	63	37	
Marital status				0.803
Single	192 (43.8)	132	60	
Married	233 (53.2)	167	66	
Unknown	13 (3.0)	9	4	
Stage				0.773
Localized	194 (44.3)	134	60	
Regional	87 (19.9)	63	24	
Distant	50 (11.4)	33	17	
Unknown	107 (24.4)	78	29	
Grade				0.947
Well	103 (23.5)	70	33	
Moderately	152 (34.7)	109	43	
Poorly	42 (9.6)	31	11	
Undifferentiated	38 (8.7)	26	12	
Unknown	103 (23.5)	72	31	
Laterality				0.938
Left	211 (48.2)	148	63	
Right	227 (51.8)	160	67	
Primary site				0.221
Main bronchus	61 (13.9)	38	23	
Upper lobe, lung	147 (33.6)	105	42	
Middle lobe, lung	41 (9.4)	30	11	
Lower lobe, lung	139 (31.7)	101	38	
Overlapping lesion of lung	18 (4.1)	9	9	
Unknown	32 (7.3)	25	7	

Table 1 (*continued*)

Table 1 (*continued*)

Demographic/clinical characteristic	N (%)	Training cohort (N)	Validation cohort (N)	P value
T				0.853
≤2.7 cm	170 (38.8)	121	49	
2.7–4.8 cm	86 (19.6)	62	24	
≥4.8 cm	50 (11.4)	36	14	
Unknown	132 (30.1)	89	43	
N				0.496
N0	251 (57.3)	174	77	
N+	64 (14.6)	49	15	
Unknown	123 (28.1)	85	38	
M				0.907
M0	289 (66.0)	202	87	
M1	42 (9.6)	29	13	
Unknown	107 (24.4)	77	30	
Surgery				0.924
No	99 (22.6)	70	29	
Yes	339 (77.4)	238	101	
Radiation				0.623
No	350 (79.9)	248	102	
Yes	88 (20.1)	60	28	
Chemotherapy				0.382
No	377 (86.1)	268	109	
Yes	61 (13.9)	40	21	

PMEC, pulmonary mucoepidermoid carcinoma.

Table 2 Univariate and multivariate analyses of overall survival

Variables	Univariate analysis		Multivariate analysis	
	HR (95% CI)	P value	HR (95% CI)	P value
Age		<0.001		<0.001
≤44 years	Ref		Ref	
44–69 years	7.72 (4.74, 12.59)	<0.001	4.97 (2.99, 8.26)	<0.001
≥69 years	13.74 (8.15, 23.16)	<0.001	8.41 (4.82, 14.65)	<0.001
Sex		0.364	NI	
Male	Ref			
Female	0.87 (0.65, 1.17)	0.364		

Table 2 (*continued*)

Table 2 (*continued*)

Variables	Univariate analysis		Multivariate analysis	
	HR (95% CI)	P value	HR (95% CI)	P value
Race		0.609	NI	
White	Ref			
Other	1.10 (0.77, 1.57)	0.609		
Marital status		0.385	NI	
Single	Ref			
Married	1.14 (0.84, 1.55)	0.385		
Stage		<0.001		<0.001
Localized	0.14 (0.09, 0.24)	<0.001	0.25 (0.14, 0.46)	<0.001
Regional	Ref		Ref	
Distant	4.12 (2.67, 6.36)	<0.001	1.47 (0.76, 2.86)	0.255
Grade		<0.001		<0.001
Well	0.06 (0.03, 0.12)	<0.001	0.22 (0.11, 0.46)	<0.001
Moderately	0.09 (0.05, 0.15)	<0.001	0.32 (0.17, 0.60)	<0.001
Poorly	Ref		Ref	
Undifferentiated	0.58 (0.34, 0.98)	0.041	0.93 (0.51, 1.72)	0.824
Laterality		0.715	NI	
Left	Ref			
Right	1.06 (0.79, 1.42)	0.715		
Primary Site		0.002		0.065
Main Bronchus	Ref		Ref	
Upper lobe, lung	1.65 (0.99, 2.78)	0.057	1.27 (0.71, 2.26)	0.419
Middle lobe, lung	0.74 (0.35, 1.57)	0.437	0.96 (0.43, 2.13)	0.922
Lower lobe, lung	0.85 (0.49, 1.48)	0.555	0.78 (0.43, 1.43)	0.432
Overlapping lesion of lung	1.89 (0.87, 4.10)	0.106	2.00 (0.89, 4.53)	0.095
T		<0.001		0.003
≤2.7 cm	Ref		Ref	
2.7–4.8 cm	2.35 (1.34, 4.11)	0.003	2.06 (1.13, 3.77)	0.018
≥4.8 cm	8.19 (4.72, 14.22)	<0.001	3.36 (1.65, 6.87)	0.001
N		<0.001		<0.001
N0	Ref		Ref	
N+	0.08 (0.05, 0.13)	<0.001	3.54 (1.94, 6.44)	<0.001
M		<0.001		<0.001
M0	Ref		Ref	
M1	0.08 (0.05, 0.12)	<0.001	9.87 (4.06, 24.00)	<0.001

Table 2 (*continued*)

Table 2 (*continued*)

