

# Nomogram for predicting cancer-specific survival of patients with buccal squamous cell carcinoma: a population-based study

**Wei-Ming Wang**

Xiangya Hospital Central South University

**xiuhong chu**

yantai yeda hospital

**sisi yang**

Xiangya Hospital Central South University

**ying wang**

Guangdong University of technology

**chuanyu hu** (✉ [huchuanyu@tjh.tjmu.edu.cn](mailto:huchuanyu@tjh.tjmu.edu.cn))

Tongji Hospital of Tongji Medical College of Huazhong University of Science and Technology

---

## Research article

**Keywords:** nomogram, SEER, buccal squamous cell carcinoma, cancer-specific survival

**Posted Date:** April 30th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-23785/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Background

The use of the traditional American Joint Committee on Cancer (AJCC) staging system alone has limitations in predicting the survival of buccal squamous cell carcinoma (BSCC) patients. We aimed to establish a comprehensive prognostic nomogram whose prognostic value compared with AJCC system.

## Methods

Patients were obtained from the SEER (Surveillance, Epidemiology, and End Result) database. The Cox regression model was used to select variables using a backward stepwise selection method. Based on the predictive model for determining prognostic factors, a nomogram was developed to predict the cancer-specific survival rates of BSCC patients at 3, 5 and 8 years. We used several standard model validation methods to evaluate the performance of the survival model: concordance index (C-index), area under the time-dependent receiver operating characteristic curve (AUC), net reclassification improvement (NRI), calibration plots, decision curve analyses (DCAs), and integrated discrimination improvement (IDI).

## Results

Multivariate analysis showed that age, AJCC stage, N stage, M stage, pathological grade, and surgical status were risk factors for survival. The c-index chart, the AUC chart, and the calibration chart show that the performance of the nomogram is better, and the NRI and IDI values show that the performance of the nomogram is better than that of the AJCC staging system. The 3-year, 5-year and 8-year DCA curves of nomogram show that it has more net benefit than traditional AJCC staging system, both in terms of training and validation queues.

## Conclusion

We developed and validated the first BSCC prognosis nomogram, which has a better prognostic value than the separate AJCC staging system. Our BSCC prognosis nomogram is a valuable tool for explaining the risks of the next 3, 5 and 8 years in clinical practice.

## Background

Buccal mucosal squamous cell carcinoma (BSCC) is uncommon in Western countries, accounting for about 10% of all squamous cell carcinomas in the oral cavity[1]. However, it is reported to be more aggressive than other malignant tumors in the oral cavity[2]. BSCC is more common in Asian countries, where its incidence doubles to more than 20%[3]. We therefore considered it important to determine whether race influences the prognosis. The currently accepted indicator for assessing the prognosis of

BSCC is the American Joint Committee on Cancer (AJCC) staging system, which is the prognostic determinant of primary tumor size, depth of invasion, lymph node spread, and distant metastasis[4]. However, even in the same AJCC staging, survival outcomes can vary widely by using only the AJCC staging system. Given the clinical uniqueness of the BSCC, to improve the accuracy of survival predictions need new prognostic tools [5].

Our aim is to create a graphical representation of a convenient and simple mathematical prediction model that includes multiple factors that contribute to the prognosis of patients with BSCC. Nomogram is a model that combines various important factors to predict a particular endpoint[6]. Nomogram has become a reliable and convenient risk quantification tool and is widely used in cancer prognosis [7, 8]. In this study, we established a comprehensive prognostic evaluation system and its prognostic value was compared with that of AJCC staging system. We found that age, AJCC stage, N stage, M stage, pathological grade, and surgical status were important risk factors for survival. The accuracy of survival prediction may decrease if these important prognostic parameters are ignored [9].

## Methods

### Patient Selection

We use SEER\* Stat software (version 8.3.5; <https://seer.cancer.gov/>). We searched patients using 3449 histological type codes (for BSCC) in the third edition of the International Classification of Tumors. Patients under the age of 15 was excluded. We excluded the cases with unknown or missing AJCC staging or incomplete variables, as well as cases that were not validated by microscopy or autopsy. Subjects included age, gender, marital status, AJCC staging, TNM staging, pathological grade, surgery, radiotherapy, and chemotherapy. We used version 6 of the AJCC staging system and limited our search to 2004–2015 because it was released in 2004.

Our study identified 1,663 qualified patients in the SEER database. We randomly assigned 70% of patients to the training queue (n = 1164) and 30% to the validation queue (n = 499) for nomogram construction and validation. All data in the SEER database is free.

### Statistical analysis

Continuous variables that conform to the normal distribution are expressed as mean  $\pm$  SD values, with the remainder being the median (25–75 percentile). The categorical variables are expressed as a percentage. The training cohort was chosen in the Cox regression model by the backward stepwise selection method. A nomogram predicting 3-, 5-, and 8-year survival rates of BSCC patients was established by predictive models with well-defined prognostic factors.

The area under the C-index and the time-dependent receiver operating characteristic curve (AUC) evaluated the prediction accuracy of the nomogram. The consistency of the predicted probability with the

actual result was evaluated using a calibration plots. Five hundred resampled bootstrap methods evaluated the identification and calibration.

