Nomogram to Predict the Occurrence and Prognosis of Distant Metastasis in T1n0 Colon Cancer: A Seer Data-Based Study

Yunxiao Liu
the Second Affiliated Hospital of Harbin Medical University
Hao Zhang
the Second Affiliated Hospital of Harbin Medical University
Mingyu Zheng
the Second Affiliated Hospital of Harbin Medical University
Chunlin Wang
the Second Affiliated Hospital of Harbin Medical University
Zhiqiao Hu
the Second Affiliated Hospital of Harbin Medical University
Yang Wang
the Second Affiliated Hospital of Harbin Medical University
Huan Xiong
the Second Affiliated Hospital of Harbin Medical University
Boyang Fan
the Second Affiliated Hospital of Harbin Medical University
Yuliuming Wang
the Second Affiliated Hospital of Harbin Medical University
Hanqing Hu
the Second Affiliated Hospital of Harbin Medical University
Qingchao Tang
the Second Affiliated Hospital of Harbin Medical University
Guiyu Wang (guiywang@163.com)
the Second Affiliated Hospital of Harbin Medical University

Research Article

Keywords: T1N0 colon cancer, Distant Metastasis, Nomogram, SEER

Posted Date: April 27th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-451250/v1

Abstract

Distant metastasis (DM) is relatively rare in T1 colon cancer (CC) patients, especially in those with negative lymph node metastasis. The aim of this study was to explore the main clinical factors and build nomogram for predicting the occurrence and prognosis of DM in T1N0 colon cancer patients. Patients with T1N0 stage colon cancer were collected from the Surveillance, Epidemiology, and End Result (SEER) database. All patients were divided into development and validation cohorts with the 3:1 ratio. A total of 6770 patients were enrolled in this study, including 428 patients (6.3%) with DM. Age, size, grade, CEA were independent risk factors associated with DM. Age, grade, CEA, surgery and chemotherapy were independent prognostic factors for CSS. Nomogram were applied and C-index, calibration curves, ROC curves and DCA curves proved good discrimination, calibration and clinical practicability of the nomogram in predicting the occurrence and prognosis of DM in T1N0 colon cancer patients. The population-based nomogram could help clinicians predict the occurrence and prognosis of DM in T1N0 CC patients and provide a reference to perform appropriate metastatic screening plans and rational therapeutic options for the special population.

Introduction

Colorectal cancer (CRC) is one of the most common cancers in the whole world, causing a large number of deaths every year and forming a huge burden on both family and society. In 2019, colorectal cancer was estimated to be the highest cancer incidence and death rate in the United States [1]. In general, surgical resection, endoscopic therapy, and neoadjuvant therapy have become the main treatment methods for different stages of colorectal cancer [2, 3-5]. With the widespread application of endoscopic screening, more colon cancer patients have been identified at an early stage in recent years. According to the American Joint Committee on Cancer (AJCC) staging manual, T1 colon cancer refers to submucosal invasive carcinoma. The 2019 Guidelines Chinese Society of Clinical Oncology (CSCO) recommended that endoscopic therapy is considered as the preferred treatment strategy for T1N0 colon cancer patients without distant metastasis, which is associated with higher quality of life and less postoperative morbidity compared with extensive surgical resection [6]. Approximately 10% of T1 cancers are stage III or IV after primary endoscopic resection based on population studies [7-9]. Notwithstanding that the risk of DM in CC patients is relatively rare, the open question that “whether the T1N0 CC patient has developed a metastatic disease” should also be considered in the clinical management for this special population without doubt. However, in some cases, because of the limited manifestations for patients with early-staged colon cancer, although simultaneous DM has occurred in their bodies, they might fail to be diagnosed as stage IV disease and subsequently receive an unreasonable endoscopic therapy. Additionally, although the effects of some multidisciplinary therapeutic strategies for stage IV colon cancer patients have been explored by many scholars in recent years, there is still no consensus on the optimal treatment for these patients. As we know, an accurate prognostic assessment is of vital significance for clinicians to develop personalized treatments for patients. Hence, a further prognostic

model is imminently needing to be explored to provide a potential reference for the better treatment of T1N0 colon cancer patients.

Therefore, in the current study, we built a nomogram to predict the probability of T1N0 patients developing DM and also established a nomogram to determine prognosis of T1N0M1 patients using the Surveillance, Epidemiology, and End Result (SEER) database.

Methods

Patients

The data of patients with T1N0Mx CC in the SEER database between January 2004 and December 2015 were extracted with the SEER*Stat software (version 8.3.8; www.seer.cancer.gov) using a private ID (account number: 25213-Nov2019), and treatment data were acquired from SEER custom data via further application. Informed consent was not required because the SEER database is publicly available.

Inclusion criteria included: 1) The patient was diagnosed as T1N0 colon cancer; 2) aged ≥ 18 years old; 3) patients with complete records of cancer-specific survival months; 4) colon cancer was the only primary malignancy. Exclusion criteria included: 1) patients underwent neoadjuvant therapy; 2) patients with unknown race, histological type, grade, T stage, N stage, tumor size and CEA level and 3) patients without complete follow-up.

Variables

In this study, the following variables were selected from the SEER database: patient ID, sex, age at diagnosis, TNM stage, tumor size, tumor site, histology, grade, CEA level, surgery, chemotherapy, CCS.

According to our study, age was regrouped into < 40, 40-59, 60-79 and ≥ 80 years old; sex was classified as male or female; race was recorded as black, white, or other; tumor size was divided into three groups; ≤ 3cm, ≤ 5cm (3 cm < tumor size ≤ 5 cm), > 5 cm. The tumor site was grouped into right-sided colon (cecum, ascending colon, hepatic flexure and transverse colon) and left-sided colon (splenic flexure, descending colon and sigmoid colon). The histology variable was classified as “adenocarcinoma”, “mucinous adenocarcinoma” or “other”; the grade variable was classified as “well differentiated”, moderately differentiated, “poorly differentiated” and “undifferentiated” and the CEA level variable was classified as “positive” and “negative”. Chemotherapy was classified as “yes” or “no”; surgery was classified as “yes” or “no” according to the SEER database. CSS was defined as the time from diagnosis to the date of death due to CC.

Statistical analysis

All the statistical analyses were calculated in statistical software package SPSS 22.0 (IBM Corp, Armonk, NY, USA) and R software (version 3.6.1, https://www.r-proje ct.org/). In this study, in order to ensure the accuracy of the nomogram, patients were randomly (random: 3:1 ratio) divided into development cohort

and validation cohort. Univariate and multivariate Logistic regression analyses were performed to determine risk factors for DM in T1N0Mx patients and Cox regression models were utilized to analyze prognostic factors in T1N0 patients with DM. Nomogram was constructed based on the results of multivariate regression, and its performance was further evaluated by C index, calibration and AUC. Calibration is performed graphically by plotting the correlation between predicted probabilities and actual results. In addition, clinical impact curves were drawn based on DCA to help us more intuitively understand the significant value of the nomogram. Besides, with the X-tile program [10], T0N0M1 patients were classified into low-risk and high-risk groups according to their total scores derived from the survival nomogram. Kaplan-Meier curves were generated and analyzed using log-rank tests. The difference was considered statistically significant for a two-sided $P < 0.05$.

Results

Patient characteristics

According to the screening criteria, a total of 6,770 T1N0Mx patients were in the study predicting distant metastasis and 428 T1N0 patients with DM were in the study predicting CSS. In all T1N0 patients, most proportions were found in 60-79 years old, white, male, right colon, tumor size ≤ 3cm, moderately differentiated, adenocarcinoma, CEA negative and absence of DM. And in patients with DM, most cases were found to be associated with 60-79 years old, white, male, right colon, 3 < tumor size ≤ 5cm, moderately differentiated, adenocarcinoma, CEA positive, absence of surgery and presence of chemotherapy. The baseline characteristics of patients after 3: 1 ratio randomly stratified were calculated in Tables 1 and 2.

