Am J Cancer Res 2024;14(5):2272-2286
www.ajcr.us /ISSN:2156-6976/ajcr0157119

Original Article

Development and validation of nomogram models for predicting overall survival and cancer-specific survival in gastric cancer patients with liver metastases: a cohort study based on the SEER database

Ning Meng[1,2,3,4*], Xiaoman Niu[1,2,3*], 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]The 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]Department of General Surgery, Shijiazhuang People's Hospital, Shijiazhuang 050050, Hebei, China. [*]Equal contributors.

Received April 9, 2024; Accepted May 4, 2024; Epub May 15, 2024; Published May 30, 2024

Abstract: Objective: To establish nomogram models for predicting the overall survival (OS) and cancer-specific survival (CSS) of gastric cancer liver metastasis (GCLM) patients. Methods: Data from the Surveillance, Epidemiology, and End Results (SEER) database for 5,451 GCLM patients diagnosed between 2010 and 2015 were analyzed. The cohort was divided into a training set (3,815 cases) and an internal validation set (1,636 cases). External validation included 193 patients from the Fourth Hospital of Hebei Medical University and 171 patients from the People's Hospital of Shijiazhuang City, spanning 2016-2018. Multivariable Cox regression analysis identified eight independent prognostic factors for OS and CSS in GCLM patients, including age, histological type, grade, tumor size, surgery, chemotherapy, bone metastasis, and lung metastasis. Two nomogram models were developed based on these factors and evaluated using time-dependent receiver operating characteristic curve analysis, calibration curves, and decision curve analysis. Results: Internal validation showed that the nomogram models outperformed the American Joint Committee on Cancer (AJCC) tumor-node-metastasis (TNM) staging system in predicting 1-year, 2-year, and 3-year OS and CSS in GCLM patients (1-year OS: 0.801 vs. 0.593, $P < 0.001$; 1-year CSS: 0.807 vs. 0.598, $P < 0.001$; 2-year OS: 0.803 vs. 0.630, $P < 0.001$; 2-year CSS: 0.802 vs. 0.633, $P < 0.001$; 3-year OS: 0.824 vs. 0.691, $P < 0.001$; 3-year CSS: 0.839 vs. 0.692, $P < 0.001$). Conclusion: This study developed and validated nomogram models using SEER database data to predict OS and CSS in GCLM patients. These models offer improved prognostic accuracy over traditional staging systems, aiding in clinical decision-making.

Keywords: Gastric cancer with liver metastasis, nomogram, overall survival, cancer specific survival, SEER

Introduction

Gastric cancer, a prevalent malignant tumor of the digestive tract, poses a significant global health burden [1-3]. According to the GLOBOCAN 2020 report from the International Agency for Research on Cancer, it ranks fifth in incidence and second in mortality among all cancers globally [4]. The late detection of gastric cancer often results in unfavorable outcomes, with liver metastasis being a frequent complication [5-7]. Between 5-14% of patients develop liver metastasis, and this rate rises to 13.5-30% among those undergoing radical surgery, leading to high mortality [8-10].

Despite advancements in tumor therapy, the prognosis for patients with gastric cancer liver metastasis (GCLM) remains grim, emphasizing the need for effective, personalized treatment strategies and comprehensive prognostic assessments [11]. While systemic chemotherapy is a treatment option, primary tumor resection has shown benefits for the survival of GCLM patients. Additionally, liver resection and radio-

https://doi.org/10.62347/ZPPK5664

frequency ablation have also demonstrated certain advantages [12, 13].

Current clinical prognostic assessments predominantly adhere to the American Joint Committee on Cancer (AJCC) staging system, primarily focused on pathological features such as tumor invasion depth and lymph node metastases [14-16]. However, these guidelines overlook crucial individual factors like age, gender, tumor differentiation, and additional treatments, limiting their ability to provide personalized prognoses.

The nomogram model, leveraging big data and multivariate analysis, emerges as a promising clinical forecasting tool [17]. By integrating various independent prognostic factors, it quantifies individual survival risks, offering more nuanced assessments. Nevertheless, research on prognostic factors for GCLM remains in flux, often constrained by limited sample sizes. Dong et al. developed two nomogram models to predict overall survival (OS) and cancer-specific survival (CSS) in GCLM patients, yet their data was confined to the SEER database without external validation [18]. Existing models lack the efficiency and comprehensiveness to integrate multiple prognostic factors for accurate survival prognosis in GCLM patients [19].

A comprehensive and precise prognostic model is urgently needed. This model should integrate traditional pathological indicators with emerging biomarkers, taking into account individual patient variations such as genetic background, lifestyle, and socioeconomic status [20, 21]. Such an approach would foster a holistic understanding of factors influencing GCLM patient survival, thereby enhancing clinical treatment precision and improving patient outcomes.

This study aims to investigate the clinical characteristics and survival-related factors of GCLM patients. Utilizing the extensive Surveillance, Epidemiology, and End Results (SEER) database, renowned for its broad clinical coverage and rich data, we systematically analyzed various clinical features of GCLM patients. Our primary objective is to develop nomogram-based prognostic models for GCLM, thereby enhancing the accuracy of individualized prognosis assessments and informing clinical decision-making.

Methods and materials

Patient selection

The SEER database, maintained by the National Cancer Institute, is a publicly available repository that has been collecting data on cancer patients from various U.S. states and counties since 1973. It comprises comprehensive patient demographics (e.g., age, gender) and detailed tumor-related information (such as stage, pathology, and post-treatment survival). The comprehensive scope and vast sample size render the SEER database an invaluable resource for cancer research [22]. Given the public and anonymized nature of the data, ethics committee approval was not required for this study.

We utilized the SEER*Stat software version 8.4.2 to extract and analyze data, focusing on GCLM cases diagnosed between 2010 and 2015 via imaging or pathology. The analyzed data encompassed gender, age, race, diagnosis year, ICD-O-3 malignant behavior code, primary site, differentiation grade, AJCC 7th edition staging, tumor size, reasons for avoiding cancer-directed surgery, treatment modalities (chemotherapy, radiotherapy), marital status, and metastasis status in bone, lung, liver, and brain. Vital status, survival months, and SEER-specific cause of death were also included. Histological types were grouped into diffuse (including signet ring cell carcinoma, diffuse carcinoma, and leathery stomach) and intestinal types (gastric carcinoma, adenocarcinoma, tubular adenocarcinoma, and intestinal type).

Inclusion and exclusion criteria

The detailed inclusion and exclusion criteria for GCLM patients in this study are:

Inclusion criteria: 1. Initial tumor located in the stomach. 2. Diagnosis between 2010 and 2015. 3. Gastric cancer with liver metastases (SEER Combined Mets at DX-liver). 4. Complete clinical and pathological information, including primary site, histology, differentiation degree, primary tumor resection, and AJCC 7th edition staging. 5. ICD-O-3 malignant behavior classification: C16.0-C16.9. 6. Comprehensive clinical, pathological, follow-up, and prognostic data.

Exclusion criteria: 1. Multiple primary tumors. 2. Metastatic gastric cancer originating from non-gastric sites. 3. Absence of liver metastasis. 4. Cases solely confirmed by autopsy or death certificate. 5. Incomplete or significantly missing data.

In addition, this study incorporated an external test set comprising 193 cases from the Fourth Hospital of Hebei Medical University and 171 cases from Shijiazhuang People's Hospital, all diagnosed with GCLM between 2016 and 2018. The diagnoses in both the SEER database and the external cohort were confirmed through imaging or pathology. The inclusion and exclusion criteria for these additional cases were identical to those applied to the SEER dataset, except for the year of diagnosis. Follow-up data for these patients encompassed postoperative survival status, OS, CSS, cause of death, and treatment efficacy. This study represents a population cohort study leveraging data from the SEER database, the Fourth Hospital of Hebei Medical University, and Shijiazhuang People's Hospital.