The accuracy of the two models was compared using net reclassification improvement (NRI) and integrated discrimination improvement (IDI) to determine the improvement achieved by applying the new prediction model. Finally, we perform decision curve analysis (DCAs) to test the clinical value of the predictive model.

SPSS software (version 24.0, SPSS, Chicago, IL, USA) and R software were used to analysis the data. Bilateral  $P \leq 0.05$  was considered statistically significant.

## Results

### Patient Characteristics

The median age of training cohort and validation cohort at the time of diagnosis was 68 years and 67 years respectively. Most of the patients were male (57.5%), married (53.4%), and white (79.7%). The AJCC stage was I, II, III, and IV in 30.3%, 20.0%, 16.6%, and 33.1% of the patients, respectively, while the pathological grade was I, II, III, and IV in 29.4%, 53.5%, 16.8%, and 0.3% of the patients, respectively. Across both cohorts, 76.6% of the patients had insurance and 84.7% of them underwent surgery. In the training and validation groups, the mean follow-up time was 32 months and 34 months, respectively. Table 1 showed the demographic characteristics and tumor characteristics of patients.

Table 1  
Patient characteristics in the study

<b>Variable</b>	<b>Training Cohort</b>	<b>Validation Cohort</b>
<b>Age at diagnosis</b>	68(53–74)	67(54–75)
<b>Race(%)</b>		
White	942(80.9)	385(96.5)
Black	75(6.5)	39(0.7)
Others	147(12.6)	75(2.8)
<b>Sex n(%)</b>		
Male gender	681(75.4)	275(77.7)
Female gender	483(24.6)	224(22.3)
<b>Marital status n(%)</b>		
Married	621(73.2)	267(74.1)
Unmarried	166(13.7)	73(13.1)
Others	377(13.1)	159(12.8)
<b>AJCC n(%)</b>		
I	363(78.9)	142(78.5)
II	239(12.5)	94(12.3)
III	187(4.4)	89(4.9)
IV	375(121)	174(4.3)
<b>Grade stage</b>		
Grade I	329(28.3)	160(32.1)
Grade II	631(55.1)	259(51.9)
Grade III	201(17.3)	78(15.6)
Grade IV	3(0.3)	2(0.4)
<b>Surgery</b>		
Yes	992(85.2)	418(83.8)
No	172(14.8)	81(16.2)
<b>Radiotherapy</b>		
Yes	533(45.8)	250(50.1)

<b>Variable</b>	<b>Training Cohort</b>	<b>Validation Cohort</b>
No	631(54.2)	249(49.9)
<b>Chemotherapy</b>		
Yes	242(20.8)	124(24.8)
No	922(79.2)	375(75.2)

## Variable Screening

Using the backward stepwise selection method, the patient's diagnostic age, AJCC staging, TNM staging, pathological grade, and surgical status were analyzed by multivariate Cox regression analysis to establish a model predicting 3-, 5-, and 8-year survival rates (Table 2).

Table 2  
Cox proportional hazards regression model based on selected variables (Training Cohort)

Variables	Multivariate analysis		
	HR	95%CI	P-value
Age at diagnosis	1.007	0.9990–1.015	0.084
<b>AJCC</b>			
I		Reference	
II	1.269	0.8274–1.947	0.275
III	1.840	1.1306–2.995	0.014
IV	2.839	1.8124–4.449	0.001
<b>N</b>			
N0		Reference	
N1	1.322	0.8861–1.972	0.171
N2	2.267	1.5722–3.270	0.001
N3	2.561	0.6105–10.742	0.199
<b>M</b>			
M0		Reference	
M1	1.586	0.9479–2.652	0.079
<b>Grade</b>			
Grade I		Reference	
Grade II	0.943	0.7033–1.265	0.696
Grade III	1.350	0.9606–1.897	0.840
Grade IV	7.578	1.8078–31.767	0.006
<b>Surgery</b>			
Yes		Reference	
No	3.212	2.4673–4.181	0.001

## Prognostic Nomogram For Survival

Based on selected variables with hazard ratios, in the training cohort a nomogram that incorporated all of the significant independent factors was established for predicting the 3-, 5-, and 8-year survival rates. Nomogram showed that pathological grade had the greatest impact on prognosis, followed by surgical status, AJCC stage, N stage, age, and M stage. Use a nomogram to give each variable its corresponding subscale, and all the variables of the score are counted in the total score, with the vertical line then descending from the total score, estimating the survival probability of 3–5-, 8 years (Fig. 1 )

## Performance Of The Nomogram

The c-index of the nomogram group (0.792) and the validation group (0.794) were higher than the AJCC staging system (0.726 and 0.723). 3–5 and 8 year auc's nomogram training queues (0.821, 0.825, and 0.780, respectively) and validation cohorts (0.840, 0.839, and 0.793) indicates that the model had good recognition ability, and this is better than ratio and staging System (Fig. 2). The calibration block nomogram shows that the expected survival probability of the SEER training and verification corps is almost the same as the actual observation of the 3–5-, 8-year survival probability (Fig. 3).

