

## RESEARCH ARTICLE

# A prognostic nomogram for predicting overall survival in colorectal mucinous adenocarcinoma patients based on the SEER database

Qian Wu , Suqin Zhang , Huan Wang , Yifei Zeng , Wei Yang , Wenjun Pan , Guodai Hong , and Wenbin Gao \*

A nomogram was constructed to predict the survival of patients with colorectal mucinous adenocarcinoma based on data extracted from the Surveillance, Epidemiology and End Result (SEER) database. Data collected between 2010 and 2018 were obtained from the SEER database. The log-rank test and multivariate Cox regression were performed to identify the independent prognostic factors for overall survival, which were further used to construct a nomogram model to predict 1-, 3-, and 5-year overall survival. In total, 10,846 patients diagnosed with colorectal mucinous adenocarcinoma were enrolled in the study. The following 11 variables were associated with survival and were further incorporated into the nomogram: age at diagnosis, primary site, grade, tumour size, lymph node dissection, T stage, N stage, M stage, surgery for primary site, chemotherapy, and household income. The concordance index (C-index) value was 0.725 (95% confidence interval 0.716–0.734), and the receiver operating characteristic curves and calibration curves showed satisfactory predictive accuracy. Both the C-index and time-independent area under the curve values were greater than those of the American Joint Committee on Cancer 7th TNM classification system (both  $P < 0.001$ ). In the validation group, the results were consistent with those of the training group, with a C-index value of 0.726 (95% confidence interval 0.713–0.739). This study constructed a practical nomogram to predict 1-, 3-, and 5-year OS for patients with colorectal mucinous adenocarcinoma based on the SEER data.

**Keywords:** Colorectal cancer (CRC), mucinous adenocarcinoma, nomogram, overall survival (OS), Surveillance, Epidemiology and End Result (SEER) program.

## Introduction

Colorectal cancer (CRC) is the third most common cancer and the second leading cause of cancer-related death worldwide [1]. The majority of CRC cases are adenocarcinomas, accounting for approximately 85% of cases. Mucinous adenocarcinoma (MAC) is a less common subtype that accounts for 8%–19% of CRC cases and is defined by the World Health Organization (WHO) as the presence of extracellular mucin in >50% of the tumor area [2, 3]. Many studies have demonstrated the distinct clinical and pathological features of MAC, which is regarded as being more advanced at diagnosis and has a poorer prognosis than nonmucinous adenocarcinoma (NMAC) [4, 5].

Despite advances in treatment strategies, such as targeted therapy and immunotherapy, the 5-year survival of locally advanced and metastatic CRC patients is still unsatisfactory [6]. Effective methods for precision medicine and prognosis prediction are in high demand. Currently, the most prevalent method for prognosis and treatment direction is the tumour-node-metastasis (TNM) staging system, which is classified mainly by the depth of primary tumour invasion (T), numbers of regional lymph node metastases (N), and distant metastasis (M) [7–9].

However, many other clinical, pathologic, and economic factors that have been proven to be related to survival, such as age, race, tumour size, histological subtype, grade, household income, and information about treatment, such as surgery, radiotherapy, and chemotherapy, are not included in this system, influencing the accuracy of its predictive ability [5, 10, 11]. Thus, it is imperative to explore a novel program that can comprise all prognostic factors to predict the outcomes of cancer patients.

A nomogram is a simple visual predictive model that is widely used for prognosis prediction. It is more comprehensive in terms of the inclusion of all effective prognostic factors and more intuitive because the estimated survival of each patient can be conveniently calculated by combining the score of each parameter and matching the corresponding percentage [12]. To the best of our knowledge, there is currently no nomogram specialized for patients with colorectal MAC. Thus, the aim of this study was to establish a nomogram to predict the survival of colorectal MAC patients based on data extracted from the Surveillance, Epidemiology and End Result (SEER) database.

Department of Oncology, Shenzhen Luohu People's Hospital, Shenzhen, China.

\*Correspondence to Wenbin Gao: [drwenbingao@163.com](mailto:drwenbingao@163.com)

DOI: 10.17305/bjbm.2022.8297

© 2022 Wu et al. This article is available under a Creative Commons License (Attribution 4.0 International, as described at <https://creativecommons.org/licenses/by/4.0/>).

## Materials and methods

### Data source and patient selection

Eligible patients were extracted from the database “Incidence-SEER Research Data, 18 Registries, Nov 2020 Sub(2000–2018)” using the SEER\*Stat program (v8.4.0). The inclusion criteria were as follows: (1) patients who were diagnosed with primary colorectal adenocarcinoma according to the third version of the International Classification of Disease for Oncology (ICD-O-3) and (2) patients who had histologically confirmed MAC subtype (ICD-O-3 coded as 8480/8481). The exclusion criteria were as follows: (1) patients with more than one primary cancer; (2) patients diagnosed by clinical criteria or based on autopsy or the death certificate; (3) patients with incomplete survival information; (4) patients with a follow-up of less than 1 month; and (5) patients with incomplete clinicopathological and treatment information (including primary site, histological type, grade, tumour size, lymph node dissection number, American Joint Committee on Cancer [AJCC] 7th TNM stage, surgery, radiotherapy, chemotherapy, and median household income). Eligible patients were randomly divided into a training group and a validation group in a 70:30 ratio.

### Variables and endpoints

In this study, 14 variables were collected from the database, including sex, age at diagnosis, race, primary site, grade, tumour size, lymph node dissection, T stage, N stage, M stage, surgery, radiotherapy, chemotherapy, and median household income. All tumours were staged according to the TNM staging system of the AJCC (7th version, 2009). The primary tumour site was divided into the caecum-ascending colon (including the appendix, caecum, ascending colon, and hepatic flexure), transverse colon, descending colon-sigmoid (including the descending colon, sigmoid colon, and splenic flexure), and rectum. For continuous variables, patients were divided into two groups, with the age of 65 years as the cutoff. The cutoff value for tumour size was 3 cm. The endpoint of this study was overall survival (OS), which was defined as the duration between diagnosis and death due to all causes.

### Ethical statement

All patient data were obtained from the SEER database, which records cancer data for approximately 30% of the American population across different regions [13]. A SEER Research Data Agreement (No. 12068-Nov2021) was signed for data acquisition. Given that the data were publicly accessible and deidentified, patient informed consent was not needed, and no approval from an ethics committee was demanded. This research was performed in accordance with the Declaration of Helsinki.

