Medicine®

OPEN

Applying nomograms based on the surveillance, epidemiology and end results database to predict long-term overall survival and cancer-specific survival in patients with oropharyngeal squamous cell carcinomas

A case–control research

Fengze Wang, MD[a,b], Jiao Wen, BS[c], Xinjie Yang, PhD[a], Tingting Jia, PhD[d], Fangchong Du, MD[b], Jianhua Wei, PhD[a,*]

Abstract
Few models regarding to the individualized prognosis assessment of oropharyngeal squamous cell carcinoma (OPSCC) patients were documented. The purpose of this study was to establish nomogram model to predict the long-term overall survival (OS) and cancer-specific survival (CSS) of OPSCC patients. The detailed clinical data for the 10,980 OPSCC patients were collected from the surveillance, epidemiology and end results (SEER) database. Furthermore, we applied a popular and reasonable random split-sample method to divide the total 10,980 patients into 2 groups, including 9881 (90%) patients in the modeling cohort and 1099 (10%) patients in the external validation cohort. Among the modeling cohort, 3084 (31.2%) patients were deceased at the last follow-up date. Of those patients, 2188 (22.1%) patients died due to OPSCC. In addition, 896 (9.1%) patients died due to other causes. The median follow-up period was 45 months (1–119 months). We developed 2 nomograms to predict 5- and 8- year OS and CSS using Cox Proportional Hazards model. The nomograms' accuracy was evaluated through the concordance index (C-index) and calibration curves by internal and external validation. The C-indexes of internal validation on the 5- and 8-year OS and CSS were 0.742 and 0.765, respectively. Moreover, the C-indexes of external validation were 0.740 and 0.759, accordingly. Based on a retrospective cohort from the SEER database, we succeeded in constructing 2 nomograms to predict long-term OS and CSS for OPSCC patients, which provides reference for surgeons to develop a treatment plan and individual prognostic evaluations.

Abbreviations: AJCC = American Joint Committee on Cancer, C-index = concordance index, CSS = cancer-specific survival, NCCN = National Comprehensive Cancer Network, OPSCC = oropharyngeal squamous cell carcinoma, OS = overall survival, SEER = surveillance, epidemiology and end results, TNM = tumor node metastasis.

Keywords: calibration curve, cancer-specific survival, C-index, nomogram, oropharyngeal squamous cell carcinoma, overall survival, surveillance, epidemiology and end results

[a] State Key Laboratory of Military Stomatology & National Clinical Research Center for Oral Diseases & Shaanxi Clinical Research Center for Oral Diseases, Department of Oral and Maxillofacial Surgery School of Stomatology, The Fourth Military Medical University, Xi'an, China, [b] Department of Stomatology, The eighth medical center of Chinese PLA General Hospital, Beijing, China, [c] State Key Laboratory of Military Stomatology & National Clinical Research Center for Oral Diseases & Shaanxi Engineering Research Center for Dental Materials and Advanced Manufacture, Department of Anesthesiology, School of Stomatology, The Fourth Military Medical University, Xi'an, [d] Department of Stomatology, The Chinese PLA General Hospital, Haidian District, Beijing, China.

* Correspondence: Jianhua Wei, State Key Laboratory of Military Stomatology & National Clinical Research Center for Oral Diseases & Shaanxi Clinical Research Center for Oral Diseases, Department of Oral and Maxillofacial Surgery, School of Stomatology, The Fourth Military Medical University, Xi'an, China (e-mail: weijianhuaprof@163.com).

1. Introduction

Oropharyngeal squamous cell carcinoma (OPSCC) is the most common malignant head and neck cancer[1] and it is mainly located in the pharynx, tongue root, pharyngeal tonsil, and soft palate. The annual incidence of oropharyngeal cancers is approximately 400,000 new OPSCC patients in the world, and the incidence of OPSCC has increased sharply in developed countries and nearly 46,000 cases in the United States.[2–4] Currently, National Comprehensive Cancer Network (NCCN) guidelines recommend applying the American Joint Committee on Cancer (AJCC) Staging Manual (7th edition) to evaluate the prognosis of OPSCC patients.[5,6] However, the prognosis of OPSCC patients is influenced by numerous factors, such as age, cigarette and alcohol consumption, tumor site, TNM stage, radiation, and human papilloma-virus (HPV).[1,7] Thus, the consideration of additional relevant elements should provide a more accurate and credible prediction of prognosis than the AJCC staging system. Therefore, we sought to establish a "nomogram" to identify additional relevant factors including age, sex, tumor site, race, different origin, grade, T stage, N stage, M stage, surgery, and radiotherapy to perform a comprehensive analysis. To verify precision and credibility, many researchers recommend using the split-sample method and bootstrap to evaluate a given model.[8–11] Specifically speaking, the nomogram is internally validated by bootstrap re-sampling and externally validated by appraising model's accuracy in split-sample cohorts.[12] The size of the validation cohort (the split-ratio) depends on the coherence and accuracy between the predicted and actual outcomes rather than a fixed value.[13] The accuracy of the nomograms is determined via C-indexes and calibration curves.