Clinical and pathological parameters

For the purposes of analysis, the study comprehensively evaluated several key indicators, encompassing fundamental demographic data, tumor attributes such as location, number, histology, differentiation grade, and AJCC TNM staging, treatment modalities, and prognostic factors like OS, CSS, cause of mortality, and duration of survival.

Statistical analyses

Statistical analyses were conducted using R software version 4.3.2. Variable comparisons were made using the chi-square test and independent sample t-tests. For survival analyses, we employed Cox proportional hazards regression models for both univariate and multivariate assessments of OS and CSS. Utilizing the results from the multivariable Cox regression analysis, we developed nomogram models to predict 1-year, 2-year, and 3-year OS and CSS for GCLM patients. These nomogram models integrated independent prognostic factors identified during the analysis.

Our patient cohort from the SEER database was randomly split into a training set (70%) and a validation set (30%). The training set was used for constructing the nomogram models, while the validation set assessed the models' accuracy. Additionally, we collected GCLM patient data from the Fourth Hospital of Hebei Medical University and Shijiazhuang People's Hospital as an external test set to validate the constructed models externally.

We evaluated model performance using the area under the curve (AUC) and calibration curves. Calibration curves were employed to compare predicted and actual risks, thereby assessing predictive accuracy. Furthermore, decision curve analysis measured the model's clinical utility by determining the net benefit at various thresholds. We also compared the constructed nomogram models with the AJCC staging system. The Delong test was utilized to compare AUC values between the training set and internal validation set, as well as between the internal validation set and external test set. Statistical significance was set at a *P*-value less than 0.05.

Results

Comparison of clinical baseline characteristics of patients with GCLM

In this study, we applied the defined inclusion and exclusion criteria to identify GCLM patients registered in the SEER database from 2010 to 2015. The selection process is illustrated in **Figure 1**. Ultimately, 5,451 GCLM patients were selected from SEER and randomly assigned to a training set (3,815 cases) and an internal validation set (1,636 cases) in a 7:3 ratio. Additionally, GCLM patients admitted between 2016 and 2018 from the Fourth Hospital of Hebei Medical University (193 cases) and Shijiazhuang People's Hospital (171 cases) were compiled as external test sets, using the same criteria (**Figure 1**).

Demographic analysis revealed a male-to-female ratio of 2,667:1,148 (69.9% vs. 30.1%) in the training set, 1,166:470 (71.3% vs. 28.7%) in the validation set, and 246:118 (67.6% vs. 32.4%) in the test set. Statistical analysis indicated no significant gender disparities among the three groups (P=0.601). **Table 1** summarizes the detailed baseline clinical characteristics of all patients.

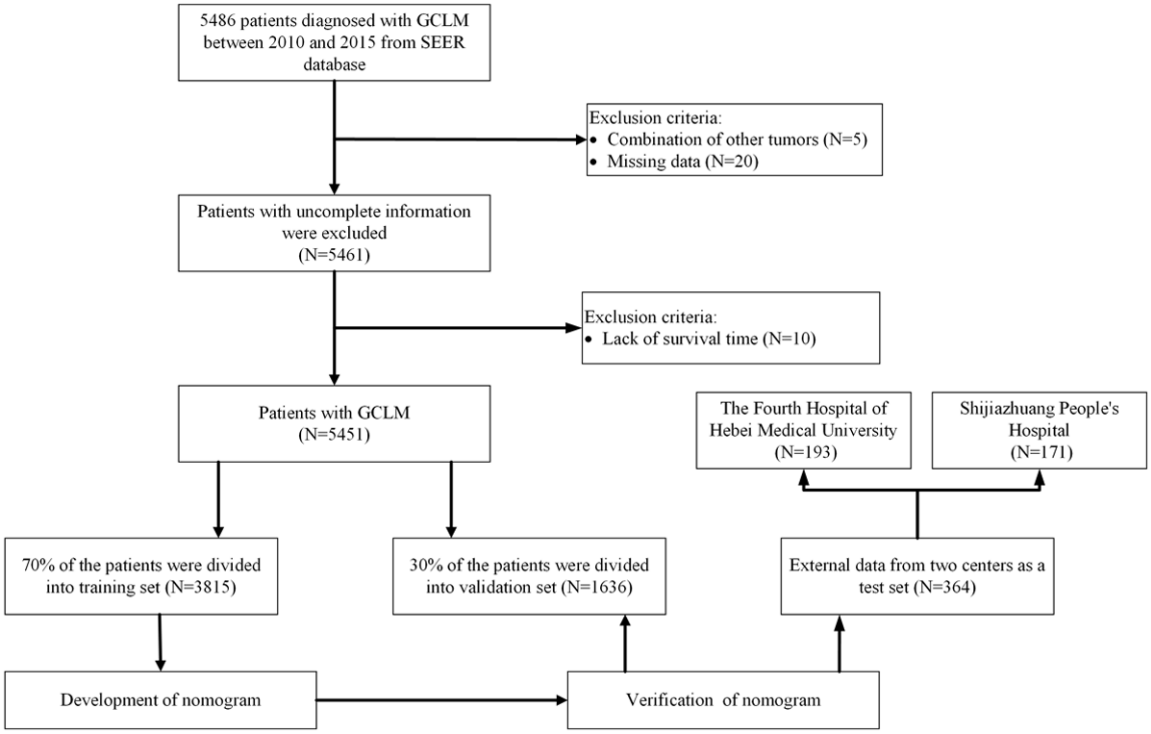

Figure 1. Flowchart of the screening procedure for patients with GCLM. GCLM, gastric cancer liver metastasis; SEER, the Surveillance, Epidemiology and End Results.

Analysis of multivariate effects on patients' OS and establishment of nomogram models

Utilizing multivariate Cox regression analysis, we identified several independent factors significantly affecting the OS of patients with GCLM. Specifically, age (HR=1.356, 95% CI: 1.133-1.624, P=0.001), type of pathology (HR=0.593, 95% CI: 0.523-0.673, P < 0.001), pathological grade (HR=1.393, 95% CI: 1.152-1.683, P=0.001), tumor size (HR=0.795, 95% CI: 0.650-0.973, P=0.026), surgical removal of the primary focus (HR=0.508, 95% CI: 0.453-0.568, P < 0.001), administration of chemotherapy (HR=0.324, 95% CI: 0.305-0.344, P < 0.001), presence of bone metastasis (HR=1.145, 95% CI: 1.043-1.256, P=0.004), and coexistence with lung metastasis (HR=1.335, 95% CI: 1.239-1.437, P < 0.001) were found to be significant predictors (**Table 2**).

Based on these statistically significant variables, we constructed nomogram models for predicting OS in GCLM patients. These nomograms integrate all the independent predictors and provide prognostic estimates for 1-year, 2-year, and 3-year OS (**Figure 2A-C**).

Analysis of multivariate effects on patients' CSS and establishment of nomogram models

Using multivariate Cox regression analysis, we identified several independent factors that significantly impact CSS in patients with GCLM. These factors include age (HR=1.362, 95% CI: 1.136-1.633, P=0.001), type of pathology (HR=0.585, 95% CI: 0.515-0.665, P < 0.001), pathological grade (HR=1.386, 95% CI: 1.146-1.677, P=0.001), tumor size (HR=0.794, 95% CI: 0.648-0.972, P=0.026), surgical removal of the primary focus (HR=0.509, 95% CI: 0.454-0.571, P < 0.001), administration of chemotherapy (HR=0.323, 95% CI: 0.304-0.343, P < 0.001), presence of bone metastasis (HR=1.139, 95% CI: 1.038-1.250, P=0.006), and coexistence with lung metastasis (HR=1.359, 95% CI: 1.262-1.465, P < 0.001) (**Table 3**).