The NRI values for 3, 5, and 8 years of follow-up were 0.476 (95% CI = 0.209–0.965), 0.487 (95% CI = 0.100–0.967), and 0.518 (95% CI = 0.126–0.959), respectively, in the validation cohort. These numerical results showed that the nomogram represents a significant improvement in prediction performance. Similarly, the IDI values for 3, 5, and 8 years of follow-up were 0.100, 0.080, and 0.079, respectively, in the validation cohort (all  $P < 0.001$ ), which also validates the improvement in predictive performance provided by the nomogram.

## Decision-curve Analysis

Although both models produced net benefits, DCA curves of 3, 5 and 8 years showed that the net benefits of nomogram were higher than those of traditional AJCC staging system in training and validation cohort (Fig. 4).

## Discussion

The incidence of BSCC varies between different regions. It is low in Europe and the United States (10%), but more than doubles in Asian countries (23–27%)[1, 2]. The survival rate of BSCC is worse than that of other oral cancers[10, 11], and it is reportedly more aggressive than other malignant tumors in the oral cavity[1, 3]. However, there is little information available in the literature for predicting the cancer-specific survival in patients with BSCC[12]. Currently the traditionally accepted indicator for evaluating the prognosis of BSCC is the AJCC stage. There are still limitation that it does not include important risk factors such as pathological grade, age, and treatment measures for the AJCC staging system to predict the prognosis of BSCC.[13]. Therefore, this study aimed to established a more comprehensive predictive model that included system demographics and other clinical parameters, rather than AJCC staging. Our

analysis provides new data for clinicians to more accurately predict patient outcomes. Nomogram possessed of high recognition ability is an excellent risk assessment tool. It can provide quantitative prognosis for individual patients through simple graphical representation of statistical prediction models, making them more sensitive and informative[14].

As with most previous studies, our multivariate Cox regression analysis suggested that age is a significant risk factor for survival in BSCC patients. Our nomogram indicates that pathological grade is the most important factor, while age is the second most important one[15, 16], and there is no significant difference between grades I and II. Multivariate analyses indicated that older age was an independent risk variable, with older patients having a lower survival rate[9, 17, 18]. Meanwhile, a higher AJCC stage was associated with worse survival[9]. The N and M stages are also significant factors affecting cancer-specific survival. Higher N and M stages indicate more-serious lymph node metastasis and distant metastasis, and are associated with worse prognoses[19]. The score on the nomogram therefore increases with the pathological grade, AJCC stage, N stage, and M stage.

Among the three main treatments applied in BSCC, surgery was found to be the only significantly effective treatment, and so the radiotherapy and chemotherapy statuses were not included in the final nomogram. Surgical treatment should be the first choice of treatment for BSCC[18, 20, 21], while avoiding unnecessary radiotherapy and chemotherapy can reduce the burden on medical resources, alleviate the additional injury and pain experienced by patients, and reduce costs. This information can further help clinicians to make rational clinical decisions.

An easy-to-use nomogram was developed and validated for predicting the survival rates of BSCC patients after 3, 5, and 8 years. Our nomogram model contains risk factors that are readily obtainable from historical records. Its clinical applicability and ease of use are the most-attractive features of the nomogram that we have constructed.

We used several common parameters to assess the performance of the survival model: identification, calibration, NRI, IDI, and DCA to further determine whether the prognostic model is superior to the traditional AJCC staging system[22–24]. The model has good discriminating ability. The c-index of the training queue is 0.792, and the c-index of the verification queue is 0.794. The corresponding value of the AJCC staging system is lower. The AUCs of the nomogram of the validation cohort were also higher (0.840, 0.839, and 0.793, respectively, 3, 5, and 8 years). Compared with the training queue (0.821, 0.825, 0.780) (Fig. 2), it has good application value. The discriminating power of nomogram is also significantly higher than that of the AJCC staging system. A similar 45-degree line plot in the training and validation queues indicates that the nomogram prediction is well calibrated (Fig. 3).

NRI and IDI are more-sensitive indicators than the C-index[25, 26]. NRI reflects how nomogram can better reclassify risk probability than AJCC staging system, while IDI reflects the improvement of distinguishing ability between nomogram and AJCC staging system. The empirical results further show that the method of the nomogram has good performance. We also included one of the latest analytical techniques, DCA, which reportedly has benefits and has been recommended for use[27, 28]. The 3-, 5-, and 8-year DCA

curves for the new model show that the net benefit of the new model, both in terms of training and verification queues, is higher than that of the traditional AJCC staging system.

The proposed nomogram is an easy-to-use clinical tool for promoting patient counseling and personalized treatments. Our findings may have pragmatic implications. This can be demonstrated by considering two BSCC patients at AJCC stage IV: Patient 1 was 69 years old and had pathological grade III, TNM stages N1 and M0, and had not received surgery; while Patient 2 was 51 years old and had pathological grade II, TNM stages N0 and M0, and had received surgery. Applying the AJCC staging system would produce the same prognosis for these two patients[29]. However, the results obtained by applying our nomogram are markedly different, with 3-, 5-, and 8-year predicted cancer-specific survival rates of 23.3%, 16.8%, and 12.5%, respectively, for Patient 1, and 78.7%, 74.5%, and 71.2% for Patient 2.