Construction and validation of nomogram to predict DM probability

In order to further explore the risk factors for DM in T1N0 patients, univariate and multivariate logistic regression analyses were performed to determine the independent risk factors for DM step by step. In univariate analysis, the candidate predictors for the model were age, race, sex, tumor size, tumor site, grade, histology and CEA. All the predictors except for sex, tumor site and histology were significantly different between subgroups, which were then further analyzed by multivariate logistic regression model. And the results indicated that age (OR = 0.647, 95%CI = 0.243-1.719 for 40-59 years old, P = 0.382; OR =0.332, 95%CI = 0.126-0.874 for 60-79 years old, P = 0.026; OR = 0.202, 95%CI = 0.073-0.557, P = 0.002 for ≥ 80 years old; using < 40 years old as the reference), tumor size (OR = 0.092, 95%CI = 0.064-0.134 for 3cm < tumor size ≤ 5 cm, P < 0.001; OR = 0.654, 95%CI = 0.446-0.958 for tumor size > 5cm, P = 0.029; using tumor size ≤ 3 cm as the reference), grade (OR = 2.323, 95%CI = 1.511-3.574 for moderately differentiated, P < 0.001; OR = 5.686, 95%CI = 3.236-9.99 for poorly differentiated, P < 0.001; OR = 7.159, 95%CI = 2.462-20.821 for undifferentiated, P < 0.001; using well differentiated as the reference), CEA level (OR = 0.054, 95%CI = 0.040-0.074 for CEA negative, P < 0.001, using CEA positive as the reference) were independent risk factors in predicting the occurrence of DM (Table 3).

Based on the independent risk factors in the multivariate analysis, we construct a nomogram to predict DM in T1N0Mx patients (Figure 1). The AUCs for development cohort and validation cohort were 0.901 (95%CI = 0.879-0.922) and 0.899 (95%CI =0.865-0.940), respectively (Figure 2). The calibration curves (development cohort: S: p = 0.712; validation cohort: S: p = 0.681) showed the relatively satisfactory prediction accuracy of the nomogram (Figure 3). In addition, the DCA curve also indicated good clinical practicability in both cohorts (Figure 4).

Construction and validation of nomogram to predict CSS in T1N0 patients with DM

After analyzing the risk factors of DM in T1N0Mx patients, we also explored CSS in DM patients using Kaplan-Meier method and Cox regression model. Univariate analysis revealed that sex and grade were not important factors for CSS in DM patients. The risk factors in the univariate analysis were further analyzed by Cox multivariate regression model. And the results indicated that age (HR = 0.587, 95%CI = 0.272-11.264 for < 40 years old, P = 0.174; HR =0.439, 95%CI = 0.296-0.651 for 40-59 years old, P < 0.001; HR = 0.603, 95%CI = 0.418-0.871 for 60-79 years old, P = 0.007; using ≥ 80 years old as the reference), histology (HR = 1.487, 95%CI = 0.794-2.784 for mucinous, P = 0.215; HR = 4.682, 95%CI = 1.672-13.115 for other, P = 0.003; using adenocarcinoma as the reference), surgery (HR = 3.450, 95%CI = 2.313-5.145 for no, P < 0.001; using yes as the reference), chemotherapy (HR = 2.032, 95%CI = 1.512-2731 for no, P < 0.001; using yes as the reference), CEA level (HR = 0.454, 95%CI = 0.293-0.703 for CEA negative, P < 0.001, using CEA positive as the reference) were independent prognosticators in predicting CSS with DM patients (Table4).

We constructed a nomogram to predict 1-, 2- and 3-year survival in DM patients with T1N0 colon cancer, incorporating age, histology, CEA, surgery and chemotherapy (Figure 5). The C-indices of the development and validation cohort were 0.718 (95%CI=0.639-0.737) and 0.712 (95%CI=0.681-0.743). The area under the ROC curves of the CSS nomogram were shown in Figure 6. The AUCs of the nomogram at 1-, 2-, and 3-year were 0.763 (95%CI=0.744-0.782), 0.794 (95%CI=0.775-0.813), and 0.822 (95%CI=0.803-0.841) for the development cohort, and 0.785 (95%CI=0.754-0.816), 0.748 (95%CI=0.717-0.779) and 0.896 (95%CI=0.865-0.927) for the validation cohort. The calibration plot showed a satisfactory predictive accuracy between 1-, 2-, and 3-year predicted CSS and observed CSS in both cohorts (Figure 7). In addition, clinical impact curves were drawn based on DCA to help us more intuitively understand the significant value of the nomogram model (Figure 8).

Using the nomogram derived scores, all DM patients were classified into two subgroup low-risk (risk score ≤ 171) and high-risk groups (risk score > 171) by the X-tile program (Figure 9). And we found there were significant differences in Kaplan-Meier curves between the high risk and low risk groups in development cohort (P < 0.001) and validation cohort (P < 0.001) (Figure 10).

Discussion

CRC is one of the most common cancers worldwide and the major causes of cancer-related mortality [1,11]. Thanks to the prevalence of screening programs and advancement of endoscopic techniques, more and more CRC patients are diagnosed at an early stage (T1) [12]. About 90% of T1 CRC patients have been diagnosed at stage I, and endoscopic resection of the lesion is an attractive therapeutic strategy for these patients which allows comparable prognosis to surgery less surgical complications, better functional recovery and improved quality of life. The risk of DM harbors a crucial part in the consideration of reasonable treatment for T1N0 patients. As far as we know, due to the low probability of DM in patients with stage T1N0, no studies have been conducted in this regard. Our study has for the first time identified the main clinical risk indicators and prognostic factors of diabetes in patients with T1N0 colon cancer, which has great guiding significance for clinical workers.

In the current study, we found age, tumor size, tumor grade and CEA level were associated with the development of DM in T1N0 patients. Compared with elderly patients, young patients are related to more aggressive histopathologic features and advanced stage, so age has been always deemed as an important factor of metastasis [13]. Mostly, large tumor size is usually associated with stronger potential to develop metastasis [14]. Poor histological grade is linked to more invasive ability of tumor cells and would be prone to distant metastasis. Luo et al revealed that CEA level is a significant serum tumor marker of metastasis in T1 patients [15]. Subsequently, factors abovementioned were used to build the nomogram to predict the probability of DM. The AUCs were 0.901 and 0.899 for the development and validation cohorts, respectively. The clinical power was also proved by the calibration and DCA curves in the two cohorts. Previously, some studies have analyzed the risk of lymph node metastasis in T1 CC in order to conduct a reasonable endoscopic treatment[6]. But DM in this special population has been described rarely. T1 stage tumors seldom present with DM, especially in the absence of lymph node metastasis. CC could metastasize in several ways, including lymphatic and hematology. The latter is a pattern in which tumor cells invade blood vessels and travel directly to distant organs. Patients with and without metastasis have completely different outcomes and therapeutic methods, so it is necessary to identify risk factors for distant metastasis in T1N0 patients. Routine imaging studies, such as computed tomography, could reveal significant disseminated lesions, and enhanced computed tomography is a more reliable method for screening the DM in CC, especially in detecting micrometastases. However, the cost and availability limit their application. Therefore, constructing an economical and convenient nomogram is crucial for clinicians to early determine high-risk T1N0 patients susceptible to distant metastasis and perform targeted screening and reasonable treatments for these special population.