Variables	Univariate analysis		Multivariate analysis	
	HR (95% CI)	P value	HR (95% CI)	P value
Surgery		<0.001		<0.001
No	Ref		Ref	
Yes	0.10 (0.07, 0.14)	<0.001	0.34 (0.22, 0.52)	<0.001
Radiation		<0.001		0.001
No	Ref		Ref	
Yes	6.53 (4.74, 8.99)	<0.001	2.00 (1.36, 2.93)	0.001
Chemotherapy		<0.001		0.471
No	Ref		Ref	
Yes	4.25 (3.01, 5.60)	<0.001	1.16 (0.78, 1.72)	0.471
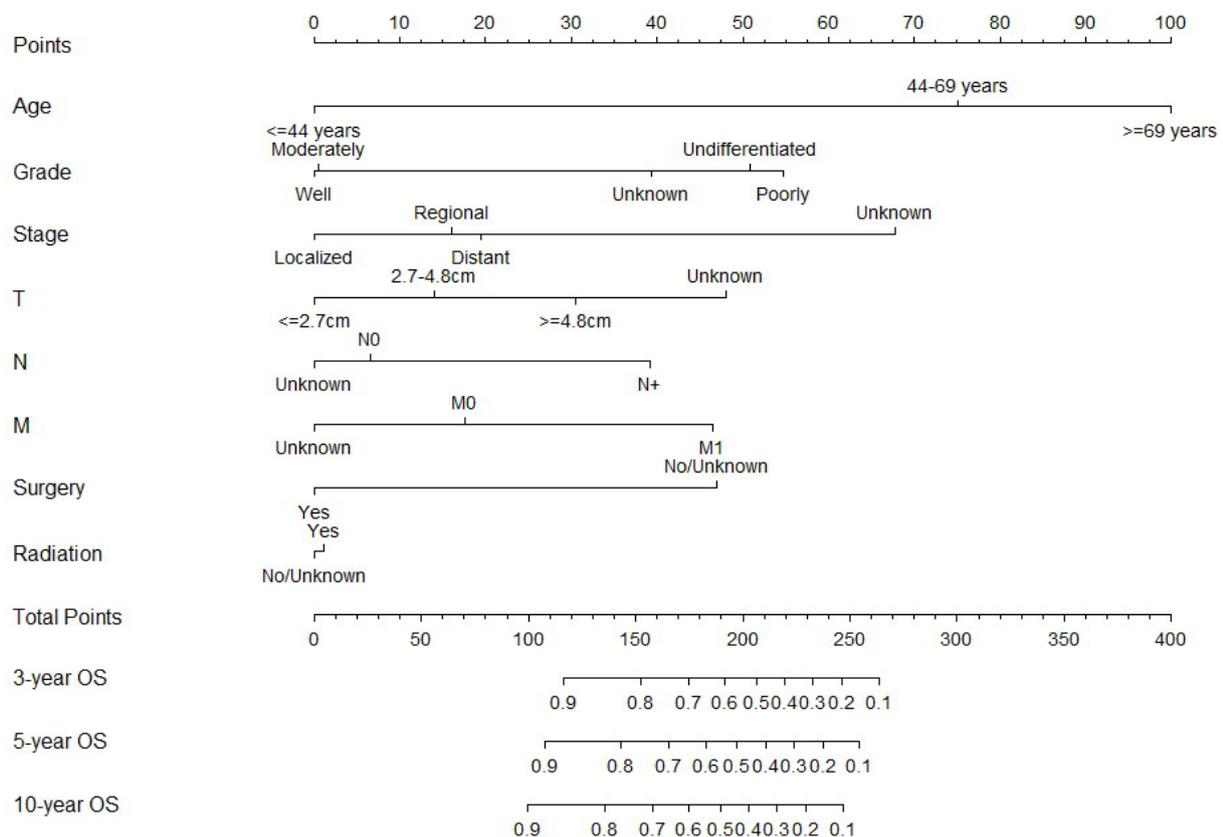
Figure 1 Prediction nomogram involved independent variables for 3-, 5-, 10-year overall survival in PMEC patients. OS, overall survival; PMEC, pulmonary mucoepidermoid carcinoma.
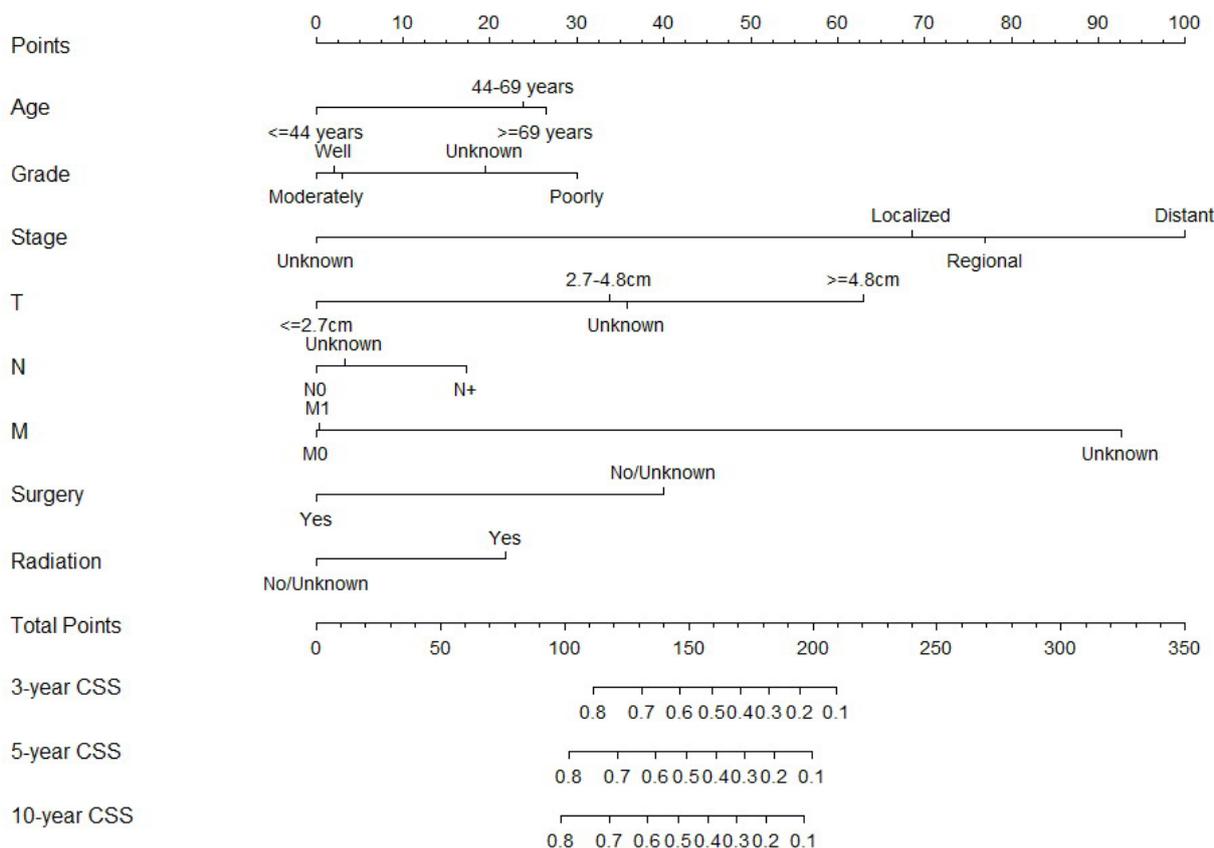
Figure 2 Prediction nomogram involved independent variables for 3-, 5-, 10-year cancer-specific survival in PMEC patients. PMEC, pulmonary mucoepidermoid carcinoma; CSS, cancer-specific survival.

