Predictive Nomogram for the Prediction of Early Recurrence of Colorectal Cancer

Shangjun Tang1,*
Yongjun Chen1,*
Shan Tian2
Yumei Wang1

1Department of Gastroenterology, Qianjiang Central Hospital of Chongqing Municipality, Chongqing, 409099, People’s Republic of China; 2Department of Infectious Disease, Wuhan Union Hospital, Wuhan, 430030, People’s Republic of China

*These authors contributed equally to this work

Aim: The prognosis of colorectal cancer (CRC) individuals after curative resection is not satisfactory due to the early recurrence. We sought to identify the affecting features of early recurrence in CRC patients.

Methods: A total of 3500 CRC patients underwent curative resection were retrospectively incorporated into our study. Among them, 246 patients exhibited tumor recurrence: 121 had early recurrence (≤1 year after operation) and 125 had late recurrence (>1 year after operation). A total of 246 CRC patients with recurrence were randomly assigned into the training group (N=177) or validation group (N=69) based on the ratio of 7:3. LASSO COX regression and support vector machine (SVM) were utilized to screen for the significant clinical indexes associated with the presence of early recurrence. Recurrent nomogram was created based on the above informative parameters to predict the probability of early recurrence.

Results: Proportion of advanced TNM stage, platelet count, systemic immune-inflammation index (SII), mean corpuscular hemoglobin concentration (MCHC), CA-199, CA-125, lactate dehydrogenase, total bile acid (TBA), urea nitrogen were significantly higher in early recurrence group compared with that in late recurrence group. Results from LASSO COX regression and support vector machine (SVM) revealed that TNM stage, CA-199, CA-125, SII and TBA were strong predictors for the presence of early recurrence among postoperative CRC patients in the training group. The recurrent nomogram based on the five predictors exhibited good predictive performance as calculated by C-index (0.846, 95% CI 0.789–0.902 in the training group and 0.799, 95% CI 0.697–0.902 in the validation group) for the prediction of early recurrence. Moreover, the recurrent nomogram exhibited not only encouraging calibration ability, but also great clinical utility both in the training group and validation group.

Conclusion: TNM stage, CA-199, CA-125, SII and TBA were closely correlated with the presence of early recurrence of CRC patients. The recurrent nomogram held well predictive ability for the identification of CRC patients with early recurrence.

Keywords: colorectal cancer, early recurrence, recurrent nomogram, predictive model

Introduction
Colorectal cancer (CRC) continues to be the third most common cancer with regard to incidence, and constitutes the second leading cause of cancer-associated death in adults.1,2 Although multimodality treatment with surgical resection or chemoradiation has improved substantially, there is still a significant cancer-associated mortality due to the early recurrence.3–6 Nowadays, CRC is a major health problem with increasing prevalence worldwide, early recurrence and unfavorable prognosis. Therefore, precise identification of early recurrence is critical for the optimal management of individuals with CRC.
The period from the surgical treatment to the presence of recurrence has been demonstrated to be strongly correlated with survival of CRC people, particularly among patients within one year of their curative surgery. Quite a few efforts have been attempted to search for reliable biomarkers for the early recurrence of CRC, but no nomogram has been specifically developed for predicting the early recurrence of CRC. Hence, the present study aimed to search for the most related indexes with the early recurrence of CRC based on the LASSO COX regression and support vector machine (SVM). Then, we created the recurrent nomogram for the prediction of early recurrence based on the above clinical features in the training group. Finally, we also verified the predictive efficiency of the recurrent nomogram and its calibration ability in the validation group.

### Patients and Methods

#### Study Population

We retrospectively collected data of 3500 CRC patients from Wuhan Union Hospital cohort during 2013 to 2017. The clinical study was executed in line with the Helsinki Declaration. Patients with CRC all gave their informed consent to our clinical research, and our study plan was approved by the clinical research ethics committee of Wuhan Union Hospital (No. 2018-S377). The inclusion criteria were: (1) CRC as the only cancer diagnosis; (2) individuals with CRC who accepted the curative intestinal resection; (3) CRC individuals with tumor recurrence. On the contrary, the exclusion criteria were: (1) CRC patients who are younger than 18 years old or over 80 years old; (2) CRC individuals accompanied by infectious or hematological diseases before surgery; (3) CRC patients who received anti-inflammatory agents before curative surgery; (4) CRC patients lacking significant clinical data or follow-up information. Finally, we recurred 246 cases of postoperative CRC patients with tumor recurrence into our analysis according to the strict standard of nanofiltration. Early recurrence and late recurrence of CRC were defined as recurrence at ≤1 and >1 year, respectively, after curative intestinal resection. This study included 121 cases of CRC with early recurrence and 125 cases of CRC with late recurrence. Then, these participants were randomly assigned into the training cohort (N=177) or validation cohort (N=69) based on the ratio of 7:3.