We then analyzed CSS in 428 patients with DM. In the CSS nomogram, we included not only patients' basic information (age, histological type and CEA), but also treatment information such as surgery and chemotherapy. As a palliative treatment modality for CC patients with M1 stage, the potential benefit of primary tumor resection is diffusely discussed in many scholars. Some people argue that stage IV patients have lost the chance of surgery, and recommend systemic chemotherapy alone to prolong survival. Ichikawa et al and Shimomura et al suggested that the benefit of primary tumor resection was unclear and needing to be explored by more clinical works [16-18]. By contrast, Park et al conducted an

analysis on 1015 stage IV CRC patients and found that patients receiving palliative surgery without residual disease and chemotherapy harbored better prognosis compared those with chemotherapy alone[19]. Yeom et al also found that the survival time of patients receiving surgery and adjuvant therapy was significantly longer than that of patients receiving adjuvant therapy alone [20]. Similar results were achieved by some other reports [21-23]. In the current research, primary tumor resection was associated with improved survival time according to the multivariate analyses. The potential reasons might be that the surgery of primary tumor could avoid tumor-related symptoms and alleviate tumor load of the patients, thus prolonging the patients' survival time. Besides, age, histological type, CEA and chemotherapy were also determined as important factors for CSS of T1N0M1 patients, which have been linked with considerable importance in predicting survival of CC patients in recent years [24-26]. Therefore, a nomogram was built to determine cancer-specific survival using these five factors (age, histological type, CEA, chemotherapy and surgery). The C-indices of the development and validation cohort were 0.718 and 0.712. Moreover, patients were stratified into low- and high-risk groups according to their total scores and we found that the high-risk group had significantly improved survival times than the low-risk group. The proposed results could provide a potential reference to better manage T1N0M1 patients and guide the rational application of medical treatment resource. If the patient's basic conditions permit, surgery plus chemotherapy could be used to pursue the highest possible survival rate with informed consent. Besides, some novel regimens such as conversion and targeted therapeutic strategies should also be investigated in the near future.

This study has several limitations. Firstly, because our study was a retrospective study and included patients from 2004 to 2015, there may be the possibility of inaccurate data and it is inevitable to have observer and confusion bias. The current result requires further validation by some prospective clinical researches. Second, despite the internal validation of our model, there is a lack of external validation to further determine the accuracy of the model. Thirdly, some potential prognosticators such as BRAF and RAS mutational status, surgical methods, and more detailed information about chemotherapy protocols are not available in the SEER database. Incorporating these important factors may further improve the validity of nomograms. Finally, due to the limited number of cases and missing data in the SEER database, further stratification of metastatic sites was not possible.

In conclusion, this study performed prediction of DM and survival analysis in patients with stage T1N0 colon cancer. Age, histology, tumor size, and CEA were independent predictors of DM. Age, grade, CEA, surgery and chemotherapy were independent prognostic factors for CSS. DM nomogram and CSS nomogram based on the above factors has favorable accuracy and superior predictive power. The proposed nomogram could help clinicians predict the risk and prognosis of distant metastasis in T1N0 colon cancer patients and provide a reference to perform appropriate metastatic screening plans and rational therapeutic options for the special population.

Declarations

Acknowledgments

The authors acknowledge the efforts of the Surveillance, Epidemiology, and End Results (SEER) Program tumor registries in the creation of the SEER database.

Conflicts of interest

None declared.

Data availability

The data sets used and/or analyzed during the current study are available from the SEER database.

Funding

None

Ethical statements

This study was approved by the Second Affiliated Hospital of Harbin Medical University and Zhejiang Cancer Hospital. The study used de-identified data and adhered to World Medical Association's Declaration of Helsinki for Ethical Human Research. The informed consent was not required according to personal identifying information was not included.

Author Contributions

YXL and HZ conceived and designed the study. YXL, HZ, MYZ and CLW collected clinical data and performed the statistical analysis. ZQH and YW performed the research and wrote the paper. YXL, HX, BYF, YLMW, HQH, QCT and GYW reviewed and edited the manuscript. All authors read and approved the manuscript.

References

1. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2019. *CA Cancer J Clin.* **69**, 7–34 (2019).
2. Yang, Y. *et al.* Individualized prediction of survival benefit from primary tumor resection for patients with unresectable metastatic colorectal cancer. *World J Surg Oncol.* **18**, 193 (2020).
3. Nozawa, H. *et al.* Conversion to Resection in Patients Receiving Systemic Chemotherapy for Unresectable and/or Metastatic Colorectal Cancer-Predictive Factors and Prognosis. *Clin Colorectal Cancer.* **17**, e91–e7 (2018).
4. Dang, H., de Cappel, V. T. N. & van der Zwaan, W. H. SMS, et al. Quality of life and fear of cancer recurrence in T1 colorectal cancer patients treated with endoscopic or surgical tumor resection. *Gastrointest Endosc.* **89**, 533–544 (2019).
5. Senore, C. *et al.* Management of Pt1 tumours removed by endoscopy during colorectal cancer screening: Outcome and treatment quality indicators. *Eur J Surg Oncol.* **44**, 1873–1879 (2018).

6. Guo, K. *et al.* Risk factors and predictors of lymph nodes metastasis and distant metastasis in newly diagnosed T1 colorectal cancer. *Cancer Med.* **9**, 5095–5113 (2020).
7. Sun, Z. Q. *et al.* Prognostic value of lymph node metastasis in patients with T1-stage colorectal cancer from multiple centers in China. *World J Gastroenterol.* **23**, 8582–8590 (2017).
8. Saitoh, Y. *et al.* Management of colorectal T1 carcinoma treated by endoscopic resection. *Dig Endosc.* **28**, 324–329 (2016).
9. Wada, H. *et al.* Systematic review and meta-analysis of histopathological predictive factors for lymph node metastasis in T1 colorectal cancer. *J Gastroenterol.* **50**, 727–734 (2015).
10. Camp, R., Dolled-Filhart, M. & Rimm, D. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clinical cancer research: an official journal of the American Association for Cancer Research.* **10**, 7252–7259 (2004).
11. Malvezzi, M. *et al.* European cancer mortality predictions for the year 2018 with focus on colorectal cancer. *Ann Oncol.* **29**, 1016–1022 (2018).
12. Chiu, H. M. *et al.* Effectiveness of fecal immunochemical testing in reducing colorectal cancer mortality from the One Million Taiwanese Screening Program. *Cancer.* **121**, 3221–3229 (2015).
13. Yantiss, R. *et al.* Clinical, pathologic, and molecular features of early-onset colorectal carcinoma. *The American journal of surgical pathology.* **33**, 572–582 (2009).
14. Kornprat, P. *et al.* Value of tumor size as a prognostic variable in colorectal cancer: a critical reappraisal. *American journal of clinical oncology.* **34**, 43–49 (2011).
15. Lou, Z. *et al.* Preoperative carcinoembryonic antibody is predictive of distant metastasis in pathologically T1 colorectal cancer after radical surgery. *World J Gastroenterol.* **19**, 389–393 (2013).
16. Kim, C. W. *et al.* The role of primary tumor resection in colorectal cancer patients with asymptomatic, synchronous unresectable metastasis: Study protocol for a randomized controlled trial. *Trials.* **17**, 34 (2016).
17. Ichikawa, Y. *et al.* Does resection of primary lesions show survival benefit for stage IV colorectal cancer patients with unresectable metastases? *Hepatogastroenterology.* **60**, 1945–1949 (2013).
18. Shimomura, M. *et al.* Identification of patients likely to benefit from metastasectomy in stage IV colorectal cancer. *Int J Colorectal Dis.* **27**, 1339–1346 (2012).
19. Park, J. H. *et al.* The beneficial effect of palliative resection in metastatic colorectal cancer. *Br J Cancer.* **108**, 1425–1431 (2013).
20. Yeom, S. S. *et al.* The outcome of primary tumor resection in the unresectable stage IV colorectal cancer patients who received the bevacizumab-containing chemotherapy. *Med. (Baltim).* **99**, e19258 (2020).
21. Ha, G. W., Kim, J. H. & Lee, M. R. Meta-analysis of oncologic effect of primary tumor resection in patients with unresectable stage IV colorectal cancer in the era of modern systemic chemotherapy. *Ann Surg Treat Res.* **95**, 64–72 (2018).

