Am J Cancer Res 2024;14(4):1747-1767
www.ajcr.us /ISSN:2156-6976/ajcr0155768

Original Article

Nomogram models for predicting overall and cancer-specific survival in early-onset gastric cancer patients: a population-based cohort study

Xiaoyan Wang[1,2,3,4*], Xiaoman Niu[1,2,3*], Fengbin Zhang[1,2,5*], Jiaxiang Wu[1,2,3], Haotian Wu[1,2,3], Tongkun Li[1,2,3], Jiaxuan Yang[1,2,3], Ping'an Ding[1,2,3], Honghai Guo[1,2,3], Yuan Tian[1,2,3], Peigang Yang[1,2,3], Zhidong Zhang[1,2,3], Dong Wang[1,2,3], Qun Zhao[1,2,3]

[1]Third Department of Surgery, The Fourth Hospital of Hebei Medical University, Shijiazhuang 050011, Hebei, China; [2]Hebei Key Laboratory of Precision Diagnosis and Comprehensive Treatment of Gastric Cancer, Shijiazhuang 050011, Hebei, China; [3]Big Data Analysis and Mining Application for Precise Diagnosis and Treatment of Gastric Cancer Hebei Provincial Engineering Research Center, Shijiazhuang 050011, Hebei, China; [4]Medical Oncology, Shijiazhuang People's Hospital, Shijiazhuang 050050, Hebei, China; [5]Department of Gastroenterology, The Fourth Hospital of Hebei Medical University, Shijiazhuang 050011, Hebei, China. [*]Equal contributors.

Received February 4, 2024; Accepted April 3, 2024; Epub April 15, 2024; Published April 30, 2024

Abstract: To develop nomogram models for predicting the overall survival (OS) and cancer-specific survival (CSS) of early-onset gastric cancer (EOGC) patients. A total of 1077 EOGC patients from the Surveillance, Epidemiology, and End Results (SEER) database were included, and an additional 512 EOGC patients were recruited from the Fourth Hospital of Hebei Medical University, serving as an external test set. Univariate and multivariate Cox regression analyses were performed to identify independent prognostic factors. Based on these factors, two nomogram models were established, and web-based calculators were developed. These models were validated using receiver operating characteristics (ROC) curve analysis, calibration curves, and decision curve analysis (DCA). Multivariate analysis identified gender, histological type, stage, N stage, tumor size, surgery, primary site, and lung metastasis as independent prognostic factors for OS and CSS in EOGC patients. Calibration curves and DCA curves demonstrated that the two constructed nomogram models exhibited good performance. These nomogram models demonstrated superior performance compared to the 7th edition of the AJCC tumor-node-metastasis (TNM) classification (internal validation set: 1-year OS: 0.831 vs 0.793, P = 0.072; 1-year CSS: 0.842 vs 0.816, P = 0.190; 3-year OS: 0.892 vs 0.857, P = 0.039; 3-year CSS: 0.887 vs 0.848, P = 0.018; 5-year OS: 0.906 vs 0.880, P = 0.133; 5-year CSS: 0.900 vs 0.876, P = 0.109). In conclusion, this study developed two nomogram models: one for predicting OS and the other for CSS of EOGC patients, offering valuable assistance to clinicians.

Keywords: Early-onset gastric cancer, nomogram, overall survival, cancer-specific survival, SEER

Introduction

Gastric cancer is a significant global health challenge, with high incidence and mortality rates observed in East Asia, Eastern Europe, and South America [1]. East Asia has the highest age-standardized incidence rate (ASR) of 22.4/100,000 population, followed by Central and Eastern Europe (ASR 11.3/100,000) and South America (ASR 8.7/100,000) [2].

Early-onset gastric cancer (EOGC), which accounts for 2.7-10% of all gastric cancers, has distinct clinical, pathological, and genetic characteristics compared to late-onset gastric cancer, including diffuse lesions, genetic alterations, and poor differentiation grade [3]. According to the Surveillance, Epidemiology, and End Results (SEER) program of the National Cancer Institute, the incidence of non-cardia gastric cancer declined in all racial and age groups from 1977 to 2006, except for an increase in the incidence among white individuals aged 25 to 39 years [4]. The treatment for early-stage gastric cancer currently doesn't differ much from the treatment for common gas-

tric cancer. It mainly includes complete surgical resection, supplemented by neoadjuvant or adjuvant chemotherapy, with or without radiotherapy [5].

Nomogram models are valuable tools for predicting outcomes. The Memorial Sloan Kettering Cancer Center (MSKCC) nomogram model, established in 2003, is a well-known and classic model for predicting gastric cancer outcomes [6]. Nomogram models have gained popularity in oncology due to their accuracy, practicality, and nuanced ability to discriminate between different outcomes [7, 8]. While the American Joint Committee on Cancer (AJCC) tumor-node-metastasis (TNM) staging system is prevalent in clinical practice, its predictive scope is somewhat constrained, which can hinder its effectiveness in prognostic prediction [8]. Although Liu et al. previously developed a nomogram model to predict cancer-specific survival (CSS) in postoperative EOGC patients, their model was specific to patients who underwent surgery without addressing overall survival (OS) in EOGC patients [9]. Therefore, there is an urgent need to construct widely applicable nomogram prediction models for OS and CSS in EOGC patients to further improve their prognosis and treatment outcomes. Based on the SEER database and data from the Fourth Hospital of Hebei Medical University, this study aims to establish two more precise nomogram models for predicting the OS and CSS prognosis of EOGC patients. Additionally, we sought to create two web-based calculators to offer reliable decision-making support for personalized clinical treatment. Furthermore, we compared these nomogram models with the AJCC staging system to assess the advantages of these nomogram models in prognostic prediction.

Methods and materials

Patient selection

This study sourced data on EOGC patients from the SEER program of the National Cancer Institute in the United States during the period of 2010-2015. SEER collects and publishes population-based cancer registry data, including incidence rates, treatment information, and survival data, covering over 28% of the U.S. population. The collected data includes detailed information on demographics, diagnosis, and tumor characteristics [10]. Since the SEER database is publicly available, analyzing the collected patient data does not require informed consent or institutional review.

For data extraction and analysis, this study utilized SEER*Stat software version 8.4.2. Histologically confirmed signet ring cell carcinoma, diffuse carcinoma, and linitis plastica were classified as diffuse type; gastric adenocarcinoma, intestinal type; and other types included neoplasm, malignant; large cell neuroendocrine carcinoma; small cell carcinoma; squamous cell carcinoma; carcinoid tumor; neuroendocrine carcinoma; and neuroendocrine carcinoma [11].

Inclusion and exclusion criteria

Inclusion criteria for this study were: (1) patients diagnosed with gastric cancer (Site recode of ICD-0-3/WHO 2008: C16.0-C16.9); (2) age at diagnosis < 40 years; (3) diagnosed between 2010 and 2015; (4) complete survival and follow-up data. Exclusion criteria were: (1) cases diagnosed through autopsy or death certificate confirmation; and (2) cases with incomplete data. In total, 1077 cases were screened out from the SEER database.

Additionally, we collected data on EOGC patients diagnosed at the Fourth Hospital of Hebei Medical University and aged younger than 40 between 2016 and 2018 as an external test dataset. The inclusion and exclusion criteria for this dataset were the same as those for the SEER dataset. This study is a population-based cohort study based on data from the SEER database and the Fourth Hospital of Hebei Medical University.

Clinical and pathological parameters

We used the R software to randomly split the dataset of SEER database into a training set and a validation set with a ratio of 7:3. The training set was used for risk factor analysis and constructing nomogram models, while the validation set was used to evaluate the models. Nineteen major clinical and pathological factors were extracted from the SEER database, including age, sex, race, year of diagnosis, histological type, grade, stage, T stage, N stage, M stage, tumor size, surgery, chemotherapy, radiation, tumor site, bone metastasis, lung metastasis, brain metastasis,

and liver metastasis. In the external test dataset, 15 major clinical and pathological factors were extracted, including age, sex, histological type, grade, stage, T stage, N stage, M stage, tumor size, surgery, tumor site, bone metastasis, lung metastasis, brain metastasis, and liver metastasis. The OS and CSS of the collected data, were set as the endpoint time of this study.

Statistical analyses

Statistical analyses were performed using R software version 4.3.2. Variable comparisons were conducted using chi-square tests and independent sample t-tests. Cox proportional hazards regression models were used for univariate analysis of OS and CSS, and variables with significant differences in the univariate Cox regression analysis were included in the multivariate Cox regression analysis to identify independent prognostic factors. Based on the independent prognostic factors, nomogram models were constructed, and web-based calculators were developed for visualization. The performance of these prediction models was evaluated using time-dependent receiver operating characteristics (ROC) curves, and the consistency between predicted survival rates and actual survival rates was assessed through calibration curves. The clinical utility of models was measured using decision curve analysis (DCA) to quantify the net benefit at different threshold probabilities. Comparisons were made between the constructed nomogram models and the AJCC staging system. The Delong test was employed to compare the AUC values between the training set and internal validation set, as well as between the internal validation set and external test set. Subsequently, we utilized X-tile software to determine the optimal cutoff value based on the predicted overall scores from the nomogram models of the training set. The patients were then divided into "high-risk", "intermediate-risk", and "low-risk" groups. Kaplan-Meier (KM) survival analysis and Log-rank test were employed to plot the survival curves for OS and CSS, stratified by risk level, for both the internal validation set and the external test set of EOGC patients. A p-value < 0.05 was considered statistically significant.

Results

Clinical pathological baseline characteristics comparison of EOGC patients

In this study, we screened EOGC patients registered in the SEER database from 2010 to 2015 using predefined inclusion and exclusion criteria. The process of conducting research and statistical analysis is depicted in **Figure 1**. From the SEER database, a total of 1,331 gastric cancer patients under 40 years old diagnosed between 2010 and 2015 were initially identified. After excluding cases with missing data on various study indicators (249 cases) and cases confirmed through autopsy or death certificate (5 cases), a final dataset of 1,077 patients was included. The patients from the SEER database were randomly divided into a training set (753 cases) and a validation set (324 cases) in a 7:3 ratio. In the training set, there were 587 patients (78.0%) aged 30-39, 489 patients (64.9%) with Grade III tumors, and 293 patients (38.9%) who underwent surgical treatment. In the validation set, there were 251 patients (77.5%) aged 30-39, with Grade III tumors being the most prevalent (227 cases, 70.1%), and 124 patients (38.3%) underwent surgical treatment. Results indicated no statistical differences in age, gender, race, age at diagnosis, histological subtype, etc., between the internal training and internal validation sets (all P > 0.05). In the external test set, there were 442 patients (86.3%) aged 30-39, 489 patients (64.9%) with Grade III tumors, and 460 patients (61.1%) who underwent surgical treatment. Comparison between the internal training and external testing sets revealed statistical differences in histological subtype, stage, tumor size, surgery, tumor location (**Table 1**).