#### Data Collection

The demographic data, tumor features and laboratory examination indexes were retrospectively collected, including gender, onset age, current smoker, primary site of tumor, family history of CRC, histological grade, tumor size, vascular invasion, TNM stage, radiotherapy, chemotherapy, time of recurrence, blood routine, tumor biomarkers, liver function, renal function and inflammatory indexes. The optimal grouping cut-off values of continuous features were measured by X-tile software to evaluate the true effect of clinical features on survival outcomes in CRC patients. The 8th edition of American Joint Committee on Cancer’s TNM staging system was applied for CRC staging.

#### Construction of the Recurrent Nomogram

The recurrent nomogram was built using a three-step approach. First, we employed LASSO Cox regression to identify the potentially informative differential indexes which were closely correlated to early recurrence of CRC in the training group. Then, we also adopted SVM regression model to rank the informative differential features on the basis of their permutation importance in the training group. In order to avoid the bias caused by single regression Cox model, we only selected the overlapping features to develop the recurrent nomogram which could provide the oncologists with an intuitive and quantitative prediction tool to identify the early recurrent CRC patients. Finally, we internally verified the predictive efficiency and calibration ability of the recurrent nomogram in the validation cohort.

#### Statistical Analysis

Collected data were analyzed using SPSS (version 21.0) in combination with R software. Measurement data conforming to normal distribution were represented with mean±standard deviation and assessed by independent sample t-test, while enumeration data were shown as number with percentage and evaluated by Pearson chi-square or Fisher exact test. X-tile software (version 3.6.1) was utilized to determine the optimal cut-off value of the selected variables. Kaplan-Meier curves in combination with Log rank tests were exploited to compare the differences in survival time between the early recurrent and late recurrent groups.
Results

Patient Characteristics

Among 3500 patients with CRC from Wuhan Union Hospital, we finally selected 246 postoperative individuals with tumor recurrence. The detailed selection process is vividly listed in Figure 1. In order to identify the clinical parameters associated with early recurrence, we divided the included individuals into early recurrence and late recurrence groups based on the cut-off value of 1 year. The student t test and chi-square test were utilized to compare the difference of clinical variables between the early recurrence and late recurrence groups. As exhibited in Table 1, the proportions of TNM stage and adjuvant chemotherapy were statistically different between the early recurrence and late recurrence groups. In addition, platelet count, systemic immune-inflammation index (SII: platelets count×neutrophil count/lymphocyte), mean corpuscular hemoglobin concentration (MCHC), CA-199, CA-125, lactate dehydrogenase, total bile acid (TBA), blood urea nitrogen (BUN) were significantly higher in early recurrence group compared with that in late recurrence group. We noticed that CRC patients with early recurrence had higher values of these markers.

