Heliyon 10 (2024) e35720

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Clinical and pathological characteristics of and predictive model for colorectal neuroendocrine tumors

Jiuyue Ma, Xiaoqian Ma, Jie Xing, Ruyun Song, Yang Zhang, Mo Liu, Shuilong Guo, Qian Zhang *, Jing Wu **

Department of Gastroenterology, Beijing Friendship Hospital, Capital Medical University, National Clinical Research Center for Digestive Disease, Beijing Digestive Disease Center, Beijing Key Laboratory for Precancerous Lesions of Digestive Disease, Beijing, 100050, China

ARTICLE INFO

Keywords:
Colorectal neuroendocrine tumor
Prognostic model
Nomogram

ABSTRACT

Background: The incidence of colorectal neuroendocrine tumors (NETs) is increasing, causing a social burden. At present, there is no specific prognostic model for colorectal NETs. Thus, an accurate model is needed to predict the prognosis of patients with colorectal NETs.

Aim: We aimed to create a new nomogram to predict the prognosis of patients with colorectal NETs. Furthermore, we compared nomogram we established and the 8th edition of the AJCC TNM staging system in terms of prediction ability and accuracy.

Methods: A total of 3353 patients with colorectal NETs were selected from the Surveillance, Epidemiology, and End Results (SEER) database. Kaplan-Meier analyses were used to assess overall survival (OS) and cancer-specific survival (CSS). Additionally, LASSO regression was used to select variables for constructing the nomogram. Furthermore, the C-index and time-dependent receiver operating characteristic (tdROC) curve were used to evaluate the nomogram. Decision curve analysis (DCA) was performed to compare the clinical utility of the nomogram with that of the TNM system. An external validation cohort (N = 61) was established to evaluate the nomogram's prediction accuracy.

Results: A total of 9 factors (age, sex, marital status, tumor size, T stage, M stage, N stage, grade, and surgery) were selected based on the results of LASSO analysis. The C-indexes of the nomogram in the training and validation sets were 0.807 and 0.775, respectively, which indicated that the nomogram had better prediction accuracy than TNM staging (C-index = 0.700 in the training set and 0.652 in the validation set). The C-index of the nomogram in the external validation cohort was 0.954, indicating that the nomogram had satisfactory prediction accuracy. The results of DCA revealed that the survival nomogram possessed greater utility in clinical practice.

Conclusion: We determined the OS and CSS of patients with colorectal NETs and developed a robust and clinically useful survival nomogram.

* Corresponding author. Department of Gastroenterology, Beijing Friendship Hospital, Capital Medical University, National Clinical Research Center for Digestive Disease, Beijing Digestive Disease Center, Beijing Key Laboratory for Precancerous Lesions of Digestive Disease, Beijing, 100050, China.

** Corresponding author. Department of Gastroenterology, Beijing Friendship Hospital, Capital Medical University, National Clinical Research Center for Digestive Disease, Beijing Digestive Disease Center, Beijing Key Laboratory for Precancerous Lesions of Digestive Disease, Beijing, 100050, China.

E-mail addresses: zhangqian200104@ccmu.edu.cn (Q. Zhang), wujing36youyi@ccmu.edu.cn (J. Wu).

https://doi.org/10.1016/j.heliyon.2024.e35720
Received 2 June 2023; Received in revised form 26 July 2024; Accepted 2 August 2024
Available online 2 August 2024

1. Introduction

Colorectal neuroendocrine tumors (NETs) are rare colorectal neoplasms that originate from neuroendocrine cells and can be malignant. Some NETs exhibit neuroendocrine functions, while others do not. The incidence of NETs in the rectum is higher than that in the colon [1], and the total incidence and prevalence of NETs have increased in recent years [2,3]. A population-based study also revealed that the age-adjusted incidence rate of NETs increased 6.4-fold from 1973 (1.09 per 100 000) to 2012 (6.98 per 100 000) [4]. Furthermore, an additional study showed that an increase in NETs is associated with an increase in the risk of early-onset colorectal cancer, suggesting that the threat of NETs may be underestimated and that they may cause a social burden [5].

The Surveillance, Epidemiology, and End Results Program (SEER) is an extensive population-based cancer database constructed by the National Cancer Institute. The SEER database contains 22 registration points in the U.S. and more than 10 million data records. Clinical and pathological data such as age, sex, race, tumor pathological type, and histological grade can be obtained from the SEER. stat software. The SEER database also contains survival data, including patient outcomes, survival time, and reasons for death, which make this database very suitable for survival analysis and the establishment of prognostic models. A nomogram integrates multiple prediction indicators and expresses the relationships between variables in a prediction model. Nomograms also transform complex regression equations into a visual graph, thus making prediction models more readable and convenient in terms of evaluating patients. Due to these advantages, we used the SEER database to construct a predictive nomogram model.

The 8th American Joint Cancer Committee (AJCC) staging system defines colorectal NETs with clinicopathological features such as T, N, and M. Piqing Gong et al. revealed that the 8th AJCC staging system is impractical for determining the survival outcomes of patients with colorectal NETs, especially patients with stage II and III disease [6]. Yu Zhang et al. reported that the c-indexes of AJCC staging for predicting survival among NET and NEC patients in their SEER cohort were 0.722 and 0.673, respectively, and there was no significant difference between the AJCC and the European Neuroendocrine Tumor Society (ENETS) staging systems [7]. Thus, more optimized staging methods for NETs should be developed.

Some rare models that aim to predict the prognosis of patients with colorectal NETs have limitations. Zihan Xu et al. reported a nomogram for predicting the prognosis of patients with gastroenteropancreatic NETs (GEP-NETs), but they did not consider the updated WHO classification of colorectal NETs. Studies have also shown that G3 NETs and NECs have different clinical behaviours and survival outcomes [8]. Hence, the WHO updated the classification system in 2019 and divided NECs and NETs based on pathological type. NETs are divided into G1, G2, and G3 based on Ki-67, and G3 NETs are no longer classified as NECs [9,10]. Using data collected between 1973 and 2004, Jinmao Zou et al. established a nomogram based SEERdatabase; this nomogram possessed better predictive ability than the 7th edition of the AJCC staging system. However, their model had a similar problem: they included “neuroendocrine carcinoma” as an NET in the sample [1]. However, since the SEER database and AJCC TNM system have been updated, their nomogram is unsuitable for clinical use. No current prediction model for colorectal NET meets the new standards of the WHO.

The clinicopathological features and prognostic factors of colorectal NENs include age, pathological grade, tumor size, primary tumor location, marital status, surgery, and liver metastasis [11–13]. Considering that the epidemiological characteristics and classification of colorectal NETs have changed, no predictive model has been specifically designed for colorectal NETs since the SEER

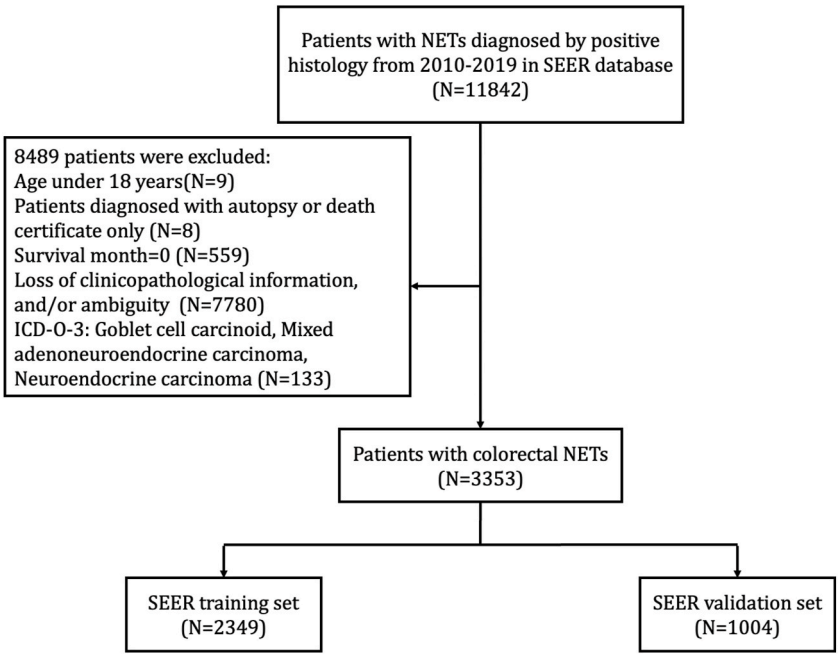

Fig. 1. Flow chart for selecting patients from the SEER database.

database has been updated. Thus, new prediction models should be developed.