Survival across clinicopathological risk factors

Based on the Kaplan-Meier and Log-rank test methods, survival curves for overall survival (OS) (Figure S1) and cancer-specific survival (CCS) (Figure S2) were plotted for various variables including grade, stage, T stage, N stage, M stage, histological subtype, lung metastasis, liver metastasis, bone metastasis, brain metastasis, tumor location, tumor size, gender, age, surgery, chemotherapy, and radiotherapy. The results showed that older age and later stage were associated with higher OS risk and CSS

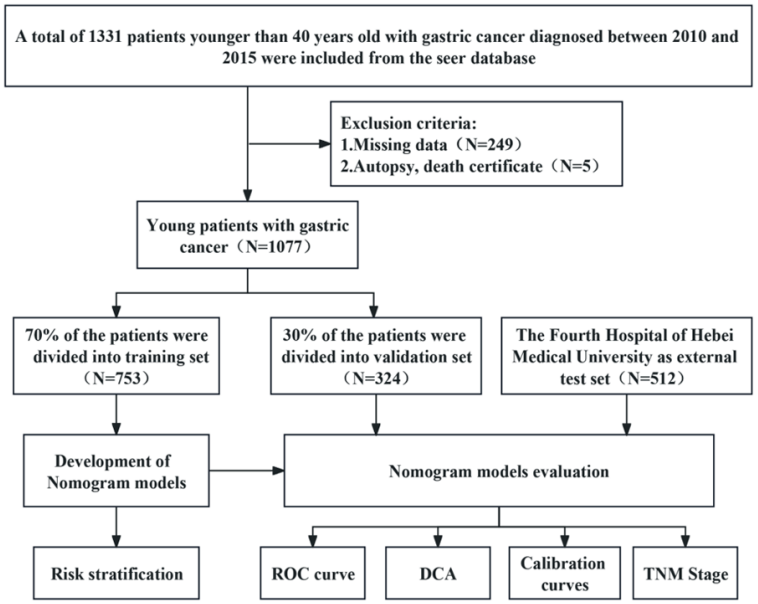

Figure 1. Flowchart of the study and statistical analysis process. SEER, Surveillance, Epidemiology and End Results; ROC, receiver operating characteristics; DCA, decision curve analysis; TNM, tumor-node-metastasis.

risk. Gastric cancer patients with bone, lung, brain, or liver metastasis had significantly lower OS and CCS. Patients who underwent radiotherapy or surgical resection of the primary tumor had better OS and CCS, while chemotherapy did not show this advantage. Tumors sized 0-1 cm displayed the worst OS and CCS, whereas tumors sized 2.1-4.0 cm showed the best OS and CCS. Females had better OS and CCS compared to males.

Cox regression analysis of the OS of the training set

Univariate regression analysis showed that 15 variables, including gender, histological subtype, grade, stage, T stage, N stage, M stage, tumor size, surgery, radiotherapy, tumor location, bone metastasis, lung metastasis, brain metastasis, and liver metastasis, were significantly associated with OS. Multivariate Cox regression analysis revealed that sex (male, P < 0.001, HR = 1.305, 95% CI = 1.126-1.514), histological type (other, P = 0.001, HR = 0.592, 95% CI = 0.438-0.799), stage (II, P < 0.001, HR = 2.918, 95% CI = 1.683-5.06; III, P < 0.001, HR = 3.788, 95% CI = 2.233-6.428; IV, P < 0.001, HR = 6.435, 95% CI = 3.918-10.568; Unknown, P = 0.011, HR = 2.122, 95% CI = 1.188-3.788), N stage (N2, P = 0.027, HR = 1.458, 95% CI = 1.043-2.037; N3, P = 0.001, HR = 1.698, 95% CI = 1.246-2.313), tumor size (2.1-4.0, P = 0.012, HR = 0.539, 95% CI = 0.333-0.871; 4.1-6.0, P = 0.04, HR = 0.607, 95% CI = 0.377-0.977), surgery (Yes, P < 0.001, HR = 0.37, 95% CI = 0.288-0.476), primary tumor site (Lesser curvature, P = 0.004, HR = 0.592, 95% CI = 0.413-0.848), and lung metastasis (Yes, P < 0.001, HR = 1.964, 95% CI = 1.481-2.606) were independent prognostic factors for OS in patients with EOGC (**Table 2**). For more detailed information, please refer to **Table 2**.

Cox regression analysis of the CSS of the training set

The single-factor regression analysis revealed that 15 variables, including gender, histological subtype, grade, stage, T stage, N stage, M stage, tumor size, surgery, radiotherapy, primary tumor site, bone metastasis, lung metastasis, brain metastasis, and liver metastasis, were significantly associated with CSS. The multi-factor Cox regression analysis showed that the following eight variables were independent prognostic factors for CSS in EOGC patients: sex (male, P = 0.001, HR = 1.281, 95% CI = 1.102-1.489), histological type (other, P < 0.001, HR = 0.567, 95% CI = 0.417-0.771), stage (II, P < 0.001, HR = 3.069, 95% CI = 1.763-5.343; III, P < 0.001, HR = 4.005, 95% CI = 2.352-6.819; IV, P < 0.001, HR = 6.804, 95% CI = 4.13-11.21; Unknown, P = 0.008, HR = 2.208, 95% CI = 1.23-3.963), N stage (N2, P = 0.046, HR = 1.41, 95% CI = 1.006-1.975; N3, P = 0.001, HR = 1.665, 95% CI = 1.218-2.276), tumor size (2.1-4.0, P = 0.017, HR = 0.551, 95% CI = 0.338-0.9), surgery (Yes, P < 0.001, HR = 0.371, 95% CI = 0.288-0.477), primary tumor site (Lesser curvature, P = 0.005, HR = 0.594, 95% CI = 0.414-0.852), and lung metastasis (Yes, P < 0.001, HR = 1.898, 95% CI = 1.426-2.526) (**Table 3**). For more detailed information, please refer to **Table 3**.

Construction of nomogram prediction models for survival prognosis in EOGC patients

We constructed two nomogram models to predict the 1-year, 3-year, and 5-year overall sur-

Table 1. Baseline clinical and pathological characteristics of early-onset gastric cancer (EOGC) patients

Variable	Training (N = 753)	Validation (N = 324)	*P*-value	Training (N = 753)	Test (N = 512)	*P*-value
Age group (years old)						
14-19	17 (2.3%)	7 (2.2%)	1	17 (2.3%)	7 (1.4%)	0.959
20-29	149 (19.8%)	66 (20.4%)		149 (19.8%)	63 (12.3%)	
30-39	587 (78.0%)	251 (77.5%)		587 (78.0%)	442 (86.3%)	
Sex						
Female	377 (50.1%)	157 (48.5%)	0.889	377 (50.1%)	208 (40.6%)	0.093
Male	376 (49.9%)	167 (51.5%)		376 (49.9%)	304 (59.4%)	
Race						
Black	80 (10.6%)	35 (10.8%)	0.322	80 (10.6%)	/	
Other	127 (16.9%)	73 (22.5%)		127 (16.9%)	/	
Unknown	10 (1.3%)	1 (0.3%)		10 (1.3%)	/	
White	536 (71.2%)	215 (66.4%)		536 (71.2%)	/	
Year						
2010-2011	241 (32.0%)	114 (35.2%)	0.522	241 (32.0%)	/	
2012-2013	258 (34.3%)	93 (28.7%)		258 (34.3%)	/	
2014-2015	254 (33.7%)	117 (36.1%)		254 (33.7%)	/	
Histologic type						
Diffuse type	383 (50.9%)	169 (52.2%)	0.196	383 (50.9%)	380 (74.2%)	0.036
Intestinal type	291 (38.6%)	136 (42.0%)		291 (38.6%)	42 (8.2%)	
Others	79 (10.5%)	19 (5.9%)		79 (10.5%)	90 (17.6%)	
Grade						
Grade I	17 (2.3%)	6 (1.9%)	0.777	17 (2.3%)	11 (2.1%)	0.728
Grade II	72 (9.6%)	21 (6.5%)		72 (9.6%)	56 (10.9%)	
Grade III	489 (64.9%)	227 (70.1%)		489 (64.9%)	437 (85.4%)	
Grade IV	13 (1.7%)	8 (2.5%)		13 (1.7%)	8 (1.6%)	
Unknown	162 (21.5%)	62 (19.1%)		162 (21.5%)	/	
Stage						
I	61 (8.1%)	30 (9.3%)	0.857	61 (8.1%)	58 (11.3%)	0.029
II	64 (8.5%)	31 (9.6%)		64 (8.5%)	109 (21.3%)	
III	140 (18.6%)	53 (16.4%)		140 (18.6%)	282 (55.1%)	
IV	414 (55.0%)	188 (58.0%)		414 (55.0%)	63 (12.3%)	
Unknown	74 (9.8%)	22 (6.8%)		74 (9.8%)	/	
T Stage						
T1-T2	171 (22.7%)	74 (22.8%)	0.993	171 (22.7%)	146 (28.5%)	0.082
T3-T4	343 (45.6%)	152 (46.9%)		343 (45.6%)	366 (71.5%)	
Unknown	239 (31.7%)	98 (30.2%)		239 (31.7%)	/	
N Stage						
N0	257 (34.1%)	118 (36.4%)	0.73	257 (34.1%)	114 (22.3%)	0.288
N1-N3	358 (47.5%)	158 (48.8%)		358 (47.5%)	398 (77.7%)	
Unknown	138 (18.3%)	48 (14.8%)		138 (18.3%)	/	
M Stage						
M0	339 (45.0%)	136 (42.0%)	0.653	339 (45.0%)	425 (83.0%)	0.086
M1	414 (55.0%)	188 (58.0%)		414 (55.0%)	87 (17.0%)	
Tumor size (cm)						
0-1	28 (3.7%)	8 (2.5%)	0.906	28 (3.7%)	2 (0.39%)	0.007
1.1-2.0	73 (9.7%)	23 (7.1%)		73 (9.7%)	4 (0.78%)	

2.1-4.0	96 (12.7%)	44 (13.6%)		96 (12.7%)	93 (18.16%)	
4.1-6.0	79 (10.5%)	42 (13.0%)		79 (10.5%)	337 (65.82%)	
6.1+	65 (8.6%)	33 (10.2%)		65 (8.6%)	76 (14.84%)	
Unknown	412 (54.7%)	174 (53.7%)		412 (54.7%)	/	
Surgery						
No	460 (61.1%)	200 (61.7%)	0.981	460 (61.1%)	99 (19.3%)	0.022
Yes	293 (38.9%)	124 (38.3%)		293 (38.9%)	413 (80.7%)	
Chemotherapy						
No	206 (27.4%)	83 (25.6%)	0.84	206 (27.4%)	/	
Yes	547 (72.6%)	241 (74.4%)		547 (72.6%)	/	
Radiotherapy						
No	582 (77.3%)	241 (74.4%)	0.588	582 (77.3%)	/	
Yes	171 (22.7%)	83 (25.6%)		171 (22.7%)	/	
Primary Site						
Lower	135 (17.9%)	62 (19.1%)	0.881	135 (17.9%)	353 (68.9%)	0.021
Middle	186 (24.7%)	67 (20.7%)		186 (24.7%)	123 (24.0%)	
Other	252 (33.5%)	109 (33.6%)		180 (23.9%)	/	
Upper	180 (23.9%)	86 (26.5%)		252 (33.5%)	36 (7.0%)	
Bone metastasis						
No	697 (92.6%)	295 (91.0%)	0.7	697 (92.6%)	479 (93.6%)	0.282
Yes	56 (7.4%)	29 (9.0%)		56 (7.4%)	33 (6.4%)	
Lung metastasis						
No	710 (94.3%)	294 (90.7%)	0.105	710 (94.3%)	504 (98.4%)	0.165
Yes	43 (5.7%)	30 (9.3%)		43 (5.7%)	8 (1.6%)	
Liver metastasis						
No	661 (87.8%)	275 (84.9%)	0.432	661 (87.8%)	468 (91.4%)	0.451
Yes	92 (12.2%)	49 (15.1%)		92 (12.2%)	44 (8.6%)	
Brain metastasis						
No	746 (99.1%)	318 (98.1%)	0.446	746 (99.1%)	510 (99.6%)	0.736
Yes	7 (0.9%)	6 (1.9%)		7 (0.9%)	2 (0.4%)	

vival (OS) and cancer-specific survival (CSS) of EOGC patients based on independent prognostic factors identified through multifactor Cox regression analysis on the training dataset (**Figures 2A-C**, **3A-C**).