22. Chan, T. W. *et al.* Primary tumor resection in patients presenting with metastatic colorectal cancer: analysis of a provincial population-based cohort. *Am J Clin Oncol.* **33**, 52–55 (2010).
23. Alawadi, Z. *et al.* Comparative effectiveness of primary tumor resection in patients with stage IV colon cancer. *Cancer.* **123**, 1124–1133 (2017).
24. Hasegawa, S. *et al.* [A study of 10 patients with unresectable colorectal cancer who achieved long-term survival]. *Gan To Kagaku Ryoho.* **40**, 1978–1980 (2013).
25. Park, E. J. *et al.* The Role of Primary Tumor Resection in Colorectal Cancer Patients with Asymptomatic, Synchronous, Unresectable Metastasis: A Multicenter Randomized Controlled Trial.Cancers (Basel). 2020; 12.
26. Kim, N. K. Paradigm Shift in the Treatment of Elderly Patients With Unresectable Stage IV Colorectal Cancer. *Ann Coloproctol.* **30**, 155–156 (2014).

Tables

Table 1 Baseline clinical characteristics of T1N0 patients in our study

Characteristics	Development cohort (n = 5077)	Validation cohort (n = 1693)	P-value
Age (years), n (%)			0.076
< 40	67(1.3)	16(0.9)	
40-59	1374(27.1)	419(24.7)	
60-79	2770(54.6)	980(57.9)	
≥ 80	866(17.1)	278(16.4)	
Race, n (%)			0.869
Black	655(12.9)	210(12.4)	
White	3921(77.2)	1315(77.7)	
Other	501(9.9)	168(9.9)	
Sex, n (%)			0.528
Male	2609(51.4)	885(52.3)	
Female	2468(48.6)	808(74.7)	
Tumor site, n (%)			0.663
Right colon	2830(55.7)	954(56.3)	
Left colon	2247(44.3)	739(43.7)	
Tumor size (cm), n (%)			0.283
≤ 3	4066(80.1)	1344(79.4)	
≤ 5	694(13.7)	254(15.0)	
> 5	317(6.2)	95(5.6)	
Grade, n (%)			0.175
Well differentiated	1054(20.8)	347(20.5)	
Moderately differentiated	3632(71.5)	1206(71.2)	
Poorly differentiated	341(6.7)	131(7.7)	
Undifferentiated	50(1.0)	9(0.5)	
Histology, n (%)			
Adenocarcinoma	4896(96.4)	1632(96.4)	0.752
Mucinous	166(3.3)	54(3.2)	

Other	15(0.3)	7(0.4)	
CEA, n (%)			0.205
Positive	987(19.4)	313(18.5)	
Negative	4090(80.6)	1380(81.5)	
Distant metastasis, n (%)			0.567
Yes	316(6.2)	112(6.6)	
No	4761(93.8)	1581(93.4)	

Table 2 Baseline clinical characteristics of T1N0 patients with DM in our study

Characteristics	Development cohort (n = 321)	Validation cohort (n = 107)	P-value
Age (years), n (%)			0.080
< 40	9(2.8)	3(2.8)	
40-59	110(34.3)	43(40.2)	
60-79	149(46.4)	54(50.5)	
≥ 80	53(16.5)	7(6.5)	
Race, n (%)			0.335
Black	70(21.8)	18(16.8)	
White	224(69.8)	76(71.0)	
Other	27(8.4)	13(12.1)	
Sex, n (%)			0.911
Male	175(54.5)	59(55.1)	
Female	146(45.5)	48(44.9)	
Tumor site, n (%)			0.615
Right colon	165(51.4)	58(54.2)	
Left colon	156(48.6)	49(45.8)	
Tumor size (cm), n (%)			0.169
≤ 3	110(34.3)	30(28.0)	
≤ 5	120(37.4)	51(47.7)	
> 5	91(28.3)	26(24.3)	
Grade, n (%)			0.483
Well differentiated	32(10.0)	7(6.5)	
Moderately differentiated	230(71.7)	84(78.5)	
Poorly differentiated	52(16.2)	15(14.0)	
Undifferentiated	7(2.2)	1(0.9)	
Histology, n (%)			0.590
Adenocarcinoma	302(94.1)	98(91.6)	
Mucinous	15(4.7)	7(6.5)	

Other	4(1.2)	2(1.9)	
CEA, n (%)			1.000
Positive	261(81.3)	87(81.3)	
Negative	60(18.7)	20(18.7)	
Surgery, n (%)			0.160
Yes	68(21.2)	16(15.0)	
No	253(78.8)	91(85.0)	
Chemotherapy, n (%)			0.418
Yes	199(62.0)	71(66.4)	
No	122(38.0)	36(33.6)	

Table 3 Logistic regression analysis of the risk factors for DM in T1N0 patients

	Univariate analysis		Multivariate analysis	
Characteristics	**OR [95% CI]**	**P-value**	**OR [95% CI]**	**P-value**
Age(years)				
< 40	Ref		Ref	
40-59	0.600[0.290-1.241]	0.168	0.647[0.243-1.719]	0.382
60-79	0.359[0.174-0.738]	**0.005**	0.332[0.126-0.874]	**0.026**
≥ 80	0.345[0.161-0.741]	**0.006**	0.202[0.073-0.557]	**0.002**
Race				
Black	Ref		Ref	
White	0.547[0.410-0.731]	**< 0.001**	0.880[0.612-1.265]	0.489
Other	0.537[0.339-0.851]	**0.008**	0.709[0.403-1.248]	0.223
Sex				
Male	Ref			
Female	0.866[0.689-1.089]	0.218		
Tumor site				
Right colon	Ref			
Left colon	1.245[0.991-1.564]	0.059		
Tumor size(cm)				
≤ 3	Ref		Ref	
≤ 5	8.543[6.469-11.282]	**< 0.001**	0.092[0.064-0.134]	**< 0.001**
> 5	16.466[12.043-22.513]	**< 0.001**	0.654[0.446-0.958]	**0.029**
Grade				
Well differentiated	Ref		Ref	
Moderately differentiated	2.149[1.475-3.132]	**< 0.001**	2.323[1.511-3.574]	**< 0.001**
Poorly differentiated	5.232[3.284-8.336]	**< 0.001**	5.686[3.236-9.99.]	**< 0.001**
Undifferentiated	5.199[2.172-12.446]	**< 0.001**	7.159[2.462-20.821]	**< 0.001**
Histology				
Adenocarcinoma	Ref			

Mucinous	1.185[0.651-2.157]	0.578		
Other	2.340[0.526-10.418]	0.264		
CEA				
Positive	Ref		Ref	
Negative	0.043[0.033-0.058]	**< 0.001**	0.054[0.040-0.074]	**< 0.001**

Table 4 COX regression analysis of the prognostic factors for CSS in T1N0 patients with DM

	Univariate analysis		Multivariate analysis	
Characteristics	**HR [95% CI]**	**P-value**	**HR [95% CI]**	**P-value**
Age (years)				
< 40	0.578[0.272-1.229]	0.154	0.587[0.272-1.264]	0.174
40-59	0.438[0.305-0.630]	**< 0.001**	0.439[0.296-0.651]	**< 0.001**
60-79	0.535[0.377-0.760]	**< 0.001**	0.603[0.418-0.871]	**0.007**
≥ 80	Ref		Ref	
Race				
Black	Ref		Ref	
White	0.669[0.496-0.903]	**0.009**	0.801[0.584-1.098]	0.168
Other	0.585[0.348-0.982]	**0.043**	0.645[0.377-1.102]	0.109
Sex				
Male	Ref		Ref	
Female	0.786[0.610-1.012]	0.061		
Tumor site				
Right colon	Ref		Ref	
Left colon	0.737[0.573-0.948]	**0.017**	0.857[0.659-1.114]	0.249
Tumor size (cm)				
≤ 3	Ref		Ref	
≤ 5	1.588[1.172-2.154]	**0.003**	1.101[0.789-1.536]	0.573
> 5	1.978[1.420-2.756]	**< 0.001**	1.366[0.960-1.945]	0.083
Grade				
Well differentiated	Ref			
Moderately differentiated	1.168[0.741-1.840]	0.504		
Poorly differentiated	1.587[0.934-2.694]	0.088		
Undifferentiated	1.709[0.643-4.541]	0.282		
Histology				
Adenocarcinoma	Ref		Ref	
Mucinous	1.590[0.866-2.917]	0.135	1.487[0.794-2.784]	0.215