2. Materials and methods

2.1. Study population

Colorectal NET patients from the SEER database (https://seer. cancer.gov) between 2010 and 2019 were initially screened. All medical records were obtained solely by SEER*Stat software (version 8.4.0.1; https://seer.cancer.gov/seerstat/). The inclusion criteria were as follows: 1) diagnosed with NETs according to the AYA site Recode 2020 Revision; and 2) primary tumor site was limited to the colon and rectum. The exclusion criteria were as follows: 1) age under 18 years, 2) diagnosed via autopsy or death certificate only, 3) survival month = 0, 4) loss of vital clinical information and/or information ambiguity, and 4) International Classification of Disease-O-3 (ICD-O-3) codes: Goblet cell carcinoid, mixed adenoneuroendocrine carcinoma, or neuroendocrine carcinoma. A total of 3353 patients were ultimately included in the study (Fig. 1). A total of 61 patients who had medical records in our hospital and who met the inclusion and exclusion criteria were enrolled in the external validation cohort. The validation cohort study was approved by the Bioethics Committee of Beijing Friendship Hospital, Capital Medical University (2020-P2-290-01).

2.2. Data Description

All demographic and clinical data from the training and validation sets were obtained from the SEER database, and essential data from the external validation cohort were obtained from medical records and follow-up examinations. This version of the SEER database was released in November 2021. Previous studies have shown that age, primary tumor site, tumor size, marital status, distant metastasis status, TNM stage, and race could be potential risk factors for survival [1,5,7,11,12]. For colorectal NETs, the tumors located in the caecum, ascending colon, hepatic flexure of the colon, and transverse colon were divided into right-sided or proximal colon tumors; the tumors located in the splenic flexure of the colon, descending colon, sigmoid colon, and rectosigmoid junction were divided into left-sided or distal colon tumors. We used X-tile software based on risk stratification to convert continuous variables such as age and tumor size into classification variables. Age was divided into 18–59 years, 60–70 years, and >70 years, and tumor size was divided into ≤30 mm and >30 mm. Although only 5 % of the patients were widowed, one study showed that being widowed was a risk factor for NET patient survival, while any other marital status was not [14]. Therefore, we classified marital status as widowed or other (divorced, unmarried, single, separated, or married). Since there are no Ki-67 data in the SEER database, we could not obtain grade data using the latest WHO NETs grading system. According to other published studies on NETs, the histopathological grade was recorded as grades I, II, and III, representing well differentiated, moderately differentiated, and poorly differentiated/undifferentiated, respectively [8]. Race was divided into White and other races. Distant metastases included bone and liver metastases. Since most patients with pathological findings of carcinoid tumors have the best prognosis, the pathology was divided into carcinoid tumor and other (including enterochromaffin cell carcinoid, adenocarcinoid tumor, and atypical carcinoid tumor). Although “carcinoid tumor” is now replaced by “neuroendocrine tumor”, the SEER database still includes carcinoid tumor as one of the pathological types. Because this was a retrospective study based on the SEER database, we still used the expression of carcinoid tumors in our study. The TNM stages were restaged against the 7th and 8th AJCC TNM editions. In the 7th AJCC TNM staging guidance, colorectal NETs use the staging criteria of gastroenteropancreatic neuroendocrine tumors (GEP-NETs); however, in the 8th edition of the AJCC TNM guidelines, the staging criteria for colorectal NETs alone are set. Moreover, stages IV B and IV C have been added to the 8th edition of the AJCC staging guidelines. The major difference between stages IV B and IV C was only M1b and M1c, and we applied this to the lack of M1c data in the data section. Stages IV B and IV C were combined into IV B + C [6]. Furthermore, we wanted to determine the impact of sex, economic status, residence in urban and rural areas, chemotherapy, systemic treatment, and summary stage on survival.

2.3. Outcome ascertainment

The primary outcome of interest was overall survival (OS). The secondary outcome of interest was patients with NET-related mortality, which can indicate cancer-specific survival (CSS).

2.4. Statistical analysis

The categorical variables are represented as numbers with percentages (n%) and were compared using the chi-square test. We used the Kaplan–Meier method to plot the survival curves of OS characteristics and CSS. All the colorectal NET patients who met the inclusion and exclusion criteria from the SEER database were randomly divided into training set and validation set at a ratio of 7:3 using R software. Specifically, the training set included more data to establish the nomogram model, and the validation set was used to verify the accuracy of the nomogram. Least absolute shrinkage and selection operator (LASSO) Cox regression analyses were used to screen statistically significant features associated with predicting OS in the training set among multiple variables. We used the R package glmnet to perform LASSO analysis in this section. The variables filtered by LASSO regression were incorporated into the multivariate Cox survival nomogram. All independent risk factors were assigned scores. The calculated total score predicted 1-year, 3-year, and 5-year OS. The C-index and calibration curves were applied to evaluate the differentiation and calibration ability of the survival nomogram. We also performed time-dependent receiver operating characteristic (tdROC) curve analysis to assess the model’s predictive accuracy and compared the accuracy between the nomogram and the 8th edition of the AJCC TNM. Additionally, decision

curve analysis (DCA) curves were used to evaluate the clinical utility of the nomogram. We used the R packages timeROC and dcurves to perform tdROC and DCA analyses. The reliability and accuracy of the nomogram were verified in both the SEER validation set and the validation cohort. Furthermore, we compared the clinical application value between the 8th edition of the AJCC staging system and our nomogram by using DCA.

All the statistical analyses were conducted using R statistical software, version 4.2.0. A two-sided P value < 0.05 was considered to indicate statistical significance. X-tile software was used to divide continuous variables into classification variables.

3. Results

3.1. Patients' baseline characteristics

A total of 3353 colorectal NETs patients in the SEER database who met the inclusion criteria were enrolled in the study. They were randomly divided into a training set (N = 2349) and a validation set (N = 1004). The overall median age of the patients was 55 (interquartile range: 50, 63) years. Whites and males accounted for 60 % and 40 % of the sample, respectively. Only 5 % of patients were widowed. A total of 60 % of patients lived in metropolitan areas with a population greater than 1 million, and 66 % had a median household income of ≤ $75,000. The tumor size of most patients was ≤30 cm (91 %). Most of the tumors were in the rectum (77 %), and others were in the distal (7 %) or proximal colon (16 %). The vast majority of patients underwent surgery (99 %). Two percent of patients received chemotherapy, and 4 % had liver metastasis. Eight percent of patients died (N = 254), 3 % of whom died due to NETs (N = 101). The median survival time was 45 (interquartile range: 21, 69) months. Table 1 shows other clinical and pathological features. Furthermore, there was no significant difference in any of the variables between the training and validation sets.

