World Journal of
Clinical Cases

Contents

Contents

ABOUT COVER

Chin-Hsiao Tseng, MD, PhD, Full Professor, Department of Internal Medicine, National Taiwan University College of Medicine, No. 1 Jen Ai Road Section 1, Taipei 100, Taiwan. ccktsh@ms6.hinet.net

AIMS AND SCOPE

The primary aim of *World Journal of Clinical Cases* (*WJCC*, *World J Clin Cases*) is to provide scholars and readers from various fields of clinical medicine with a platform to publish high-quality clinical research articles and communicate their research findings online.

   WJCC mainly publishes articles reporting research results and findings obtained in the field of clinical medicine and covering a wide range of topics, including case control studies, retrospective cohort studies, retrospective studies, clinical trials studies, observational studies, prospective studies, randomized controlled trials, randomized clinical trials, systematic reviews, meta-analysis, and case reports.

INDEXING/ABSTRACTING

The *WJCC* is now indexed in Science Citation Index Expanded (also known as SciSearch®), Journal Citation Reports/Science Edition, Scopus, PubMed, and PubMed Central. The 2020 Edition of Journal Citation Reports® cites the 2019 impact factor (IF) for *WJCC* as 1.013; IF without journal self cites: 0.991; Ranking: 120 among 165 journals in medicine, general and internal; and Quartile category: Q3. The *WJCC*'s CiteScore for 2019 is 0.3 and Scopus CiteScore rank 2019: General Medicine is 394/529.

RESPONSIBLE EDITORS FOR THIS ISSUE

ORIGINAL ARTICLE

Retrospective Study

Construction of a clinical survival prognostic model for middle-aged and elderly patients with stage III rectal adenocarcinoma

Hao Liu, Yu Li, Yi-Dan Qu, Jun-Jiang Zhao, Zi-Wen Zheng, Xue-Long Jiao, Jian Zhang

ORCID number: Hao Liu 0000-0002-3625-0437; Yu Li 0000-0002-0815-3023; Yi-Dan Qu 0000-0002-8818-9201; Jun-Jiang Zhao 0000-0002-5334-4662; Zi-Wen Zheng 0000-0001-5916-629X; Xue-Long Jiao 0000-0002-0348-9732; Jian Zhang 0000-0003-2998-873X.

Author contributions: Liu H, Jiao XL, and Zhang J participated in the design of the study, acquisition of data, and article revisions; Liu H, Li Y, Qu YD, Zhao JJ, Zheng ZW, and Zhang J analyzed the data and drafted the article. All authors have approved the final version of the article.

Institutional review board statement: This study was reviewed and approved by the Ethics Committee of the Affiliated Hospital of Qingdao University.

Informed consent statement: As this study is based on a publicly available database without identifying patient information, informed consent was not needed.

Conflict-of-interest statement: All authors declare no conflicts-of-interest related to this article.

Hao Liu, Yu Li, Jun-Jiang Zhao, Zi-Wen Zheng, Xue-Long Jiao, Jian Zhang, Department of Colorectal Surgery, The Affiliated Hospital of Qingdao University, Qingdao 266000, Shandong Province, China

Yi-Dan Qu, Rheumatology and Immunology Department, The Affiliated Hospital of Qingdao University, Qingdao 266000, Shandong Province, China

Corresponding author: Jian Zhang, PhD, Chief Doctor, Postdoc, Professor, Surgeon, Department of Colorectal Surgery, The Affiliated Hospital of Qingdao University, No. 19 Jiangsu Road, Shinan District, Qingdao 266000, Shandong Province, China. zhangjian@qduhospital.cn

Abstract

BACKGROUND
Nomograms for prognosis prediction in colorectal cancer patients are few, and prognostic indicators differ with age.

AIM
To construct a new nomogram survival prediction tool for middle-aged and elderly patients with stage III rectal adenocarcinoma.

METHODS
A total of 2773 eligible patients were divided into the training cohort (70%) and the validation cohort (30%). Optimal cutoff values were calculated using the X-tile software for continuous variables. Univariate and multivariate Cox proportional hazards regression analyses were used to determine overall survival (OS) and cancer-specific survival (CSS)-related prognostic factors. Two nomograms were successfully constructed. The discriminant and predictive ability and clinical usefulness of the model were also assessed by multiple methods of analysis.

RESULTS
The 95%CI in the training group was 0.719 (0.690-0.749) and 0.733 (0.702-0.74), while that in the validation group was 0.739 (0.696-0.782) and 0.750 (0.701-0.800) for the OS and CSS nomogram prediction models, respectively. In the validation group, the AUC of the three-year survival rate was 0.762 and 0.770, while the AUC of the five-year survival rate was 0.722 and 0.744 for the OS and CSS nomograms, respectively. The nomogram distinguishes all-cause mortality from cancer-specific mortality in patients with different risk grades. The time-

dependent AUC and decision curve analysis showed that the nomogram had good clinical predictive ability and decision efficacy and was significantly better than the tumor-node-metastases staging system.

CONCLUSION
The survival prediction model constructed in this study is helpful in evaluating the prognosis of patients and can aid physicians in clinical diagnosis and treatment.

Key Words: Rectal adenocarcinoma; Lymph node positive rate; Nomogram; Prognostic model; Predictive model; Survival time

Core Tip: This investigation was based on a large-scale population study of middle-aged and elderly patients with stage III rectal adenocarcinoma. In this study, we analyzed the clinical data of thousands of patients with stage III rectal adenocarcinoma aged 45 years or older and determined the relevant prognostic factors and the degree of impact. New cutoff values were identified and used to construct nomograms. The nomograms showed excellent clinical predictive ability and decision power. The nomograms constructed in this study have clinical utility.

INTRODUCTION

Colorectal cancer (CRC) is the second most common malignant tumor worldwide, and its incidence is increasing[1]. Rectal adenocarcinoma is a common type of rectal cancer[2]. Seventy percent of patients with rectal cancer have no evidence of distant metastasis at the time of diagnosis[3]. Middle-aged and elderly patients are often susceptible to rectal adenocarcinoma, and stage III rectal adenocarcinoma is the most prevalent type in this population[4]. Therefore, providing this population with the tools necessary for prognosis can help physicians make correct clinical decisions regarding relevant treatment, benefiting physicians and patients.

A statistical model based on the Cox proportional hazards model for individualized prediction analysis of clinical events, can comprehensively include multiple prognostic indicators to construct a survival prediction model and visualize the risk of prognostic factors and outcomes. A nomogram is an effective tool for visual assessment and quantifies individual risk by using important prognostic factors. In the diagnosis and treatment of a variety of cancers, nomograms currently show good predictive performance, such as for pancreatic cancer[5] and liver cancer[6]. Nomograms can improve the accuracy of survival prediction and provide physicians with more accurate diagnosis and treatment.

In previous studies related to clinical survival, the sample sizes were small, and the evaluations were incomplete. Simultaneously, changes and advances in medical technology also affect the prognosis of patients. However, it is still necessary to study the survival prognostic factors in middle-aged and elderly patients with stage III rectal adenocarcinoma. Presently, there are few prognostic models for the population. The nomogram helps to assess the clinical risks and benefits for the patient and assists the physician in making clinical decisions[7]. Therefore, we assessed the five-year overall survival (OS) and cancer-specific survival (CSS) of these patients. We also constructed a nomogram based on multiple prognostic indicators to predict survival.

METHODS AND MATERIALS

Collection of patient data

Data were obtained from the Surveillance, Epidemiology, and End Results Program (SEER) database (https://seer.cancer.gov/data/). The data used in this study have been licensed from the SEER database (accession number: 12285-Nov2019). This study was approved by the Ethics Committee of the Affiliated Hospital of Qingdao University (ethical approval number: QYFY WZLL 25780) The SEER database has a large clinical multicenter data sample that provides baseline patient data and a variety of clinical indicators. This provided a strong basis for us to analyze the prognostic survival factors of rectal cancer patients from various aspects and construct nomogram models[8].