The eight variables that demonstrated statistical significance in the multivariate Cox regression analysis were incorporated into a nomogram model for predicting CSS in GCLM patients. This nomogram integrates all the independent predictors and enables prognostic

Table 1. Baseline characteristics of patients with GCLM

Variable	Train (N=3815)	Validation (N=1636)	Test (N=364)	P
Age group (years old)				1.000
15-39	116 (3.0%)	49 (3.0%)	17 (4.7%)	
40-59	981 (25.7%)	421 (25.7%)	101 (27.7%)	
60-79	1991 (52.2%)	853 (52.1%)	196 (53.8%)	
80+	727 (19.1%)	313 (19.1%)	50 (13.7%)	
Sex				0.601
Female	1148 (30.1%)	470 (28.7%)	118 (32.4%)	
Male	2667 (69.9%)	1166 (71.3%)	246 (67.6%)	
Histologic type				0.876
Diffuse	323 (8.5%)	133 (8.1%)	40 (11.0%)	
Intestinal	2867 (75.2%)	1252 (76.5%)	222 (61.0%)	
Others	625 (16.4%)	102 (28.0%)	102 (28.0)	
Grade				0.680
I	86 (2.3%)	53 (3.2%)	14 (3.8%)	
II	930 (24.4%)	384 (23.5%)	88 (24.2%)	
III	1793 (47.0%)	752 (46.0%)	153 (42.0%)	
IV	77 (2.0%)	38 (2.3%)	38 (10.5%)	
Unknown	929 (24.3%)	409 (25.0%)	71 (19.5%)	
N stage				0.102
N0	1462 (38.3%)	564 (34.5%)	109 (29.9%)	
N1-N3	1660 (43.5%)	744 (45.5%)	255 (70.1%)	
Unknown	693 (18.2%)	328 (20.0%)	0 (0.0%)	
T stage				0.815
T1-T2	879 (23.0%)	357 (21.8%)	98 (26.9%)	
T3-T4	1127 (29.5%)	507 (31.0%)	266 (73.1%)	
Unknown	1809 (47.4%)	772 (47.2%)	0 (0.0%)	
Tumor size (cm)				0.949
0-1	98 (2.6%)	51 (3.1%)	12 (3.3%)	
1.1-2.0	230 (6.0%)	99 (6.1%)	42 (11.5%)	
2.1-4.0	454 (11.9%)	195 (11.9%)	95 (26.1%)	
4.1-6.0	480 (12.6%)	228 (13.9%)	128 (35.2%)	
6.1+	409 (10.7%)	161 (9.8%)	87 (23.9%)	
Unknown	2144 (56.2%)	902 (55.1%)	0 (0.0%)	
Surgery				0.138
None/Unknown	3507 (91.9%)	1477 (90.3%)	1477 (90.3%)	
Yes	308 (8.1%)	159 (9.7%)	159 (9.7%)	
Chemotherapy				1.000
No/Unknown	1737 (45.5%)	745 (45.5%)	62 (17.0%)	
Yes	2078 (54.5%)	891 (54.5%)	302 (83.0%)	
Radiotherapy				0.824
None/Unknown	3289 (86.2%)	1400 (85.6%)	342 (94.0%)	
Yes	526 (13.8%)	236 (14.4%)	22 (6.0%)	
Marital status				0.999
Married	2167 (56.8%)	922 (56.4%)	284 (78.0%)	
Other	1047 (27.4%)	454 (27.8%)	65 (17.9%)	
Singled	601 (15.8%)	260 (15.9%)	15 (4.1%)	

Primary Site				0.446
Lower	564 (14.8%)	204 (12.5%)	106 (29.1%)	
Middle	656 (17.2%)	298 (18.2%)	81 (22.2%)	
Upper	1681 (44.1%)	748 (45.7%)	158 (43.4%)	
Other	914 (24.0%)	386 (23.6%)	19 (5.2%)	
Bone metastasis				0.829
No/Unknown	3433 (90.0%)	1481 (90.5%)	352 (96.7%)	
Yes	382 (10.0%)	155 (9.5%)	12 (3.3%)	
Lung metastasis				0.574
No/Unknown	3181 (83.4%)	1345 (82.2%)	359 (98.6%)	
Yes	634 (16.6%)	291 (17.8%)	5 (1.4%)	
Brain metastasis				0.947
No/Unknown	3752 (98.3%)	1611 (98.5%)	363 (98.5%)	
Yes	63 (1.7%)	25 (1.5%)	1 (0.3%)	

GCLM, gastric cancer liver metastasis.

Table 2. Cox analysis of factors for OS in patients with GCLM

Variable	OS-Univariate				OS-Multivariate			
	HR	95% CI lower	95% CI upper	*P* value	HR	95% CI lower	95% CI upper	*P* value
Age group (years old)								
15-39	1				1.000			
40-59	1.181	0.995	1.401	0.057	1.048	0.882	1.245	0.595
60-79	1.429	1.209	1.688	< 0.001	1.181	0.997	1.400	0.054
80+	2.153	1.809	2.561	< 0.001	1.356	1.133	1.624	0.001
Sex								
Female	1							
Male	1.060	0.998	1.125	0.058				
Race								
American Indian	1							
Asian or Pacific Islander	0.992	0.757	1.300	0.955				
Black	0.955	0.731	1.246	0.733				
Unknown	0.522	0.259	1.054	0.070				
White	1.028	0.793	1.332	0.834				
Histologic type								
Diffuse	1				1.000			
Intestinal	0.877	0.794	0.968	0.009	0.934	0.843	1.034	0.188
Others	0.503	0.446	0.566	< 0.001	0.593	0.523	0.673	< 0.001
Grade								
I	1				1.000			
II	1.511	1.255	1.820	< 0.001	1.393	1.152	1.683	0.001
III	1.949	1.624	2.338	< 0.001	1.879	1.560	2.264	< 0.001
IV	1.375	1.06	1.784	0.016	2.205	1.695	2.869	< 0.001
Unknown	1.462	1.214	1.761	< 0.001	1.450	1.202	1.750	< 0.001
N Stage								
N0	1				1.000			
N1-N3	1.090	1.026	1.159	0.006	1.063	0.997	1.134	0.061
Unknown	1.585	1.468	1.713	< 0.001	1.246	1.149	1.351	< 0.001

T Stage								
T1-2	1				1.000			
T3-T4	0.852	0.79	0.92	< 0.001	1.011	0.934	1.095	0.785
Unknown	1.125	1.050	1.205	0.001	0.978	0.909	1.051	0.544
Tumor size (cm)								
0-1	1				1.000			
1.1-2.0	0.686	0.562	0.838	< 0.001	0.795	0.650	0.973	0.026
2.1-4.0	0.762	0.635	0.915	0.004	0.724	0.602	0.869	0.001
4.1-6.0	0.868	0.724	1.041	0.126	0.841	0.701	1.009	0.063
6.1+	0.792	0.658	0.954	0.014	0.767	0.636	0.924	0.005
Unknown	1.075	0.908	1.273	0.403	0.879	0.741	1.042	0.138
Surgery								
None/Unknown	1				1.000			
Yes	0.493	0.444	0.546	< 0.001	0.508	0.453	0.568	< 0.001
Chemotherapy								
None/Unknown	1				1.000			
Yes	0.351	0.332	0.371	< 0.001	0.324	0.305	0.344	< 0.001
Radiotherapy								
None/Unknown	1							
Yes	0.990	0.916	1.070	0.807				
Marital status								
Married	1				1.000			
Other	1.137	1.068	1.211	< 0.001	0.950	0.891	1.014	0.121
Singled	1.077	0.996	1.164	0.062	1.066	0.984	1.155	0.115
Primary Site								
Lower	1				1.000			
Middle	0.953	0.864	1.051	0.330	0.956	0.866	1.056	0.377
Other	1.055	0.963	1.157	0.250	1.057	0.964	1.160	0.239
Upper	0.974	0.896	1.059	0.538	0.981	0.902	1.067	0.656
Bone metastasis								
No	1				1.000			
Yes	1.318	1.204	1.442	< 0.001	1.145	1.043	1.256	0.004
Lung metastasis								
No	1				1.000			
Yes	1.446	1.346	1.554	< 0.001	1.335	1.239	1.437	< 0.001
Brain metastasis								
No	1							
Yes	1.144	0.923	1.418	0.218				

GCLM, gastric cancer liver metastasis; OS, overall survival.

estimates for 1-year, 2-year, and 3-year CSS survival (**Figure 3A-C**).