## Limitations

The biggest advantage of this study is that a large population, the high quality of the SEER database was covered, and population-based research results are more universal than a single study. However, certain limitations need to be addressed. First, a nomogram is created based on retrospective data from the SEER database, which has the potential for selection bias. Second, the differences in surgical procedures (if with or without neck dissection) were not reported in the SEER database, and some important clinical pathological parameters associated with prognosis, including vascular invasion and surgical margin status, were not reported[30]. Third, there are no relevant information for certain potentially important prognostic factors, such as the type of chemotherapy applied and the status of tumor markers (such as P53, EGFR, and HPV)[31–33]. Finally, the predicted values calculated by nomogram are only suitable for clinician reference and do not represent an absolutely accurate prognosis. We intend to conduct forward-looking research in the future to address these limitations.

## Conclusions

We have developed and validated a highly accurate nomogram for the prognosis of BSCC. The prognostic value of the nomogram is better than that of the AJCC staging system alone. Our nomogram may be a valuable tool for use in clinical practice when helping patients to understand their risk over the next 3, 5, and 8 years.

## Declarations

Ethics approval and consent to participate N/A.

Consent for publication: N/A.

Availability of data and materials: N/A.

Competing interests: The authors declare that they have no competing interests.

Authors' contributions: Chuan-Yu Hu and Wei-Ming Wang participated in the design of this study; and they both performed the statistical analysis. Chuan-Yu Hu carried out the study and collected important background information. Wei-Ming Wang drafted the manuscript. Xiu-Hong Chu carried out literature search, data acquisition and manuscript editing; Si-si yang performed manuscript review. All authors read and approved the final manuscript.

Funding: This study was funded by the National Natural Science Foundation of China (grant no.81702708), Natural Science Foundation of Hunan (grant no. 2018JJ3862), Natural Science Foundation of Hubei (grant no. 2019CFB657).

Acknowledgments: We appreciated Xue-Jin Tao for manuscript editing.

## References

1. DeConde A, et al. Squamous cell carcinoma of buccal mucosa: a 40-year review. *Am J Otolaryngol.* 2012;33(6):673–7.
2. Strome SE, et al. Squamous cell carcinoma of the buccal mucosa. *Otolaryngol Head Neck Surg.* 1999;120(3):375–9.
3. Liao CT, et al. Comparative clinical outcomes of Taiwanese patients with resected buccal and tongue squamous cell carcinomas. *Oral Oncol.* 2017;67:95–102.
4. Amin MB, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J Clin.* 2017;67(2):93–9.
5. Weiss SA, et al. Revisiting determinants of prognosis in cutaneous melanoma. *Cancer.* 2015;121(23):4108–23.
6. Zhang G, et al. Nomograms for predicting long-term overall survival and disease-specific survival of patients with clear cell renal cell carcinoma. *Onco Targets Ther.* 2018;11:5535–44.
7. Cao J, et al. Clinical Nomogram for Predicting Survival of Esophageal Cancer Patients after Esophagectomy. *Sci Rep.* 2016;6:26684.
8. Sun W, et al. Nomograms to estimate long-term overall survival and breast cancer-specific survival of patients with luminal breast cancer. *Oncotarget.* 2016;7(15):20496–506.
9. Wang F, et al. Nomograms forecasting long-term overall and cancer-specific survival of patients with oral squamous cell carcinoma. *Cancer Med.* 2018;7(4):943–52.
10. Lubek JE, et al. A retrospective analysis of squamous carcinoma of the buccal mucosa: an aggressive subsite within the oral cavity. *J Oral Maxillofac Surg.* 2013;71(6):1126–31.
11. Bachar G, et al. Squamous cell carcinoma of the buccal mucosa: outcomes of treatment in the modern era. *Laryngoscope.* 2012;122(7):1552–7.
12. Liu J, et al. Development and external validation of a prognostic nomogram for gastric cancer using the national cancer registry. *Oncotarget.* 2016;7(24):35853–64.

