



Nomograms-based prediction of overall and cancer-specific survivals for patients diagnosed with major salivary gland carcinoma

Zhiyong Guo[^], Zilin Wang, Yige Liu, Jing Han, Jiannan Liu, Chenping Zhang

Department of Oromaxillofacial-Head & Neck Oncology, Shanghai Ninth People's Hospital, College of Stomatology, Shanghai Jiao Tong University School of Medicine, National Clinical Research Center for Oral Diseases, Shanghai Key Laboratory of Stomatology & Shanghai Research Institute of Stomatology, Shanghai, China

Contributions: (I) Conception and design: C Zhang, J Liu; (II) Administrative support: Z Wang, Y Liu, J Han; (III) Provision of study materials or patients: Z Guo; (IV) Collection and assembly of data: Z Guo, Z Wang; (V) Data analysis and interpretation: Z Guo; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Chenping Zhang; Jiannan Liu. Department of Oral & Maxillofacial-Head & Neck Oncology, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Zhi-zao-ju Road No. 639, Shanghai 200011, China. Email: zhang.chenping@hotmail.com; laurence_ljn@163.com.

Background: Major salivary glands carcinoma (MSGC) is a relatively rare cancer with diverse histological types and biological behavior. The treatment planning and prognosis prediction are challenging for clinicians. The aim of the current study was to establish a reliable and effective nomogram to predict the overall survival (OS) and cancer-specific survival (CSS) for MSGC patients.

Methods: Patients pathologically diagnosed with MSGC were recruited from Surveillance, Epidemiology, and End Results (SEER) database and randomly divided into training and validation groups (7:3 ratio). Univariate, multivariate Cox proportional hazard models, and least absolute shrinkage and selection operator (LASSO) regression were adopted for the selection of risk factors. Nomograms were developed using R software. The model performance was evaluated by drawing receiver operating characteristic (ROC), overtime C-index curves, and calibration curves. Harrell C-index, areas under the curves (AUC), and Brier score were also calculated. The decision curve analysis (DCA) was conducted to measure the net clinical benefit.

Results: A total of 11,362 patients were identified and divided into training (n=7,953) and validation (n=3,409) dataset. Sex, age, race, marital status, site, differentiation grade, American Joint Committee on Cancer (AJCC) stage, T/N/M stage, tumor size, surgery, and histological type were incorporated into the Cox hazard model for OS prediction after variable selection, while all predictors, except for marital status and site, were selected for CSS prediction. For 5-year prediction, the AUC of the nomogram for OS and CSS was 83.5 and 82.7 in the training and validation dataset, respectively. The C-index was 0.787 for OS and 0.798 for CSS in the validation group. The Brier score was 0.0153 and 0.0130 for OS and CSS, respectively. The calibration curves showed that the nomogram had well prediction accuracy. From the perspective of DCA, a nomogram was superior to the AJCC stage and TNM stage in net benefit. In general, the performance of the nomogram was consistently better compared to the AJCC stage and TNM stage across all settings.

Conclusions: The performance of the novel nomogram for predicting OS and CSS of MSGC patients was further verified, revealing that it could be used as a valuable tool in assisting clinical decision-making.

Keywords: Major salivary glands carcinoma (MSGC); nomogram; prognosis; survival; Surveillance, Epidemiology, and End Results (SEER)

[^] ORCID: 0000-0001-8791-0020.

Submitted Apr 08, 2021. Accepted for publication Jun 20, 2021.

doi: 10.21037/atm-21-1725

View this article at: <https://dx.doi.org/10.21037/atm-21-1725>

Introduction

Major salivary glands carcinoma (MSGC) is a rare malignant tumor, which accounts for 11% of oropharyngeal neoplasms in the United States (U.S.) (1,2). MSGC comprises over 20 different histological subtypes with different biological behavior and therapy responses. Due to the low number of patients in each subtype, there are no randomized phase III studies (3). The management of MSGC is still challenging. It there are no randomized phase III studies is reported that the 5-year overall survival (OS) is only about 50% (4). The rarity of MSGC, combined with complex classification schema and non-standard therapy, makes the prediction of the survival of MSGC patients difficult. Therefore, developing an effective prognostic tool to improve clinical treatment strategies for MSGC is of great clinical importance.

Although there are many prognosis factors for cancers such as sex, race, age, and other characteristics, tumor staging systems are most commonly applied for the assessment of risk and the decision-making of cancer management in clinical practice, especially American Joint Committee on Cancer (AJCC) classification system and TNM staging system (5,6). Still, these systems have limitations: only tumor size, lymph nodes, and distant metastasis were considered. Moreover, MSGC patients classified as in the same stage may have completely different survival outcomes (7). The nomogram is an effective and convenient tool based on a statistical model for quantitative prediction of clinical events using multiple selected prognostic factors. It has been widely applied for predicting survival in patients with malignancies, such as lung cancer, breast cancer, and colon cancer (8-10). In their recent study, Lukovic *et al.* developed a predictive model for the prediction of distant metastasis (11), thus revealing the broad application prospect of predictive models in clinical practice for MSGC management. Ali *et al.* created a nomogram for predicting survival and recurrence of major salivary gland carcinoma based on the characteristics of 301 patients recruited at Memorial Sloan Kettering Cancer Centre between 1985 and 2009; however, the sample size of the development cohort was limited (12,13). Furthermore, histological type, which is a key factor in determining the biological behavior of the tumor, is not considered in the

nomogram. Li *et al.* constructed a nomogram for survival prediction of MSGC (14). Nonetheless, improper study design and non-transparent reporting of results significantly damaged the reliability of the study. Therefore, an updated and reliable nomogram for predicting the survival of MSGC patients is urgently needed.

In the present study, we used data of MSGC patients from the Surveillance, Epidemiology, and End Results (SEER) database to develop a predictive model, which was presented as a nomogram, and to test the performance for OS and cancer-specific survival (CSS). We present the following article in accordance with the TRIPOD reporting checklist (available at <https://dx.doi.org/10.21037/atm-21-1725>) (15).

Methods

Data source and selection criteria

Data for this study were obtained from the SEER database (approval number: 21476-Nov2019), which is an authoritative source for cancer statistics in the United States (<https://seer.cancer.gov/>). The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The ethical approval was waived since the study was based on secondary data analysis.

The inclusion criteria were the following: (I) patients diagnosed with major salivary gland carcinoma since 2004; (II) the survival time of patients is known. Patients were excluded if they met the following conditions: (I) survival time <1 month; (II) patients with missing data; (III) patients only diagnosed by autopsy or death certificate.

Fourteen predictor variables included in the present study were: sex, age at diagnosis, race, marital status, location of the tumor, grade of differentiation, T/N/M stage, AJCC stage, size of the tumor, histological type of the tumor, surgery, and previously having/not having surgery of lymph nodes. Vital status and cancer-specific death, which were primary outcomes in this study, were also extracted. For radiotherapy and chemotherapy, patients were divided into two groups—"Yes" and "No/Unknown" in the SEER. Combining "No" and "Unknown" as a group would probably reduce the effect of the predictor and lead to confusion in the clinical practice. Radiotherapy and

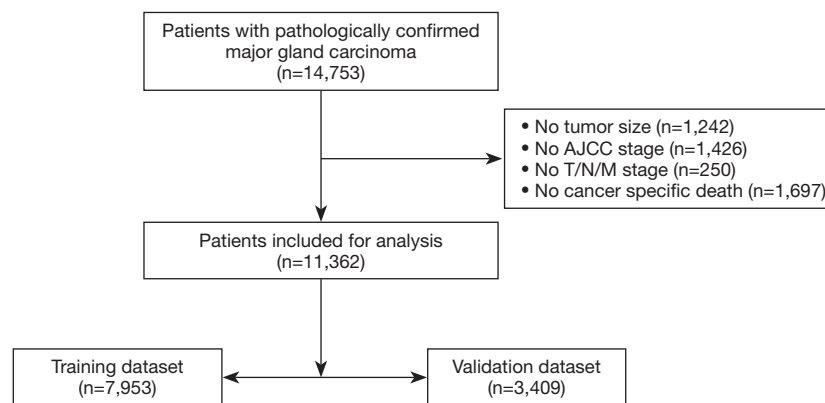


Figure 1 Flow chart of patient selection.

chemotherapy were not extracted in this study. All data were extracted using SEER*Stat 8.3.5 software.

Statistical analysis

The whole population extracted from the SEER database was randomly divided into training dataset and validation dataset at ratio 7:3. Categorical variables were analyzed by Chi-square test, and continuous variables (survival months) were analyzed by rank-sum test. Univariate and multivariate Cox proportional hazards regression were performed to assess the relationship between predictor variables and survival outcomes. Hazard ratio (HR) was applied as the effect size. To select predictor variables, the least absolute shrinkage and selection operator (LASSO) regression was adopted (16).

The performance of the predictive model was evaluated in both the training dataset and validation data set. The discriminative performance of the predictive model was evaluated by Harrell C-index (C-index) at a given time point and over time C-index (17). The C-index estimates the probability that the model can correctly sort the sequence of events in a randomly selected pair of cases. The value of the C-index ranges from 0.5 to 1. A higher C-index means better predict performance (18). The receiver operating characteristic (ROC) curves were drawn, and the areas under the curves (AUC) were also calculated. AUC is another parameter for measuring the discriminative power of the model. Calibration curves were drawn to further evaluate the concordance of the predicted survival probability and the observed probability. Brier score, which is squared prediction error, was calculated. A lower Brier score means a lower prediction error (19). The decision curve analysis (DCA) was conducted for clinical significance

assessment by calculating net benefit (20). All statistical analyses were performed using R version 4.0.3 (<https://www.r-project.org/>). An α level was set as 0.05.

Results

Characteristics of patients

A total of 14,753 eligible patients recorded in the SEER databases were initially recruited. Patients whose survival time was less than 1 month were excluded ($n=718$). Since the design of this study implied complete case analysis, patients with missing values were also excluded ($n=2,673$). Finally, 11,362 patients were included. Patients were randomly assigned into training dataset ($n=7,953$) and validation dataset ($n=3,409$) with a ratio 7:3. The flowchart of patient selection is shown in *Figure 1*. There was no statistical difference in patients' distribution between the training and validation dataset ($P>0.05$). The median survival time for the whole population was 48 months. The majority of the patients were male ($n=6,400$, 56%), white ($n=9,214$, 81%), married when diagnosed with MSGC, with an age between 60 and 80 ($n=4,592$, 40%). Most of lesions were located at parotid gland ($n=9,191$, 81%) and were classified as AJCC IV stage ($n=3,139$, 28%). For TNM stage, T1 accounted for 29% ($n=3,305$), N0 accounted for 67% ($n=7,568$), and M0 accounted for 89% ($n=10,082$). The size of most tumors was 0–1 cm ($n=9,651$, 85%). Mucoepidermoid carcinoma was the most common histological type ($n=2,637$, 23%). There were 86% patients ($n=9,758$, 86%) who underwent surgery, while 61% of total population ($n=6,896$) underwent removal of lymph nodes. A total of 3,616 death cases were reported in this cohort, including 2,338 cancer-related

death cases and 1,278 competing death cases. The detailed characteristics of patients are shown in *Table 1*.