Internal validation of the prediction models of the nomogram

Based on the randomized 7:3 split of GCLM patients from the SEER database into training and internal validation sets, the developed nomogram models underwent rigorous internal validation. The results demonstrated that the nomogram models for predicting OS and CSS in GCLM patients exhibited high predictive accuracy (**Figures 2D-I**, **3D-I**). Specifically, the training set showed significantly higher AUC values for 1-, 2-, and 3-year OS and CSS compared to the traditional TNM staging system (1-year OS: 0.788 vs. 0.604, P < 0.001; 1-year CSS: 0.785 vs. 0.592, P < 0.001; 2-year OS: 0.795 vs. 0.644, P < 0.001; 2-year CSS: 0.792 vs. 0.631,

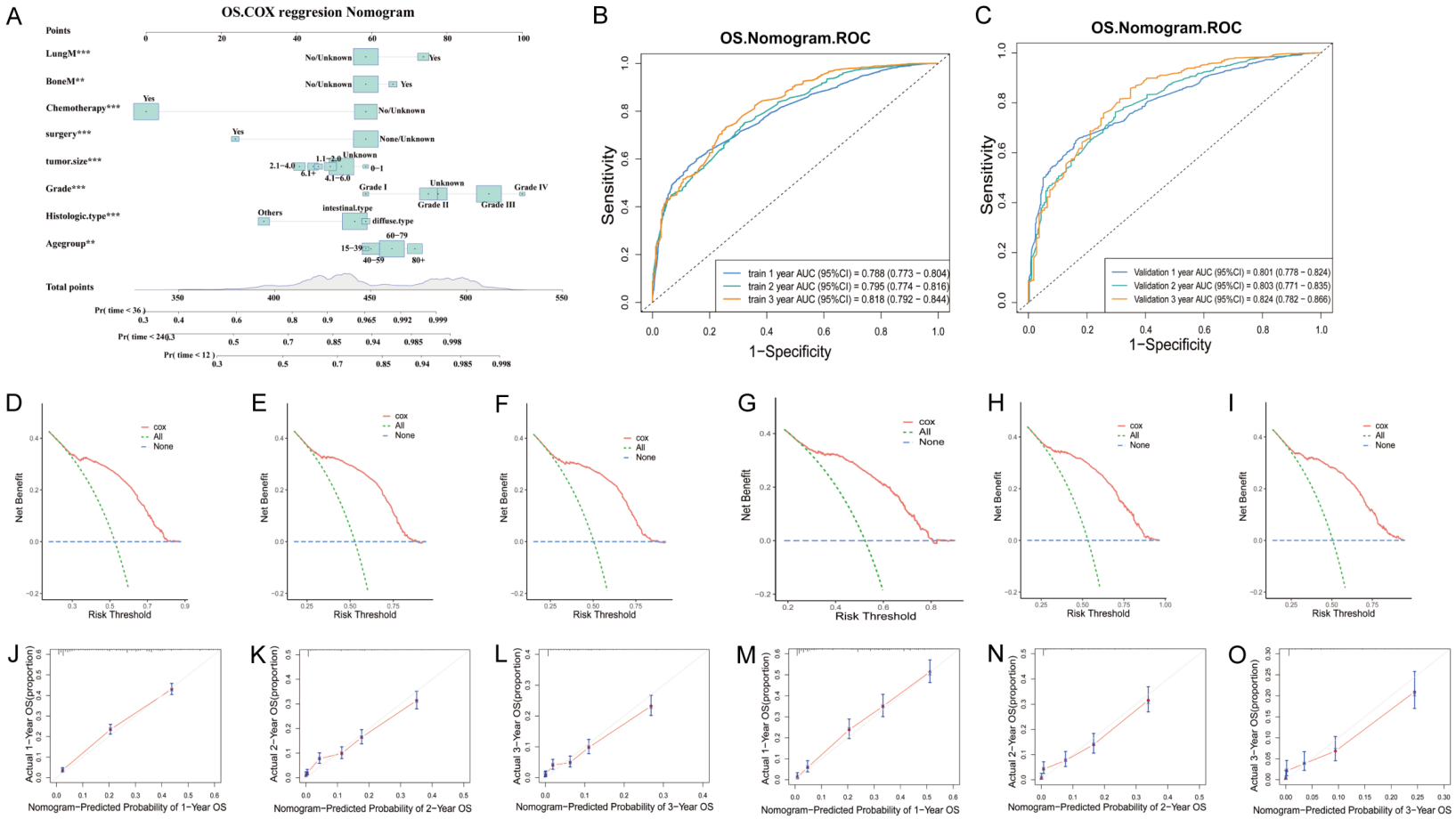

Figure 2. Construction and validation of a diagnostic nomogram. A prognostic nomogram for predicting the OS of GCLM patients for the 1 year, 2 years, and 3 years (A); the receiver operating characteristic curve of the nomogram for the 1 year, 2 years, and 3 years in the training set (B); calibration cure of the nomogram predicting 1 year, 2 years and 3 years OS (C) in the validation set; Training DCA curves for 1-year (D), 2-year (E), and 3-year (F) survival for the nomogram model of OS in patients with centralized GCLM. DCA curves for 1-year (G), 2-year (H), and 3-year (I) survival for the nomogram model of OS in GCLM patients in the validation set. Calibration curves for 1-year (J), 2-year (K), and 3-year (L) survival for the nomogram model of OS in patients with GCLM in the training set. Calibration curves for 1-year (M), 2-year (N), and 3-year (O) survival for the nomogram model of OS in GCLM patients in the validation set. GCLM, gastric cancer liver metastasis; OS, overall survival; CSS, cancer-specific survival; ROC, receiver operating characteristics; DCA, decision curve analysis; AUC, the area under the curve.