13. Li Y, 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(5):1002–13.
14. Bobdey S, et al. A Nomogram based prognostic score that is superior to conventional TNM staging in predicting outcome of surgically treated T4 buccal mucosa cancer: Time to think beyond TNM. *Oral Oncol.* 2018;81:10–5.
15. Eder-Czembirek C, et al. P16 positivity and regression grade predict survival after neoadjuvant radiotherapy of OSCC. *Oral Dis.* 2018;24(4):544–51.
16. Doll C, et al. Expression of Estrogen Receptors in OSCC in Relation to Histopathological Grade. *Anticancer Res.* 2015;35(11):5867–72.
17. Hu CY, et al. Application of Delayed Surgical Managements in Patients with Stensen's Duct Injury. *Curr Med Sci.* 2018;38(3):519–23.
18. Mohiyuddin SMA, et al. Clinicopathological study of surgical margins in squamous cell carcinoma of buccal mucosa. *World J Otorhinolaryngol Head Neck Surg.* 2016;2(1):17–21.
19. Torrecillas V, et al. Adjuvant radiation for T1-2N1 oral cavity cancer survival outcomes and utilization treatment trends: Analysis of the SEER database. *Oral Oncol.* 2018;85:1–7.
20. Mair MD, et al. Longitudinal and cross-sectional assessment of quality of life in surgically treated advanced (T4) cancer of the buccal mucosa. *Oral Surg Oral Med Oral Pathol Oral Radiol.* 2017;124(6):529–36.
21. Zaman SU, Aqil S, Sulaiman MA. Predictors of locoregional recurrence in early stage buccal cancer with pathologically clear surgical margins and negative neck. *Acta Otorrinolaringol Esp.* 2018;69(4):226–30.
22. Lip GYH, et al. Long-term bleeding risk prediction in 'real world' patients with atrial fibrillation: Comparison of the HAS-BLED and ABC-Bleeding risk scores. *Thromb Haemost.* 2017;117(10):1848–58.
23. Garcia-Fernandez A, et al. Does von Willebrand factor improve the predictive ability of current risk stratification scores in patients with atrial fibrillation? *Sci Rep.* 2017;7:41565.
24. Rodrigues G, et al. A statistical comparison of prognostic index systems for brain metastases after stereotactic radiosurgery or fractionated stereotactic radiation therapy. *Clin Oncol (R Coll Radiol).* 2013;25(4):227–35.
25. Tan X, et al. Radiomics nomogram outperforms size criteria in discriminating lymph node metastasis in resectable esophageal squamous cell carcinoma. *Eur Radiol.* 2019;29(1):392–400.
26. Chen LD, et al. Multiparametric radiomics improve prediction of lymph node metastasis of rectal cancer compared with conventional radiomics. *Life Sci.* 2018;208:55–63.
27. Talluri R, Shete S. Using the weighted area under the net benefit curve for decision curve analysis. *BMC Med Inform Decis Mak.* 2016;16:94.
28. Rousson V, Zumbo T. Decision curve analysis revisited: overall net benefit, relationships to ROC curve analysis, and application to case-control studies. *BMC Med Inform Decis Mak.* 2011;11:45.

29. 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(6): p. 1471-4.
30. Song W, et al. Development and validation of a nomogram for predicting survival in patients with gastrointestinal stromal tumours. Eur J Surg Oncol. 2018;44(10):1657–65.
31. Larsen SR, et al. The prognostic significance of histological features in oral squamous cell carcinoma. J Oral Pathol Med. 2009;38(8):657–62.
32. Perrone F, et al. TP53 mutations and pathologic complete response to neoadjuvant cisplatin and fluorouracil chemotherapy in resected oral cavity squamous cell carcinoma. J Clin Oncol. 2010;28(5):761–6.
33. Zanotti L, et al. Epidermal growth factor receptor detection in serum and saliva as a diagnostic and prognostic tool in oral cancer. Laryngoscope. 2017;127(11):E408–14.

