A Retrospective Study from a Single Center to Identify Hematological Factors that Distinguish Between Patients with Colorectal Carcinoma and Colorectal Adenoma

Jun Huang
Pingping Xu
Youxiang Chen

Corresponding Author: Youxiang Chen, e-mail: chenyx_gastro@163.com
Financial support: None declared
Conflict of interest: None declared

Background: This retrospective study from a single center in China was conducted using data from medical records between 2012 and 2020, to identify hematological factors that distinguish between patients with colorectal carcinoma (CRC) and colorectal adenoma.

Material/Methods: In this case-control study, 856 eligible patients were randomly divided into a training set (n=600) and a testing set (n=256). Routine blood parameters, blood coagulation, and liver and kidney function parameters were collected. Univariate and multivariate Cox regression analyses were used to explore diagnostic indicators. The values of the area under the curve and calibration curves were used to evaluate the model.

Results: The study included 325 colorectal adenoma and 531 CRC patients. The prediction model for diagnosing CRC using hemoglobin-to-platelet ratio, fibrinogen-albumin ratio (FAR), albumin-globulin ratio (A/G), platelet-lymphocyte ratio, carcinoembryonic antigen (CEA), and thrombin time (TT) was developed on the basis of the patients grouped into the CRC and colorectal adenoma groups. The prediction model for diagnosing CRC stage was developed using prothrombin time (PT), TT, CEA, A/G, FAR, and HPR. The prediction model for diagnosing CRC grade was developed using PT, TT, A/G, plateletcrit, FAR, and HPR. The AUCs of the 3 prediction models were [0.848, 95% CI: (0.800-0.896)], [0.806, 95% CI: (0.775-0.836)], and [0.829, 95% CI: (0.797-0.860)] in the testing set.

Conclusions: Three diagnostic prediction models for early screening of CRC, stage of CRC, and grade of CRC were established through a panel of readily available hematological parameters, which could provide auxiliary tools for early screening of CRC.

Keywords: Colorectal Neoplasms • Diagnosis • Polyposis Syndrome, Hereditary Mixed, 1

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/936745
**Background**

Colorectal carcinoma (CRC) is the fourth most common cancer in women and the third most common cancer in men, and CRC is the leading cause of cancer deaths worldwide [1,2]. In 2020, there were an estimated 0.55 million new cases of CRC, and 0.28 million deaths, in China [3,4]. CRC disease progresses slowly, and many patients do not have clinical manifestations such as abdominal pain or intestinal bleeding until the cancer metastasizes, leading to delay in treatment [5]. CRC mostly results from colon or rectal epithelial polyps, of which colorectal adenoma is the most common neoplastic polyp, and is associated with an increased risk of CRC [6,7]. According to the National Comprehensive Cancer Network recommendations, early screening could prevent CRC by detecting colorectal lesions early and treating them appropriately [8]. Colonoscopy is the gold standard for screening for CRC and distinguishing between CRC and colorectal adenoma [2,9]. However, its application in early screening is limited due to its high cost and inconvenience, and the fact that it can easily lead to unnecessary treatments [10,11]. Therefore, the development of non-invasive and easily available markers to distinguish CRC from colorectal adenoma is of great significance to avoid unnecessary treatments and improve the accuracy of CRC detection [12]. Blood is invariably the most desirable analyte for cancer biomarkers due to its availability and high patient acceptance [12].

Studies have shown several hematological factors to be correlated with poor survival rates of CRC and other cancers, including platelet-to-lymphocyte (LYM) ratio (PLR) [13], monocyte (MONO)-to-LYM ratio (MLR) [14], neutrophil (NEUT)-to-LYM ratio (NLR) [15], systemic immune-inflammatory index (SII) [16], and others. Recently, the relationship between platelet-related parameters (platelets, mean platelet volume (MPV), and plateletcrit, the product of platelets and MPV) and CRC has also been reported [17,18]. Wu et al [19] stated that the MPV-to-platelets ratio could be used as an effective diagnostic index to distinguish benign from malignant CRC. In addition, carcinomaembryonic antigen (CEA) is a common tumor marker that has assisted in the diagnosis of CRC [20]. A machine-learning model using biochemical and hematological markers was built to identify CRC patients [21]. Several studies have established models of diagnosing CRC based on inflammation factors, platelet-related markers, or CEA [22,23], but most studies have excluded people with inflammatory diseases. Clinically, some patients with suspected CRC have inflammatory diseases. Excluding patients with inflammatory diseases may limit the application of the prediction model to the population. The performance of the existing prediction model of CRC diagnosis is limited in its application in clinical practice, and needs to be further improved.

Therefore, this retrospective study from a single center in China was conducted using data from patient medical records between 2012 to 2020, to identify hematological factors that distinguish between patients with colorectal carcinoma and CRC.