Internal validation of the nomogram models for survival prognosis in EOGC patients

We performed the internal validation of the constructed nomogram models using the training set and internal validation set established by 7:3 randomization in the SEER database. DCA revealed that the nomogram models exhibited substantial clinical utility in predicting OS and CSS prognosis for EOGC patients (**Figures 2D-I**, **3D-I**). In the training dataset, the nomogram models showed significantly higher AUC values for 1-year OS (0.835 vs 0.780, P < 0.001), 1-year CSS (0.835 vs 0.777, P < 0.001), 3-year OS (0.881 vs 0.836, P = 0.0003), 3-year CSS (0.890 vs 0.837, P < 0.001), 5-year OS (0.900 vs 0.853, P = 0.0002), and 5-year CSS (0.904 vs 0.848, P < 0.001) compared to the TNM staging system (Table S1). In the internal validation dataset, the nomogram models also showed significantly higher AUC values for 1-year OS (0.831 vs 0.793, P = 0.072), 1-year CSS (0.842 vs 0.816, P = 0.190), 3-year OS (0.892 vs 0.857, P = 0.039), 3-year CSS (0.887 vs 0.848, P = 0.018), 5-year OS (0.906 vs 0.880, P = 0.133), and 5-year CSS (0.900 vs 0.876, P = 0.109) compared to the TNM staging system (Table S1). Furthermore, the Delong analysis of the ROC curves for the training and internal validation datasets revealed that the nomogram models had larger areas under the curve (AUC) for evaluating OS and CSS in patients (all > 0.8), indicating better accuracy of these models (Table S1).

Table 2. Cox regression analysis of overall survival (OS) in the training set

Variable	OS-Univariate		OS-Multivariate	
	HR.95 CI	*p*_value	HR.95 CI	*p*_value
Age group (years old)				
14-19	1 (reference)			
20-29	1.575 (0.927-2.676)	0.093		
30-39	1.636 (0.98-2.731)	0.059		
Sex				
Female	1 (reference)		1 (reference)	
Male	1.186 (1.034-1.359)	0.015	1.305 (1.126-1.514)	P < 0.001
Race				
Black	1 (reference)			
Other	0.774 (0.598-1.002)	0.052		
Unknown	0.401 (0.148-1.092)	0.074		
White	0.875 (0.703-1.088)	0.230		
Year				
2010-2011	1 (reference)			
2012-2013	1.031 (0.869-1.222)	0.729		
2014-2015	1.066 (0.902-1.26)	0.452		
Histologic type				
Diffuse type	1 (reference)		1 (reference)	
Intestinal type	0.996 (0.863-1.149)	0.955	0.876 (0.739-1.038)	0.126
Others	0.563 (0.428-0.741)	P < 0.001	0.592 (0.438-0.799)	0.001
Grade				
Grade I	1 (reference)		1 (reference)	
Grade II	1.82 (0.961-3.446)	0.066	0.879 (0.455-1.7)	0.702
Grade III	2.558 (1.408-4.649)	0.002	1.285 (0.69-2.394)	0.43
Grade IV	2.045 (0.949-4.408)	0.068	1.43 (0.648-3.154)	0.376
Unknown	2.832 (1.539-5.213)	0.001	1.327 (0.697-2.526)	0.389
Stage				
Stage I	1 (reference)		1 (reference)	
Stage II	2.477 (1.509-4.068)	P < 0.001	2.918 (1.683-5.06)	P < 0.001
Stage III	4.359 (2.801-6.783)	P < 0.001	3.788 (2.233-6.428)	P < 0.001
Stage IV	13.188 (8.636-20.139)	P < 0.001	6.435 (3.918-10.568)	P < 0.001
Unknown	3.927 (2.422-6.366)	P < 0.001	2.122 (1.188-3.788)	0.011
N Stage				
N0	1 (reference)		1 (reference)	
N1	1.514 (1.273-1.801)	P < 0.001	1.196 (0.981-1.458)	0.077
N2	1.013 (0.776-1.322)	0.927	1.458 (1.043-2.037)	0.027
N3	1.269 (1.004-1.605)	0.046	1.698 (1.246-2.313)	0.001
Unknown	1.622 (1.323-1.99)	P < 0.001	1.059 (0.833-1.345)	0.64
T Stage				
T0	1 (reference)		1 (reference)	
T1	0.132 (0.032-0.537)	0.005	0.533 (0.119-2.388)	0.411
T2	0.139 (0.034-0.573)	0.006	0.526 (0.116-2.377)	0.404
T3	0.15 (0.037-0.608)	0.008	0.405 (0.091-1.806)	0.236
T4	0.27 (0.067-1.089)	0.066	0.596 (0.134-2.64)	0.495
Tis	0.164 (0.015-1.817)	0.141	0.791 (0.065-9.602)	0.854
Unknown	0.323 (0.08-1.304)	0.113	0.517 (0.116-2.298)	0.386

M Stage				
M0	1 (reference)		1 (reference)	
M1	4.292 (3.677-5.01)	P < 0.001	NA(NA-NA)	*NA*
Tumor size (cm)				
0-1	1 (reference)		1 (reference)	
1.1-2.0	0.699 (0.441-1.107)	0.127	0.842 (0.517-1.371)	0.489
2.1-4.0	0.478 (0.305-0.749)	0.001	0.539 (0.333-0.871)	0.012
4.1-6.0	0.779 (0.499-1.215)	0.27	0.607 (0.377-0.977)	0.04
6.1+	0.766 (0.485-1.208)	0.251	0.68 (0.418-1.104)	0.119
Unknown	1.426 (0.954-2.131)	0.084	0.748 (0.482-1.161)	0.196
Primary Site				
Antrum	1 (reference)		1 (reference)	
Body	1.241 (0.953-1.616)	0.109	0.993 (0.757-1.303)	0.96
Cardia	1.119 (0.892-1.404)	0.33	0.772 (0.597-1)	0.05
Fundus of stomach	1.035 (0.695-1.541)	0.866	0.886 (0.587-1.339)	0.566
Greater curvature	1.202 (0.845-1.709)	0.306	0.9 (0.625-1.295)	0.569
Lesser curvature	0.782 (0.554-1.106)	0.164	0.592 (0.413-0.848)	0.004
Overlapping	1.487 (1.206-1.833)	0	0.947 (0.758-1.182)	0.628
Pylorus	0.619 (0.362-1.057)	0.079	0.713 (0.413-1.23)	0.224
Surgery				
None/Unknown	1 (reference)		1 (reference)	
Yes	0.264 (0.226-0.308)	P < 0.000	0.37 (0.288-0.476)	P < 0.001
Chemotherapy				
None/Unknown	1 (reference)			
Yes	1.084 (0.92-1.278)	0.334		
Radiotherapy				
None/Unknown	1 (reference)		1 (reference)	
Yes	0.734 (0.624-0.864)	P < 0.001	1.046 (0.867-1.261)	0.642
Bone metastasis				
NO	1 (reference)		1 (reference)	
Yes	2.974 (2.341-3.778)	P < 0.001	1.195 (0.922-1.55)	0.178
Lung metastasis				
NO	1 (reference)		1 (reference)	
Yes	3.888 (3.021-5.002)	P < 0.001	1.964 (1.481-2.606)	P < 0.001
Liver metastasis				
NO	1 (reference)		1 (reference)	
Yes	2.472 (2.049-2.983)	P < 0.001	1.136 (0.909-1.419)	0.262
Brain metastasis				
NO	1 (reference)		1 (reference)	
Yes	2.581 (1.491-4.47)	0.001	0.986 (0.543-1.789)	0.963

HR, hazard ratio; CI, confidence interval.

To minimize the overfitting of AUC values, calibration curves were constructed for 1-year, 3-year, and 5-year OS and CSS in EOGC patients. The calibration curves demonstrated a high degree of concordance between the predicted survival rates by the nomogram models and the observed survival rates, as indicated by their proximity to the diagonal line. This suggests good consistency between the actual observations and the survival probabilities pre-

Table 3. Cox regression analysis of cancer-specific survival (CSS) in the training set

Variable	CSS-Univariate		CSS-Multivariate	
	HR.95 CI	*p*_value	HR.95 CI	*p*_value
Age group (years old)				
14-19	1 (reference)			
20-29	1.542 (0.907-2.622)	0.109		
30-39	1.613 (0.966-2.693)	0.067		
Sex				
Female	1 (reference)		1 (reference)	
Male	1.18 (1.027-1.356)	0.019	1.281 (1.102-1.489)	0.001
Race				
Black	1 (reference)			
Other	0.774 (0.595-1.006)	0.055		
Unknown	0.318 (0.101-1.003)	0.051		
White	0.876 (0.703-1.093)	0.241		
Year				
2010-2011	1 (reference)			
2012-2013	1.033 (0.869-1.229)	0.71		
2014-2015	1.079 (0.911-1.278)	0.38		
Histologic type				
Diffuse type	1 (reference)		1 (reference)	
Intestinal type	0.994 (0.859-1.149)	0.931	0.858 (0.721-1.02)	0.082
Others	0.553 (0.418-0.731)	P < 0.001	0.567 (0.417-0.771)	P < 0.001
Grade				
Grade I	1 (reference)		1 (reference)	
Grade II	1.806 (0.952-3.427)	0.071	0.922 (0.476-1.787)	0.81
Grade III	2.528 (1.391-4.593)	0.002	1.304 (0.699-2.432)	0.404
Grade IV	1.997 (0.917-4.348)	0.082	1.58 (0.708-3.527)	0.264
Unknown	2.742 (1.488-5.053)	0.001	1.351 (0.708-2.577)	0.362
Stage				
Stage I	1 (reference)		1 (reference)	
Stage II	2.474 (1.507-4.063)	P < 0.001	3.069 (1.763-5.343)	P < 0.001
Stage III	4.328 (2.781-6.736)	P < 0.001	4.005 (2.352-6.819)	P < 0.001
Stage IV	13.13 (8.594-20.06)	P < 0.001	6.804 (4.13-11.21)	P < 0.001
Unknown	3.856 (2.373-6.267)	P < 0.001	2.208 (1.23-3.963)	0.008
N Stage				
N0	1 (reference)		1 (reference)	
N1	1.501 (1.26-1.789)	P < 0.0001	1.166 (0.955-1.424)	0.131
N2	1.005 (0.768-1.314)	0.971	1.41 (1.006-1.975)	0.046
N3	1.258 (0.993-1.594)	0.058	1.665 (1.218-2.276)	0.001
Unknown	1.585 (1.285-1.955)	P < 0.001	1.042 (0.813-1.336)	0.745
T Stage				
T0	1 (reference)		1 (reference)	
T1	0.128 (0.031-0.523)	0.004	0.527 (0.117-2.37)	0.404
T2	0.138 (0.033-0.568)	0.006	0.497 (0.11-2.251)	0.364
T3	0.151 (0.037-0.611)	0.008	0.384 (0.086-1.717)	0.21
T4	0.267 (0.066-1.078)	0.064	0.56 (0.126-2.485)	0.445
Tis	0.165 (0.015-1.826)	0.142	0.742 (0.061-9.04)	0.815
Unknown	0.322 (0.08-1.301)	0.112	0.492 (0.111-2.191)	0.352

M Stage				
M0	1 (reference)		1 (reference)	
M1	4.31 (3.686-5.04)	P < 0.001	NA(NA-NA)	*NA*
Tumor size (cm)				
0-1	1 (reference)		1 (reference)	
1.1-2.0	0.727 (0.456-1.161)	0.182	0.866 (0.527-1.423)	0.571
2.1-4.0	0.493 (0.313-0.779)	0.002	0.551 (0.338-0.9)	0.017
4.1-6.0	0.804 (0.512-1.264)	0.345	0.619 (0.381-1.005)	0.053
6.1+	0.793 (0.498-1.262)	0.327	0.716 (0.436-1.176)	0.187
Unknown	1.466 (0.972-2.209)	0.068	0.778 (0.496-1.219)	0.274
Primary Site				
Antrum	1 (reference)		1 (reference)	
Body	1.244 (0.951-1.627)	0.111	1.018 (0.772-1.341)	0.902
Cardia	1.133 (0.9-1.425)	0.287	0.785 (0.606-1.019)	0.069
Fundus of stomach	1.039 (0.69-1.567)	0.853	1.002 (0.656-1.53)	0.992
Greater curvature	1.147 (0.797-1.649)	0.461	0.848 (0.582-1.234)	0.389
Lesser curvature	0.794 (0.561-1.123)	0.193	0.594 (0.414-0.852)	0.005
Overlapping	1.482 (1.199-1.832)	P < 0.001	0.938 (0.749-1.175)	0.576
Pylorus	0.596 (0.343-1.036)	0.067	0.691 (0.394-1.212)	0.197
Surgery				
None/Unknown	1 (reference)		1 (reference)	
Yes	0.264 (0.226-0.309)	P < 0.001	0.371 (0.288-0.477)	P < 0.001
Chemotherapy				
None/Unknown	1 (reference)			
Yes	1.095 (0.927-1.295)	0.286		
Radiotherapy				
None/Unknown	1 (reference)		1 (reference)	
Yes	0.734 (0.623-0.866)	P < 0.001	1.05 (0.867-1.271)	0.62
Bone metastasis				
NO	1 (reference)		1 (reference)	
Yes	2.968 (2.329-3.783)	P < 0.001	1.179 (0.907-1.533)	0.219
Lung metastasis				
NO	1 (reference)		1 (reference)	
Yes	3.84 (2.973-4.959)	P < 0.001	1.898 (1.426-2.526)	P < 0.001
Liver metastasis				
NO	1 (reference)		1 (reference)	
Yes	2.486 (2.057-3.005)	P < 0.001	1.127 (0.899-1.414)	0.3
Brain metastasis				
NO	1 (reference)		1 (reference)	
Yes	2.62 (1.513-4.538)	0.001	1.005 (0.554-1.824)	0.987

dicted by the nomogram models, indicating the models possess robust discriminatory capabilities (**Figures 2J-O**, **3J-O**).