A nomogram is an accurate scoring and graphical instrument that can convert the results of multivariate Cox regression into an understandable linear graph. Nomograms are widely used to assist doctors in formulating a therapeutic regimen and have been shown to predict the prognosis of several cancers, including adenoid cystic carcinoma,[14] hepatocellular carcinoma,[15,16] gastric cancer,[17] head and neck cancers,[18] nasopharyngeal cancer,[19] and breast cancer.[20] Most importantly, the application of a nomogram in the early detection of prostate cancers has been included in the NCCN guidelines.[21] Additionally, it is worth noting that the American Joint Committee on Cancer Eighth Edition Cancer Staging Manual indicated that a future version will embrace nomograms and individualized treatment strategies.[22] One previous study used a nomogram to assess the prognosis and progression of OPSCC.[23] However, this research didn't take race, origin, pathological grade, and surgery into account to predict the overall survival (OS) and cancer-specific survival (CSS). In our research, we collected the detailed OS and CSS information of the OPSCC patients. We used the Kaplan–Meier univariate and Cox Proportional Hazard Model multivariate survival analysis to determine the final independent risk clinicopathological parameters influencing the prognosis ($P < .05$). Hence, we sought to establish an OPSCC nomograms to predict long-term OS and CSS based on relevant clinicopathological parameters mentioned above by means of R software R.3.2.4 (Lucent, New Jersey, USA), which should help doctors develop rational and personalized treatments.

2. Methods

2.1. Patient clinical data collection

All 10,980 OPSCC patients from the years 2004 to 2012 were included in this retrospective research, and the clinical data were gathered from the SEER program of the National Cancer Institute.[24] We collected and sorted the detailed clinicopathological parameters, including age, sex, tumor site, race, ethnic origin, pathological grade, surgery or no surgery, radiation or no radiation, T stage, N stage, and M stage (Table 1). The pathological grade consists of Grade I, II, III, IV. Grade I, II, III, IV represent well differentiated, moderately differentiated, poorly differentiated and undifferentiated respectively. Based on the split-sample method discussed in the Introduction, we randomly split the 10,980 OPSCC patients into a modeling cohort and an external validation cohort. The modeling cohort consisted of 9881 OPSCC patients to build a nomogram model. Another 1099 OPSCC patients were included in the external validation cohort to verify the performance and credibility of nomogram model. The study was approved by the Ethical Review Committee of the Fourth Military Medical University.

2.2. Survival analysis

We concentrated on the indicators of OS and CSS to assess the prognosis of modeling cohort. OS was defined from the date of diagnosis to death from any reason or censored if patients were alive at the last follow-up. CSS was calculated from diagnosis to death due to OPSCC or censored if patients were alive or dead because of other causes.

We conducted univariate survival analysis using a Kaplan–Meier and log-rank test. The variables that were statistically significant were included into the multivariate Cox Proportional Hazards analysis to confirm the independent prognostic factors from indicators, such as age, sex, tumor site, race, ethnic origin, pathological grade, surgery, radiation, T stage, N stage, and M stage (Tables 2 and 3). $P < .05$ was considered statistically significant.

2.3. Nomogram development

After conscientious survival analysis using SPSS 21.0 (IBM, Armonk, USA) for Windows software or the rms package of R version 3.2.4 software, we used the independent prognostic indicators for establishing a nomogram.[11,25] The nomogram transformed the clinicopathological data into visual graphics.

2.4. Nomogram validation

The nomogram was required to validate its accuracy through internal and external validation conducted by 1000 times bootstrapping and 10-fold cross-validation measures. The predict reliability was examined in concordance indexes (C-index) and calibration plot.[11] The C-index was used to appraise the difference between predicted and actual situations.[11] The C-index results were acquired through the "rcorrcens" command in R software. Additionally, the calibration plot was composed of two lines: one was a 45° diagonal line representing a reference line, and the other line was the actual line. The distance between the 2 lines reflected the precision of the nomogram. The calibration plots were obtained through the "calibrate" command in R software. All statistical analyses were performed adopting a 2-sided P value, and $P < .05$ was considered to be statistically significant.

3. Results

3.1. Patient clinicopathological data

After applying a strict filter, 9881 and 1099 OPSCC patients from SEER database were included in the modeling and

validation cohorts respectively, using the popular and reasonable random split-sample method (the split ratio was 9:1). In the modeling cohort, the patients' ages ranged from 15 to 98 years (median, 57). Of these 9881 OPSCC patients, 7936 (80.3%) were men. In total, 8503 (86.1%) patients were white, and 9223 (93.3%) patients were non-Spanish-Hispanic-Latino. Of the tumor locations, 7957 (80.5%) cases were primarily located on the oropharynx and 1555 (15.7%) were found on the tongue base. Additionally, 5085 (51.4%) were poorly defined or undifferentiated. Of the studied cases, 5964 (60.4%) received surgery and 8172 (82.7%) received radiotherapy. The proportion of T1–T2 tumors was 68.6% (6778/9881). The N1–3 and M1 tumors accounted for 75.4% and 2.9% of all cases, respectively. The general data for the validation cohort are shown in Table 1.

According to the SAS variable "sur_time_mon" from the SEER database, we found that the median follow-up times for the modeling and validation cohorts were 45 months (1–119 months) and 57 months (1–119 months). According to the SAS variables "STAT_REC," "VSRTSADX," and "ODTH-CLASS" in SEER database, we obtained accurate information on the outcomes for 10,980 OPSCC patients. A total, of 3084 (31.2%) patients in the modeling cohort were deceased at the last follow-up date. Of those patients, 2188 (22.1%) patients died due to OPSCC. Additionally, 896 (9.1%) patients died from causes other than OPSCC.

3.2. Survival analysis and nomogram establishment

The result of the survival analysis with regard to OS and CSS is shown in Tables 2 and 3. For the modeling cohort, the results of the univariate Kaplan–Meier survival analysis revealed that age, sex, race, pathological grade, surgery, radiation, T stage, N stage, and M stage were relevant factors influencing OS ($P < .05$). Multivariate Cox Proportional Hazards analysis showed that all the above elements from the univariate analysis were independent prognostic indicators ($P < .05$), which were shown in Table 2.

### Table 1

Patients' clinicopathological data.