Figure 1 Detailed flow path of CRC patient selection based on the standard of nanofiltration.
| Characteristics                  | Early Recurrence (n=121) | Late Recurrence (n=125) | P value |
|---------------------------------|--------------------------|-------------------------|---------|
| Age (years), n (%)              |                          |                         |         |
| ≥60                             | 46 (38.0)                | 61 (48.8)               | 0.089   |
| <60                             | 75 (62.0)                | 64 (51.2)               |         |
| Sex, male, n (%)                |                          |                         | 0.263   |
|                                 | 65 (53.7)                | 76 (60.8)               |         |
| Primary site, n (%)             |                          |                         | 0.360   |
| Left colon                      | 47 (38.8)                | 45 (36.0)               |         |
| Right colon                     | 28 (23.1)                | 23 (18.4)               |         |
| Rectum                          | 46 (38.0)                | 57 (45.6)               |         |
| Family history of cancer, n (%) |                          |                         | 0.761   |
|                                 | 11 (9.1)                 | 10 (8.0)                |         |
| Current smoker, n (%)           |                          |                         | 0.249   |
|                                 | 10 (8.3)                 | 16 (12.8)               |         |
| Histological grade, n (%)       |                          |                         | 0.534   |
| Well differentiated             | 17 (14.0)                | 18 (14.4)               |         |
| Moderately differentiated       | 98 (81.0)                | 97 (77.6)               |         |
| Poorly differentiated           | 6 (5.0)                  | 10 (8.0)                |         |
| Tumor size, n (%)               |                          |                         |         |
| <2cm                            | 9 (7.4)                  | 6 (4.8)                 | 0.724   |
| 2–5cm                           | 67 (55.4)                | 70 (56.0)               |         |
| ≥5cm                            | 45 (37.2)                | 49 (39.2)               |         |
| Vascular invasion, n (%)        |                          |                         | 0.777   |
| Yes                             | 22 (18.2)                | 21 (16.8)               |         |
| No                              | 99 (81.8)                | 104 (83.2)              |         |
| T stage, n (%)                  |                          |                         | 0.950   |
| T1                              | 16 (13.2)                | 17 (13.6)               |         |
| T2                              | 27 (22.3)                | 20 (16.0)               |         |
| T3                              | 52 (43.0)                | 69 (55.2)               |         |
| T4                              | 26 (21.5)                | 19 (15.2)               |         |
| N stage, n (%)                  |                          |                         | 0.580   |
| N1                              | 73 (60.3)                | 67 (53.6)               |         |
| N2                              | 27 (22.3)                | 38 (30.4)               |         |
| N3                              | 21 (17.4)                | 20 (16.0)               |         |
| TNM stage, n (%)                |                          |                         | <0.001  |
| Stage I                         | 0 (0.0)                  | 7 (5.6)                 |         |
| Stage II                        | 9 (7.4)                  | 28 (22.4)               |         |
| Stage III                       | 39 (32.2)                | 46 (36.8)               |         |
| Stage IV                        | 73 (60.4)                | 44 (35.2)               |         |
| Adjuvant chemotherapy, n (%)    |                          |                         | 0.003   |
| Yes                             | 70 (57.9)                | 49 (39.2)               |         |
| No                              | 51 (42.1)                | 76 (60.8)               |         |
| Radiotherapy, n (%)             |                          |                         | 0.315   |
| Yes                             | 12 (9.9)                 | 8 (6.4)                 |         |
| No                              | 109 (90.1)               | 117 (93.6)              |         |
| Laboratory results              |                          |                         | 0.724   |
| Leukocyte, (× 10^9/L)           | 6.9±2.8                  | 7.1±3.0                 |         |
| Neutrophil, (× 10^9/L)          | 6.1±2.8                  | 4.6±3.0                 | 0.101   |

(Continued)
recurrence experienced shorter overall survival time than those with late recurrence (P < 0.0001). For the sake of building a recurrent nomogram and verifying its predictive efficiency and calibration ability, we randomly assigned these CRC patients to the training group (N=177) or validation group (N=69) according to the ratio of 7:3.

Identification of Significant Features

LASSO Cox regression was performed to identify factors that were significantly correlated with early recurrence of CRC in the training group. As vividly shown in Figure 2A, basophil, monocytes to lymphocytes ratio (MLR), triglyceride, MCHC, alanine aminotransferase to lymphocyte ratio index (ALRI), CA-724, GGT, BUN, TNM stage,
CA-199, CA125, SII and TBA were all influencing factors of early recurrence. In order to construct an easy-to-use predictive model with relatively high accuracy, we also applied the SVM model to screen for the significant indexes associated with early recurrence of CRC. Results from SVM algorithm showed that six clinical parameters were screened out by this regression model, including CEA, TNM stage, CA-199, CA125, SII and TBA (Figure 2B).

