World Journal of

Gastrointestinal Surgery

Contents

Contents

ABOUT COVER

Editorial Board Member of *World Journal of Gastrointestinal Surgery*, Osman Nuri Dilek, FACS, Professor, Department of Surgery, Division of Hepatopancreatobiliary Surgery, Izmir Katip Çelebi University School of Medicine, İzmir 35150, Turkey. osmannuridilek@gmail.com

AIMS AND SCOPE

The primary aim of *World Journal of Gastrointestinal Surgery* (*WJGS*, *World J Gastrointest Surg*) is to provide scholars and readers from various fields of gastrointestinal surgery with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

WJGS mainly publishes articles reporting research results and findings obtained in the field of gastrointestinal surgery and covering a wide range of topics including biliary tract surgical procedures, biliopancreatic diversion, colectomy, esophagectomy, esophagostomy, pancreas transplantation, and pancreatectomy, *etc*.

INDEXING/ABSTRACTING

The *WJGS* is now abstracted and indexed in Science Citation Index Expanded (SCIE, also known as SciSearch®), Current Contents/Clinical Medicine, Journal Citation Reports/Science Edition, PubMed, PubMed Central, Reference Citation Analysis, China National Knowledge Infrastructure, China Science and Technology Journal Database, and Superstar Journals Database. The 2023 Edition of Journal Citation Reports® cites the 2022 impact factor (IF) for *WJGS* as 2.0; IF without journal self cites: 1.9; 5-year IF: 2.2; Journal Citation Indicator: 0.52; Ranking: 113 among 212 journals in surgery; Quartile category: Q3; Ranking: 81 among 93 journals in gastroenterology and hepatology; and Quartile category: Q4.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: *Rui-Rui Wu*; **Production Department Director:** *Xiang Li*; **Editorial Office Director:** *Jia-Ru Fan*.

ORIGINAL ARTICLE

Retrospective Study

Advantage of log odds of positive lymph nodes in prognostic evaluation of patients with early-onset colon cancer

Heng-Bo Xia, Chen Chen, Zhi-Xing Jia, Liang Li, A-Man Xu

Heng-Bo Xia, Chen Chen, A-Man Xu, Department of General Surgery, The First Affiliated Hospital of Anhui Medical University, Hefei 230032, Anhui Province, China

Heng-Bo Xia, Chen Chen, A-Man Xu, Department of General Surgery, Anhui Public Health Clinical Center, Hefei 230032, Anhui Province, China

Zhi-Xing Jia, Liang Li, Department of Surgery, The Second People's Hospital of Hefei, Hefei 230011, Anhui Province, China

Corresponding author: A-Man Xu, MD, Doctor, Department of General Surgery, The First Affiliated Hospital of Anhui Medical University, Jixi Road, Hefei 230032, Anhui Province, China. xuaman@ahmu.edu.cn

Abstract

BACKGROUND

Colon cancer (CC) is one of the most common cancers of the digestive tract, the third most common cancer worldwide, and the second most common cause of cancer-related deaths. Previous studies have demonstrated a higher risk of lymph node metastasis (LNM) in young patients with CC. It might be reasonable to treat patients with early-onset locally advanced CC with extended lymph node dissection. However, few studies have focused on early-onset CC (ECC) patients with LNM. At present, the methods of predicting and evaluating the prognosis of ECC patients with LNM are controversial.

AIM

To compare the prognostic values of four lymph node staging indices and establish the best nomogram for patients with ECC.

METHODS

From the data of patients with CC obtained from the Surveillance, Epidemiology, and End Results (SEER) database, data of young patients with ECC (≤ 50 years old) was screened. Patients with unknown data were excluded from the study, while the remaining patients were included. The patients were randomly divided into a training group (train) and a testing group (test) in the ratio of 7:3, while building the model. The model was constructed by the training group and verified by the testing group. Using multiple Cox regression models to compare the prediction efficiency of LNM indicators, nomograms were built based on the best model selected for overall survival (OS) and cause-specific survival (CSS). In

the two groups, the performance of the nomogram was evaluated by constructing a calibration plot, time-dependent area under the curve (AUC), and decision curve analysis. Finally, the patients were grouped based on the risk score predicted by the prognosis model, and the survival curve was constructed after comparing the survival status of the high and low-risk groups.

RESULTS
Records of 26922 ECC patients were screened from the SEER database. N classification, positive lymph nodes (PLN), lymph node ratio (LNR) and log odds of PLN (LODDS) were considered to be independent predictors of OS and CSS. In addition, independent risk factors for OS included gender, race, marital status, primary site, histology, grade, T, and M classification, while the independent prognostic factors for CSS included race, marital status, primary site, grade, T, and M classification. The prediction model including LODDS is composed of minimal Akaike information criterion, maximal concordance indexes, and AUCs. Factors including gender, race, marital status, primary site, histology, grade, T, M classification, and LODDS were integrated into the OS nomogram, while race, marital status, primary site, grade, T, M classification, and LODDS were included into the CSS nomogram. The nomogram representing both cohorts had been successfully verified in terms of prediction accuracy and clinical practicability.

CONCLUSION
LODDS is superior to N-stage, PLN, and LNR of ECC. The nomogram containing LODDS might be helpful in tumor evaluation and clinical decision-making, since it provides an appropriate prediction of ECC.

Key Words: Early-onset colon cancer; Log odds of positive lymph nodes; Lymph node metastasis; Nomogram; Prognosis; Surveillance, Epidemiology, and End Results

Core Tip: Few studies have focused on early-onset colon cancer (ECC) patients with lymph node metastasis. This study compared the prognostic value of four lymph node staging indexes. It is shown that log odds of positive lymph nodes (LODDS) is superior to N-stage, positive lymph nodes, and lymph node ratio of ECC. Subsequently, the nomogram containing LODDS was established and provides an appropriate prediction of ECC, which may be helpful in tumor evaluation and clinical decision-making.

INTRODUCTION

Colon cancer (CC) is one of the most common cancers of the digestive tract, the third most common cancer in the world, and the second leading cause of cancer-related deaths[1]. Although medical technology and prevention policies in addition to advances in colonoscopy screening and treatment have enabled a reduction in the morbidity and mortality associated with colorectal cancer (CRC) in elderly patients[2], an opposite trend has been observed in young people under the age of 50[3]. The incidence of early-onset CRC (EOCRC), defined as CRC diagnosed before the age of 50 is increasing worldwide. A previously conducted study has reported that about 11% of CRC cases registered in the National Cancer Database had been diagnosed in adults between the ages 18-49 years[4]. Similarly, the latest data from Europe indicates that the incidence of CRC in subjects aged 4-9 years, 1-6 years, and 20-29 years has increased by 30.39%, 40.49%, and 2004.20%, respectively in last seven to nine years[5]. Although the prevalence of CRC is still relatively limited in the younger population (0.12%), the alarming increase in EOCRC patients cannot be ignored[1]. Compared with late-onset CRC, most early-onset CC (ECC) patients tend to ignore the occult incidence of CRC, which leads to a late-stage diagnosis and poor prognosis.

The first choice for the treatment for locally advanced CC is radial resection. Colectomy has been shown to be associated with a greater survival advantage[6], and complete mesocolic excision has become the preferred treatment option for colorectal surgeons[7]. In addition, with the emergence and development of endoscopic technology, CRC surgery is further benefited due to the application of laparoscopy and robot-assisted laparoscopy[8]. With regards to lymph node dissection, there is a consensus that, specimens after radical surgery for CC should contain at least 12 regional lymph nodes in accordance with the recommendations in the NCCN guidelines[9]. However, previously conducted studies have demonstrated that a higher risk of lymph node metastasis (LNM) is observed in young patients with CC[10], because of which extended lymph node dissection might be a more reasonable choice for treating patients

with early-onset locally advanced CC. One of the main causes of poor prognosis and frequent recurrence is LNM. Large population-based cohort studies have demonstrated a high incidence of LNM in patients with ECC. These studies have also indicated that 60% of patients with stages III or IV-ECC develop LNM[11]. Previously conducted studies have considered factors such as the location of the primary tumor, histopathological grade, tumor-node-metastasis (TNM) stage, carcinoembryonic antigen level, and tumor size for evaluating the prognosis[12]. However, studies focusing on ECC patients with LNM are rare. The current methods of predicting and evaluating the prognosis of ECC patients with LNM are controversial[13]. Previous prognostic models ignored a different number of anatomical regional lymph nodes, which could compromise the accuracy of the prognostic predictions. In addition, along with individual differences in the pattern of regional LNM, there is a lack of consensus on the optimum extent of intraoperative lymph node dissection. Therefore, new LNM indicators are urgently required to develop better prognostic nomograms that would enable the prediction of overall survival (OS) and cause-specific survival (CSS) in patients with ECC.