Establishment of predict model

First, univariate Cox regression was constructed. With reference to OS, all variables were significantly correlated with outcome ($P < 0.001$). The results of multivariate Cox regression showed that sex, age, race, grade, AJCC stage, T/N stage, tumor size, surgery, and histological type were significantly correlated with OS ($P < 0.001$) (*Table 2*). Both M stage and site also showed a significant correlation ($P < 0.01$), except for marital status and removal of lymph nodes ($P > 0.05$). As for CSS, similar results were obtained in the univariate Cox regression, while marital status and M stage were not significantly associated with CSS in the multivariate Cox regression ($P > 0.05$) (*Table 3*). Then the LASSO model was applied for variable selection (*Figure S1*). For OS, only one variable (removal of lymph nodes) was not incorporated in the model. A total of 12 predictors, except for marital status and site, were selected for CSS prediction. The regression coefficients are provided in *Table S1*.

The nomogram was drawn using the “rms” package (*Figure 2*). When using the nomogram to predict the survival of patients, a perpendicular line was drawn towards the point line from each predictor to identify the specific value of each variable. The total points were the sum of values for each variable. Then, the survival likelihood was determined by drawing a line from the total points' line down along the survival axis.

Evaluation of model performance

The performance of the predictive model was compared with the AJCC stage and TNM stage system. The values of AUC, C-index, and Brier score are presented in *Table 4* and *Table S2*. For OS, AUC of nomogram in the training dataset was 83.5 (82.4–84.7), while the AUC of TNM stage and AJCC stage was 72.7 (71.3–74.1) and 71.6 (70.2–73.0), respectively. In the validation dataset, the AUC of the nomogram was 82.7 (81.0–84.3), while the AUC of the TNM stage and AJCC stage were 71.0 (68.9–73.2) and 69.8 (67.7–71.9), respectively. Similar results were found for

CSS: nomogram yielded the highest value of AUC of 83.9 (82.6–85.2), followed by TNM stage (75.9, 74.4–77.3) and AJCC stage (73.9, 72.4–75.3) in the training cohort. The AUC values of the nomogram, TNM stage, AJCC stage were 82.8 (80.9–84.7), 74.5 (72.2–76.7), 72.6 (70.4–74.8), respectively, in the validation cohort. As shown in *Figure 3* and *Figure S2*, the AUC values of the nomogram were the best across all settings.

The C-index of the nomogram in predicting OS was 0.796 (0.788–0.804) and 0.787 (0.774–0.799), showing a greater predictive performance compared with TNM stage (0.724, 0.713–0.734; 0.711, 0.695–0.726) and AJCC stage (0.751, 0.739–0.762; 0.730, 0.712–0.748) in the training group and validation group. For CSS, the C-index of nomogram was 0.806 (0.797–0.815) in training group and 0.798 (0.78–0.81) in validation group. The C-index of TNM stage was 0.75 (0.74–0.76) and 0.74 (0.784–0.812), while the C-index of AJCC stage was 0.781 (0.769–0.793) and 0.769 (0.750–0.788) in training set and validation set, respectively. The over-time C-indexes of nomogram were consistently better than TNM stage and AJCC stage throughout the investigated period, as shown in *Figure 4*.

The Brier scores of the nomogram in predicting OS, which was 0.0148 and 0.0153 in the training set and validation set, respectively, were better than the TNM stage and AJCC stage. Concerning the TNM stage and AJCC stage, the Brier scores were 0.0188, 0.0187 in the training group, and 0.0194, 0.0194 in the validation group. As for CSS, the Brier score of the nomogram, TNM stage, AJCC stage in the training dataset were 0.0128, 0.0153, 0.0160, and 0.0130, 0.0155, 0.0160 in the validation set, respectively. The calibration curves displayed by the nomogram were closer to the reference line, which indicated that the nomogram exhibited a higher accuracy than other systems (*Figure 5*, *Figure S3*).

The DCA was adopted for the clinical significance assessment of the nomogram. The performance of the nomogram was also compared with the TNM stage and AJCC stage. As shown in *Figure 6* and *Figure S4*, the net benefits of nomogram were higher compared to the other two systems, no matter in training dataset or validation dataset, for OS or CSS. The results of DCA indicated that the nomogram established by the present study was applicable for predicting the 3- and 5-year CSS and OS.

Table 1 Baseline demographic and clinical characteristics of the patients

Characteristic	Total (n=11,362), n [%]	Training dataset (n=7,953), n [%]	Validation dataset (n=3,409), n [%]	P value
Sex				–
Female	4,962 [44]	3,497 [44]	1,465 [43]	
Male	6,400 [56]	4,456 [56]	1,944 [57]	
Age				0.298
<40 years	1,573 [14]	1,084 [14]	489 [14]	
40–59 years	3,562 [31]	2,531 [32]	1,031 [30]	
60–80 years	4,592 [40]	3,211 [40]	1,381 [41]	
>80 years	1,635 [14]	1,127 [14]	508 [15]	
Race				0.234
White	9,214 [81.1]	6,468 [81.3]	2,746 [80.6]	
Black	1,047 [9.2]	705 [8.9]	342 [10.0]	
Others	971 [8.5]	690 [8.7]	281 [8.2]	
Unknown	130 [1.1]	90 [1.1]	40 [1.2]	
Marital status				0.379
Divorced	862 [7.6]	610 [7.7]	252 [7.4]	
Married	6,321 [55.6]	4,416 [55.5]	1,905 [55.9]	
Separated	91 [0.8]	73 [0.9]	18 [0.5]	
Single	2,110 [18.6]	1,465 [18.4]	645 [18.9]	
Widowed	1,190 [10.5]	839 [10.5]	351 [10.3]	
Unknown	788 [6.9]	550 [6.9]	238 [7.0]	
Site				0.306
Parotid gland	9,191 [80.9]	6,434 [80.9]	2,757 [80.9]	
Submandibular gland	1,478 [13.0]	1,028 [12.9]	450 [13.2]	
Sublingual gland	124 [1.1]	93 [1.2]	31 [0.9]	
Overlapping of major gland	11 [<0.1]	5 [<0.1]	6 [0.2]	
Unspecified major gland	558 [4.9]	393 [4.9]	165 [4.8]	
Grade				0.575
I	1,509 [13.3]	1,066 [13.4]	443 [13.0]	
II	2,673 [23.5]	1,895 [23.8]	778 [22.8]	
III	2,342 [20.6]	1,617 [20.3]	725 [21.3]	
IV	1,007 [8.9]	694 [8.7]	313 [9.2]	
Unknown	3,831 [33.7]	2,681 [33.7]	1,150 [33.7]	

Table 1 (continued)

Table 1 (continued)

Characteristic	Total (n=11,362), n [%]	Training dataset (n=7,953), n [%]	Validation dataset (n=3,409), n [%]	P value
AJCC stage				0.630
I	2,885 [25]	1,988 [25]	897 [26]	
II	2,048 [18]	1,450 [18]	598 [18]	
III	1,946 [17]	1,370 [17]	576 [17]	
IV	3,139 [28]	2,207 [28]	932 [27]	
Unknown	1,344 [12]	938 [12]	406 [12]	
T stage				0.320
0	40 [0.4]	28 [0.4]	12 [0.4]	
1	3,305 [29.1]	2,268 [28.5]	1,037 [30.4]	
2	2,708 [23.8]	1,920 [24.1]	788 [23.1]	
3	2,110 [18.6]	1,481 [18.6]	629 [18.5]	
4	1,805 [15.9]	1,288 [16.2]	517 [15.2]	
X	1,394 [12.3]	968 [12.2]	426 [12.5]	
N stage				0.743
0	7,568 [66.6]	5,317 [66.9]	2,251 [66.0]	
1	1,389 [12.2]	978 [12.3]	411 [12.1]	
2	1,565 [13.8]	1,084 [13.6]	481 [14.1]	
3	69 [0.6]	46 [0.6]	23 [0.7]	
X	771 [6.8]	528 [6.6]	243 [7.1]	
M stage				0.741
0	10,082 [88.7]	7,069 [88.9]	3,013 [88.4]	
1	573 [5.0]	396 [5.0]	177 [5.2]	
X	707 [6.2]	488 [6.1]	219 [6.4]	
Tumor size				0.547
≤1 cm	9,651 [84.9]	6,762 [85.0]	2,889 [84.7]	
>1, ≤2 cm	90 [0.8]	63 [0.8]	27 [0.8]	
>2, ≤3 cm	16 [0.1]	13 [0.2]	3 [<0.1]	
>3, ≤4 cm	22 [0.2]	19 [0.2]	3 [<0.1]	
>4, ≤5 cm	12 [0.1]	8 [0.1]	4 [0.1]	
>5 cm	4 [<0.1]	2 [<0.1]	2 [<0.1]	
Unknown	1,567 [13.8]	1,086 [13.7]	481 [14.1]	

Table 1 (continued)

Table 1 (continued)

Characteristic	Total (n=11,362), n [%]	Training dataset (n=7,953), n [%]	Validation dataset (n=3,409), n [%]	P value
Histological type				0.886
Unclassified carcinoma	172 [1.5]	116 [1.5]	56 [1.6]	
Secondary carcinoma	662 [5.8]	468 [5.9]	194 [5.7]	
Large cell carcinoma	41 [0.4]	27 [0.3]	14 [0.4]	
Undifferentiated carcinoma	45 [0.4]	32 [0.4]	13 [0.4]	
Small cell carcinoma	87 [0.8]	58 [0.7]	29 [0.9]	
Squamous cell carcinoma	1,706 [15.0]	1,189 [15.0]	517 [15.2]	
Lymphoepithelial carcinoma	86 [0.8]	60 [0.8]	26 [0.8]	
Adenocarcinoma	885 [7.8]	617 [7.8]	268 [7.9]	
Basal cell adenocarcinoma	175 [1.5]	126 [1.6]	49 [1.4]	
Adenoid cystic carcinoma	1,210 [10.6]	860 [10.8]	350 [10.3]	
Neuroendocrine carcinoma	79 [0.7]	52 [0.7]	27 [0.8]	
Oxyphilic adenocarcinoma	88 [0.8]	60 [0.8]	28 [0.8]	
Mucoepidermoid carcinoma	2,637 [23.2]	1,817 [22.8]	820 [24.1]	
Duct carcinoma	246 [2.2]	161 [2.0]	85 [2.5]	
Acinar cell carcinoma	1,380 [12.1]	980 [12.3]	400 [11.7]	
Epithelial-myoepithelial carcinoma	243 [2.1]	174 [2.2]	69 [2.0]	
Carcinoma in pleomorphic adenoma	495 [4.4]	346 [4.4]	149 [4.4]	
Malignant myoepithelioma	174 [1.5]	130 [1.6]	44 [1.3]	
Others	951 [8.4]	680 [8.6]	271 [7.9]	
Surgery				0.538
Yes	9,758 [85.9]	6,845 [86.1]	2,913 [85.5]	
No	1,500 [13.2]	1,033 [13.0]	467 [13.7]	
Unknown	104 [0.9]	75 [0.9]	29 [0.9]	
Removal of lymph nodes				0.713
Yes	6,896 [60.7]	4,826 [61]	2,070 [60.7]	
Biopsy only	155 [1.4]	115 [1.4]	40 [1.2]	
No	4,109 [36.2]	2,872 [36]	1,237 [36.3]	
Unknown	202 [1.8]	140 [1.8]	62 [1.8]	
Survival months, median [inter-quartile range]	48 [21, 89]	47 [21, 88]	49 [21, 90]	0.250
Cancer-specific death				0.561
Alive	9,024 [79]	6,305 [79]	2,719 [80]	
Dead	2,338 [21]	1,648 [21]	690 [20]	
Vital status				0.480
Alive	7,746 [68]	5,438 [68]	2,308 [68]	
Dead	3,616 [32]	2,515 [32]	1,101 [32]	

AJCC, American Joint Committee on Cancer; CI, confidence interval.