The inclusion criteria were: (1) Patients registered during 2010-2015; (2) Surgical treatment; (3) The pathological diagnosis was rectal adenocarcinoma; (4) Clinical stage III disease and (5) Relevant follow-up data were complete. The exclusion criteria were: (1) Pathological diagnosis and grading were unclear; (2) Ethnicity unknown; (3) Unknown tumor-node-metastases (TNM) stage; (4) Carcinoembryonic antigen (CEA), tumor engraftment (TD), perineural invasion (PNI), and circumferential resection margin (CRM) data were unknown; (5) Tumor size was unknown; and (6) The number of positive lymph nodes (pLN), number of examined lymph nodes (DLNs), and lymph node-positive rate (LNR) were unknown.

We collected the following demographic information: patient age, ethnicity, and sex. We also collected the following clinical information: year of diagnosis, tumor location, CEA, TD, CRM, PNI, histological grade, clinical stage, tumor size, pLN, DLNs, LNR, metastasis status, histopathological type (malignant behavior based on International Classification of Diseases for Oncology, 3rd Edition), survival time, survival status, and cause of death. Clinical staging was based on the American Joint Committee on Cancer (AJCC) 8th edition staging system. The highest CEA value in the preoperative tests was considered for evaluation.

The clinical data of 260833 patients were obtained from the SEER database. After exclusion, 2773 eligible patients were enrolled in the study (Figure 1). Eligible patients were randomly divided into the training and validation cohorts at 70% and 30%, respectively, using R software. The training cohort was used to construct a nomogram for the clinical survival prediction model, and the validation cohort was used for internal validation.

Study endpoints

OS and CSS were study endpoints. At the same time, survival outcomes at three and five years were assessed. According to the C-index, receiver operating characteristic (ROC) curve, and calibration curve, the effectiveness, accuracy, and predictive ability of the nomogram were evaluated. The Kaplan-Meier curves were used to assess the discriminatory ability of the nomograms for patients with different risk grades. The area under the curve (AUC) of the time-dependent curve and decision curve helped us to judge the predictive performance and clinical application of the model. OS was defined as the time from the first day of last follow-up or diagnosis to death[9]. CSS was defined as the time from diagnosis to death due to rectal adenocarcinoma and last follow-up[10].

Statistical analysis

Based on the relationship between the predictor and survival outcome, X-tile software creates the optimal segmentation point for this variable and helps visualize the image[11]. Statistical analysis was performed using SPSS software (SPSS 24.0, Chicago, IL, United States) and R software (R software, version 4.0). R packages "rms," "foreign," and "caret" were used for randomization, while "rms," "foreign," and "survival" were used to perform Cox regression analysis, calculate risk score and C-index, and draw the calibration curve, decision curve and survival curve of the prediction model. The R packages "regplot" and "timeROC" were used to draw the nomogram plot, the ROC and the time-dependent AUC, respectively. All statistical tests were two-way.

Nomogram construction and performance evaluation

Optimal cutoff values for OS were assessed using the X-tile software for the continuous variables including age, tumor size, pLN, DLNs, and LNR.

Frequency and percentage were used to evaluate the baseline data of the patients in both groups. Differences between the two groups were analyzed using the chi-square

SEER1973-2015 (260882)

Age < 45 yr (14452)

Gender, race unknown (924)

Non-adenocarcinoma (73387)

TNM stage unknown (144525)

Not stage III (19421)

Tumor size unknown (755)

CEA/TD/CRM/PNI unknown (4422)

Pathological grade unknown/
un explained death/
incomplete follow-up information/
no surgical treatment (496)

Final (2773)
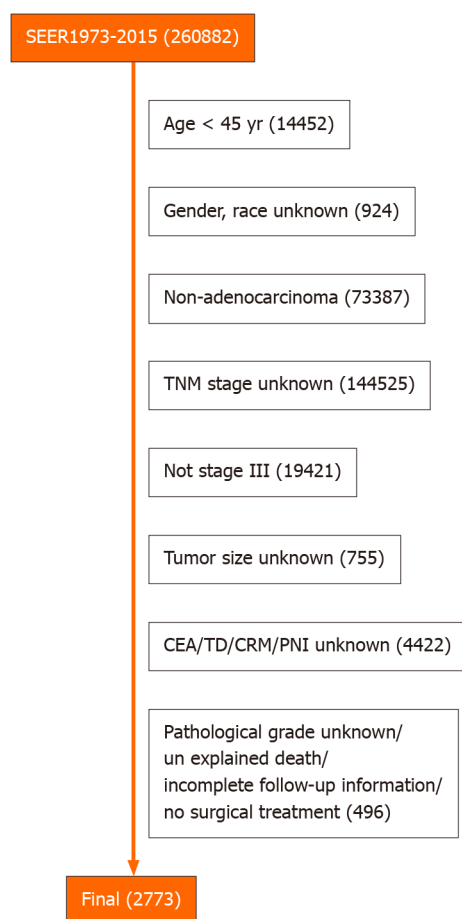
Figure 1 Flow diagram. From the 260882 patients with rectal cancer in the Surveillance, Epidemiology, and End Results Program database, a total of 2773 eligible patients were screened. TNM: Tumor-node-metastases; CEA: Carcinoembryonic antigen; TD: Tumor engraftment; CRM: Circumferential resection margin; PNI: Perineural invasion.

test. Univariate *vs* multivariate Cox analysis was used to screen for independent prognostic factors, which were then used to construct the nomograms. Hazard ratios (HR) and 95%CI were also calculated. The predictive ability of the model was assessed by C-index and AUC. The similarity between the predicted and actual results was compared by a calibration curve. *P* values < 0.05 were considered significant.

Verification of nomogram discrimination

The "prediction" function of the R software was used to calculate a risk score for each patient. The optimal cutoff value for the risk score in the training cohort was clarified by the X-tile software. Through the optimal cutoff value, patients were divided into three groups: high risk group, medium risk group, and low risk group. The log-rank test was used to assess survival differences between patients with different degrees of risk and to plot survival curves based on risk scores.

Evaluation of predictive efficacy of the nomogram

The AUC values of the prediction model at different time points can be visualized by time-dependent AUC[12]. The predictive ability of TNM stage was compared with the nomogram using this method.

Clinical efficacy evaluation of the nomogram

Clinical utility and net benefit were assessed by decision curve analysis[13]. The clinical efficacy of the nomogram was tested and compared with the TNM stage using the decision curve.

RESULTS

Baseline patient characteristics

The differences in baseline patient characteristics between the two groups are shown in Table 1 and Supplementary Figure 1. The randomization was satisfactory. In the univariate OS analysis, except for race, gender, and tumor location, other factors were significant. In the univariate CSS analysis, except for race, sex, tumor location, and DLNs, other factors were significant (Table 2). The variables with significant differences were included in the multivariate analysis, which showed that there were significant differences in age, stage, CEA, CRM, TD, PNI, LNR, and tumor size in OS and CSS (Table 3).

Construction of nomogram models

Factors with minor effects were excluded based on the results of Cox regression analysis. We used age, stage, tumor size, CEA, TD, CRM, PNI, and LNR as predictor variables of OS and CSS. Two nomogram models for survival prediction were constructed and calibration curves were plotted for the validation cohort (Figure 2). From the calibration curve, we can see that the predicted results are very similar to the actual results. The 3- and 5-year OS/CSS probabilities of the target population were predicted by calculating the total score of the predictors in the nomogram (Supplementary Table 1).