Variables	Modeling group (n = 9881)		Validation group (n = 1099)	
	n	%	n	%
Age				
15–35	118	1.2	105	9.6
36–45	667	6.8	182	16.6
46–55	3182	32.2	478	43.5
56–65	3768	38.1	118	10.7
66–75	1569	15.9	104	9.5
76–85	482	4.9	78	7.1
85+	95	0.9	34	3.0
Sex				
Male	7939	80.3	825	75.1
Female	1942	19.7	274	24.9
Site				
Oropharynx	7957	80.5	742	67.5
TR	1555	15.7	258	23.5
SP	256	2.6	87	7.9
PT	113	1.2	12	1.1
Race				
White	8503	86.1	875	79.6
Black	971	9.8	142	12.9
Others	407	4.1	82	7.5
Origin				
NSHL	9223	93.3	955	86.9
SHL	658	6.7	144	13.1
Grade				
I	589	6.0	78	7.1
II	4207	42.6	479	43.6
III	4886	49.4	488	44.4
IV	199	2.0	54	4.9
Surgery				
Performed	5964	60.4	692	63.0
None	3917	39.6	407	37.0
Radiation				
Yes	8172	82.7	895	81.4
No	1709	17.3	204	18.6
T stage				
T1	2929	29.6	325	29.6
T2	3849	39.0	354	32.2
T3	1034	10.5	158	14.4
T4	1808	20.9	257	23.8
N stage				
N0	2434	24.6	231	21.0
N1	2163	21.9	254	23.1
N2	4828	48.9	553	50.3
N3	456	4.6	71	6.5
M stage				
M0	9597	97.1	993	90.4
M1	284	2.9	106	9.6

Grade I: Well differentiated. II: Moderately differentiated. III: Poorly differentiated. IV: Undifferentiated.
Site: NSHL = Non-Spanish-Hispanic-Latino; PT = pharyngeal tonsil; SP = soft palate; TR = tongue root. Others: American Indian/AK Native, Asian/Pacific Islander.

Table 2

Univariate and multivariate analyses of OS in nomogram cohort.

Variables	Univariate analysis	Multivariate analysis	
	P value	HR (95% CI)	P value
Age	<.001		<.001
15–35		0.145 (0.086–0.244)	<.001
36–45		0.170 (0.126–0.229)	<.001
46–55		0.207 (0.161–0.266)	<.001
56–65		0.282 (0.220–0.361)	<.001
66–75		0.393 (0.306–0.505)	<.001
76–85		0.763 (0.586–0.994)	.045
85+		Reference	
Sex	.001		.010
Male		Reference	
Female		1.121 (1.025–1.225)	.012
Site	.071		
Oropharynx			
TR			
SP			
PT			
Race	<.001		<.001
White		Reference	
Black		1.592 (1.443–1.755)	<.001
Others		0.848 (0.699–1.030)	.096
Origin	.688		
NSHL			
SHL			
Grade	<.001		<.001
I		1.099 (0.823–1.470)	.522
II		1.220 (0.942–1.581)	.132
III		0.955 (0.737–1.238)	.728
IV		Reference	
Surgery	<.001		<.001
Performed		Reference	
None		1.742 (1.610–1.885)	<.001
Radiation	<.001		<.001
Yes		Reference	
No		1.858 (1.691–2.041)	<.001
T stage	<.001		<.001
T1		0.339 (0.303–0.379)	<.001
T2		0.467 (0.427–0.510)	<.001
T3		0.707 (0.632–0.790)	<.001
T4		Reference	
N stage	<.001		<.001
N0		0.611 (0.521–0.717)	<.001
N1		0.713 (0.608–0.835)	<.001
N2		0.737 (0.636–0.853)	<.001
N3		Reference	
M stage	<.001		<.001
M0		0.319 (0.277–0.367)	<.001
M1		Reference	

Grade I: Well differentiated. II: Moderately differentiated. III: Poorly differentiated. IV: Undifferentiated.
Site: NSHL = Non-Spanish-Hispanic-Latino; PT = pharyngeal tonsil; SP = soft palate; TR = tongue root. Others: American Indian/AK Native, Asian/Pacific Islander.
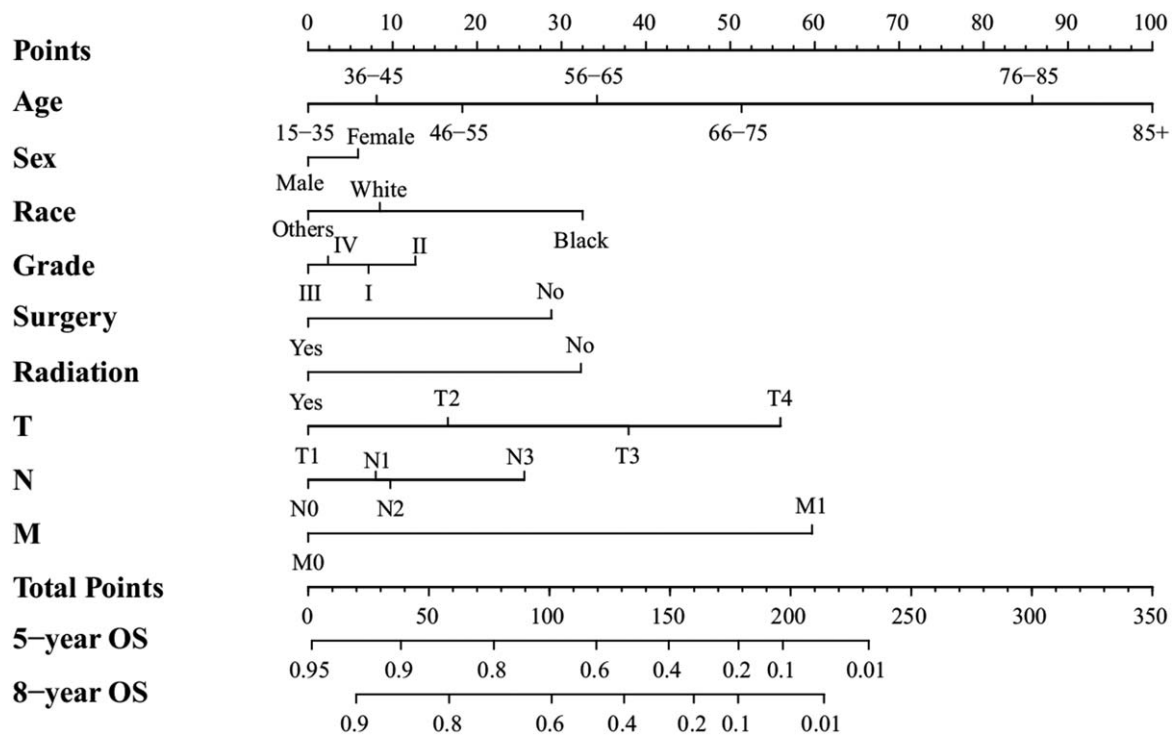
Figure 1. Nomogram predicting 5-year and 8-year OS. Others: American Indian/Alaska Native/Asian or Pacific Islander. Grade I: well differentiated. II: moderately differentiated. III: poorly differentiated. IV: Undifferentiated.