External validation of the nomogram models for survival prognosis in EOGC patients

We included EOGC patients from the Fourth Hospital of Hebei Medical University as an external test set to validate these constructed nomogram models. The time-dependent ROC curve showed lower accuracy in predicting OS and CSS prognosis for EOGC patients using these nomogram models (**Figure 4A**, **4H**). AUC for the internal validation set and the external test set showed no statistical differences between the external testing set and the internal validation set (1-year OS: 0.831 vs 0.686, P

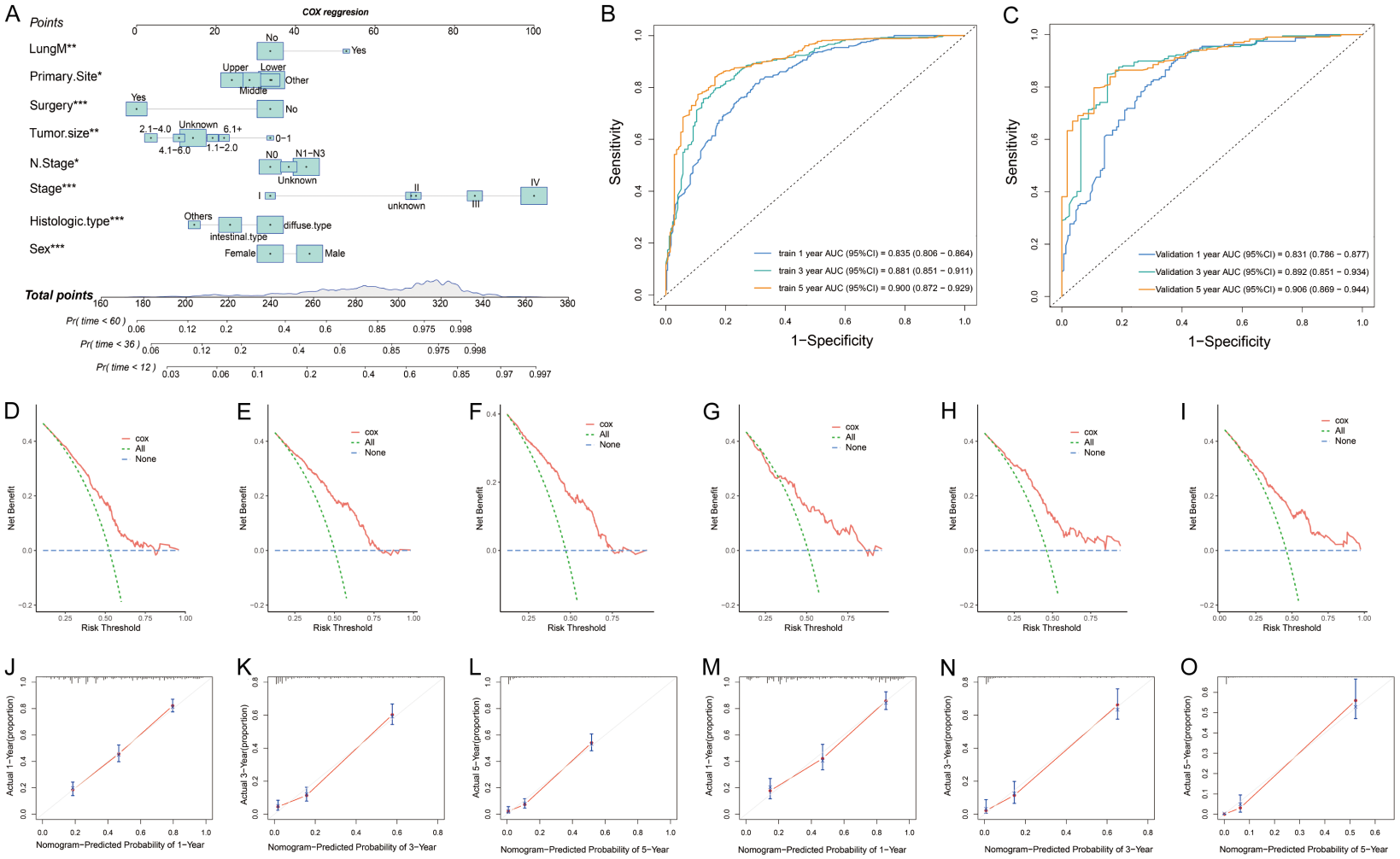

Figure 2. Construction and validation of the nomogram model for predicting overall survival (OS). A prognostic nomogram for predicting the 1-year, 3-year, and 5-year OS of Early-onset gastric cancer (EOGC) patients (A). The receiver operating characteristics (ROC) curves of the nomogram for the 1-year, 3-year, and 5-year OS in the training set (B). The ROC curves of the nomogram for the 1-year, 3-year, and 5-year OS in the validation set (C). The decision curve analysis (DCA) of 1-year (D), 3-year (E), and 5-year (F) OS in the training set. The DCA of 1-year (G), 3-year (H), and 5-year (I) OS in the validation set. The calibration curves for 1-year (J), 3-year (K), and 5-year (L) OS in the training set. The calibration curves for 1-year (M), 3-year (N), and 5-year (O) OS in the validation set.

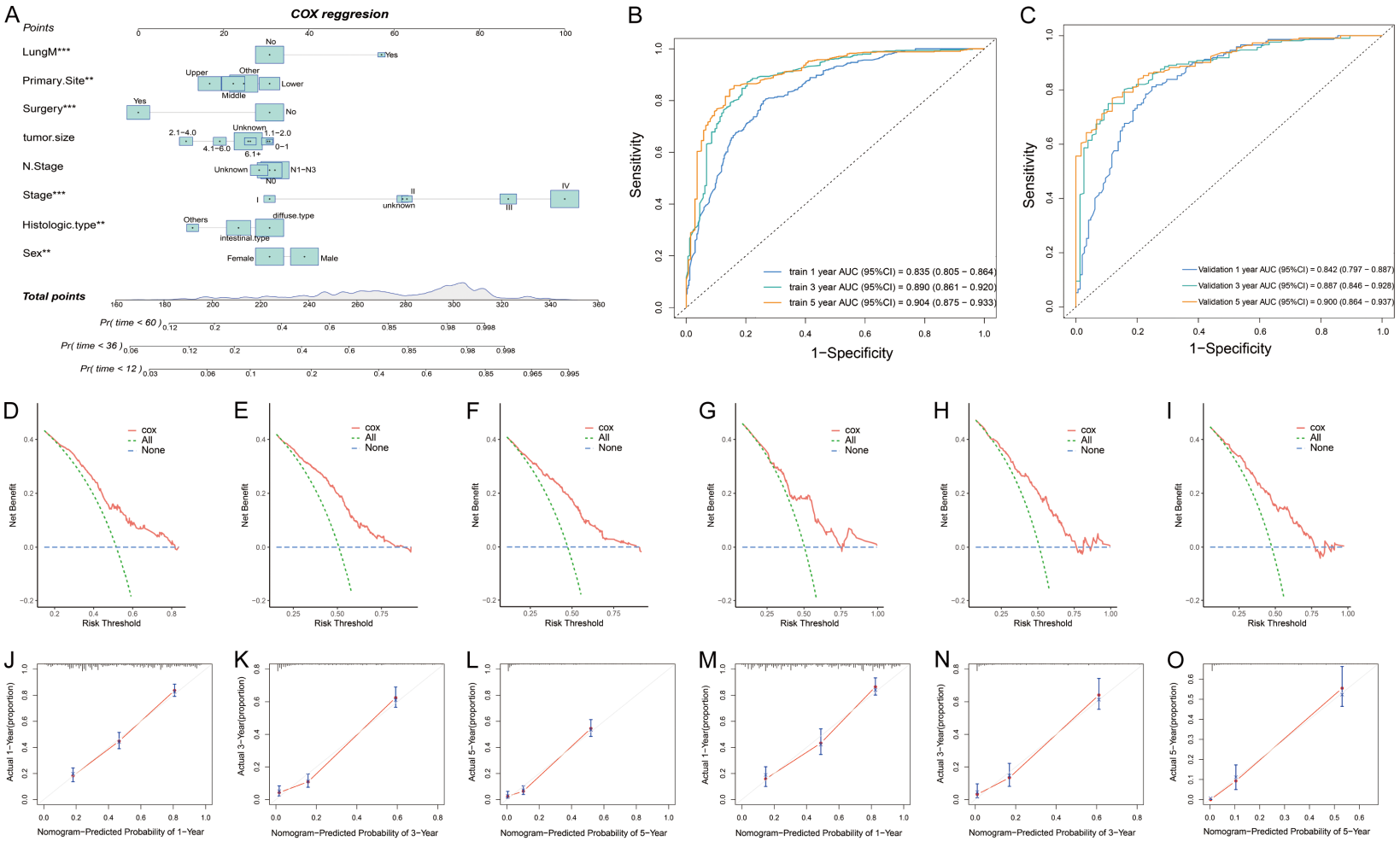

Figure 3. Construction and validation of the nomogram model for predicting CSS. A prognostic nomogram for predicting the 1-year, 3-year, and 5-year CSS of EOGC patients (A). The ROC curves of the nomogram for the 1-year, 3-year, and 5-year CSS in the training set (B). The ROC curves of the nomogram for the 1-year, 3-year, and 5-year CSS in the validation set (C). The DCA of 1-year (D), 3-year (E), and 5-year (F) CSS in the training set. The DCA of 1-year (G), 3-year (H), and 5-year (I) CSS in the validation set. The calibration curves for 1-year (J), 3-year (K), and 5-year (L) CSS in the training set. The calibration curves for 1-year (M), 3-year (N), and 5-year (O) CSS in the validation set.