Table 3. Cox analysis of factors for CSS in patients with GCLM

Variable	CSS-Univariate				CSS-Multivariate			
	HR	95% CI lower	95% CI upper	*P* value	HR	95% CI lower	95% CI upper	*P* value
Age group (years old)								
15-39	1.000				1.000			
40-59	1.190	1.002	1.414	0.048	1.052	0.884	1.251	0.568
60-79	1.444	1.221	1.708	< 0.001	1.193	1.005	1.414	0.043
80+	2.172	1.823	2.587	< 0.001	1.362	1.136	1.633	0.001
Sex								
Female	1.000							
Male	1.055	0.993	1.120	0.082				
Race								
American Indian	1.000							
Asian or Pacific Islander	1.002	0.765	1.314	0.987				
Black	0.954	0.730	1.245	0.727				
Unknown	0.436	0.199	0.954	0.038				
White	1.026	0.792	1.330	0.843				
Histologic type								
Diffuse type	1.000				1.000			
Intestinal type	0.871	0.788	0.963	0.007	0.930	0.839	1.031	0.167
Others	0.500	0.443	0.564	< 0.001	0.585	0.515	0.665	< 0.001
Grade								
I	1.000				1.000			
II	1.497	1.242	1.805	< 0.001	1.386	1.146	1.677	0.001
III	1.936	1.612	2.325	< 0.001	1.871	1.552	2.255	< 0.001
IV	1.365	1.052	1.772	0.019	2.202	1.692	2.866	< 0.001
Unknown	1.446	1.200	1.743	< 0.001	1.452	1.202	1.753	< 0.001
N Stage								
N0	1.000				1.000			
N1-N3	1.087	1.022	1.156	0.008	1.058	0.992	1.129	0.086
Unknown	1.575	1.457	1.702	< 0.001	1.238	1.141	1.344	< 0.001
T Stage								
T1-2	1.000				1.000			
T3-T4	0.855	0.792	0.922	< 0.001	1.013	0.935	1.098	0.749
Unknown	1.129	1.053	1.210	0.001	0.980	0.911	1.054	0.589
Tumor size (cm)								
0-1	1.000				1.000			
1.1-2.0	0.698	0.570	0.854	< 0.001	0.794	0.648	0.972	0.026
2.1-4.0	0.771	0.641	0.928	0.006	0.719	0.597	0.865	< 0.001
4.1-6.0	0.869	0.723	1.044	0.133	0.825	0.686	0.992	0.041
6.1+	0.798	0.662	0.963	0.019	0.755	0.626	0.912	0.004
Unknown	1.080	0.910	1.281	0.380	0.868	0.73	1.032	0.108
Surgery								
None/Unknown	1.000				1.000			
Yes	0.492	0.443	0.546	< 0.001	0.509	0.454	0.571	< 0.001
Chemotherapy								
None/Unknown	1.000				1.000			
Yes	0.350	0.330	0.370	< 0.001	0.323	0.304	0.343	< 0.001

Radiotherapy								
None/Unknown	1.000							
Yes	0.990	0.915	1.070	0.791				
Marital status								
Married	1.000				1.000			
Other	1.147	1.077	1.222	< 0.001	0.962	0.901	1.026	0.239
Singled	1.079	0.997	1.167	0.059	1.074	0.990	1.164	0.085
Primary Site								
Lower	1.000							
Middle	0.956	0.863	1.053	0.372				
Other	1.057	0.964	1.160	0.234				
Upper	0.981	0.902	1.067	0.656				
Bone metastasis								
No	1.000				1.000			
Yes	1.317	1.204	1.442	< 0.001	1.139	1.038	1.250	0.006
Lung metastasis								
No	1.000				1.000			
Yes	1.462	1.360	1.572	< 0.001	1.359	1.262	1.465	< 0.001
Brain metastasis								
No	1.000							
Yes	1.144	0.923	1.418	0.218				

GCLM, gastric cancer liver metastasis; CSS, cancer-specific survival.

P < 0.001; 3-year OS: 0.818 vs. 0.694, P < 0.001; 3-year CSS: 0.809 vs. 0.679, P < 0.001), as detailed in Table S1). Similarly, the internal validation sets also displayed superior AUC values (1-year OS: 0.801 vs. 0.593, P < 0.001; 1-year CSS: 0.807 vs. 0.598, P < 0.001; 2-year OS: 0.803 vs. 0.630, P < 0.001; 2-year CSS: 0.802 vs. 0.633, P < 0.001; 3-year OS: 0.824 vs. 0.691, P < 0.001; 3-year CSS: 0.839 vs. 0.692, P < 0.001).

Delong's analysis of the receiver operating characteristic (ROC) curves further confirmed that both OS and CSS nomogram models exhibited larger AUC values, indicating superior accuracy compared to the TNM staging system (Table S1). To mitigate potential overfitting of the AUC, calibration curves were constructed using the 1-, 2-, and 3-year OS and CSS data of GCLM patients. These curves revealed a high degree of agreement between the predicted and observed survival probabilities, indicating excellent discriminatory power and accuracy of the developed nomogram models (**Figures 2J-O**, **3J-O**).

External validation of nomogram models

An external test set comprising GCLM patients from the Fourth Hospital of Hebei Medical University and the People's Hospital of Shijiazhuang was established to validate the constructed nomogram models. The results confirm that both nomogram models, intended for predicting the OS and CSS of GCLM patients, exhibit considerable external validity (**Figure 4A**, **4H**). Analysis of the AUC values of the ROC curves for OS and CSS at 3.5 years (OS: 0.816 vs. 0.648, P=0.268; CSS: 0.833 vs. 0.637, P=0.155), 4 years (OS: 0.820 vs. 0.657, P= 0.110; CSS: 0.834 vs. 0.659, P=0.141), and 4.5 years (OS: 0.850 vs. 0.672, P=0.100; CSS: 0.855 vs. 0.672, P=0.056) on the external test set revealed no significant statistical differences between the external test set and the internal validation set This finding suggests that the two nomogram models possess a certain level of generalization capability (Table S2).

To mitigate potential overfitting, the AUC was recalibrated using the OS and CSS data of GCLM patients at 1, 2, and 3 years. The calibration curves revealed a high degree of agreement between the predicted and observed survival probabilities, indicating that the nomogram models accurately predict both OS and CSS (**Figure 4B-G**, **4I-N**). This demonstrates the excellent discriminatory ability and accuracy of the constructed models.