Thus, these factors were included to construct a nomogram to predict 5- and 8-year OS (Fig. 1). The prognosis survival analysis of OS was conducted using SPSS 21.0 (IBM, Armonk, USA) software for Windows.

The results identified that age, race, tumor site, pathological grade, surgery, radiation, T stage, N stage, and M stage as

Table 3

**Univariate and multivariate analyses of CSS in nomogram cohort.**

Variables	Univariate analysis	Multivariate analysis	
	P value	HR (95% CI)	*P* value
Age	<.001		<.001
15–35		0.228 (0.131–0.398)	<.001
36–45		0.200 (0.141–0.285)	<.001
46–55		0.238 (0.176–0.321)	<.001
56–65		0.300 (0.224–0.403)	<.001
66–75		0.367 (0.271–0.496)	<.001
76–85		0.672 (0.488–0.926)	.015
85+		Reference	
Sex	.08		
Male			
Female			
Site	<.001		
Oropharynx		0.605 (0.527–0.694)	<.001
PT		0.240 (0.034–1.718)	.155
SP		0.466 (0.341–0.638)	<.001
TR		Reference	
Race	<.001		<.001
White		Reference	
Black		1.564 (1.395–1.753)	<.001
Others		0.798 (0.628–1.014)	.065

(continued)

Table 3

(continued).

Variables	Univariate analysis	Multivariate analysis	
	P value	HR (95% CI)	*P* value
Origin	.608		
NSHL			
SHL			
Grade	<.001		<.001
I		1.274 (0.895–1.814)	.179
II		1.364 (0.993–1.873)	.055
III		1.034 (0.763–1.420)	.836
IV		Reference	
Surgery	<.001		<.001
Performed		Reference	
None		2.117 (1.900–2.359)	<.001
Radiation	<.001		<.001
Yes		Reference	
No		1.902 (1.699–2.129)	<.001
T stage	<.001		<.001
T1		0.273 (0.238–0.314)	<.001
T2		0.397 (0.357–0.440)	<.001
T3		0.641 (0.564–0.729)	<.001
T4		Reference	
N stage	<.001		<.001
N0		0.518 (0.432–0.623)	<.001
N1		0.662 (0.554–0.791)	<.001
N2		0.683 (0.580–0.804)	<.001
N3		Reference	
M stage	<.001		<.001
M0		0.277 (0.238–0.322)	<.001
M1		Reference	

CSS = cancer-specific survival.
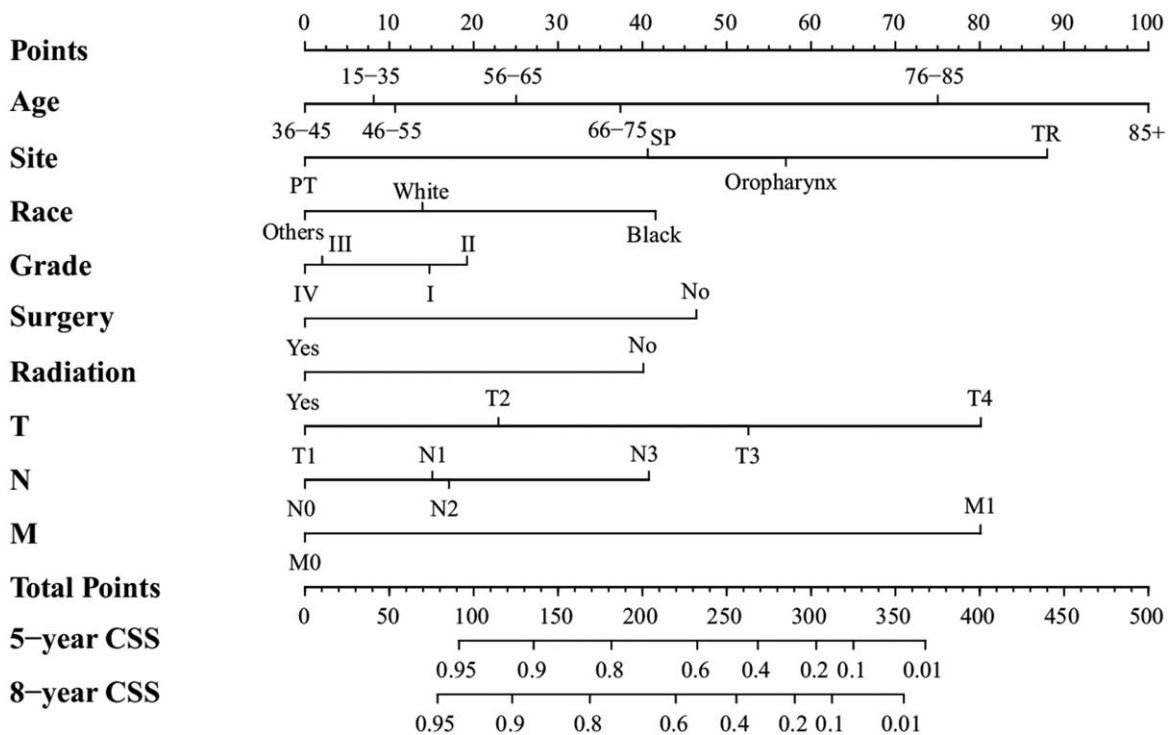
Figure 2. Nomogram predicting 5-year and 8-year CSS. Others: American Indian/Alaska Native/Asian or Pacific Islander. Grade I: well differentiated. II: moderately differentiated. III: poorly differentiated. IV: undifferentiated. PT: pharyngeal tonsil. SP: soft palate. TR: tongue root.