Table 2 Cox univariate and multivariate analysis of OS rates in training dataset

Characteristic	Univariate analysis			Multivariate analysis		
	HR	95% CI	P value	HR	95% CI	P value
Sex						
Female	–	–		–	–	
Male	1.97	1.81, 2.14	<0.001	1.35	1.23, 1.48	<0.001
Age						
<40 years	–	–		–	–	
40–59 years	3.43	2.68, 4.38	<0.001	2.26	1.75, 2.91	<0.001
60–80 years	7.05	5.56, 8.95	<0.001	3.50	2.72, 4.50	<0.001
>80 years	19.20	15.10, 24.50	<0.001	7.78	5.98, 10.10	<0.001
Race						
White	–	–		–	–	
Black	0.78	0.68, 0.91	0.001	1.10	0.95, 1.28	0.200
Others	0.49	0.41, 0.59	<0.001	0.72	0.60, 0.86	<0.001
Unknown	0.13	0.05, 0.34	<0.001	0.17	0.06, 0.45	<0.001
Marital status						
Divorced	–	–		–	–	
Married	0.84	0.73, 0.98	0.025	0.87	0.74, 1.01	0.061
Separated	1.06	0.70, 1.60	0.800	1.01	0.66, 1.54	>0.9
Single	0.67	0.56, 0.80	<0.001	1.08	0.90, 1.29	0.400
Widowed	2.10	1.78, 2.48	<0.001	1.26	1.06, 1.50	0.009
Unknown	0.92	0.75, 1.13	0.400	0.84	0.68, 1.04	0.110
Site						
Parotid gland	–	–		–	–	
Submandibular gland	1.13	1.01, 1.27	0.034	1.14	1.01, 1.29	0.032
Sublingual gland	0.67	0.43, 1.02	0.064	1.11	0.72, 1.72	0.600
Overlapping of major gland	1.46	0.37, 5.85	0.600	2.49	0.62, 10.00	0.200
Unspecified major gland	1.45	1.23, 1.71	<0.001	1.00	0.84, 1.19	>0.9
Grade						
I	–	–		–	–	
II	2.18	1.75, 2.71	<0.001	1.56	1.25, 1.94	<0.001
III	7.05	5.75, 8.66	<0.001	2.37	1.91, 2.95	<0.001
IV	6.10	4.88, 7.62	<0.001	2.82	2.24, 3.55	<0.001
Unknown	3.78	3.09, 4.64	<0.001	1.96	1.59, 2.42	<0.001

Table 2 (continued)

Table 2 (continued)

Characteristic	Univariate analysis			Multivariate analysis		
	HR	95% CI	P value	HR	95% CI	P value
AJCC stage						
I	–	–		–	–	
II	1.75	1.44, 2.13	<0.001	1.03	0.74, 1.41	0.900
III	4.23	3.57, 5.01	<0.001	1.36	1.02, 1.82	0.035
IV	9.77	8.38, 11.4	<0.001	1.96	1.45, 2.65	<0.001
Unknown	4.69	3.93, 5.59	<0.001	1.51	1.08, 2.11	0.015
T stage						
0	–	–		–	–	
1	0.24	0.13, 0.46	<0.001	1.11	0.57, 2.18	0.800
2	0.51	0.27, 0.95	0.035	1.66	0.87, 3.17	0.130
3	1.10	0.59, 2.05	0.800	2.10	1.11, 3.98	0.023
4	1.79	0.96, 3.34	0.068	2.20	1.16, 4.16	0.016
X	1.09	0.58, 2.03	0.800	1.27	0.65, 2.47	0.500
N stage						
0	–	–		–	–	
1	2.80	2.51, 3.12	<0.001	1.33	1.17, 1.52	<0.001
2	4.11	3.73, 4.54	<0.001	1.49	1.29, 1.72	<0.001
3	4.23	2.89, 6.20	<0.001	1.84	1.23, 2.75	0.003
X	2.86	2.50, 3.27	<0.001	1.31	1.06, 1.62	0.013
M stage						
0	–	–		–	–	
1	5.84	5.17, 6.60	<0.001	2.15	1.84, 2.50	<0.001
X	1.89	1.65, 2.17	<0.001	1.12	0.90, 1.40	0.300
Tumor size						
≤1 cm	–	–		–	–	
>1, ≤2 cm	1.90	1.31, 2.74	<0.001	1.74	1.20, 2.52	0.004
>2, ≤3 cm	1.01	0.38, 2.69	>0.9	0.93	0.35, 2.50	0.900
>3, ≤4 cm	0.92	0.38, 2.21	0.900	0.81	0.33, 1.97	0.600
>4, ≤5 cm	2.06	0.86, 4.96	0.110	1.19	0.48, 2.95	0.700
>5 cm	2.10	0.30, 14.9	0.500	0.60	0.08, 4.28	0.600
Unknown	1.74	1.57, 1.92	<0.001	0.98	0.84, 1.15	0.800

Table 2 (continued)

Table 2 (continued)

Characteristic	Univariate analysis			Multivariate analysis		
	HR	95% CI	P value	HR	95% CI	P value
Histological type						
Unclassified carcinoma	–	–	–	–	–	–
Secondary carcinoma	0.85	0.65, 1.11	0.200	0.81	0.61, 1.08	0.200
Large cell carcinoma	1.08	0.64, 1.79	0.800	0.68	0.40, 1.15	0.150
Undifferentiated carcinoma	0.34	0.17, 0.69	0.003	0.52	0.25, 1.04	0.066
Small cell carcinoma	0.82	0.54, 1.25	0.400	0.79	0.51, 1.21	0.300
Squamous cell carcinoma	0.83	0.65, 1.08	0.200	0.88	0.67, 1.15	0.300
Lymphoepithelial carcinoma	0.16	0.08, 0.33	<0.001	0.29	0.14, 0.60	<0.001
Adenocarcinoma	0.58	0.45, 0.77	<0.001	0.76	0.57, 1.02	0.066
Basal cell adenocarcinoma	0.16	0.10, 0.26	<0.001	0.34	0.20, 0.57	<0.001
Adenoid cystic carcinoma	0.29	0.22, 0.38	<0.001	0.69	0.51, 0.93	0.013
Neuroendocrine carcinoma	0.69	0.44, 1.09	0.110	0.55	0.34, 0.88	0.013
Oxyphilic adenocarcinoma	0.55	0.35, 0.86	0.009	0.96	0.60, 1.54	0.900
Mucoepidermoid carcinoma	0.17	0.13, 0.23	<0.001	0.58	0.44, 0.78	<0.001
Duct carcinoma	0.46	0.32, 0.65	<0.001	0.54	0.37, 0.79	0.002
Acinar cell carcinoma	0.11	0.08, 0.15	<0.001	0.49	0.35, 0.68	<0.001
Epithelial-myoepithelial carcinoma	0.10	0.06, 0.17	<0.001	0.30	0.17, 0.53	<0.001
Carcinoma in pleomorphic adenoma	0.31	0.23, 0.43	<0.001	0.61	0.44, 0.86	0.004
Malignant myoepithelioma	0.25	0.16, 0.39	<0.001	0.50	0.31, 0.78	0.003
Others	0.62	0.47, 0.81	<0.001	0.92	0.69, 1.22	0.500
Surgery						
Yes	–	–	–	–	–	–
No	4.40	4.03, 4.82	<0.001	1.78	1.56, 2.04	<0.001
Unknown	3.25	2.38, 4.43	<0.001	1.50	1.06, 2.12	0.022
Removal of lymph nodes						
Yes	–	–	–	–	–	–
Biopsy only	3.44	2.69, 4.40	<0.001	1.41	1.08, 1.84	0.012
No	1.33	1.23, 1.44	<0.001	1.30	1.16, 1.45	<0.001
Unknown	2.27	1.80, 2.86	<0.001	1.84	1.40, 2.40	<0.001

OS, overall survival; HR, hazard ratio; CI, confidence interval.

Table 3 Cox univariate and multivariate analysis of CSS rates in training dataset

Characteristic	Univariate analysis			Multivariate analysis		
	HR	95% CI	P value	HR	95% CI	P value
Sex						
Female	–	–	–	–	–	–
Male	1.81	1.64, 2.01	<0.001	1.16	1.03, 1.30	0.011
Age						
<40 years	–	–	–	–	–	–
40–59 years	2.95	2.26, 3.85	<0.001	1.71	1.30, 2.25	<0.001
60–80 years	5.00	3.87, 6.47	<0.001	2.10	1.60, 2.76	<0.001
>80 years	10.6	8.14, 13.90	<0.001	3.77	2.82, 5.05	<0.001
Race						
White	–	–	–	–	–	–
Black	0.90	0.76, 1.07	0.200	1.13	0.94, 1.35	0.200
Others	0.54	0.43, 0.67	<0.001	0.73	0.58, 0.91	0.005
Unknown	0.15	0.05, 0.46	<0.001	0.20	0.06, 0.62	0.005
Marital status						
Divorced	–	–	–	–	–	–
Married	0.81	0.68, 0.97	0.022	0.89	0.74, 1.07	0.200
Separated	1.16	0.72, 1.87	0.500	1.04	0.64, 1.69	0.900
Single	0.71	0.58, 0.87	0.001	1.04	0.84, 1.29	0.700
Widowed	1.78	1.45, 2.18	<0.001	1.26	1.01, 1.56	0.038
Unknown	0.78	0.60, 1.01	0.057	0.76	0.58, 0.99	0.043
Site						
Parotid gland	–	–	–	–	–	–
Submandibular gland	1.31	1.14, 1.50	<0.001	1.24	1.07, 1.43	0.003
Sublingual gland	0.70	0.42, 1.19	0.200	1.06	0.62, 1.82	0.800
Overlapping of major gland	1.17	0.17, 8.35	0.900	1.95	0.27, 13.9	0.500
Unspecified major gland	1.67	1.38, 2.02	<0.001	1.04	0.85, 1.29	0.700
Grade						
I	–	–	–	–	–	–
II	4.50	3.08, 6.58	<0.001	3.13	2.13, 4.59	<0.001
III	16.50	11.40, 23.80	<0.001	5.06	3.47, 7.37	<0.001
IV	14.80	10.10, 21.60	<0.001	6.15	4.17, 9.07	<0.001
Unknown	8.45	5.86, 12.20	<0.001	3.76	2.59, 5.47	<0.001