**Material and Methods**

**Study Population**

This case-control study was approved by the Ethics Committee of Jiangxi Cancer Hospital (Affiliated Cancer Hospital of Nanchang University), and informed consent was obtained from all the participants. Demographic and clinical data for 980 patients were collected from the electronic medical records of Jiangxi Cancer Hospital (Affiliated Cancer Hospital of Nanchang University) from May 2012 to September 2020. The inclusion criteria for the CRC and colorectal adenoma groups were patients who had not received anti-tumor therapy and were diagnosed with CRC or colorectal adenoma by histopathology. The exclusion criteria were: (1) patient age less than 18 years, (2) patients diagnosed with distant metastasis of CRC, (3) patients with hematological disease, and (4) patients with a history of blood transfusion within 3 months. According to the inclusion and exclusion criteria, 856 CRC and colorectal adenoma patients were finally included in this study (Figure 1). Then, these patients were randomly divided into 2 groups: the training set (n=600) and the testing set (n=256), at a ratio of 7: 3.

**Diagnosis of CRC and Colorectal Adenoma**

CRC and colorectal adenoma were diagnosed into distinct pathological types assigned after surgical resection or endoscopic resection. The histological types were categorized into either colorectal adenoma, on the basis of the International Classification of Diseases (ICD-0) (8210/0 and 8210/2), ICD-11 (E292.4Y & XH7S1), or CRC [ICD-0 (8140/3), ICD-11 (2890.Y & XH7S1)]. The diagnosis of CRC was according to the eighth edition of the American Joint Committee on Cancer/Tumor Node, Metastases tumor staging (including pathological staging, degree of differentiation, lymph node metastasis, vascular invasion, and other values).

**Data Collection**

We collected age, gender, body mass index [BMI, (kg/m²)], routine blood parameters, blood coagulation parameters, liver and kidney function parameters, and tumor biomarkers for eligible patients. The routine blood parameters and blood coagulation parameters were hemoglobin [Hb (g/L)], red blood cell count [RBC (10¹²/L)], white blood cell count [WBC (10⁹/L)], NEUT (10⁹/L), LYM (10⁹/L), MONO (10⁹/L), platelets (10¹²/L), MPV (FL), plateletcrit (%), red blood cell distribution width [RDW (%)], prothrombin time [PT (s)], activated partial thromboplastin time [APTT (s)], thrombin time [TT (s)], fibrinogen...
[FIB (g/L)], and D-dimer (mg/L). Routine blood and blood coagulation parameters were measured by a Mindray 5800 plus automatic blood cell analyzer and Sysmex CS5001 blood coagulation analyzer, respectively. Liver and kidney function parameters included glutamic pyruvic transaminase [GPT (U/L)], glutamic oxaloacetic transaminase [GOT (U/L)], total bilirubin [TBil (μmol/L)], albumin [ALB (g/L)], globulin [GLB (g/L)], ALB/ GLB (A/G), blood urea nitrogen [BUN (mmol/L)], creatinine [Cr (μmol/L)], and uric acid [UA (μmol/L)]. A Hitachi 7600 automatic biochemical analyzer was used for the determination of liver and kidney function parameters. Tumor biomarkers collected were: CEA (ng/mL), carbohydrate antigen 199 (U/mL), and carbohydrate antigen 724 (U/mL). Tumor biomarkers were measured using the Siemens ADVIA Centaur XPT. The values of NLR, PLR, MLR, FIB-to-ALB ratio (FAR), 10×ALB+0.005×LYM (PNI), SII, NEUT/(WBC-NEUT) (dNLR) and hemoglobin-to-platelets ratio (HPR) were calculated. In addition, we also collected data on whether eligible patients had diabetes or inflammatory disease. There were 600 patients in the training set and 256 patients in the testing set. No statistical differences between groups (P>0.05) were included in logistic regression analysis for stepwise regression to screen out diagnostic indicators of CRC. Three prediction models were developed by using the diagnostic indicators of CRC in the different groups. The receiver operating characteristic (ROC) curves, calibration curves, and the value of area under the curves (AUCs) were used to validate the performance of 3 prediction models to differentiate CRC from colorectal adenoma. We also established a nomogram for CRC diagnosis in patients grouped into CRC or colorectal adenoma. Three prediction models were validated in participants with and without inflammatory disease using AUCs. Power analysis was calculated using the existing sample size to test the efficiency of the AUC. The statistical power value achieved was 0.89, indicating that the sample size was adequate to support the results.

All statistical tests used two-sided tests, and P<0.05 was considered to be a statistically significant difference. SAS v. 9.4 (SAS Institute, Cary, North Carolina) was used for data sorting and construction of prediction models, and R v. 4.20 (Foundation for Statistical Computing, Vienna, Austria) was used for model validation and image production.