Prediction ability and clinical usefulness of the prediction models

The 95%CI of the OS and CSS prediction models for the training cohort were 0.719 (0.690-0.749) and 0.733 (0.702-0.746), respectively. The 95%CI of the OS and CSS prediction models for the validation cohort was 0.739 (0.696-0.782) and 0.750 (0.701-0.800), respectively. In the validation group, the AUC of the three-year survival rate was 0.762 and 0.770, and the AUC of the five-year survival rate was 0.722 and 0.744 for the OS and CSS nomograms, respectively. The prediction model showed excellent predictive ability (Figure 3A and B).

Time-dependent AUCs are not only useful for evaluating the predictive ability of nomograms, but they can also be compared with the predictive abilities of different clinical stages. Using this curve, we found that the AUC values of the predictive models for OS and CSS were higher than the TNM stage between 0 and 60 mo. We can consider the predictive power of this model to be superior to TNM stage (Figure 3C and D).

Predictive model patient discrimination validation

The optimal OS risk score cutoffs were 1.0 and 2.7. The optimal CSS risk score cutoffs were 1.0 and 2.8. The study population was divided into three risk groups according to the size of the cut-off value: high, medium, and low, and the corresponding survival curves were plotted. The survival curve showed that there was a significant survival difference between the training cohort and the validation cohort among the three groups (Figure 4). Therefore, the OS prediction model was found to successfully distinguish low, medium, and high all-cause mortality, and the CSS prediction model was found to successfully distinguish low, medium, and high cancer-specific mortality in patients.

Clinical efficacy evaluation of the nomogram

Decision curves were constructed for the OS and CSS nomogram validation cohorts (Figure 5) with threshold probabilities of < 71% and < 72% at 3 years, respectively. With a threshold probability of < 89% and < 82% at 5 years, the OS and CSS nomograms showed a net benefit over the "all treatment" and "no treatment" strategies, respectively. Therefore, the proposed nomogram aids in good clinical decision-making ability. Comparing the nomogram with TNM staging, we can see that its use in clinical decision efficacy is significantly better than that using TNM staging.

DISCUSSION

In this study, the prognostic factors of OS and CSS in the study population were found by analyzing the relevant clinical data, and two nomogram prediction models were constructed. The original intention of the clinical prediction models was to predict the status and prognosis of a disease with readily available practical predictors[14]. We

Table 1 Baseline demographic and clinical characteristics of middle-aged and elderly patients with stage III rectal adenocarcinoma							
	Training cohort (1944)		Validation cohort (829)		Overall (2773)		
Variable	Quantity	SCALE	Quantity	SCALE	Quantity	SCALE	*P* value
Age (yr)							0.596
45-62	1107	56.94%	456	55.01%	1563	56.36%	
62-75	588	30.25%	258	31.12%	846	30.51%	
> 75	249	12.81%	115	13.87%	364	13.13%	
Race							0.425
Other	146	7.51%	66	7.96%	212	7.65%	
White	1511	77.73%	656	79.13%	2167	78.15%	
Black	287	14.76%	107	12.91%	394	14.21%	
Sex							0.051
Female	794	40.84%	305	36.79%	1099	39.63%	
Male	1150	59.16%	524	63.21%	1674	60.37%	
Grade							0.859
I/II	1642	84.47%	698	84.20%	2340	84.39%	
III/IV	302	15.53%	131	15.80%	433	15.61%	
Site							0.692
Rectosigmoid Junction	660	33.95%	275	33.17%	935	33.72%	
Rectum	1284	66.05%	554	66.83%	1838	66.28%	
Stage							0.138
III A	199	10.24%	90	10.86%	289	10.42%	
III B	1394	71.71%	615	74.19%	2009	72.45%	
III C	351	18.06%	124	14.96%	475	17.13%	
Tumor size							0.204
≤ 68	1626	83.64%	677	81.66%	2303	83.05%	
> 68	318	16.36%	152	18.34%	470	16.95%	
CEA							0.558
Low	1074	55.25%	468	56.45%	1542	55.61%	
High	870	44.75%	361	43.55%	1231	44.39%	
TD							0.473
Neg	1481	76.18%	621	74.91%	2102	75.80%	
Pos	463	23.82%	208	25.09%	671	24.20%	
CRM							0.409
Pos	354	18.21%	162	19.54%	516	18.61%	
Neg	1590	81.79%	667	80.46%	2257	81.39%	
PNI							0.197
Neg	1551	79.78%	679	81.91%	2230	80.42%	
Pos	393	20.22%	150	18.09%	543	19.58%	
pLN							0.152
0	508	26.13%	204	24.61%	712	25.68%	
1-6	1212	62.35%	546	65.86%	1758	63.40%	
> 6	224	11.52%	79	9.53%	303	10.93%	

DLNs							0.241
≤ 26	1675	86.16%	728	87.82%	2403	86.66%	
> 26	269	13.84%	101	12.18%	370	13.34%	
LNR							0.246
0-0.12	1146	58.95%	477	57.54%	1623	58.53%	
< 0.31	463	23.82%	221	26.66%	684	24.67%	
> 0.31	335	17.23%	131	15.80%	466	16.80%	

CEA: Carcinoembryonic antigen; TD: Tumor engraftment; CRM: Circumferential resection margin; PNI: Perineural invasion; pLN: Positive lymph nodes; DLNs: Examined lymph nodes; LNR: Lymph node-positive rate.

constructed new nomograms that aid in effectively judging the prognosis of patients using more comprehensive and practical clinical indicators in this study.

In this study, the OS and CSS nomogram prediction models were successfully constructed. The model shows the extent to which different predictors affect the study population and provides an accurate score. In previous studies by Brenner *et al*[15] and Fu *et al*[16], age was a significant prognostic factor, similar to the results of this study. However, we found that race and sex were not significant prognostic factors in this group.

Among the clinical features, previous studies have shown that the stage, grade, and size of the tumor are prognostic factors in patients with rectal cancer[17,18]. In this study population, clinical stage and tumor size were known prognostic factors for OS and CSS. However, grade was not a significant prognostic factor in the current study. Massarweh *et al*[19] showed that CRM is an important prognostic factor for rectal adenocarcinoma and that high CEA levels indicate a poor prognosis, similar to our study results. TD refers specifically to a solitary tumor nodule present within the lymphatic drainage area of the primary tumor. Most TD originate from angiolymphatic invasion, in which there are no identifiable lymph nodes, blood vessels, or neural structures[20]. Wang *et al*[21] concluded that TD did not influence the clinical prognostic impact, unlike in our study and that by Liu *et al*[22] wherein TD was a significant prognostic factor. PNI is an important factor in the prognosis of rectal cancer, and the prognosis of node-negative patients with PNI-positive tumors is significantly lower than that of node-positive patients[23]. We also obtained similar results. In the abovementioned studies on prognostic factors of CRC, the common indicators CEA, TD, CRM, and PNI have clinical utility and are the key to determining prognostic factors of the disease. The nomogram showed that poor prognostic factors were as follows: high preoperative CEA levels, positive TD, positive PNI, and positive CRM. These key predictors have gradually become standard reporting practice in clinicopathological reports, and our nomograms show the important predictive ability and scoring proportion of these indicators.