Table 3 (continued)

Table 3 (continued)

Characteristic	Univariate analysis			Multivariate analysis		
	HR	95% CI	P value	HR	95% CI	P value
AJCC stage						
I	–	–	–	–	–	–
II	1.93	1.43, 2.61	<0.001	0.99	0.64, 1.53	>0.9
III	6.11	4.73, 7.89	<0.001	1.69	1.15, 2.50	0.008
IV	17.60	13.90, 22.20	<0.001	2.67	1.80, 3.97	<0.001
Unknown	7.28	5.60, 9.46	<0.001	1.61	1.03, 2.51	0.035
T stage						
0	–	–	–	–	–	–
1	0.24	0.10, 0.60	0.002	1.81	0.71, 4.60	0.200
2	0.65	0.27, 1.58	0.30	3.12	1.26, 7.72	0.014
3	1.49	0.62, 3.61	0.400	3.66	1.49, 9.01	0.005
4	2.69	1.11, 6.48	0.028	3.96	1.62, 9.71	0.003
X	1.61	0.66, 3.90	0.300	2.55	1.01, 6.38	0.046
N stage						
0	–	–	–	–	–	–
1	3.73	3.27, 4.27	<0.001	1.60	1.37, 1.87	<0.001
2	5.75	5.10, 6.47	<0.001	1.70	1.44, 2.01	<0.001
3	7.01	4.70, 10.50	<0.001	2.37	1.55, 3.62	<0.001
X	3.77	3.19, 4.44	<0.001	1.38	1.07, 1.79	0.012
M stage						
0	–	–	–	–	–	–
1	7.95	6.97, 9.08	<0.001	2.38	2.01, 2.81	<0.001
X	2.23	1.90, 2.63	<0.001	1.46	1.12, 1.89	0.005
Tumor size						
≤1 cm	–	–	–	–	–	–
>1, ≤2 cm	2.05	1.33, 3.16	0.001	1.66	1.07, 2.58	0.025
>2, ≤3 cm	1.19	0.38, 3.71	0.800	0.92	0.29, 2.88	0.900
>3, ≤4 cm	1.11	0.42, 2.97	0.800	0.89	0.33, 2.42	0.800
>4, ≤5 cm	1.26	0.31, 5.03	0.700	0.72	0.17, 2.96	0.600
>5 cm	3.39	0.48, 24.1	0.200	1.12	0.16, 8.02	>0.9
Unknown	1.86	1.65, 2.10	<0.001	0.93	0.77, 1.12	0.400

Table 3 (continued)

Table 3 (continued)

Characteristic	Univariate analysis			Multivariate analysis		
	HR	95% CI	P value	HR	95% CI	P value
Histological type						
Unclassified carcinoma	–	–	–	–	–	–
Secondary carcinoma	0.85	0.61, 1.18	0.300	0.72	0.51, 1.02	0.067
Large cell carcinoma	0.92	0.47, 1.78	0.800	0.55	0.28, 1.09	0.086
Undifferentiated carcinoma	0.23	0.08, 0.63	0.004	0.30	0.11, 0.85	0.024
Small cell carcinoma	0.67	0.39, 1.17	0.200	0.63	0.36, 1.11	0.110
Squamous cell carcinoma	0.75	0.55, 1.03	0.074	0.79	0.57, 1.10	0.200
Lymphoepithelial carcinoma	0.17	0.07, 0.39	<0.001	0.28	0.12, 0.66	0.004
Adenocarcinoma	0.72	0.52, 1.00	0.050	0.83	0.59, 1.17	0.300
Basal cell adenocarcinoma	0.11	0.06, 0.23	<0.001	0.29	0.14, 0.60	<0.001
Adenoid cystic carcinoma	0.33	0.24, 0.46	<0.001	0.71	0.50, 1.02	0.065
Neuroendocrine carcinoma	0.76	0.45, 1.31	0.300	0.53	0.30, 0.92	0.025
Oxyphilic adenocarcinoma	0.43	0.24, 0.79	0.006	0.76	0.41, 1.41	0.400
Mucoepidermoid carcinoma	0.16	0.11, 0.22	<0.001	0.53	0.37, 0.75	<0.001
Duct carcinoma	0.51	0.34, 0.78	0.002	0.57	0.36, 0.89	0.013
Acinar cell carcinoma	0.09	0.06, 0.14	<0.001	0.46	0.30, 0.70	<0.001
Epithelial-myoepithelial carcinoma	0.05	0.02, 0.12	<0.001	0.16	0.06, 0.42	<0.001
Carcinoma in pleomorphic adenoma	0.34	0.23, 0.49	<0.001	0.64	0.42, 0.95	0.029
Malignant myoepithelioma	0.21	0.12, 0.38	<0.001	0.46	0.26, 0.83	0.010
Others	0.56	0.40, 0.78	<0.001	0.82	0.58, 1.15	0.200
Surgery						
Yes	–	–	–	–	–	–
No	5.02	4.52, 5.59	<0.001	1.98	1.67, 2.33	<0.001
Unknown	3.90	2.74, 5.54	<0.001	1.72	1.15, 2.57	0.008
Removal of lymph nodes						
Yes	–	–	–	–	–	–
Biopsy only	3.91	2.97, 5.15	<0.001	1.27	0.94, 1.73	0.130
No	1.29	1.16, 1.42	<0.001	1.26	1.09, 1.46	0.002
Unknown	2.59	1.97, 3.39	<0.001	1.85	1.34, 2.56	<0.001

CSS, cancer-specific survival; HR, hazard ratio; CI, confidence interval.

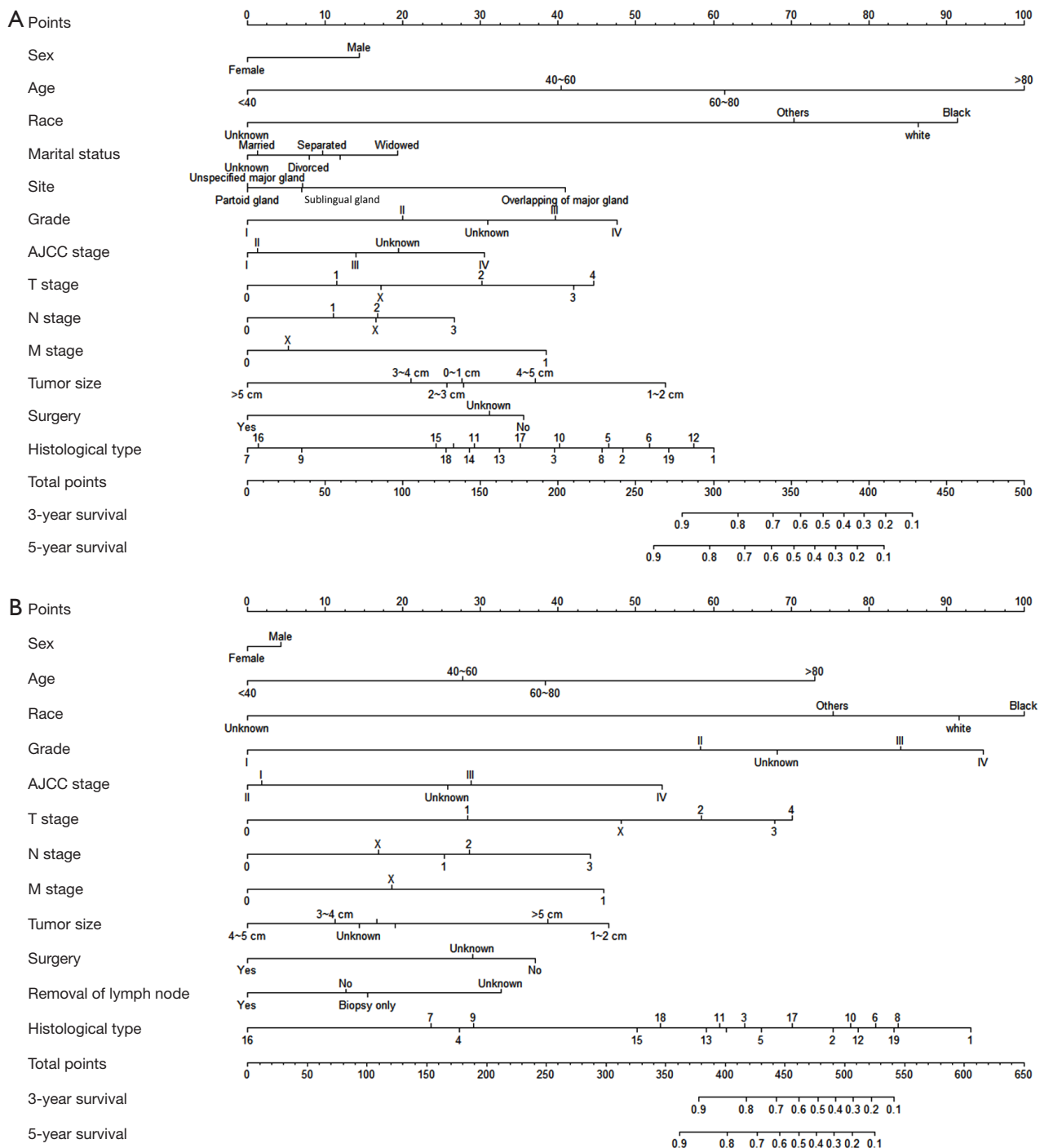


Figure 2 Nomograms for prediction of 3-year and 5-year OS (A) and CSS (B) rate of MSGC patients. For histological type, the number 1 to 19 represents: (I) unclassified carcinoma; (II) secondary carcinoma; (III) large cell carcinoma; (IV) undifferentiated carcinoma; (V) small cell carcinoma; (VI) squamous cell carcinoma; (VII) lymphoepithelial carcinoma; (VIII) adenocarcinoma; (IX) basal cell adenocarcinoma; (X) adenoid cystic carcinoma; (XI) neuroendocrine carcinoma; (XII) oxyphilic adenocarcinoma; (XIII) mucoepidermoid carcinoma; (XIV) duct carcinoma; (XV) acinar cell carcinoma; (XVI) epithelial-myoeplithelial carcinoma; (XVII) carcinoma in pleomorphic adenoma; (XVIII) malignant myoeplithelioma; (XIX) others. AJCC, American Joint Committee on Cancer; OS, overall survival; CSS, cancer-specific survival; MSGC, major salivary glands carcinoma.

