Research in Health Science
ISSN 2470-6205 (Print) ISSN 2470-6213 (Online)
Vol. 9, No. 4, 2024
www.scholink.org/ojs/index.php/rhs

Original Paper

Construction and Validation of a Prognostic Prediction Model for Gastric Mucinous Adenocarcinoma Based on the SEER Database

Peng Yan & Zhiliang Jin*

Jingzhou Hospital Affiliated to Yangtze University, Hubei Jingzhou 434100, China

* Corresponding author: 623357741@qq.com

Received: October 2, 2024 Accepted: October 15, 2024 Online Published: October 31, 2024

doi:10.22158/rhs.v9n4p10 URL: http://dx.doi.org/10.22158/rhs.v9n4p10

Abstract

This study focused on the development and validation of a prognostic model to predict overall survival (OS) in patients with gastric mucinous adenocarcinoma (GMA), utilizing data from the SEER database. A cohort of 537 GMA patients was analyzed to identify factors affecting survival, using both univariate and multivariate Cox regression analyses on demographic and clinicopathological variables. The analysis led to the creation of a nomogram model. Findings revealed that TNM staging and the administration of radiotherapy or chemotherapy were significant factors impacting survival outcomes. Specifically, patients at stages T4, N3, and M1 faced a notably higher mortality risk, while those who received radiotherapy and chemotherapy experienced a significant reduction in mortality risk. The model's performance, as indicated by C-indices of 0.73 in the training set and 0.66 in the validation set, suggests robust predictive capability. The area under the curve (AUC) analysis confirmed high accuracy in predicting 1-year, 3-year, and 5-year survival rates, and calibration curves displayed strong agreement between predicted and actual outcomes. Decision curve analysis (DCA) demonstrated that the model exhibited significant clinical utility across various thresholds. Overall, this study presents a reliable prognostic tool for individualized risk stratification and clinical treatment in GMA patients. Future work should focus on enhancing the model by incorporating molecular biological data to increase its predictive accuracy and applicability.

Keywords

Gastric Mucinous Adenocarcinoma (GMA), Prognostic Model, SEER Database, Nomogram

1. Introduction

Gastric mucinous adenocarcinoma (GMA) is recognized as a highly aggressive subtype of gastric cancer, constituting approximately 3-5% of all gastric cancer cases. In recent years, the incidence of

gastric cancer has been increasing annually, influenced by the global aging population and lifestyle changes, rendering research on gastric mucinous adenocarcinoma increasingly significant[1–3].However, due to its distinct pathological characteristics and variability in treatment response, the survival rate of gastric mucinous adenocarcinoma remains significantly lower than that of other gastric cancer subtypes[4,5].According to the most recent global cancer statistics, the five-year survival rate for gastric cancer ranges between 20% and 30% worldwide; conversely, the five-year survival rate for patients with gastric mucinous adenocarcinoma is even lower, falling below 20%. This represents a considerable challenge to global public health[6,7].

With the advancement of bioinformatics and database technologies, extensive databases have unveiled new opportunities for developing prognostic models for tumors. The Surveillance, Epidemiology, and End Results (SEER) database, maintained by the National Cancer Institute, contains extensive information on gastric cancer patients, thereby offering a valuable resource for investigating prognostic factors in gastric mucinous adenocarcinoma. Through the analysis of data from 537 patients diagnosed with gastric mucinous adenocarcinoma in the SEER database, we can investigate key factors influencing the survival of these patients from various dimensions, including demographic characteristics, pathological grading, and TNM staging, to achieve precise prognostic predictions.

This study developed and validated a prognostic prediction model for patients diagnosed with gastric mucinous adenocarcinoma using the SEER database. Utilizing Cox regression analysis, we identified significant factors influencing overall survival (OS), which ultimately facilitated the development of a nomogram model. The predictive accuracy of the model was validated through the consistency index (C-index) and the area under the time-dependent ROC curve (AUC), while its clinical applicability was assessed via calibration curves and decision curve analysis (DCA). The results indicate that the model provides individualized predictions for 1-year, 3-year, and 5-year survival rates for patients diagnosed with gastric mucinous adenocarcinoma, offering clinicians a scientific tool for prognostic assessment and aiding in the formulation of personalized treatment strategies.

2. Method

2.1 Data Collection

This study incorporated clinical case data obtained from the Surveillance, Epidemiology, and End Results (SEER) database of the National Cancer Institute. Utilizing SEER*Stat version 8.4.2, a total of 1,590 clinical cases of gastric mucinous adenocarcinoma from 2000 to 2019 were extracted. All patient information in the SEER database has been de-identified, thereby eliminating the need for informed consent from patients or approval from an ethics committee.

2.2 Inclusion and Exclusion Criteria

Inclusion criteria included: (1) Patients with a pathologically confirmed diagnosis of gastric mucinous adenocarcinoma; (2) complete follow-up information; and (3) staging of cases according to the 6th edition of the UICC/AJCC TNM staging system.

Exclusion criteria included: (1) Patients with multiple primary tumors; (2) patients with an unknown cause of death; and (3) incomplete clinical case data regarding tumor type, treatment information, degree of differentiation, and TNM staging. Based on these criteria, a total of 537 patients with gastric mucinous adenocarcinoma were ultimately included in the study.

2.3 Indicators Included in the Analysis

The study incorporated data regarding gender, age, ethnicity, race, marital status, tumor location, pathological grade, TNM staging, surgical status of the primary lesion, lymph node surgery, radiotherapy, chemotherapy, the sequence of surgery and radiotherapy, waiting time for surgery, months of survival, and survival status for analysis.

2.4 Statistical Analysis

Statistical analyses were conducted using R software (version 4.2.3). Chi-square tests were utilized to compare patient characteristics between the modeling and validation groups. Univariate and multivariate Cox regression analyses were performed to identify demographic and clinicopathological features significantly associated with survival. Subsequently, a nomogram was developed using the selected variables to predict the 1-year, 3-year, and 5-year overall survival (OS) of patients with gastric mucinous adenocarcinoma. The model's discriminative ability was assessed using the concordance index (C-index) and the area under the time-dependent receiver operating characteristic (ROC) curve (time-dependent AUC). Calibration curves were employed to assess the consistency between the probabilities generated by the nomogram and the observed actual probabilities. Decision curve analysis (DCA) assessed the clinical utility of the nomogram, while the accuracy of the risk scores derived from the nomogram was evaluated using receiver operating characteristic (ROC) curves. In this study, a p-value of less than 0.05 was considered statistically significant.