independent factors influencing CSS (Table 3). Furthermore, these factors were used to establish another nomogram to forecast the 5-year and 8-year CSS (Fig. 2).

3.3. Nomogram validation

The nomogram's credibility are internally validated by 1000 times bootstrap re-sampling and externally validated by evaluating model's accuracy in split-sample cohort of 1099 patients. The concordance index (C-index) and calibration curves were used to assess the precision of the nomograms. Because the value of C-index was >0.7, the predicted OS and CSS were consistent with the actual OS and CSS. Our results of internal validation showed that the C-index values of OS and CSS were 0.742 and 0.765, respectively. External validation revealed that the C-index of OS and CSS were 0.740 and 0.759, respectively. Moreover, the internal and external calibrations were close to the 45° ideal straight line (Figs. 3 and 4).

4. Discussion

Due to the combined effects of clinicopathological factors, an increasing number of patients have been diagnosed with OPSCC.[26] Generally, surgery, radiation and chemotherapy are familiar therapeutic methods for OPSCC, which would influence the patients' prognosis to a great extent.[27] Additionally, other relevant factors including age, sex, smoking, TNM stages also affected the prognosis of patients with OPSCC. In order to provide individualized, patient-specific prediction of prognosis, we have developed and validated two nomograms to forecast 5- and 8-year OS and CSS of patients with OPSCC via popular random split-sample method (split ratio=9:1), which has been applied

extensively.[28] Nomogram could integrate the clinical and pathological factors together. Notably, the 8th Head Neck Cancer AJCC staging system showed that in the future version they would incorporate the nomogram to evaluate the prognosis.