Calibration and validation of nomogram

OS and CSS nomograms had the acceptance C-index in training cohort, 0.887 (95% CI: 0.863–0.911) and 0.764 (95% CI: 0.713–0.815), respectively. In internal validation, the nomograms of OS (*Figure 3*) and CSS (*Figure S1*) had considerable C-index values, 0.912 (95% CI: 0.883–0.941) and 0.794 (95% CI: 0.727–0.861), respectively. The calibration curves of 3-, 5- and 10-year OS (*Figures 4,5*) and CSS (*Figures S2,S3*) indicated considerable coherence between estimation and actual in the training and validation cohort.

Compared to the TNM system, our nomograms have better accuracy in estimating survival time in the training cohort. The IDI (*Table 3*) of our nomogram comparing with the TNM system is 0.167 for 3-year OS (P<0.001). For 5-, 10-year OS and 3-,5-, 10-year CSS, the IDI indicated that our models are superior to the TNM system significantly.

Risk stratification of OS and CSS

According to each case's total scores, we further explored risk stratification for OS and CSS based on this study's training cohort. All MEC training cohort patients were divided into a high-risk group or low-risk group. As *Figure 6* shown, both OS and CSS, the Kaplan-Meier curves, indicated that the low-risk group had a significantly better survival (P<0.001 and P<0.001, respectively).