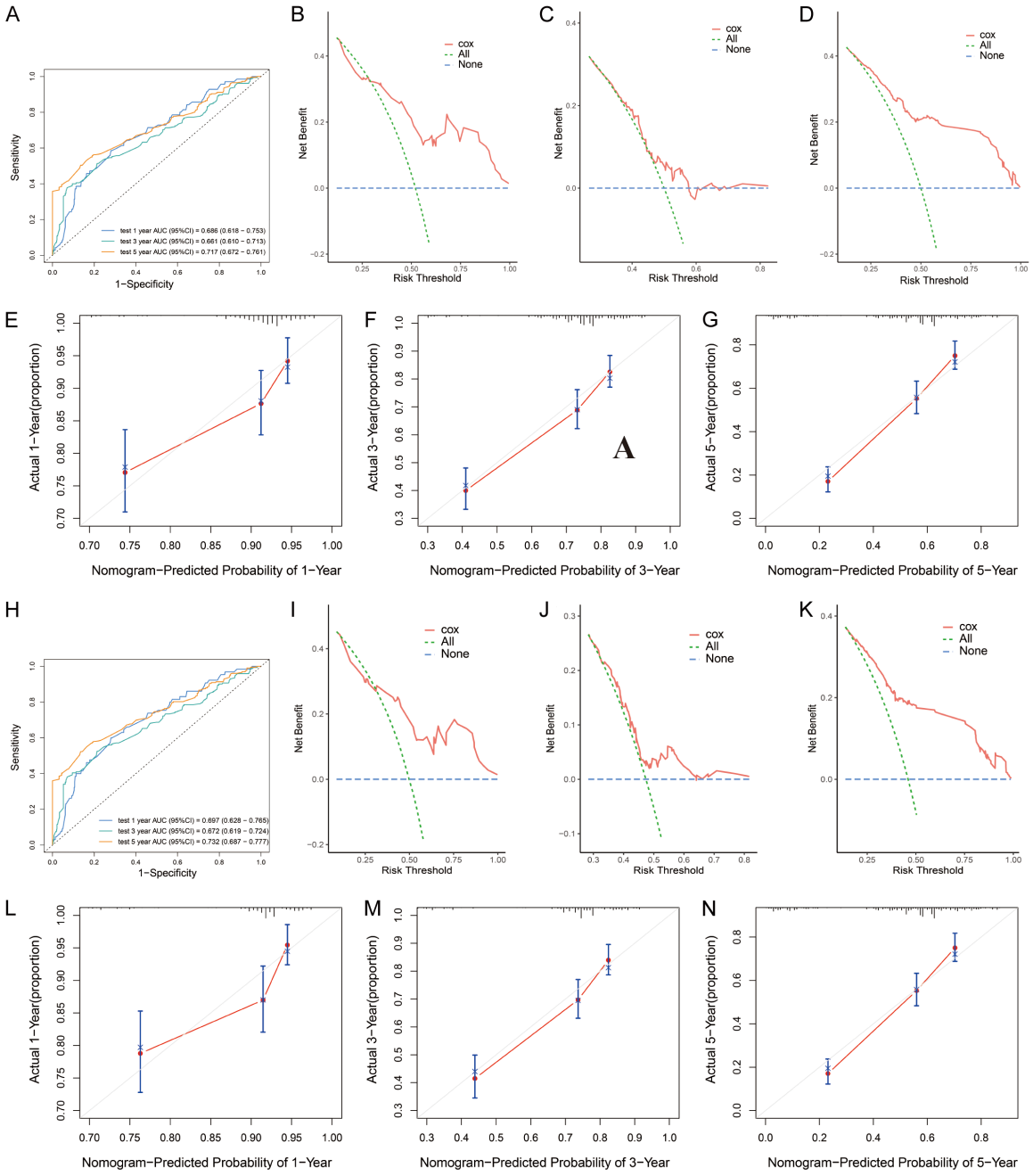

Figure 4. External testing of nomograms. ROC curves of the nomogram predicting 1-, 3-, and 5-year OS in the external test set (A). The DCA of 1-year (B), 3-year (C), and 5-year (D) OS in the external test set. The calibration curves for 1-year (E), 3-year (F), and 5-year (G) OS in the external test set. ROC curves of the nomogram predicting 1-, 3-, and 5-year CSS in the external test set (H). The DCA of 1-year (I), 3-year (J), and 5-year (K) CSS in the external test set. The calibration curves for 1-year (L), 3-year (M), and 5-year (N) CSS in the external test set.

= 0.483; CSS: 0.842 vs 0.697, P = 0.469. 3-year OS: 0.892 vs 0.661, P = 0.896; CSS: 0.887 vs 0.672, P = 0.868. 5-year OS: 0.906 vs 0.717, P = 0.336; CSS: 0.900 vs 0.732, P = 0.381), indicating that the external adaptability of the models needs improvement (Table S2). The DCA curves for the external testing set demonstrated that the constructed nomogram models possess notable clinical utility. Calibration curves for 1-year, 3-year, and 5-year OS and CSS were developed using external testing set data. The nomogram models exhibited strong consistency between observed and predicted survival probabilities in the external

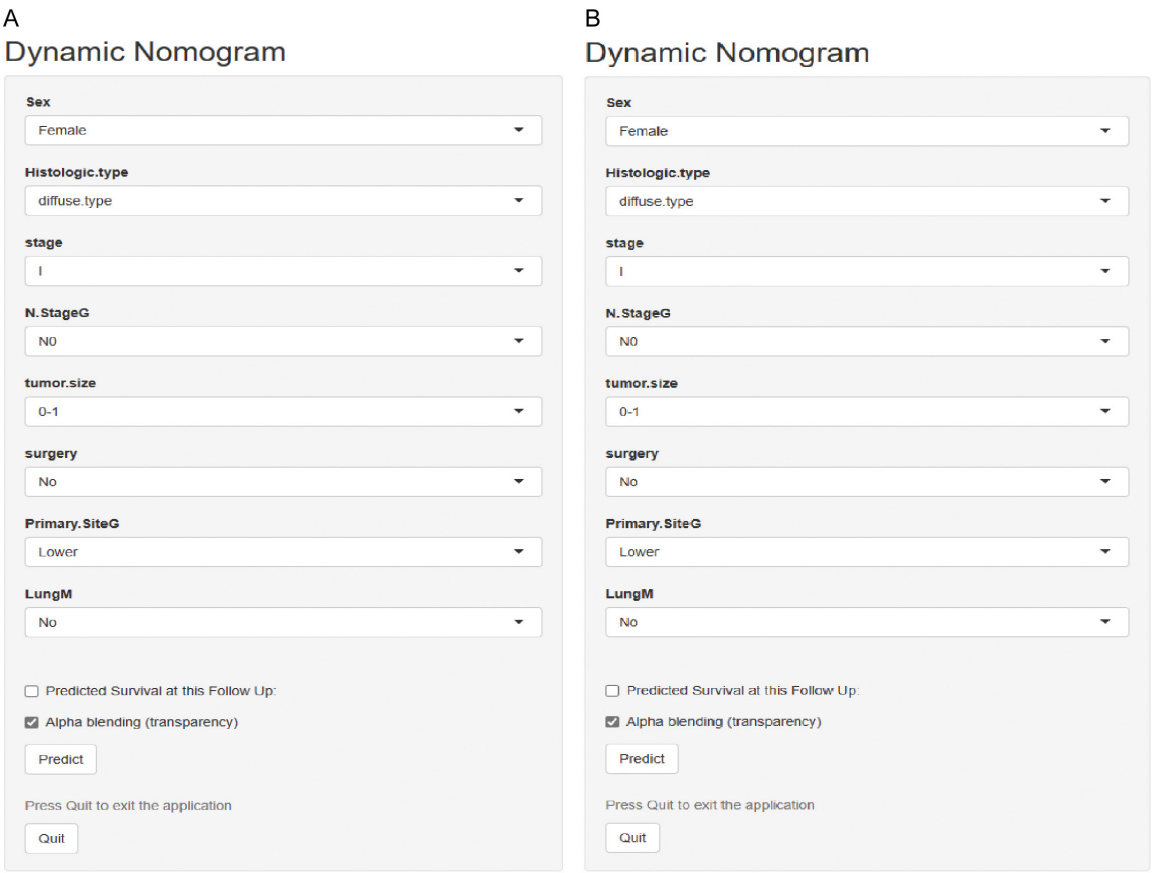

Figure 5. Online calculator of nomogram model for predicting OS in EOGC (A); Online calculator of nomogram model for predicting CSS in EOGC (B).

testing set. This indicates that the constructed nomogram models maintain robust discriminatory capabilities for external data as well (**Figure 4B-G**, **4I-N**).

Construction of online calculators for the survival prognosis nomogram models for EOGC patients

We developed an online calculator for the OS nomogram model of EOGC patients (https://nxm156.shinyapps.io/DynNomYoungGCOS/) (**Figure 5A**) to visualize the model. Similarly, we created an online calculator for the CSS nomogram model of EOGC patients (https://nxm156.shinyapps.io/DynNomapp/) (**Figure 5B**) to provide a more intuitive display of the model.

Risk stratification system for predicting the prognosis of EOGC patients using nomogram models

Using X-tile software, EOGC patients in the training set were stratified into three risk groups (**Figure 6A**) based on the total score of overall survival (OS) predicted by the nomogram model: low risk (score < 142.5), intermediate risk (score: 142.5-207.0), and high risk (> 207.0) groups. Kaplan-Meier curves for these groups were then generated to provide both the overall and stratified analysis across the validation set and external test sets (**Figure 6B-O**). Kaplan-Meier survival analysis revealed that the low-risk and intermediate-risk groups had better OS compared to the high-risk group. Additionally, the nomogram model demonstrated excellent discriminatory ability among the three risk groups ($P < 0.05$). Similarly, using X-tile software, EOGC patients in the training set were divided into three risk groups (**Figure 7A**) based on the cancer-specific survival (CSS) scores predicted by the nomogram model: low risk (score < 131.1), intermediate risk (score: 131.1-217.3), and high risk (> 217.3) groups. Kaplan-Meier curves were plotted for the overall and stratified analysis of the validation set

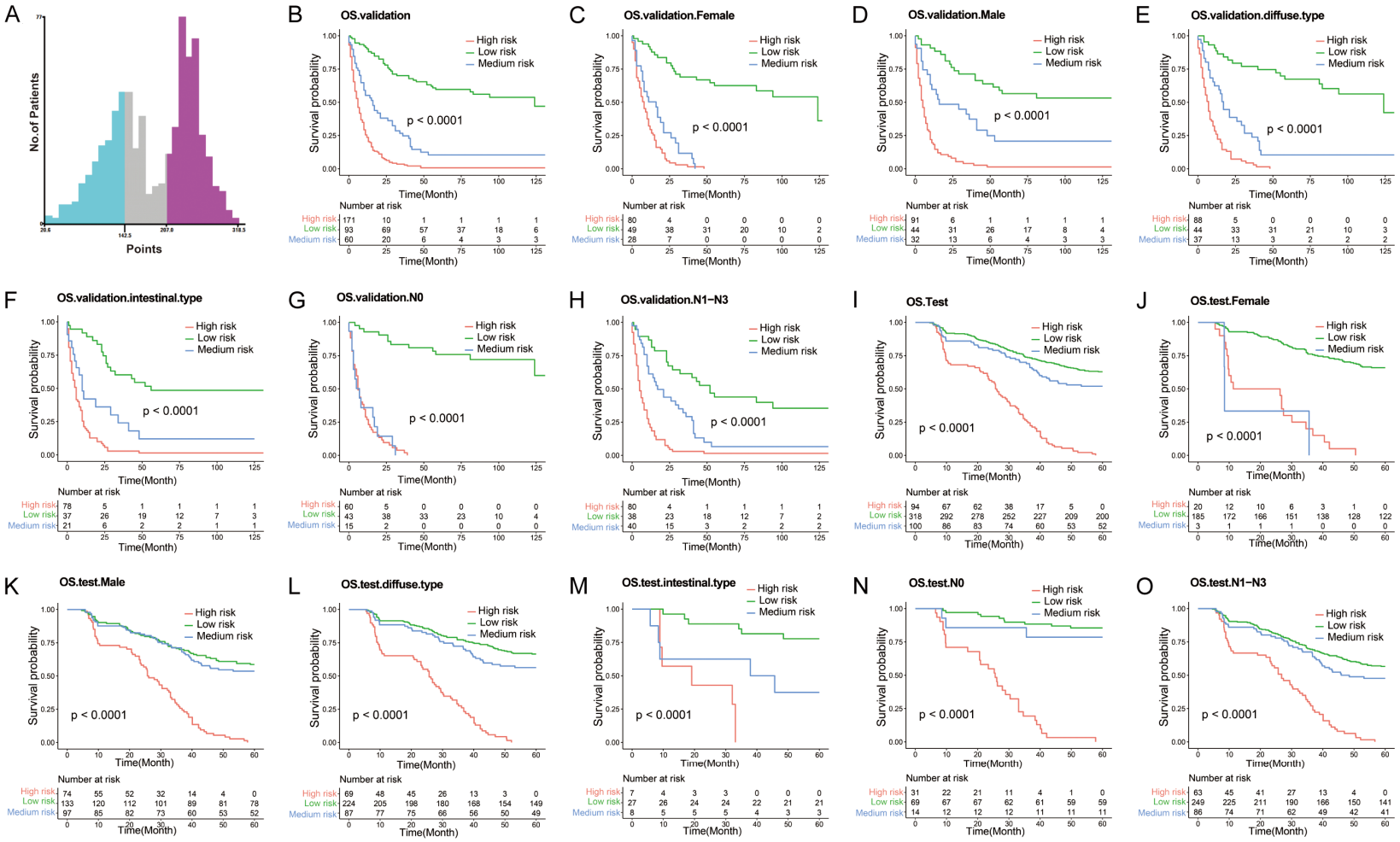

Figure 6. Risk stratification based on the total score of the nomogram model in the training set (A). Using the Kaplan-Meier (KM) and log-rank test methods, survival curves for the OS in different risk groups in the internal validation set were plotted, including all patients with EOGC (B), female patients (C), male patients (D), diffuse-type patients (E), intestinal-type patients (F), N0 stage patients (G), and N1-N3 stage patients (H). Similarly, survival curves for the OS in different risk groups in the external test set were plotted, including all patients with EOGC (I), female patients (J), male patients (K), diffuse-type patients (L), intestinal-type patients (M), N0 stage patients (N), and N1-N3 stage patients (O).

and external test set, and showed that the low-risk and intermediate-risk groups had better CSS compared to the high-risk group (**Figure 7B-0**). Furthermore, the nomogram model exhibited excellent discriminatory ability among the three risk groups ($P < 0.05$).