## Figures

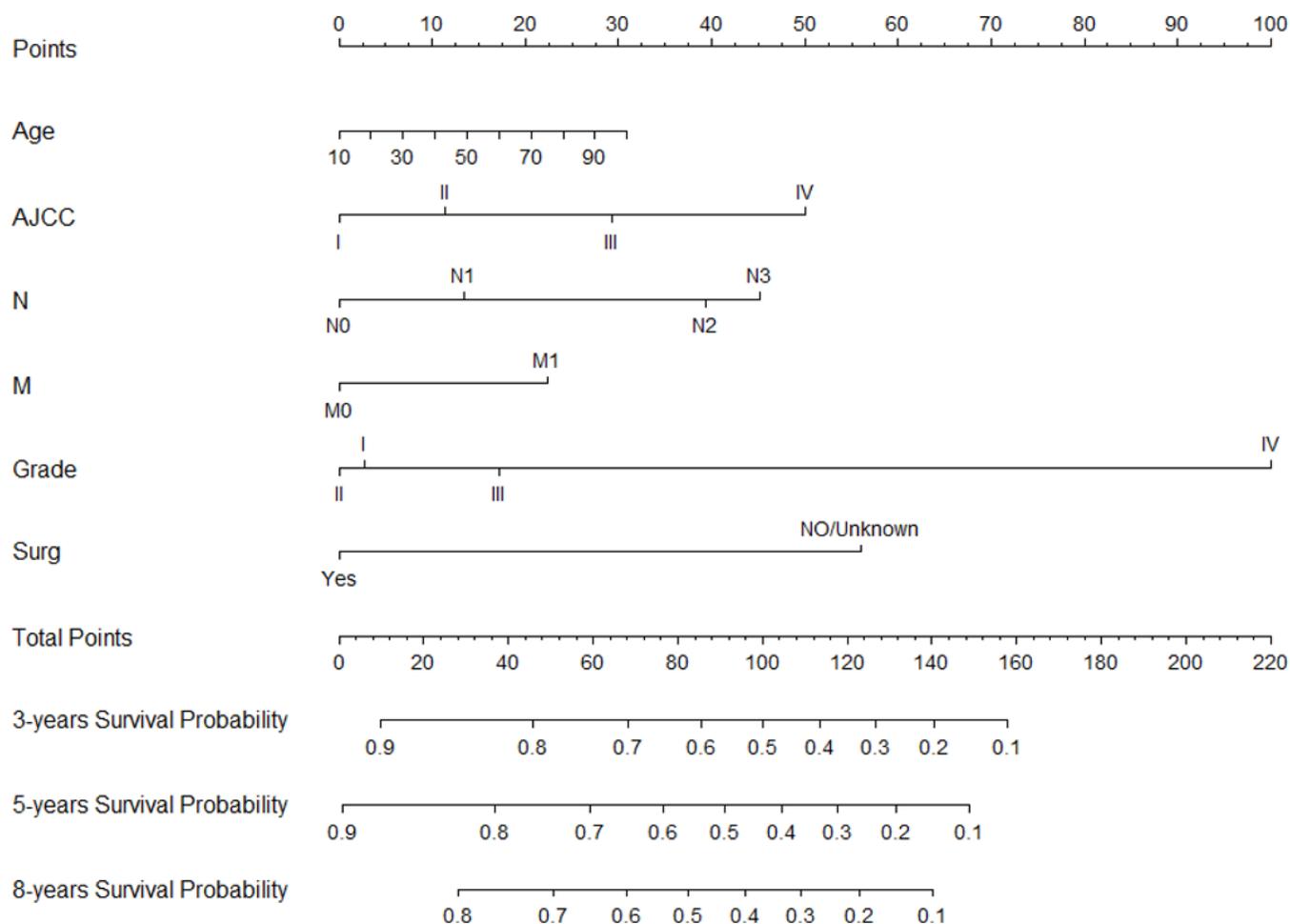
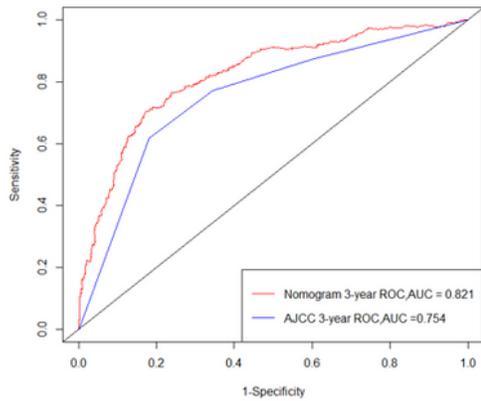
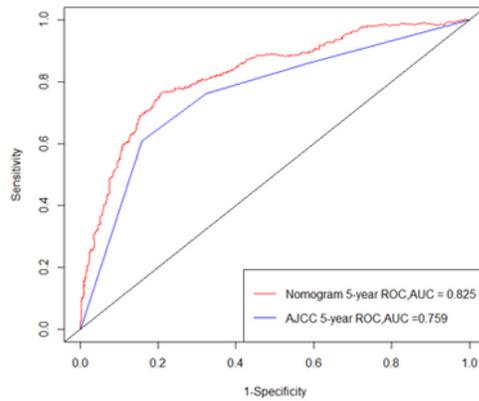


Figure 1

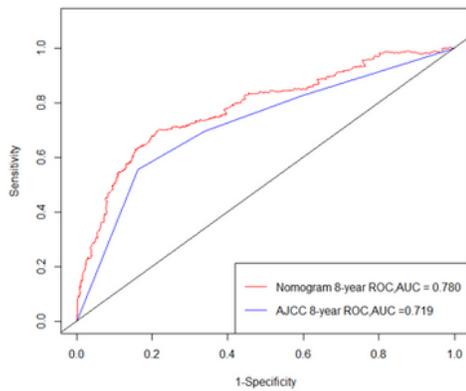
Nomogram predicting 3- 5- and 8-year survival. Mari, marital status: M, married; SD, single/domestic partner; DSW, divorced & separated & widowed. SEER Stage: L, Localized; R, Regional; D, distant. Insur, insurance status: In, insured; AM, any medical; Unin, uninsured. LND, lymph node density.