Statistical Analysis
Quantitative data were described by mean±standard deviation (mean±SD) or median (interquartile range) [M (Q1, Q3)]. Categorical variable data were described by frequency and proportion [n (%)]. The statistical differences between the 2 groups of quantitative data were compared and analyzed using the Mann-Whitney U test or t-test as appropriate, and the categorical variables were compared using the chi-square test. In the training set, participants were divided into 2 groups to compare CRC and colorectal adenoma; 4 groups to compare CRC stage I, II, and III, and colorectal adenoma; and 3 groups for CRC: poorly differentiated, moderately and well-differentiated, and colorectal adenoma. We conducted univariate analyses using the data from the training set. The factors with statistical differences between groups (P<0.05) were included in logistic regression analysis for stepwise regression to screen out diagnostic indicators of CRC. Three prediction models were developed by using the diagnostic indicators of CRC in the different groups. The receiver operating characteristic (ROC) curves, calibration curves, and the value of area under the curves (AUCs) were used to validate the performance of 3 prediction models to differentiate CRC from colorectal adenoma. We also established a nomogram for CRC diagnosis in patients grouped into CRC or colorectal adenoma. Three prediction models were validated in participants with and without inflammatory disease using AUCs. Power analysis was calculated using the existing sample size to test the efficiency of the AUC. The statistical power value achieved was 0.89, indicating that the sample size was adequate to support the results.

Results
Characteristics of the Study Population
A total of 856 patients [325 (37.97%) with colorectal adenoma, 531 (62.03%) with CRC] were included in the study. Among the patients with CRC, 98 (11.45%) were poorly differentiated, 419 (48.95%) were moderately differentiated, and 14 (1.64%) were well differentiated; and 68 (7.94%) had stage I disease, 249 (29.09%) had stage II disease, and 214 (25.00%) had stage III disease. There were 600 patients in the training set and 256 patients in the testing set. No statistical differences between the training and testing sets were observed (Table 1). Of all the
selected patients, 511 (59.70%) were male and 345 (40.30%) were female. The average age was 56.39±12.15 years old, and the average BMI was 22.45±3.38 kg/m². There were 38 (4.44%) participants who had diabetes, and 380 (44.39%) participants who had other inflammatory diseases. More detailed information about the demographic and clinicopathological characteristics of the selected patients is shown in Table 1. There was no statistical difference between the data before and after the missing values were interpolated (Supplementary Table 1).

Comparison of CRC and Colorectal Adenoma Characteristics in the Training Set

As shown in Table 2, significant differences were found between the CRC group and colorectal adenoma group in the following parameters: BMI (P<0.004), Hb (P<0.001), RBC (P<0.001), WBC (P=0.038), NEUT (P<0.004), LYMPH (P<0.001), platelets (P<0.001), MPV (P<0.001), plateletcrit (P<0.001), RDW (P<0.001), PT (P<0.001), APTT (P=0.007), TT (P<0.001), FIB (P<0.001), D-dimer (P<0.001), CEA (P<0.001), carbohydrate antigen 199 (P=0.023), GPT (P<0.001), GOT (P<0.001), THBIL (P<0.001), ALB+GLB (P<0.001), ALB (P=0.004), GLB (P<0.001), A/G (P<0.001), BUN (P<0.001), UA (P<0.001), NLR (P<0.001), PLR (P<0.001), MLR (P=0.002), MPV/platelets (P<0.001), FAR (P<0.001), PNI (P=0.004), SII (P<0.001), dNLR (P<0.001), and HPR (P<0.001).

When comparing the characteristics of CRC stage I, II, and III, and colorectal adenoma, the variables with significant differences were basically the same as the variables with differences between the groups of CRC and colorectal adenoma, except that WBC, APTT, MPV, and plateletcrit were not significantly different (Supplementary Table 2). There were significant differences in the values of WBC and carbohydrate antigen 199 when comparing the characteristics of CRC (poorly differentiated, moderately differentiated, and well-differentiated) and colorectal adenoma (Supplementary Table 3).

Diagnostic Indicators for CRC Patients

In the patients grouped into CRC and colorectal adenoma (Table 3), the variable of HPR was associated with a decreased risk of being diagnosed with CRC [odds ratio (OR) =0.178, 95% confidence interval (CI): 0.065-0.487, (P<0.001)]. The variables of FAR [OR=1.590, 95% CI: 1.406-1.798, (P<0.001)], A/G [OR=1.410, 95% CI: 1.301-1.528, (P<0.001)], PLR [OR=1.004, 95% CI: 1.001-1.007, (P=0.010)], CEA [OR=1.064, 95% CI: 1.016-1.115, (P=0.009)], and TT [OR=1.585, 95% CI: 1.365-1.840, (P<0.001)] were associated with increased risks of being diagnosed with CRC. In the patients grouped into CRC stage I, II, and III, and colorectal adenoma, PT, TT, CEA, A/G, FAR, and HPR were diagnostic indicators (Table 4). In the patients grouped into poorly differentiated CRC, moderately and well-differentiated CRC, and colorectal adenoma, PT, TT, A/G, plateletcrit, FAR, and HPR were the diagnostic indicators (Table 5).