In the past, several scholars have proposed several prognostic factors of lymph nodes including the number of positive lymph nodes (PLN)[14], the number of negative lymph nodes[15,16], and lymph node ratio (LNR) for estimating the prognosis of ECC patients[17-19]. In recent years, the log odds of PLN (LODDS) have proven reliable for many tumor types[20,21]. Some studies have found LODDS to be more efficient in predicting the prognosis of CC patients compared with American joint committee of cancer (AJCC)-N classification and LNR[22-24]. However, prediction of CSS were unexplored in these studies, nor did they further establish clinical prognostic nomograms.

The study analyzed the data of the Surveillance, Epidemiology, and End Results (SEER) database to compare the predictive values of different lymph nodes indicators in ECC patients. Subsequently, establishing a new nomogram including LODDS for predicting OS and CSS, and successfully verifying it on the testing group.

MATERIALS AND METHODS

Data source

Data of CC patients was obtained from the SEER database of the National Cancer Institute program, which is one of the most representative oncology databases. Since the data was downloaded from public databases, ethical approval for this study was exempted. The exclusion criteria for data extraction included: (1) Patients aged < 18 years or > 50 years at the time of diagnosis; (2) Patients with lymph nodes (ELN) and PLN without resection of lymph nodes or unclear lymph nodes examination; (3) Patients with unclear T-stage, N-stage, stage, and grade; and (4) Patients with unclear survival data. Finally, 26922 patients diagnosed with ECC following radical resection of CC were recruited for the study, and they were randomly divided into a training group (*n* = 18845, 70%) and a testing group (*n* = 8077, 30%) in a ratio of 7:3 (Figure 1).
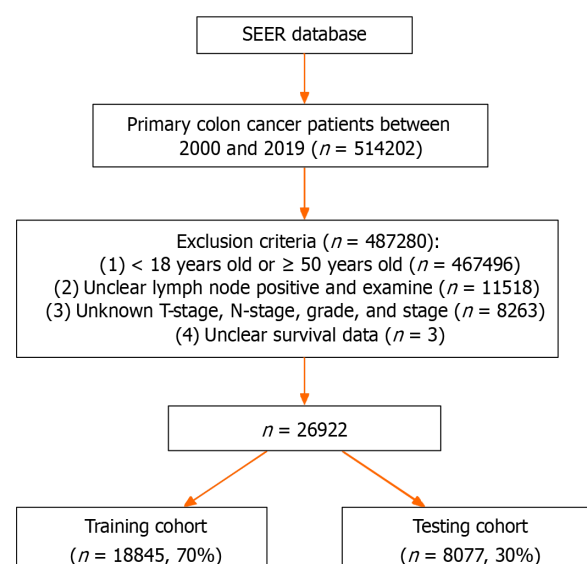
DOI: 10.4240/wjgs.v15.i11.2430 **Copyright** ©The Author(s) 2023.

Figure 1 The flowchart of data collection and grouping for early-onset colon cancer patients. 26922 patients with ECC after radical resection of colon cancer were enrolled and randomly divided into a training group (n = 18845, 70%) and a testing group (n = 8077, 30%).

Measurement of variables

Patient variables including age, sex, primary site and pathology, tumor size and grade, TNM classification, number of regional ELN and PLN, survival time, and status were collected for the study. Race was divided into four subgroups: White, black, other, and unknown. The "Other" subgroup included American Indian, AK Native, Asian, and Pacific Islanders. Marital status was divided into two subgroups, namely married and others/unknown. The primary site was

divided into right colon (transverse colon), left colon, and large intestine (NOS). Histology was divided into three subgroups: Adenomas/adenocarcinomas, cystic/mucinous/serous neoplasms, and others. The grade was divided into four subgroups, namely well, moderate, poor, and undifferentiated. The stage was divided into: Stage 0, stage I, stage II, stage III, and stage IV. T-stage was divided into five subgroups, namely T0, T1, T2, T3, and T4. N-stage was divided into three subgroups: N0, N1, and N2. M-stage was divided into three subgroups: M0, and M1. The size was divided into three subgroups: < 5 cm, ≥ 5 cm, and unknown.

Staging was determined in accordance with the sixth edition of the AJCC staging system. LNR was defined as the ratio of the amount of PLN to ELN. LODDS is determined in accordance with the reaction: $\log[(PLN + 0.5)/(ELN - PLN + 0.5)]$. 0.5 was added to both numerator and the denominator, to avoid division by the zero error. The primary and secondary endpoints are OS and CSS, which have been indicated in the SEER database as "COD to site recording" and "SEER cause-specific death classification", respectively.

Independence and prediction performance comparison of lymph nodes status indicators

Clinicopathological predictors with univariate Cox regression analysis were selected through the "survival" R package of OS and CSS in the training group. Each lymph nodes status factor (including N classification, PLN, LNR, and LODDS), was integrated with other risk variables into a multivariate regression model for further evaluation of the predicted value. The predicted value for the univariate analysis was $P < 0.05$. The above models respectively use the Akaike information criterion (AIC) was employed by the above models as the stop rule, which adopted backward step-by-step selection (through the "MASS" R packet). The model with the minimum AIC was selected as the best model. The prediction efficiencies of these filtering models with different lymph nodes factors were compared by the "risk regression" R package using AIC, bootstrapped concordance index (C-index), and the area under the curve (AUC).

Construction and verification of nomograms

The nomograms, which were developed by integrating variables with the highest precision from the filtering model, were used to predict the OS and CSS in the training group (through the "rms" R package). The C-index, AUC, and calibration plots in the training group and testing group were used to evaluate the efficiency of the nomogram. The "ggDCA" R package was used to evaluate the net income and clinical performance of the nomograms based on a decision curve analysis (DCA) generated in advance.

Survival risk classifiers established by nomograms

The "nomogram formula" R package was used to apply the multivariate Cox regression formula of OS and CSS nomograms formed in the training group to the patients in the two groups. All patients were divided into a high-risk group and a low-risk group based on the total score calculated by the "survminer" R package. The survival differences of OS and CSS between the two risk groups were evaluated using the Kaplan-Meier method.

Statistical analysis

The counting data were expressed as an example (%), and the comparison between groups was conducted using χ^2 test. The measurement data of normal distribution were expressed by mean ± SD, and the comparison between groups was done using the independent sample *t*-test. The measurement data of non-normal distribution were expressed by median (interquartile range), and the comparison between the groups was done with the Mann-Whitney *U* nonparametric test with $P < 0.05$ indicating statistical significance. The R software was used for all statistical analyses.

RESULTS

Patient characteristics and survival

The characteristics of the training group and testing group of the patients have been depicted in Table 1. The median follow-up times of the training group and the testing group in the whole SEER database were 51 mo [95% confidence interval (CI): 21-106] and 50 mo (95%CI: 21-103), respectively. Additionally, there was insignificant difference in the indices of the training group and the testing group (*P* > 0.05).

Prognostic analyses for OS and CSS

Table 2 depicts the detailed results of the univariate Cox regression analysis in the training group. The important risk factors for OS were gender, race, marital status, primary site, histology, grade, T-stage, N-stage, M-stage, PLN, LNR, and LODDS. The important prognostic factors for CSS included race, marital status, primary site, histology, grade, T-stage, N-stage, M-stage, PLN, LNR, and LODDS.

We further generated prognostic models, including different lymph nodes indicators after conducting amultivariate analysis. In short, as evident from Tables 3 and 4, N classification, PLN, LNR, and LODDS are independent risk factors for OS and CSS, respectively. In addition, the independent prognostic factors for OS were gender, race, marital status, primary site, histology, grade, T, and M classification; while the independent risk factors for CSS were race, marital status, primary site, grade, T, and M classification.