We calculated the estimated OS and CSS by means of Kaplan–Meier method. After the Kaplan–Meier univariate and Cox multivariate survival analysis, we obtained the independent prognostic risk factors and established 2 nomograms. The nomograms were well validated internally and externally. The C-index of all nomograms were >0.7, and there was a good consistency between the calibration curve and the 45° straight line, showing a potential advantage than Shoultz-Henley model which was short of validation.[29] There had quantitative axis for each index, corresponding to the axises representing 5- and 8-year OS and CSS (see Figs. 1 and 2). The more left the axis was, the higher the survival rate was. We found that OS and CSS gradually declined after the age of 55. Therefore, the age groups of "15 to 35" and "36 to 45" were at the far left of the age axises, demonstrating the best OS and CSS respectively. Many studies have found that age was a significant element influencing survival.[26,27,30,31] Another study showed that patients <45 years old had the best CSS compared with the age groups of "45 to 64" and ">65", conforming to our research.[26] Compared with white patients and patients of other races, black patients have demonstrated relatively lower survival, which is in agreement with the outcomes of the current research.[32] As a mechanism, one study hypothesized that melanin might contribute to tumorigenesis and cancer development.[18] Notably, the patients with pathological grades III and IV disease showed improved OS and lower cancer-specific death compared with patients with grades I and II disease (Table 3). Moreover, the results were also verified in the nomograms constructed. This may be because that radiotherapy and surgery
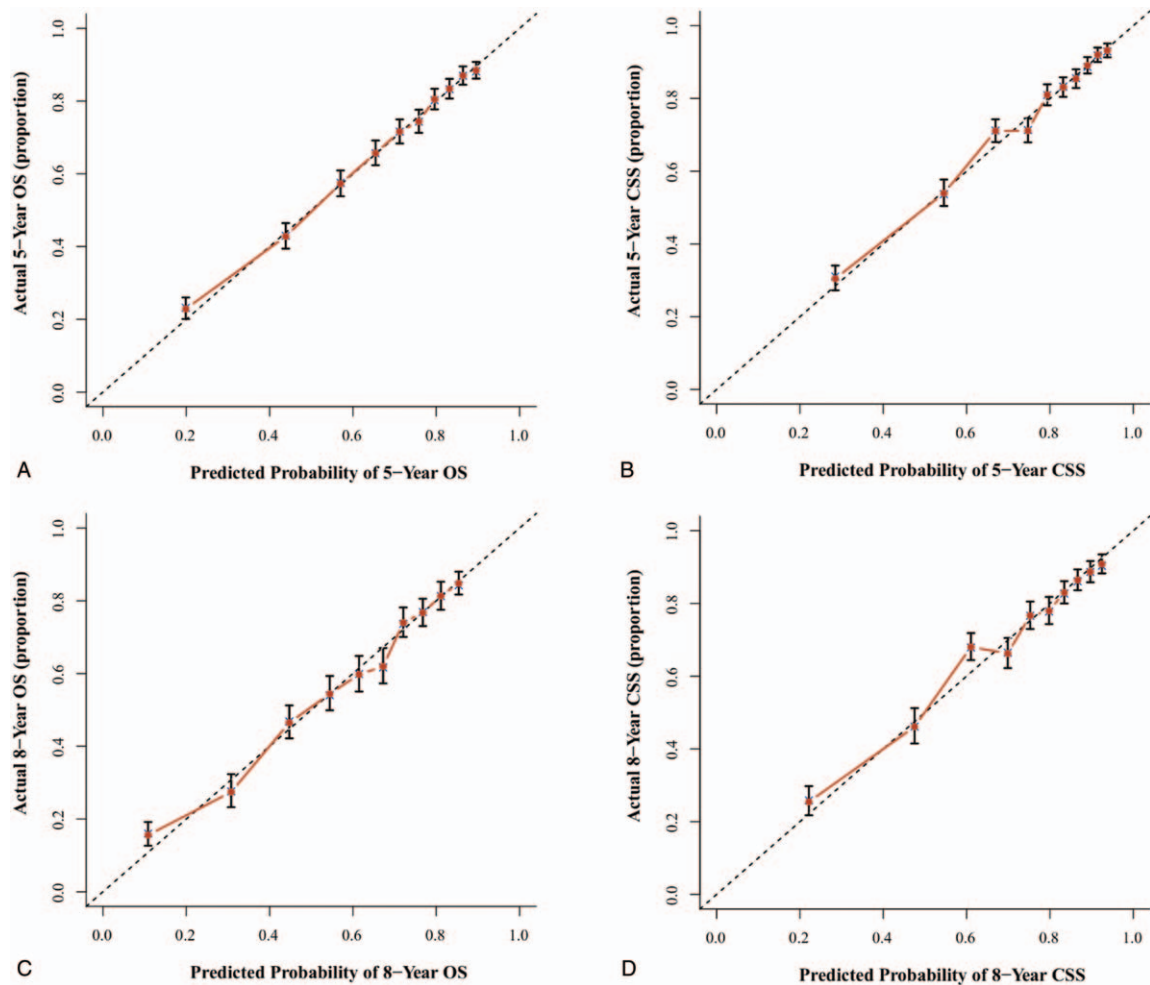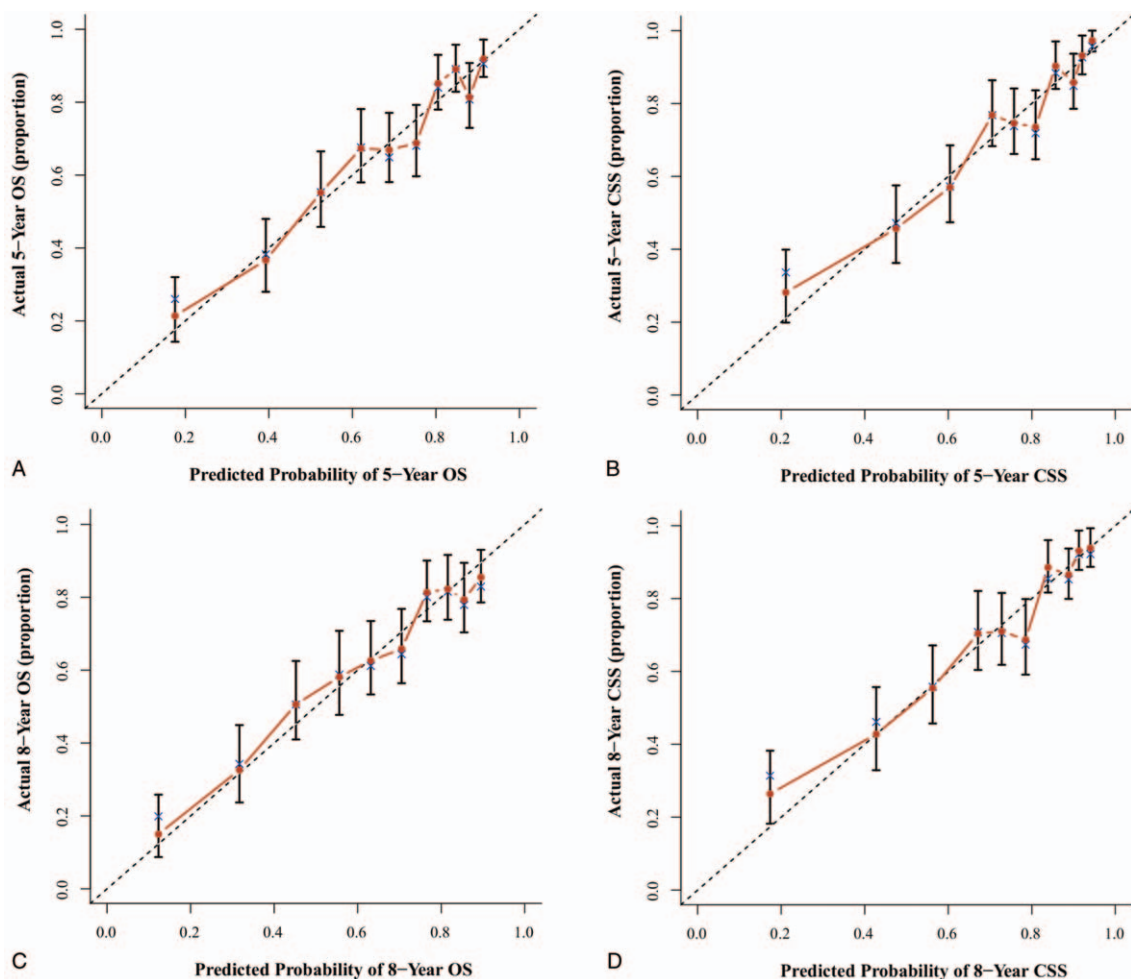
Figure 3. Internal validation via calibration curves for 5-year and 8-year OS (A, C) and 5-year and 8-year CSS (B, D). The 45° line embodys an perfect match between the actual survival (*y* axis) and nomogram-predicted survival (*x* axis). The perpendicular line means 95% confidence intervals. CSS= cancer-specific survival; OS= overall survival.