### Statistical analysis

#### Nomogram construction

Categorical variables are shown as proportions and frequencies and were compared by the Chi-square test or Fisher’s exact test. The associations between each variable and survival were first evaluated with univariate analysis using the log-rank test. Variables with a  $P$  value  $\leq 0.1$  during univariate analysis were further examined by multivariate backward stepwise Cox proportional hazard regression analysis. Statistically

significant variables in the multivariate Cox regression analysis ( $P \leq 0.05$ ) were determined to be independent prognostic factors to predict the survival outcome. Additionally, variance inflation factor examinations of the effective prognostic factors were evaluated to exclude multicollinearity problems. Then, these selected factors were used to establish a nomogram model to predict 1-, 3-, and 5-year OS in the training and validation groups performed by the rms package in R (version 4.1.2).

#### Nomogram validation

The performance of the nomogram was evaluated by detecting its discrimination and calibration abilities both internally (in the training group) and externally (in the validation group). The bootstrapping resampling approach (1000 repetitions) was applied to interval validation. The concordance index (C-index) and the receiver operating characteristic (ROC) curve were used to validate discrimination performance. A higher C-index value and a larger area under the curve (AUC) in the ROC curve represented better discrimination ability. In addition, we calculated the C-index and the ROC curve using the AJCC 7th TNM classification system and then compared the results with our nomogram to identify differences. The calibration curves were used to evaluate the calibration ability of the nomogram. A 45-degree plot represented an optimal model. All statistical analyses were conducted using SPSS 26.0 (SPSS Inc., Armonk, NY, USA) and R (version 4.1.2, www.r-project.org). A difference of  $P < 0.05$  (two-tailed) was considered statistically significant.

## Results

### Patient characteristics

In total, 10,846 patients diagnosed with colorectal MAC were enrolled in the study and were assigned to the training group ( $n = 7527$ ) or the validation group ( $n = 3319$ ). A detailed flowchart of patient selection is presented in Figure 1. Among all the included patients, the majority were female ( $n = 5639$ , 52.0%), aged  $\geq 65$  years ( $n = 6158$ , 56.8%) and white ( $n = 8793$ , 81.1%). A total of 71.7% of patients had a household income higher than 75,000 dollars. The most common primary site was the right colon ( $n = 6645$ , 61.3%), followed by the left colon ( $n = 2324$ , 21.4%), rectum ( $n = 983$ , 9.1%), and transverse colon ( $n = 894$ , 8.2%). The tumour grade proportions of all patients were well ( $n = 1374$ , 12.7%), moderate ( $n = 7171$ , 66.1%), poor ( $n = 1832$ , 16.9%), and undifferentiated ( $n = 469$ , 4.3%). More than 80% of the patients had a tumour larger than 3 cm ( $n = 8958$ , 82.6%). Stages T1-4 accounted for 3.8%, 11.2%, 55.8%, and 29.2% of the population, respectively. Regarding the treatment methods, only 1.2% of the patients did not undergo surgery, and 92.7% of the patients had more than 4 lymph node dissections. Radiotherapy and chemotherapy were administered to 9.3% and 44.8% of the patients, respectively. The 1-, 3-, and 5-year OS rates of all enrolled patients were 87.4%, 68.3%, and 57.2%, respectively. The detailed clinicopathological features of the two groups are listed in Table 1. There were no significant differences between the training and the validation group.

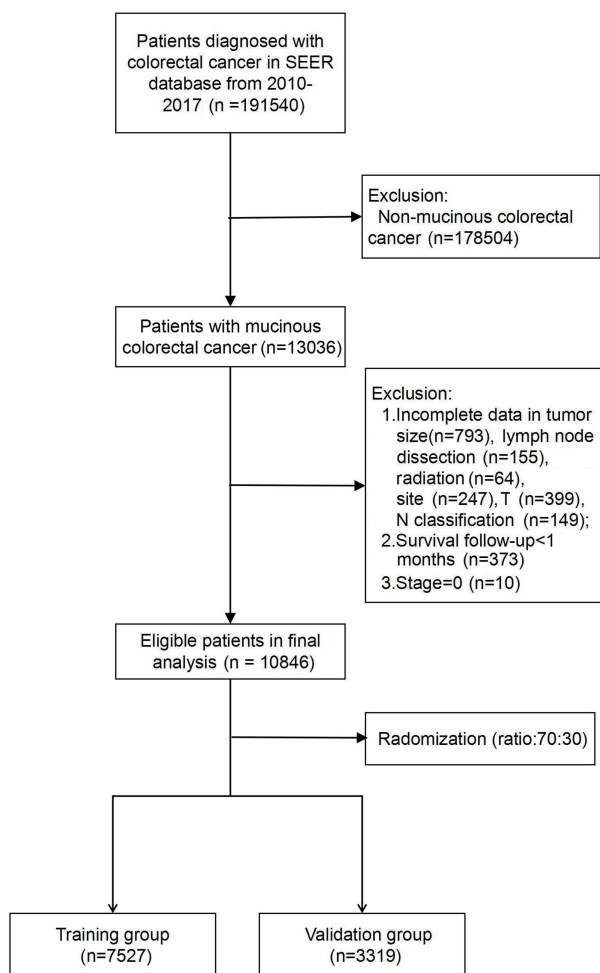


Figure 1. Flowchart of patient selection.

#### Variable selection and nomogram construction

In the univariate analysis, 11 variables were associated with survival: age at diagnosis, primary site, grade, tumour size, lymph node dissection, T stage, N stage, M stage, surgery for primary site, chemotherapy, and median household income. The above variables were then included in the stepwise model, and multivariate analysis showed that all these variables were independent prognostic factors for survival ( $P < 0.05$ ). The variance inflation factor exam suggested that multicollinearity issues did not exist (Figure S1). More details about the results of univariate and multivariate Cox regression analyses are presented in Table 2. A nomogram for predicting 1-, 3-, and 5-year survival was constructed based on the independent prognostic factors (Figure 2). As shown in the nomogram, the survival possibility of individual patients can be easily calculated by adding the scores of each variable.

#### Nomogram validation

The nomogram was validated internally using data from the training group and externally using data from the validation group. The C-indices of the nomogram in the training group and the validation group were 0.725 (95% confidence interval

[CI] 0.716–0.734) and 0.726 (95% CI 0.713–0.739), respectively. In this study, we also calculated the C-index for both groups using the AJCC 7th TNM classification system to compare the values with our nomogram. The results showed that the prognosis-predicting ability of the nomogram was significantly better than that of the AJCC 7th TNM classification system ( $P < 0.05$ ), as shown in Table 3. Figure 3 shows the ROC curves of the training group and the validation group. In the training group, the time-independent AUCs (tAUCs) of 1-, 3-, and 5-year OS were 0.807 (95% CI 0.783–0.831), 0.801 (95% CI 0.733–0.819), and 0.795 (95% CI 0.776–0.813), respectively. In the validation group, the tAUCs of 1-, 3-, and 5-year OS were 0.806 (95% CI 0.783–0.829), 0.797 (95% CI 0.779–0.814), and 0.790 (95% CI 0.773–0.806), respectively, which were all greater than those of the AJCC staging system. Bootstrapping with 1000 resamples in the training group exhibited a C-index of 0.723 (95% CI 0.710–0.736), which reflected consistent discrimination of the training group. The calibration curves of both the training and validation groups were close to the 45-degree line, demonstrating good consistency between the predictions and practical results. The calibration curves are shown in Figure 4.