**Construction and Validation of Recurrent Nomogram**

We only included the overlapping features (TNM stage, CA-199, CA125, SII and TBA) selected by LASSO Cox regression model and SVM algorithm into the formation of recurrent nomogram (Figure 2C). As displayed in Figure 3, all the five features exhibited great prognostic significance among CRC patients with recurrent tumor. Based on the five informative indexes, we created a recurrent nomogram, which could intuitively predict the probability of early recurrence (Figure 4). The predictive performance of the recurrent nomogram as measured by C-index was 0.846 (95% CI 0.789–0.902) in the training group and 0.799 (95% CI 0.697–0.902) in the validation group for the prediction of early recurrence in patients with CRC. As far as we know, this is the first clinical study to explore the affecting features of early recurrence among CRC patients. For the first time, we utilized LASSO Cox and SVM models to build a recurrent nomogram based on TNM stage, CA-199, CA125, SII and TBA. This nomogram possessed well predictive ability for the identification of CRC patients with early recurrence.

Although TNM stage, CA-199, CA125 and SII were well-recognized prognostic factors in CRC patients, we also identified that total bile acid was a strong prognostic factor for patients with CRC. Tilman et al20 reported that

**Discussion**

Nomogram, an easy-to-use predictive model, is widely applied in the diagnosis of cancer and prediction of survival.15 However, most of them focused on overall survival,16–19 and no nomogram is specifically created for the prediction of early recurrence in patients with CRC. As far as we know, this is the first clinical study to explore the affecting features of early recurrence among CRC patients. For the first time, we utilized LASSO Cox and SVM models to build a recurrent nomogram based on TNM stage, CA-199, CA125, SII and TBA. This nomogram possessed well predictive ability for the identification of CRC patients with early recurrence.
levels of certain conjugated primary and secondary bile acids were positively associated with risk of CRC. Jia et al. demonstrated that bile acids were involved in the pathogenesis of CRC. Liu et al. summarized the critical role of bile acids in the occurrence and progression of CRC, and also highlighted the rationale of multiple interventions for the management of CRC patients by regulating bile acids-microbiota axis. In our study, we emphasized the prognostic role of bile acid in patients with CRC for the first time.

Early recurrence after surgical resection is common and negatively correlated with the survival among patients...
Early recurrence of CRC may result from inadequate surgical resection, aggressive tumor biology, and failure of systemic chemotherapy or radiotherapy. Early recurrence of CRC may also be an indicator of suboptimal preoperative staging. Colonoscopy is the endoscopic modality for postoperative monitoring, especially for tumor recurrence. However, except for the relatively high cost of colonoscopy, quite a few CRC patients underwent curative surgery cannot bear colonoscopy. Hence, searching for the noninvasive indexes for the precise prediction of early recurrence for postoperative individuals with CRC is quite significant.

Considerable number of clinical studies have investigated the independently affecting features of early recurrence in CRC patients with colorectal liver metastases (CLM). A recent French study with 6025 cases of participants had shown that advanced T stage, limited resection margins and synchronous CLM increase the risk of early recurrence. Additionally, Malik and its coworkers indicated that 20% patients with CLM developed early recurrence after hepatic resection, and the presence of eight or more metastases was a strong prognostic factor of early recurrence via multivariate regression analysis. Bhogal et al reported that 11% patients with CLM developed early recurrence after hepatic resection, and tumor number (>2) and tumor size (≥3.6 cm) were significantly correlated with early recurrence of CLM. Lan et al concluded that CRC patients with early recurrence possessed advanced N stage, more liver metastases and higher mutation rate of APC mutation than those with late recurrence. Bozkurt et al investigated the clinical characteristics in CRC patients with recurrence within 1 year and over 1 year after surgery, and they reported that inadequate lymph node dissection and K-ras mutation was significantly correlated with the presence of early recurrence. To our knowledge, no study has systemically depicted the prognostic variables of early recurrence in postoperative patients with CRC. Through LASSO Cox and SVM
models, we found that TNM stage, CA-199, CA125, SII and TBA were strong prognostic variables of early recurrence in postoperative patients with CRC.

Although CRC patients with early recurrence usually signify unfavorable survival,\textsuperscript{10} it is not easy to accurately identify those patients with high risk of early recurrence. Quite a few features could affect the presence of early recurrence in patients with CRC, and construction of the recurrence predictive model incorporating highly correlated factors is of great clinical significance. On the one hand, use of the recurrent nomogram is expected to improve the survival of postoperative CRC patients who might benefit from systematic chemotherapy or radiotherapy.\textsuperscript{23,28} On the other hand, clinicians should shorten the follow-up period for the early recurrent patients with CRC, as these individuals might progress more rapidly than those with late recurrence.