The Model Developed to Diagnose CRC

The 6 diagnostic indicators (HPR, FAR, A/G, PLR, CEA, and TT) of CRC were used to jointly construct a prediction model for diagnosing CRC in the patients grouped into CRC and colorectal adenoma. A nomogram was generated based on the identified diagnostic indicators (Figure 2), which was built on HPR, FAR, A/G, PLR, CEA, and TT. As shown in Figure 3, we randomly selected an observation, the 140th patient. The value of each diagnostic indicator is indicated by a red dot in Figure 3. Each value corresponded to the score of the first row, and then the scores of the diagnostic indicators were added. The resulting total score, denoted Total Points, was 281 in the penultimate row, and the probability that this patient was diagnosed with CRC was 0.986. In the patients grouped into CRC stage I, II, and III, and those with colorectal adenoma, PT, TT, CEA, A/G, FAR, and HPR were used to construct a prediction model for diagnosing CRC. In the patients grouped into poorly differentiated CRC, moderately and well-differentiated CRC, and colorectal adenoma, PT, TT, A/G, plateletcrit, FAR, and HPR were used to construct a prediction model for diagnosing CRC.

Model Evaluation and Verification

The AUCs of the training set [0.864, (95% CI: 0.834-0.895)] and testing set [0.848, (95% CI: 0.800-0.896)] in patients grouped into CRC and colorectal adenoma (Table 6, Figure 4A), the AUCs of the training set [0.818, (95% CI: 0.799-0.838)] and testing set [0.806, (95% CI: 0.775-0.836)] in patients grouped into CRC stage I, II, and III, and colorectal adenoma (Table 6, Figure 4B), the AUCs of the training set [0.857, (95% CI: 0.838-0.876)] and testing set [0.829, (95% CI: 0.797-0.860)] in patients grouped into the poorly differentiated CRC, moderately and well-differentiated CRC, and colorectal adenoma (Table 6, Figure 4C), suggest the prediction model distinguished between these groups well. The values of accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) in the prediction model are also shown in Table 6. The calibration curves in Figure 4D-4F showed that the prediction model was well-calibrated. Moreover, the results in Table 7 of patients grouped into CRC and colorectal adenoma showed that our prediction model has higher AUC and PPV in patients with [0.879 (95% CI: 0.760-0.997)] of AUC, 0.906 (95% CI: 0.805-0.999) of PPV or without [0.835 (95% CI: 0.741-0.928)] of AUC, 0.773 (95% CI: 0.649-0.897) of PPV inflammatory diseases. The prediction models we developed showed good performance in patients divided into CRC stage I, II, or III, and in colorectal adenomas and patients divided into poorly differentiated CRC, moderately to well-differentiated CRC, and colorectal adenomas.
Table 1. Demographic and clinicopathological characteristics of selected patients.