Other	4.846[1.785-13.153]	**0.002**	4.682[1.672-13.115]	**0.003**
CEA				
Positive	Ref		Ref	
Negative	0.428[0.287-0.638]	**< 0.001**	0.454[0.293-0.703]	**< 0.001**
Surgery				
Yes	Ref		Ref	
No	3.671[2.543-5.299]	**< 0.001**	3.450[2.313-5.145]	**< 0.001**
Chemotherapy				
Yes	Ref		Ref	
No	1.341[1.027-1.751]	**0.031**	2.032[1.512-2.731]	**< 0.001**

Figures

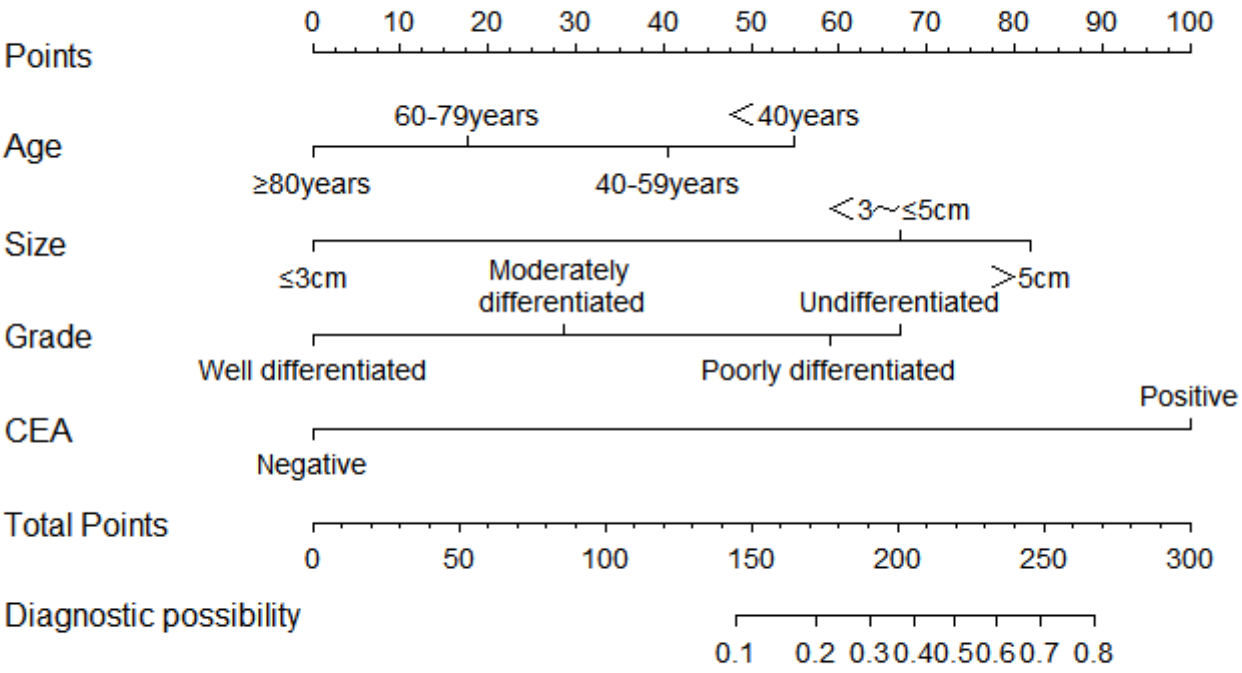

Figure 1

Nomogram for predicting the probability of distant metastasis

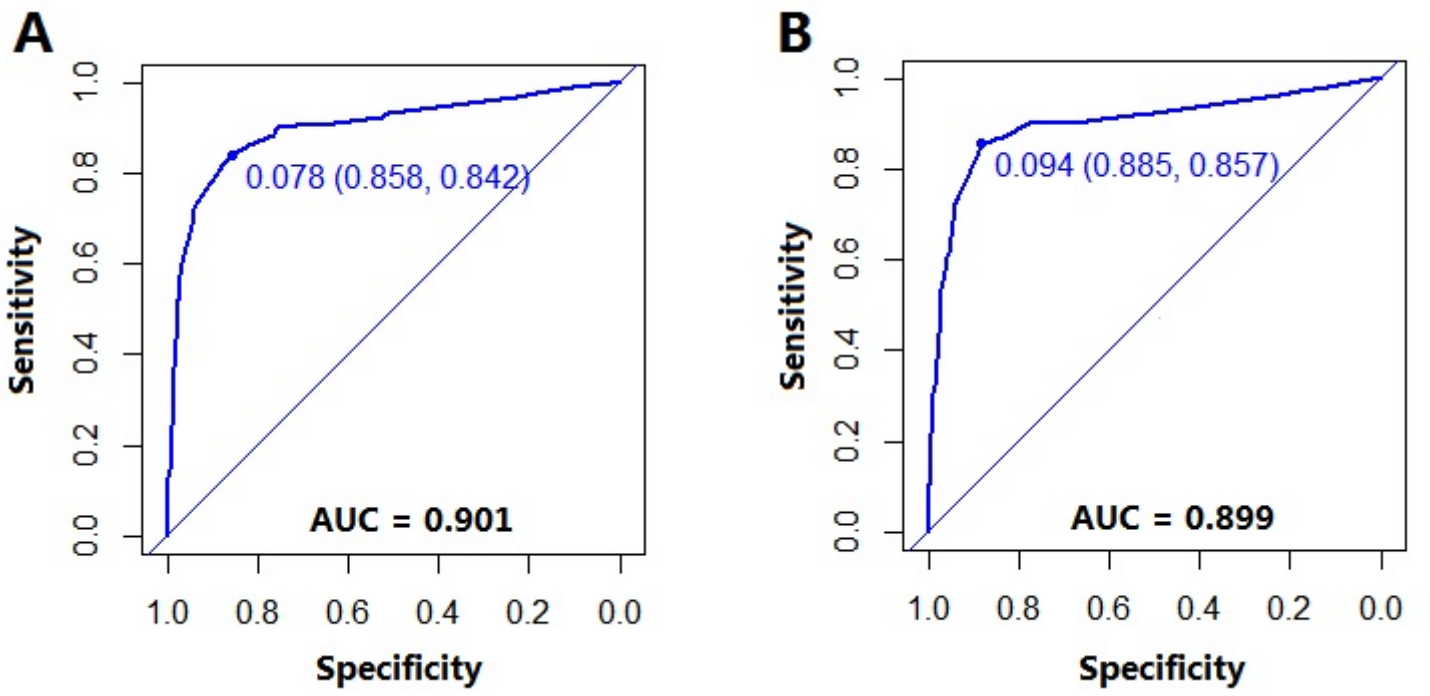

Figure 2

The ROC curves of nomogram for predicting DM in the development cohort (A) and validation cohort (B)