3. Results

3.1 Basic Characteristics of Included Cases

This study comprised 537 patients diagnosed with gastric mucinous adenocarcinoma from the SEER database. To ensure the reliability and broad applicability of the model, the study data were partitioned into a training set and a validation set, comprising 375 and 162 patients, respectively (Table 1). A balance test was conducted on the baseline characteristics of both the training and validation sets, revealing no statistically significant differences between the two groups ($P > 0.05$), thus indicating good comparability of the data. There was no significant difference in age between the two groups ($P = 0.361$). Among the patients, those younger than 60 years constituted 21.60% (n = 116), whereas those aged 60 years or older comprised 78.40% (n = 421). The proportion of male patients was higher, constituting 70.39% (n = 378), whereas female patients represented 29.61% (n = 159). There was no significant difference in gender distribution between the two groups ($P = 0.101$). In terms of racial distribution, white patients constituted 69.83% (n = 375), black patients 14.53% (n = 78), and patients of other races 15.64% (n = 84). There was no significant difference in racial distribution between the

training and validation sets (P = 0.459). Married patients constituted 62.01% (n = 333), single patients represented 12.66% (n = 68), and patients with other marital statuses accounted for 25.33% (n = 136), with no significant difference in marital status between the two groups (P = 0.973). The tumor locations were primarily concentrated in the body of the stomach, fundus, and other regions, constituting 7.82%, 4.28%, and 87.90%, respectively. There was no significant difference in the distribution of tumor locations between the two groups (P = 0.896). The distribution of pathological grades was as follows: grade I, 7.45%; grade II, 37.99%; and grade III, 54.56%. There was no significant difference in the distribution of tumor grades between the two groups (P = 0.240). In the T staging, the proportions of patients at T1 to T4 were 11.17%, 15.08%, 34.08%, and 39.66%, respectively. In the N staging, the proportions of patients at N0 to N3 were 33.33%, 43.20%, 15.46%, and 8.01%, respectively. For the M staging, M0 constituted 82.68% and M1 17.32%. There were no significant differences in TNM staging between the training and validation sets (P > 0.05). The proportion of patients undergoing surgery for the primary lesion was 88.08%, while the proportion undergoing lymph node surgery was 78.03%. Radiotherapy and chemotherapy were administered to 37.99% and 54.38% of patients, respectively. The use of surgery, radiotherapy, and chemotherapy demonstrated no significant differences between the two groups (P > 0.05).

Table 1. Balance Test for Training and Validation Sets

Variables	Total (n = 537)	1 (n = 375)	2 (n = 162)	Statistic	P
< 60y	116 (21.60)	85 (22.67)	31 (19.14)		
≥60y	421 (78.40)	290 (77.33)	131 (80.86)		
Sex, n(%)				χ^2=2.69	0.101
Female	159 (29.61)	119 (31.73)	40 (24.69)		
Male	378 (70.39)	256 (68.27)	122 (75.31)		
Race, n(%)				χ^2=1.56	0.459
Black	78 (14.53)	59 (15.73)	19 (11.73)		
Other	84 (15.64)	59 (15.73)	25 (15.43)		
White	375 (69.83)	257 (68.53)	118 (72.84)		
Marriage, n(%)				χ^2=0.05	0.973
Married	333 (62.01)	233 (62.13)	100 (61.73)		
others	136 (25.33)	94 (25.07)	42 (25.93)		
Single	68 (12.66)	48 (12.80)	20 (12.35)		
Location, n(%)				χ^2=0.22	0.896
Body of stomach	42 (7.82)	28 (7.47)	14 (8.64)		
Fundus of stomach	23 (4.28)	16 (4.27)	7 (4.32)		

other	472 (87.90)	331 (88.27)	141 (87.04)		
Grade, n(%)				χ^2=2.85	0.240
Grade I	40 (7.45)	25 (6.67)	15 (9.26)		
Grade II	204 (37.99)	137 (36.53)	67 (41.36)		
Grade III	293 (54.56)	213 (56.80)	80 (49.38)		
T, n(%)				χ^2=2.01	0.571
T1	60 (11.17)	44 (11.73)	16 (9.88)		
T2	81 (15.08)	59 (15.73)	22 (13.58)		
T3	183 (34.08)	121 (32.27)	62 (38.27)		
T4	213 (39.66)	151 (40.27)	62 (38.27)		
N, n(%)				χ^2=1.76	0.624
N0	179 (33.33)	121 (32.27)	58 (35.80)		
N1	232 (43.20)	169 (45.07)	63 (38.89)		
N2	83 (15.46)	56 (14.93)	27 (16.67)		
N3	43 (8.01)	29 (7.73)	14 (8.64)		
M, n(%)				χ^2=0.07	0.793
M0	444 (82.68)	309 (82.40)	135 (83.33)		
M1	93 (17.32)	66 (17.60)	27 (16.67)		
Primary focus surgery, n(%)				χ^2=0.04	0.841
0	64 (11.92)	44 (11.73)	20 (12.35)		
1	473 (88.08)	331 (88.27)	142 (87.65)		
Lymph node surgery, n(%)				χ^2=2.12	0.347
1 to 3	36 (6.70)	29 (7.73)	7 (4.32)		
4 or more	419 (78.03)	289 (77.07)	130 (80.25)		
None	82 (15.27)	57 (15.20)	25 (15.43)		
Radiotherapy, n(%)				χ^2=2.09	0.149
Beam radiation	204 (37.99)	135 (36.00)	69 (42.59)		
None/Unknown	333 (62.01)	240 (64.00)	93 (57.41)		
Surgical radiotherapy, n(%)				χ^2=4.81	0.090
No radiation	364 (67.78)	260 (69.33)	104 (64.20)		
Radiation after surgery	128 (23.84)	90 (24.00)	38 (23.46)		
Radiation prior to surgery	45 (8.38)	25 (6.67)	20 (12.35)		
Chemotherapy, n(%)				χ^2=0.04	0.837
No/Unknown	245 (45.62)	170 (45.33)	75 (46.30)		
Yes	292 (54.38)	205 (54.67)	87 (53.70)		
Surgical waiting time, n(%)				χ^2=0.21	0.649