| Characteristic                              | Total (n=856) | Groups                     | Statistics | P    |
|---------------------------------------------|---------------|----------------------------|------------|------|
|                                             |               | Training set (n=600)       | Testing set (n=256) |      |
| Gender, n (%)                               |               |                           |            |      |
| Male                                        | 511 (59.70)   | 356 (59.33)                | 155 (60.55) |      |
| Female                                      | 345 (40.30)   | 244 (40.67)                | 101 (39.45) |      |
| Age, years, Mean±SD                         | 56.39±12.15   | 56.16±12.43                | 56.94±11.46 |      |
| BMI, kg/m², Mean±SD                         | 22.45±3.38    | 22.40±3.42                 | 22.56±3.30 |      |
| Diabetes, n (%)                             |               |                           |            |      |
| No                                          | 818 (95.56)   | 575 (95.83)                | 243 (94.92) |      |
| Yes                                         | 38 (4.44)     | 25 (4.17)                  | 13 (5.08)  |      |
| Inflammatory disease, n (%)                 |               |                           |            |      |
| No                                          | 706 (82.48)   | 498 (83.00)                | 208 (81.25) |      |
| Yes                                         | 150 (17.52)   | 102 (17.00)                | 48 (18.75) |      |
| Hb, g/L, Mean±SD                            | 125.14±24.58  | 125.48±24.67               | 124.35±24.38 | t=0.620 | 0.538 |
| RBC, 10¹¹/L, Mean±SD                        | 4.21±0.65     | 4.24±0.63                  | 4.14±0.68 | t=1.520 | 0.094 |
| WBC, 10⁹/L, M (Q₁, Q₃)                      | 5.87 (4.77, 7.34) | 5.91 (4.82, 7.22)         | 5.78 (4.66, 7.63) | Z=-0.011 | 0.992 |
| NEUT, 10⁹/L, M (Q₁, Q₃)                     | 3.67 (1.07, 1.80) | 3.69 (1.09, 1.79)         | 3.66 (1.06, 1.80) | Z=-0.742 | 0.458 |
| MONO, 10⁹/L, M (Q₁, Q₃)                     | 0.35 (0.26, 0.46) | 0.35 (0.26, 0.47)         | 0.35 (0.25, 0.46) | Z=-0.411 | 0.681 |
| PLT, 10⁹/L, M (Q₁, Q₃)                      | 227.00 (181.00, 281.00) | 229.00 (181.00, 284.00) | 221.00 (180.50, 272.50) | Z=-0.687 | 0.492 |
| MPV, fl, Mean±SD                            | 10.14±1.49    | 10.20±1.49                 | 9.98±1.49 | t=1.950 | 0.051 |
| PCT, %, M (Q₁, Q₃)                          | 0.23 (0.19, 0.27) | 0.23 (0.19, 0.28)         | 0.22 (0.19, 0.27) | t=1.600 | 0.101 |
| RDW, %, Mean±SD                             | 13.51±2.29    | 13.59±2.52                 | 13.32±2.04 | t=0.040 | 0.686 |
| PT, s, Mean±SD                              | 11.25±1.01    | 11.25±0.98                 | 11.28±1.08 | t=0.730 | 0.463 |
| APTT, s, Mean±SD                            | 26.76±4.17    | 26.83±4.18                 | 26.60±4.12 | t=0.040 | 0.686 |
| TT, s, Mean±SD                              | 18.31±1.79    | 18.27±1.72                 | 18.41±1.95 | t=0.103 | 0.301 |
| FIB, g/L, Mean±SD                           | 3.07±0.84     | 3.06±0.84                  | 3.10±0.84 | t=0.710 | 0.479 |
| D-dimer, mg/L, M (Q₁, Q₃)                   | 0.30 (0.17, 0.60) | 0.31 (0.17, 0.62)         | 0.30 (0.17, 0.56) | Z=-0.548 | 0.584 |
| CEA, ng/mL, M (Q₁, Q₃)                      | 1.86 (0.90, 3.51) | 1.82 (0.90, 3.80)         | 1.93 (0.88, 3.26) | Z=-0.688 | 0.492 |
| Carbohydrate antigen 199, U/mL, M (Q₁, Q₃)  | 10.94 (6.80, 19.42) | 10.84 (6.89, 20.45) | 11.03 (6.61, 17.54) | Z=-0.109 | 0.917 |
| Carbohydrate antigen 125, U/mL, M (Q₁, Q₃)  | 8.50 (5.90, 12.65) | 8.60 (6.30, 12.50) | 8.40 (5.60, 13.20) | Z=-0.958 | 0.338 |
| GPT, U/L, M (Q₁, Q₃)                        | 16.00 (11.30, 24.00) | 16.00 (11.00, 24.00) | 16.00 (12.00, 25.10) | Z=0.587 | 0.557 |
| GOT, U/L, M (Q₁, Q₃)                        | 21.00 (17.00, 26.00) | 21.00 (17.00, 26.00) | 21.00 (17.00, 27.00) | Z=-0.230 | 0.818 |
| TBil, μmol/L, M (Q₁, Q₃)                    | 13.00 (9.25, 17.80) | 13.00 (9.30, 17.80) | 12.90 (9.20, 17.90) | Z=0.650 | 0.518 |
| ALB+GLB, g/L, Mean±SD                       | 66.61±7.48    | 66.72±7.47                 | 66.36±7.53 | t=0.650 | 0.518 |
### Table 1 continued. Demographic and clinicopathological characteristics of selected patients.

| Characteristic | Total (n=856) | **Groups** | **Statistics** | **P** |
|----------------|--------------|------------|----------------|------|
|                |              | **Training set** | **Testing set** |       |
|                |              | (n=600) | (n=256) |       |
| ALB, g/L, Mean±SD | 41.63±4.90 | 41.70±4.78 | 41.45±5.18 | t=0.690 | 0.493 |
|                | 24.99±4.56 |            |              |       |
| A/G, Mean±SD   | 1.71±0.33 | 1.71±0.32 | 1.71±0.33 | t=0.220 | 0.829 |
|                | 24.99±4.56 |            |              |       |
| BUN, mmol/L, M (Q₁, Q₃) | 4.80 (3.80, 5.80) | 4.80 (3.75, 5.80) | 4.80 (3.90, 5.80) | Z=0.637 | 0.524 |
| Cr, μmol/L, M (Q₁, Q₃) | 71.00 (61.00, 84.00) | 71.00 (61.00, 84.00) | 73.00 (61.00, 84.50) | Z=0.547 | 0.584 |
| UA, μmol/L, M (Q₁, Q₃) | 288.00 (221.00, 359.00) | 286.00 (220.00, 360.00) | 291.00 (225.50, 355.00) | Z=0.561 | 0.575 |
| MLR, M (Q₁, Q₃) | 0.24 (0.18, 0.35) | 0.24 (0.18, 0.35) | 0.24 (0.17, 0.36) | Z=0.437 | 0.662 |
| MPV/PLT, M (Q₁, Q₃) | 0.04 (0.03, 0.06) | 0.04 (0.03, 0.06) | 0.04 (0.03, 0.06) | Z=0.561 | 0.575 |
| FAR, M (Q₁, Q₃) | 0.07 (0.06, 0.09) | 0.07 (0.06, 0.09) | 0.07 (0.06, 0.09) | Z=0.437 | 0.662 |
| Pathological diagnosis, n (%) | 416.26±49.03 | 417.01±47.83 | 414.50±51.76 | t=0.69 | 0.493 |
| CRC | 531 (62.03) | 371 (61.83) | 160 (62.50) |       |
| Poorly differentiated | 325 (37.97) | 229 (38.17) | 96 (37.50) |       |
| Moderately differentiated | 249 (29.09) | 173 (28.83) | 76 (29.69) |       |
| Well differentiated | 214 (25.00) | 150 (25.00) | 64 (25.00) |       |