Comparison of N classification, PLN, LNR, and LODDS

The comparison of lymph nodes status indicators in the training group is shown in Table 5. The C-index of the filter

Table 1 Clinical and pathologic characteristics of patients with early-onset colon cancer in two cohorts

Characteristics	Overall	Training set	Testing set	P value
	n = 26922	*n* = 18845	*n* = 8077	
Age	44.00 (39.00, 47.00)	44.00 (39.00, 47.00)	44.00 (39.00, 47.00)	0.730
LN examined	19.00 (14.00, 27.00)	19.00 (14.00, 27.00)	19.00 (14.00, 27.00)	0.959
LN positive	1.00 (0.00, 4.00)	1.00 (0.00, 4.00)	1.00 (0.00, 4.00)	0.800
LNR	0.03 (0.00, 0.19)	0.03 (0.00, 0.19)	0.03 (0.00, 0.20)	0.851
LODDS	-2.66 (-3.66, -1.29)	-2.66 (-3.61, -1.30)	-2.71 (-3.66, -1.24)	0.881
Survival months	51.00 (21.00, 105.00)	51.00 (21.00, 106.00)	50.00 (21.00, 103.00)	0.169
Gender (%)				0.856
Female	13367 (49.7)	9364 (49.7)	4003 (49.6)	
Male	13555 (50.3)	9481 (50.3)	4074 (50.4)	
Race (%)				0.898
White	19659 (73.0)	13783 (73.1)	5876 (72.7)	
Blake	4050 (15.0)	2829 (15.0)	1221 (15.1)	
Others	2985 (11.1)	2073 (11.0)	912 (11.3)	
Unknown	228 (0.8)	160 (0.8)	68 (0.8)	
Marital status (%)				0.383
Married	15296 (56.8)	10674 (56.6)	4622 (57.2)	
Others/unknown	11626 (43.2)	8171 (43.4)	3455 (42.8)	
Primary site (%)				0.783
Right colon	13053 (48.5)	9139 (48.5)	3914 (48.5)	
Left colon	13342 (49.6)	9330 (49.5)	4012 (49.7)	
Large intestine, NOS	527 (2.0)	376 (2.0)	151 (1.9)	
Histology (%)				0.995
Adenomas/adenocarcinomas	23406 (86.9)	16384 (86.9)	7022 (86.9)	
Cystic/mucinous/serous neoplasms	3324 (12.3)	2326 (12.3)	998 (12.4)	
Others	192 (0.7)	135 (0.7)	57 (0.7)	
Grade (%)				0.055
Well	2775 (10.3)	1992 (10.6)	783 (9.7)	
Moderately	18265 (67.8)	12708 (67.4)	5557 (68.8)	
Poorly	5094 (18.9)	3603 (19.1)	1491 (18.5)	
Undifferentiated	788 (2.9)	542 (2.9)	246 (3.0)	
Stage (%)				0.198
Stage 0	204 (0.8)	136 (0.7)	68 (0.8)	
Stage I	3939 (14.6)	2792 (14.8)	1147 (14.2)	
Stage II	7360 (27.3)	5103 (27.1)	2257 (27.9)	
Stage III	9806 (36.4)	6914 (36.7)	2892 (35.8)	
Stage IV	5613 (20.8)	3900 (20.7)	1713 (21.2)	
T stage (%)				0.437
T0	210 (0.8)	142 (0.8)	68 (0.8)	
T1	2595 (9.6)	1845 (9.8)	750 (9.3)	
T2	2801 (10.4)	1985 (10.5)	816 (10.1)	

T3	15121 (56.2)	10537 (55.9)	4584 (56.8)	
T4	6195 (23.0)	4336 (23.0)	1859 (23.0)	
N stage (%)				0.259
N0	12410 (46.1)	8645 (45.9)	3765 (46.6)	
N1	7886 (29.3)	5576 (29.6)	2310 (28.6)	
N2	6626 (24.6)	4624 (24.5)	2002 (24.8)	
M stage (%)				0.351
M0	21309 (79.2)	14945 (79.3)	6364 (78.8)	
M1	5613 (20.8)	3900 (20.7)	1713 (21.2)	
Size (%)				0.871
< 5 cm	9238 (34.3)	6476 (34.4)	2762 (34.2)	
≥ 5 cm	9179 (34.1)	6434 (34.1)	2745 (34.0)	
Unknown	8505 (31.6)	5935 (31.5)	2570 (31.8)	

LNR: Lymph node ratio; LODDS: Log odds of positive lymph node.

Table 2 Univariate Cox regression analyses for predicting overall survival and cause-specific survival in the training cohort

Characteristics	OS		CSS	
	HR (95%CI)	*P* value	HR (95%CI)	*P* value
Age	1.004 (0.999-1.008)	0.090	1.000 (0.995-1.005)	0.896
Gender	1.096 (1.040-1.156)	0.001[a]	1.058 (0.996-1.123)	0.067
Race	1.050 (1.013-1.090)	0.009[a]	1.073 (1.030-1.118)	0.001[a]
Marital status	1.330 (1.262-1.402)	< 0.001[a]	1.282 (1.208-1.362)	< 0.001[a]
Primary site	0.926 (0.881-0.974)	0.003[a]	0.926 (0.875-0.980)	0.007[a]
Histology	1.501 (1.409-1.599)	< 0.001[a]	1.498 (1.394-1.609)	< 0.001[a]
Grade	1.696 (1.631-1.764)	< 0.001[a]	1.774 (1.698-1.853)	< 0.001[a]
Stage	3.267 (3.149-3.389)	< 0.001[a]	4.102 (3.922-4.290)	< 0.001[a]
T stage	2.258 (2.169-2.350)	< 0.001[a]	2.566 (2.448-2.690)	< 0.001[a]
N stage	2.347 (2.271-2.426)	< 0.001[a]	2.653 (2.552-2.757)	< 0.001[a]
M stage	7.380 (6.988-7.793)	< 0.001[a]	8.790 (8.262-9.351)	< 0.001[a]
LN examined	0.989 (0.987-0.991)	< 0.001[a]	0.988 (0.985-0.991)	< 0.001[a]
LN positive	1.085 (1.082-1.087)	< 0.001[a]	1.088 (1.085-1.090)	< 0.001[a]
LNR	14.557 (13.403-15.811)	< 0.001[a]	17.876 (16.324-19.575)	< 0.001[a]
LODDS	1.562 (1.541-1.584)	< 0.001[a]	1.630 (1.605-1.656)	< 0.001[a]
Size	1.006 (0.970-1.043)	0.759	0.999 (0.959-1.040)	0.955

[a]$P < 0.05$.

OS: Overall survival; CSS: Cause-specific survival; HR: Hazard ratio; CI: Confidence interval; LNR: Lymph node ratio; LODDS: Log odds of positive lymph node.

model containing LODDS was higher than that of N classification, PLN, and LNR; when compared with the above prognostic models. In addition, the selected model containing LODDS has the least AIC. Additionally, the 1-year, 3-year, 5-year, and 10-year AUCs of the selected models, including LODDS, were higher than those of other models. To sum up, the selected model including LODDS is more efficient in the predictions of OS and CSS, and LODDS might be the strongest predictor of N classification, PLN, and LNR.

Table 3 Multivariate Cox regression analyses for predicting overall survival in the training cohort

Characteristics	N-stage		PLN		LNR		LODDS	
	HR (95%CI)	*P* value	HR (95%CI)	*P* value	HR (95%CI)	*P* value	HR (95%CI)	*P* value
Gender	1.131 (1.073-1.193)	< 0.001[a]	1.103 (1.046-1.164)	< 0.001[a]	1.120 (1.062-1.181)	< 0.001[a]	1.120 (1.062-1.181)	< 0.001[a]
Race	1.046 (1.008-1.086)	0.017[a]	1.058 (1.020-1.098)	0.003[a]	1.048 (1.010-1.088)	0.013[a]	1.045 (1.007-1.085)	0.019[a]
Marital status	1.300 (1.233-1.371)	< 0.001[a]	1.289 (1.223-1.360)	< 0.001[a]	1.307 (1.239-1.378)	< 0.001[a]	1.310 (1.243-1.382)	< 0.001[a]
Primary site	0.901 (0.856-0.948)	< 0.001[a]	0.934 (0.888-0.983)	0.008[a]	0.891 (0.847-0.938)	< 0.001[a]	0.873 (0.830-0.919)	< 0.001[a]
Histology	1.184 (1.108-1.264)	< 0.001[a]	1.089 (1.019-1.164)	0.012[a]	1.120 (1.048-1.197)	0.001[a]	1.125 (1.052-1.202)	0.001[a]
Grade	1.256 (1.205-1.309)	< 0.001[a]	1.284 (1.232-1.338)	< 0.001[a]	1.238 (1.188-1.290)	< 0.001[a]	1.230 (1.180-1.281)	< 0.001[a]
T-stage	1.433 (1.373-1.495)	< 0.001[a]	1.485 (1.425-1.549)	< 0.001[a]	1.456 (1.396-1.518)	< 0.001[a]	1.439 (1.380-1.501)	< 0.001[a]
M-stage	4.645 (4.378-4.928)	< 0.001[a]	5.074 (4.783-5.382)	< 0.001[a]	4.484 (4.220-4.764)	< 0.001[a]	4.254 (4.003-4.519)	< 0.001[a]
N-stage	1.636 (1.578-1.697)	< 0.001[a]	/	/	/	/	/	/
PLN	/	/	1.048 (1.044-1.052)	< 0.001[a]	/	/	/	/
LNR	/	/	/	/	4.736 (4.303-5.213)	< 0.001[a]	/	/
LODDS	/	/	/	/	/	/	1.309 (1.288-1.330)	< 0.001[a]

[a]*P* < 0.05.