Table 4 Comparison of AUC, C-index, and Brier score for 5-year prediction between the nomogram, TNM stage, and AJCC stage in MSGC patients

Survival types	Training dataset					Validation dataset				
	AUC	95% CI	C-index	95% CI	Brier score	AUC	95% CI	C-index	95% CI	Brier score
OS										
Nomogram	83.5	82.4–84.7	0.796	0.788–0.804	0.0148	82.7	81.0–84.3	0.787	0.774–0.799	0.0153
TNM stage	72.7	71.3–74.1	0.724	0.713–0.734	0.0188	71.0	68.9–73.2	0.711	0.695–0.726	0.0194
AJCC stage	71.6	70.2–73.0	0.751	0.739–0.762	0.0187	69.8	67.7–71.9	0.730	0.712–0.748	0.0194
CSS										
Nomogram	83.9	82.6–85.2	0.806	0.797–0.815	0.0128	82.8	80.9–84.7	0.798	0.784–0.812	0.0130
TNM stage	75.9	74.4–77.3	0.753	0.742–0.764	0.0153	74.5	72.2–76.7	0.742	0.725–0.759	0.0155
AJCC stage	73.9	72.4–75.3	0.781	0.769–0.793	0.0160	72.6	70.4–74.8	0.769	0.750–0.788	0.0160

MSGC, major salivary gland carcinoma; OS, overall survival; CSS, cancer-specific survival; AJCC, American Joint Committee on Cancer; AUC, area under curve; CI, confidence interval.

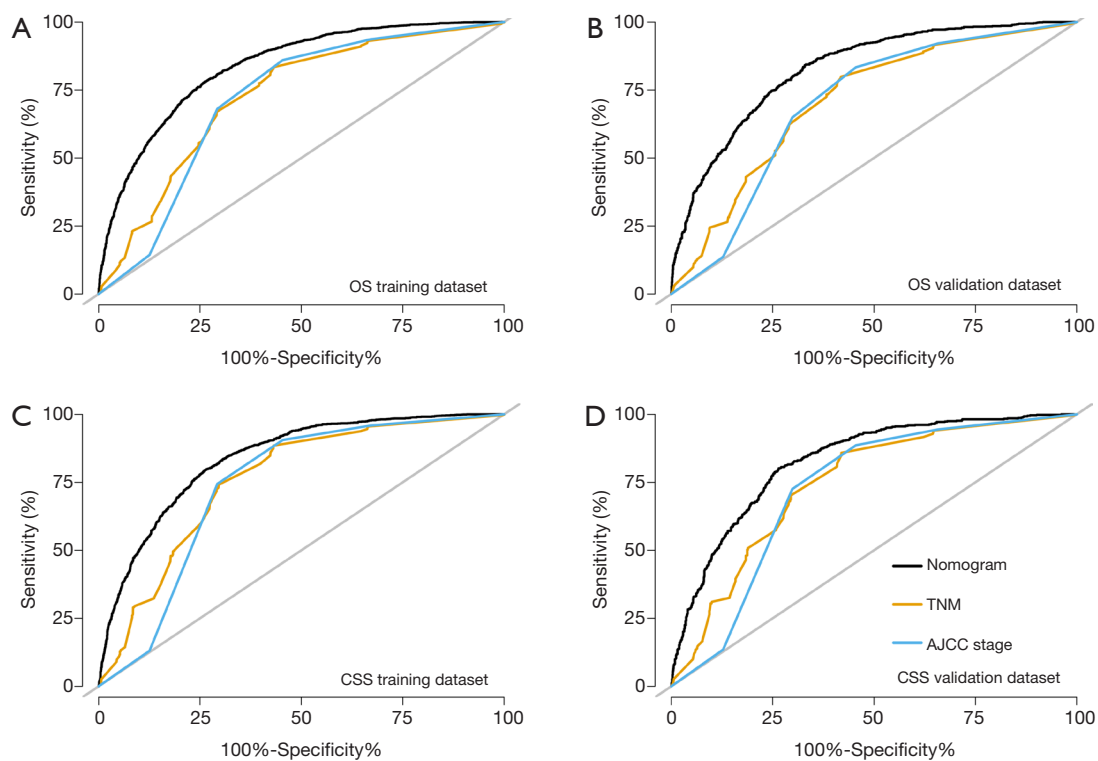


Figure 3 The ROC curves of nomograms, AJCC stage, and TNM stage for 5-year prediction in training and validation dataset. (A,B) OS in the training dataset (A) and validation dataset (B). (C,D) CSS in the training dataset (C) and validation dataset (D). ROC, receiver operating characteristic; AJCC, American Joint Committee on Cancer; OS, overall survival; CSS, cancer-specific survival.

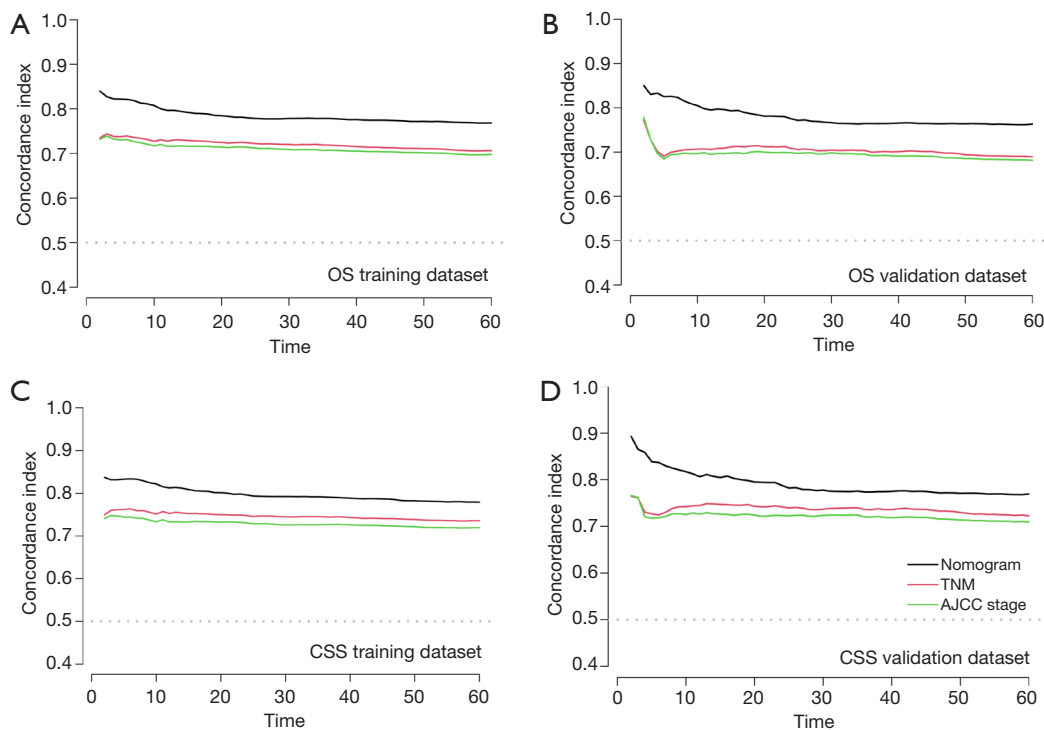


Figure 4 The overtime C-index of nomograms, AJCC stage, and TNM stage for predicting OS (A,B) and CSS (C,D) in training and validation dataset. C-index, Harrell C-index; AJCC, American Joint Committee on Cancer; OS, overall survival; CSS, cancer-specific survival.

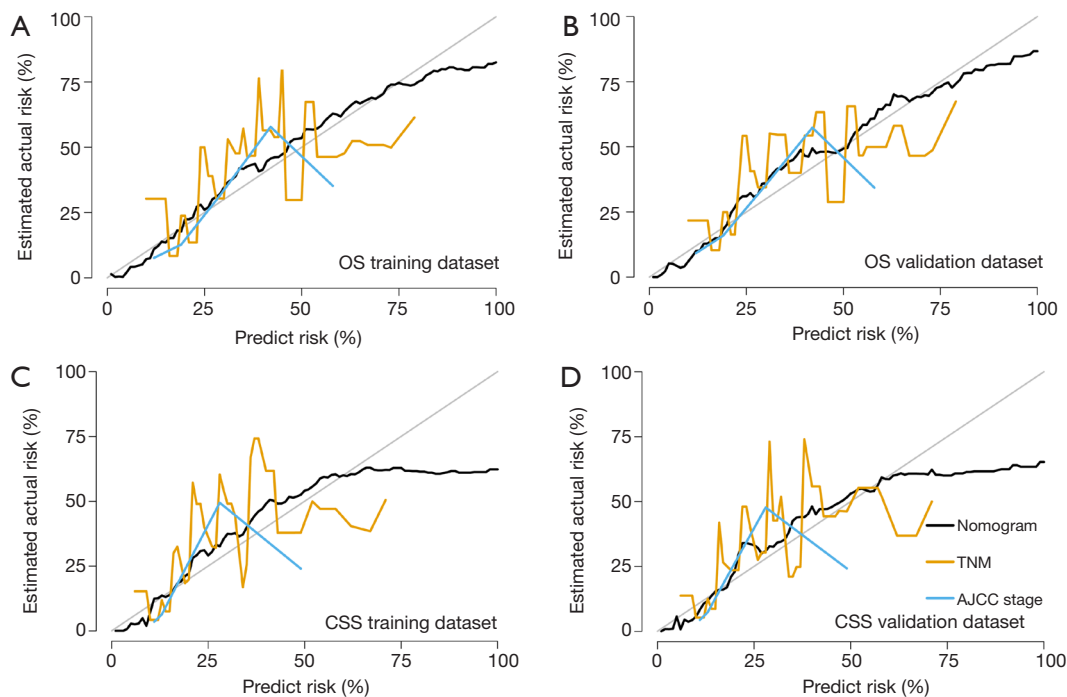


Figure 5 The calibration curves of nomograms, AJCC stage, and TNM stage for 5-year prediction in training and validation dataset. (A,B) OS in the training dataset (A) and validation dataset (B). (C,D) CSS in the training dataset (C) and validation dataset (D). AJCC, American Joint Committee on Cancer; OS, overall survival; CSS, cancer-specific survival.