Lymph node dissection is a crucial step in surgery, and the completeness and number of dissections often affect the prognosis of cancer patients[24,25]. Lymph node status is a key factor used to predict the prognosis of cancer patients and guide postoperative treatment[26]. The higher the pLN, the greater the chance of recurrence, metastasis, and poor prognosis[27]. More DLNs are helpful to clarify the pathological stage of CRC[28,29]. Although DLNs and pLNs influence treatment outcome with survival outcome in patients with rectal adenocarcinoma, DLNs and pLNs were not prognostic factors in our study population. However, we found that LNR played an important role in the prognosis of this group of patients. Although the N-staging system of the 8th edition of the AJCC is the most widely used lymph node staging method, the accuracy of this system remains controversial[30]. Some scholars have proposed that LNR can be used as a prognostic factor in patients with rectal cancer and, thus, constructed a new staging system[31]. LNR synthesizes the effects of DLNs and pLN and avoids the drawbacks of insufficient DLNs. This has a more practical guiding significance in clinical application. We found that LNR was an important factor affecting the independent prognosis of surgical patients in this study population. The LNR-based staging method is promising as a new staging or adjuvant staging method for the AJCC staging system, which can more accurately predict the prognosis of patients with stage III rectal cancer presenting with postoperative recurrence, thus providing an important basis for further treatment and planning of postoperative management. We derived an LNR cutoff of 0.05 and 0.43, but the

Table 2 Univariate Cox regression analysis						
Variable	OS			CSS		
	HR	95%CI	*P* value	HR	95%CI	*P* value
Age (yr)			< 0.001			< 0.001
45-62						
62-75	1.684	1.32-2.148	< 0.001	1.649	1.649-1.255	< 0.001
> 75	3.102	2.372-4.057	< 0.001	2.687	2.687-1.964	< 0.001
Race			0.3			0.4
Other						
White	0.752	0.519-1.091	0.134	0.7399	0.739-0.485	0.163
Black	0.682	0.432-1.076	0.009	0.7474	0.747-0.449	0.262
Sex			0.08			0.3
Female						
Male	1.209	0.9731-1.502	0.087	1.139	1.139-0.8914	0.297
Grade			< 0.001			< 0.001
I/II						
III/IV	1.214	2.02-1.214	< 0.001	1.788	1.788-1.352	< 0.001
Site			0.3			0.4
Rectosigmoid Junction						
Rectum	0.894	0.718-1.114	0.319	0.889	0.889-0.693	0.356
Stage			< 0.001			< 0.001
III A						
III B	2.399	1.446-3.979	< 0.001	3.225	3.225-1.651	< 0.001
III C	3.941	2.316-6.706	< 0.001	6.037	6.037-3.027	< 0.001
Tumor size size			< 0.001			< 0.001
≤ 68						
> 68	1.651	1.279-2.131	< 0.001	1.663	1.663-1.245	< 0.001
CEA			< 0.001			< 0.001
Low						
High	1.724	1.395-2.131	< 0.001	1.576	1.576-1.239	< 0.001
TD			< 0.001			< 0.001
Neg						
Pos	2.009	1.612-2.504	< 0.001	2.072	2.072-1.614	< 0.001
CRM			< 0.001			< 0.001
Neg						
Pos	2.205	1.756-2.77	< 0.001	2.551	2.551-1.981	< 0.001
PNI			< 0.001			< 0.001
Neg						
Pos	2.092	1.665-2.629	< 0.001	2.222	2.222-1.718	<0.001
pLN			< 0.001			<0.001
0						
1-6	1.379	1.039-1.831	0.026	1.711	1.711-1.21	0.002
> 6	2.27	1.594-3.233	< 0.001	2.924	2.924-1.927	< 0.001

DLNs			0.02			0.07
≤ 26						
> 26	1.412	1.072-1.858	0.014	1.348	1.348-0.981	0.066
LNR			< 0.001			< 0.001
0-0.12						
< 0.31	1.405	1.088-1.815	0.009	1.64	1.64-1.228	< 0.001
> 0.31	2.261	1.755-2.914	< 0.001	2.594	2.594-1.944	< 0.001

OS: Overall survival; CCS: Cancer-specific survival; HR: Hazard ratios; CEA: Carcinoembryonic antigen; TD: Tumor engraftment; CRM: Circumferential resection margin; PNI: Perineural invasion; pLN: Positive lymph nodes; DLNs: Examined lymph nodes; LNR: Lymph node-positive rate.

optimal cutoff still needs further confirmation and consensus.

By comprehensively analyzing the effect of predictors on the prognosis of the study population, nomograms based on clinically practical prognostic factors were constructed to assess the three- and five-year CSS and OS of patients. This predictive model has shown excellent predictive ability with clinical decision-making ability, as validated in a number of ways. Moreover, this nomogram was superior to TNM staging in survival prediction. These findings still need further validation in clinical practice.

In this study, we evaluated the prognostic impact of multiple factors on the study population due to the large sample size and long overall follow-up time of patients in multiple centers. Through validation, the constructed model may have good application value and evaluation advantages in clinical practice. First, the prediction model constructed in this study is based on the results of a large sample study of the target population. Second, the data used to construct the prediction model for this study are clinically easy to obtain and have good clinical utility. Third, we found that the optimal cutoff values for some factors differed from previous studies. Instead of using a common cutoff, we need to find relevant cutoff values for different target populations for clinical application.

The study has some limitations. First, the study was retrospective in nature and there was an inherent selection bias due to the strict inclusion and exclusion criteria used. Second, the included patients were treated by different surgeons, which may have an impact due to the different experience and level of surgeons as well as pathologists. Third, data on some novel predictors were not recorded, such as tumor sprouting and molecular changes. Fourth, the current database information does not contain the latest clinical data of patients diagnosed with rectal adenocarcinoma, and other databases are needed to further validate the findings of this paper. Finally, some indicators, including nutritional status, detailed chemoradiotherapy information, and complications, were not recorded in the database.

CONCLUSION

We constructed three-year and five-year CSS and OS nomogram models for middle-aged and elderly patients with stage III rectal adenocarcinoma. Both prediction models demonstrated excellent clinical decision-making ability and survival prediction ability. The results indicate that LNR is more important for the prognosis of the target population than pLN and DLNs. This nomogram can be used for individualized survival prediction in the target population. It is a convenient tool for general practitioners and surgeons as it can help in more accurate evaluation of the patient's status.

Table 3 Multivariate Cox regression analysis.

Variable	OS			CSS		
	HR	95%CI	*P* value	HR	95%CI	*P* value
Age (yr)						
45-62						
62-75	1.773	1.384-2.272	< 0.001	1.731	1.311-2.286	< 0.001
> 75	3.427	2.601-4.516	< 0.001	2.976	2.157-4.106	< 0.001
Grade						
I/II						
III/IV	1.188	0.914-1.545	0.198	1.281	0.96-1.709	0.092
Stage						
III A						
III B	1.680	1.005-2.809	0.047	2.369	1.204-4.662	0.013
III C	1.770	0.968-3.24	0.064	2.960	1.391-6.297	0.005
Tumor size						
≤ 68						
> 68	1.669	1.277-2.18	< 0.001	1.638	1.209-2.219	0.001
CEA						
Low						
High	1.426	1.148-1.772	0.001	1.261	0.986-1.612	0.064
TD						
Neg						
Pos	1.525	1.204-1.933	< 0.001	1.520	1.164-1.985	0.002
CRM						
Neg						
Pos	1.538	1.204-1.963	< 0.001	1.773	1.351-2.328	< 0.001
PNI						
Neg						
Pos	1.550	1.204-1.996	< 0.001	1.520	1.147-2.015	0.004
pLN						
0						
1-6	1.096	0.784-1.532	0.592	1.378	0.925-2.053	0.115
> 6	0.895	0.497-1.612	0.712	1.076	0.57-2.031	0.821
DLNs						
≤ 26						
> 26	1.254	0.928-1.694	0.141			
LNR						
0-0.12						
< 0.31	1.279	0.953-1.717	0.101	1.330	0.963-1.839	0.084
> 0.31	1.871	1.298-2.697	< 0.001	1.809	1.22-2.684	0.003