HR: Hazard ratio; CI: Confidence interval; PLN: Positive lymph node; LNR: Lymph node ratio; LODDS: Log odds of positive lymph node.

Table 4 Multivariate Cox regression analyses for predicting cause-specific survival in the training cohort

Characteristics	N-stage		PLN		LNR		LODDS	
	HR (95%CI)	*P* value	HR (95%CI)	*P* value	HR (95%CI)	*P* value	HR (95%CI)	*P* value
Race	1.066 (1.023-1.111)	0.003	1.080 (1.036-1.126)	< 0.001	1.068 (1.025-1.114)	0.002	1.066 (1.022-1.111)	0.003
Marital status	1.252 (1.178-1.329)	< 0.001	1.240 (1.167-1.317)	< 0.001	1.258 (1.185-1.336)	< 0.001	1.262 (1.189-1.341)	< 0.001
Primary site	0.885 (0.836-0.938)	< 0.001	0.925 (0.873-0.979)	0.008	0.876 (0.826-0.928)	< 0.001	0.858 (0.809-0.909)	< 0.001
Histology	1.155 (1.071-1.245)	< 0.001	1.044 (0.967-1.126)	0.272	1.074 (0.996-1.159)	0.065	1.077 (0.999-1.162)	0.054
Grade	1.270 (1.212-1.331)	< 0.001	1.311 (1.252-1.373)	< 0.001	1.256 (1.200-1.316)	< 0.001	1.244 (1.188-1.303)	< 0.001
T-stage	1.546 (1.471-1.626)	< 0.001	1.609 (1.532-1.690)	< 0.001	1.574 (1.499-1.654)	< 0.001	1.551 (1.477-1.630)	< 0.001
M-stage	5.132 (4.800-5.487)	< 0.001	5.772 (5.399-6.170)	< 0.001	5.017 (4.684-5.373)	< 0.001	4.716 (4.404-5.051)	< 0.001
N-stage	1.783 (1.710-1.860)	< 0.001	/	/	/	/	/	/
PLN	/	/	1.050 (1.046-1.054)	< 0.001	/	/	/	/
LNR	/	/	/	/	5.304 (4.773-5.895)	< 0.001	/	/
LODDS	/	/	/	/	/	/	1.345 (1.321-1.368)	< 0.001

HR: Hazard ratio; CI: Confidence interval; PLN: Positive lymph node; LNR: Lymph node ratio; LODDS: Log odds of positive lymph node.

Construction and validation of nomograms

In this study, the nomogram was based on the selected model containing LODDS in the training group. As a result, the final nomogram for predicting the OS included the gender, race, marital status, primary site, histology, grade, T, M classification, and LODDS in the (Figure 2); while the CSS nomogram included factors such as race, marital status, primary site, grade, T, M classification, and LODDS (Figure 3). The calibration plots of the two groups are shown in the figure, which demonstrates the consistency of predicted observations of OS and CSS with the actual observations. The time-dependent AUC values of the OS nomograms (Figure 4) and CSS (Figure 5) show more stable accuracy and better prediction efficiency. DCA, which has more advantages over AUC, is a new method for evaluating alternative prognostic strategies. The DCA of nomogram is more beneficial compared to the TNM staging system, indicating that it has better clinical application value than TNM staging. The detailed C-index of the nomogram in each group was evaluated along with the 1-year, 3-year, 5-year, and 10-year AUC values. The results demonstrate the reliability and clinical practicability of the prognostic nomograms.
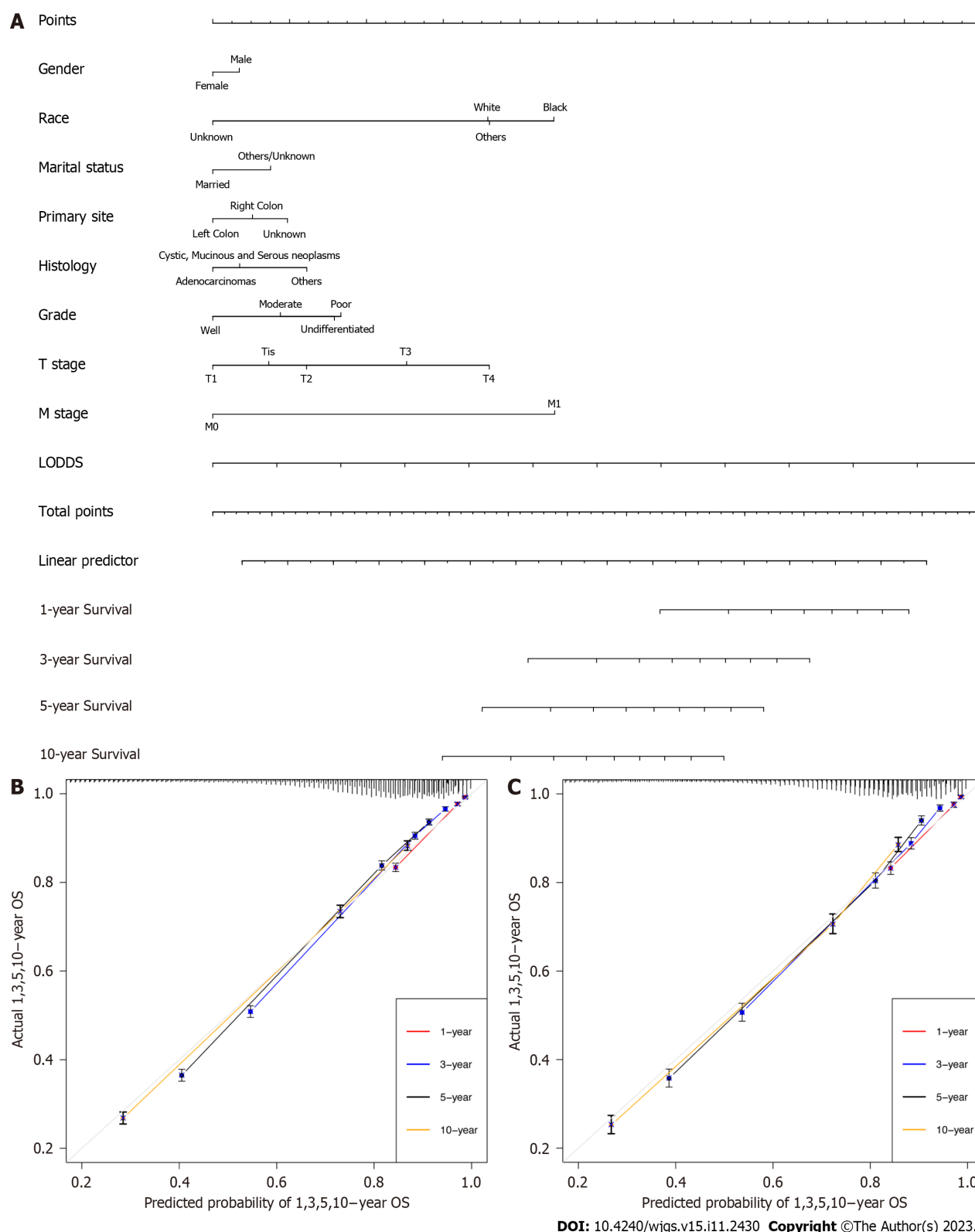
Figure 2 Nomogram for overall survival of early-onset colon cancer patients. A: Prediction for 1-, 3-, 5- and 10-year overall survival of nomogram; B and C: Calibration plots for 1-, 3- 5- and 10-year in training (B), internal validation (C). LODDS: Log odds of positive lymph node; OS: Overall survival.