## Discussion

In this study, we analyzed 10,846 US colorectal MAC patients from the SEER database who were diagnosed during 2010–2018. After randomly dividing the patients into the training group and the validation group in a 70:30 ratio, we established an effective nomogram to predict the 1-, 3-, and 5-year OS in the training group and then internally and externally validated the discrimination and calibration of the model. Additionally, we compared the predictive capacity of our nomogram to the TNM staging system by the C-index as well as ROC values, and both presented a significant increase ( $P < 0.001$ ), reflecting the superiority of this nomogram as a predictive tool for the prognosis of colorectal MAC patients.

MAC is a histological subtype of CRC with significant molecular differences in comparison to NMAC, for instance, overexpression of the mucin 2 and MUC5AC proteins and a high frequency of microsatellite instability (MSI-H). The relationship between molecular differences and clinical features is still undefined [4, 14, 15]. MAC has been reported to have worse clinical characteristics than NMAC, including a larger size and deeper invasion in primary lesions and higher rates of nodal metastasis and peritoneal metastasis. MAC also occurs more frequently in younger patients and females, located in the right colon and is less sensitive to radiotherapy and chemotherapy [5, 16, 17]. In our study, among all 10,846 included patients, 52% were female, the majority of the patients were initially diagnosed as having a cancer stage of T3 (56%) or T4 (29.2%), and more than half of the tumours (61.3%) occurred in the right colon, which was in agreement with the previous studies. Our study did not summarize the data of patients with peritoneal metastasis because of data deficiency. The reason why MAC tends to spread to the lymph nodes and peritoneum may be because of the pressure of mucus,

**Table 1. Demographics and pathological characteristics of enrolled colorectal mucinous adenocarcinoma patients**

<b>Variable</b>	<b>All patients, N (%)</b>	<b>Training group, N (%)</b>	<b>Validation group, N (%)</b>	<b>P value</b>
<i>Age (years)</i>				0.986
< 65	4688 (43.2)	3253 (43.2)	1435 (43.2)	
≥ 65	6158 (56.8)	4274 (56.8)	1884 (56.8)	
<i>Sex</i>				0.548
Male	5207 (48.0)	3628 (48.2)	1579 (47.6)	
Female	5639 (52.0)	3899 (51.8)	1740 (52.4)	
<i>Race</i>				0.904
White	8793 (81.1)	6106 (81.1)	2687 (81.0)	
Black	1144 (10.5)	796 (10.6)	348 (10.5)	
Other	909 (8.4)	625 (8.3)	284 (8.6)	
<i>Primary site</i>				0.958
Right-side colon	6645 (61.3)	4622 (61.4)	2023 (61.0)	
Transverse	894 (8.2)	617 (8.2)	277 (8.3)	
Left-side colon	2324 (21.4)	1604 (21.3)	720 (21.7)	
Rectum	983 (9.1)	684 (9.1)	299 (9.0)	
<i>Grade</i>				0.206
Well	1374 (12.7)	948 (12.6)	426 (12.8)	
Moderate	7171 (66.1)	5002 (66.5)	2169 (65.4)	
Poor	1832 (16.9)	1278 (17.0)	554 (16.7)	
Undifferentiated	469 (4.3)	299 (4.0)	170 (5.1)	
<i>Tumour size</i>				0.518
< 3 cm	1888 (17.4)	1322 (17.6)	566 (17.1)	
≥ 3 cm	8958 (82.6)	6205 (82.4)	2753 (82.9)	
<i>Lymph node dissection</i>				0.794
Less than 4	790 (7.3)	545 (7.2)	245 (7.4)	
4 or more	10056 (92.7)	6982 (92.8)	3074 (92.6)	
<i>T</i>				0.457
Tis-T1	414 (3.8)	275 (3.7)	139 (4.2)	
T2	1220 (11.2)	835 (11.1)	385 (11.6)	
T3	6048 (55.8)	4218 (56.0)	1830 (55.1)	
T4	3164 (29.2)	2199 (29.2)	965 (29.1)	
<i>N</i>				0.132
N0	5807 (53.5)	4016 (53.4)	1791 (54.0)	
N1	3076 (28.4)	2174 (28.9)	902 (27.2)	
N2	1963 (18.1)	1337 (17.8)	626 (18.8)	
<i>M</i>				0.806
M0	8962 (82.6)	6224 (82.7)	2738 (82.5)	
M1	1884 (17.4)	1303 (17.3)	581 (17.5)	
<i>Surgery</i>				0.532
No	133 (1.2)	89 (1.2)	44 (1.3)	
Yes	10713 (98.8)	7438 (98.8)	3275 (98.7)	
<i>Radiotherapy</i>				0.952
No/Unknown	9839 (90.7)	6829 (90.7)	3010 (90.7)	
Yes	1007 (9.3)	698 (9.3)	309 (9.3)	
<i>Chemotherapy</i>				0.797
No/Unknown	5983 (55.2)	4146 (55.1)	1837 (55.3)	
Yes	4863 (44.8)	3381 (44.9)	1482 (44.7)	
<i>Household income, \$</i>				0.958
< 75000	7777 (71.7)	5396 (71.7)	2381 (71.7)	
≥ 75000	3069 (28.3)	2131 (28.3)	938 (28.3)	