Clinical use

Scores of each variable were obtained by drawing a vertical line to the top points row; then, the sum of the points of each variable is located on the total points row, and a vertical line is drawn from the total point row to 3-, 5-, 10-year OS or CSS row to acquire the survival probability of patients (*Figure 1*). For example, one patient was a
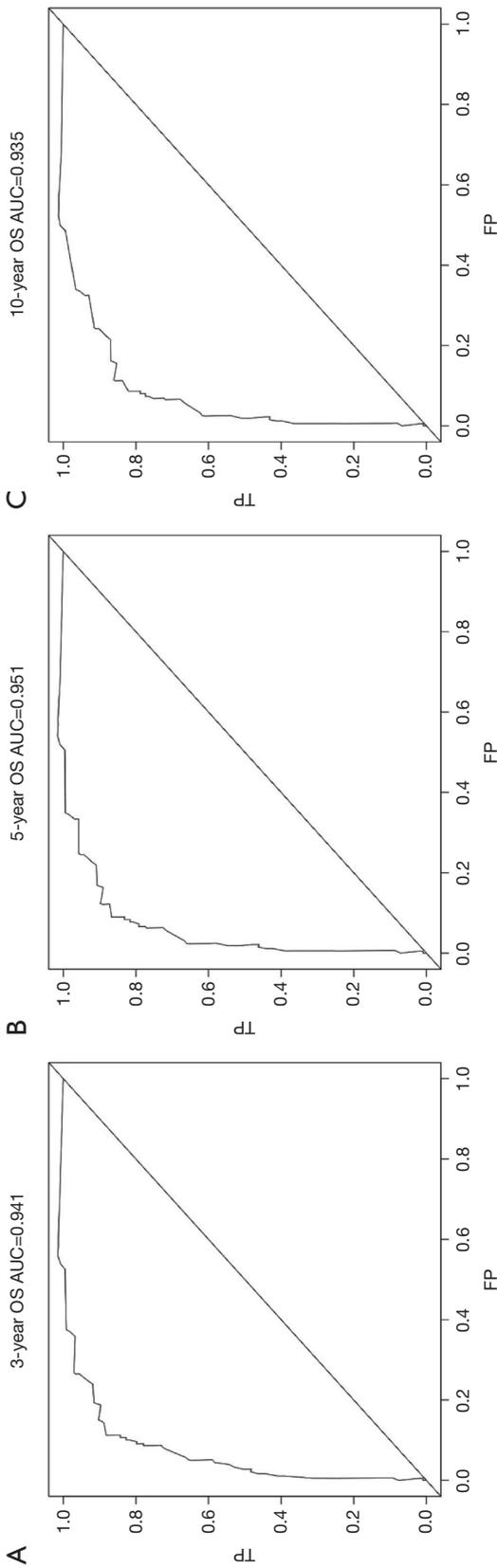
Figure 3 ROC curves for (A) 3–, (B) 5–, (C) 10-year overall survival of the nomogram.
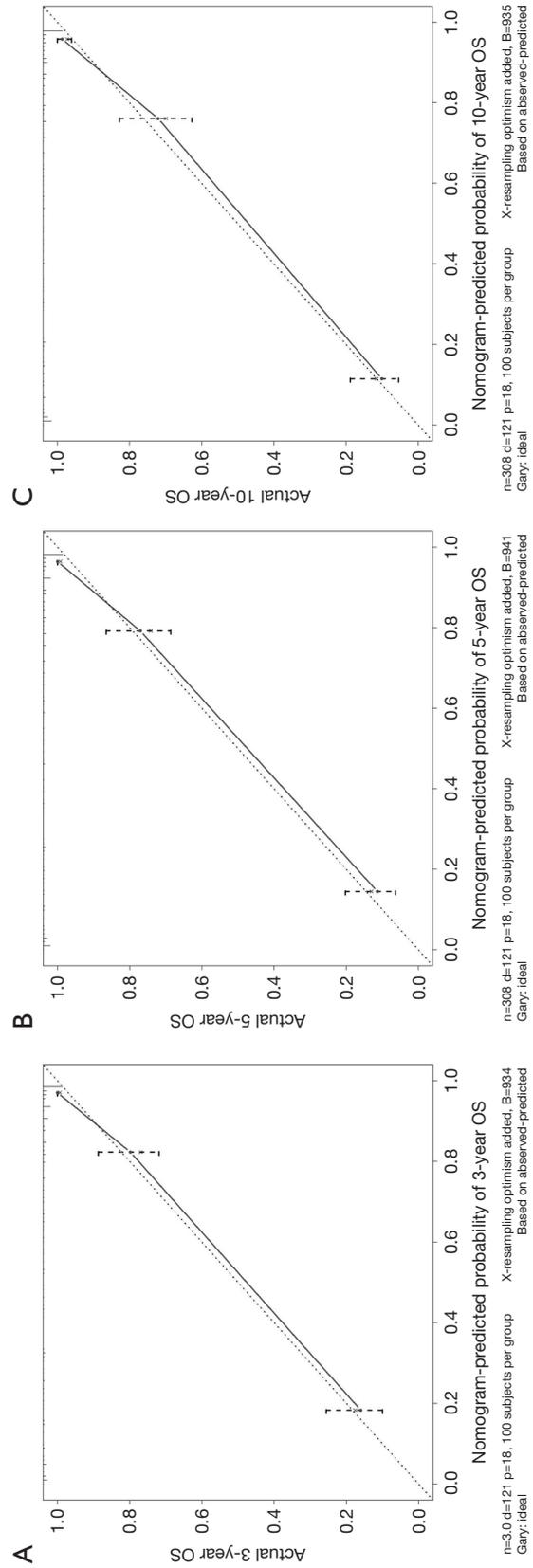

Figure 4 The calibration plots for (A) 3–, (B) 5–, (C) 10-year overall survival in training cohort.
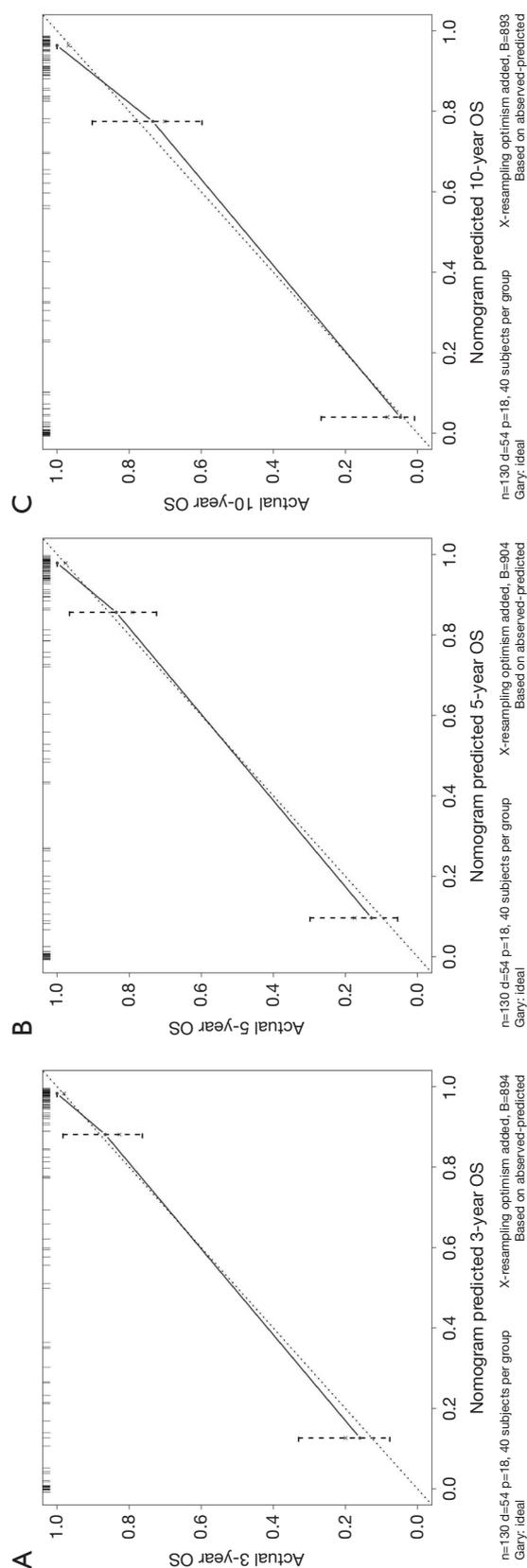
Figure 5 The calibration plots for (A) 3-, (B) 5-, (C) 10-year overall survival in validation cohort.