Discussion

The average age at gastric cancer diagnosis is typically above 60, with only a small percentage (6-7%) of patients being diagnosed before the age of 50, and an even smaller percentage (2%) diagnosed before the age of 40 [12]. While the global incidence of gastric cancer is decreasing, there is a rising trend in the incidence of EOGC [4]. Studies have indicated that most EOGC patients are diagnosed at an advanced stage, often with a diffuse histological subtype [13], which aligns with the findings of our study (diffuse type: 59.3% vs intestinal type: 39.6%). Younger patients with gastric cancer tend to exhibit faster tumor growth, increased metastasis, poorer prognosis, and higher resistance to traditional chemotherapy. Pathologically, these patients often present with poorly differentiated tumors, signet ring cell carcinoma, and the diffuse type [14]. In our study, older age at diagnosis was associated with later stage, higher OS risk, and higher CSS risk. Patients with bone, lung, brain, and liver metastases had significantly lower OS and CSS rates. Interestingly, patients with tumors measuring 0-1 cm showed the worst OS and CSS rates, emphasizing the importance of early detection in this subgroup. A study [15] reported a global increase in gender disparity in gastric cancer from 1990 to 2017, with the male-to-female age-standardized incidence ratio (ASR) rising from 1.86 to 2.20. However, our study did not observe significant gender differences in EOGC patients (males: 50.4% vs females: 49.6%).

The nomogram model, incorporating several informative variables, serves as an intuitive and easily accessible tool for physicians to make accurate diagnoses and predict survival outcomes [16]. In our study, we developed two nomogram models specifically for EOGC patients, utilizing independent prognostic factors identified through Cox regression analysis. The internal validation of these models conducted using ROC curves, calibration curves, and decision curve analysis (DCA), demonstrated high accuracy. When compared to the AJCC guidelines, the nomogram models exhibited superior predictive performance. External validation of the OS and CSS nomogram models on test sets showed moderate adaptability, although the AUC values were slightly lower than those obtained in the internal validation sets, suggesting the potential for further improvement in predictive accuracy.

Studies have reported that in addition to common risk factors shared with ordinary gastric cancer, such as genetic susceptibility expressed through single nucleotide polymorphisms, various acquired mutations (chromosomal instability, microsatellite instability, somatic gene mutations, epigenetic changes), and environmental factors (e.g., Helicobacter pylori infection, diet, smoking, and Epstein-Barr virus infection), EOGC may be more closely associated with molecular susceptibility and genetic background [17]. Another study reported a significant association between the rs10052016 locus on 5p15 and gastric cancer risk, particularly early-onset gastric cancer, with a greater protective effect of the rs10052016-G allele observed in young individuals [18]. Sugimoto et al. [19] first described a de novo large genomic deletion in the CDH1 gene associated with early onset diffuse gastric cancer, suggesting that clinicians should consider CDH1 germline mutations when encountering relatively young patients with multiple signet ring cell carcinomas. Bacani et al. [20] discovered microsatellite instability in at least one marker in 30% of EOGC cases, with approximately 1% of EOGC patients having germline MMR mutations. In a study by Milne et al. [21], it was suggested that genetic factors may have a greater influence on young patients compared to older patients, with a relatively smaller impact from environmental carcinogens. Therefore, future research should further investigate the inclusion of genetic factors in prognostic analysis.

Compared to elderly patients, there is still limited data available regarding the treatment of early-onset gastric cancer (EOGC) patients. Joshi et al. mentioned in their review that treatment options for gastric cancer include systemic chemotherapy, radiation therapy, surgery, immunotherapy, and targeted therapy. However, the choice of treatment depends primarily on the patient's physical condition, comorbidities, and the potential toxic side effects of the treatment plan [22, 23]. Analyzing the SEER data-

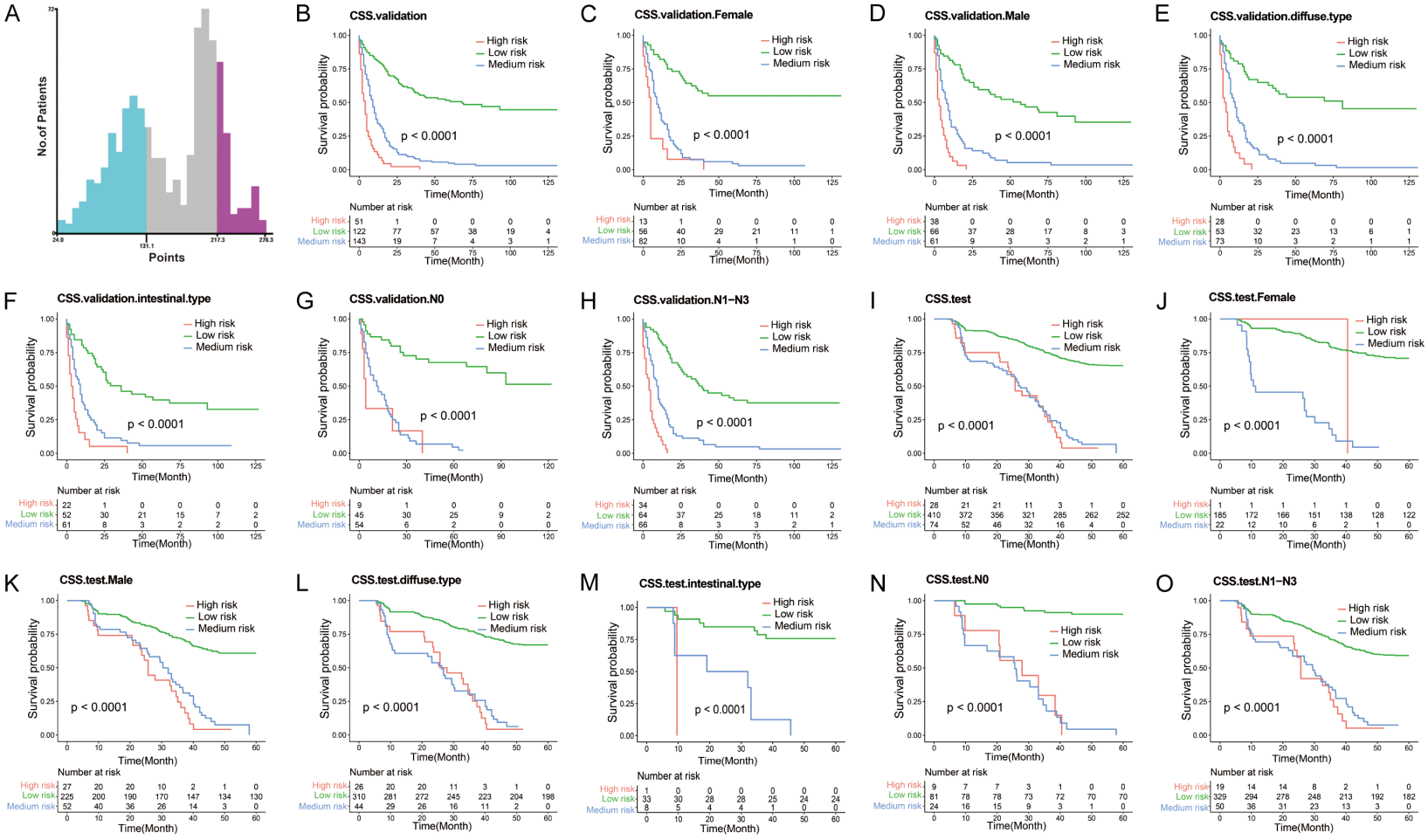

Figure 7. Risk stratification based on the total score of the nomogram model in the training set (A). Using the Kaplan-Meier and log-rank test methods, survival curves for the CSS in different risk groups in the validation set were plotted, including all patients with EOGC (B), female patients (C), male patients (D), diffuse-type patients (E), intestinal-type patients (F), N0 stage patients (G), and N1-N3 stage patients (H). Similarly, survival curves for the CSS in different risk groups in the external test set were plotted, including all patients with EOGC (I), female patients (J), male patients (K), diffuse-type patients (L), intestinal-type patients (M), N0 stage patients (N), and N1-N3 stage patients (O).

base, Al-Refaie et al. [24] found that EOGC patients were more likely to undergo surgical treatment compared to elderly gastric cancer patients. They also observed a significant correlation between increasing age and decreased use of adjuvant radiotherapy among patients who underwent surgery. A study conducted in China [25] analyzed data and showed that chemotherapy provided greater survival benefits for elderly patients compared to younger and middle-aged patients with stage II and III gastric cancer, indicating potential overuse of chemotherapy in the young group. Another study [26] found that EOGC patients had similar long-term survival outcomes with surgery alone, surgery plus radiation therapy or chemotherapy, and surgery plus chemoradiotherapy. There was no significant statistical difference in survival outcomes among these three treatment approaches, indicating that additional radiation therapy, chemotherapy, or chemo-radiotherapy did not provide coordinated survival benefits. Consistent with previous research conclusions, our study revealed that patients who underwent radiation therapy or surgical resection of the primary lesion had better OS and CSS, while chemotherapy did not demonstrate this advantage. Therefore, the choice of treatment for EOGC patients, particularly chemotherapy, should be thoroughly evaluated based on the patient's physical condition and the potential toxic side effects of chemotherapy regimens. Therefore, when it comes to the choice of treatment for EOGC patients, especially regarding chemotherapy, it is crucial to evaluate the patient's physical condition and the toxic side effects of chemotherapy regimens comprehensively and thoroughly. This evaluation is necessary to provide individualized treatment and improve the patient's overall survival benefits.

It has been reported [27] that when tumors in stage IV gastric cancer patients are unresectable, receiving various combination chemotherapy regimens followed by radical gastrectomy can lead to favorable survival outcomes. Radical conversion surgery for unresectable stage IV gastric cancer, when R0 resection can be achieved, significantly improves survival rates [27, 28]. Cox regression analysis conducted by Morgagni et al. [29] found that radical conversion surgery is an independent factor positively correlated with survival rates. Solaini et al. also found that unresectable stage IV gastric cancer patients who undergo radical surgery after chemotherapy achieve longer survival periods [30]. Crew et al. mentioned that patients with tumors in higher positions often have a worse prognosis and higher surgical mortality rates [31]. This may be attributed to the surgical procedure itself. In total gastrectomy, the difficulty of esophago-jejunal anastomosis is higher, and postoperative complications such as anastomotic stenosis and leakage may occur, leading to poorer prognosis for these patients [32]. Therefore, for patients with tumors located in higher positions, thorough preoperative assessment, careful selection of surgical approaches, and proactive management of postoperative complications are crucial to achieve better survival outcomes.