**Table 2. Univariate and multivariate analyses of overall survival for patients in the training group**

Variable	Univariate analysis		Multivariate analysis	
	HR (95% CI)	P value	HR (95% CI)	P value
<i>Age (years)</i>				
< 65	1 (Reference)	1	1 (Reference)	
≥ 65	1.596 (1.504–1.694)	< 0.001	1.884 (1.769–2.006)	< 0.001
<i>Sex</i>				
Male	1 (Reference)		–	NI
Female	1.007 (0.952–1.066)	0.803	–	
<i>Race</i>				
White	1 (Reference)		–	NI
Black	1.016 (0.927–1.114)	0.735	–	
Other	0.090 (0.819–1.015)	0.090	–	
<i>Primary site</i>				
Right-side colon	1 (Reference)		1 (Reference)	
Transverse	1.014 (0.914–1.125)	0.796	1.028 (0.926–1.141)	0.603
Left-side colon	1.109 (1.034–1.189)	0.004	1.164 (1.085–1.250)	< 0.001
Rectum	1.019 (0.923–1.125)	0.713	1.217 (1.094–1.355)	< 0.001
<i>Grade</i>				
Well	1 (Reference)		1 (Reference)	
Moderate	1.287 (1.168–1.419)	< 0.001	1.307 (1.183–1.445)	< 0.001
Poor	2.022 (1.814–2.254)	< 0.001	1.584 (1.415–1.774)	< 0.001
Undifferentiated	2.058 (1.770–2.394)	< 0.001	1.567 (1.342–1.829)	< 0.001
<i>Tumour size</i>				
< 3 cm	1 (Reference)		1 (Reference)	
≥ 3 cm	1.547 (1.424–1.681)	< 0.001	1.212 (1.110–1.324)	< 0.001
<i>Lymph node dissection</i>				
0–3	1 (Reference)		1 (Reference)	
4 or more	0.684 (0.620–0.756)	< 0.001	0.830 (0.737–0.936)	0.001
<i>T</i>				
Tis-T1	1 (Reference)		1 (Reference)	
T2	0.941 (0.768–1.153)	0.556	0.972 (0.790–1.195)	0.786
T3	1.482 (1.240–1.771)	< 0.001	1.219 (1.010–1.472)	0.039
T4	2.610 (2.246–3.219)	< 0.001	1.803 (1.489–2.183)	< 0.001
<i>N</i>				
N0	1 (Reference)	0	1 (Reference)	
N1	1.668 (1.559–1.784)	< 0.001	1.812 (1.682–1.952)	< 0.001
N2	2.778 (2.591–2.978)	< 0.001	2.664 (2.456–2.890)	< 0.001
<i>M</i>				
M0	1 (Reference)		1 (Reference)	
M1	3.451 (3.240–3.675)	< 0.001	3.159 (2.938–3.396)	< 0.001
<i>Surgery</i>				
No	1 (Reference)		1 (Reference)	
Yes	0.219 (0.182–0.264)	< 0.001	0.322 (0.258–0.402)	< 0.001
<i>Radiotherapy</i>				
No/Unknown	1 (Reference)		–	NI
Yes	1.057 (0.963–1.160)	0.242	–	
<i>Chemotherapy</i>				
No/Unknown	1 (Reference)	0	1 (Reference)	
Yes	0.702 (0.656–0.788)	< 0.001	0.588 (0.549–0.630)	< 0.001
<i>Household income, \$</i>				
< 75000	1 (Reference)		1 (Reference)	
≥ 75000	0.904 (0.848–0.964)	0.002	0.895 (0.839–0.955)	< 0.001

NI: Not included in the multivariate survival analysis; HR: Hazard ratio; CI: Confidence interval.

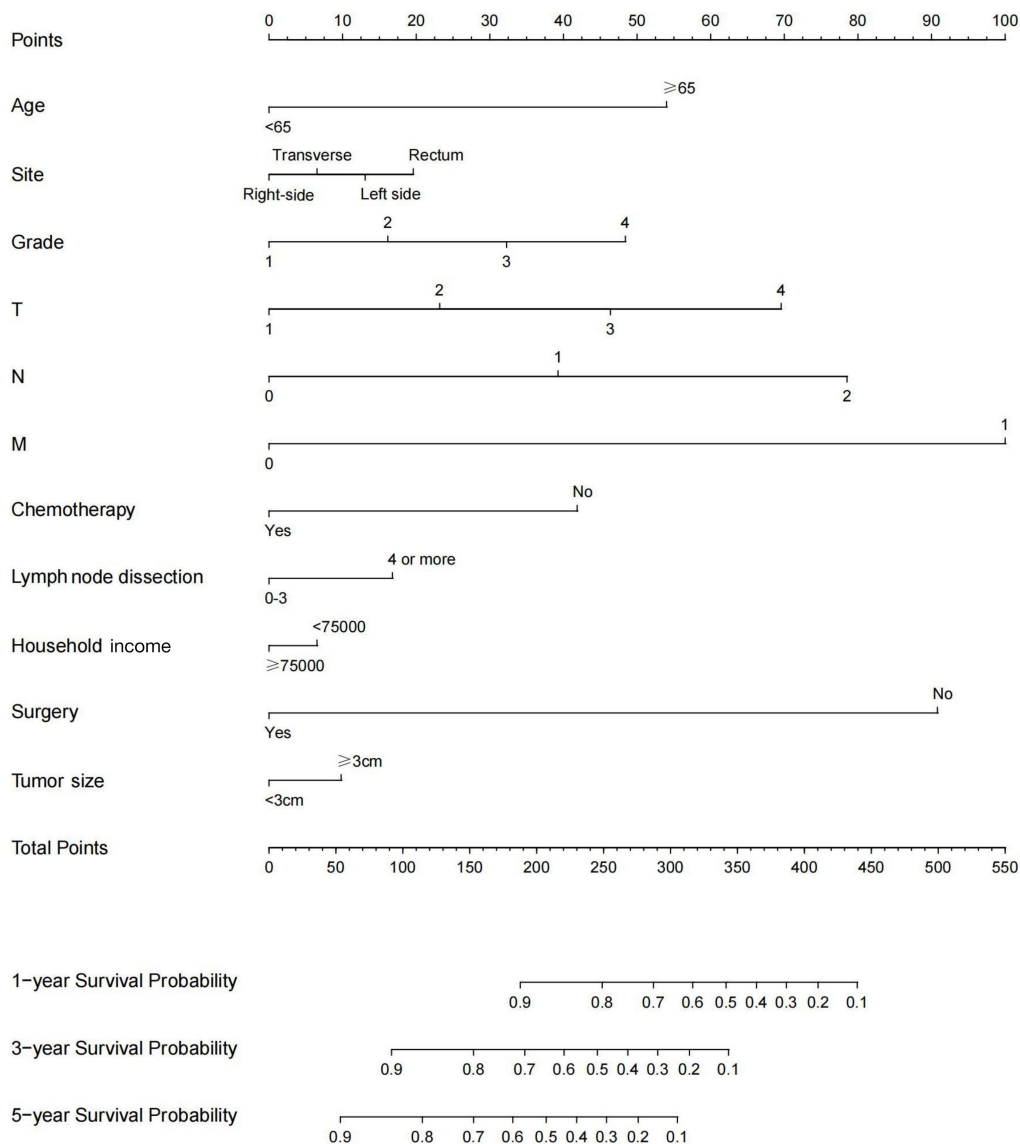


Figure 2. Nomogram model to predict the 1-, 3-, and 5-year overall survival of colorectal mucinous adenocarcinoma patients.