≤1 month	421 (78.40)	292 (77.87)	129 (79.63)
> 1 month	116 (21.60)	83 (22.13)	33 (20.37)

χ^2: Chi-square test

3.2 Univariate and Multivariate Analysis of OS in Gastric Mucinous Adenocarcinoma Patients

Univariate analysis revealed (Table 2) that the mortality risk for patients aged ≥60 years was significantly higher than that for those aged <60 years (HR = 1.45, 95% CI: 1.08–1.94, P = 0.013). Compared to married patients, individuals with other marital statuses had a significantly increased mortality risk (HR = 1.33, 95% CI: 1.03–1.74, P = 0.031). The mortality risk for patients at stage T4 was significantly higher than that for those at stage T1 (HR = 1.59, 95% CI: 1.08–2.35, P = 0.019). The mortality risk for patients at stage N3 was significantly higher than that for those at stage N0 (HR = 2.10, 95% CI: 1.34–3.30, P = 0.001). The mortality risk for patients at stage M1 was significantly higher than that for those at stage M0 (HR = 2.65, 95% CI: 1.98–3.55, P < 0.001). Patients who did not receive radiotherapy had a significantly higher mortality risk compared to those who did (HR = 1.62, 95% CI: 1.26–2.07, P < 0.001). Patients who did not receive chemotherapy had a significantly higher mortality risk compared to those who did (HR = 0.61, 95% CI: 0.48–0.76, P < 0.001). Patients with a surgical waiting time of ≥1 month had a significantly higher mortality risk compared to those with a waiting time of ≤1 month (HR = 0.70, 95% CI: 0.53–0.94, P = 0.017). In the multivariate Cox regression analysis (Table 2), after adjusting for other clinical variables, it was found that the mortality risk for patients at stages T3 and T4 was significantly increased, with HRs of 2.13 (95% CI: 1.37–3.30, P < 0.001) for T3 and 2.51 (95% CI: 1.63–3.88, P < 0.001) for T4. The mortality risk for patients at stages N1, N2, and N3 gradually increased, with an HR of 3.31 (95% CI: 1.96–5.58, P < 0.001) for those at N3. The mortality risk for patients at stage M1 was significantly elevated (HR = 1.73, 95% CI: 1.24–2.42, P = 0.001). The mortality risk for patients who did not receive radiotherapy remained significantly elevated (HR = 1.36, 95% CI: 1.01–1.82, P = 0.042). The mortality risk for patients who did not receive chemotherapy was significantly elevated (HR = 0.40, 95% CI: 0.30–0.54, P < 0.001). The absence of lymph node surgery also significantly elevated the mortality risk (HR = 2.40, 95% CI: 1.19–4.83, P = 0.015). Both univariate and multivariate analyses indicated that T staging, N staging, M staging, radiotherapy, and chemotherapy are significant prognostic factors influencing overall survival in patients with gastric mucinous adenocarcinoma. In particular, patients at stages T4, N3, and M1 have a poorer prognosis, indicating an elevated mortality risk. Furthermore, the administration of radiotherapy and chemotherapy significantly reduced the mortality risk for these patients.

Table 2. Results of Univariate and Multivariate Cox Proportional Hazards Model

Variables	Single factor					Multi-factor				
	β	S.E	Z	P	hR (95%CI)	β	S.E	Z	P	hR (95%CI)
Age										
＜60 years					1.00 (Reference)					1.00 (Reference)
≥ 60 years	0.37	0.15	2.47	0.013	1.45 (1.08 ~ 1.94)	0.29	0.17	1.73	0.084	1.34 (0.96 ~ 1.86)
Sex										
Female					1.00 (Reference)					
Male	0.09	0.13	0.70	0.485	1.09 (0.85 ~ 1.41)					
Race										
Black					1.00 (Reference)					
Other	-0.31	0.22	-1.40	0.163	0.74 (0.48 ~ 1.13)					
White	-0.02	0.16	-0.10	0.923	0.98 (0.71 ~ 1.36)					
Marriage										
Married					1.00 (Reference)					1.00 (Reference)
Single	-0.16	0.19	-0.82	0.411	0.85 (0.59 ~ 1.24)	-0.37	0.21	-1.77	0.076	0.69 (0.46 ~ 1.04)
others	0.29	0.13	2.15	0.031	1.33 (1.03 ~ 1.74)	0.07	0.14	0.51	0.608	1.07 (0.82 ~ 1.42)
Location										
Body of stomach					1.00 (Reference)					1.00 (Reference)
Fundus of stomach	0.62	0.34	1.85	0.065	1.86 (0.96 ~ 3.60)	0.37	0.35	1.07	0.285	1.45 (0.73 ~ 2.89)
other	0.06	0.23	0.27	0.785	1.07 (0.68 ~ 1.68)	-0.22	0.24	-0.91	0.361	0.80 (0.50 ~ 1.28)
Grade										
Grade I					1.00					