Survival risk classifiers based on nomograms

To further verify the performance of the nomogram, the patients were divided based on the total scores calculated by OS and CSS nomograms into high- and low-risk groups. Kaplan-Meier curves demonstrated a significant difference in the survival outcomes of the risk classifiers of OS and CSS between the two cohorts (Figure 6).
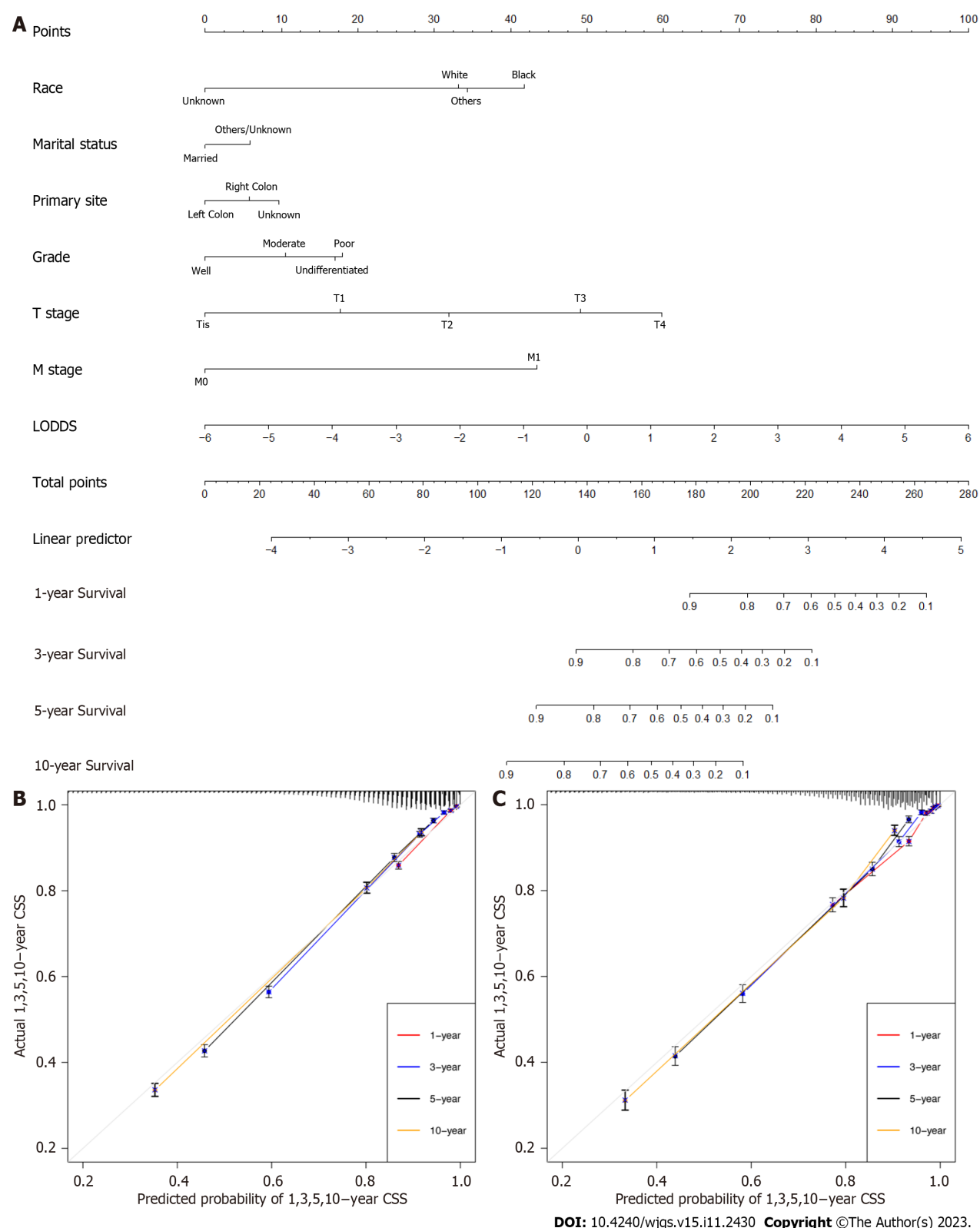
Figure 3 Nomogram for cause-specific survival of early-onset colon cancer patients. A: Prediction for 1-, 3-, 5- and 10-year cause-specific survival of the nomogram; B and C: Calibration plots for 1-, 3- 5- and 10-year in training (B), internal validation (C). ECC: LODDS: Log odds of positive lymph node; CSS: Cause-specific survival.

DISCUSSION

The screening of CC in the elderly is getting better due to the development and popularity of colonoscopy in recent years. CC is now detected in the early stages of tumorigenesis, enabling early intervention, and subsequently better prognosis for patients. However, studies have indicated that CRC in young patients is more invasive than that in the elderly population[25]. Despite the rapid developments in medical technology, ECC remains a malignant tumor of the digestive

Table 5 Prognostic efficiency of different lymph node status indicators in the training cohort

Endpoint	Filtered model	C-index	AIC	AUC			
				1-yr	3-yr	5-yr	10-yr
OS	N-stage	0.799	96473.30	0.835	0.844	0.842	0.818
	LN-positive	0.796	96726.02	0.835	0.845	0.838	0.811
	LNR	0.802	96299.43	0.840	0.849	0.842	0.817
	LODDS	0.806	96143.26	0.842	0.850	0.845	0.820
CSS	N-stage	0.826	74358.02	0.862	0.871	0.869	0.855
	LN-positive	0.822	74674.34	0.861	0.872	0.863	0.846
	LNR	0.829	74253.59	0.868	0.876	0.870	0.853
	LODDS	0.834	74084.86	0.870	0.877	0.872	0.857

AIC: Akaike information criterion; AUC: Area under the curve; C-index: Concordance index; OS: Overall survival; CSS: Cause-specific survival.
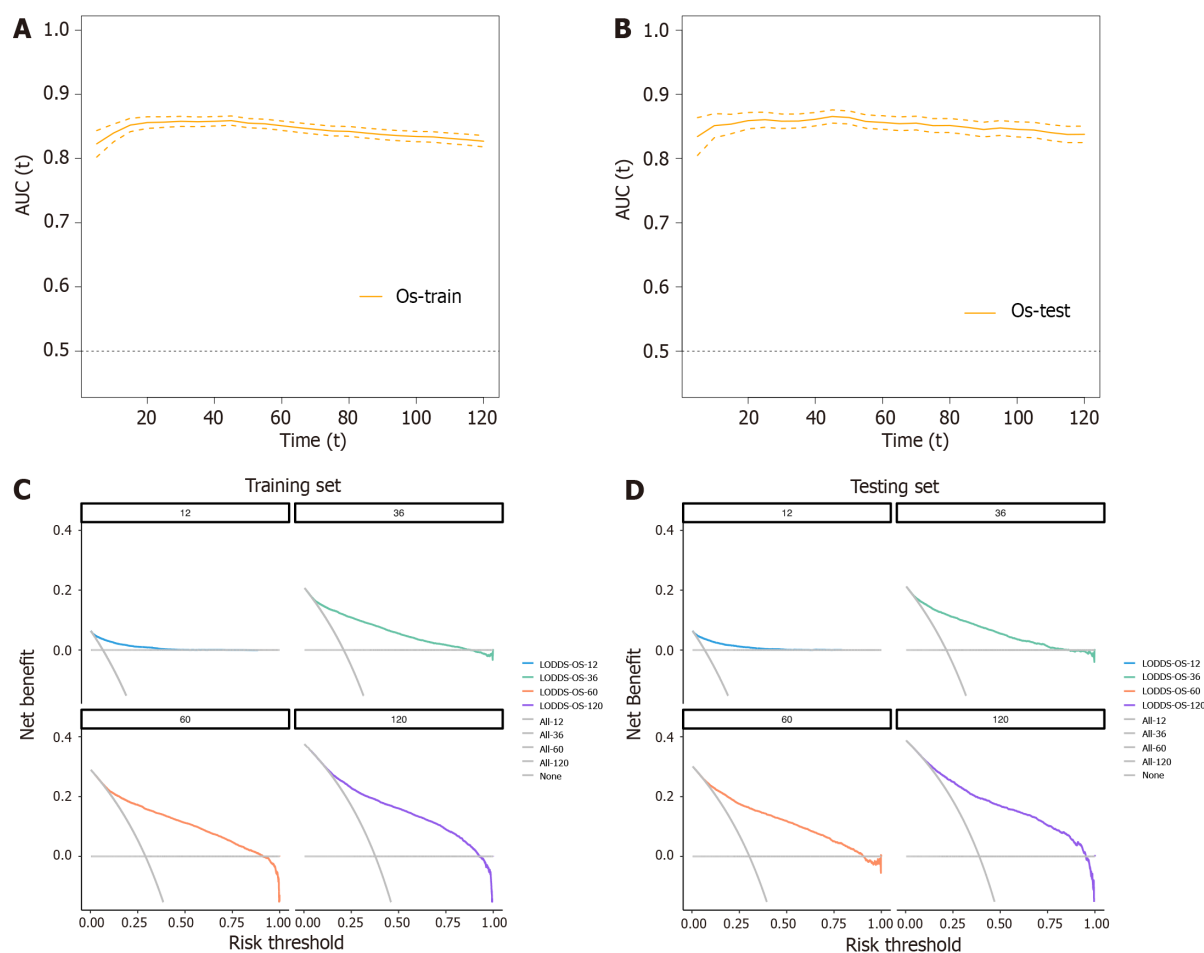
DOI: 10.4240/wjgs.v15.i11.2430 **Copyright** ©The Author(s) 2023.