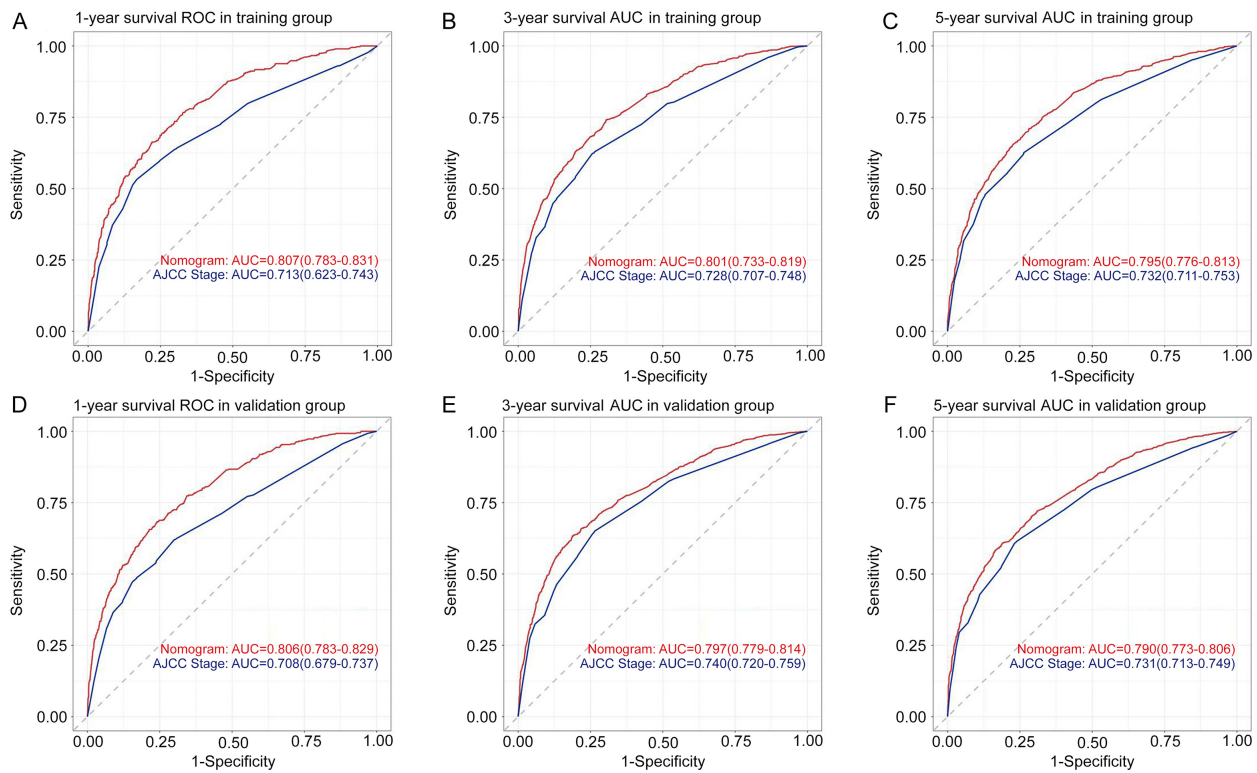
Table 3. C-index for the nomogram and TNM stage system in patients with colorectal mucinous adenocarcinoma

Classification	Training group		Validation group	
	C-index (95% CI)	P value	C-index (95% CI)	P value
Nomogram	0.725 (0.716–0.734)	1 (Reference)	0.726 (0.713–0.739)	1 (Reference)
AJCC 8 <sup>th</sup> stage	0.674 (0.664–0.684)	< 0.001	0.679 (0.664–0.696)	< 0.001

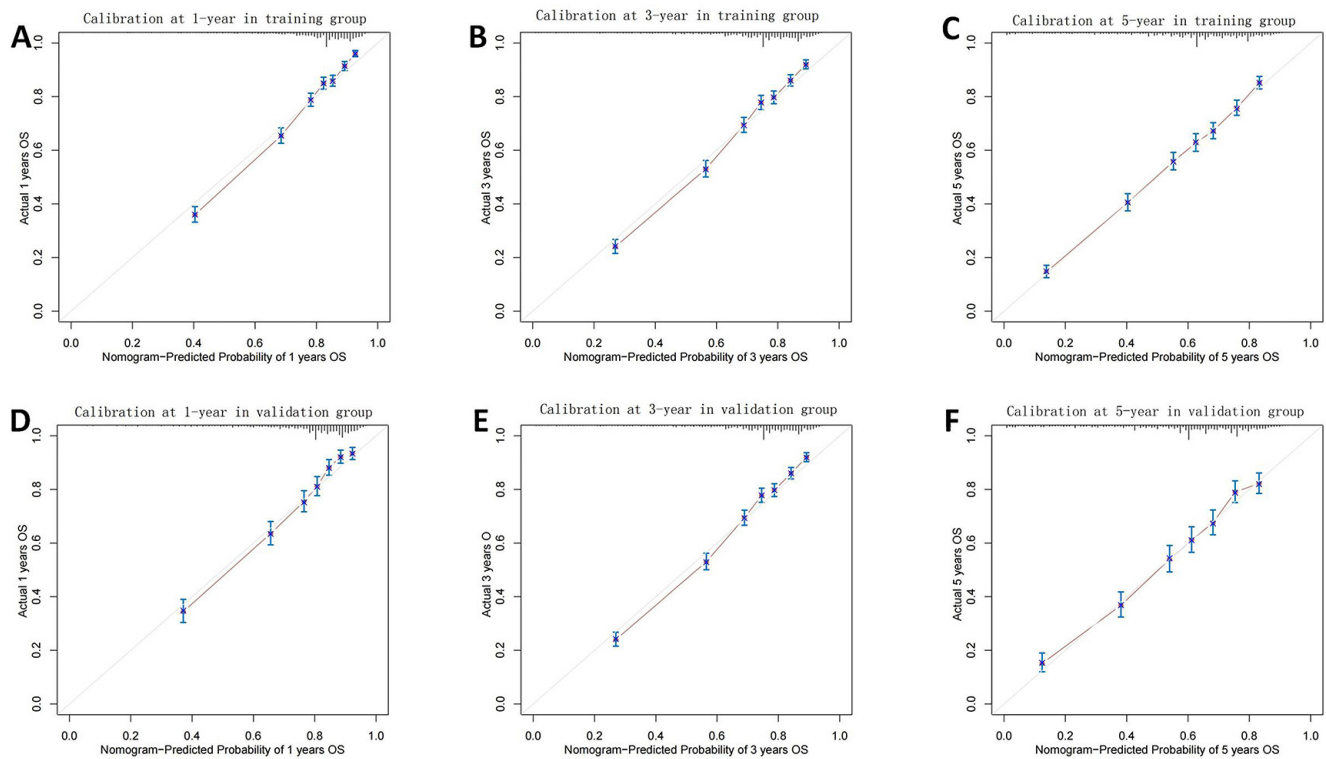
C-index: Concordance index; CI: Confidence interval; AJCC: American Joint Committee on Cancer.

which can be taken up by the lymphatic system and can spread into regional lymph nodes or push the MAC cells to the peritoneal cavity [5, 18]. The consensus is that the location of the primary tumour is significantly related to prognosis. The outcomes of patients with adenocarcinoma of the left colon were better than those of patients with adenocarcinoma of the right colon [19, 20]. In our study, the primary site was also an independent factor for prognosis, but only the comparison

between cancer of the left and right colon was significantly different. The hazard ratio results indicated that the prognosis of left colon cancer is worse than that of right colon cancer. There are two possible explanations for this condition. First, appendix MAC, which was assigned to the right colon in our study, is considered to have a good prognosis [21]. Second, studies have shown that the relationship between primary location and prognosis in MAC is different from that in NMAC.