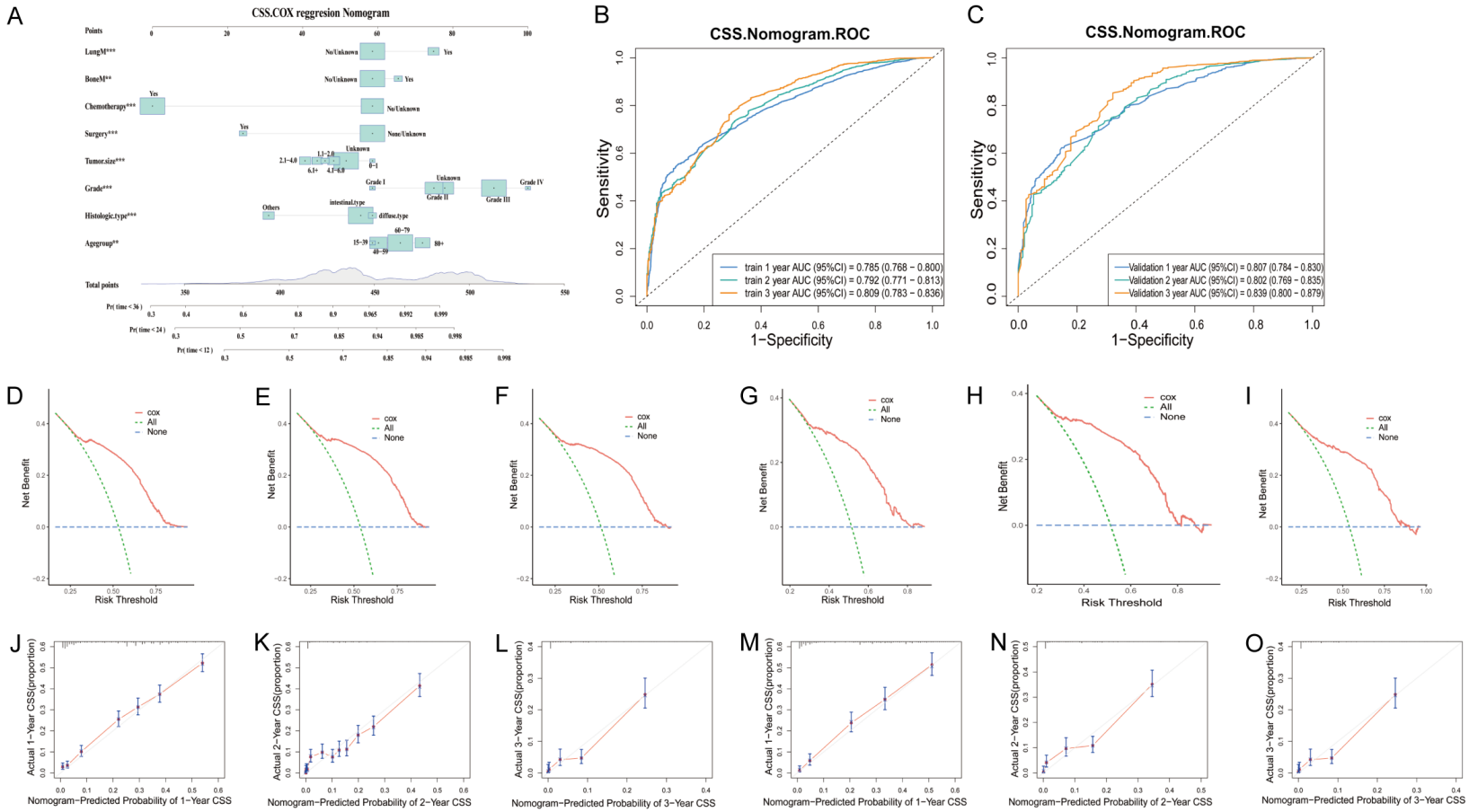

Figure 3. Construction and validation of a diagnostic nomogram. A prognostic nomogram for predicting the CSS of GCLM patients for the 1 year, 2 years, and 3 years (A); the receiver operating characteristic curve of the nomogram for the 1 year, 2 years, and 3 years in the training set (B); calibration cure of the nomogram predicting 1 year, 2 years and 3 years overall survival (C) in the validation set; Training DCA curves for 1-year (D), 2-year (E), and 3-year (F) survival for the nomogram model of CSS in patients with centralized GCLM. DCA curves for 1-year (G), 2-year (H), and 3-year (I) survival for the nomogram model of CSS in GCLM patients in the validation set. Calibration curves for 1-year (J), 2-year (K), and 3-year (L) survival for the nomogram model of CSS in patients with GCLM in the training set. Calibration curves for 1-year (M), 2-year (N), and 3-year (O) survival for the nomogram model of CSS in GCLM patients in the validation set. GCLM, gastric cancer liver metastasis; OS, overall survival; CSS, cancer-specific survival; ROC, receiver operating characteristics; DCA, decision curve analysis; AUC, the area under the curve.

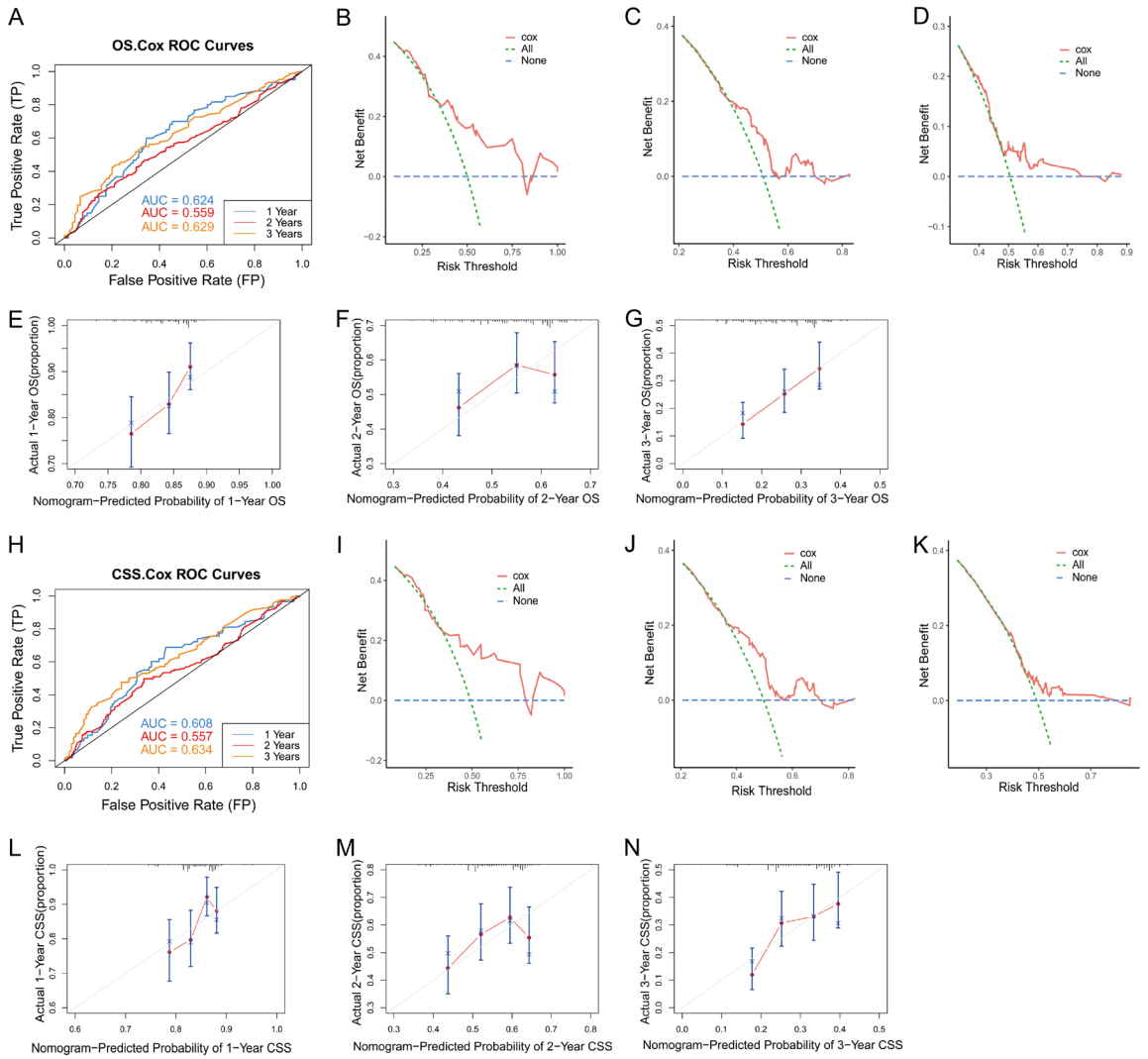

Figure 4. External testing of diagnostic nomograms. Receiver operating characteristic curves of the nomogram predicting 1-, 2-, and 3-year OS in the test set (A); DCA curves of the OS histogram predicting 1-year (B), 2-year (C), and 3-year (D) survival for the model in the test set GCLM patients. Calibration curves for 1-year (E), 2-year (F), and 3-year (G) survival for the OS histogram prediction model for GCLM patients in the test set. Receiver operating characteristic curves (H) for the nomogram prediction of 1-, 2-, and 3-year CSS in the test set; DCA curves for 1-year (I), 2-year (J), and 3-year (K) survival for the CSS histogram prediction model in the GCLM patients in the test set. Calibration curves for 1-year (L), 2-year (M), and 3-year (N) survival for the CSS histogram prediction model for GCLM patients in the test set. GCLM, gastric cancer liver metastasis; OS, overall survival; CSS, cancer-specific survival; ROC, receiver operating characteristics; DCA, decision curve analysis; AUC, the area under the curve.

Discussion

GCLM is often associated with a poor prognosis. However, advancements in medical technology, the integration of multiple disciplines in treatment, and the development of new adjuvant drugs have significantly improved the survival period of GCLM patients [23, 24]. Previous studies have linked survival predictions to clinical characteristics, pathological features, and tumor biology [22, 23]. Therefore, considering these indicators is crucial for optimizing treatment and enhancing prognostic accuracy. This study leverages extensive clinical data to develop a prognostic model for GCLM, offering personalized treatment recommendations.