OS: Overall survival; CCS: Cancer-specific survival; HR: Hazard ratios; CEA: Carcinoembryonic antigen; TD: Tumor engraftment; CRM: Circumferential resection margin; PNI: Perineural invasion; pLN: Positive lymph nodes; DLNs: Examined lymph nodes; LNR: Lymph node-positive rate.
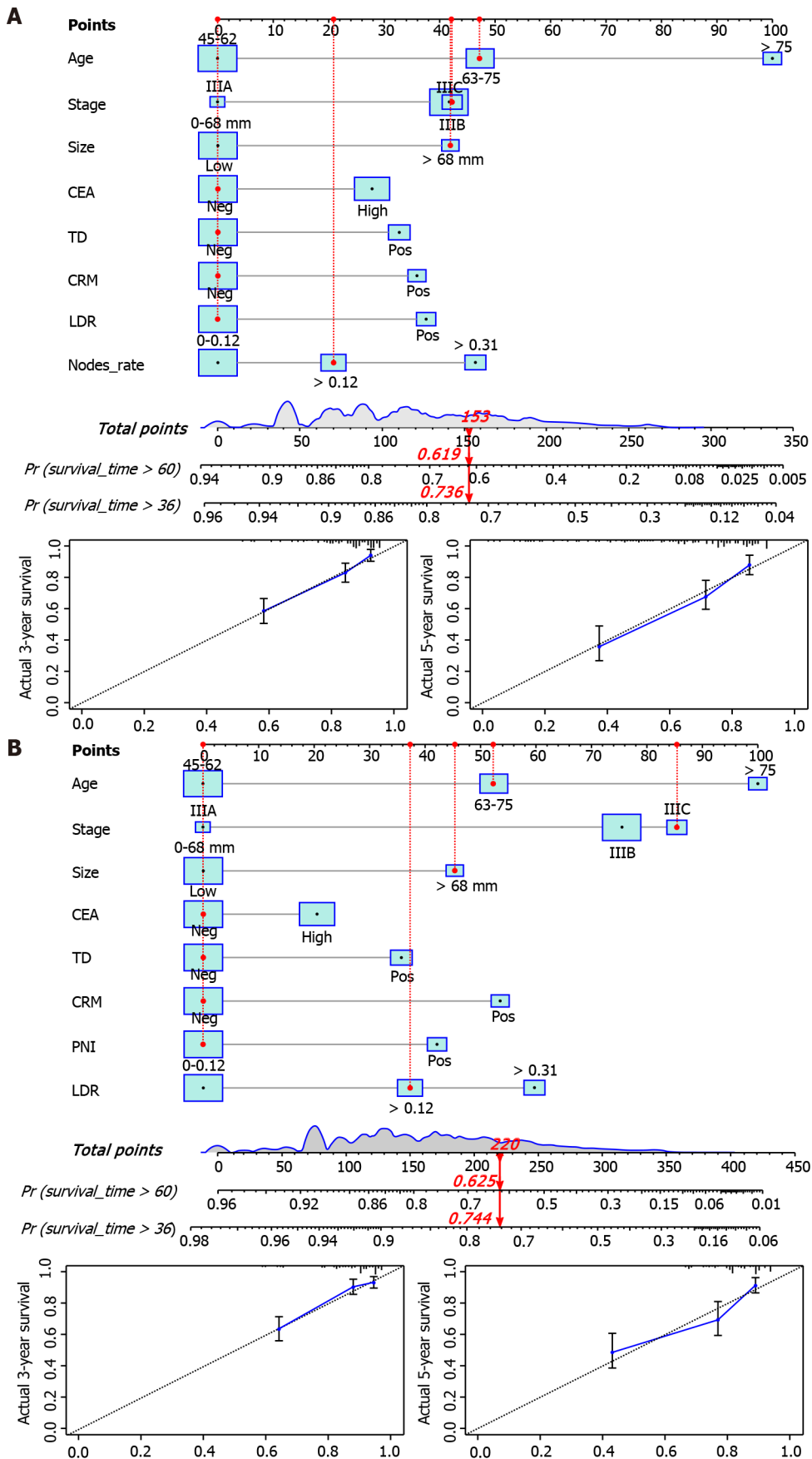
Figure 2 Overall survival nomogram, cancer-specific survival nomogram and verification cohorts calibration curve. The total score obtained by summing the individual scores of the predictors was used to predict the 3- and 5-year survival rates of the patients. The calibration curve showed a high degree of agreement between the predicted and actual values of the overall survival (OS) nomogram and the cancer-specific survival (CSS) nomogram. A: OS nomogram and

calibration curve B: CSS nomogram and calibration curve. (Example: 70-year-old male patient with rectal adenocarcinoma, tumor size 80 mm, circumferential resection margin (-), tumor engraftment (-), carcinoembryonic antigen (-), perineural invasion (-), clinical stage IIIC, LNR: 0.18. The red indicator line in the figure represents the score of patients: OS total score 153, 3-year OS: 0.619, 5-year OS: 0.736; CSS total score 220, 3-year CSS: 0.744, 5-year CSS: 0.625). CEA: Carcinoembryonic antigen; TD: Tumor engraftment; CRM: Circumferential resection margin; PNI: Perineural invasion; LDR: Low-dose-rate.
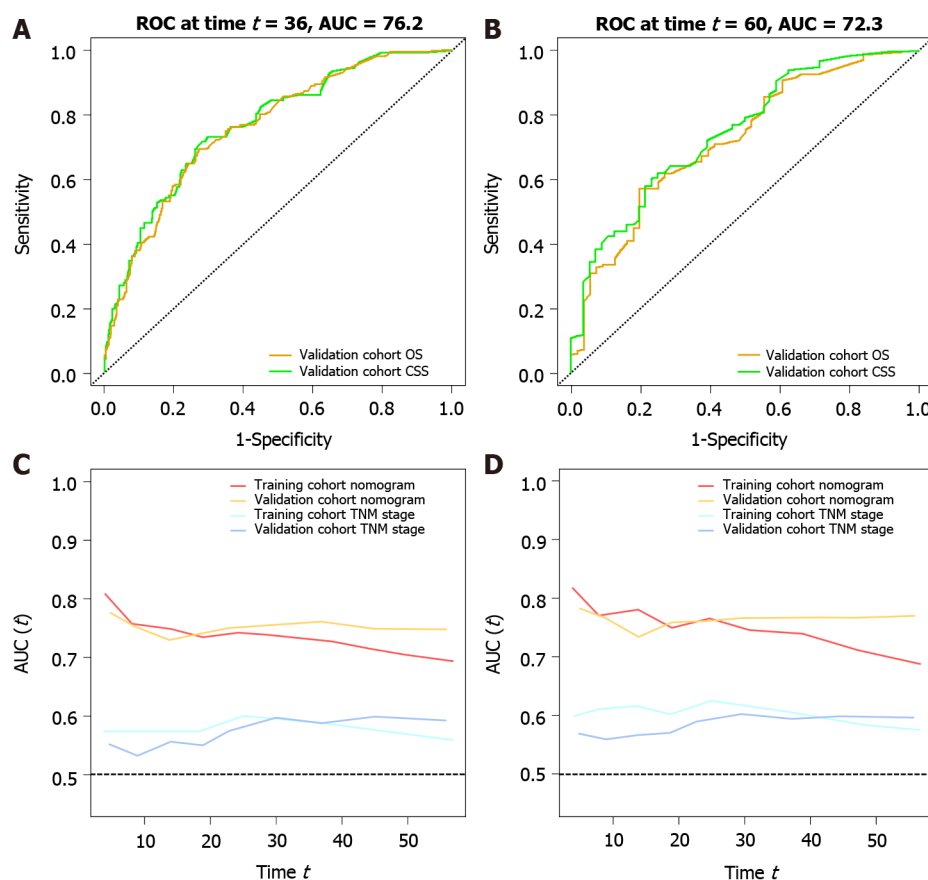