					(Reference)					
Grade II	-0.15	0.25	-0.59	0.552	0.86 (0.53 ~ 1.41)					
Grade III	0.11	0.24	0.43	0.665	1.11 (0.69 ~ 1.79)					
T										
T1					1.00 (Reference)				1.00 (Reference)	
T2	-0.39	0.25	-1.59	0.112	0.68 (0.42 ~ 1.10)	-0.02	0.26	-0.09	0.926	0.98 (0.59 ~ 1.61)
T3	0.19	0.20	0.92	0.359	1.21 (0.81 ~ 1.80)	0.76	0.22	3.37	<.001	2.13 (1.37 ~ 3.30)
T4	0.47	0.20	2.35	0.019	1.59 (1.08 ~ 2.35)	0.92	0.22	4.16	<.001	2.51 (1.63 ~ 3.88)
N										
N0					1.00 (Reference)				1.00 (Reference)	
N1	0.17	0.14	1.20	0.230	1.18 (0.90 ~ 1.55)	0.64	0.16	3.95	<.001	1.90 (1.38 ~ 2.61)
N2	0.59	0.18	3.23	0.001	1.80 (1.26 ~ 2.56)	1.07	0.21	5.07	<.001	2.92 (1.93 ~ 4.41)
N3	0.74	0.23	3.23	0.001	2.10 (1.34 ~ 3.30)	1.20	0.27	4.48	<.001	3.31 (1.96 ~ 5.58)
M										
M0					1.00 (Reference)				1.00 (Reference)	
M1	0.97	0.15	6.52	<.001	2.65 (1.98 ~ 3.55)	0.55	0.17	3.20	0.001	1.73 (1.24 ~ 2.42)
Primary focus surgery										
0					1.00 (Reference)				1.00 (Reference)	
1	-0.89	0.17	-5.25	<.001	0.41 (0.29 ~ 0.57)	-0.48	0.33	-1.47	0.143	0.62 (0.32 ~ 1.18)
Lymph node surgery										
1 to 3					1.00 (Reference)				1.00 (Reference)	

4 or more	-0.30	0.22	-1.39	0.164	0.74 (0.49 ~ 1.13)	-0.32	0.23	-1.39	0.165	0.73 (0.47 ~ 1.14)
None	0.60	0.25	2.45	0.014	1.83 (1.13 ~ 2.97)	0.87	0.36	2.44	0.015	2.40 (1.19 ~ 4.83)
Radiotherapy										
Beam radiation					1.00 (Reference)					1.00 (Reference)
None/Unknown	0.48	0.13	3.83	<.001	1.62 (1.26 ~ 2.07)	0.31	0.15	2.04	0.042	1.36 (1.01 ~ 1.82)
Surgical/radiotherapy										
No radiation					1.00 (Reference)					
Radiation after surgery	-0.70	0.15	-4.76	<.001	0.50 (0.37 ~ 0.66)					
Radiation prior to surgery	-0.63	0.26	-2.42	0.015	0.53 (0.32 ~ 0.89)					
Chemotherapy										
No/Unknown					1.00 (Reference)					1.00 (Reference)
Yes	-0.50	0.12	-4.25	<.001	0.61 (0.48 ~ 0.76)	-0.92	0.15	-5.98	<.001	0.40 (0.30 ~ 0.54)
Surgical waiting time										
≤1 moths					1.00 (Reference)					
≥1 moths	-0.35	0.15	-2.40	0.017	0.70 (0.53 ~ 0.94)					

HR: Hazards Ratio, CI: Confidence Interval

3.3 Building Nomogram Models

Initially, univariate and multivariate Cox regression analyses were performed to identify clinical variables significantly associated with overall survival (OS) in patients with gastric mucinous adenocarcinoma. The results of the analysis indicated that T staging, N staging, M staging, and whether or not the patient received radiotherapy or chemotherapy were significant prognostic factors influencing OS. For example, the mortality risk for patients at stage T4 (HR = 2.51, 95% CI: 1.63–3.88, P < 0.001), N3 (HR = 3.31, 95% CI: 1.96–5.58, P < 0.001), and M1 (HR = 1.73, 95% CI: 1.24–2.42, P = 0.001) was significantly higher than that for patients at earlier stages. Furthermore, patients who did not receive radiotherapy (HR = 1.36, 95% CI: 1.01–1.82, P = 0.042) and those who did not receive

chemotherapy (HR = 0.40, 95% CI: 0.30–0.54, P < 0.001) exhibited significantly elevated mortality risks. Based on these significant prognostic factors, a nomogram model was developed using the "rms" package in R software (Figure 1). This model generates individualized risk scores for each patient and predicts their 1-year, 3-year, and 5-year overall survival (OS) probabilities based on these scores.

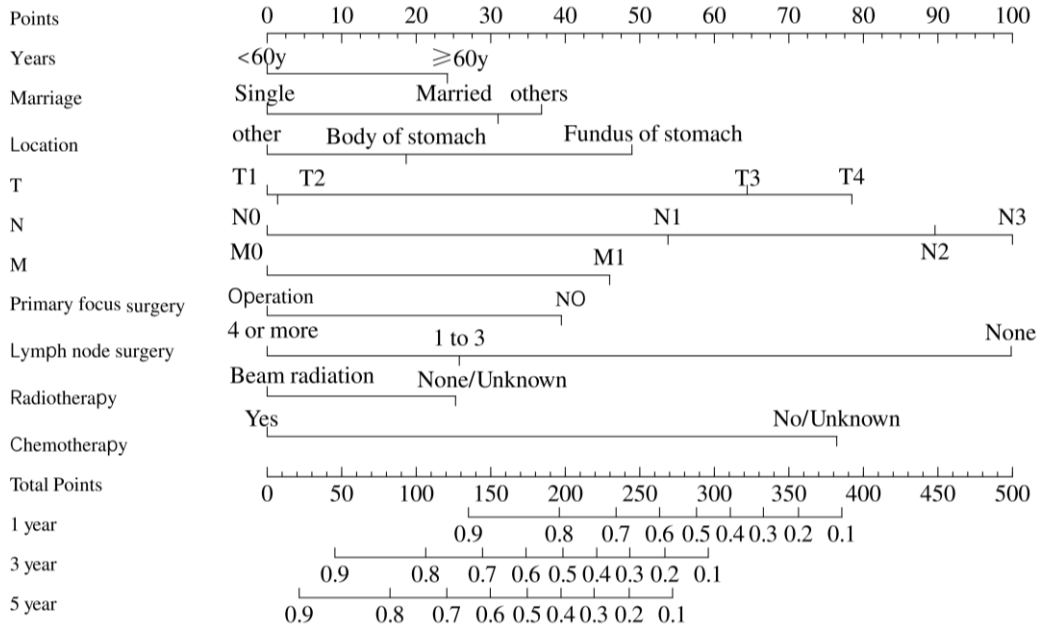

Figure 1. Nomogram

This nomogram illustrates the key factors influencing the prognosis of patients with gastric cancer, along with their corresponding scoring system. The figure encompasses the basic characteristics of patients (such as age and marital status), tumor location, tumor staging (T staging, N staging, M staging), surgical status of the primary lesion, lymph node surgery status, and the status of radiotherapy and chemotherapy. Each variable corresponds to a specific score, and the total score is derived by summing these individual scores. The three axes at the bottom display the survival probabilities at 1 year, 3 years, and 5 years corresponding to the total score, thereby predicting the prognosis of patients based on different total scores.