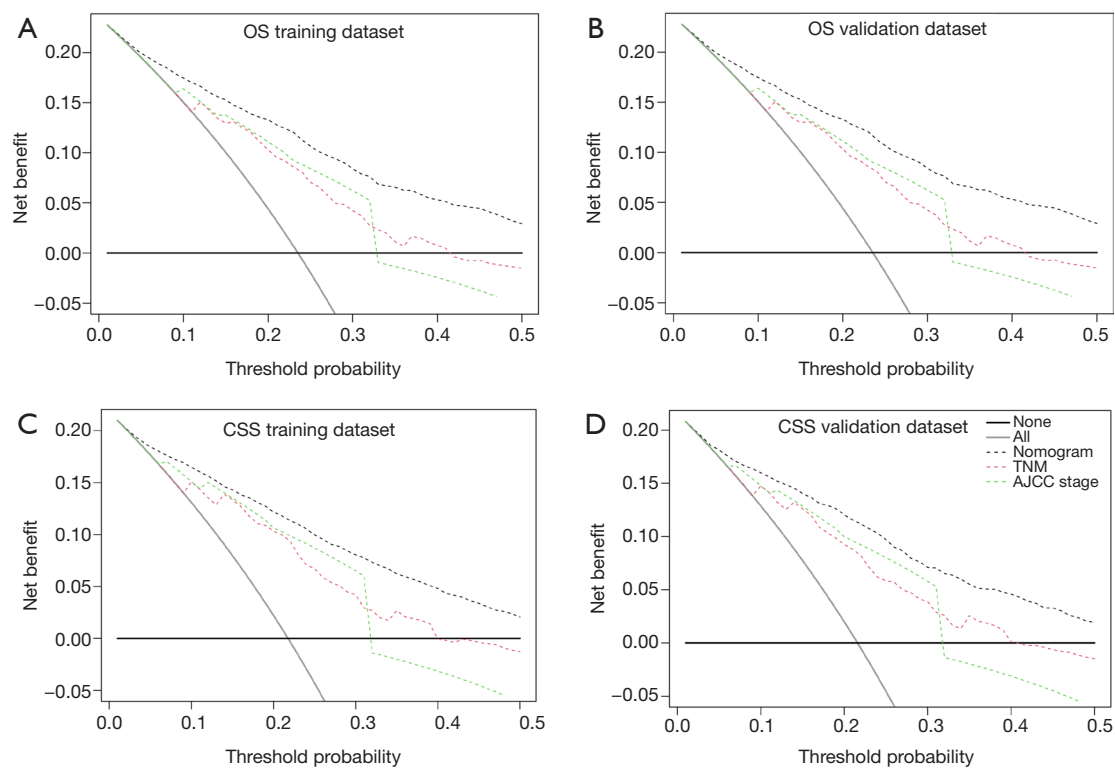


Figure 6 The DCA of nomograms, AJCC stage, and TNM stage for 5-year prediction in training and validation dataset. (A,B) OS in the training dataset (A) and validation dataset (B). (C,D) CSS in the training dataset (C) and validation dataset (D). DCA, decision curve analysis; AJCC, American Joint Committee on Cancer; OS, overall survival; CSS, cancer-specific survival.

Discussion

The incidence rate of MSGC ranges from 1.1 to 1.3/100,000/year, which is quite low (21,22). The limited sample size greatly impacts the related research, as it increases the research error and may lead to a biased result. SEER database is an authoritative source of information on cancer incidence, diagnosis, therapy, and survival in the United States. This database collects and publishes cancer incidence and survival data from eighteen cancer registries covering approximately 35% of the U.S. population. The SEER database has been widely adopted in cancer-related research (23-25). In this study, a large cohort of patients ($n=14,753$) diagnosed with MSGC was recruited from SEER. To the best of our knowledge, this is the largest sample size used so far in the research on the prognosis of MSGC.

Given the morphologic diversity and rarity of MSGC, there are no high-quality randomized data defining predictive factors. In a relatively large cohort, site, histological type, race, sex, and age at diagnosis were identified as predictors for MSGC OS (26). Li *et al.* failed to find a prognostic,

predictive factor for lymphoepithelial carcinoma in the major salivary gland in a cohort containing 37 patients (27). In a study comprising 95 patients recruited from a single institution, male sex and >10 pack-year smoking history, grade, N stage, squamous cell carcinoma, perineural invasion, and neck dissection significantly reduced OS (28). In the present study, multivariate Cox analysis identified sex, age, race, grade, AJCC stage, T/N/M stage, tumor size, surgery, histological type, and site as significant factors predicting OS. All predictors, except for marital status and M stage, were significant in multivariate Cox analysis for CSS.

Sex was proven to be an independent risk factor of MSGC survival. Females have been reported to have better survival outcomes compared with males, which is consistent with our result (26,29). Not surprisingly, the risk increases with the patients' age. As for the race, black people were at the highest risk, followed by white, and then other races, which was consistent with some previous studies (30-32). It is important to note that this may be due to the socioeconomic differences between black people and other races in some areas (33). Poor differentiation grade and

high AJCC stage are two independent indicators for worse survival outcomes. T stage and tumor size reflected the condition of the primary site. Interestingly, the patient with a larger primary lesion might have better survival compared to a smaller lesion, which was contradictory to the results of other studies (29). On the one hand, the larger tumor was often associated with better differentiation grade and histological type, which led to better survival outcomes. Patients who suffered from highly malignant tumors usually died before their tumor grew larger. On the other hand, this may be due to the small sample size with large tumors (>1 cm) in this cohort (n=1,171, 15.1%). Whether in univariate, multivariate Cox regression, or LASSO model, tumor size was significantly related to the OS and CSS; thus, it was incorporated in the final model. M stage, which reflects distant tumor metastasis, was one of the most important predictors for cancer survival (34). In the present study, the M stage showed a significant association with OS but not CSS (P>0.05). The important role of distant metastasis in MSGC has been well-recognized (35,36). Considering the clinical significance, the M stage was finally included. Surgery was the primary approach for MSGC management. Patients who underwent surgery had better survival compared to those who did not. In the 2017 World Health Organization classification, there were over 20 different histological subtypes of MSGC, with specific features and outcomes. Consequently, it is important to consider the histological type for prediction survival of MSGC. In the present study, when the number of patients diagnosed as a specific subtype was less than 10, it was put into the category of "Other". Epithelial-myoepithelial carcinoma and lymphoepithelial carcinoma had better survival, while patients diagnosed with unclassified carcinoma had worse OS and CSS. Radiotherapy and chemotherapy were not incorporated in the model. For radiotherapy and chemotherapy, patients were divided into two groups—"Yes" and "No/Unknown" in the SEER, while "Unknown" was set as a separated group for other variables, such as race, marital status, grade, AJCC stage. After discussion, we thought that combining "No" and "Unknown" as a group would reduce the effect of these two variables and might lead to confusion in the clinical practice.

The key strengths in the present study are the following: first, the SEER database provided a total of 11,362 MSGC patients. Compared with a study by Li *et al.*, the sample size in the present study was larger (4,218 vs. 14,753). This cohort, which to the best of our knowledge is the largest one used thus far, is large enough to ensure the reliability and effectiveness of the nomogram. Second, the histological

type is considered in the nomogram. The Memorial Sloan Kettering Cancer Center (MSKCC) a salivary carcinoma nomogram, is a common tool used for MSGC survival prediction (12,13,37). Still, unlike the histological type, the only histological grade was considered for the development of the MSKCC nomogram. As an important factor, the survival outcomes in different histological types greatly vary. Third, treatment planning and prognosis prediction are currently dependent on the AJCC stage and TNM stage systems to a great extent. To evaluate the value of using the nomogram, the performance of the nomogram was compared with the AJCC stage and TNM stage. C-index, AUC, Brier score was calculated. Calibration curves, ROC curves, and DCA were also drawn. The result showed that the performance of the nomogram was consistently better compared to the AJCC stage and TNM stage. Finally, even though Li *et al.* and our nomograms were constructed based on the SEER database, only the present study was conducted according to the TRIPOD statement (30). Moreover, the performance of the nomogram was evaluated all-around to ensure the reliability of the prediction tool.

The present study also has some limitations. The variables extracted from the SEER database were limited. Records related to prognosis, such as lymphovascular invasion, perineural invasion, tobacco/alcohol use, were inaccessible. In this cohort, the seventh edition of the AJCC stage was adopted as only a small portion of patients had information about the eighth edition of the AJCC stage. As a result, the extranodal extension of lymph nodes was not considered in the present study, which is a common limitation in research collecting information from SEER. Moreover, target therapy has an increasingly important role in MSGC treatment (38). Once the patient accepts target therapy, the survival will certainly change. Furthermore, different drugs can also have different roles. Due to the limitation of SEER, the information on target therapy remains unknown. Although the nomogram performed well in internal validation, external validation is needed in future studies.

Conclusions

A reliable nomogram for OS and CSS prediction in MSGC patients based on the Cox regression model was developed using the SEER database. The nomogram showed good predictive discrimination and accuracy compared with the AJCC stage and TNM stage. Consequently, the nomogram can become a practical tool to help clinicians to make a treatment plan and prognosis predictions.

Acknowledgments

The authors would like to thank the SEER program for open access to the database.

Funding: This work was supported by grants from the Natural Science Fund of China (grant No. 81901049), Outstanding Young Scholars of Shanghai Health Commission (2018YQ34), and Young Scholars and Innovative Team of Shanghai Ninth People's Hospital (QC201901).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at <https://dx.doi.org/10.21037/atm-21-1725>

Peer Review File: Available at <https://dx.doi.org/10.21037/atm-21-1725>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/atm-21-1725>). The 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). Since individual information has been removed from all the SEER databases, informed consent from patients and approval by the institutional review board were exempted.

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. Boukheris H, Curtis RE, Land CE, et al. Incidence of carcinoma of the major salivary glands according to the WHO classification, 1992 to 2006: a population-based study in the United States. *Cancer Epidemiol Biomarkers Prev* 2009;18:2899-906.
2. McKenna RJ. Tumors of the major and minor salivary glands. *CA Cancer J Clin* 1984;34:24-39.
3. Andry G, Hamoir M, Locati LD, et al. Management of salivary gland tumors. *Expert Rev Anticancer Ther* 2012;12:1161-8.
4. Cruz A, Magalhães H, Pereira FF, et al. A 10-year review of primary major salivary gland cancers. *Ecancermedicalscience* 2020;14:1055.
5. Egger EK, Stope MB, Recker F, et al. Lower Genital Tract Melanomas: Staging, Predictors of Outcome, and New Therapeutic Options. *Anticancer Res* 2021;41:999-1004.
6. Pu N, Li J, Xu Y, et al. Comparison of prognostic prediction between nomogram based on lymph node ratio and AJCC 8th staging system for patients with resected pancreatic head carcinoma: a SEER analysis. *Cancer Manag Res* 2018;10:227-38.
7. Lombardi D, Tomasoni M, Paderno A, et al. The impact of nodal status in major salivary gland carcinoma: A multicenter experience and proposal of a novel N-classification. *Oral Oncol* 2021;112:105076.
8. Peng F, Li J, Mu S, et al. Epidemiological features of primary breast lymphoma patients and development of a nomogram to predict survival. *Breast* 2021;57:49-61.
9. Liu C, Hu C, Huang J, et al. A Prognostic Nomogram of Colon Cancer With Liver Metastasis: A Study of the US SEER Database and a Chinese Cohort. *Front Oncol* 2021;11:591009.
10. Lin G, Qi K, Liu B, et al. A nomogram prognostic model for large cell lung cancer: analysis from the Surveillance, Epidemiology and End Results Database. *Transl Lung Cancer Res* 2021;10:622-35.
11. Lukovic J, Alfaraj FA, Mierzwa ML, et al. Development and validation of a clinical prediction-score model for distant metastases in major salivary gland carcinoma. *Ann Oncol* 2020;31:295-301.
12. Ali S, Palmer FL, Yu C, et al. Postoperative nomograms predictive of survival after surgical management of malignant tumors of the major salivary glands. *Ann Surg Oncol* 2014;21:637-42.
13. Ali S, Palmer FL, Yu C, et al. A predictive nomogram for recurrence of carcinoma of the major salivary glands. *JAMA Otolaryngol Head Neck Surg* 2013;139:698-705.
14. Li Y, Ju J, Liu X, et al. Nomograms for predicting long-term overall survival and cancer-specific survival in patients with major salivary gland cancer: a population-based study. *Oncotarget* 2017;8:24469-82.
15. Collins GS, Reitsma JB, Altman DG, et al. Transparent

- Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* 2015;162:55-63.
16. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med* 1997;16:385-95.
 17. Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Stat Med* 2004;23:2109-23.
 18. Huitzil-Melendez FD, Capanu M, O'Reilly EM, et al. Advanced hepatocellular carcinoma: which staging systems best predict prognosis? *J Clin Oncol* 2010;28:2889-95.
 19. Graf E, Schmoor C, Sauerbrei W, et al. Assessment and comparison of prognostic classification schemes for survival data. *Stat Med* 1999;18:2529-45.
 20. Vickers AJ, Cronin AM, Elkin EB, et al. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. *BMC Med Inform Decis Mak* 2008;8:53.
 21. Pinkston JA, Cole P. Incidence rates of salivary gland tumors: results from a population-based study. *Otolaryngol Head Neck Surg* 1999;120:834-40.
 22. Spiro RH. Salivary neoplasms: overview of a 35-year experience with 2,807 patients. *Head Neck Surg* 1986;8:177-84.
 23. Dess RT, Suresh K, Zelefsky MJ, et al. Development and Validation of a Clinical Prognostic Stage Group System for Nonmetastatic Prostate Cancer Using Disease-Specific Mortality Results From the International Staging Collaboration for Cancer of the Prostate. *JAMA Oncol* 2020;6:1912-20.
 24. Kurian AW, Ward KC, Abrahamse P, et al. Time Trends in Receipt of Germline Genetic Testing and Results for Women Diagnosed With Breast Cancer or Ovarian Cancer, 2012-2019. *J Clin Oncol* 2021;39:1631-40.
 25. Sung H, Hyun N, Leach CR, et al. Association of First Primary Cancer With Risk of Subsequent Primary Cancer Among Survivors of Adult-Onset Cancers in the United States. *JAMA* 2020;324:2521-35.
 26. Spitz MR, Batsakis JG. Major salivary gland carcinoma. Descriptive epidemiology and survival of 498 patients. *Arch Otolaryngol* 1984;110:45-9.
 27. Li F, Zhu G, Wang Y, et al. A clinical analysis of 37 cases with lymphoepithelial carcinoma of the major salivary gland treated by surgical resection and postoperative radiotherapy: a single institution study. *Med Oncol* 2014;31:957.
 28. Gutschenritter T, Machiorlatti M, Vesely S, et al. Outcomes and Prognostic Factors of Resected Salivary Gland Malignancies: Examining a Single Institution's 12-year Experience. *Anticancer Res* 2017;37:5019-25.
 29. Taylor ZC, Kaya EA, Bunn JD, et al. Overall and cause-specific survival for mucoepidermoid carcinoma of the major salivary glands: Analysis of 2210 patients. *World J Clin Oncol* 2020;11:1029-44.
 30. Li Y, Zhao Z, Liu X, et al. Nomograms to estimate long-term overall survival and tongue cancer-specific survival of patients with tongue squamous cell carcinoma. *Cancer Med* 2017;6:1002-13.
 31. Wang F, Zhang H, Wen J, et al. Nomograms forecasting long-term overall and cancer-specific survival of patients with oral squamous cell carcinoma. *Cancer Med* 2018;7:943-52.
 32. Booth JR, Unsal AA, Tadros S, et al. Salivary Gland Cancers of the Nasopharynx: A Population-Based Analysis of 383 Cases. *Otolaryngol Head Neck Surg* 2019;161:442-9.
 33. Weng Y, Korte JE. Racial disparities in being recommended to surgery for oral and oropharyngeal cancer in the United States. *Community Dent Oral Epidemiol* 2012;40:80-8.
 34. Gennari A, Conte P, Rosso R, et al. Survival of metastatic breast carcinoma patients over a 20-year period: a retrospective analysis based on individual patient data from six consecutive studies. *Cancer* 2005;104:1742-50.
 35. Fushimi C, Tada Y, Takahashi H, et al. A prospective phase II study of combined androgen blockade in patients with androgen receptor-positive metastatic or locally advanced unresectable salivary gland carcinoma. *Ann Oncol* 2018;29:979-84.
 36. Locati LD, Galbiati D, Calareso G, et al. Patients with adenoid cystic carcinomas of the salivary glands treated with lenvatinib: Activity and quality of life. *Cancer* 2020;126:1888-94.
 37. Hay A, Migliacci J, Zanoni DK, et al. Validation of nomograms for overall survival, cancer-specific survival, and recurrence in carcinoma of the major salivary glands. *Head Neck* 2018;40:1008-15.
 38. Uijen MJM, Lassche G, van Engen-van Grunsven ACH, et al. Systemic therapy in the management of recurrent or metastatic salivary duct carcinoma: A systematic review. *Cancer Treat Rev* 2020;89:102069.

Cite this article as: Guo Z, Wang Z, Liu Y, Han J, Liu J, Zhang C. Nomograms-based prediction of overall and cancer-specific survivals for patients diagnosed with major salivary gland carcinoma. *Ann Transl Med* 2021;9(15):1230. doi: 10.21037/atm-21-1725

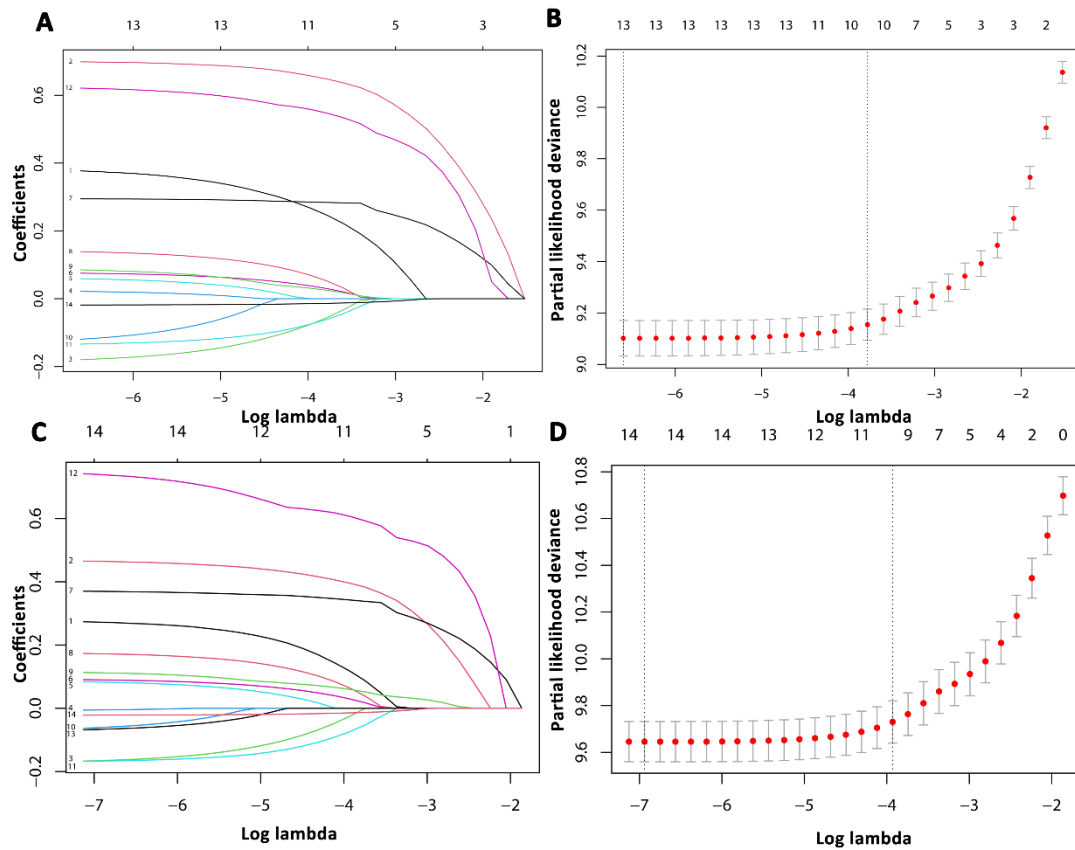


Figure S1 Prognostic factor selection using LASSO regression model. The LASSO coefficient profiles of the 14 prognostic factors for OS (A) and CSS (C). The coefficient profile plots were produced against the log lambda. Numbers represent prognostic factors, i.e., (I) sex; (II) surgery; (III) removal of lymph nodes; (IV) marital status; (V) M stage; (VI) race; (VII) age; (VIII) site; (IX) differentiation grade; (X) histological type; (XI) tumor size; (XII) AJCC stage; (XIII) N stage; (XIV) T stage. Tuning parameter (Lambda) selection in the LASSO model used tenfold cross-validation for OS (B) and CSS (D). The partial likelihood deviance curve was plotted versus log lambda. The two vertical dashed lines represent the minimum value and one standard deviation from the minimum value. The minimum value of lambda, 0.00136 and 0.00080, was chosen for OS and CSS, respectively. LASSO, least absolute shrinkage, and selection operator; OS, overall survival; CSS, cancer-specific survival; AJCC, American Joint Committee on Cancer.

Table S1 Coefficients of each predictor returned by Cox regression

Variables	Exp(coef)	Exp(-coef)	Lower 95	Upper 95
OS				
Sex	1.4745	0.6782	1.3525	1.6074
Age	2.0265	0.4935	1.9218	2.1369
Race	0.8267	1.2097	0.7691	0.8886
Marital status	1.0254	0.9752	0.9988	1.0528
Site	1.0664	0.9377	1.0257	1.1088
Grade	1.0821	0.9241	1.0498	1.1154
AJCC stage	1.3433	0.7444	1.2639	1.4278
T stage	1.1565	0.8646	1.0976	1.2186
N stage	1.0956	0.9127	1.0512	1.1418
M stage	0.8740	1.1442	0.8060	0.9477
Tumor size	0.8704	1.1489	0.8517	0.8895
Surgery	1.8844	0.5307	1.7169	2.0682
Histological type	0.9813	1.0190	0.9735	0.9893
CSS				
Sex	1.3085	0.7642	1.1787	1.4526
Age	1.5929	0.6278	1.4960	1.6960
Race	0.8397	1.1909	0.7710	0.9146
Grade	1.1009	0.9084	1.0596	1.1438
AJCC stage	1.4514	0.6890	1.3453	1.5658
T stage	1.1952	0.8367	1.1233	1.2717
N stage	1.1260	0.8881	1.0721	1.1826
M stage	0.9311	1.0740	0.8467	1.0239
Tumor size	0.8445	1.1841	0.8228	0.8668
Surgery	2.1491	0.4653	1.8977	2.4337
Removal of lymph nodes	0.9329	1.0719	0.8784	0.9908
Histological type	0.9793	1.0211	0.9697	0.9891

OS, overall survival; CSS, cancer-specific survival; AJCC, American Joint Committee on Cancer.