Figure 3 Receiver operating characteristic and time-dependent area under the curve. Receiver operating characteristic (ROC) values for the validation cohort are shown in the figure. Area under the curve (AUC) values of the nomogram *vs* tumor-node-metastases (TNM) stage based on temporal changes are shown. In the training and validation cohorts, the AUC of the nomogram was higher than that of TNM stage. A: Validation cohort overall survival (OS), 3-year survival ROC of cancer-specific survival (CSS) (OS AUC: 0.762; CSS AUC: 0.770); B: Validation cohort OS, 5-year survival ROC of CSS (OS AUC: 0.722; CSS AUC: 0.744); C: Time-dependent AUC curve of OS between the training cohort and validation cohort (nomogram and TNM stage); D: Time-dependent AUC curve of CSS between the training cohort and validation cohort (nomogram and TNM stage). ROC: Receiver operating characteristic; AUC: Area under the curve; OS: Overall survival; CCS: Cancer-specific survival; TNM: Tumor-node-metastases.
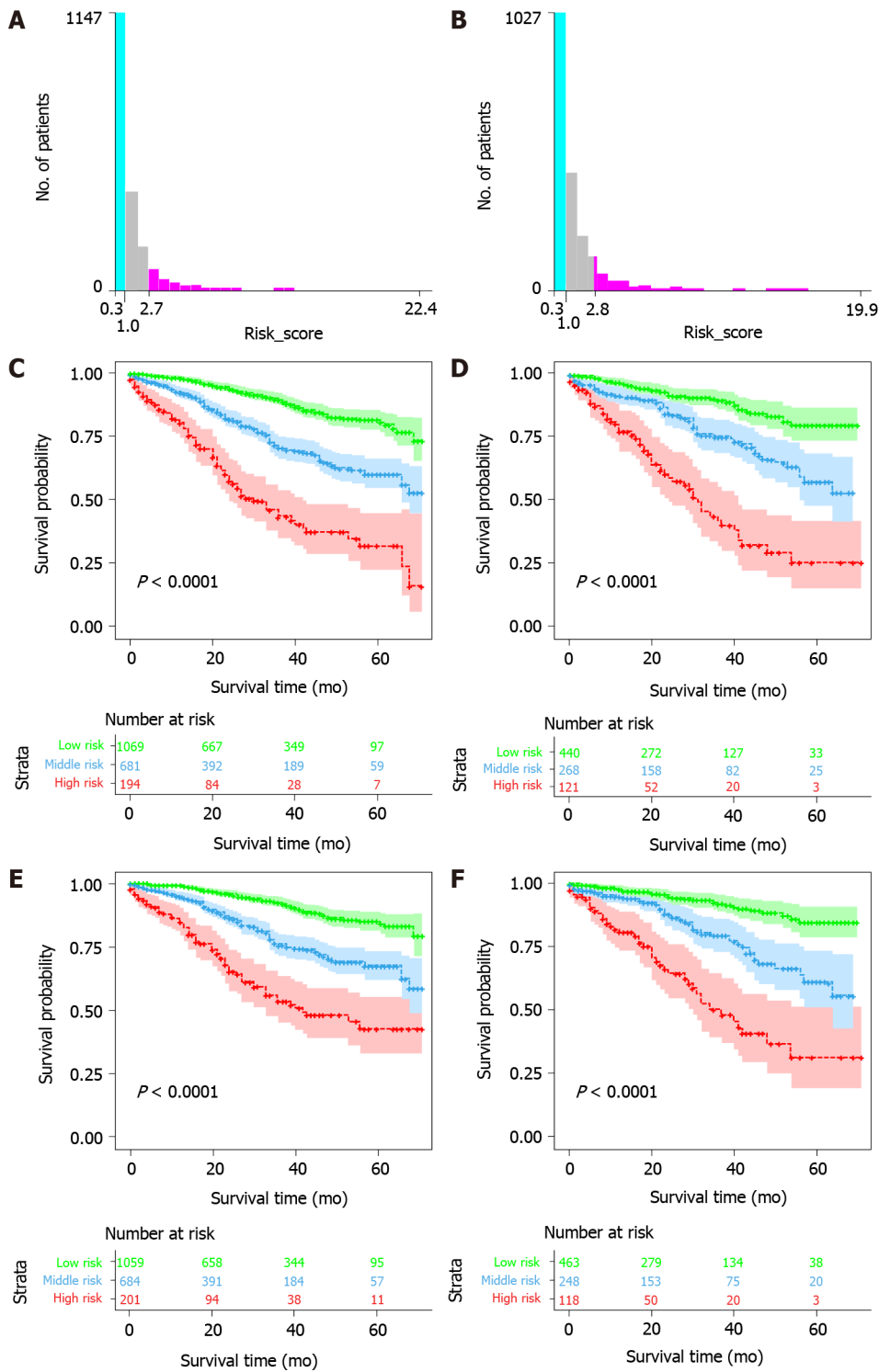
Figure 4 Kaplan-Meier survival curves for low-, medium-, and high-risk groups based on risk scores. The optimal overall survival (OS) risk score cutoffs were 1 and 2.7. Cancer-specific survival (CSS) risk score cutoffs were 1 and 2.8. Significant differences in OS and CSS were observed between low-risk, intermediate-risk, and high-risk patients in the training and validation cohorts. A: OS risk score cutoff value; B: CSS risk score cutoff value; C: Risk-stratified survival curve for OS in the training cohort; D: Risk-stratified survival curve for OS in the validation cohort; E: Risk-stratified survival curve for CSS in the training cohort; F: Risk-stratified survival curve for CSS in the validation cohort.