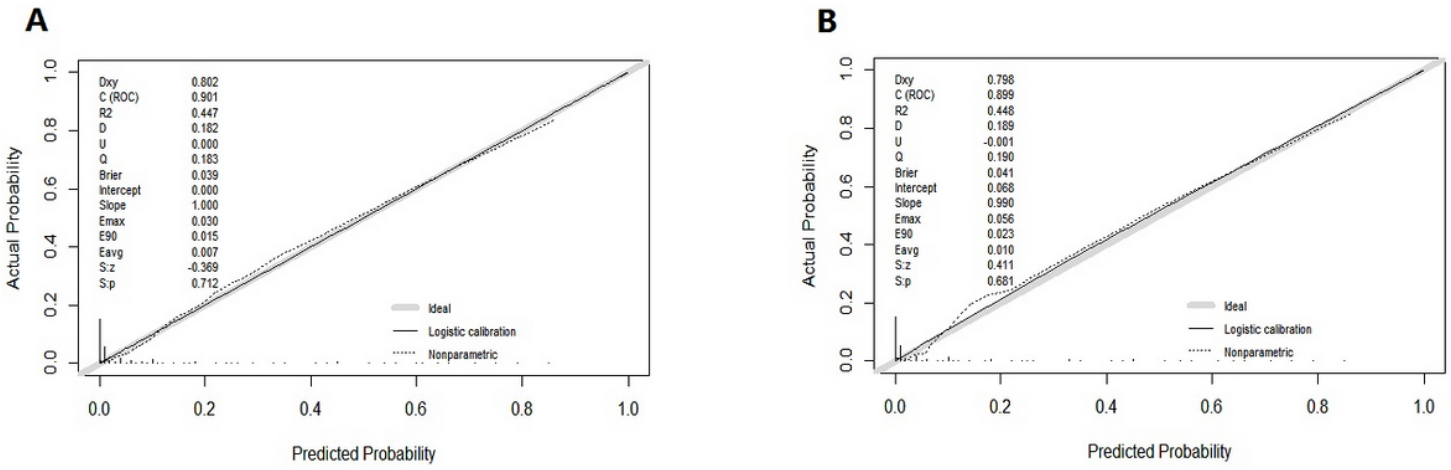

Figure 3

The calibration curves of the nomogram for predicting DM in the development cohort (A) and validation cohort (B)

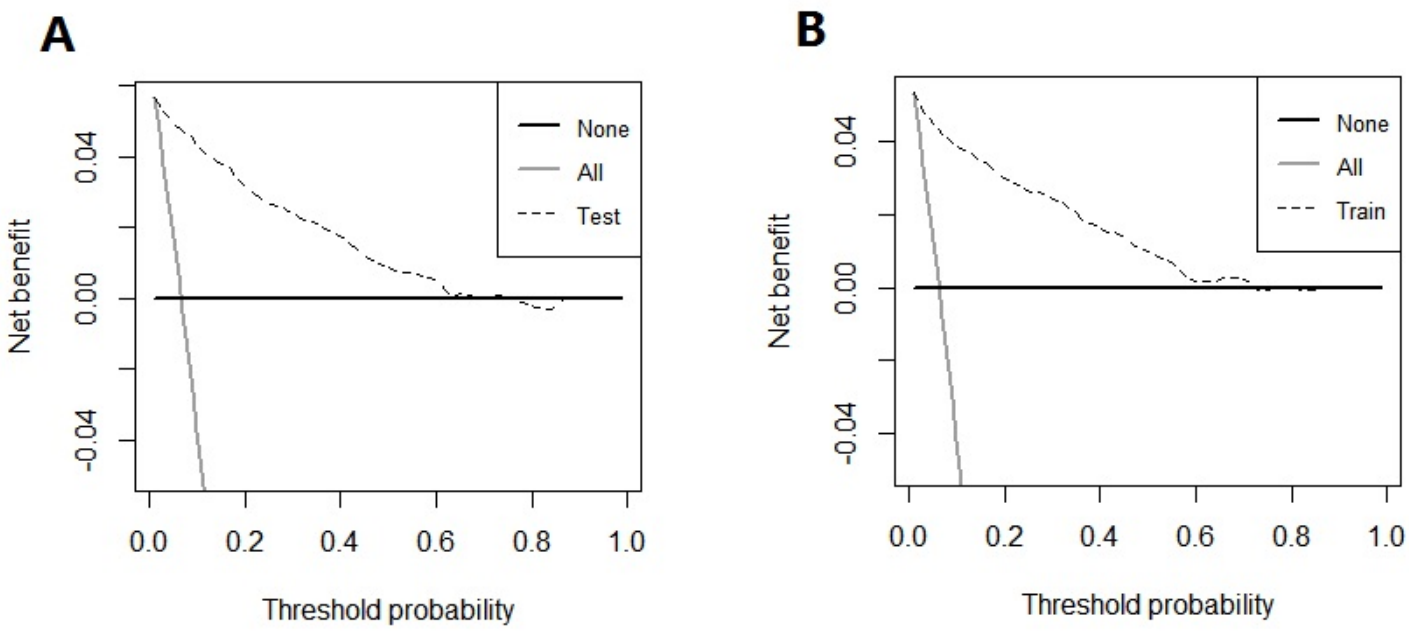

Figure 4

The DCA curves of the nomogram for predicting the occurrence of DM in the development cohort (A) and validation cohort (B)