3.4 Model Performance Verification

The predictive performance of the model was assessed using the concordance index (C-index) and the area under the time-dependent receiver operating characteristic (ROC) curve (AUC). The C-index for the training set was 0.73 (95% CI: 0.70–0.76), while that for the validation set was 0.66 (95% CI: 0.61–0.72), indicating that the model demonstrates good predictive accuracy. The time-dependent ROC curve exhibited AUC values of 0.78, 0.82, and 0.81 for 1-year, 3-year, and 5-year survival, respectively, further validating the model's discriminative ability (Figure 2). Calibration curve analysis revealed that the predicted values from the nomogram model were in good agreement with the actual observed

values, indicating high predictive accuracy at various time points (Figure 3). Additionally, decision curve analysis (DCA) evaluated the clinical utility of the nomogram, revealing that the model offers net clinical benefit across various threshold probabilities (Figure 4).

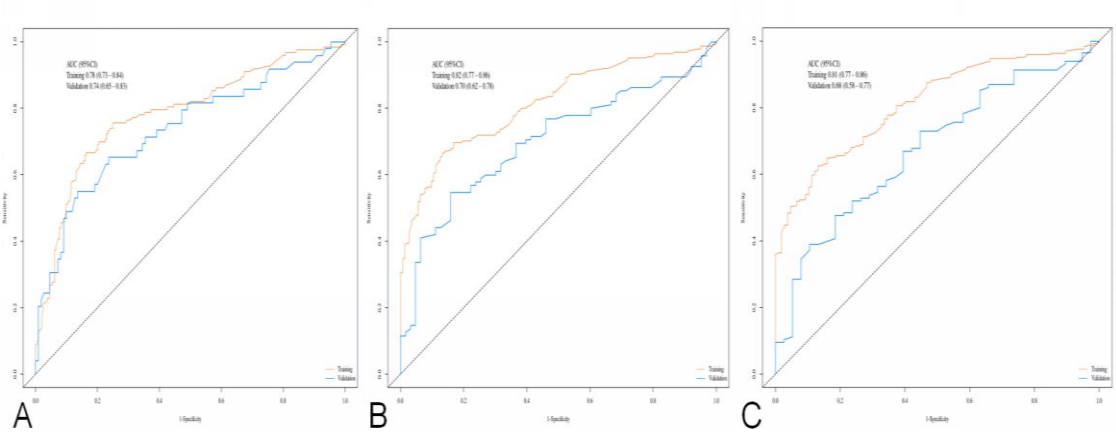

Figure 2. ROC Curves for Training and Validation Sets

This figure illustrates the ROC curve performance of the survival prediction models for patients with gastric mucinous adenocarcinoma at 1 year (A), 3 years (B), and 5 years (C) in both the training and validation sets. The area under the curve (AUC) is utilized to evaluate the model's discriminative ability. The AUC values for the training and validation sets are as follows: for the 1-year prediction model (Training set AUC: 0.78 [95% CI: 0.73–0.84], Validation set AUC: 0.74 [95% CI: 0.65–0.83]), for the 3-year prediction model (Training set AUC: 0.82 [95% CI: 0.77–0.86], Validation set AUC: 0.70 [95% CI: 0.62–0.78]), and for the 5-year prediction model (Training set AUC: 0.81 [95% CI: 0.77–0.86], Validation set AUC: 0.68 [95% CI: 0.58–0.77]).