Figure 4 Evaluation of the nomograms for overall survival of early-onset colon cancer patients with area under the curve and decision curve analysis. A and B: The time-dependent area under the curve for overall survival (OS) in the training (A) and testing (B) cohort; C and D: Decision curves for predicting 1-, 3-, 5- and 10-year OS in training (C) and testing (D) cohort. LODDS: Log odds of positive lymph node; OS: Overall survival; AUC: Area under the curve.

tract and is associated with a poor prognosis because of its location and difficulty in detection. Although the TNM staging system is the most widely used system for prognosis evaluation and determination of the course of treatment for patients with CC, it is associated with several hidden defects limiting its application. Related studies indicate the importance of LNM for the prognosis of ECC[26]; however, the N classification based on the AJCC staging system[27] is not accurate enough to evaluate LNM. Therefore, new lymph nodes status indicators are urgently required to evaluate lymph nodes
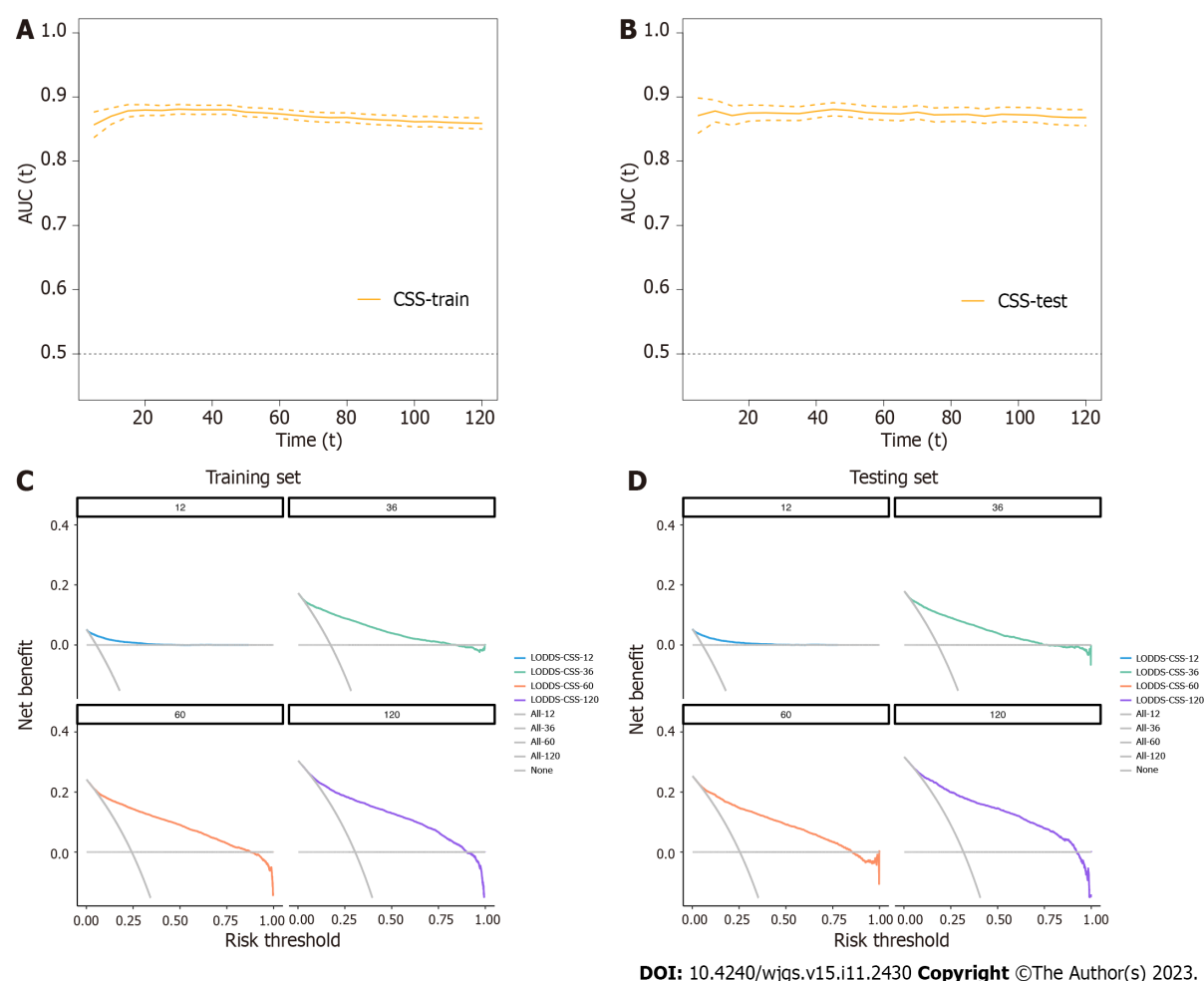
Figure 5 Evaluation of the nomograms for cause-specific survival of early-onset colon cancer patients with area under the curve and decision curve analysis. A and B: The time-dependent area under the curve for cause-specific survival (CSS) in the train (A) and testing (B) cohort; C and D: Decision curves for predicting 1-, 3-, 5- and 10-year CSS in training (C) and testing (D) cohort. LODDS: Log odds of positive lymph node; CSS: Cause-specific survival; AUC: Area under the curve.

involvement and to stratify ECC patients for individualized treatment.

Several studies have proposed modified lymph nodes state factors including PLN[28], LNR[29], and LODDS[30] to predict the prognosis of ECC patients, but there are no uniform results to confirm, which one is the best. One study has demonstrated the reliability of LODDS in predicting the prognosis of elderly patients with CRC[31]. Our study demonstrated that the four lymph nodes indices are independent prognostic factors of OS and CSS in patients with ECC, and further compared the predictions between N classification, PLN, LNR, and LODDS. The results indicate that the LODDS model can be considered to be the best prognostic model for OS and CSS, since it encompasses minimum AIC, maximum C-index, and AUC, The results indicate that LODDS is better for predicting the prognosis of ECC patients. However, to ensure its clinical applicability, we constructed two nomograms combined with LODDS, using the training group data to predict the OS and CSS of ECC patients, and then verified the accuracy of the nomogram using the testing group data. The calibration curve demonstrates stable linearity and appropriate validity of the nomogram, and the calculated C-index and AUC are the highest in the two groups. With regards to clinical utility, the DCA curve reveals consistently large net benefits of the nomogram over a wide range of thresholds, leading us to trust the satisfactory applicability of the nomogram in predicting the survival of ECC patients. To sum up, our nomogram has better prediction accuracy and clinical effectiveness compared with AJCC and other lymph nodes state systems.

In addition, the risk classifiers of OS and CSS have been established according to the total score of the nomogram, and the patients with ECC have been divided into different risk groups. The results demonstrate the poor survival rate of the two high-risk groups in each cohort. It is worth noting that the high-risk group had a higher matrix score, which was consistent with previous studies on ECC patients.