Inevitably, several limitations still existed in our study. First, the clinical information is from a retrospective cohort rather than a prospective cohort, so the inherent bias of selection might limit our study. Second, the clinical data were from a single cohort (Wuhan Union Hospital), no data from other hospitals were available for the external validation. Finally, although some immunohistochemical indexes, such as Ki-67, P53, HER-2, VEGF, might be correlated with the recurrence of CRC. Unfortunately, they were unavailable in our study. Hence, prospective clinical trials from multicenter are needed to verify the recurrent nomogram in the near future.

Conclusion

TNM stage, CA-199, CA125, SII and TBA were closely correlated with the presence of early recurrence of CRC patients. The recurrent nomogram held encouraging predictive ability for the identification of CRC patients with early recurrence. The recurrent nomogram might provide the oncologists significant reference for appropriate post-operative treatment plans for CRC patients.

Disclosure

The authors declared no conflicts of interest in this work.

References

1. Thanikachalam K, Khan G. Colorectal cancer and nutrition. Nutrients. 2019;11(1):164. doi:10.3390/nu11010164
2. Kudo SE, Ichimasa K, Mori Y. Artificial intelligence and prediction of the risk of lymph node metastasis of endoscopically removed T1 colorectal cancer. Gastroenterology. 2021;160:1075–1084.e2. doi:10.1053/j.gastro.2020.09.027
3. Chang LC, Shun CT, Lin BR, et al. Recurrence outcomes less favorable in T1 rectal cancer than in T1 colon cancer. Oncolocist. 2021. doi:10.1002/onco.13815
4. Vega EA, Salehi O, Nicolausco D, et al. Failure to cure patients with colorectal liver metastases: the impact of the liver surgeon. Ann Surg Oncol. In Press 2021. doi:10.1245/s10434-021-10030-0
5. Tie J, Wang Y, Cohen J, et al. Circulating tumor DNA dynamics and recurrence risk in patients undergoing curative intent resection of colorectal cancer liver metastases: a prospective cohort study. PLoS Med. 2021;18(5):e1003620. doi:10.1371/journal.pmed.1003620
6. Westberg K, Othman B, Suzuki C, et al. Magnetic resonance imaging as a predictor of surgical outcome in patients with pelvic recurrence of colorectal cancer. Eur J Surg Oncol. 2021;47:2119–2124. doi:10.1016/j.ejso.2021.04.021
7. Kobayashi H, Mochizuki H, Sugihara K, et al. Characteristics of recurrence and surveillance tools after curative resection for colorectal cancer: a multicenter study. Surgery. 2007;141(1):67–75. doi:10.1016/j.surg.2006.07.020
8. Yang IP, Tsai HL, Huang CW, et al. The functional significance of microRNA-29c in patients with colorectal cancer: a potential circulating biomarker for predicting early relapse. PLoS One. 2013;8(6):e66842. doi:10.1371/journal.pone.0066842
9. Imai K, Allard MA, Benitez CC, et al. Early recurrence after hepatectomy for colorectal liver metastases: what optimal definition and what predictive factors? Oncologist. 2016;21(7):887–894. doi:10.1634/theoncologist.2015-0468
10. Bozkurt O, Inanc M, Turkmen E, et al. Clinicopathological characteristics and prognosis of patients according to recurrence time after curative resection for colorectal cancer. Asian Pac J Cancer Prev. 2014;15(21):9277–9281. doi:10.7314/APJCP.2014.15.21.9277
11. Gong P, Chen C, Wang Z, et al. Prognostic significance for colorectal carcinoid tumors based on the 8th edition TNM staging system. Cancer Med. 2020;9(21):7979–7987. doi:10.1002/cam4.3431
12. Heyne K, Lasson U. [Congenital syphilis plus BCG immunization–clinico-immunological aspects of a case]. Die Medizinische Welt. 1977;28(36):1442–1443. German.
13. Liang F, Guan P, Wu W, et al. Forecasting influenza epidemics by integrating internet search queries and traditional surveillance data with the support vector machine regression model in Liaoning, from 2011 to 2015. PEARJ. 2018;6:e5134. doi:10.7717/pearj.e5134
14. Pan C, Wang X, Chen W, et al. Reevaluation of glypican-3 as a prognostic marker in HCC using X-tile software. Med Oncol. 2015;32(1):359. doi:10.1007/s12032-014-0359-z
15. Wang S, Tian S, Li Y, et al. Development and validation of a novel scoring system developed from a nomogram to identify malignant pleural effusion. EBioMed. 2020;58:102924. doi:10.1016/j.ebiom.2020.102924
16. Zhu J, Hao J, Ma Q, et al. A novel prognostic model and practical nomogram for predicting the outcomes of colorectal cancer: based on tumor biomarkers and log odds of positive lymph node scheme. Front Oncol. 2021;11:661040. doi:10.3389/fonc.2021.661040
17. Ma Y, Lu P, Liang X, et al. Models based on dynamic clinicopathological indices for predicting prognosis during the perioperative period for patients with colorectal cancer. J Inflamm Res. 2021;14:1591–1601. doi:10.2147/JIR.S302435
18. Kou FR, Zhang YZ, Xu WR. Prognostic nomograms for predicting overall survival and cause-specific survival of signet ring cell carcinoma in colorectal cancer patients. World J Clin Cases. 2021;9(11):2503–2518. doi:10.2998/wjcc.v9.i11.2503
19. Liu Z, Xu Y, Xu G, et al. Nomogram for predicting overall survival in colorectal cancer with distant metastasis. BMC Gastroenterol. 2021;21(1):103. doi:10.1186/s12876-021-01692-x
20. Kuhn T, Stepien M, Lopez-Nogueroles M, et al. Prediagnostic plasma bile acid levels and colon cancer risk: a prospective study. J Natl Cancer Inst. 2020;112(5):516–524. doi:10.1093/jnci/djz166
21. Jia W, Xie G, Jia W. Bile acid-microbiota crosstalk in gastrointestinal inflammation and carcinogenesis. Nat Rev Gastroenterol Hepatol. 2018;15(2):111–128. doi:10.1038/nrgastro.2017.119