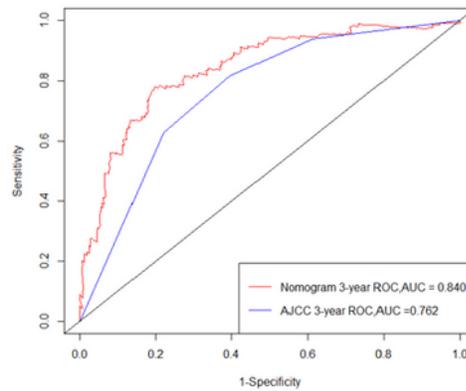
A Training 3-year ROC curve



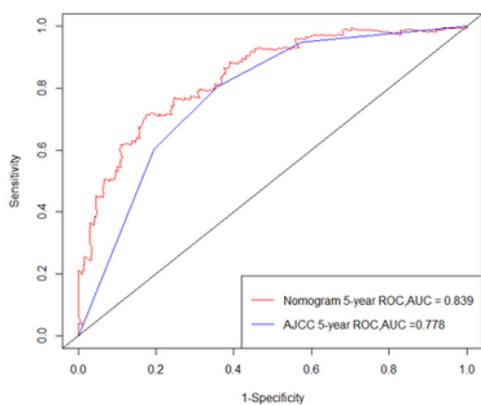
B Training 5-year ROC curve



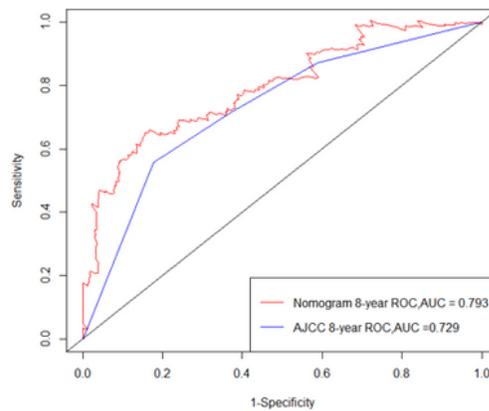
C Training 8-year ROC curve



D Validation 3-year ROC curve



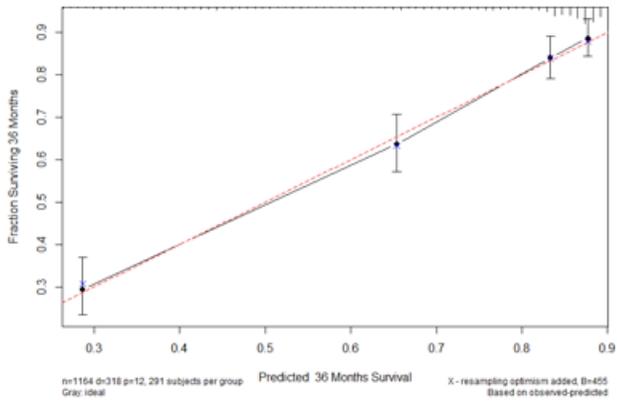
E Validation 5-year ROC curve



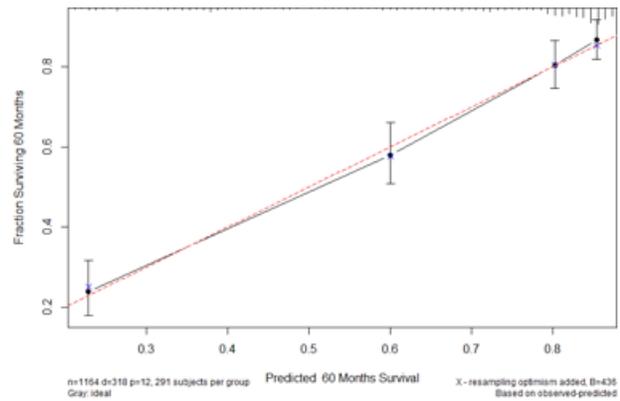
F Validation 8-year ROC curve

## Figure 2

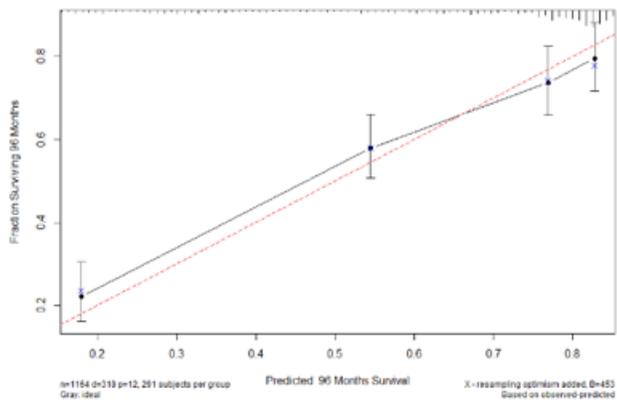
ROC curves. The ability of the model to be measured by the C index. A,B,C came from the training set, and D,E,F came from the validation set.