Although the nomogram models developed in this study exhibited favorable prognostic predictive performance, there are certain limitations. Firstly, due to the retrospective nature of our study, there may be some missing key information that could impact the accuracy of prognostic predictions. Additionally, the SEER database lacks specific details regarding surgical procedures and chemotherapy regimens, which could further affect the accuracy of prognostic predictions for gastric cancer patients. Secondly, it has been demonstrated that patients who undergo more extensive lymph node examination tend to have higher survival rates, and the examination of examined lymph nodes (ELN) is an independent prognostic factor for gastric cancer [33]. However, our prognostic prediction models did not incorporate postoperative clinical and pathological indicators such as lymph node examination. Thirdly, factors such as marital status, economic conditions, and living environment (rural or urban) were not included in the regression analysis, potentially introducing bias into our study. Fourthly, studies have indicated that early-onset gastric cancer is more likely to be associated with Epstein-Barr virus or genomic stable subtype, while late-onset gastric cancer is more likely to be associated with microsatellite instability subtype [34]. Furthermore, alternative splicing has been reported to play an important role in EOGC, including the regulation of specific protein modifications and the reshaping of the cancer immune microenvironment [35]. However, the predictive models developed in this study did not incorporate genetic and molecu-

lar biology characteristics. Therefore, in future research, we aim to incorporate social factors, specific treatment information, postoperative clinical and pathological indicators, molecular biology information, and other relevant factors into the predictive models to further enhance its accuracy.

Conclusion

In conclusion, we found that gender, histological subtype, stage, lymph node involvement, tumor size, surgery, tumor location, and lung metastasis are independent prognostic factors for overall survival (OS) and cancer-specific survival (CSS) in patients with EOGC. Additionally, we have developed stable nomogram models using the SEER database that can predict the 1-year, 3-year, and 5-year OS and CSS of EOGC patients. These models have demonstrated good predictive performance in internal validation and have shown some adaptability in external validation, providing more personalized and convenient assistance for survival prediction and clinical decision-making for EOGC patients.

Acknowledgements

This work was supported by the Cultivating Outstanding Talents Project of Hebei Provincial Government Fund (2019012, ZF2023047); and Hebei public health committee county-level public hospitals suitable health technology promotion and storage project (No. 2019024).

Informed consent was performed for all the study subjects.

Disclosure of conflict of interest

None.

Abbreviations

OS, overall survival; CSS, cancer-specific survival; EOGC, early-onset gastric cancer; SEER, the Surveillance, Epidemiology and End Results; ROC, receiver operating characteristics; DCA, decision curve analysis; AJCC, the American Joint Committee on Cancer; ASR, age-standardized incidence rate; MSKCC, the Memorial Sloan Kettering Cancer Center; AUC, the area under the curve; KM, Kaplan-Meier; HR, Hazard ratio; ELN, examined lymph nodes; TNM, tumor-node-metastasis; CI, confidence interval.

Address correspondence to: Qun Zhao, Third Department of Surgery, The Fourth Hospital of Hebei Medical University, Shijiazhuang 050011, Hebei, China. Tel: +86-311-86095688; Fax: +86-311-86095688; E-mail: zhaoqun@hebmu.edu.cn

References

[1] Smyth EC, Nilsson M, Grabsch HI, van Grieken NC and Lordick F. Gastric cancer. Lancet 2020; 396: 635-648.

[2] López MJ, Carbajal J, Alfaro AL, Saravia LG, Zanabria D, Araujo JM, Quispe L, Zevallos A, Buleje JL, Cho CE, Sarmiento M, Pinto JA and Fajardo W. Characteristics of gastric cancer around the world. Crit Rev Oncol Hematol 2023; 181: 103841.

[3] Milne AN and Offerhaus GJ. Early-onset gastric cancer: learning lessons from the young. World J Gastrointest Oncol 2010; 2: 59-64.

[4] Anderson WF, Camargo MC, Fraumeni JF Jr, Correa P, Rosenberg PS and Rabkin CS. Age-specific trends in incidence of noncardia gastric cancer in US adults. JAMA 2010; 303: 1723-1728.

[5] Thrumurthy SG, Chaudry MA, Chau I and Allum W. Does surgery have a role in managing incurable gastric cancer? Nat Rev Clin Oncol 2015; 12: 676-682.

[6] Kattan MW, Karpeh MS, Mazumdar M and Brennan MF. Postoperative nomogram for disease-specific survival after an R0 resection for gastric carcinoma. J Clin Oncol 2003; 21: 3647-3650.

[7] Mo H, Li P and Jiang S. A novel nomogram based on cardia invasion and chemotherapy to predict postoperative overall survival of gastric cancer patients. World J Surg Oncol 2021; 19: 256.

[8] Dong Z, Zhang Y, Geng H, Ni B, Xia X, Zhu C, Liu J and Zhang Z. Development and validation of two nomograms for predicting overall survival and cancer-specific survival in gastric cancer patients with liver metastases: a retrospective cohort study from SEER database. Transl Oncol 2022; 24: 101480.

[9] Liu H, Li Z, Zhang Q, Li Q, Zhong H, Wang Y, Yang H, Li H, Wang X, Li K, Wang D, Kong X, He Z, Wang W, Wang L, Zhang D, Xu H, Yang L, Chen Y, Zhou Y and Xu Z. Multi-institutional development and validation of a nomogram to predict prognosis of early-onset gastric cancer patients. Front Immunol 2022; 13: 1007176.

[10] Dubecz A, Gall I, Solymosi N, Schweigert M, Peters JH, Feith M and Stein HJ. Temporal trends in long-term survival and cure rates in esophageal cancer: a SEER database analysis. J Thorac Oncol 2012; 7: 443-447.

[11] Li ZY, Zhang QW, Teng LM, Zhang CH and Huang Y. Comparable rates of lymph node metastasis and survival between diffuse type and intestinal type early gastric cancer patients: a large population-based study. Gastrointest Endosc 2019; 90: 84-95, e10.
[12] Kluijt I, Sijmons RH, Hoogerbrugge N, Plukker JT, de Jong D, van Krieken JH, van Hillegersberg R, Ligtenberg M, Bleiker E and Cats A; Dutch Working Group on Hereditary Gastric Cancer. Familial gastric cancer: guidelines for diagnosis, treatment and periodic surveillance. Fam Cancer 2012; 11: 363-369.
[13] Pocurull A, Herrera-Pariente C, Carballal S, Llach J, Sánchez A, Carot L, Botargues JM, Cuatrecasas M, Ocaña T, Balaguer F, Bujanda L and Moreira L. Clinical, molecular and genetic characteristics of early onset gastric cancer: analysis of a large multicenter study. Cancers (Basel) 2021; 13: 3132.
[14] Huang Q, Zheng X, Jiao Y, Lei Y, Li X, Bi F, Guo F, Wang G and Liu M. A distinct clinicopathological feature and prognosis of young gastric cancer patients aged ≤ 45 years old. Front Oncol 2021; 11: 674224.
[15] Lou L, Wang L, Zhang Y, Chen G, Lin L, Jin X, Huang Y and Chen J. Sex difference in incidence of gastric cancer: an international comparative study based on the Global Burden of Disease Study 2017. BMJ Open 2020; 10: e033323.
[16] Liao F, Guo X, Lu X and Dong W. A validated survival nomogram for early-onset diffuse gastric cancer. Aging (Albany NY) 2020; 12: 13160-13171.
[17] Skierucha M, Milne AN, Offerhaus GJ, Polkowski WP, Maciejewski R and Sitarz R. Molecular alterations in gastric cancer with special reference to the early-onset subtype. World J Gastroenterol 2016; 22: 2460-2474.
[18] Du J, Xu Y, Dai J, Ren C, Zhu C, Dai N, Ma H, Shi Y, Hu Z, Lin D, Shen H and Jin G. Genetic variants at 5p15 are associated with risk and early onset of gastric cancer in Chinese populations. Carcinogenesis 2013; 34: 2539-2542.
[19] Sugimoto S, Yamada H, Takahashi M, Morohoshi Y, Yamaguchi N, Tsunoda Y, Hayashi H, Sugimura H and Komatsu H. Early-onset diffuse gastric cancer associated with a de novo large genomic deletion of CDH1 gene. Gastric Cancer 2014; 17: 745-749.
[20] Bacani J, Zwingerman R, Di Nicola N, Spencer S, Wegrynowski T, Mitchell K, Hay K, Redston M, Holowaty E, Huntsman D, Pollett A, Riddell R and Gallinger S. Tumor microsatellite instability in early onset gastric cancer. J Mol Diagn 2005; 7: 465-477.
[21] Milne AN, Sitarz R, Carvalho R, Carneiro F and Offerhaus GJ. Early onset gastric cancer: on the road to unraveling gastric carcinogenesis. Curr Mol Med 2007; 7: 15-28.
[22] LaPelusa M, Shen C, Gillaspie EA, Cann C, Lambright E, Chakravarthy AB, Gibson MK and Eng C. Variation in treatment patterns of patients with early-onset gastric cancer. Cancers (Basel) 2022; 14: 3633.
[23] Joshi SS and Badgwell BD. Current treatment and recent progress in gastric cancer. CA Cancer J Clin 2021; 71: 264-279.
[24] Al-Refaie WB, Hu CY, Pisters PW and Chang GJ. Gastric adenocarcinoma in young patients: a population-based appraisal. Ann Surg Oncol 2011; 18: 2800-2807.
[25] Jiang Y, Xie J, Huang W, Chen H, Xi S, Li T, Chen C, Sun Z, Hu Y, Liu W, Yu J, Zhou Z, Cai S and Li G. Chemotherapy use and survival among young and middle-aged patients with gastric cancer. Clin Transl Gastroenterol 2020; 11: e00253.
[26] Zhang C, Tang R, Zhu H, Ge X, Wang Y, Wang X and Miao L. Comparison of treatment strategies and survival of early-onset gastric cancer: a population-based study. Sci Rep 2022; 12: 6288.
[27] Zhang F, Huang X, Song Y, Gao P, Zhou C, Guo Z, Shi J, Wu Z and Wang Z. Conversion surgery for stage IV gastric cancer. Front Oncol 2019; 9: 1158.
[28] Yamaguchi K, Yoshida K, Tanahashi T, Takahashi T, Matsuhashi N, Tanaka Y, Tanabe K and Ohdan H. The long-term survival of stage IV gastric cancer patients with conversion therapy. Gastric Cancer 2018; 21: 315-323.
[29] Morgagni P, Solaini L, Framarini M, Vittimberga G, Gardini A, Tringali D, Valgiusti M, Monti M and Ercolani G. Conversion surgery for gastric cancer: a cohort study from a western center. Int J Surg 2018; 53: 360-365.
[30] Solaini L, Ministrini S, Bencivenga M, D'Ignazio A, Marino E, Cipollari C, Molteni B, Mura G, Marrelli D, Graziosi L, Roviello F, De Manzoni G, Tiberio GAM and Morgagni P. Conversion gastrectomy for stage IV unresectable gastric cancer: a GIRCG retrospective cohort study. Gastric Cancer 2019; 22: 1285-1293.
[31] Crew KD and Neugut AI. Epidemiology of gastric cancer. World J Gastroenterol 2006; 12: 354-362.
[32] Guo Z, Deng C, Zhang Z, Liu Y, Qi H and Li X. Safety and effectiveness of overlap esophagojejunostomy in totally laparoscopic total gastrectomy for gastric cancer: a systematic review and meta-analysis. Int J Surg 2022; 102: 106684.
[33] Li H, Lin D, Yu Z, Li H, Zhao S, Hainisayimu T, Liu L and Wang K. A nomogram model based on the number of examined lymph nodes-related signature to predict prognosis and guide

clinical therapy in gastric cancer. Front Immunol 2022; 13: 947802.