For the validation cohort, 61 patients were enrolled based on the inclusion and exclusion criteria, among whom 4 patients died, and the median survival time was 33 (interquartile range: 23, 53) months. Male patients accounted for 41 % of the cohort, and the majority of patients underwent surgery (98 %). Consistent with the SEER set, the tumor size of most patients was ≤30 cm (95 %). Most of the tumors were located in the rectum (93 %), and others were in the distal (5 %) or proximal colon (2 %) (Supplementary Table 1).

3.2. OS and CSS of colorectal NET patients in the SEER database

Fig. 2(A–H) show the 5-year OS curves. Among all patients, age was an important indicator of OS. Patients aged 18–59, 60–70, and >70 years had 5-year OS rates of 94.5 %, 87.8 %, and 70.9 %, respectively. Widowed patients had a lower OS rate than did patients with other marital statuses (79.9 % vs. 91.6 %). The 5-year OS rates of whites and other races were 90.2 % and 92.0 %, respectively (p = 0.13). Patients with a tumor size >30 mm had worse OS (74.1 % vs. 92.7 %, respectively), and those with a proximal colon tumor had a worse 5-year OS than those with a distal colon tumor or rectum tumor (81.2 % vs. 89.8 %/93.1 %, respectively). Pathology grade was also a significant indicator of OS (grade I/II/III, 93.0 %/84.5 %/46.1 %). Surgery significantly improved the OS rate (91.5 % vs. 27.8 %). The OS rates of M0 and M1 patients were 92.8 % and 54.6 %, respectively. According to the Kaplan–Meier analysis, sex was not a significant factor for survival (p = 0.24).

Age was also a significant indicator of CSS. Patients aged 18–59, 60–70, and >70 years had 5-year CSS rates of 97.9 %, 95.0 %, and 87.6 %, respectively. The CSS rate of widowed patients was 93.0 %, while that of patients with other marital statuses was 96.4 %. Tumor size, tumor location, and pathological grade were important factors that influenced the CSS rate (>30 mm/≤30 mm, 82.4 %/97.6 %; rectum/distal colon/proximal colon, 97.8 %/94.8 %/88.9 %; grade I/II/III, 98.0 %/91.1 %/47.2 %) (Fig. 3(A–H)). The worst CSS rate was observed in TNM stage 4 patients (60.4 %) as opposed to stage 1/2/3 patients (99.2 %/94.8 %/91.8 %). The CSS rates of M0 and M1 patients were 98.0 % and 60.4 %, respectively. White patients had a CSS rate of 95.4 %, and patients of other races (Black, Asian, etc.) had a CSS rate of 97.3 % (p = 0.082). Sex did not significantly affect the CSS rate (p = 0.14), and the median household income per year and residence area did not significantly impact CSS or OS.

3.3. Establishment and validation of the prediction model

Based on the LASSO regression results (Fig. 4(A and B)) and clinical features, we included 9 statistically significant factors (age, sex, marital status, tumor size, T stage, M stage, N stage, grade, histology, and surgery) for OS in the nomogram; these factors can predict the 1-year, 3-year and 5-year OS rates of patients with NETs, and the regplot package was used to visualize the nomogram. For a specific patient, the total point equals the summary of the 9 factors' points, with a straight line from the total point site to the bottom line of 5-year, 3-year, and 1-year death risks, indicating the risk of death (Fig. 5). As shown in Fig. 5, the total score was 429, and the 5-year, 3-year, and 1-year mortality rates were 66.4 %, 41.4 %, and 18.2 %, respectively. The C-indexes of the nomogram in the training set and validation set were 0.807 (95 % CI: 0.774–0.840) and 0.775 (95 % CI: 0.706–0.844), respectively. Time-dependent operator curves had AUC values of 0.840, 0.830, and 0.810 for 1-year, 3-year, and 5-year survival, respectively, in the training set. In the SEER validation set, the AUC values of the tdROC curves for 1-year, 3-year, and 5-year survival were 0.840, 0.780, and 0.770, respectively (Fig. 6(A and B)) (Supplementary Tables 2 and 3). The c-index of the nomogram in the validation cohort was 0.954 (95 % CI: 0.901–1.007), and the AUC values of the tdROC curves for 3-year and 5-year survival were 0.910 and 0.970, respectively (Fig. 7) (Supplementary Table 6). The calibration curves of the nomogram in the training and SEER validation sets are shown in Fig. 8(A–F). All curves were close to the diagonal, indicating good agreement between the nomogram-predicted and actual OS probabilities.

Table 1
Baseline features of NET patients in the SEER database.

Characteristics	Total (n = 3353)	Training set (n = 2349)	Validation set (n = 1004)	P value
Age (years), Median (Q1,Q3)	55 (50, 63)	55 (50, 63)	56 (50, 63)	0.582
Sex, n (%)				0.927
Female	1739 (52)	1220 (52)	519 (52)	
Male	1614 (48)	1129 (48)	485 (48)	
Race, n (%)				0.486
Others	1331 (40)	942 (40)	389 (39)	
White	2022 (60)	1407 (60)	615 (61)	
Marital status, n (%)	0.816			
Widow	184 (5)	127 (5)	57 (6)	
Divorced	3169 (95)	2222 (95)	947 (94)	
Rural–Urban, n (%)	0.312			
Others	1348 (40)	958 (41)	390 (39)	
Metropolitan[a]	2005 (60)	1391 (59)	614 (61)	
Median household income, n (%)	0.502			
≤ $75,000	2201 (66)	1533 (65)	668 (67)	
> $75,000	1152 (34)	816 (35)	336 (33)	
Tumor size, n (%)				0.458
≤30 mm	3052 (91)	2132 (91)	920 (92)	
>30 mm	301 (9)	217 (9)	84 (8)	
Tumor location n (%)			0.496	
Rectum	2585 (77)	1806 (77)	779 (78)	
Proximal colon	531 (16)	369 (16)	162 (16)	
Distal colon	237 (7)	174 (7)	63 (6)	
Grade, n (%)				0.416
I	2813 (84)	1965 (84)	848 (84)	
II	497 (15)	350 (15)	147 (15)	
III	43 (1)	34 (1)	9 (1)	
Histology, n (%)				0.818
Carcinoid tumor	3225 (96)	2261 (96)	964 (96)	
Others	128 (4)	88 (4)	40 (4)	
T stage, n (%)				0.965
T0/T1	2628 (78)	1845 (79)	783 (78)	
T2	275 (8)	191 (8)	84 (8)	
T3	311 (9)	215 (9)	96 (10)	
T4	128 (4)	91 (4)	37 (4)	
Tx	11 (0)	7 (0)	4 (0)	
N stage, n (%)				0.879
N0	2842 (85)	1987 (85)	855 (85)	
N1	509 (15)	360 (15)	149 (15)	
N2	2 (0)	2 (0)	0 (0)	
M stage, n (%)				0.686
M0	3190 (95)	2232 (95)	958 (95)	
M1	163 (5)	117 (5)	46 (5)	
8th AJCC stage, n (%)			0.830	
1	2583 (77)	1809 (77)	774 (77)	
2	217 (6)	147 (6)	70 (7)	
3	390 (12)	276 (12)	114 (11)	
4	163 (5)	117 (5)	46 (5)	
Summary stage, n (%)			0.871	
Distant	163 (5)	116 (5)	47 (5)	
Localized	2788 (83)	1948 (83)	840 (84)	
Reginal	402 (12)	285 (12)	117 (12)	
Chemotherapy, n (%)			0.232	
No/Unknown	3279 (98)	2292 (98)	987 (98)	
Yes	74 (2)	57 (2)	17 (2)	
Surgery, n (%)		0.836		
No	30 (1)	20 (1)	10 (1)	
Performed	3323 (99)	2329 (99)	994 (99)	
Systemic therapy n (%)			0.084	
No✣	3267 (97)	2281 (97)	986 (98)	
Yes	86 (3)	68 (3)	18 (2)	
Bone metastases, n (%)	>0.999			
No	3334 (99)	2336 (99)	998 (99)	
Yes	19 (1)	13 (1)	6 (1)	
Liver metastases, n (%)	0.808			
No	3227 (96)	2259 (96)	968 (96)	
Yes	126 (4)	90 (4)	36 (4)	