**Figure 3.** ROC curves for predicting 1-, 3-, and 5-year overall survival in the training group (A)–(C) and validation group (D)–(F). ROC: Receiver operating characteristics; AUC: Area under the curve; AJCC: American Joint Committee on Cancer.



**Figure 4.** Calibration curves for predicting 1-, 3-, and 5-year OS in the training group (A)–(C) and validation group (D)–(F). A 45-degree plot represented an optimal model. Vertical bars indicate 95% confidence interval. AUC: Area under the curve; AJCC: American Joint Committee on Cancer; OS: Overall survival.

One study aimed to determine the prognosis and molecular differences in MAC and NMAC in CRC and demonstrated that right colon MAC was associated with more MSI-H tumours and a similar 5-year OS rate compared with NMAC. On the other hand, the left colon and rectal MAC were related to a worse 5-year OS rate [22]. Another study retrospectively analyzed 244,794 patients from the National Cancer Database and concluded that MAC of the rectum is associated with poorer survival [23].

The TNM stages are the most important and generally acknowledged prognostic factors for patients with malignant solid tumours. Our study indicated a consistent result: TNM stages, especially the M stage, were the most valuable variables among all factors. There was another pivotal variable, surgery, which remarkably impacted the prognosis according to our nomogram. Additionally, four or more regional lymph node dissections during surgery were found to be a beneficial factor for colorectal MAC patients. These results have been proven in numerous studies and illustrate the importance of standard surgical treatment, which is the only chance for a radical cure or long-term survival [10, 24].

Whether MAC is related to poor outcomes in CRC patients is controversial. Unlike signet ring cell CRC, MAC is not identified as an independent factor for worse prognosis by the AJCC, and neither the National Comprehensive Cancer Network (NCCN) nor the European Society for Medical Oncology (ESMO) guidelines suggest disparate standard treatments particularly designed for MAC [25, 26]. Several studies indicated that no significant differences were shown between MAC and NMAC survival [27, 28]. Nevertheless, most of these studies had insufficient sample sizes, and more studies with large scales have demonstrated that MAC is an independent significant factor compared with NMAC, especially in the specific groups. Numata et al. [16] identified that MAC was associated with worse survival than NMAC in patients with stages III and IV disease. In a study including 6475 patients with stages I to III CRC, MAC was not an independent prognostic factor of disease-free survival in the entire cohort but was a significant factor in the colon subgroup analysis ( $P = 0.026$ ) [5]. A meta-analysis including 44 studies also showed worse survival in colorectal MAC vs adenocarcinoma patients [29].

Despite the ambiguity in MAC prognosis, a consensus has been reached regarding the impaired response to chemotherapy. It has been observed that compared with NMAC, MAC is less responsive to neoadjuvant chemotherapy, adjuvant chemotherapy, and palliative chemotherapy, expressed by lower disease-free survival rates and survival rates [4, 30, 31]. Although our nomogram verified chemotherapy as an independent factor of prognosis, patients who underwent chemotherapy had better survival. Under these circumstances, it is in close agreement that patients with MAC could require more consideration during follow-up or even intensified adjuvant therapy.

There was a nomogram for the prognosis of colorectal MAC patients constructed by Lian et al. Compared to our nomogram study, this previous study was mainly focused on the pattern of distant metastases, and external validation was absent [32]. Our

study was larger in scale and more comprehensive. We included 11 independent factors that were selected by both univariate analysis and multivariate backward stepwise Cox proportional hazard regression analysis. After excluding multicollinearity problems, a nomogram was successfully constructed. Moreover, both the C-index and ROC results demonstrated a significant advantage in prognosis compared with TNM stage classification. Additionally, the employment of our nomogram is convenient and practical in clinical work. The prognosis of each patient can be simply estimated by adding the scores of each factor and finding the corresponding possibility of the total score.

Certain limitations existed in this research. First, CRC is a heterogeneous disease, and the prognosis of CRC is affected by many factors. In addition to the characteristics mentioned in our study, clinical information was not reported or was reported incompletely in the SEER database, for instance, gene status, such as *KRAS*, *NRAS*, *BRAF* mutations, and MSI status, details of treatments, such as regimens of chemotherapy, target therapy, and immunotherapy, family history of disease, and comprehensive pathology descriptions, which may influence the effectiveness of our nomogram [33]. Further studies including these aspects are needed. Second, since this was a retrospective study and patients with incomplete data were excluded according to the study design, selection bias was inevitable. To avoid this bias, a prospective randomized controlled study should be performed. Third, data for external verification were also extracted from the SEER database, which may be less convincing. All these conditions should be considered during the application of our nomogram.

## Conclusion

In conclusion, this study constructed a practical and user-friendly nomogram based on the SEER database to predict 1-, 3-, and 5-year OS in patients with colorectal MAC. The nomogram was validated both internally and externally and demonstrated a significant advantage in predictive accuracy compared to TNM stage classification. Due to the retrospective design and the absence of more clinical and gene information, prospective studies are required in the future.

**Conflicts of interest:** Authors declare no conflicts of interest.

**Funding:** Authors received no specific funding for this work.

**Data availability:** The dataset from the SEER database generated and/or analyzed during the current study is available in the SEER dataset repository (<https://seer.cancer.gov/>).

Submitted: 09 October 2022

Accepted: 08 November 2022

Published online: 10 November 2022

## References

- [1] Baidoun F, Elshiyw K, Elkeraie Y, Merjaneh Z, Khoudari G, Sarmini MT, et al. Colorectal cancer epidemiology: recent trends and impact on outcomes. *Curr Drug Targets* 2021;22(9):998–1009. <https://doi.org/10.2174/1389450121999201117115717>.