Table S2 Comparison of AUC, C-index, and Brier score for 3-year prediction between the nomogram, TNM stage, and AJCC stage in MSGC patients

Survival types	Training dataset			Validation dataset		
	AUC	95% CI	Brier score	AUC	95% CI	Brier score
OS						
Nomogram	84.1	83.1–85.1	0.0131	81.8	80.2–83.5	0.0138
TNM stage	74.3	73.0–75.6	0.0161	72.0	70.0–74.0	0.0165
AJCC stage	72.7	71.5–74.0	0.0413	70.8	68.9–72.8	0.0165
CSS						
Nomogram	84.6	83.5–85.7	0.0114	82.3	80.5–84.2	0.0116
TNM stage	77.0	75.7–78.3	0.0133	75.7	73.6–77.8	0.0131
AJCC stage	74.7	73.5–76.0	0.0434	73.7	71.7–75.6	0.0135

MSGC, major salivary gland carcinoma; OS, overall survival; CSS, cancer-specific survival; AJCC, American Joint Committee on Cancer; AUC, area under curve; CI, confidence interval.

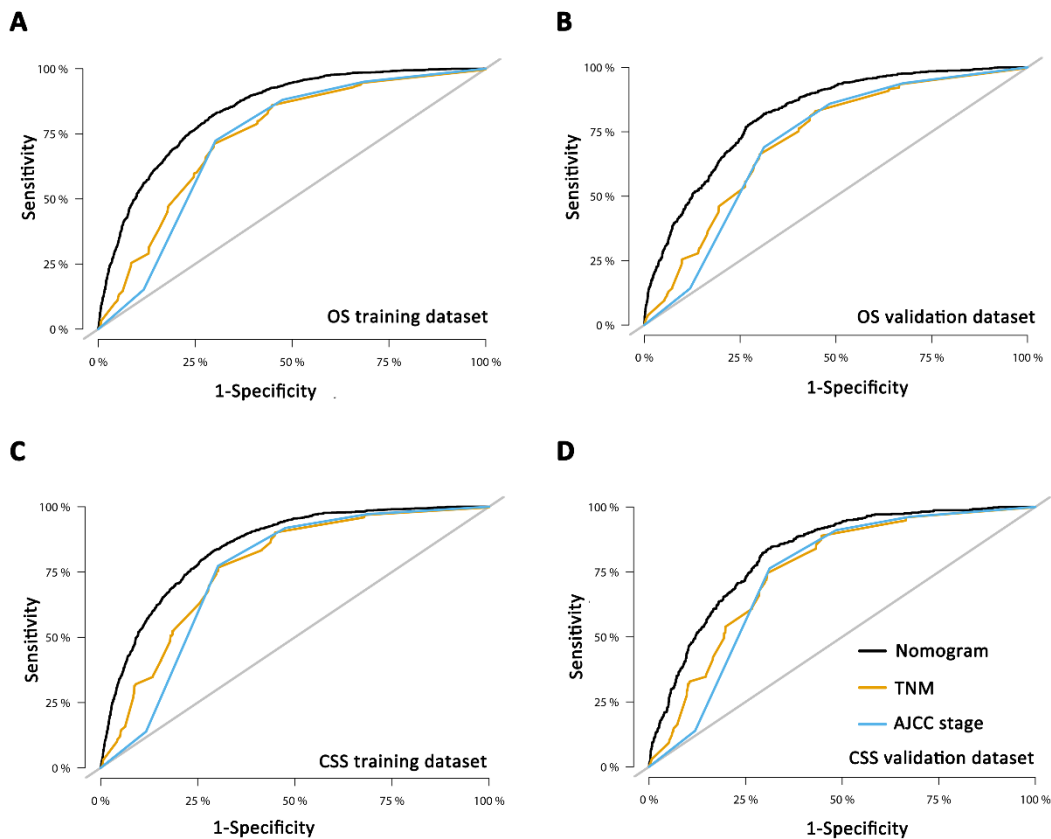


Figure S2 The ROC curves of nomograms, AJCC stage, and TNM stage for 3-year prediction in training and validation dataset. OS in the training dataset (A) and validation dataset (B). CSS in the training dataset (C) and validation dataset (D). ROC, receiver operating characteristic; AJCC, American Joint Committee on Cancer; OS, overall survival; CSS, cancer-specific survival.

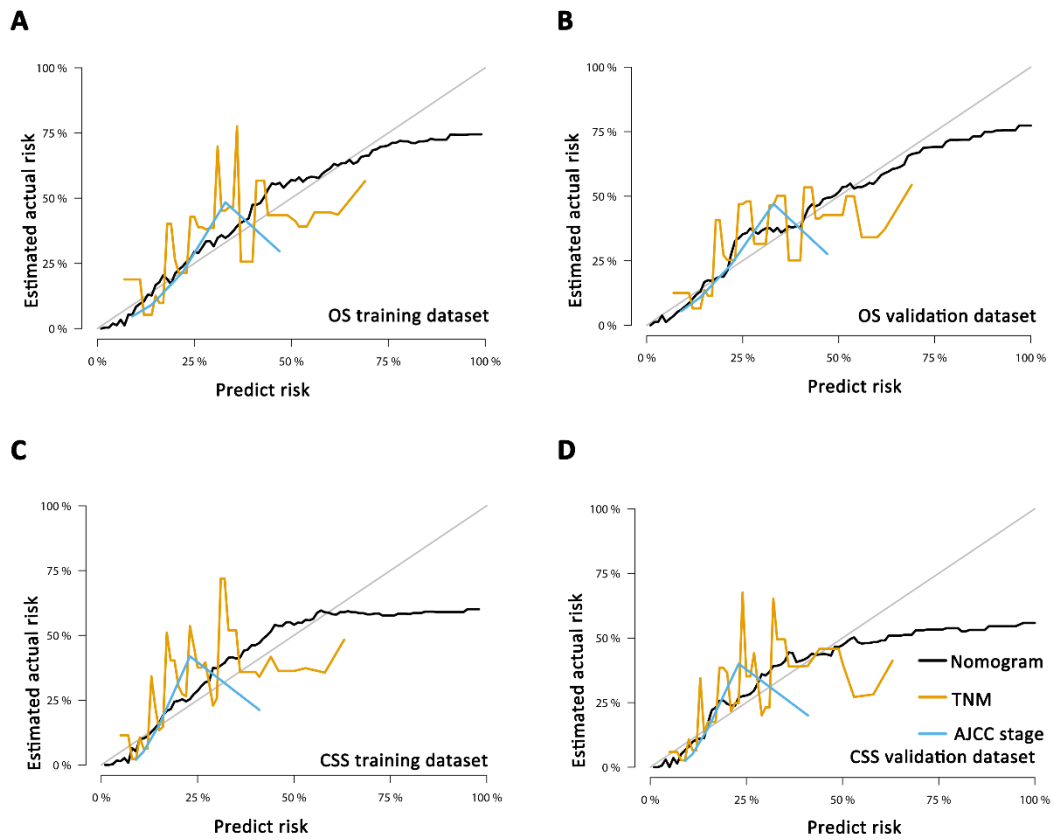


Figure S3 The calibration curves of nomograms, AJCC stage, and TNM stage for 3-year prediction in training and validation dataset. OS in the training dataset (A) and validation dataset (B). CSS in the training dataset (C) and validation dataset (D). AJCC, American Joint Committee on Cancer; OS, overall survival; CSS, cancer-specific survival.

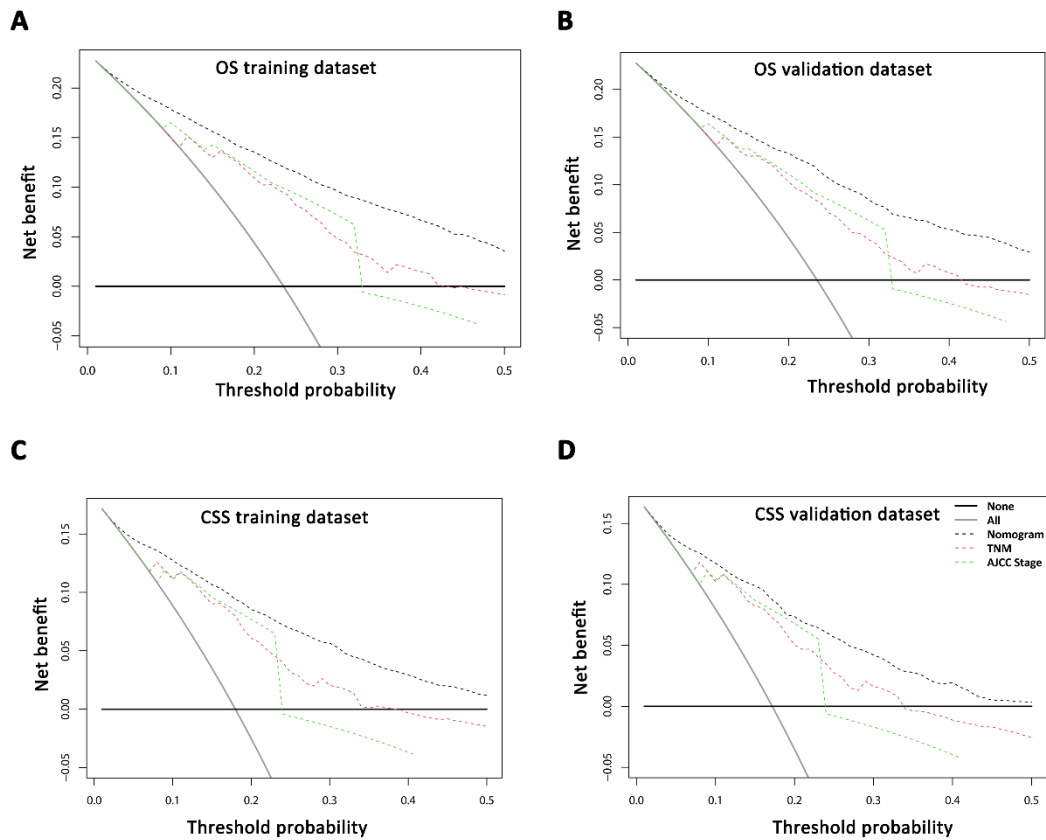


Figure S4 The DCA of nomograms, AJCC stage, and TNM stage for 5-year prediction in training and validation dataset. OS in the training dataset (A) and validation dataset (B). CSS in the training dataset (C) and validation dataset (D). DCA, decision curve analysis; AJCC, American Joint Committee on Cancer; OS, overall survival; CSS, cancer-specific survival.