(*continued on next page*)

Table 1 (*continued*)

Characteristics	Total (n = 3353)	Training set (n = 2349)	Validation set (n = 1004)	P value
Status, n (%)			0.100	
Alive	3099 (92)	2159 (92)	940 (94)	
Dead	254 (8)	190 (8)	64 (6)	
Dead (attributable to NET/NEC), n (%)	0.409			
No	3252 (97)	2274 (97)	978 (97)	
Yes	101 (3)	75 (3)	26 (3)	
Survival months,				
Median (Q1,Q3)	45 (21, 69)	45 (21, 69)	43 (20, 69)	0.584

✢No systematic therapy or surgical procedures were performed.
[a] Metropolitan: Counties in metropolitan areas with a population of more than 1 million.

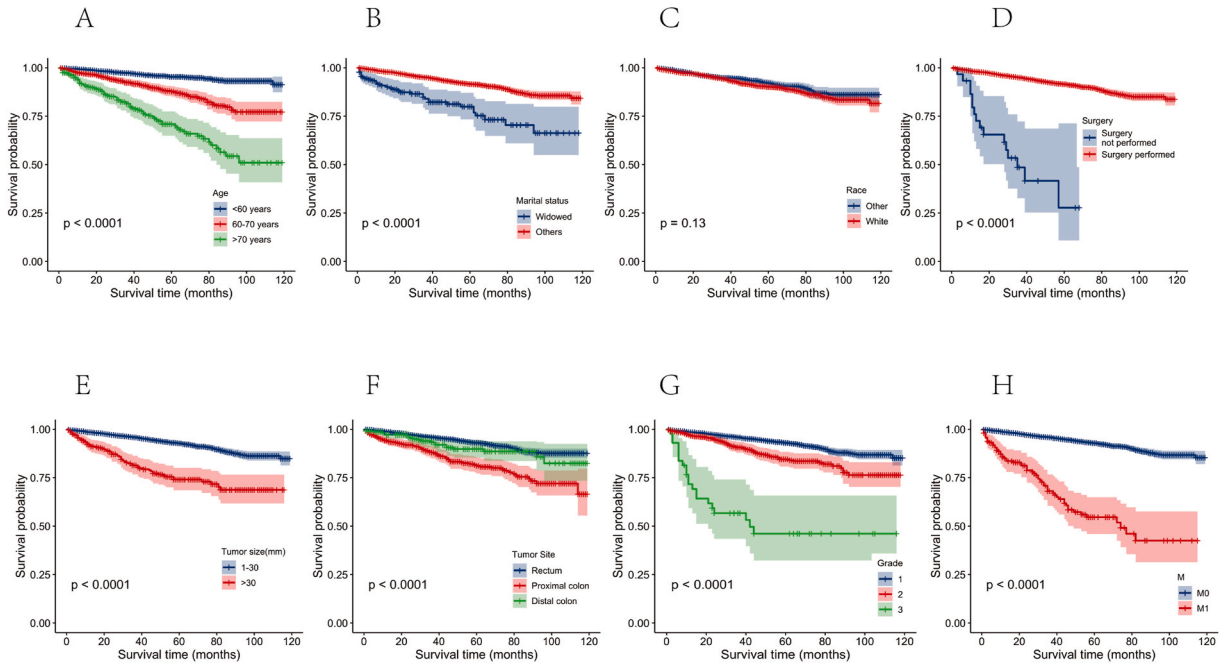

Fig. 2. Kaplan-Meier analysis of overall survival. Comparison of clinicopathological factor-specific OS according to (A) age, (B) marital status, (C) race, (D) surgery performed, (E) tumor size, (F) tumor site, (G) tumor grade, and (H) AJCC 8th M stage.

3.4. Comparison of the prediction model with the 8th edition of the AJCC TNM system

The 8th edition of the AJCC TNM staging system had C-indexes of 0.700 (95 % CI: 0.659–0.741) and 0.652 (95 % CI: 0.581–0.723) in the training and validation sets, respectively. In the training and validation sets, we tested the nomogram and the TNM model using ANOVA. There was a significant difference between these two models ($p < 0.001$), indicating that the nomogram had better predictive accuracy in both the training and validation sets. The tdROC curves of TNM in the training and validation sets are shown in Fig. 6 (C, D) (Supplementary Tables 4 and 5). Fig. 9 shows the DCA analysis of the nomogram and TNM stage. Thus, our nomogram was more powerful in predicting the clinical prognosis of patients with colorectal NETs.

4. Discussion

A considerable amount of literature has been published regarding the prognosis of colorectal NETs. Satya et al. reported an observed 5-year OS rate of 88.5 % for patients with rectum NETs and 54.6 % for patients with colorectal NETs. The median OS in patients with NETs was also 7.4 years [2]. In this paper, we revealed the OS and CSS of patients with colorectal NETs. The 5-year OS rate of patients with NETs was 90.9 % (95 % CI: 89.7%–92.2 %), and the 5-year CSS rate was 96.2 % (95 % CI: 95.4%–97.0 %). Increasing evidence has suggested that G3 NETs and NECs are considered two distinct diseases in terms of tissue origin, molecular markers, and clinical manifestations [15]. In general, the prognosis for patients with colorectal NETs is relatively positive. G3 NETs and G1 and G2 NETs have different clinical manifestations. Similarly, our study revealed that the prognosis of patients with G3 NETs was significantly worse (5-year OS rate: grade I/II/III, 93.0 %/84.5 %/46.1 %), confirming that G3 NETs may have different clinical outcomes.