- [2] Nagtegaal ID, Odze RD, Klimstra D, Paradis V, Rugge M, Schirmacher P, et al. The 2019 WHO classification of tumours of the digestive system. *Histopathology* 2020 Jan;76(2):182–8. <https://doi.org/10.1111/his.13975>.
- [3] Debunne H, Ceelen W. Mucinous differentiation in colorectal cancer: molecular, histological and clinical aspects. *Acta Chir Belg* 2013 Nov-Dec;113(6):385–90. <https://doi.org/10.1080/00015458.2013.11680951>.
- [4] Luo C, Cen S, Ding G, Wu W. Mucinous colorectal adenocarcinoma: clinical pathology and treatment options. *Cancer Commun (Lond)* 2019 Mar 29;39(1):1–13. <https://doi.org/10.1186/s40880-019-0361-0>.
- [5] Park JS, Huh JW, Park YA, Cho YB, Yun SH, Kim HC, et al. Prognostic comparison between mucinous and nonmucinous adenocarcinoma in colorectal cancer. *Medicine (Baltimore)* 2015 Apr;94(15):e658. <https://doi.org/10.1097/MD.0000000000000658>.
- [6] Mahar AL, Compton C, Halabi S, Hess KR, Weiser MR, Groome PA. Personalizing prognosis in colorectal cancer: a systematic review of the quality and nature of clinical prognostic tools for survival outcomes. *J Surg Oncol* 2017 Dec;116(8):969–82. <https://doi.org/10.1002/jso.24774>.
- [7] Gunderson LL, Jessup JM, Sargent DJ, Greene FL, Stewart AK. Revised TN categorization for colon cancer based on national survival outcomes data. *J Clin Oncol* 2010;28:264–71. <https://doi.org/10.1200/JCO.2009.24.0952>.
- [8] Shia J, Klimstra DS, Bagci P, Basturk O, Adsay NV. TNM staging of colorectal carcinoma: issues and caveats. *Semin Diagn Pathol* 2012 Aug;29(3):142–53. <https://doi.org/10.1053/j.semdp.2012.02.001>.
- [9] Yamano T, Yamauchi S, Igeta M, Takenaka Y, Song J, Kimura K, et al. Combination of preoperative tumour markers and lymphovascular invasion with TNM staging as a cost and labour efficient subtyping of colorectal cancer. *Sci Rep* 2020 Jun 24;10(1):10238. <https://doi.org/10.1038/s41598-020-66652-z>.
- [10] Haraldsdottir S, Einarsdottir HM, Smaradottir A, Gunnlaugsson A, Halfdanarson TR. Colorectal cancer—review. *Laeknabladid* 2014 Feb;100(2):75–82. <https://doi.org/10.17992/lbl.2014.02.531>.
- [11] Luo X-J, Zhao Q, Liu J, Zheng J-B, Qiu M-Z, Ju H.-Q, et al. Novel genetic and epigenetic biomarkers of prognostic and predictive significance in stage II/III colorectal cancer. *Mol Ther* 2021 Feb 3;29(2):587–96. <https://doi.org/10.1016/j.yimthe.2020.12.017>.
- [12] Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol* 2008 Mar 10;26(8):1364–70. <https://doi.org/10.1200/JCO.2007.12.9791>.
- [13] Doll KM, Rademaker A, Sosa JA. Practical guide to surgical data sets: surveillance, epidemiology, and end results (SEER) database. *JAMA Surg* 2018 Jun 1;153(6):588–9. <https://doi.org/10.1001/jamasurg.2018.0501>.
- [14] Imai Y, Yamagishi H, Fukuda K, Ono Y, Inoue T, Ueda Y. Differential mucin phenotypes and their significance in a variation of colorectal carcinoma. *World J Gastroenterol* 2013;19(25):3957–68. <https://doi.org/10.3748/wjg.v19.i25.3957>.
- [15] Huang A, Yang Y, Shi J-Y, Li Y-K, Xu J.-X, Cheng Y, et al. Mucinous adenocarcinoma: a unique clinicopathological subtype in colorectal cancer. *World J Gastrointest Surg* 2021 Dec 27;13(12):1567–83. <https://doi.org/10.4240/wjgs.v13.i12.1567>.
- [16] Numata M, Shiozawa M, Watanabe T, Tamagawa H, Yamamoto N, Morinaga S, et al. The clinicopathological features of colorectal mucinous adenocarcinoma and a therapeutic strategy for the disease. *World J Surg Oncol* 2012 Jun 15;10:109. <https://doi.org/10.1186/1477-7819-10-109>.
- [17] Nozoe T, Anai H, Nasu S, Nasu S, Sugimachi K. Clinicopathological characteristics of mucinous carcinoma of the colon and rectum. *J Surg Oncol* 2000;75:103–7. [https://doi.org/10.1002/1096-9098\(200010\)75:2<103::AID-JSO6>protect\\$relax>\\$3.0.CO;2-C](https://doi.org/10.1002/1096-9098(200010)75:2<103::AID-JSO6>protect$relax>$3.0.CO;2-C).
- [18] Sugarbaker PH. Mucinous colorectal carcinoma. *J Surg Oncol* 2001;77:282–3. <https://doi.org/10.1002/jso.1111>.
- [19] Stintzing S, Tejpar S, Gibbs P, Thiebach L, Lenz H-J. Understanding the role of primary tumour localisation in colorectal cancer treatment and outcomes. *Eur J Cancer* 2017 Oct;84:69–80. <https://doi.org/10.1016/j.ejca.2017.07.016>.
- [20] Loree JM, Pereira AAL, Lam M, Willauer AN, Raghav K, Dasari A, et al. Classifying colorectal cancer by tumor location rather than sidedness highlights a continuum in mutation profiles and consensus molecular subtypes. *Clin Cancer Res* 2018 Mar 1;24(5):1062–72. <https://doi.org/10.1158/1078-0432.CCR-17-2484>.
- [21] Shaib WL, Assi R, Shamseddine A, Alese OB, Staley C, Memis B, et al. Appendiceal mucinous neoplasms: diagnosis and management. *Oncologist* 2017 Sep;22(9):1107–16. <https://doi.org/10.1634/theoncologist.2017-0081>.
- [22] Lan Y-T, Chang S-C, Lin P-C, Lin C-C, Lin H-H, Huang S-C, et al. Clinicopathological and molecular features of colorectal cancer patients with mucinous and non-mucinous adenocarcinoma. *Front Oncol* 2021 Mar 2;11:620146. <https://doi.org/10.3389/fonc.2021.620146>.
- [23] Hyngstrom JR, Hu C-Y, Xing Y, You YN, Feig BW, Skibber JM, et al. Clinicopathology and outcomes for mucinous and signet ring colorectal adenocarcinoma: analysis from the national cancer data base. *Ann Surg Oncol* 2012 Sep;19(9):2814–21. <https://doi.org/10.1245/s10434-012-2321-7>.
- [24] Akgül Ö, Çetinkaya E, Ersöz Ş, Tez M. Role of surgery in colorectal cancer liver metastases. *World J Gastroenterol* 2014 May 28;20(20):6113–22. <https://doi.org/10.3748/wjg.v20.i20.6113>.
- [25] Benson AB, Venook AP, Al-Hawary MM, Arain MA, Chen Y-J, Ciombor KK, et al. Colon cancer, version 2.2021, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw* 2021 Mar 2;19(3):329–59. <https://doi.org/10.6004/jnccn.2021.0012>.
- [26] De Falco V, Napolitano S, Roselló S, Huerta M, Cervantes A, Ciardiello F, et al. How we treat metastatic colorectal cancer. *ESMO Open* 2020 Aug;4(Suppl 2):e000813. <https://doi.org/10.1136/esmoopen-2020-000813>.
- [27] Nitsche U, Friess H, Agha A, Angele M, Eckel R, Heitland W, et al. Prognosis of mucinous and signet-ring cell colorectal cancer in a population-based cohort. *J Cancer Res Clin Oncol* 2016 Nov;142(11):2357–66. <https://doi.org/10.1007/s00432-016-2224-2>.
- [28] Jia X, Li B, Wang H, Yan Z. Clinical features, molecular alterations and prognosis of colorectal adenocarcinoma with mucinous component in Chinese patients. *Appl Immunohistochem Mol Morphol* 2021 Nov-Dec 01;29(10):765–72. <https://doi.org/10.1097/PAI.0000000000000950>.
- [29] Verhulst J, Ferdinand L, Demetter P, Ceelen W. Mucinous subtype as prognostic factor in colorectal cancer: a systematic review and meta-analysis. *J Clin Pathol* 2012 May;65(5):381–8. <https://doi.org/10.1136/jclinpath-2011-200340>.
- [30] Reynolds IS, O'Connell E, Fichtner M, McNamara DA, Kay EW, Prehn JHM, et al. Mucinous adenocarcinoma is a pharmacogenomically distinct subtype of colorectal cancer. *Pharmacogenomics J* 2020 Jun;20(3):524–32. <https://doi.org/10.1038/s41397-019-0137-6>.
- [31] Ott C, Gerken M, Hirsch D, Fichtner-Feigl S, Munker S, Schnoy E, et al. Advanced mucinous colorectal cancer: epidemiology, prognosis and efficacy of chemotherapeutic treatment. *Digestion* 2018;98(3):143–52. <https://doi.org/10.1159/000487710>.
- [32] Lian L, Xu X-F, Shen X-M, Huang T-A, Li X-M, Han S.-G, et al. Pattern of distant metastases and predictive nomograms in colorectal mucinous adenocarcinoma: a SEER analysis. *J Gastrointest Oncol* 2021 Dec;12(6):2906–18. <https://doi.org/10.21037/jgo-21-824>.
- [33] Sagaert X, Vanstapel A, Verbeek S. Tumor heterogeneity in colorectal cancer: what do we know so far? *Pathobiology* 2018;85(1–2):72–84. <https://doi.org/10.1159/000486721>.