Although our research has proved that the prediction model including LODDS has obvious advantages in the prediction of OS and CSS, there are still some limitations. First, the presence of some unknown indicators may reduce the predictive ability of the model. Second, use of data from only a single database (SEER) may reduce the credibility of the model. Third, the study was a retrospective study and more prospective and multicenter studies are needed to verify the prognostic value of LODDS. Finally, the details of the surgical approach, such as the degree of lymph nodes anatomy at a specific lymph node level, have not been recorded in detail, and further research is warranted. Despite these limitations, our study has successfully demonstrated better predictive values of LODDS, and included it in the prognostic
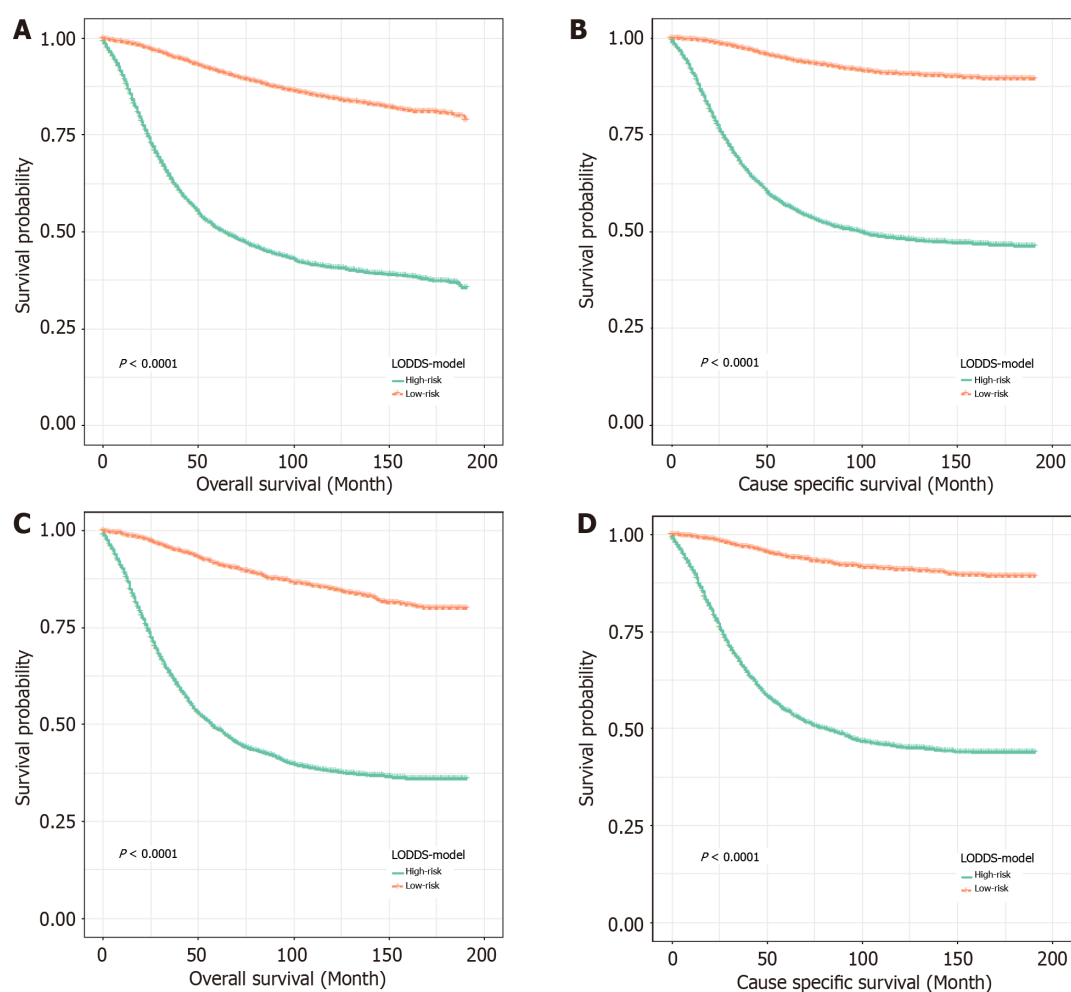
Figure 6 Kaplan-Meier analyses for early-onset colon cancer patients classified by nomograms. A and B: Kaplan-Meier curves for overall survival (OS) (A) and cause-specific survival (CSS) (B) in the training cohort; C and D: Kaplan-Meier curves for OS (C) and CSS (D) in the testing cohort. LODDS: Log odds of positive lymph node.

nomograms of OS and CSS in patients with ECC for the first time.

CONCLUSION

Our study confirmed that LODDS is more accurate than other LNM indicators in predicting the prognosis of ECC patients and established a new nomogram containing LODDS to predict OS and CSS. The applicability of the nomogram was successfully verified in the testing group. The nomogram can help physicians to design a more accurate treatment plan and personalized follow-up management for ECC patients. It is worthy of further clinical promotion.

ARTICLE HIGHLIGHTS

Research background

Colon cancer (CC) is one of the most common cancers of the digestive tract, the third most common cancer worldwide, and the second most common cause of cancer-related deaths. A higher risk of lymph node metastasis (LNM) in young patients with CC. It might be reasonable to treat patients with early-onset locally advanced CC with extended lymph node dissection. However, few studies have focused on early-onset CC (ECC) patients with LNM.

Research motivation

To compare the predictive values of different LN indicators in ECC patients.

Research objectives

The prognostic values of four lymph node staging indices were compared. And the best nomogram for patients with ECC was established.

Research methods

The patients obtained from the Surveillance, Epidemiology, and End Results database were randomly divided into a training group and a testing group. The model was constructed by the training group and verified by the testing group. Using multiple Cox regression models to compare the prediction efficiency of LNM indicators, nomograms were built based on the best model selected for overall survival (OS) and cause-specific survival (CSS). In the two groups, the performance of the nomogram was evaluated by constructing a calibration plot, time-dependent area under the curve (AUC), and decision curve analysis. Finally, the patients were grouped based on the risk score predicted by the prognosis model, and the survival curve was constructed after comparing the survival status of the high and low-risk groups.

Research results

Log odds of PLN (LODDS) were considered to be independent predictors of OS and CSS. The prediction model including LODDS is composed of minimal Akaike information criterion, maximal concordance indexes, and AUCs. The nomograms of OS and CSS were constructed, which representing both cohorts had been successfully verified in terms of prediction accuracy and clinical practicability.

Research conclusions

LODDS is superior to N-stage, PLN, and LNR of ECC. The nomogram based on LODDS might be helpful in tumor evaluation and clinical decision-making, since it provides an appropriate prediction of ECC.

Research perspectives

The nomogram containing LODDS may be helpful in tumor evaluation and clinical decision-making.

FOOTNOTES

REFERENCES

1   **Akimoto N**, Ugai T, Zhong R, Hamada T, Fujiyoshi K, Giannakis M, Wu K, Cao Y, Ng K, Ogino S. Rising incidence of early-onset colorectal cancer - a call to action. *Nat Rev Clin Oncol* 2021; **18**: 230-243 [PMID: 33219329 DOI: 10.1038/s41571-020-00445-1]

2   **Siegel RL**, Miller KD, Fedewa SA, Ahnen DJ, Meester RGS, Barzi A, Jemal A. Colorectal cancer statistics, 2017. *CA Cancer J Clin* 2017; **67**: 177-193 [PMID: 28248415 DOI: 10.3322/caac.21395]

3   **Sung H**, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021; **71**: 209-249 [PMID: 33538338 DOI: 10.3322/caac.21660]

4   **Virostko J**, Capasso A, Yankeelov TE, Goodgame B. Recent trends in the age at diagnosis of colorectal cancer in the US National Cancer Data Base, 2004-2015. *Cancer* 2019; **125**: 3828-3835 [PMID: 31328273 DOI: 10.1002/cncr.32347]

5   **Vuik FE**, Nieuwenburg SA, Bardou M, Lansdorp-Vogelaar I, Dinis-Ribeiro M, Bento MJ, Zadnik V, Pellisé M, Esteban L, Kaminski MF, Suchanek S, Ngo O, Májek O, Leja M, Kuipers EJ, Spaander MC. Increasing incidence of colorectal cancer in young adults in Europe over the last 25 years. *Gut* 2019; **68**: 1820-1826 [PMID: 31097539 DOI: 10.1136/gutjnl-2018-317592]

6   **Li Y**, Liu W, Zhou Z, Ge H, Zhao L, Liu H, Song X, Wang D, Pei Q, Tan F. Development and validation of prognostic nomograms for early-onset locally advanced colon cancer. *Aging (Albany NY)* 2020; **13**: 477-492 [PMID: 33289705 DOI: 10.18632/aging.202157]

7   **Bertelsen CA**, Neuenschwander AU, Jansen JE, Wilhelmsen M, Kirkegaard-Klitbo A, Tenma JR, Bols B, Ingeholm P, Rasmussen LA, Jepsen LV, Iversen ER, Kristensen B, Gögenur I; Danish Colorectal Cancer Group. Disease-free survival after complete mesocolic excision compared with conventional colon cancer surgery: a retrospective, population-based study. *Lancet Oncol* 2015; **16**: 161-168 [PMID: 25555421 DOI: 10.1016/S1470-2045(14)71168-4]

8   **Mushtaq HH**, Shah SK, Agarwal AK. The Current Role of Robotics in Colorectal Surgery. *Curr Gastroenterol Rep* 2019; **21**: 11 [PMID: 30840156 DOI: 10.1007/s11894-019-0676-7]

9   **Benson AB**, Venook AP, Al-Hawary MM, Arain MA, Chen YJ, Ciombor KK, Cohen S, Cooper HS, Deming D, Farkas L, Garrido-Laguna I, Grem JL, Gunn A, Hecht JR, Hoffe S, Hubbard J, Hunt S, Johung KL, Kirilcuk N, Krishnamurthi S, Messersmith WA, Meyerhardt J, Miller ED, Mulcahy MF, Nurkin S, Overman MJ, Parikh A, Patel H, Pedersen K, Saltz L, Schneider C, Shibata D, Skibber JM, Sofocleous CT, Stoffel EM, Stotsky-Himelfarb E, Willett CG, Gregory KM, Gurski LA. Colon Cancer, Version 2.2021, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* 2021; **19**: 329-359 [PMID: 33724754 DOI: 10.6004/jnccn.2021.0012]