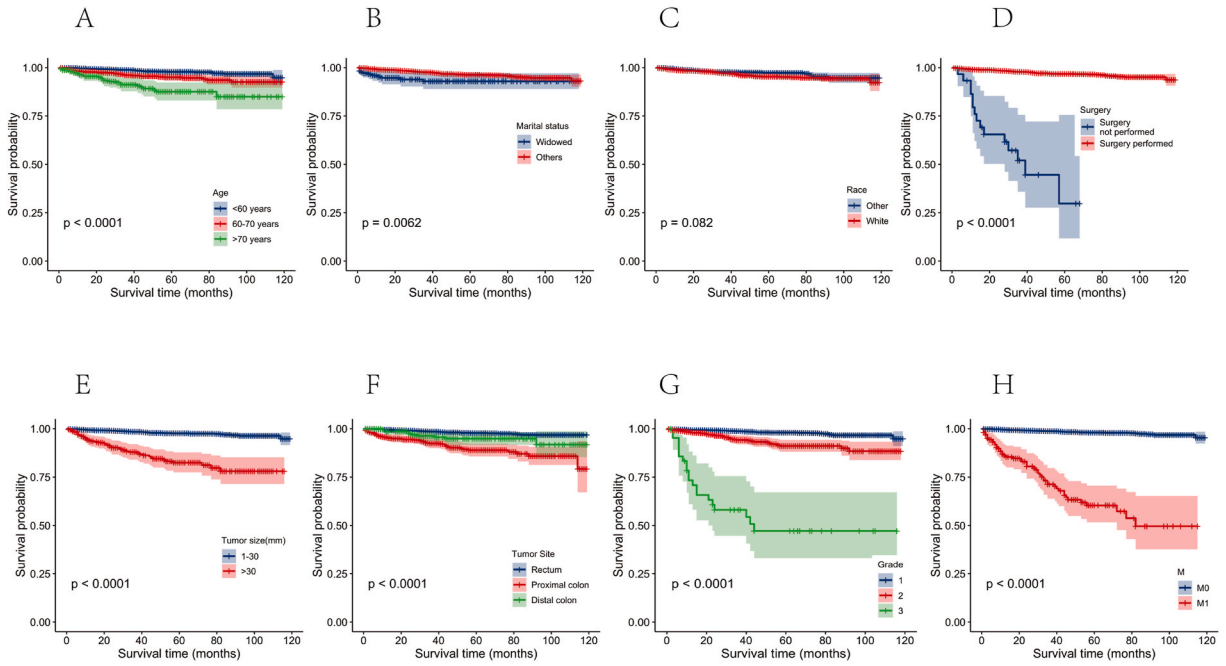

Fig. 3. Kaplan-Meier analysis of cancer-specific survival. Comparison of clinicopathological factor-specific CSS according to (A) age, (B) marital status, (C) race, (D) surgery performed, (E) tumor size, (F) tumor site, (G) tumor grade, and (H) AJCC 8th M stage.

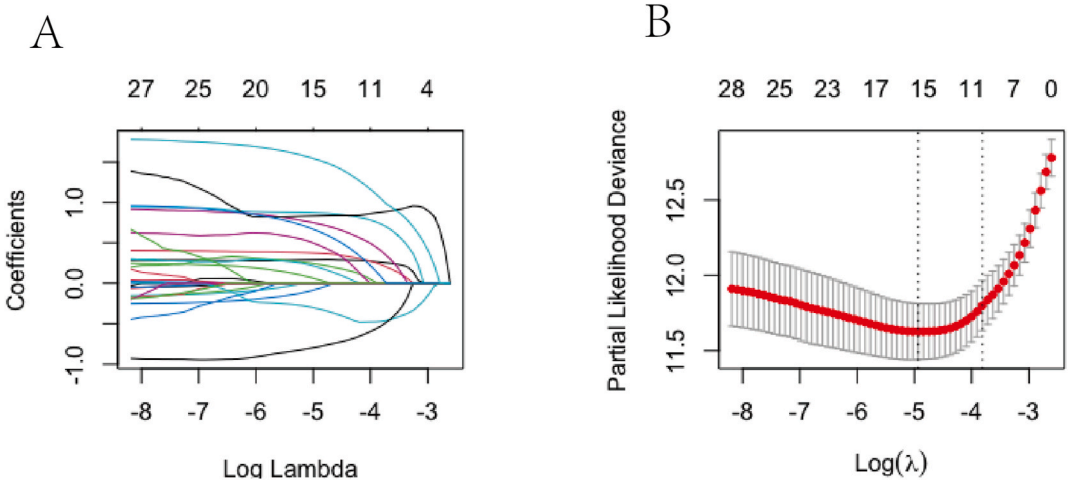

Fig. 4. Selection of predictive variables using LASSO binary Cox regression analysis. (A) Plot of the LASSO coefficient for clinical and pathological characteristics. (B) Cross-certification of the LASSO model and chosen tuning parameter (lambda) in the LASSO Cox model.

The prevalence of colorectal NETs is currently increasing, and research has shown that it may promote an increase in early-onset colorectal cancer. However, few specific models have focused on the prognosis and risk stratification of patients with colorectal NETs [5]. This is the first study to focus on establishing a prediction model for colorectal NET patients based on the SEER database and to validate the model in an external cohort of Chinese patients. A survival prediction nomogram was established to visualize the 1-year, 3-year, and 5-year OS of patients with NETs. The C-index and tdROC curve showed that the nomogram had good predictive ability (C-indexes of 0.807 and 0.775 in the training set and validation set, respectively), and the calibration curves implied that the nomogram had satisfactory predictive accuracy. Moreover, we established an external validation cohort with a C-index of 0.954 (95 % CI: 0.901–1.007). The C-index and tdROC curve of the nomogram model in the cohort indicated good prediction accuracy. In contrast, the C-indexes of the 8th edition of the AJCC TNM were 0.700 and 0.652 in the training and validation sets, respectively. Thus, we can infer that our prediction model had a better ability to predict OS. According to the DCA curves above, we can conclude that the survival nomogram had better utility than TNM in assisting clinical practice. Most previous studies have focused on analysing the prognosis of

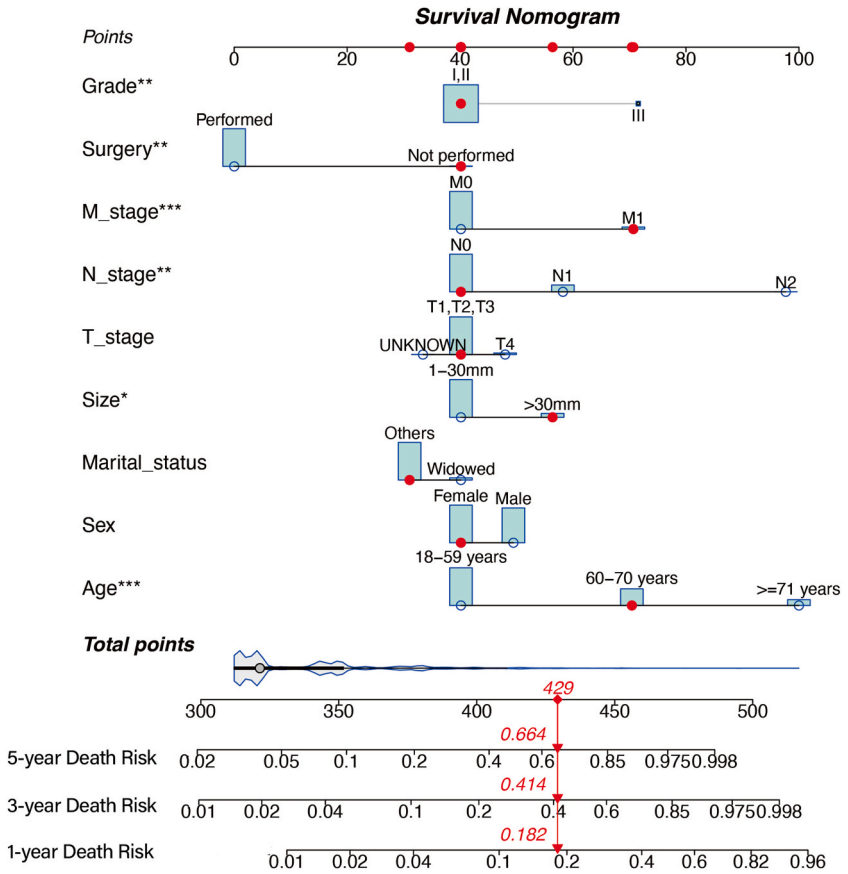

Fig. 5. Survival nomogram for predicting the 1-year, 3-year, and 5-year survival rates of colorectal NET patients.

patients with NENs according to the 2010 WHO classification criteria. Although Zihan Xu et al. reported a nomogram for predicting the prognosis of GEP-NETs, they did not consider the updated WHO classification of colorectal NETs. Moreso, Jinmao Zou et al. established a prediction nomogram for colorectal NETs. Nevertheless, they included "neuroendocrine carcinoma" as an indicator of NETs in the sample, which does not align with the current WHO classification standard [1]. Our nomogram identified colorectal NET patients using the SEER AYA code and ICD-O 3 code to ensure that all enrolled patients met the NETs definition of the WHO. Moreover, most of the current nomograms included 4–5 prognostic factors, while 9 factors were identified in our study, making the nomogram more detailed and comprehensive. Therefore, our survival nomogram still plays a vital role in predicting the prognosis of patients with colorectal NETs [3].