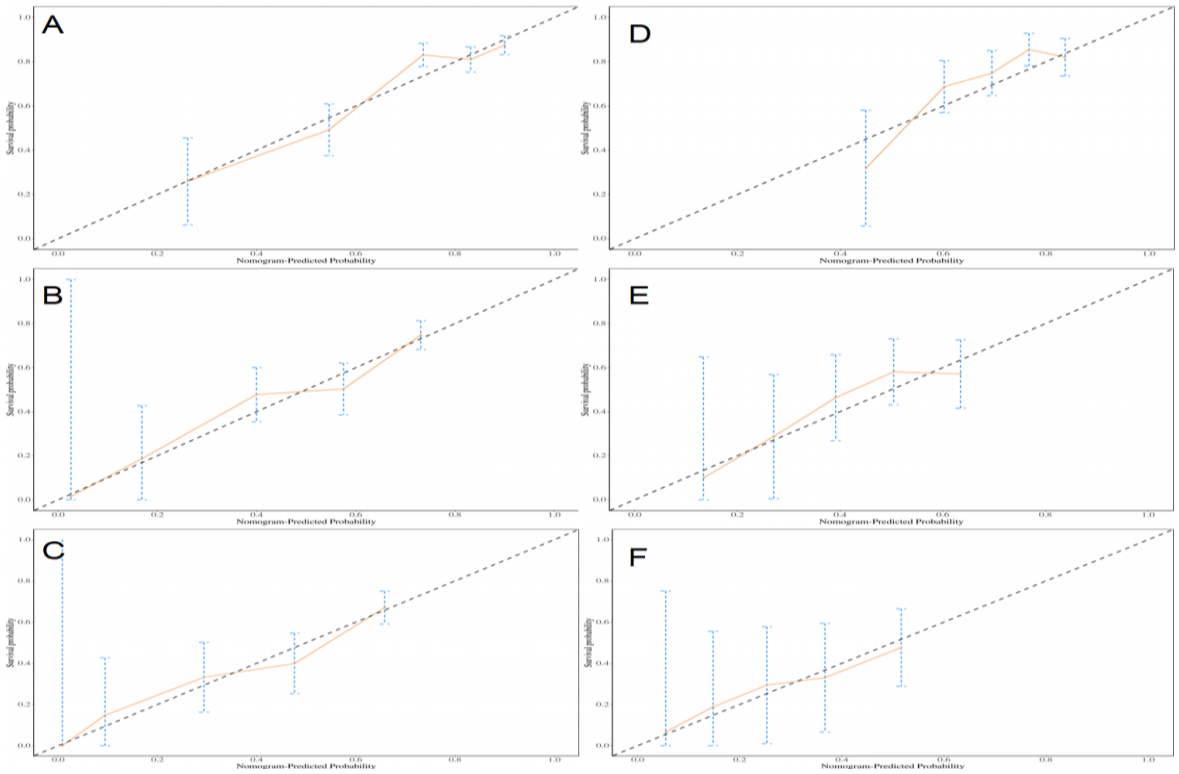

Figure 3. Calibration Plots for Training and Validation Sets

This figure illustrates the calibration performance of the survival prediction model for patients with gastric mucinous adenocarcinoma in both the training and validation sets. Panels A, B, and C depict the calibration curves for 1-year, 3-year, and 5-year survival rates in the training set, whereas panels D, E, and F depict the calibration curves for the same survival rates in the validation set. The x-axis represents the predicted survival probabilities from the nomogram, while the y-axis denotes the actual observed survival probabilities. The dashed line represents perfect calibration, wherein the predicted probabilities align precisely with the actual probabilities. The orange solid line indicates the calibration of the model, and the error bars represent the 95% confidence intervals for various probability segments. The closer the calibration curve is to the dashed line, the more accurate the model's predictions become.

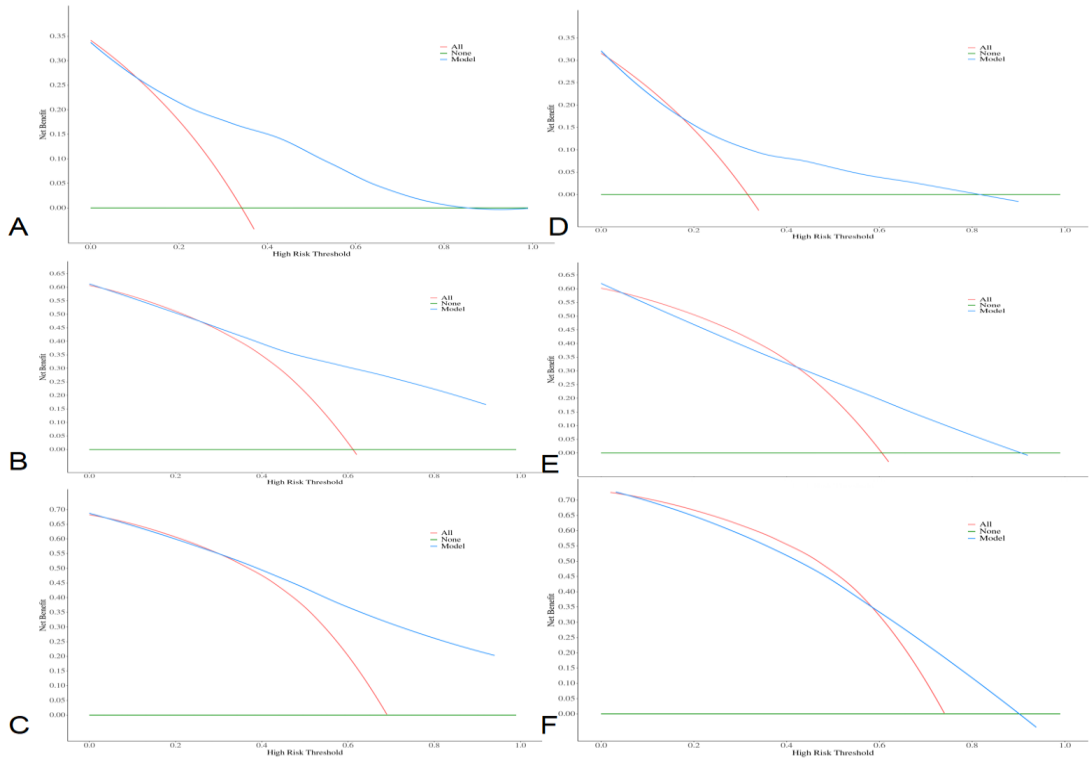

Figure 4. DCA Plot for Training and Validation Sets

This figure illustrates the decision curve analysis (DCA) for the survival prediction model of patients with gastric mucinous adenocarcinoma at 1 year, 3 years, and 5 years in both the training set (A, B, C) and validation set (D, E, F). The x-axis represents the high-risk threshold, whereas the y-axis indicates the net benefit. The green straight line (None) represents the scenario without intervention, the red curve (All) indicates the scenario in which all patients receive intervention, and the blue curve (Model) reflects the scenario where intervention is based on the predictive model.

4. Discussion

4.1 Overview of Research Results and Significance

This study analyzed clinical data from 537 patients diagnosed with gastric mucinous adenocarcinoma (GMA) within the SEER database, successfully constructing and validating a prognostic model that accurately predicts overall survival (OS) at 1, 3, and 5 years. GMA is a rare and highly aggressive subtype of gastric cancer; its unique biological characteristics and diverse treatment responses lead to significantly lower survival rates when compared to other gastric cancer types. In recent years, the global incidence of gastric cancer has been steadily increasing, particularly among the elderly population. Therefore, the development of a reliable model for predicting the survival prognosis of GMA patients is of significant clinical importance for promoting personalized treatment and optimizing therapeutic strategies.

The study employed both univariate and multivariate Cox regression analyses to identify significant prognostic factors influencing OS in GMA patients, including age, TNM staging, and treatment modalities such as radiotherapy and chemotherapy. The findings indicated that patients with advanced TNM stages (T3/T4, N3, M1), individuals aged 60 and older, and those who did not receive radiotherapy or chemotherapy experienced poorer survival outcomes. Notably, patients who did not receive radiotherapy or chemotherapy exhibited a significantly elevated risk of mortality, while the survival rates of T4, N3, and M1 stage patients were markedly lower than those of early-stage patients. These results are consistent with previous literature and further underscore the importance of staging, treatment selection, and age factors in the management of GMA patients.

The innovative aspect of this study lies in the construction of a nomogram based on the results of multivariate Cox regression analysis derived from the SEER database. The model was validated using various metrics, including the concordance index (C-index), area under the time-dependent receiver operating characteristic (ROC) curve (AUC), and calibration curves, thereby demonstrating excellent predictive performance, particularly for short-term (1 year), mid-term (3 years), and long-term (5 years) prognostic predictions.