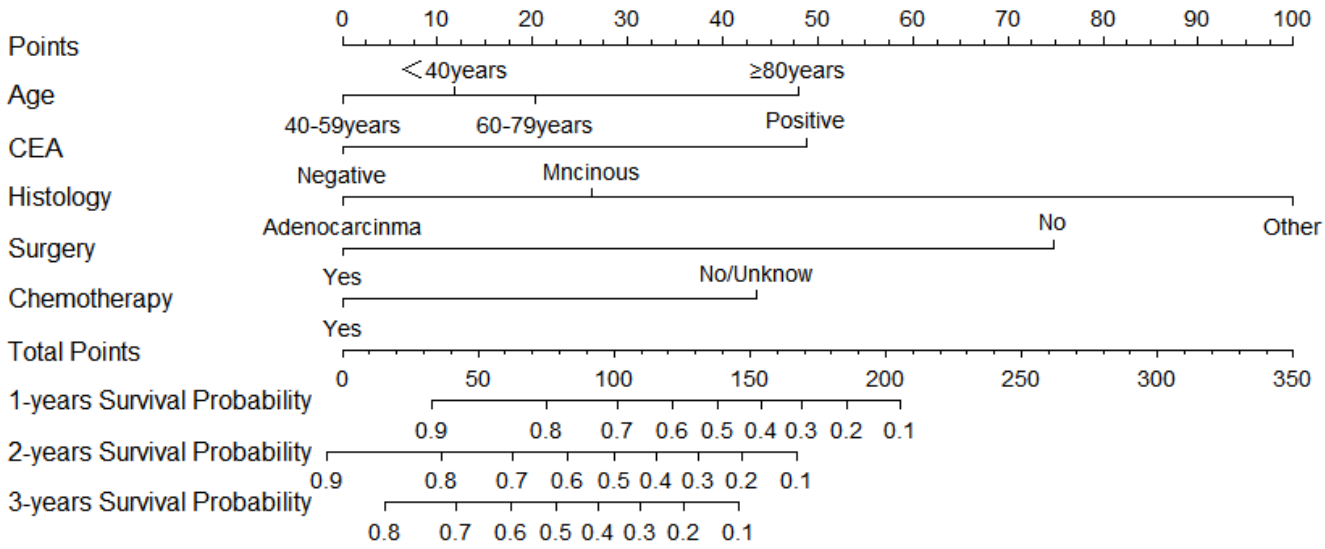

Figure 5

Nomogram for predicting 1-, 2- and 3-year CSS of T1N0 patients with DM

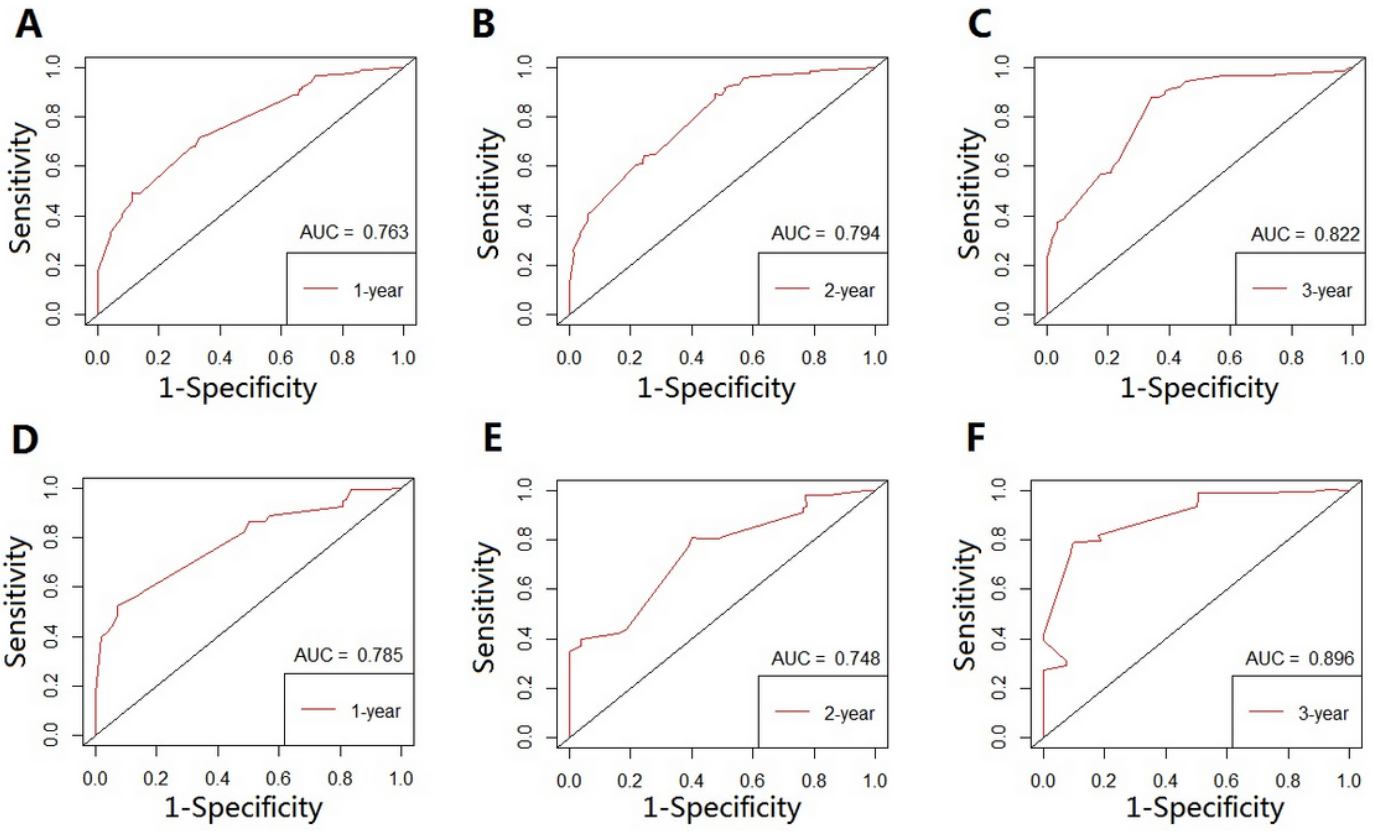

Figure 6

The ROC curves of nomogram for predicting 1-, 2- and 3-year CSS in the development cohort (A-C) and validation cohort (D-F)

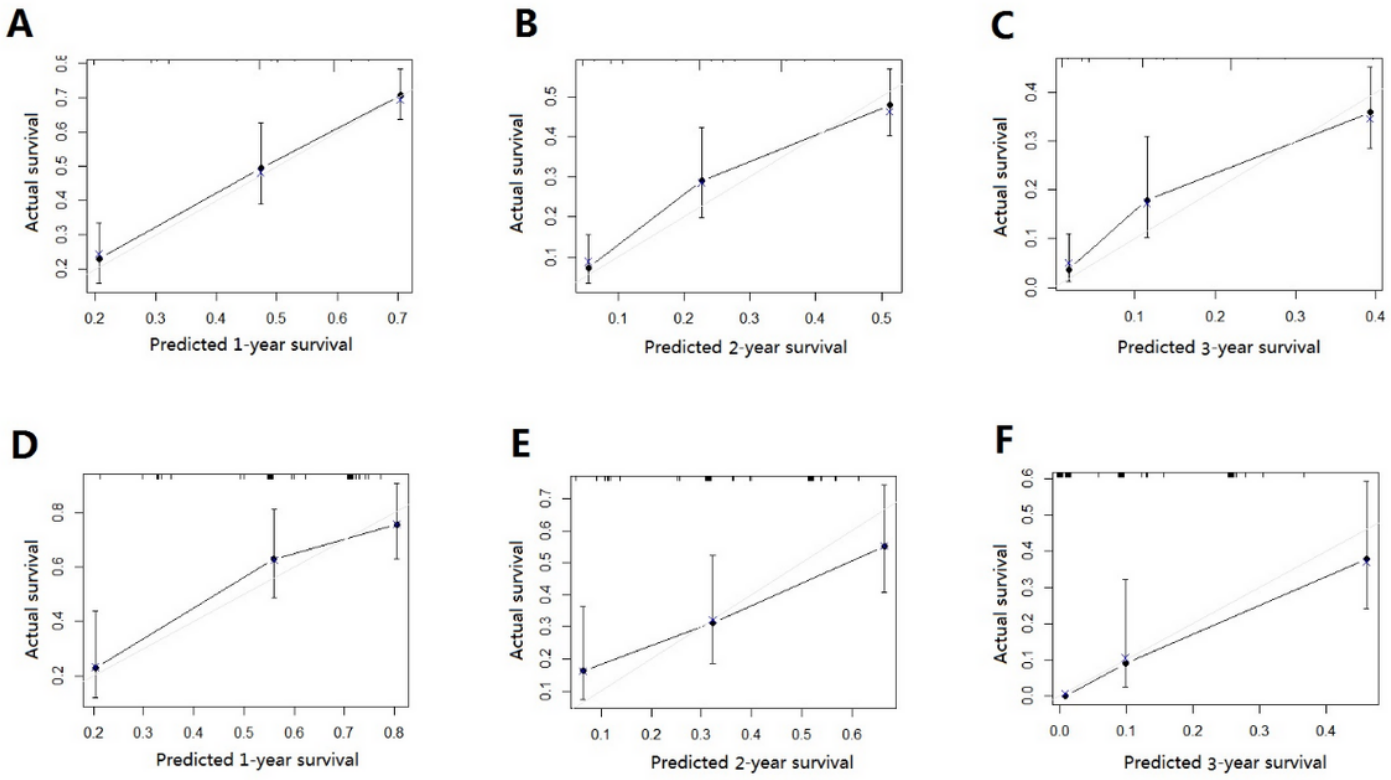

Figure 7

The calibration curves of the nomogram for predicting 1-, 2- and 3-year CSS in the development cohort (A-C) and validation cohort (D-F)