were used to treat the higher-grade patients primarily. T stage, N stage were also the significant predictive factors influencing the prognosis.[33,34]

The procedure of nomograms to predict the 5- and 8- year OS and CSS was simple and feasible. We selected clinicopathological factor sub-categories according to personalized conditions and constructed a vertical line to the point axis. Then, we added all the points acquired by each sub-category corresponding to the total points axis. Finally, we plotted vertical lines from the total points to the 5- and 8- year OS and CSS axis to obtain the predicted value. This process was completed using the "rms" package of R software.[35] Additionally, we conducted both internal and external validation; the concordance indexes were all >0.7, and they matched well with the 45° straight line (Figs. 3 and 4). The nomogram was more accurate than TNM staging for predicting the prognosis of OPSCC. As an example, compare the following 2 types of T4N0M0 patients: type 1, a 50-year-old white patient with pathological grade III disease who received only surgery, and type 2, a 60-year-old black patient with pathological grade II disease underwent surgery and radiation. If we evaluate the prognosis of those 2 types of patients according to AJCC TNM classification,[6] the 2 patients all belonged to stage IV, having the

same outcome. Yet, the results were different using the nomogram. The 5-year predicted OS for type 1 and type 2 patient was 76% and 30% accordingly. Therefore, we included the independent prognostic factors into the model to construct more credible nomograms to predict the OS and CSS.

Our research had obvious strengths. First, we collected detailed and reliable information regarding OPSCC patients from the SEER database to guarantee the credibility of the results. The data came from 18 SEER registries located in 18 different states, which was a large-sample multi-center research. Secondly, our nomograms have potential advantage over previously published models for OPSCC. Shoultz-Henley et al[29] had established a nomogram, but the model was neither internally nor externally validated. By contrast, we validated the our nomogram models via C-index and calibrations, showing a higher accurancy. Rios et al[33] also constructed a nomogram, but our larger sample capacity and longer follow-up period allowed us to develop a separate nomograms about 5- and 8-year OS and CSS in patients with OPSCC. Karadaghy et al[36] had developed prediction model using machine learning for 5-year overall survival. However, the main obstacles to the widespread application of this algorithm include convenience, regulatory, and financial considerations.

Figure 4. External validation via calibration curves for 5-year and 8-year OS (A, C) and 5-year and 8-year CSS (B, D). The 45° line embodys an perfect match between the actual survival (*y* axis) and nomogram-predicted survival (*x* axis). The perpendicular line means 95% confidence intervals. CSS = cancer-specific survival; OS = overall survival.

Our research had certain limitations. First, the SEER database didn't include other significant prognostic factors, such as anemia,[33] chemoradiation,[37] and cigarette and alcohol consumption status,[32] thrombocytosis.[29] Thus, our nomograms lacked evaluation of the above elements. Also, we couldn't assess the disease-free survival, progression-free survival and locoregional control. Second, we couldn't obtain the dose, cycle and type of radiotherapy to compare the survival differences between various radiation plans. Third, we couldn't acquire the life habits of all the patients, which may influence the prognosis. Meanwhile, we couldn't gain the TNM staging information before the year 2004.

In conclusion, we have constructed 2 successful nomograms forecasting 5- and 8- year OS and CSS via Cox regression. We also obtained favorable C-indexes through internal and external validation. These nomograms may provide surgeons with a reference to develop treatment plans and conduct individual prognostic evaluations, as the future will most certainly bring an era of personalized therapy.

Acknowledgment

Author contributions

Conceptualization: Fengze Wang, Jiao Wen, and Jianhua Wei.
Data curation: Jiao Wen, Tingting Jia, Fangchong Du.
Formal analysis: Jiao Wen, Xinjie Yang, and Tingting Jia.
Funding acquisition: Jianhua Wei.
Investigation: Fengze Wang, Jiao Wen, Tingting Jia
Methodology: Xinjie Yang, Tingting Jia, and Fangchong Du.
Project administration: Fengze Wang, Jiao Wen, and Jianhua Wei.
Resources: Tingting Jia and Fangchong Du.
Software: Jiao Wen and Tingting Jia.
Supervision: Jianhua Wei.
Writing – original draft: Fengze Wang, Jiao Wen, Xinjie Yang, Tingting Jia, and Fangchong Du.
Writing – review & editing: Fengze Wang, Xinjie Yang, and Jianhua Wei.

References

[1] Marur S, Forastiere AA. Head and neck cancer: changing epidemiology, diagnosis, and treatment. Mayo Clin Proc 2008;83:489–501.
[2] Chaturvedi AK, Anderson WF, Lortet-Tieulent J, et al. Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. J Clin Oncol 2013;31:4550–9.

[3] Rainsbury JW, Ahmed W, Williams HK, et al. Prognostic biomarkers of survival in oropharyngeal squamous cell carcinoma: systematic review and meta-analysis. Head Neck 2013;35:1048–55.

[4] Sim FW, Xiao HD, Bell RB. Margin analysis: squamous cell carcinoma of the oropharynx. Oral Maxillofac Surg Clin North Am 2017;29:269–80.

[5] Pfister DG, Spencer S, Brizel DM, et al. Head and neck cancers, Version 1.2015. J Natl Compr Canc Netw 2015;13:847–56.

[6] Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol 2010;17:1471–4.

[7] Fakhry C, Westra WH, Li S, et al. Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. J Natl Cancer Inst 2008;100:261–9.

[8] Austin PC, Steyerberg EW. Events per variable (EPV) and the relative performance of different strategies for estimating the out-of-sample validity of logistic regression models. Stat Methods Med Res 2017;26:796–808.