4.2 Model Performance Evaluation

The prognostic model for gastric mucinous adenocarcinoma (GMA) developed in this study is based on multivariate Cox regression analysis and utilizes a nomogram for individualized survival predictions. To assess the predictive performance of the model, we conducted a comprehensive evaluation utilizing the concordance index (C-index), area under the time-dependent ROC curve (AUC), and calibration curves in both the training and validation sets. This systematic evaluation not only demonstrates the reliability of the model in predicting overall survival (OS) for GMA patients at 1 year, 3 years, and 5 years but also provides a scientific basis for its application in clinical practice.

First, the C-index of the model was 0.73 in the training set and 0.66 in the validation set, indicating strong discriminative ability. The C-index is a measure of the model's ability to accurately rank survival prognoses, with higher values indicating greater predictive capability. The C-index values in this study suggest that the constructed nomogram model can effectively differentiate between patient populations with varying risk levels, demonstrating high predictive stability. In comparison to previous gastric cancer prognostic models that did not consider the specificities of the GMA subtype, this model exhibits stable discriminative ability in both the training and validation sets, further emphasizing its potential value in the individualized management of GMA.

Secondly, the analysis of time-dependent ROC curves further validated the predictive accuracy of the model. The AUC values for 1-year, 3-year, and 5-year survival predictions were 0.78, 0.82, and 0.81, respectively, indicating that the model possesses high sensitivity and specificity in predicting patient survival across different time points. Notably, for short-term and mid-term survival predictions, the AUC values exceeded 0.75, satisfying the high standards required for prognostic models in clinical settings. These AUC results not only demonstrate the robustness of the nomogram model across various survival

stages but also provide a reliable tool for clinical risk stratification in GMA patients. Furthermore, in comparison to other gastric cancer prognostic models, the model developed in this study demonstrates higher specificity for GMA patients, suggesting that it aligns more closely with the pathological and biological characteristics of the GMA subtype and can provide more targeted survival predictions for these patients.

Calibration curve analysis was performed to verify the consistency between the model's predicted survival probabilities and the actual observed values. The results indicated that for 1-year, 3-year, and 5-year survival predictions, the model's calibration curves closely aligned with the ideal 45-degree line, demonstrating a high degree of consistency between the predicted values and the actual survival outcomes at various time points. Calibration is one of the key indicators for assessing the applicability of prognostic models in individualized survival predictions. This result demonstrates that the nomogram model developed in this study not only exhibits strong discrimination in risk stratification but also facilitates accurate predictions of individual survival probabilities, further confirming the model's clinical applicability.

Additionally, decision curve analysis (DCA) indicated that the model developed in this study offers a high net benefit across various threshold probabilities.Decision curve analysis (DCA) is an evaluation method that integrates the predictive results of the model with actual clinical benefits. It considers the impacts of false positive and false negative rates, thereby providing a foundation for clinical decision-making[8].The DCA results of this study indicate that the nomogram model provides a greater net benefit for predicting 1-year, 3-year, and 5-year overall survival (OS) compared to the "treat all" or "treat none" strategies. This suggests that utilizing the model can more effectively guide individualized treatment decisions for GMA patients across various risk levels. Notably, among high-risk patient populations, the model's net benefit is particularly substantial, demonstrating its potential to guide the management of complex cases.

4.3 Clinical Significance of Prognostic Factors

In the prognostic model for gastric mucinous adenocarcinoma (GMA) developed using the SEER database, several prognostic factors of significant clinical relevance were identified, including age, TNM staging, and the administration of radiotherapy and chemotherapy. Each prognostic factor plays a crucial role in the survival prognosis of GMA patients and significantly impacts their long-term survival rates. Understanding these factors not only facilitates risk stratification but also provides a scientific foundation for individualized treatment plans for GMA patients.

First, age has been established as a significant factor influencing the survival rate of GMA patients. This study demonstrates that patients aged ≥60 years face a significantly elevated risk of mortality, a finding that aligns with previous research.Elderly patients frequently exhibit a higher incidence of comorbidities, decreased physical function, and reduced tolerance to treatment. These factors may constrain treatment options and subsequently affect survival prognosis[9].Additionally, tumors in elderly patients frequently exhibit more aggressive biological characteristics, thereby exacerbating the risk of poor

prognosis[10].Therefore, the management of elderly GMA patients in clinical practice necessitates increased caution. It is essential to conduct a comprehensive assessment based on the patient's health status to strike a balance between improving survival rates and enhancing quality of life.

Secondly, TNM staging exhibits significant prognostic value in the GMA prognostic model. The later the stage, the poorer the survival prognosis for patients, particularly for those at T4, N3, and M1 stages, where the risk of mortality is significantly elevated. Patients at T4 stage have tumors that have invaded tissues beyond the gastric wall, whereas N3 stage indicates extensive lymph node metastasis, and M1 stage denotes the presence of distant metastasis. These pathological characteristics suggest that advanced GMA exhibits a higher level of aggressiveness, thereby significantly increasing the complexity and difficulty of treatment. The lower survival rates in high-stage patients underscore the importance of accurate staging assessments during early diagnosis. For advanced patients, it is crucial to actively consider comprehensive treatment strategies, including surgery, chemotherapy, and targeted therapy, to maximize survival duration.

Radiotherapy has demonstrated significant effects in improving survival for GMA patients. In this study, patients receiving radiotherapy exhibited a markedly reduced risk of mortality, indicating the potential value of radiotherapy in the local control of GMA and in extending survival. Although GMA may be less sensitive to radiotherapy than other cancer types, it can serve as an adjunctive treatment alongside surgery and chemotherapy, particularly for patients who cannot tolerate extensive surgery, thereby providing effective local control. Future studies could further explore the optimal radiation dose and its synergistic effects with other treatment modalities to refine comprehensive treatment strategies for GMA.