SD – standard deviation; M (Q₁, Q₃) – median (interquartile range); BMI – body mass index; Hb – hemoglobin; RBC – red blood cell; WBC – white blood cell; NEUT – neutrophil; LYM – lymphocyte; MONO – monocyte; PLT – platelet; MPV – mean platelet volume; PCT – plateletcrit (the product of PLT and MPV); RDW – red blood cell distribution width; PT – prothrombin time; APTT – activated partial thromboplastin time; TT – thrombin time; FIB – fibrinogen; CEA – carcinoembryonic antigen; GPT – glutamic pyruvic transaminase; GOT – glutamic oxaloacetic transaminase; TBil – total bilirubin; ALB – albumin; GLB – globulin; A/G = ALB-to-GLB ratio; BUN – blood urea nitrogen; Cr – creatinine; UA – uric acid; NLR – NEUT-to-LYM ratio; PLR – PLT-to-LYM ratio; MLR – MONO-to-LYM ratio; FAR – FIB-to-ALB ratio; PNI – 10×ALB+0.005×LYM; SII – systemic immune-inflammatory index; dNLR – NEUT/(WBC-NEUT); HPR – Hb-to-PLT ratio; CRC – colorectal cancer.
Table 2. Comparison of characteristics between colorectal adenoma and CRC in the training set.

| Characteristic                  | Total (n=600) | Groups | Statistics | P  |
|---------------------------------|--------------|--------|------------|----|
|                                 |              | Colorectal adenoma (n=229) | CRC (n=371) |    |
| Gender, n (%)                   |              |        |            |    |
| Male                            | 356 (59.33)  | 142 (62.01) | 214 (57.68) | $\chi^2=1.099$ | 0.295 |
| Female                          | 244 (40.67)  | 87 (37.99)  | 157 (42.32)  |    |
| Age, years, Mean±SD             | 56.16±12.43  | 56.78±11.70 | 55.78±12.86 | t=0.960 | 0.338 |
| BMI, kg/m², Mean±SD             | 22.40±3.42   | 22.93±3.71  | 22.08±3.19  | t=2.900 | 0.004 |
| Diabetes, n (%)                 |              |        |            |    |
| No                              | 575 (95.83)  | 219 (95.63) | 356 (95.96) | $\chi^2=0.037$ | 0.847 |
| Yes                             | 25 (4.17)    | 10 (4.37)   | 15 (4.04)   |    |
| Inflammatory disease, n (%)     |              |        |            |    |
| No                              | 498 (83.00)  | 195 (85.15) | 303 (81.67) |    |
| Yes                             | 102 (17.00)  | 34 (14.85)  | 68 (18.33)  |    |
| Hb, g/L, Mean±SD                | 125.48±24.67 | 137.79±19.88 | 117.88±24.30 | t=10.930 | <0.001 |
| RBC, 10^{12}/L, Mean±SD         | 4.24±0.63    | 4.49±0.55  | 4.09±0.63  | t=8.150 | <0.001 |
| WBC, 10^{9}/L, M (Q1, Q3)       | 5.91 (4.82, 7.22) | 5.84 (4.81, 6.68) | 6.04 (4.82, 7.67) | Z=-2.076 | 0.038 |
| NEUT, 10^{9}/L, M (Q1, Q3)      | 3.69 (2.89, 5.01) | 3.54 (2.82, 4.50) | 3.92 (2.92, 5.56) | Z=-2.879 | 0.004 |
| LYM, 10^{9}/L, M (Q1, Q3)       | 1.43 (1.09, 1.79) | 1.54 (1.24, 1.87) | 1.33 (1.03, 1.72) | Z=4.050 | <0.001 |
| MONO, 10^{9}/L, M (Q1, Q3)      | 0.35 (0.26, 0.47) | 0.34 (0.27, 0.44) | 0.35 (0.25, 0.49) | Z=-1.121 | 0.262 |
| PLT, 10^{9}/L, M (Q1, Q3)       | 229.00 (181.00, 284.00) | 212.00 (173.00, 253.00) | 243.00 (187.00, 305.00) | Z=-4.912 | <0.001 |
| MPV, fL, Mean±SD                | 10.20±1.49   | 10.55±1.44 | 9.98±1.48  | t=6.450 | <0.001 |
| PCT, %, M (Q1, Q3)              | 0.23 (0.19, 0.28) | 0.22 (0.18, 0.26) | 0.24 (0.20, 0.29) | Z=-3.786 | <0.001 |
| RDW, %, Mean±SD                 | 13.59±2.52   | 13.12±1.61 | 13.87±2.91 | t=-4.080 | <0.001 |
| PT, s, Mean±SD                  | 11.25±0.98   | 10.80±0.90 | 11.52±0.93 | t=-9.220 | <0.001 |
| TT, s, Mean±SD                  | 26.83±4.18   | 26.28±3.40 | 27.16±4.57 | t=-2.690 | 0.007 |
| FIB, g/L, Mean±SD               | 3.06±0.84    | 2.79±0.71  | 3.23±0.87  | t=6.840 | <0.001 |
| GPT, U/L, M (Q1, Q3)            | 16.00 (11.00, 24.00) | 18.90 (14.00, 26.00) | 14.00 (10.00, 22.00) | Z=5.241 | <0.001 |
| GOT, U/L, M (Q1, Q3)            | 8.60 (6.30, 12.50) | 8.60 (6.20, 11.80) | 8.50 (6.30, 13.00) | Z=-0.961 | 0.336 |
| Carbohydrate antigen 199, U/mL, M (Q1, Q3) | 10.84 (6.89, 20.45) | 10.34 (6.50, 17.06) | 11.30 (6.99, 22.43) | Z=-2.280 | 0.023 |
| Carbohydrate antigen 125, U/mL, M (Q1, Q3) | 8.60 (6.30, 12.50) | 8.60 (6.20, 11.80) | 8.50 (6.30, 13.00) | Z=-0.961 | 0.336 |
| GPT, U/L, M (Q1, Q3)            | 16.00 (11.00, 24.00) | 18.90 (14.00, 26.00) | 14.00 (10.00, 22.00) | Z=5.241 | <0.001 |
| GOT, U/L, M (Q1, Q3)            | 21.00 (17.00, 26.00) | 23.00 (19.00, 27.00) | 20.00 (16.00, 25.00) | Z=4.688 | <0.001 |
Table 2 continued. Comparison of characteristics between colorectal adenoma and CRC in the training set.