For patients with NETs, chemotherapy is usually performed to treat liver metastasis and control hormonal symptoms [16–18]. Unlike other studies, we evaluated the effect of chemotherapy on prognosis. The Kaplan-Meier analysis revealed that patients who received chemotherapy had lower 5-year OS and CSS rates. We assumed this was because of confounding factors such as age, grade, and distant metastasis. Although it was statistically significant in the LASSO regression analysis, considering the potential confounding factors, it was not included in the survival nomogram. Keshuai et al. reported that widowed patients with NETs had worse survival outcomes, which was an independent risk factor for survival [14]. We came to the same conclusion about the research based on LASSO regression and survival analysis (5-year OS: 79.9 % vs. 91.6 %). According to the nomogram model, the widowed group had 10 more points than the other marital statuses group according to the nomogram. Surgery is a standard method for treating colorectal NETs and is a protective factor for prognosis [19]. Although we were keen to determine the prognostic impact of endoscopic procedures versus surgical procedures, there were no relevant data in the version of the SSER database we used in the current study.

Previous studies have noted that colon and rectal NETs should be considered distinct diseases since colorectal NETs are associated with worse survival outcomes [20,21]. Our study indicated similar results, with shorter survival in patients with colon NETs. One unanticipated finding was that proximal colon NETs might have worse survival outcomes than distal colon and rectum NETs (5-year OS rate: 81.2 % vs. 89.8 %/93.1 %). However, tumor location was not significantly different according to the LASSO analysis. Mengjie et al. also reported that tumor location was not a substantial factor for colorectal NENs, and more data may be needed for advanced analysis [11]. We further speculate that proximal colon NETs have a worse prognosis because most colorectal NETs are found by colonoscopy, and reaching the proximal colon is difficult. Surprisingly, sex was identified as an important factor by LASSO analysis. Considering that previous studies revealed that females had better survival outcomes than males, indicating that sex may have clinical significance, we added sex as a factor in the nomogram [22]. Furthermore, the 5-year OS rate of White patients was shorter than that of

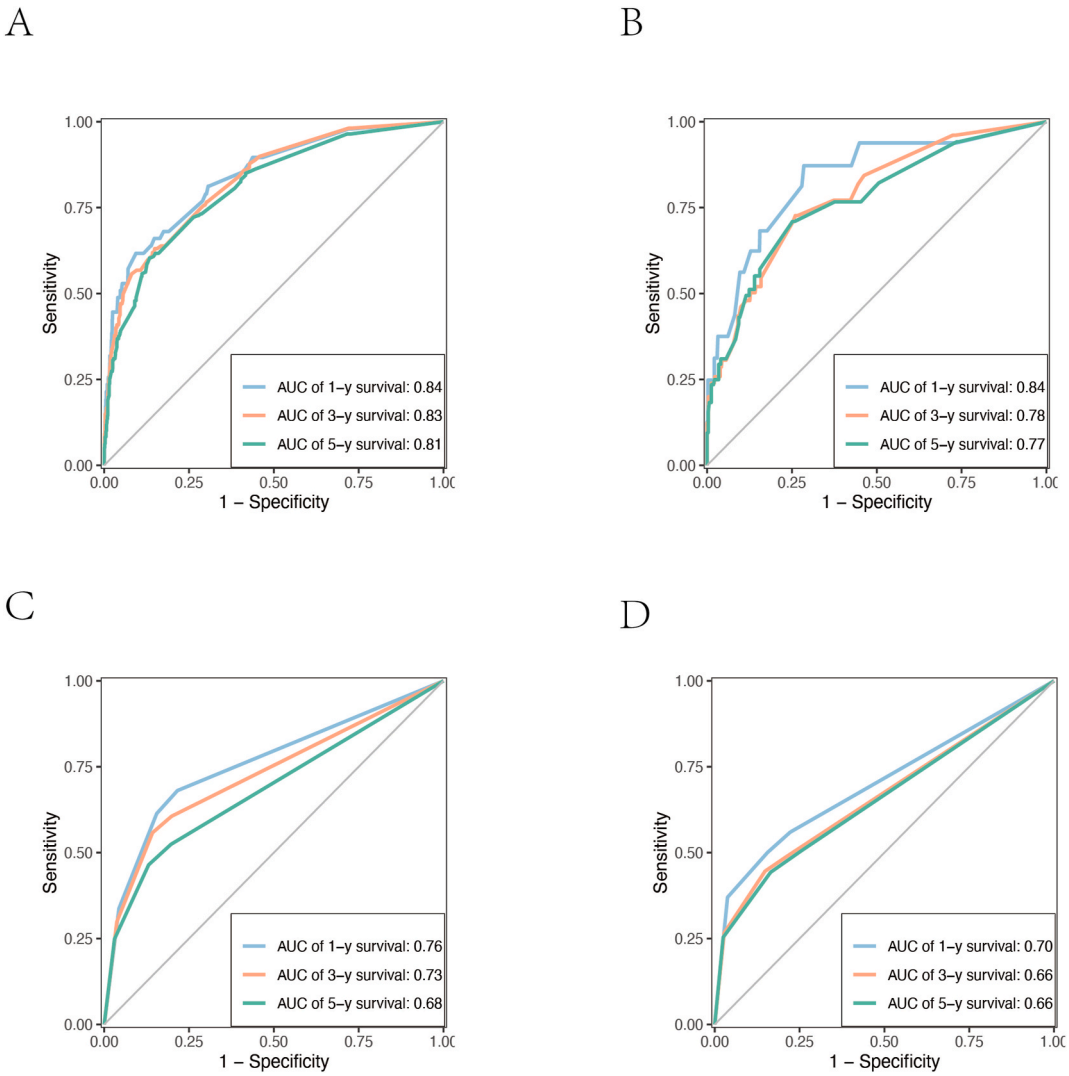

Fig. 6. Time-dependent ROC curves for predicting the 1-year, 3-year, and 5-year OS of colorectal NET patients according to the (A) survival nomogram in the training set, (B) survival nomogram in the validation set, (C) TNM staging in the training set, and (D) TNM staging in the validation set.

people of other races, but this difference was not statistically significant according to LASSO or multivariate analysis. As Kessel et al. reported, race may have little influence on prognosis [23]. T stage was also identified as an important factor by LASSO regression. However, due to the existence of confounding factors such as tumor size, we did not directly incorporate T stage into the model but reclassified it into three groups (T1+T2+T3, T4, unknown T). As mentioned earlier, Grade I and Grade II NETs are considerably different from Grade III NETs, so we divided the grades into two groups, Grade I + Grade II and Grade III, into the model. Under such grouping, the model showed good efficiency and clinical application ability. Histology (carcinoid tumor and other) was selected by LASSO regression, but considering that "carcinoid" is an outdated expression and has now been replaced by "neuroendocrine tumor" in many guidelines, it did not have much clinical significance; thus, we did not include this feature in our prediction model.