Table 3 The IDI of nomograms compare to TNM system

	IDI	P value
Overall survival		
3-year	0.167	<0.001
5-year	0.171	<0.001
10-year	0.172	<0.001
Cancer-specific survival		
3-year	0.119	<0.001
5-year	0.107	<0.001
10-year	0.103	<0.001

IDI, integrated discrimination index.

47-year-old, undifferentiated PMEC with 1.8 cm and N+M0 regional stage, and received surgical and radiation therapy; the total score is 200 points with corresponding 3-, 5-, 10-year OS of 0.50–0.60, 0.40–0.50, and 0.40–0.50, respectively.

Discussion

Primary mucoepidermoid carcinoma of pulmonary salivary gland cancer is a scarce histologic type, which accounts for 0.1-0.2% in all malignant lung cancers generally (15). Owing to the small sample size, precedent reports in the literature were case presentations and case series, while the study based on a large population is lack, and knowledge of PMEC patients' survival is limited. Moreover, its pathological staging still depends on the traditional TNM system based on general lung cancer. We hypothesized that one of the reasons for poor prognosis was the lack of a specific model specifically for PMEC patients. To provide new ideas for the treatment of PMEC, we established the nomograms and the risk classifications to predict individual PMEC patients' survival probabilities.

Our nomograms were validated with the internal cohort, C-index, ROC curve, and calibration plots, suggesting that our models were steady and accurate as the predictive device to estimate survival rate and select treatment options. We also established risk stratification as the supplement of nomograms. The TNM system is a generalized and straightforward tool to forecast survival and guide treatment. An individual predicting model should be developed to narrow further the difference between the predicted survival and the actual survival.

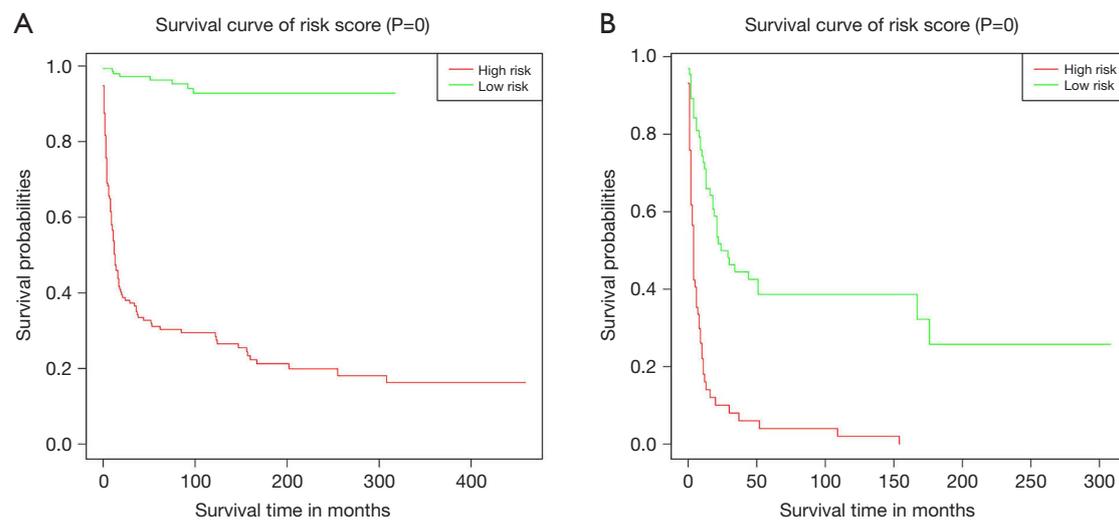Based on univariate and multivariate analyses, we

Figure 6 Kaplan-Meier curves for (A) overall survival and (B) cancer-specific survival in training cohort.

recognized these independent prognosis factors, including age, tumor stage, differentiated grade, T stage, N stage, M stage, surgery, and radiation. Most of the above factors are consistent with the previous study (16,17). Previous studies reported that young people are more likely to have PMEC, while for the past few years, the morbidity of PMEC aging (16). Like the precedent conclusion (18,19), sex was no significant relation to PMEC patients' prognosis. Our study exclusively examined PMEC samples, which is probably more representative than the precedent outcomes.

Higher degree differentiation of the PMEC enhanced the survival rate of patients. 255 (58.2%) patients are a higher differentiation degree (well and moderately) and 80 (18.3%) samples are a lower differentiation degree (poorly and undifferentiated). In both groups, there was less patients response to radiation or chemotherapy (P=0.001 and 0.471, respectively). In this study, most PMEC was moderate grade; this finding is consistent with the discovery by Chen *et al.* (20), in which 47.7% of mucoepidermoid carcinoma were moderate grade.

Complete surgical resection is essential for PMEC patients, and it will guide postoperative chemoradiotherapy according to pathological biopsy (21). Besides, bronchoplasty procedures are efficient methods for preserving pulmonary parenchyma. Several published studies regarded that surgical resection can prolong survival time (18,19,22). Endobronchial intervention can maintain airway patency, commonly applied to obstructive pneumonia patients (23,24). However, endobronchial intervention therapy does not remove the tumor entirely;

surgical resection is still preferred (25).