## Related articles published in BJBMS

1. Construction and validation of prognostic nomogram for metaplastic breast cancer  
Yongfeng Li et al., BJBMS, 2022
2. Trends of incidence and prognosis of upper tract urothelial carcinoma  
Ming Chen et al., BJBMS, 2020
3. The prognostic significance of different proportion of signet-ring cells of colorectal carcinoma  
Wei Chen et al., BJBMS, 2020
4. Development and validation of nomograms for predicting survival of elderly patients with stage I small-cell lung cancer  
Yaji Yang et al., BJBMS, 2020

## Supplemental Data

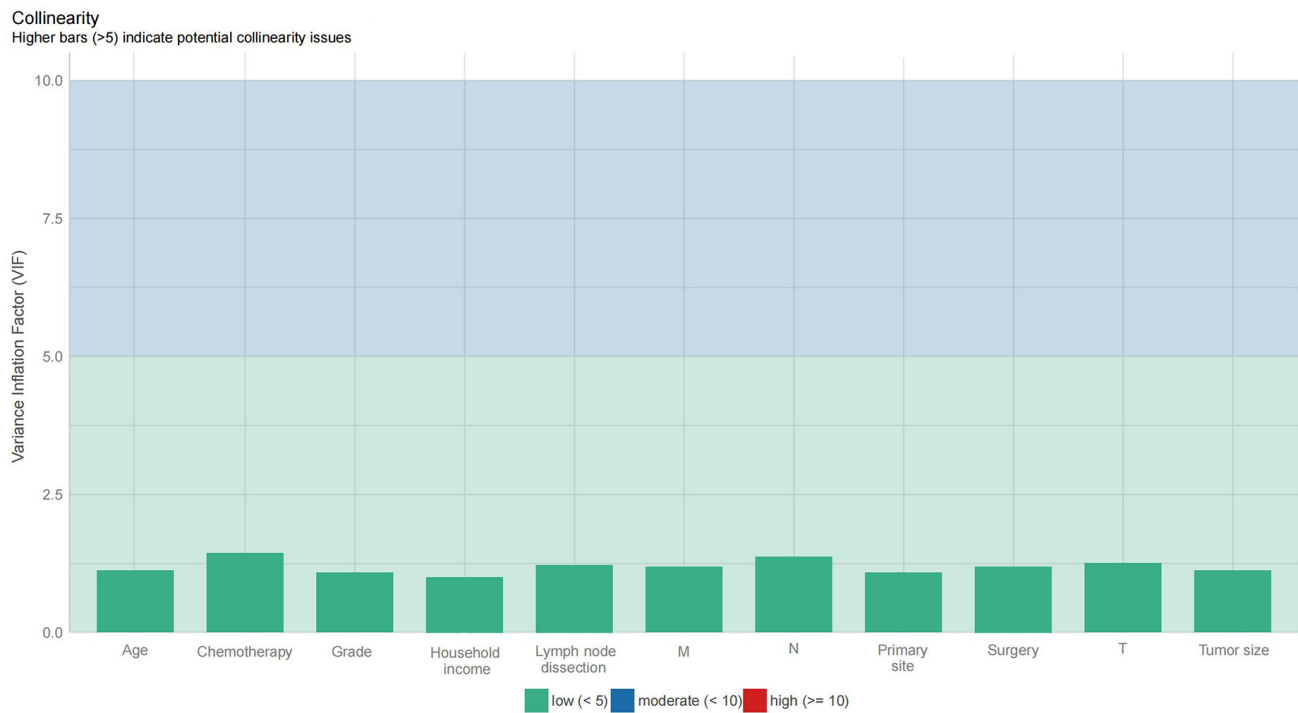


Figure S1. Variance inflation factor examination of included prognostic factors.