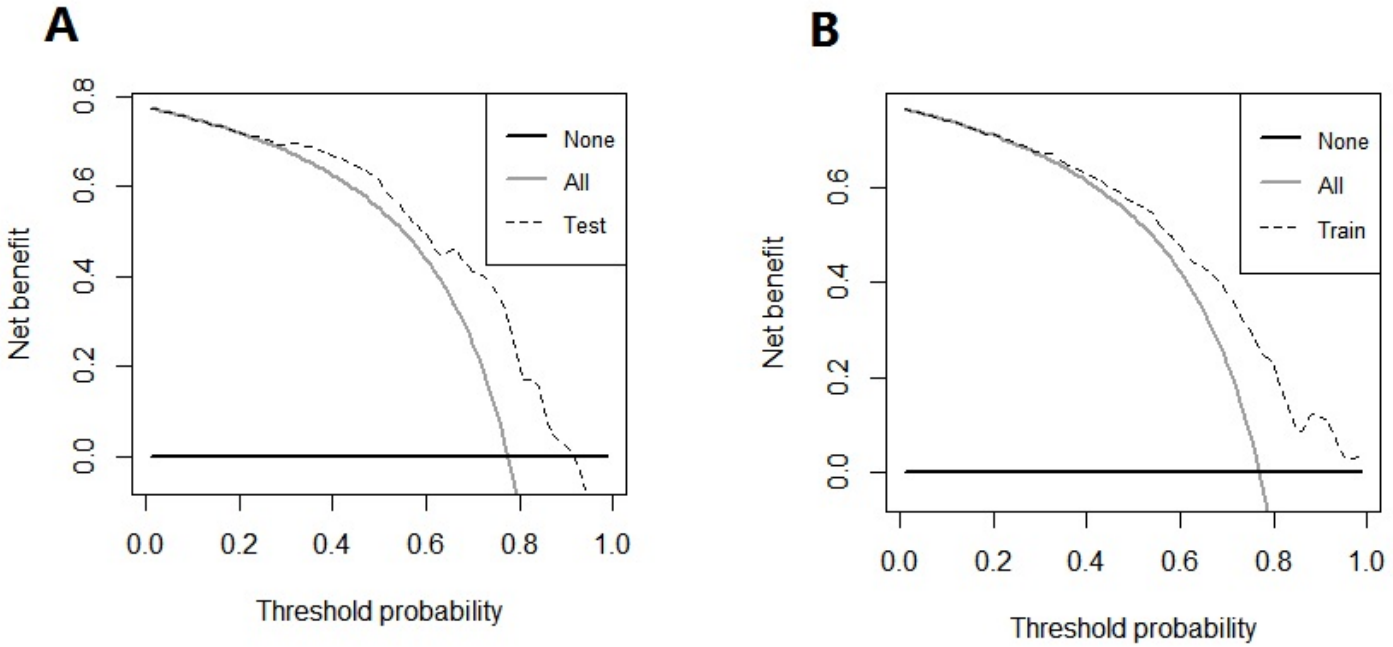

Figure 8

The DCA curves of the nomogram for predicting 1-, 2- and 3-year CSS in the development cohort (A) and validation cohort (B)

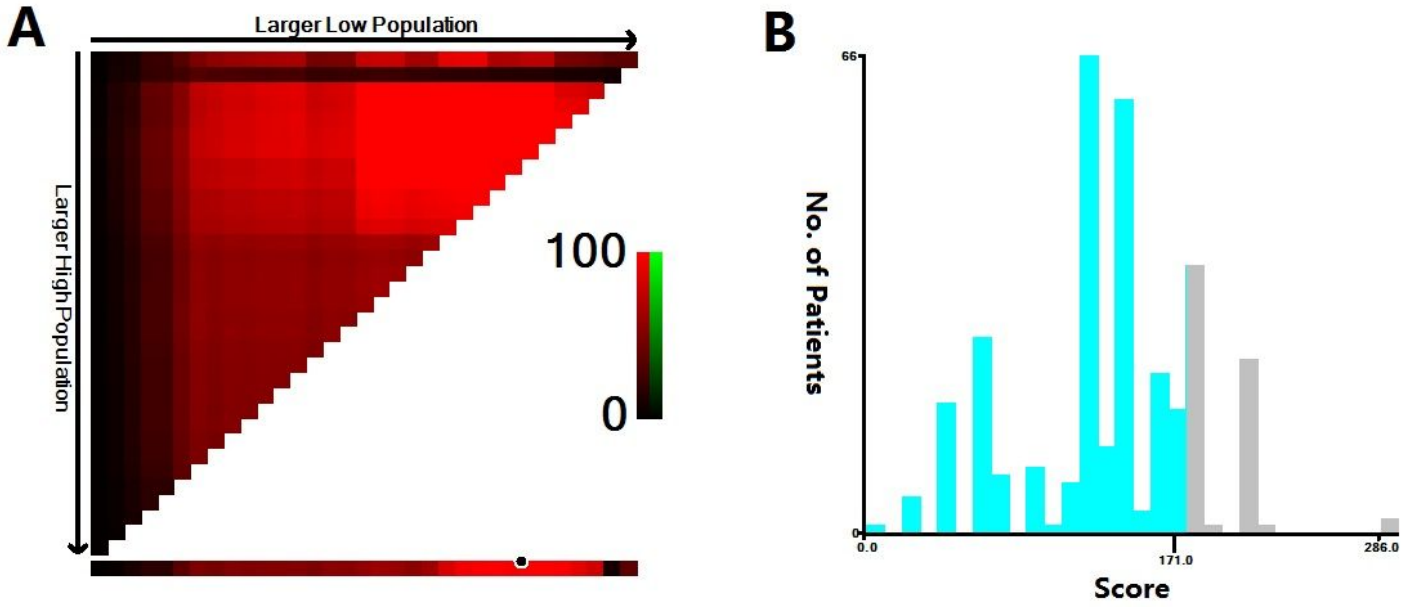

Figure 9

Calculate the cutoff value in DM patients by X-tile program

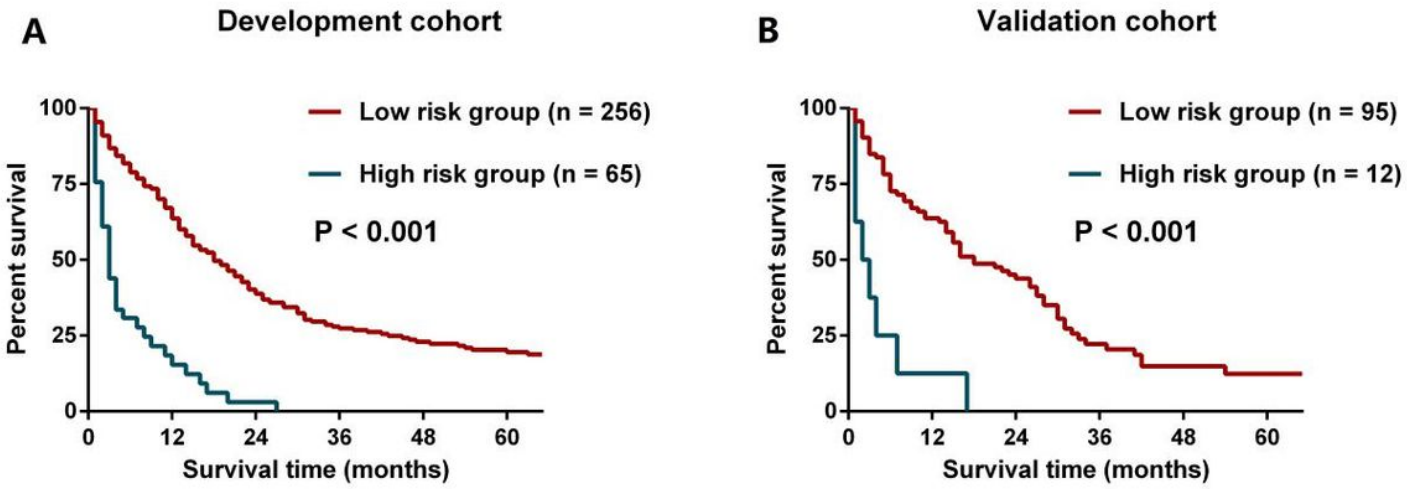

Figure 10

Kaplan-Meier curves for DM patients in the low- and high-risk groups in the development cohort (A) and validation cohort (B)