Nevertheless, high-grade tumors are challenging to resection entirely owing to aggressiveness; these patients require postoperative chemoradiotherapy (26). For these patients, adjuvant therapy should be performed when surgical resection is not radical, including chemotherapy and radiation (27). The precedent report considered that adjuvant chemotherapy is useful for high-grade PMEC but not presented to low-grade; the combination of cisplatin and vinorelbine or docetaxel is generally used. However, the benefit of adjuvant radiotherapy is still controversial (28). In this study, we noticed that patients with adjuvant therapy had poor prognosis than those without chemotherapy and (or) radiation. The reasonable explanation is that traditional adjuvant therapy is not conducive to reconstructing immune function, and the painful physical and mental also reduces the survival time. Besides, some patients can not tolerate surgery or other reasons; they received chemotherapy and (or) radiation instead of surgical in our study samples. Although our nomograms are precisely models, there are also many shortcomings to noticed.

(I) the data come from the SEER database, and this database can not provide complete variables, for instance, smoking status, surgery procedure, chemotherapy regimens.

(II) The AJCC 8th edition system can not apply to previous patients; we transformed the 432 PMEC patients' TNM status prudently, and some errors are inevitable.

(III) We failed to obtain sufficient cases as an external

validation cohort to further demonstrate our nomograms' consistency.

(IV) The data used in our study were retrospective; there is information bias. Further studies need to be performed to improve these deficiencies.

(V) The small sample size may still affect the stability of this data analysis, though we as far as possible to improve the analysis process.

Conclusions

In conclusion, we built OS (CSS) nomograms that can accurately estimate individualized survival time and identify the risk classification of PMEC. Calibration plots and ROC curves demonstrated the clinical utility of our nomograms.

Acknowledgments

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at http://dx.doi.org/10.21037/atm-20-6555

Peer Review File: Available at http://dx.doi.org/10.21037/atm-20-6555

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/atm-20-6555). JH serves as an unpaid Editors-in-Chief of *Annals of Translational Medicine*. The other authors have no conflicts of interest to declare.

Ethical statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the

original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Goode RK, Auclair PL, Ellis GL. Mucoepidermoid carcinoma of the major salivary glands: clinical and histopathologic analysis of 234 cases with evaluation of grading criteria. Cancer 1998;82:1217-24.

2. Fonseca FP, Carvalho Mde V, de Almeida OP, et al. Clinicopathologic analysis of 493 cases of salivary gland tumors in a Southern Brazilian population. Oral Surg Oral Med Oral Pathol Oral Radiol 2012;114:230-9.

3. de Almeida-Pinto YD, Costa S, de Andrade BAB, et al. t(6;9)(MYB-NFIB) in head and neck adenoid cystic carcinoma: A systematic review with meta-analysis. Oral diseases 2019;25:1277-82.

4. Turnbull AD, Huvos AG, Goodner JT, et al. Mucoepidermoid tumors of bronchial glands. Cancer 1971;28:539-44.

5. Heitmiller RF, Mathisen DJ, Ferry JA, et al. Mucoepidermoid lung tumors. Ann Thorac Surg 1989;47:394-9.

6. Smetana HF, Iverson L, Swan LL. Bronchogenic carcinoma; an analysis of 100 autopsy cases. Mil Surg 1952;111:335-51.

7. Stenner M, Klussmann JP. Current update on established and novel biomarkers in salivary gland carcinoma pathology and the molecular pathways involved. European archives of oto-rhino-laryngology: official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS): affiliated with the German Society for Oto-Rhino-Laryngology - Head Neck Surg 2009;266:333-41.

8. Moran CA. Primary salivary gland-type tumors of the lung. Semin Diagn Pathol 1995;12:106-22.

9. Wang YQ, Mo YX, Li S, et al. Low-Grade and High-Grade Mucoepidermoid Carcinoma of the Lung: CT Findings and Clinical Features of 17 Cases. AJR Am J Roentgenol 2015;205:1160-6.

10. Roden AC, García JJ, Wehrs RN, et al. Histopathologic, immunophenotypic and cytogenetic features of pulmonary mucoepidermoid carcinoma. Mod Pathol 2014;27:1479-88.

11. Li X, Zhang W, Wu X, et al. Mucoepidermoid carcinoma of the lung: common findings and unusual appearances on CT. Clin Imaging 2012;36:8-13.

12. Yu Y, Song Z, Gao H, et al. EGFR L861Q mutation is a frequent feature of pulmonary mucoepidermoid carcinoma. J Cancer Res Clin Oncol 2012;138:1421-5.

13. Park SY. Nomogram: An analogue tool to deliver digital knowledge. J Thorac Cardiovasc Surg 2018;155:1793.

14. Balachandran VP, Gonen M, Smith JJ, et al. Nomograms in oncology: more than meets the eye. Lancet Oncol 2015;16:e173-80.

15. Shen C, Che G. Clinicopathological analysis of pulmonary mucoepidermoid carcinoma. World J Surg Oncol 2014;12:33.

16. Komiya T, Perez RP, Yamamoto S, et al. Primary lung mucoepidermoid carcinoma: analysis of prognostic factors using surveillance, epidemiology and end results program. Clin Respir J 2017;11:847-53.

17. Jiang L, Li P, Xiao Z, et al. Prognostic factors of primary pulmonary mucoepidermoid carcinoma: a clinical and pathological analysis of 34 cases. Int J Clin Exp Pathol 2014;7:6792-9.