In statistical prediction, the nomogram model, an innovative approach, demonstrates superior accuracy and flexibility compared to traditional methods. This model, now prevalent in various cancer types, was employed in our study to pre-

dict OS and CSS in GCLM patients [24, 25]. We evaluated its effectiveness using a validation set and assessed model performance with consistency indices, ROC curves, and calibration curves. The models exhibited good predictive capabilities for OS and CSS and outperformed the AJCC-TNM staging system.

Our analysis identified eight key independent risk factors affecting GCLM patient survival prognosis: age, pathological classification, grade, tumor size, primary site surgery, chemotherapy, and bone and lung metastases. The nomogram models incorporated these factors. Notably, age is an important predictive factor in the OS model, indicating that older patients receive less treatment or no treatment at all, leading to factors such as a weakened immune system, which negatively impact prognosis [26, 27].

Treatment options for GCLM currently encompass surgical resection, systemic chemotherapy, ablation therapy, and radiation therapy. Although there's ongoing debate regarding their applicability, systemic chemotherapy is widely acknowledged as the primary adjuvant treatment [28-30]. Chang et al. categorized patients into palliative resection and non-palliative resection groups [31]. Their study revealed that surgical resection combined with chemotherapy might offer survival benefits, particularly when metastasis is confined to a single site [31]. Recent clinical research underscores the significance of palliative care alongside radical primary site resection, although evidence is still evolving.

The primary lesion can create a tumor microenvironment that enhances invasion and growth. For example, primary tumor stem cells can upregulate VEGF expression, promoting neovascularization and tumor progression [32, 33]. Additionally, inflammatory factors within primary lesions, like IFN-Y and IL-10, may indirectly modulate tumor progression, suggesting that primary tumor resection could modify these factors and improve survival prognosis [34].

The two nomogram models developed in this study, utilizing comprehensive clinical and pathological data from the SEER database, represent valuable clinical tools for assessing the OS and CSS of GCLM patients. These models contribute significantly to treatment decision-making. However, limitations of the SEER database, such as missing critical information and lack of detail, may introduce bias and restrict research in certain areas. Li et al.'s analysis of gene activity in a mouse model of gastric cancer revealed that MAPK4 promotes MIF degradation in gastric cancer cells. They also noted a significant correlation between MAPK4 downregulation in gastric cancer patients and liver metastasis, along with a poor prognosis [19]. Yu et al. demonstrated notable differences in gut microbiota composition between gastric cancer patients and healthy individuals, identifying Streptococcus as a potential biomarker for early gastric cancer and GCLM prediction [35].

Our study has limitations, including the absence of specific molecular markers for liver metastasis, lack of detailed descriptions of surgical and chemotherapy protocols, genetic information related to gastric cancer with liver metastasis, and relevant information on gut microbiota. Therefore, further clinical research is necessary to validate and bolster these findings.

In conclusion, this study utilized the SEER database to develop nomogram prediction models for 1-year, 2-year, and 3-year OS and CSS in GCLM patients. These models serve as valuable clinical tools for evaluating GCLM patient prognosis, enhancing the accuracy of individualized prognosis assessment, and providing a foundation for clinical decision-making by healthcare professionals.

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).

Written informed consent was performed for all the study subjects.

Disclosure of conflict of interest

None.

Abbreviations

OS, overall survival; CSS, cancer-specific survival; GCLM, gastric cancer with liver metasta-

sis; SEER, the Surveillance, Epidemiology and End Results; ROC, receiver operating characteristics; DCA, decision curve analysis; AJCC, the American Joint Committee on Cancer; AUC, the area under the curve; KM, Kaplan-Meier; HR, Hazard ratio; TNM, tumor-node-metastasis; CI, confidence interval.

Address correspondence to: Qun Zhao, The 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] Joshi SS and Badgwell BD. Current treatment and recent progress in gastric cancer. CA Cancer J Clin 2021; 71: 264-279.

[2] Norwood DA, Montalvan-Sanchez E, Dominguez RL and Morgan DR. Gastric cancer: emerging trends in prevention, diagnosis, and treatment. Gastroenterol Clin North Am 2022; 51: 501-518.

[3] Huang RJ, Laszkowska M, In H, Hwang JH and Epplein M. Controlling gastric cancer in a world of heterogeneous risk. Gastroenterology 2023; 164: 736-751.

[4] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021; 71: 209-249.

[5] Aurello P, Minervini A, Pace M, D'Angelo F, Nigri G, Antolino L, Valabrega S, Ramacciato G and Petrucciani N. The role of surgery in the treatment of metachronous liver metastasis from gastric cancer: a systematic review. Anticancer Res 2022; 42: 25-33.

[6] Jin X, Zhu Z and Shi Y. Metastasis mechanism and gene/protein expression in gastric cancer with distant organs metastasis. Bull Cancer 2014; 101: E1-12.

[7] Kumagai K, Tanaka T, Yamagata K, Yokoyama N and Shimizu K. Liver metastasis in gastric cancer with particular reference to lymphatic advancement. Gastric Cancer 2001; 4: 150-155.

[8] Moon YW, Rha SY, Zhang X, Jeung HC, Yang WI, Kwon O, Jeong JH, Cheon SH, Yoo NC and Chung HC. Increments of alpha-dystroglycan expression in liver metastasis correlate with poor survival in gastric cancer. J Surg Oncol 2009; 100: 459-465.

[9] Kassar S, Samaha R, Aoun R, Khoury M and Kattan J. The concept of oligometastatic disease in gastric cancer: reality or fiction? Future Oncol 2022; 18: 135-138.

[10] Su L, Liu X, Chai N, Lv L, Wang R, Li X, Nie Y, Shi Y and Fan D. The transcription factor FOXO4 is down-regulated and inhibits tumor proliferation and metastasis in gastric cancer. BMC Cancer 2014; 14: 378.

[11] Song JC, Ding XL, Zhang Y, Zhang X and Sun XH. Prospective and prognostic factors for hepatic metastasis of gastric carcinoma: a retrospective analysis. J Cancer Res Ther 2019; 15: 298-304.

[12] Guner A, Son T, Cho I, Kwon IG, An JY, Kim HI, Cheong JH, Noh SH and Hyung WJ. Liver-directed treatments for liver metastasis from gastric adenocarcinoma: comparison between liver resection and radiofrequency ablation. Gastric Cancer 2016; 19: 951-960.

[13] Wu J, Yu J, Chen Z, Zhu H, Zhong C, Liang Y, Mai Z, Lin Z, Wan Y and Li G. Survival benefit of primary tumor resection for gastric cancer with liver metastasis: a propensity score-matched, population-based study. Front Oncol 2022; 12: 1039086.

[14] Zhang Y and Yu C. Development and validation of a Surveillance, Epidemiology, and End Results (SEER)-based prognostic nomogram for predicting survival in elderly patients with gastric cancer after surgery. J Gastrointest Oncol 2021; 12: 278-296.

[15] Ding P, Yang P, Tian Y, Guo H, Liu Y, Zhang Z, Zheng T, Tan B, Zhang Z, Wang D, Li Y and Zhao Q. Neoadjuvant intraperitoneal and systemic paclitaxel combined with apatinib and S-1 chemotherapy for conversion therapy in gastric cancer patients with positive exfoliative cytology: a prospective study. J Gastrointest Oncol 2021; 12: 1416-1427.

[16] Liu P, Ding P, Wu H, Wu J, Yang P, Tian Y, Guo H and Zhao Q. Prediction of occult peritoneal metastases or positive cytology using CT in gastric cancer. Eur Radiol 2023; 33: 9275-9285.