Chemotherapy has also been shown to have a significant protective effect on the survival of GMA patients. Patients receiving chemotherapy exhibited a significantly reduced risk of mortality, highlighting the important role of chemotherapy in the management of GMA. Chemotherapy is particularly advantageous for controlling micrometastases, especially in advanced or metastatic patients, as it can effectively delay disease progression[11].In clinical practice, due to the potential for significant toxic side effects from chemotherapy, it is essential to tailor the chemotherapy regimen based on a comprehensive assessment of the patient's health status. Future studies should include more prospective research to determine the optimal chemotherapy regimens and durations, thereby developing more precise treatment strategies for GMA patients at various stages.

4.4 Limitations of the Study and Future Directions

This study constructed and validated a prognostic prediction model for gastric mucinous adenocarcinoma (GMA) using data from the SEER database, offering a relatively reliable tool for assessing survival risk in GMA patients. However, despite the significant findings of this study, several limitations exist in the model's practical application that must be interpreted cautiously.

First, the limitation of data sources represents one of the primary constraints of this study. The data utilized in this research were sourced exclusively from the SEER database, primarily encompassing

patient information from North America, which may limit applicability in other racial and geographical contexts. Significant differences may exist among various populations regarding the incidence, pathological characteristics, and treatment responses of gastric cancer. Therefore, a model developed using data from a single region may not be applicable on a global scale. Future research should incorporate multi-center and multi-ethnic population databases for external validation of the model to ensure its broad applicability across diverse races and regions, thereby enhancing its clinical utility worldwide.

Second, this study included only a limited set of clinical variables and did not incorporate molecular and genomic characteristics, which may influence the model's predictive accuracy. GMA exhibits unique molecular features and gene expression profiles; incorporating tumor molecular typing and genetic mutation information (such as HER2 and p53 mutations) into the prognostic model could provide a more comprehensive reflection of the disease's biological behavior, thereby enhancing the model's precision. In recent years, with the advancement of precision medicine, the role of molecular markers in cancer prognosis evaluation has gained increasing importance. Future studies could explore the potential for integrating molecular characteristics and genomic data into the GMA prognostic model to provide more personalized survival predictions and treatment guidance for patients.

Finally, the retrospective data utilized in this study may introduce potential bias, thereby affecting the reliability of the results. Although the SEER database provides a wealth of high-quality clinical data, the limitations of retrospective studies include the potential for selection bias and incomplete information, as some patients' historical data and treatment details may be missing or inaccurate. Retrospective data cannot fully account for baseline characteristic differences among patients, nor can it eliminate potential confounding factors. Future research should validate the model through prospective cohort studies to reduce bias, ensure the reliability and reproducibility of the results.

5. Conclusion

This study constructed a prognostic prediction model for patients with gastric mucinous adenocarcinoma using data from the SEER database and validated the model's performance and clinical applicability through various statistical analysis methods. Key prognostic factors, including T staging, N staging, M staging, radiotherapy, and chemotherapy, were identified through Cox regression analysis, ultimately leading to the development of a nomogram model that provides individualized predictions for patients' 1-year, 3-year, and 5-year survival rates.

The model's C-index and time-dependent ROC curve (AUC) demonstrated strong discriminative ability and predictive accuracy, while calibration curves and decision curve analysis (DCA) further validated its clinical benefit. This model provides clinicians with a reliable tool for prognostic assessment, facilitating the formulation of individualized treatment strategies to improve patients' long-term survival rates.

This study not only addresses the shortcomings of previous gastric cancer models that did not adequately consider the specificities of the gastric mucinous adenocarcinoma subtype but also provides a direction

for the future integration of molecular biological data to achieve more precise individualized management. With the accumulation and optimization of additional data, this model is expected to be widely applied in the clinical practice of gastric cancer, thereby contributing to the advancement of precision medicine.

Acknowledgement

I would like to thank my mentor and colleagues for their help in this research, as well as the SEER database for data support.

References

[1] Lee J E, Choi Y Y, An J Y, et al. Clinicopathologic and genomic characteristics of mucinous gastric adenocarcinoma. *Gastric Cancer*, 2022, *25*(4). 697-711.

[2] Ma Y, Zhang S, Wang Y, et al. Comparison of Clinicopathological Features and Prognosis of Mucinous Gastric Carcinoma and other Gastric Cancers: A Retrospective Study of 4,417 Patients. *Journal of Gastrointestinal Surgery*, 2023, *27*(11), 2352-2364.

[3] Aslani A, Soheili A, Mousavi S E, et al. Incidence trends of gastric cancer in the United States over 2000-2020: A population-based analysis. *PLOS ONE*, 2024, *19*(9), e0310040.

[4] Wang Q, Zhong J, Huang Q, et al. A survival comparison of gastric mucin-producing adenocarcinoma to conventional adenocarcinoma: a SEER database analysis. *BMC Cancer*, 2021, *21*(1), 1138.

[5] Cai L, Li Y, Yang X W, et al. Prognostic significance of mucinous component in gastric adenocarcinoma after radical D2 gastrectomy. *OncoTargets and Therapy*, 2018, *11*, 967-973.

[6] Dizon D S, Kamal A H. Cancer statistics 2024: All hands on deck. *CA: A Cancer Journal for Clinicians*, 2024, *74*(1), 8-9.

[7] van Velzen M J M, Braemer M, Nieuwenhuijzen G A P, et al. Incidence, Stage, Treatment, and Survival of Noncardia Gastric Cancer. *JAMA network open*, 2023, *6*(8), e2330018.

[8] Piovani D, Sokou R, Tsantes A G, et al. Optimizing Clinical Decision Making with Decision Curve Analysis: Insights for Clinical Investigators. *Healthcare*, 2023, *11*(16), 2244.

[9] Chen W, Altshuler R D, Daschner P, et al. Older Adults with Cancer and Common Comorbidities-Challenges and Opportunities in Improving Their Cancer Treatment Outcomes. *JNCI: Journal of the National Cancer Institute*, 2024, djae163.

[10] Song X, Liu X, Guo Q, et al. Unraveling the nexus between cellular senescence and malignant transformation: a paradigm shift in cancer research. *Cancer Biology & Medicine*, 2024, *21*(7), 1-6.

[11] Canavan M E, Wang X, Ascha M S, et al. Systemic Anticancer Therapy and Overall Survival in Patients With Very Advanced Solid Tumors. *JAMA Oncology*, 2024, *10*(7), 887.