Despite these promising results, some questions remain unaswered. First, this retrospective study had unavoidable selection bias. Nevertheless, our retrospective study provides a direction and available predictive model for further prospective research. Second, considering that our data originated from the SEER database, these results were limited to the U.S. population. In addition, the SEER databases lacks some important information, such as information on Ki-67 and CEA, that may be related to patient prognosis, thereby reducing the accuracy of the survival nomogram. Due to the change in TNM staging standards with time and the lack of clinical and pathological information, we failed to analyse the influence of detailed TNM staging on prognosis and divided TNM staging into only 4 stages. Finally, although we established an external validation cohort, due to the relatively low prevalence rate, the sample size was relatively small. Given the above limitations, large sample sizes and multicentre real-world prospective studies are required in the future.

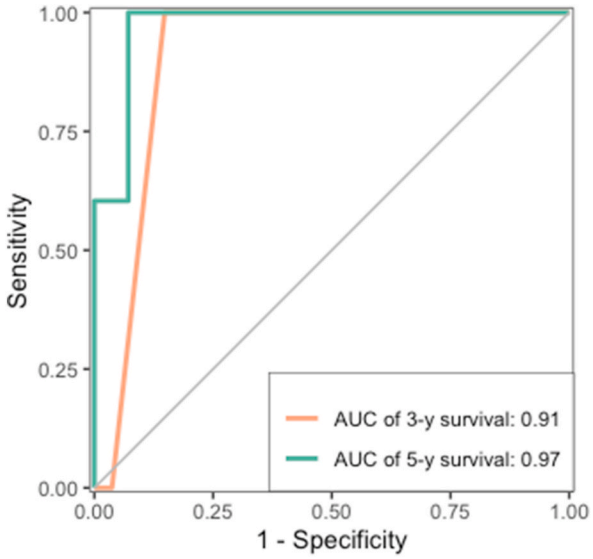

Fig. 7. Time-dependent ROC curves for predicting the 3-year and 5-year OS of patients with colorectal NETs according to the survival nomogram in the validation cohort.

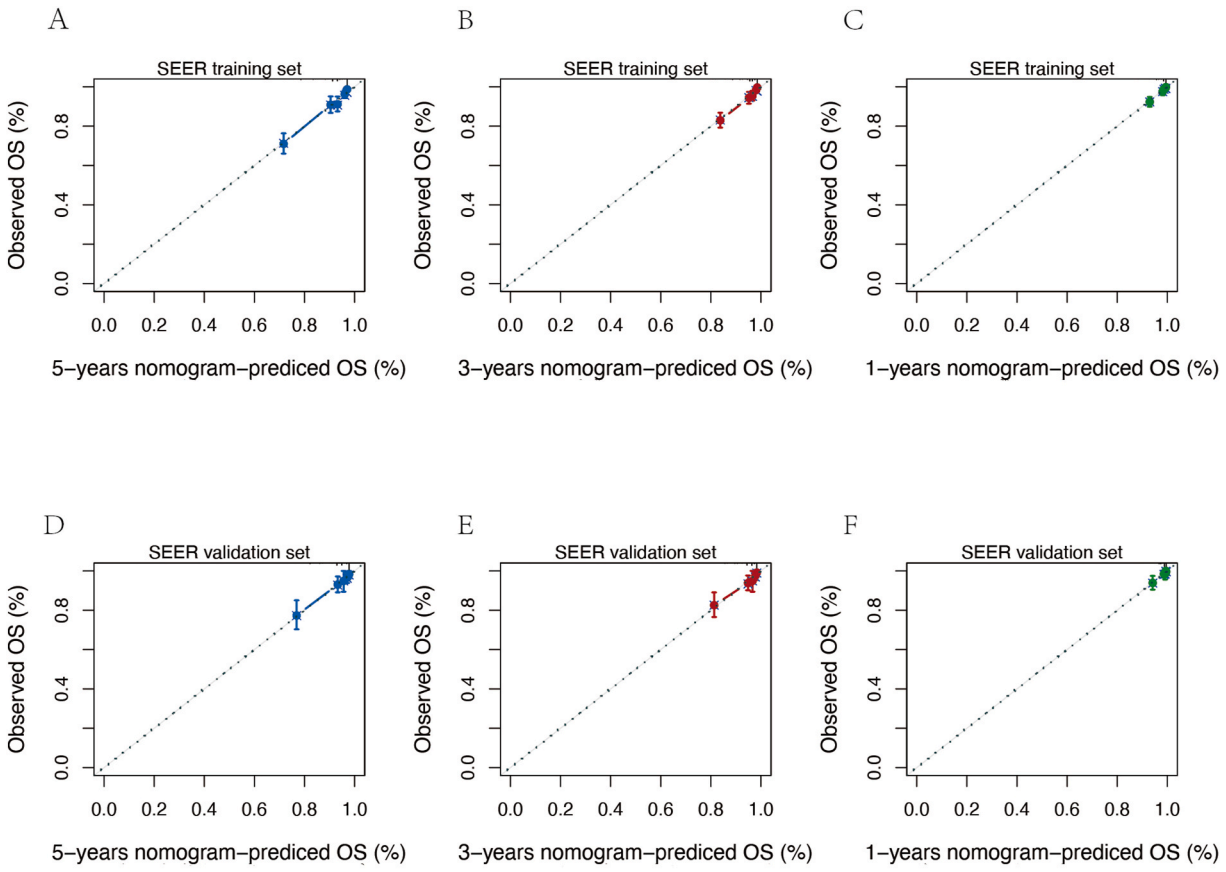

Fig. 8. The calibration curves for the nomogram model in predicting survival in the training and validation sets. (A–C) Calibration plots of 5-year, 3-year, and 1-year OS in the training set. (D–F) Calibration plots of 5-year, 3-year, and 1-year OS in the validation set.

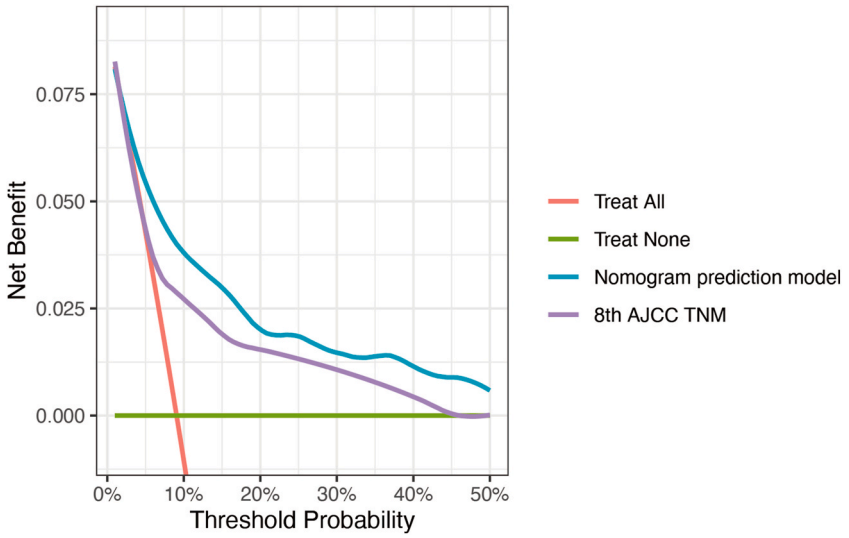

Fig. 9. Decision curve analysis of the survival nomogram for predicting OS. The DCA of the nomogram and TNM stage for OS in all patients from the SEER cohort.