[9] Ron S. Validation of regression models: methods and examples. Technometrics 1977;19:415–28.

[10] Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Intern Med 1999;130:515–24.

[11] Harrell FEJr. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York: Springer; 2001.

[12] Jeong SJ, Yeon JS, Lee JK, et al. Development and validation of nomograms to predict the recovery of urinary continence after radical prostatectomy: comparisons between immediate, early, and late continence. World J Urol 2014;32:437–44.

[13] Harrell FJ, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361–87.

[14] Ganly I, Amit M, Kou L, et al. Nomograms for predicting survival and recurrence in patients with adenoid cystic carcinoma. An international collaborative study. Eur J Cancer 2015;51:2768–76.

[15] Cho CS, Gonen M, Shia J, et al. A novel prognostic nomogram is more accurate than conventional staging systems for predicting survival after resection of hepatocellular carcinoma. J Am Coll Surg 2008;206:281–91.

[16] Li J, Liu Y, Yan Z, et al. A nomogram predicting pulmonary metastasis of hepatocellular carcinoma following partial hepatectomy. Br J Cancer 2014;110:1110–7.

[17] Liu J, Geng Q, Liu Z, et al. Development and external validation of a prognostic nomogram for gastric cancer using the national cancer registry. Oncotarget 2016;7:35853–64.

[18] Ju J, Wang J, Ma C, et al. Nomograms predicting long-term overall survival and cancer-specific survival in head and neck squamous cell carcinoma patients. Oncotarget 2016;7:51059–68.

[19] Cho JK, Lee GJ, Yi KI, et al. Development and external validation of nomograms predictive of response to radiation therapy and overall survival in nasopharyngeal cancer patients. Eur J Cancer 2015;51:1303–11.

[20] Wen J, Ye F, He X, et al. Development and validation of a prognostic nomogram based on the log odds of positive lymph nodes (LODDS) for breast cancer. Oncotarget 2016;7:21046–53.

[21] Kawachi MH, Bahnson RR, Barry M, et al. NCCN clinical practice guidelines in oncology: prostate cancer early detection. J Natl Compr Canc Netw 2010;8:240–62.

[22] Lydiatt WM, Patel SG, O'Sullivan B, et al. Head and Neck cancers-major changes in the American Joint Committee on cancer eighth edition cancer staging manual. CA Cancer J Clin 2017;67:122–37.

[23] Larsen CG, Jensen DH, Carlander AF, et al. Novel nomograms for survival and progression in HPV+ and HPV- oropharyngeal cancer: a population-based study of 1,542 consecutive patients. Oncotarget 2016;7:71761–72.

[24] National Cancer Institute: Surveillance, Epidemiology, and End Results Program. Available at: http://seer.cancer.gov. Access date: 2018.11.10.

[25] Dongsheng Yang. Build prognostic nomograms for risk assessment using SAS. SAS Global Forum 2013;103:264–9.

[26] Zumsteg ZS, Cook-Wiens G, Yoshida E, et al. Incidence of oropharyngeal cancer among elderly patients in the United States. JAMA Oncol 2016;2:1617–23.

[27] Marur S, Forastiere AA. Head and neck squamous cell carcinoma: update on epidemiology, diagnosis, and treatment. Mayo Clin Proc 2016;91:386–96.

[28] Wang YY, Xiang BD, Ma L, et al. Development and validation of a nomogram to preoperatively estimate post-hepatectomy liver dysfunction risk and long-term survival in patients with hepatocellular carcinoma. Ann Surg 2020;1: doi:10.1097/SLA.0000000000003803.

[29] Shoultz-Henley S, Garden AS, Mohamed AS, et al. Prognostic value of pretherapy platelet elevation in oropharyngeal cancer patients treated with chemoradiation. Int J Cancer 2016;138:1290–7.

[30] Skillington SA, Kallogjeri D, Lewis JJ, et al. Prognostic importance of comorbidity and the association between comorbidity and p16 in oropharyngeal squamous cell carcinoma. JAMA Otolaryngol Head Neck Surg 2016;142:568–75.

[31] Shen W, Sakamoto N, Yang L. Cancer-specific mortality and competing mortality in patients with head and neck squamous cell carcinoma: a competing risk analysis. Ann Surg Oncol 2015;22:264–71.

[32] Fakhry C, Westra WH, Wang SJ, et al. The prognostic role of sex, race, and human papillomavirus in oropharyngeal and nonoropharyngeal head and neck squamous cell cancer. Cancer 2017;123:1566–75.

[33] Rios VE, Hoebers F, Aerts HJ, et al. Externally validated HPV-based prognostic nomogram for oropharyngeal carcinoma patients yields more accurate predictions than TNM staging. Radiother Oncol 2014;113:324–30.

[34] Tham T, Machado R, Herman SW, et al. Personalized prognostication in head and neck cancer: a systematic review of nomograms according to the AJCC precision medicine core (PMC) criteria. Head Neck 2019;41:2811–22.

[35] Harrell F. Predict: compute predicted values and confidence limits. Package 'rms'; pp. 132–137. Available at: https://cran.r-project.org/web/packages/rms/rms.pdf. Access date: 2017.04.03.

[36] Karadaghy OA, Shew M, New J, et al. Development and assessment of a machine learning model to help predict survival among patients with oral squamous cell carcinoma. JAMA Otolaryngol Head Neck Surg 2019; doi:10.1001/jamaoto.2019.0981.

[37] Calais G, Bardet E, Sire C, et al. Radiotherapy with concomitant weekly docetaxel for Stages III/IV oropharynx carcinoma. Results of the 98-02 GORTEC Phase II trial. Int J Radiat Oncol Biol Phys 2004;58:161–6.