Figure 5 Decision curve analysis. Plot net benefit *vs* threshold probability. The net benefit was calculated by subtracting the proportion o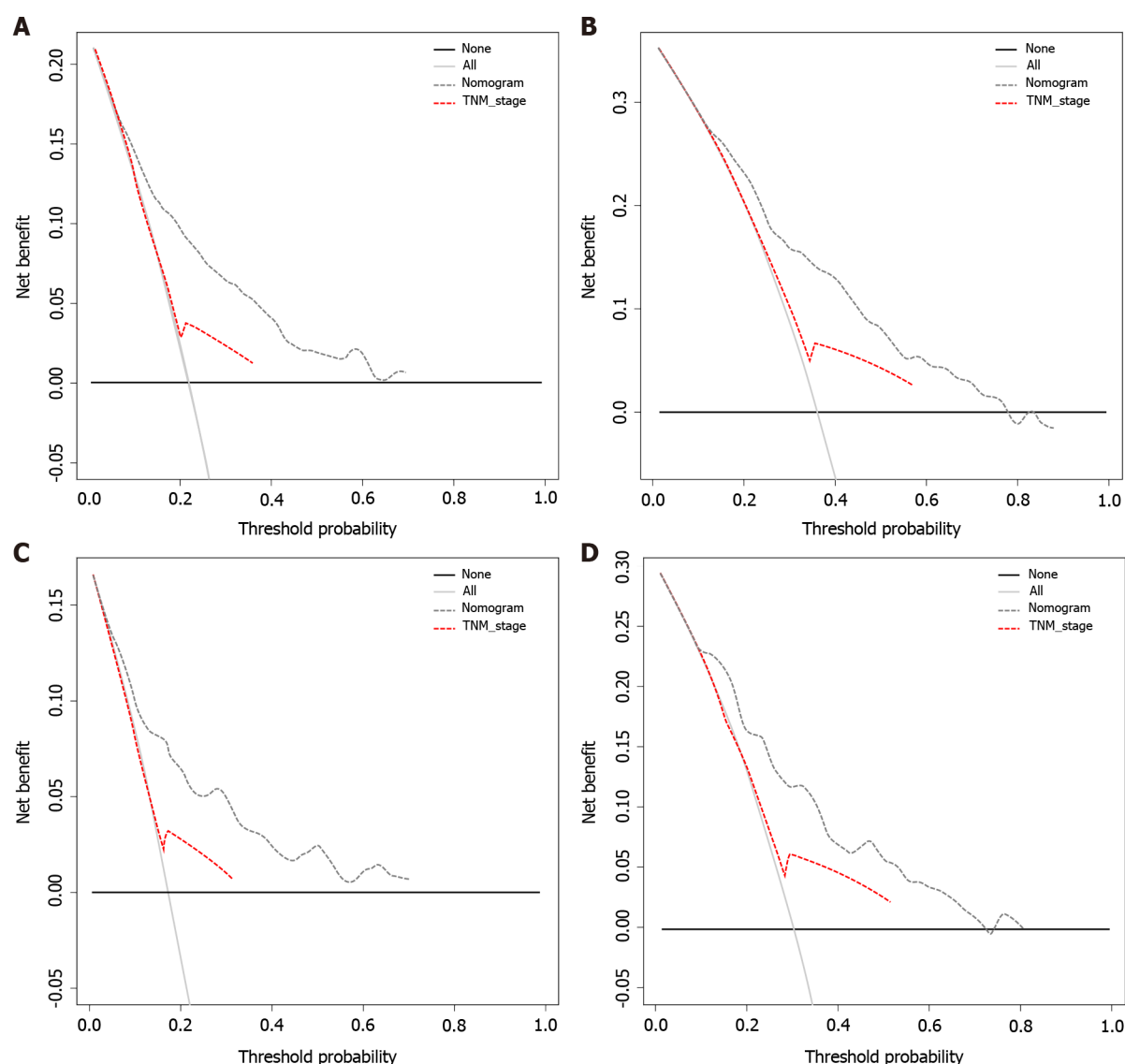f all false-positive patients from the proportion of true positives, weighing the relative harm of abandoning treatment against the adverse consequences of unnecessary treatment. The gray and black lines indicate the net benefit of treating all patients and no patient strategies, respectively. Dashed lines represent the nomograms. The results showed that the nomograms had good decision power in the validation cohort. A: 3 Year overall survival (OS) decision curve analysis (DCA); B: 5 Year OS DCA; C: 3 Year CSS DCA; D: 5 Year CSS DCA. TNM: Tumor-node-metastases.

ARTICLE HIGHLIGHTS

Research background

Patients with colorectal cancer have fewer nomograms for prognosis prediction, and prognostic indicators change with age. Middle-aged and elderly people frequently develop rectal adenocarcinoma, and clinical patients are mainly stage III patients.

Research motivation

Providing a prognostically essential tool for the target population can help physicians make correct clinical decisions regarding related treatments and benefit physicians and patients.

Research objectives

To construct a prognostic nomogram for in the target population, a prognostic tool that can predict OS and CSS in this population. This prognostic tool should have good predictive power and clinical utility.

Research methods

First, patients in the Surveillance, Epidemiology, and End Results Program database were screened according to the inclusion and exclusion criteria; second, the prognostic factors of OS and CSS in the target population were determined by univariate analysis and multivariate analysis; third, a predictor-based clinical survival prediction model was constructed; and fourth, the predictive power and clinical efficacy of the nomogram were verified.

Research results

The 95%CI was 0.719 (0.690-0.749) and 0.733 (0.702-0.74) in the OS and CSS nomogram prediction model training groups, respectively, compared with 0.739 (0.696-0.782) and 0.750 (0.701-0.800) in the validation group. In the validation, the area under the receiver operating characteristic curve (AUC) for the OS and CSS nomograms for the three-year survival rate was 0.762 and 0.770, respectively, while the AUC for the five-year survival rate was 0.722 and 0.744, respectively. Predictive models can distinguish all-cause mortality from cancer-specific mortality in patients with different risk grades. Time-dependent AUC and decision curve analysis showed that the nomogram had excellent clinical prediction and decision-making capabilities, significantly better than the tumor-node-metastases staging system.

Research conclusions

Lymph node-positive rate is more important for the prognosis of the target population than the number of positive lymph nodes and number of examined lymph nodes. The nomogram survival prediction model we constructed is helpful in assessing the clinical prognosis of this population and providing guidance for the optimization of clinical treatment plans.

Research perspectives

The nomograms constructed can be used for individualized survival prediction in the target population. It is a convenient tool for general practitioners and surgeons as it can help to evaluate the patient's status more accurately and provide help for the relevant treatment of the target population in clinical practice.

ACKNOWLEDGEMENTS

REFERENCES

1   **Xie YH**, Chen YX, Fang JY. Comprehensive review of targeted therapy for colorectal cancer. *Signal Transduct Target Ther* 2020; **5**: 22 [PMID: 32296018 DOI: 10.1038/s41392-020-0116-z]

2   **Zhu X**, Wang D, Lin Q, Wu G, Yuan S, Ye F, Fan Q. Screening key lncRNAs for human rectal adenocarcinoma based on lncRNA-mRNA functional synergistic network. *Cancer Med* 2019; **8**: 3875-3891 [PMID: 31116002 DOI: 10.1002/cam4.2236]

3   **Jia J**, Zhang P, Gou M, Yang F, Qian N, Dai G. The Role of Serum CEA and CA19-9 in Efficacy Evaluations and Progression-Free Survival Predictions for Patients Treated with Cetuximab Combined with FOLFOX4 or FOLFIRI as a First-Line Treatment for Advanced Colorectal Cancer. *Dis Markers* 2019; **2019**: 6812045 [PMID: 30805037 DOI: 10.1155/2019/6812045]

4   **Reinau D**, Surber C, Jick SS, Meier CR. Epidemiology of basal cell carcinoma in the United Kingdom: incidence, lifestyle factors, and comorbidities. *Br J Cancer* 2014; **111**: 203-206 [PMID: 24874476 DOI: 10.1038/bjc.2014.265]

5   **Song W**, Miao DL, Chen L. Nomogram for predicting survival in patients with pancreatic cancer. *Onco Targets Ther* 2018; **11**: 539-545 [PMID: 29416354 DOI: 10.2147/OTT.S154599]

6   **Du Q**, Wang Y, Guan S, Hu C, Li M, Zhou L, Zhang M, Chen Y, Mei X, Sun J, Zhou Y. Retraction Note: The diagnostic nomogram of platelet-based score models for hepatic alveolar echinococcosis and atypical liver cancer. *Sci Rep* 2020; **10**: 16405 [PMID: 32994550 DOI: 10.1038/s41598-020-73758-x]

7   **Balachandran VP**, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol* 2015; **16**: e173-e180 [PMID: 25846097 DOI: 10.1016/S1470-2045(14)71116-7]

8   **Duggan MA**, Anderson WF, Altekruse S, Penberthy L, Sherman ME. The Surveillance, Epidemiology, and End Results (SEER) Program and Pathology: Toward Strengthening the Critical Relationship. *Am J Surg Pathol* 2016; **40**: e94-e102 [PMID: 27740970 DOI:

10.1097/PAS.0000000000000749]

9   **Wang Y**, Li J, Wen S, Yang X, Zhang Y, Wang Z, Zhang Z. CHRM3 is a novel prognostic factor of poor prognosis in patients with endometrial carcinoma. *Am J Transl Res* 2015; **7**: 902-911 [PMID: 26175851]

10   **Guan X**, Chen W, Jiang Z, Liu Z, Miao D, Hu H, Zhao Z, Yang R, Wang X. Exploration of the Optimal Minimum Lymph Node Count after Colon Cancer Resection for Patients Aged 80 Years and Older. *Sci Rep* 2016; **6**: 38901 [PMID: 27941906 DOI: 10.1038/srep38901]