| Characteristic                  | Total (n=600) | Colorectal adenoma (n=229) | CRC (n=371) | Statistics | P |
|--------------------------------|---------------|----------------------------|-------------|------------|---|
| Tbil, μmol/L, M (Q₁, Q₃)       | 13.00 (9.30, 17.80) | 15.30 (11.40, 19.70)       | 11.70 (8.40, 16.10) | Z=6.389 | <0.001 |
| ALB+GLB, g/L, Mean±SD          | 66.72±7.47    | 69.57±6.86                 | 64.96±7.29  | t=7.700   | <0.001 |
| ALB, g/L, Mean±SD              | 41.70±4.78    | 42.38±4.21                 | 41.28±5.07  | t=2.880   | 0.004  |
| GLB, g/L, Mean±SD              | 25.02±4.60    | 27.19±4.80                 | 23.68±3.92  | t=9.310   | <0.001 |
| A/G, Mean±SD                   | 1.71±0.32     | 1.60±0.31                  | 1.78±0.31   | t=-6.740  | <0.001 |
| BUN, mmol/L, M (Q₁, Q₃)        | 4.80 (3.75, 5.80) | 5.20 (4.30, 6.20)          | 4.50 (3.30, 5.60) | Z=6.065 | <0.001 |
| Cr, μmol/L, M (Q₁, Q₃)         | 71.00 (61.00, 84.00) | 73.00 (62.00, 86.00)       | 70.00 (60.00, 84.00) | Z=1.660 | 0.097  |
| UA, μmol/L, M (Q₁, Q₃)         | 286.00 (220.00, 360.00) | 325.00 (258.00, 391.00)   | 260.00 (207.00, 333.00) | Z=6.864 | <0.001 |
| NLR, M (Q₁, Q₃)                | 2.55 (1.81, 3.95) | 2.22 (1.72, 3.08)          | 2.85 (1.93, 4.50) | Z=4.693 | <0.001 |
| PLR, M (Q₁, Q₃)                | 156.44 (116.62, 229.85) | 136.30 (106.14, 174.32)   | 176.36 (126.70, 264.58) | Z=6.442 | <0.001 |
| MLR, M (Q₁, Q₃)                | 0.24 (0.18, 0.35) | 0.21 (0.18, 0.30)          | 0.26 (0.18, 0.38) | Z=3.132 | 0.002  |
| MPV/PLT, M (Q₁, Q₃)            | 0.04 (0.03, 0.06) | 0.05 (0.04, 0.06)         | 0.04 (0.03, 0.06) | Z=5.243 | <0.001 |
| FAR, M (Q₁, Q₃)                | 0.07 (0.06, 0.09) | 0.06 (0.05, 0.07)         | 0.07 (0.06, 0.09) | Z=7.222 | <0.001 |
| PNI, Mean±SD                   | 417.01±47.83  | 423.84±42.09               | 412.80±50.66 | t=2.880 | 0.004  |
| SII, M (Q₁, Q₃)                | 581.29 (377.24, 1034.25) | 472.22 (339.95, 698.73)   | 675.76 (417.75, 1325.62) | Z=5.985 | <0.001 |
| dNLR, M (Q₁, Q₃)               | 1.84 (1.38, 2.61) | 1.73 (1.35, 2.15)          | 1.99 (1.42, 2.96) | Z=4.042 | <0.001 |
| HPR, M (Q₁, Q₃)                | 0.57 (0.42, 0.75) | 0.67 (0.53, 0.81)         | 0.50 (0.34, 0.70) | Z=7.778 | <0.001 |