18. Falk N, Weissferdt A, Kalhor N, et al. Primary Pulmonary Salivary Gland-type Tumors: A Review and Update. Adv Anat Pathol 2016;23:13-23.

19. Lee GD, Kang DK, Kim HR, et al. Surgical outcomes of pulmonary mucoepidermoid carcinoma: a review of 23 cases. Thorac Cardiovasc Surg 2014;62:140-6.

20. Chen MM, Roman SA, Sosa JA, et al. Histologic grade as prognostic indicator for mucoepidermoid carcinoma: a population-level analysis of 2400 patients. Head Neck 2014;36:158-63.

21. Kanemoto A, Oshiro Y, Sugahara S, et al. Proton beam therapy for inoperable recurrence of bronchial high-grade mucoepidermoid carcinoma. Jpn J Clin Oncol 2012;42:552-5.

22. Molina JR, Aubry MC, Lewis JE, et al. Primary salivary gland-type lung cancer: spectrum of clinical presentation, histopathologic and prognostic factors. Cancer 2007;110:2253-9.

23. Vadasz P, Egervary M. Mucoepidermoid bronchial tumors: a review of 34 operated cases. Eur J Cardiothorac Surg 2000;17:566-9.

24. Song Z, Liu Z, Wang J, et al. Primary tracheobronchial mucoepidermoid carcinoma--a retrospective study of 32 patients. World J Surg Oncol 2013;11:62.

25. Kitada M, Matsuda Y, Sato K, et al. Mucoepidermoid carcinoma of the lung: a case report. J Cardiothorac Surg 2011;6:132.

26. Sonobe S, Inoue K, Tachibana S, et al. A case of pulmonary mucoepidermoid carcinoma responding to carboplatin and paclitaxel. Jpn J Clin Oncol 2014;44:493-6.

27. Zhu F, Liu Z, Hou Y, et al. Primary salivary gland-type lung cancer: clinicopathological analysis of 88 cases from China. J Thorac Oncol 2013;8:1578-84.

28. Xi JJ, Jiang W, Lu SH, et al. Primary pulmonary mucoepidermoid carcinoma: an analysis of 21 cases. World J Surg Oncol 2012;10:232.