[17] Wu H, Ding P, Wu J, Sun C, Guo H, Chen S, Lowe S, Yang P, Tian Y, Liu Y and Zhao Q. A new online dynamic nomogram: construction and validation of a predictive model for distant metastasis risk and prognosis in patients with gastrointestinal stromal tumors. J Gastrointest Surg 2023; 27: 1429-1444.

[18] 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.

[19] Li S, Guo D, Sun Q, Zhang L, Cui Y, Liu M, Ma X, Liu Y, Cui W, Sun L, Teng L, Wang L, Lin A, Liu W, Zhuo W and Zhou T. MAPK4 silencing in gastric cancer drives liver metastasis by positive feedback between cancer cells and macrophages. Exp Mol Med 2023; 55: 457-469.

[20] Tuan TF, Chung CT, Tsou HH, Chen FW, Lin HL, Lai YK, Lee WS, Chao YS, Hwang LL and Chen CT. Putative tumor metastasis-associated genes in human gastric cancer. Int J Oncol 2012; 41: 1068-1084.
[21] Polkowska-Pruszyńska B, Rawicz-Pruszyński K, Ciseł B, Sitarz R, Polkowska G, Krupski W and Polkowski WP. Liver metastases from gastric carcinoma: a case report and review of the literature. Curr Probl Cancer 2017; 41: 222-230.
[22] 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.
[23] Sakamoto Y, Sano T, Shimada K, Esaki M, Saka M, Fukagawa T, Katai H, Kosuge T and Sasako M. Favorable indications for hepatectomy in patients with liver metastasis from gastric cancer. J Surg Oncol 2007; 95: 534-539.
[24] Liu SF, Lu CR, Cheng HD, Xi HQ, Cui JX, Li JY, Shen WS and Chen L. Comparison of therapeutic efficacy between gastrectomy with transarterial chemoembolization plus systemic chemotherapy and systemic chemotherapy alone in gastric cancer with synchronous liver metastasis. Chin Med J (Engl) 2015; 128: 2194-2201.
[25] Xu X, Li J, Han X, Shi C, Jin D and Lou W. Clinical characteristics and prognostic factors of patients with gastric neuroendocrine carcinoma treated with radical surgery. Chin Med J (Engl) 2014; 127: 2419-2422.
[26] Wang YY, Xiang BD, Ma L, Zhong JH, Ye JZ, Wang K, Xing BC and Li LQ. Development and validation of a nomogram to preoperatively estimate post-hepatectomy liver dysfunction risk and long-term survival in patients with hepatocellular carcinoma. Ann Surg 2021; 274: e1209-e1217.
[27] Benoit L, Balaya V, Guani B, Bresset A, Magaud L, Bonsang-Kitzis H, Ngô C, Mathevet P and Lécuru F. Nomogram predicting the likelihood of parametrial involvement in early-stage cervical cancer: avoiding unjustified radical hysterectomies. J Clin Med 2020; 9: 2121.
[28] Zhang W, Yu Y, Fang Y, Wang Y, Cui Y, Shen K and Liu T. Systemic chemotherapy as a main strategy for liver metastases from gastric cancer. Clin Transl Oncol 2015; 17: 888-894.
[29] Chen L, Song MQ, Lin HZ, Hao LH, Jiang XJ, Li ZY and Chen YX. Chemotherapy and resection for gastric cancer with synchronous liver metastases. World J Gastroenterol 2013; 19: 2097-2103.
[30] Namikawa T, Tsuda S, Fujisawa K, Iwabu J, Uemura S, Tsujii S, Maeda H, Kitagawa H, Kobayashi M and Hanazaki K. Conversion surgery after S-1 plus oxaliplatin combination chemotherapy for advanced gastric cancer with multiple liver metastases. Clin J Gastroenterol 2018; 11: 297-301.
[31] Chang YR, Han DS, Kong SH, Lee HJ, Kim SH, Kim WH and Yang HK. The value of palliative gastrectomy in gastric cancer with distant metastasis. Ann Surg Oncol 2012; 19: 1231-1239.
[32] van der Wal GE, Gouw AS, Kamps JA, Moorlag HE, Bulthuis ML, Molema G and de Jong KP. Angiogenesis in synchronous and metachronous colorectal liver metastases: the liver as a permissive soil. Ann Surg 2012; 255: 86-94.
[33] Liu W, Dong Z, Hu R and Wang C. Association of vascular endothelial growth factor (VEGF) gene polymorphisms with gastric cancer and its development, prognosis, and survival. Technol Cancer Res Treat 2018; 17: 1533034617753810.
[34] Evans C, Morrison I, Heriot AG, Bartlett JB, Finlayson C, Dalgleish AG and Kumar D. The correlation between colorectal cancer rates of proliferation and apoptosis and systemic cytokine levels; plus their influence upon survival. Br J Cancer 2006; 94: 1412-1419.
[35] Yu D, Yang J, Jin M, Zhou B, Shi L, Zhao L, Zhang J, Lin Z, Ren J, Liu L, Zhang T and Liu H. Fecal streptococcus alteration is associated with gastric cancer occurrence and liver metastasis. mBio 2021; 12: e0299421.

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

Variable	Train		Validation		*P*
	AUC	95% CI	AUC	95% CI	
1-year OS					
Nomogram	0.788	0.773 -0.804	0.801	0.778 - 0.824	0.960#
AJCC-TNM	0.604	0.583-0.625	0.593	0.560 - 0.626	0.630#
P	< 0.0001*		< 0.0001&		
1-year CSS					
Nomogram	0.785	0.768 -0.800	0.807	0.784 - 0.830	0.020#
AJCC-TNM	0.592	0.571-0.614	0.598	0.566 - 0.631	0.549#
P	< 0.0001*		< 0.0001&		
2-year OS					
Nomogram	0.795	0.774 - 0.816	0.803	0.771 - 0.835	0.941#
AJCC-TNM	0.644	0.616 -0.672	0.630	0.588 - 0.673	0.469#
P	< 0.0001*		< 0.0001&		
2-year CSS					
Nomogram	0.792	0.771 -0.813	0.802	0.769 - 0.835	0.319#
AJCC-TNM	0.631	0.602 - 0.660	0.633	0.592 - 0.675	0.951#
P	< 0.0001*		< 0.0001&		
3-year OS					
Nomogram	0.818	0.792 -0.844	0.824	0.782 - 0.866	0.934#
AJCC-TNM	0.694	0.660 -0.728	0.691	0.641 - 0.741	0.347#
P	< 0.0001*		< 0.0001&		
3-year CSS					
Nomogram	0.809	0.783 -0.836	0.839	0.800 - 0.879	0.294#
AJCC-TNM	0.679	0.643-0.714	0.692	0.645 - 0.740	0.386#
P	< 0.0001*		< 0.0001&		

Note: #, Delong test for predicting AUC area for OS and 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 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. AUC, the area under the curve; OS, overall survival; CSS, cancer-specific survival; AJCC, the American Joint Committee on Cancer; TNM, tumor-node-metastasis; CI, confidence interval.

Table S2. AUC of predictive performance between internal and external data of nomogram models

Variable	Validation		Test		*P*
	AUC	95% CI	AUC	95% CI	
3.5-year OS					
Nomogram	0.816	0.771-0.861	0.648	0.570-0.726	0.268
3.5-year CSS					
Nomogram	0.833	0.787-0.878	0.637	0.568-0.726	0.155
4-year OS					
Nomogram	0.820	0.771-0.869	0.657	0.513-0.742	0.110
4-year CSS					
Nomogram	0.834	0.784-0.884	0.659	0.504-0.734	0.141
4.5-year OS					
Nomogram	0.850	0.795-0.904	0.672	0.540-0.793	0.100
4.5-year CSS					
Nomogram	0.855	0.802-0.907	0.672	0.537-0.796	0.056