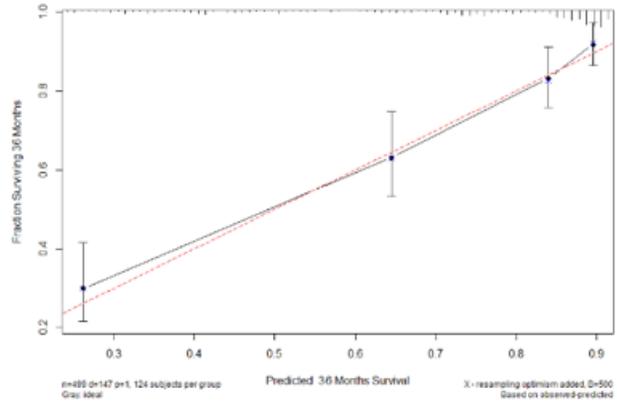
A



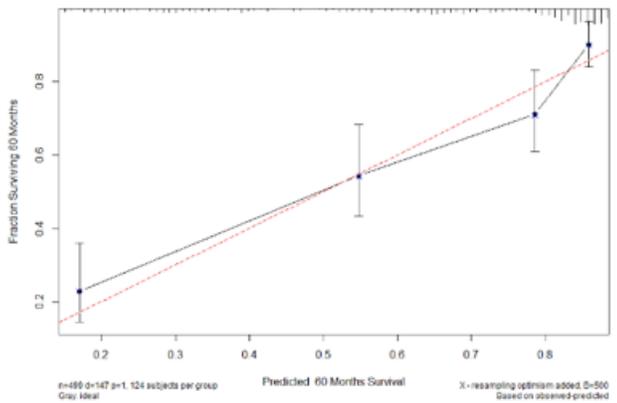
B



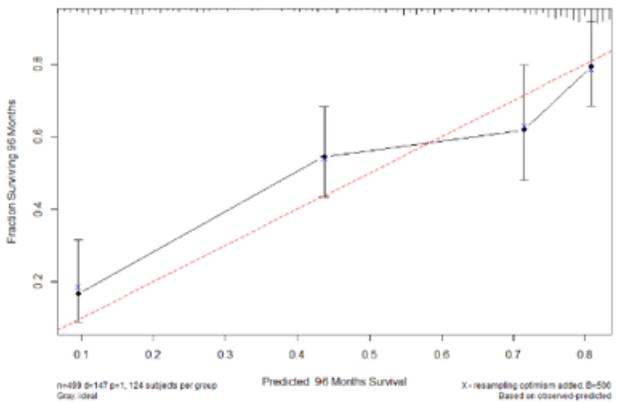
C



D



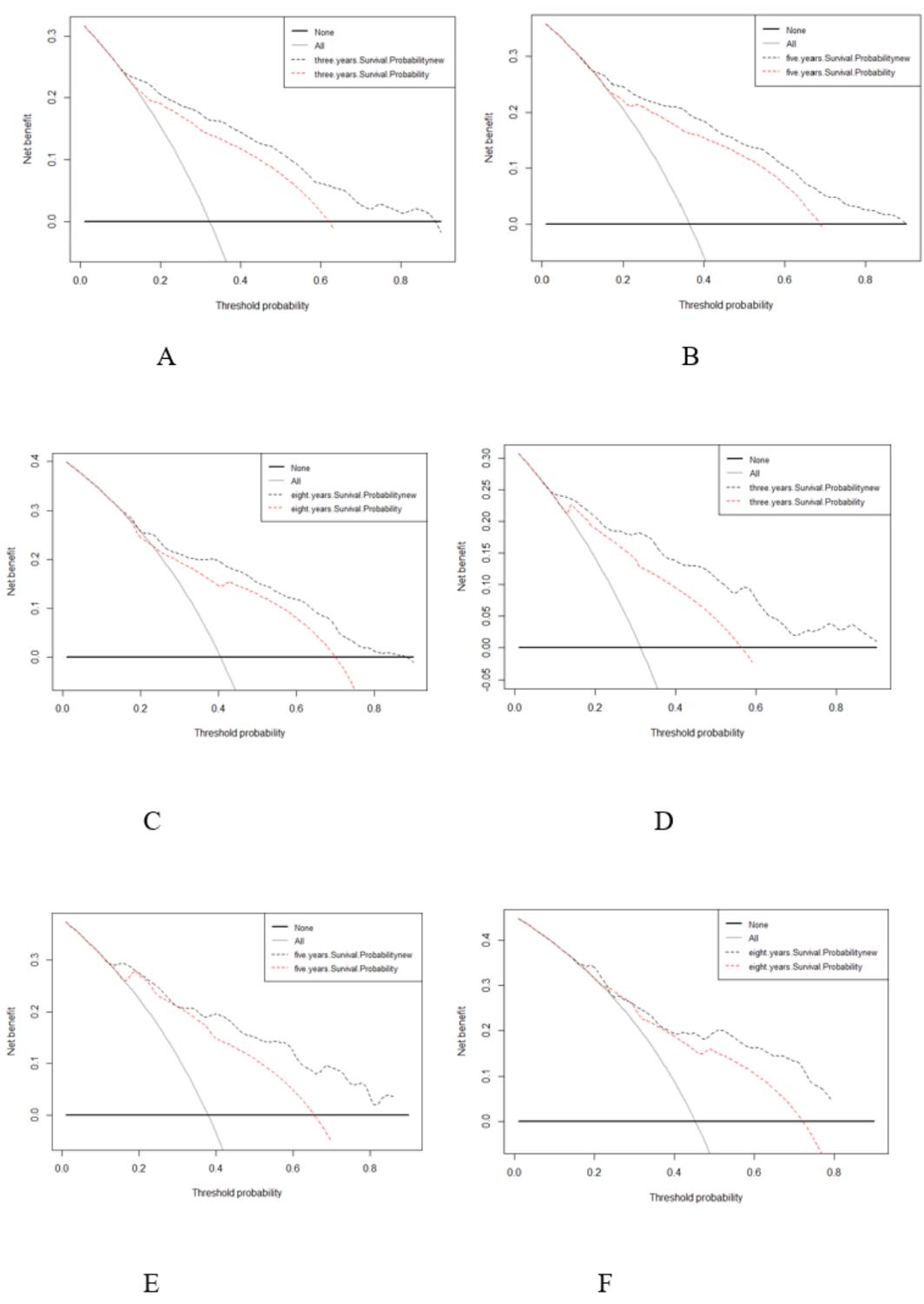
E



F

**Figure 3**

Calibration plots. Show the relationship between the predicted probabilities base on the nomogram and actual values of the train set (A, B,C) and validation set (D, E,F).



**Figure 4**

Decision curve analysis In the figure, the abscissa is the threshold probability, the ordinate is the net benefit rate. The horizontal one indicates that all samples are negative and all are not treated, with a net benefit of zero. The oblique one indicates that all samples are positive. The net benefit is a backslash with a negative slope. A, B and C came from the training set, and D, E and F came from the validation set.