Table S1 Univariate and multivariate analyses of cancer-specific survival

| Variables	Univariate analysis		Multivariate analysis	
	HR (95% CI)	P value	HR (95% CI)	P value
Age		0.438	NI	
≤44 years	Ref			
44–69 years	1.30 (0.76, 2.23)	0.338		
≥69 years	1.06 (0.59, 1.90)	0.845		
Sex		0.025		0.264
Male	Ref		Ref	
Female	0.68(0.48, 0.95)	0.025	0.82 (0.57, 1.17)	0.264
Race		0.783	NI	
White	Ref			
Other	1.06 (0.72, 1.60)	0.783		
Marital status		0.423	NI	
Single	Ref			
Married	0.87 (0.62, 1.22)	0.423		
Stage		<0.001		0.041
Localized	0.19(0.07, 0.49)	0.001	0.27 (0.09, 0.81)	0.019
Regional	Ref		Ref	
Distant	2.67 (1.67, 4.28)	<0.001	1.50 (0.72, 3.13)	0.276
Grade		0.002		0.105
Well	0.34 (0.15, 0.78)	0.011	0.40 (0.16, 0.96)	0.040
Moderately	0.35 (0.19, 0.65)	0.001	0.50 (0.24, 1.06)	0.070
Poorly	Ref		Ref	
Undifferentiated	0.50 (0.27, 0.93)	0.029	0.88 (0.44, 1.77)	0.729
Laterality		0.528	NI	
Left	Ref			
Right	0.90 (0.64, 1.26)	0.528		
Primary Site		0.449	NI	
Main Bronchus	Ref			
Upper lobe, lung	1.03 (0.56, 1.88)	0.935		
Middle lobe, lung	0.70 (0.29, 1.70)	0.433		
Lower lobe, lung	0.70 (0.37, 1.36)	0.299		
Overlapping lesion of lung	0.70 (0.28, 1.75)	0.442		
T		<0.001		0.004
≤2.7 cm	Ref		Ref	
2.7–4.8 cm	1.78 (0.85, 3.73)	0.129	1.15 (0.46, 2.87)	0.765
≥4.8 cm	4.30 (2.06, 8.97)	<0.001	3.19 (1.37, 7.41)	0.007
N		<0.001		0.233
N0	Ref		Ref	
N+	2.84 (1.65, 4.90)	<0.001	1.48 (0.78, 2.79)	0.233
M		<0.001		0.039
M0	Ref		Ref	
M1	3.69 (2.32, 5.84)	<0.001	2.39 (1.05, 5.45)	0.039
Surgery		<0.001		<0.001
No	Ref		Ref	
Yes	0.25 (0.17, 0.37)	<0.001	0.36 (0.23, 0.55)	<0.001
Radiation		<0.001		0.033
No	Ref		Ref	
Yes	2.32 (1.63, 3.31)	<0.001	1.53 (1.04, 2.27)	0.033
Chemotherapy		0.002		0.237
No	Ref		Ref	
Yes	1.79 (1.24, 2.58)	0.002	1.27 (0.85, 1.90)	0.237
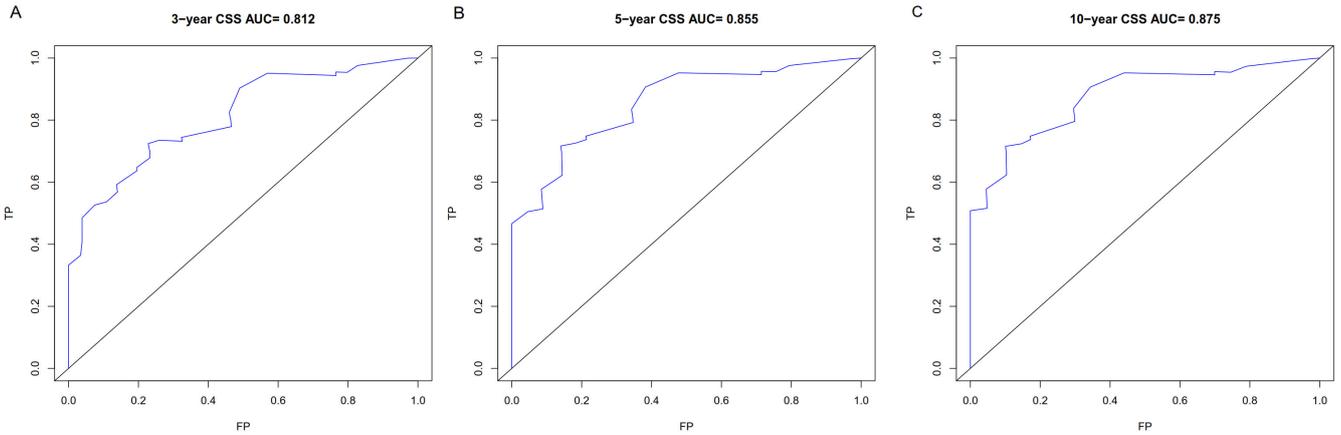
Figure S1 ROC curves for (A) 3-, (B) 5-, (C) 10-year cancer-specific survival of the nomogram.
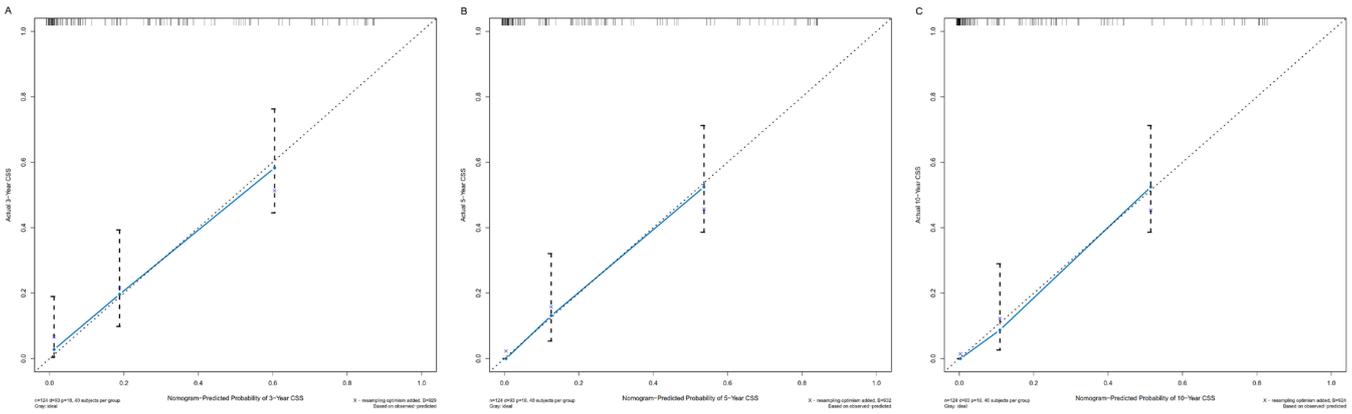

Figure S2 The calibration plots for (A) 3-, (B) 5-, (C) 10-year overall survival in training cohort.
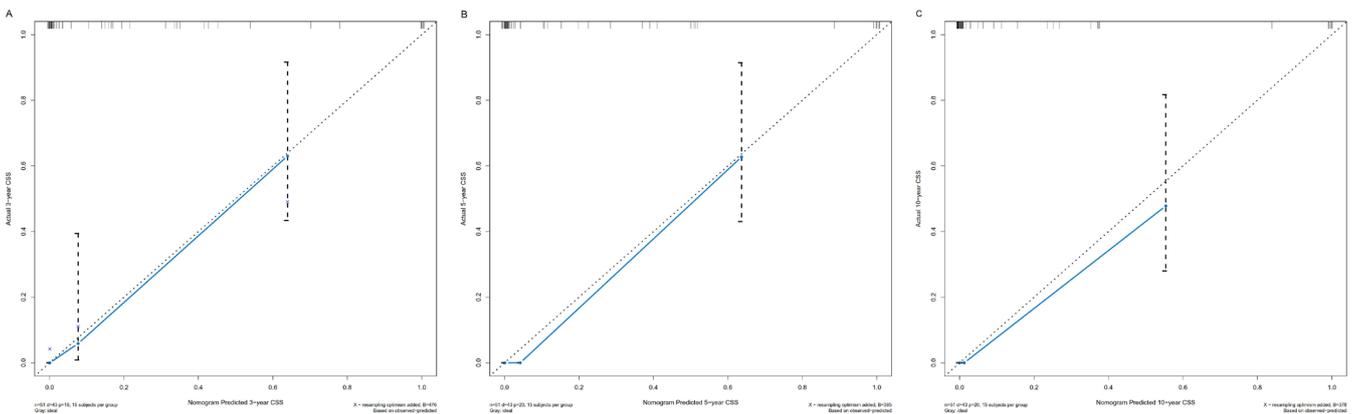

Figure S3 The calibration plots for (A) 3-, (B) 5-, (C) 10-year overall survival in validation cohort.