22. Liu T, Song X, Khan S, et al. The gut microbiota at the intersection of bile acids and intestinal carcinogenesis: an old story, yet mesmerizing. Int J Cancer. 2020;146(7):1780–1790. doi:10.1002/ijc.32563

23. Malik HZ, Gomez D, Wong V, et al. Predictors of early disease recurrence following hepatic resection for colorectal cancer metastasis. Eur J Surg Oncol. 2007;33(8):1003–1009. doi:10.1016/j.ejso.2007.01.005

24. Bhogal RH, Hodson J, Bramhall SR, et al. Predictors of early recurrence after resection of colorectal liver metastases. World J Surg Oncol. 2015;13:135. doi:10.1186/s12957-015-0549-y

25. Fong Y, Fortner J, Sun RL, et al. Clinical score for predicting recurrence after hepatic resection for metastatic colorectal cancer: analysis of 1001 consecutive cases. Ann Surg. 1999;230(3):309–318, 318–321. doi:10.1097/00000658-199909000-00004

26. Benamouzig R, Barre S, Saurin JC, et al. Cost-effectiveness analysis of alternative colorectal cancer screening strategies in high-risk individuals. Therap Adv Gastroenterol. 2021;14:1088185335.

27. Huang Y, Cai S, Li Q, et al. Six years of colorectal cancer mortality surveillance in the screening population for a risk stratified screening program. Cancer Epidemiol. 2021;73:101937. doi:10.1016/j.canep.2021.101937

28. Vigano L, Capussotti L, Lapointe R, et al. Early recurrence after liver resection for colorectal metastases: risk factors, prognosis, and treatment. A LiverMetSurvey-based study of 6025 patients. Ann Surg Oncol. 2014;21(4):1276–1286. doi:10.1245/s10434-013-3421-8

29. Lan YT, Chang SC, Lin PC, et al. Clinicopathological and molecular features of patients with early and late recurrence after curative surgery for colorectal cancer. Cancers (Basel). 2021;13(8):1883. doi:10.3390/cancers13081883