[34] Bergquist JR, Leiting JL, Habermann EB, Cleary SP, Kendrick ML, Smoot RL, Nagorney DM, Truty MJ and Grotz TE. Early-onset gastric cancer is a distinct disease with worrisome trends and oncogenic features. Surgery 2019; 166: 547-555.

[35] Zhang J, Goel A and Zhu L. Identification of novel alternative splicing events associated with tumorigenesis, protein modification, and immune microenvironment in early-onset gastric cancer. Front Oncol 2021; 11: 640272.

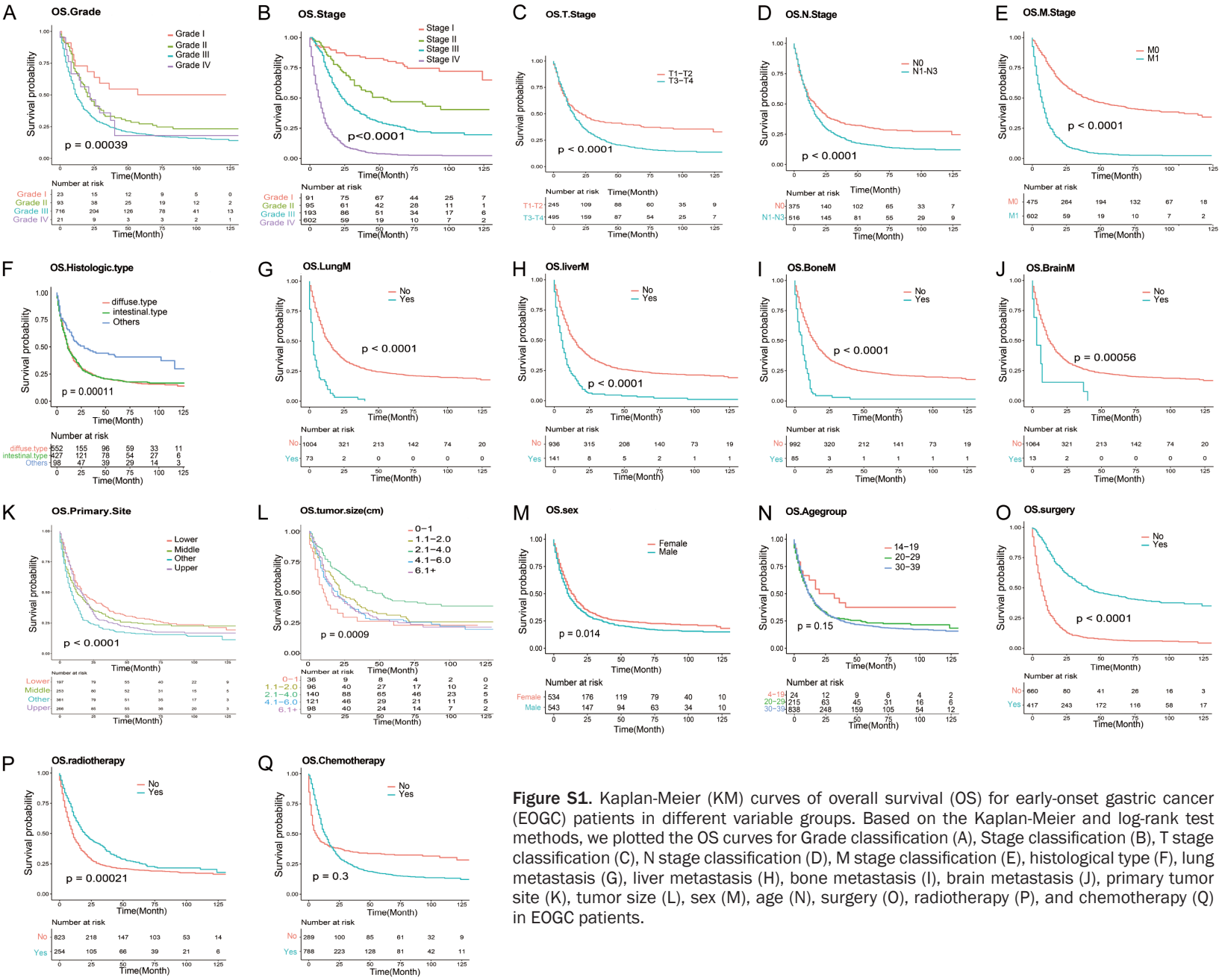

Figure S1. Kaplan-Meier (KM) curves of overall survival (OS) for early-onset gastric cancer (EOGC) patients in different variable groups. Based on the Kaplan-Meier and log-rank test methods, we plotted the OS curves for Grade classification (A), Stage classification (B), T stage classification (C), N stage classification (D), M stage classification (E), histological type (F), lung metastasis (G), liver metastasis (H), bone metastasis (I), brain metastasis (J), primary tumor site (K), tumor size (L), sex (M), age (N), surgery (O), radiotherapy (P), and chemotherapy (Q) in EOGC patients.

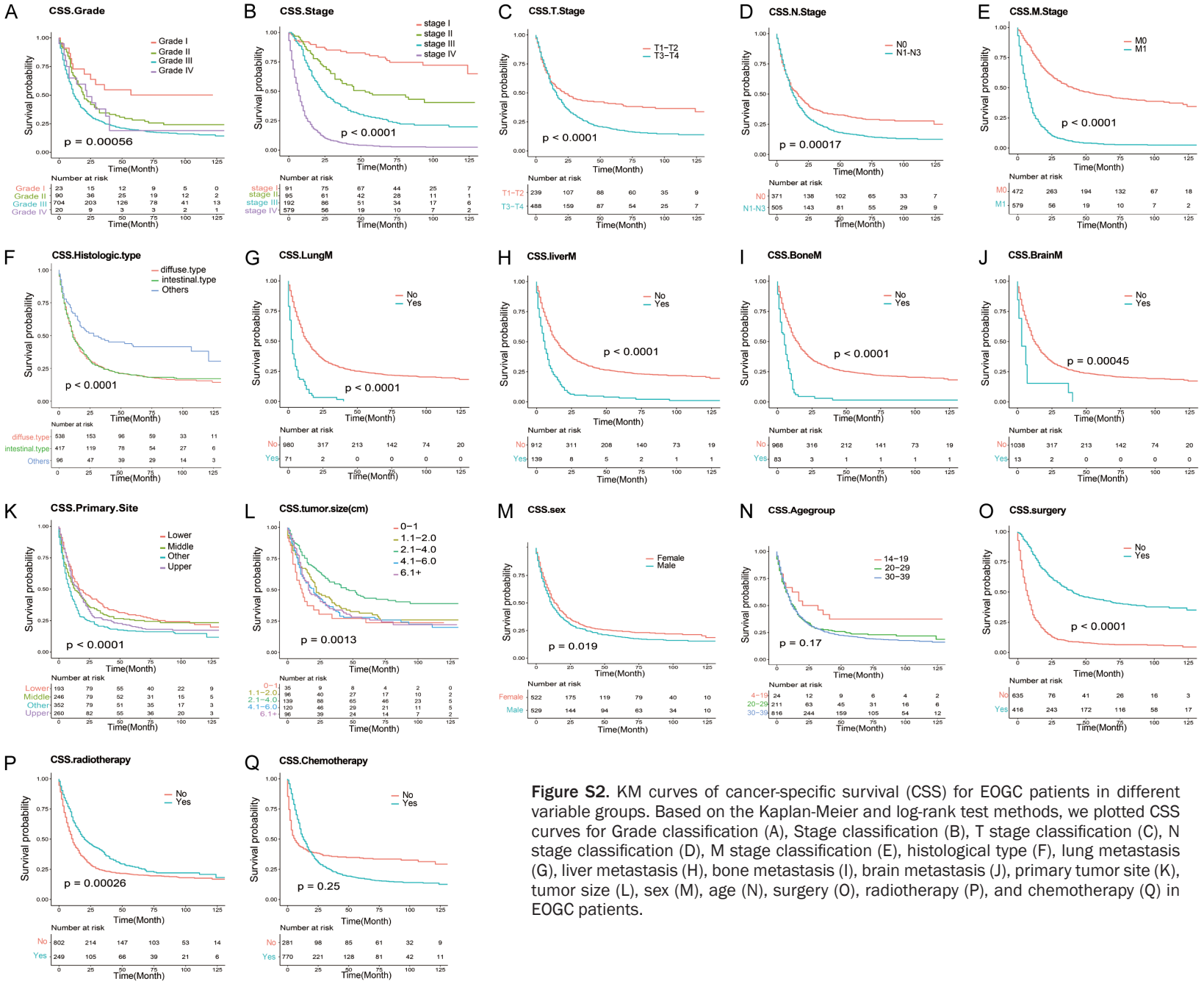

Figure S2. KM curves of cancer-specific survival (CSS) for EOGC patients in different variable groups. Based on the Kaplan-Meier and log-rank test methods, we plotted CSS curves for Grade classification (A), Stage classification (B), T stage classification (C), N stage classification (D), M stage classification (E), histological type (F), lung metastasis (G), liver metastasis (H), bone metastasis (I), brain metastasis (J), primary tumor site (K), tumor size (L), sex (M), age (N), surgery (O), radiotherapy (P), and chemotherapy (Q) in EOGC patients.

Table S1. The AUC of nomogram models and AJCC-TNM evaluation system

Variable	Train		Validation		*P*
	AUC	95% CI	AUC	95% CI	
1-year OS					
Nomogram	0.835	0.806-0.864	0.831	0.786-0.877	0.700#
AJCC-TNM	0.780	0.746-0.815	0.793	0.740-0.847	0.740#
P-Value		< 0.001*		0.072&	
1-year CSS					
Nomogram	0.835	0.805-0.864	0.842	0.797-0.887	0.965#
AJCC-TNM	0.777	0.741-0.813	0.816	0.768-0.865	0.608#
P-Value		< 0.001*		0.190&	
3-year OS					
Nomogram	0.881	0.851-0.911	0.892	0.851-0.934	0.650#
AJCC-TNM	0.836	0.800-0.871	0.857	0.808-0.907	0.743#
P-Value		0.0003*		0.039&	
3-year CSS					
Nomogram	0.890	0.861-0.920	0.887	0.846-0.928	0.570#
AJCC-TNM	0.837	0.801-0.873	0.848	0.802-0.895	0.999#
P-Value		< 0.001*		0.018&	
5-year OS					
Nomogram	0.900	0.872-0.929	0.906	0.869-0.944	0.617#
AJCC-TNM	0.853	0.817-0.888	0.880	0.833-0.927	0.674#
P-Value		0.0002*		0.133&	
5-year CSS					
Nomogram	0.904	0.875-0.933	0.900	0.864-0.937	0.496#
AJCC-TNM	0.848	0.810-0.885	0.876	0.835-0.918	0.835#
P-Value		< 0.001*		0.109&	

Note: #, Delong test for predicting the area under the curve (AUC) area for overall survival (OS) and cancer-specific survival (CSS) in patients in the training and validation sets; *, Delong test for predicting AUC area for OS and CSS of patients in the training set using Nomogram prediction model with American Joint Committee on Cancer (AJCC) tumor-node-metastasis (TNM) staging system; &, Delong test for predicting AUC area for OS and CSS of patients in the validation set using Nomogram prediction model with TNM staging system. CI, confidence interval.

Table S2. Comparison of AUCs of nomogram models between internal validation and external data

Variable	Validation		Test		*P*
	AUC	95% CI	AUC	95% CI	
1-year OS					
Nomogram	0.831	0.786-0.877	0.686	0.618-0.753	0.483#
1-year CSS					
Nomogram	0.842	0.797-0.887	0.697	0.628-0.765	0.469#
3-year OS					
Nomogram	0.892	0.851-0.934	0.661	0.610-0.713	0.896#
3-year CSS					
Nomogram	0.887	0.846-0.928	0.672	0.619-0.724	0.868#
5-year OS					
Nomogram	0.906	0.869-0.944	0.717	0.672-0.761	0.336#
5-year CSS					
Nomogram	0.900	0.864-0.937	0.732	0.687-0.777	0.381#

Note: #, Delong test for predicting AUC area for OS and CSS in patients in the validation and test sets.