5. Conclusion

We analysed data from the SEER database to better predict the prognosis of colorectal NETs. A nomogram with good predictive accuracy and clinical utility, including 9 factors (age, sex, marital status, tumor size, T stage, M stage, N stage, grade, and surgery), was constructed. Although the nomogram had better predictive accuracy than the 8th edition of the AJCC TNM staging system, it had some limitations, and a large population and multicentre real-world prospective studies are required in the future.

Consent to publication

Not applicable.

Consent to publication

The authors declare that no competing interests exist.

Funding

This work was supported by the National Natural Science Foundation of China (grant number 72104150), the Beijing Science and Technology Planning Project (grant numbers Z191100006619084 and Z211100002921028), the Digestive Medical Coordinated Development Center of Beijing Hospitals Authority (grant number XXZ0015), the Special Scientific Research Fund for Tutor (grant number YYDSZX201901), and Capital's Funds for Health Improvement and Research (grant numbers CFH2022-2-2025 and 2020-4-2085). None of the funding organizations had any role in the design and conduct of the study; in the collection, management, and analysis of the data; or in the manuscript's preparation, review, and approval.

Data availability statement

The training and validation set data relevant to the study are available in the SEER database (https://seer.cancer.gov/). External validation data can be requested from the corresponding authors.

Ethical approval

The validation cohort study was approved by the Bioethics Committee of Beijing Friendship Hospital, Capital Medical University (2020-P2-290-01).

CRediT authorship contribution statement

Jiuyue Ma: Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation. **Xiaoqian Ma:** Writing – original draft, Resources, Methodology. **Jie Xing:** Validation, Software, Resources. **Ruyun Song:** Software, Resources.

Yang Zhang: Writing – review & editing, Formal analysis. **Mo Liu:** Writing – review & editing, Conceptualization. **Shuilong Guo:** Writing – review & editing, Funding acquisition. **Qian Zhang:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration. **Jing Wu:** Writing – review & editing, Visualization, Validation, Supervision, Software, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank AJE (https://www.aje.cn/) for English language editing and review services.

List of abbreviations

NETs	neuroendocrine tumor
NENs	neuroendocrine neoplasms
NECs	neuroendocrine carcinomas
AJCC	American Joint Cancer Committee
SEER	Surveillance, Epidemiology, and End Results Program
OS	overall survival; CSS: cancer-specific survival

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e35720.

References

[1] S.C. Jinmao Zou, et al., Prognostic and metastasis-related factors in colorectal neuroendocrine tumors: a cross-sectional study based on the Surveillance, Epidemiology and End Results, Oncol. Lett. 18 (5) (2019).
[2] S. Das, A. Dasari, Epidemiology, incidence, and prevalence of neuroendocrine neoplasms: are there global differences? Curr. Oncol. Rep. 23 (4) (2021) 43.
[3] Epidemiologic Trends of and Factors Associated with Overall Survival for Patients with Gastroenteropancreatic Neuroendocrine Tumors in the United States.
[4] A. Dasari, et al., Trends in the incidence, prevalence, and survival outcomes in patients with neuroendocrine tumors in the United States, JAMA Oncol. 3 (10) (2017) 1335–1342.
[5] J.L. Petrick, et al., Racial disparities and sex differences in early- and late-onset colorectal cancer incidence, 2001-2018, Front. Oncol. 11 (2021) 734998.
[6] P. Gong, et al., Prognostic significance for colorectal carcinoid tumors based on the 8th edition TNM staging system, Cancer Med. 9 (21) (2020) 7979–7987.
[7] Y. Zhang, et al., Clinicopathological features and prognostic validity of the European Neuroendocrine Tumor Society (ENETS) and American Joint Committee on Cancer (AJCC) 8th staging systems in colonic neuroendocrine neoplasms, Cancer Med. 8 (11) (2019) 5000–5011.
[8] Z. Xu, et al., Epidemiologic trends of and factors associated with overall survival for patients with gastroenteropancreatic neuroendocrine tumors in the United States, JAMA Netw. Open 4 (9) (2021) e2124750.
[9] I.D. Nagtegaal, et al., The 2019 WHO classification of tumours of the digestive system, Histopathology 76 (2) (2020) 182–188.
[10] G. Rindi, et al., Overview of the 2022 WHO classification of neuroendocrine neoplasms, Endocr. Pathol. 33 (1) (2022) 115–154.
[11] M. Jiang, et al., Clinicopathological features and prognostic factors of colorectal neuroendocrine neoplasms, Gastroenterol Res Pract 2017 (2017) 4206172.
[12] Y.W. Jing Li, Fang Han, Zhu Wang, Lichun Xu, Jiandong Tong, Disadvantage of survival outcomes in widowed patients with colorectal neuroendocrine neoplasm: an analysis of surveillance, epidemiology and end results database, Oncotarget 7 (No. 50) (2016) 83200–83207.
[13] X. Ding, et al., Risk and prognostic nomograms for colorectal neuroendocrine neoplasm with liver metastasis: a population-based study, Int. J. Colorectal Dis. 36 (9) (2021) 1915–1927.
[14] K. Xiao, et al., The effect of marital status on the survival of patients with colorectal neuroendocrine neoplasms: an analysis of the SEER database, Rev. Esp. Enferm. Dig. 112 (2) (2020) 109–117.
[15] M. Shi, et al., Gastroenteropancreatic neuroendocrine neoplasms G3: novel insights and unmet needs, Biochim. Biophys. Acta Rev. Canc 1876 (2) (2021) 188637.
[16] M.H. Shah, et al., Neuroendocrine and adrenal tumors, version 2.2021, NCCN clinical practice guidelines in oncology, J Natl Compr Canc Netw 19 (7) (2021) 839–868.
[17] M. Pavel, et al., ENETS consensus guidelines update for the management of distant metastatic disease of intestinal, pancreatic, bronchial neuroendocrine neoplasms (NEN) and NEN of unknown primary site, Neuroendocrinology 103 (2) (2016) 172–185.
[18] M. Pavel, et al., Gastroenteropancreatic neuroendocrine neoplasms: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up, Ann. Oncol. 31 (7) (2020) 844–860.
[19] J.K. Ramage, et al., ENETS consensus guidelines update for colorectal neuroendocrine neoplasms, Neuroendocrinology 103 (2) (2016) 139–143.
[20] M. Araghi, et al., Colon and rectal cancer survival in seven high-income countries 2010-2014: variation by age and stage at diagnosis (the ICBP SURVMARK-2 project), Gut 70 (1) (2021) 114–126.
[21] R. Chagpar, et al., Neuroendocrine tumors of the colon and rectum: prognostic relevance and comparative performance of current staging systems, Ann. Surg Oncol. 20 (4) (2013) 1170–1178.
[22] O. Abdel-Rahman, N. Fazio, Sex-based differences in prognosis of patients with gastroenteropancreatic-neuroendocrine neoplasms: a population-based study, Pancreas 50 (5) (2021) 727–731.
[23] E. Kessel, et al., Racial differences in gastroenteropancreatic neuroendocrine tumor treatment and survival in the United States, Pancreas 50 (1) (2021) 29–36.