10   **Li M**, Zhang J, Dan Y, Yao Y, Dai W, Cai G, Yang G, Tong T. A clinical-radiomics nomogram for the preoperative prediction of lymph node metastasis in colorectal cancer. *J Transl Med* 2020; **18**: 46 [PMID: 32000813 DOI: 10.1186/s12967-020-02215-0]

11   **Patel SG**, Karlitz JJ, Yen T, Lieu CH, Boland CR. The rising tide of early-onset colorectal cancer: a comprehensive review of epidemiology, clinical features, biology, risk factors, prevention, and early detection. *Lancet Gastroenterol Hepatol* 2022; **7**: 262-274 [PMID: 35090605 DOI: 10.1016/S2468-1253(21)00426-X]

12   **Wu J**, Lu L, Chen H, Lin Y, Zhang H, Chen E, Lin W, Li J, Chen X. Prognostic nomogram to predict the overall survival of patients with early-onset colorectal cancer: a population-based analysis. *Int J Colorectal Dis* 2021; **36**: 1981-1993 [PMID: 34322745 DOI: 10.1007/s00384-021-03992-w]

13   **Liu Y**, Sun Z, Guo Y, Liu C, Tian S, Dong W. Construction and validation of a nomogram of risk factors and cancer-specific survival prognosis for combined lymphatic metastases in patients with early-onset colorectal cancer. *Int J Colorectal Dis* 2023; **38**: 128 [PMID: 37183238 DOI: 10.1007/s00384-023-04432-7]

14   **Vather R**, Sammour T, Kahokehr A, Connolly AB, Hill AG. Lymph node evaluation and long-term survival in Stage II and Stage III colon cancer: a national study. *Ann Surg Oncol* 2009; **16**: 585-593 [PMID: 19116751 DOI: 10.1245/s10434-008-0265-8]

15   **Johnson PM**, Porter GA, Ricciardi R, Baxter NN. Increasing negative lymph node count is independently associated with improved long-term survival in stage IIIB and IIIC colon cancer. *J Clin Oncol* 2006; **24**: 3570-3575 [PMID: 16877723 DOI: 10.1200/JCO.2006.06.8866]

16   **Kuo YH**, You JF, Hung HY, Chin CC, Chiang JM, Chang CH. Number of negative lymph nodes with a positive impact on survival of stage III colon cancer; a retrospective observation study for right side and left side colon. *BMC Cancer* 2022; **22**: 126 [PMID: 35100975 DOI: 10.1186/s12885-021-09154-z]

17   **Li Destri G**, Barchitta M, Pesce A, Latteri S, Bosco D, Di Cataldo A, Agodi A, Puleo S. Predictive Value of the Number of Harvested Lymph Nodes and Cut-Off for Lymph Node Ratio in the Prognosis of Stage II and III Colorectal Cancer Patients. *J Invest Surg* 2019; **32**: 1-7 [PMID: 28972442 DOI: 10.1080/08941939.2017.1369605]

18   **Zhang CH**, Li YY, Zhang QW, Biondi A, Fico V, Persiani R, Ni XC, Luo M. The Prognostic Impact of the Metastatic Lymph Nodes Ratio in Colorectal Cancer. *Front Oncol* 2018; **8**: 628 [PMID: 30619762 DOI: 10.3389/fonc.2018.00628]

19   **Lv Y**, Feng QY, Lin SB, Mao YH, Xu YQ, Zheng P, Yang LL, He GD, Xu JM. Exploration of exact significance of lymph node ratio and construction of a novel stage in colon cancer with no distant metastasis. *Cancer Manag Res* 2019; **11**: 6841-6854 [PMID: 31440082 DOI: 10.2147/CMAR.S203533]

20   **Sun Z**, Xu Y, Li de M, Wang ZN, Zhu GL, Huang BJ, Li K, Xu HM. Log odds of positive lymph nodes: a novel prognostic indicator superior to the number-based and the ratio-based N category for gastric cancer patients with R0 resection. *Cancer* 2010; **116**: 2571-2580 [PMID: 20336791 DOI: 10.1002/cncr.24989]

21   **Li S**, Wang Y, Hu X. Prognostic nomogram based on the lymph node metastasis indicators for patients with bladder cancer: A SEER population-based study and external validation. *Cancer Med* 2023; **12**: 6853-6866 [PMID: 36479835 DOI: 10.1002/cam4.5475]

22   **Fang HY**, Yang H, He ZS, Zhao H, Fu ZM, Zhou FX, Zhou YF. Log odds of positive lymph nodes is superior to the number- and ratio-based lymph node classification systems for colorectal cancer patients undergoing curative (R0) resection. *Mol Clin Oncol* 2017; **6**: 782-788 [PMID: 28529752 DOI: 10.3892/mco.2017.1203]

23   **Occhionorelli S**, Andreotti D, Vallese P, Morganti L, Lacavalla D, Forini E, Pascale G. Evaluation on prognostic efficacy of lymph nodes ratio (LNR) and log odds of positive lymph nodes (LODDS) in complicated colon cancer: the first study in emergency surgery. *World J Surg Oncol* 2018; **16**: 186 [PMID: 30213260 DOI: 10.1186/s12957-018-1483-6]

24   **Li T**, Yang Y, Wu W, Fu Z, Cheng F, Qiu J, Li Q, Zhang K, Luo Z, Qiu Z, Huang C. Prognostic implications of ENE and LODDS in relation to lymph node-positive colorectal cancer location. *Transl Oncol* 2021; **14**: 101190 [PMID: 34403906 DOI: 10.1016/j.tranon.2021.101190]

25   **Cheong C**, Oh SY, Kim YB, Suh KW. Differences in biological behaviors between young and elderly patients with colorectal cancer. *PLoS One* 2019; **14**: e0218604 [PMID: 31211804 DOI: 10.1371/journal.pone.0218604]

26   **Ning FL**, Pei JP, Zhang NN, Wang J, Quan HG, Mei ZB, Zeng XT, Abe M, Zhang CD. Harvest of at least 18 lymph nodes is associated with

improved survival in patients with pN0 colon cancer: a retrospective cohort study. *J Cancer Res Clin Oncol* 2020; **146**: 2117-2133 [PMID: 32285257 DOI: 10.1007/s00432-020-03212-y]

27   **Kim MJ**, Jeong SY, Choi SJ, Ryoo SB, Park JW, Park KJ, Oh JH, Kang SB, Park HC, Heo SC, Park JG. Survival paradox between stage IIB/C (T4N0) and stage IIIA (T1-2N1) colon cancer. *Ann Surg Oncol* 2015; **22**: 505-512 [PMID: 25145501 DOI: 10.1245/s10434-014-3982-1]

28   **Tsikitis VL**, Larson DL, Wolff BG, Kennedy G, Diehl N, Qin R, Dozois EJ, Cima RR. Survival in stage III colon cancer is independent of the total number of lymph nodes retrieved. *J Am Coll Surg* 2009; **208**: 42-47 [PMID: 19228501 DOI: 10.1016/j.jamcollsurg.2008.10.013]

29   **Cozzani F**, Agnesi S, Dell'abate P, Rossini M, Viani L, Pedrazzi G, Del Rio P. The prognostic role of metastatic lymph node ratio in colon cancer: a retrospective cohort study on 241 patients in a single center. *Minerva Surg* 2023; **78**: 155-160 [PMID: 36193952 DOI: 10.23736/S2724-5691.22.09619-8]

30   **Fortea-Sanchis C**, Martínez-Ramos D, Escrig-Sos J. The lymph node status as a prognostic factor in colon cancer: comparative population study of classifications using the logarithm of the ratio between metastatic and nonmetastatic nodes (LODDS) versus the pN-TNM classification and ganglion ratio systems. *BMC Cancer* 2018; **18**: 1208 [PMID: 30514228 DOI: 10.1186/s12885-018-5048-4]

31   **González N**, Loroño A, Aguirre U, Lázaro S, Baré M, Redondo M, Briones E, Sarasqueta C, Bilbao A, de Larrea NF, Quintana JM; REDISSEC-CARESS/CCR group. Risk scores to predict mortality 2 and 5 years after surgery for colorectal cancer in elderly patients. *World J Surg Oncol* 2021; **19**: 252 [PMID: 34446044 DOI: 10.1186/s12957-021-02356-6]