11   **Greaves P**, Clear A, Coutinho R, Wilson A, Matthews J, Owen A, Shanyinde M, Lister TA, Calaminici M, Gribben JG. Expression of FOXP3, CD68, and CD20 at diagnosis in the microenvironment of classical Hodgkin lymphoma is predictive of outcome. *J Clin Oncol* 2013; **31**: 256-262 [PMID: 23045593 DOI: 10.1200/JCO.2011.39.9881]

12   **Shah RV**, Truong QA, Gaggin HK, Pfannkuche J, Hartmann O, Januzzi JL Jr. Mid-regional pro-atrial natriuretic peptide and pro-adrenomedullin testing for the diagnostic and prognostic evaluation of patients with acute dyspnoea. *Eur Heart J* 2012; **33**: 2197-2205 [PMID: 22645194 DOI: 10.1093/eurheartj/ehs136]

13   **Hijazi Z**, Oldgren J, Lindbäck J, Alexander JH, Connolly SJ, Eikelboom JW, Ezekowitz MD, Held C, Hylek EM, Lopes RD, Yusuf S, Granger CB, Siegbahn A, Wallentin L; ARISTOTLE and RE-LY Investigators. A biomarker-based risk score to predict death in patients with atrial fibrillation: the ABC (age, biomarkers, clinical history) death risk score. *Eur Heart J* 2018; **39**: 477-485 [PMID: 29069359 DOI: 10.1093/eurheartj/ehx584]

14   **Zhou ZR**, Wang WW, Li Y, Jin KR, Wang XY, Wang ZW, Chen YS, Wang SJ, Hu J, Zhang HN, Huang P, Zhao GZ, Chen XX, Li B, Zhang TS. In-depth mining of clinical data: the construction of clinical prediction model with R. *Ann Transl Med* 2019; **7**: 796 [PMID: 32042812 DOI: 10.21037/atm.2019.08.63]

15   **Purim O**, Gordon N, Brenner B. Cancer of the colon and rectum: potential effects of sex-age interactions on incidence and outcome. *Med Sci Monit* 2013; **19**: 203-209 [PMID: 23511310 DOI: 10.12659/MSM.883842]

16   **Fu WJ**. Racial-Sex Disparities--A Challenging Battle Against Cancer Mortality in the USA. *J Racial Ethn Health Disparities* 2015; **2**: 158-166 [PMID: 26863334 DOI: 10.1007/s40615-014-0059-6]

17   **Li X**, Yu W, Liang C, Xu Y, Zhang M, Ding X, Cai X. INHBA is a prognostic predictor for patients with colon adenocarcinoma. *BMC Cancer* 2020; **20**: 305 [PMID: 32293338 DOI: 10.1186/s12885-020-06743-2]

18   **Cai D**, Huang ZH, Yu HC, Wang XL, Bai LL, Tang GN, Peng SY, Li YJ, Huang MJ, Cao GW, Wang JP, Luo YX. Prognostic value of preoperative carcinoembryonic antigen/tumor size in rectal cancer. *World J Gastroenterol* 2019; **25**: 4945-4958 [PMID: 31543685 DOI: 10.3748/wjg.v25.i33.4945]

19   **Massarweh NN**, Hu CY, You YN, Bednarski BK, Rodriguez-Bigas MA, Skibber JM, Cantor SB, Cormier JN, Feig BW, Chang GJ. Risk-adjusted pathologic margin positivity rate as a quality indicator in rectal cancer surgery. *J Clin Oncol* 2014; **32**: 2967-2974 [PMID: 25092785 DOI: 10.1200/JCO.2014.55.5334]

20   **Mirkin KA**, Kulaylat AS, Hollenbeak CS, Messaris E. Prognostic Significance of Tumor Deposits in Stage III Colon Cancer. *Ann Surg Oncol* 2018; **25**: 3179-3184 [PMID: 30083832 DOI: 10.1245/s10434-018-6661-9]

21   **Wang S**, Guan X, Ma M, Zhuang M, Ma T, Liu Z, Chen H, Jiang Z, Chen Y, Wang G, Wang X. Reconsidering the prognostic significance of tumour deposit count in the TNM staging system for colorectal cancer. *Sci Rep* 2020; **10**: 89 [PMID: 31919408 DOI: 10.1038/s41598-019-57041-2]

22   **Liu F**, Zhao J, Li C, Wu Y, Song W, Guo T, Chen S, Cai S, Huang D, Xu Y. The unique prognostic characteristics of tumor deposits in colorectal cancer patients. *Ann Transl Med* 2019; **7**: 769 [PMID: 32042785 DOI: 10.21037/atm.2019.11.69]

23   **Liebig C**, Ayala G, Wilks J, Verstovsek G, Liu H, Agarwal N, Berger DH, Albo D. Perineural invasion is an independent predictor of outcome in colorectal cancer. *J Clin Oncol* 2009; **27**: 5131-5137 [PMID: 19738119 DOI: 10.1200/JCO.2009.22.4949]

24   **Salehi B**, Fokou PVT, Yamthe LRT, Tali BT, Adetunji CO, Rahavian A, Mudau FN, Martorell M, Setzer WN, Rodrigues CF, Martins N, Cho WC, Sharifi-Rad J. Phytochemicals in Prostate Cancer: From Bioactive Molecules to Upcoming Therapeutic Agents. *Nutrients* 2019; **11** [PMID: 31261861 DOI: 10.3390/nu11071483]

25   **Zhou YY**, Zhang QW, Huang J, Yan XL, Chen C, Xu FF, Du XJ, Jin R. Additional lymphadenectomy might not improve survival of patients with resectable metastatic colorectal adenocarcinoma of T4 stage, proximal location, poor/undifferentiation, or N3/N4 stages: a large population-based study. *J Cancer* 2018; **9**: 2428-2435 [PMID: 30026839 DOI: 10.7150/jca.24675]

26   **Brenner H**, Kloor M, Pox CP. Colorectal cancer. *Lancet* 2014; **383**: 1490-1502 [PMID: 24225001 DOI: 10.1016/S0140-6736(13)61649-9]

27   **Govindarajan A**, Gönen M, Weiser MR, Shia J, Temple LK, Guillem JG, Paty PB, Nash GM. Challenging the feasibility and clinical significance of current guidelines on lymph node examination in rectal cancer in the era of neoadjuvant therapy. *J Clin Oncol* 2011; **29**: 4568-4573 [PMID: 21990400 DOI: 10.1200/JCO.2011.37.2235]

28   **Compton CC**, Greene FL. The staging of colorectal cancer: 2004 and beyond. *CA Cancer J Clin* 2004; **54**: 295-308 [PMID: 15537574 DOI: 10.3322/canjclin.54.6.295]

29   **Kidner TB**, Ozao-Choy JJ, Yoon J, Bilchik AJ. Should quality measures for lymph node dissection in

colon cancer be extrapolated to rectal cancer? *Am J Surg* 2012; **204**: 843-7; discussion 847 [PMID: 22981183 DOI: 10.1016/j.amjsurg.2012.05.003]

30   **Lykke J**, Jess P, Roikjaer O; Danish Colorectal Cancer Group. The prognostic value of lymph node ratio in a national cohort of rectal cancer patients. *Eur J Surg Oncol* 2016; **42**: 504-512 [PMID: 26856955 DOI: 10.1016/j.ejso.2016.01.012]

31   **Persiani R**, Cananzi FC, Biondi A, Paliani G, Tufo A, Ferrara F, Vigorita V, D'Ugo D. Log odds of positive lymph nodes in colon cancer: a meaningful ratio-based lymph node classification system. *World J Surg* 2012; **36**: 667-674 [PMID: 22270984 DOI: 10.1007/s00268-011-1415-x]