SD – standard deviation; M (Q₁, Q₃) – median (interquartile range); BMI – body mass index; Hb – hemoglobin; RBC – red blood cell; WBC – white blood cell; NEUT – neutrophil; LYM – lymphocyte; MONO – monocyte; PLT – platelet; MPV – mean platelet volume; PCT – plateletcrit (the product of PLT and MPV); RDW – red blood cell distribution width; PT – prothrombin time; APTT – activated partial thromboplastin time; TT – thrombin time; FIB – fibrinogen; CEA – carcinoembryonic antigen; GPT – glutamic pyruvic transaminase; GOT – glutamic oxaloacetic transaminase; TBil – total bilirubin; ALB – albumin; GLB – globulin; A/G – ALB-to-GLB ratio; BUN – blood urea nitrogen; Cr – creatinine; UA – uric acid; NLR – NEUT-to-LYM ratio; PLR – PLT-to-LYM ratio; MLR – MONO-to-LYM ratio; FAR – FIB-to-ALB ratio; PNI – 10×ALB+0.005×LYM; SII – systemic immune-inflammatory index; dNLR – NEUT/(WBC-NEUT); HPR – Hb-to-PLT ratio; CRC – colorectal cancer.

Discussion

In the present study, we found HPR, FAR, A/G, PLR, CEA, PT, TT, and plateletcrit to be diagnostic indicators of CRC, showing the ability to differentiate colorectal adenoma from CRC. A prediction model for diagnosing CRC using HPR, FAR, A/G, PLR, CEA, and TT was developed, and a nomogram was generated for the patients grouped into CRC. The other 2 prediction models for diagnosing CRC were developed (1) using PT, TT, CEA, A/G, FAR, and HPR in the patients grouped into CRC stages I, II, and III, and colorectal adenoma; and (2) using PT, TT, A/G, plateletcrit, FAR, and HPR in the patients grouped into poorly differentiated CRC, moderately and well-differentiated CRC, and colorectal adenoma.

In the identification of CRC and colorectal adenoma patients, Roman et al [24] created a prediction model and nomogram using Hb, FIB, ALB, and NLR to predict the risk of CRC in patients, with an AUC value of 0.744. Another study showed that the diagnostic value of plateletcrit combined with CEA
Table 3. Diagnostic indicators for CRC patients using the multivariate logistic regression model in participants grouped into CRC and colorectal adenoma.

| Variables | β   | SE  | Wald  | P      | OR   | 95% CI Lower | 95% CI Upper |
|-----------|-----|-----|-------|--------|------|--------------|--------------|
| Constant  | -16.884 | 1.875 | 81.077 | <0.001 | 1.000 | 1.000        | 1.000        |
| HPR       | -1.725 | 0.513 | 11.313 | <0.001 | 0.178 | 0.065        | 0.487        |
| FAR       | 0.463  | 0.063 | 54.558 | <0.001 | 1.590 | 1.406        | 1.798        |
| A/G       | 0.343  | 0.041 | 70.110 | <0.001 | 1.310 | 1.000        | 1.528        |
| PLR       | 0.004  | 0.001 | 6.661  | <0.001 | 1.000 | 1.000        | 1.000        |
| CEA       | 0.062  | 0.024 | 6.808  | <0.001 | 1.000 | 1.000        | 1.000        |
| TT        | 0.460  | 0.076 | 36.588 | <0.001 | 1.585 | 1.365        | 1.840        |

OR – odds ratio; CI – confidence interval; SE – standard error; Constant – intercept; β – beta coefficient; Wald – Wald Chi-square; HPR – hemoglobin-platelet ratio; FAR – fibrinogen-albumin ratio; A/G – albumin-globulin ratio; PLR – platelet-lymphocyte ratio; CEA – carcinoembryonic antigen; TT – thrombin time; CRC – colorectal cancer.

Table 4. Diagnostic indicators for CRC patients using the multivariate logistic regression model in participants grouped into stage I, II, and III CRC and colorectal adenoma.

| Variables | Model                                      | β   | SE  | Wald  | P      | OR   | 95% CI Lower | 95% CI Upper |
|-----------|--------------------------------------------|-----|-----|-------|--------|------|--------------|--------------|
| PT        | Colorectal adenoma & stage III CRC         | 0.758 | 0.152 | 24.820 | <0.001 | 2.133 | 1.583        | 2.874        |
| PT        | Colorectal adenoma & stage II CRC          | 0.702 | 0.149 | 22.157 | <0.001 | 2.019 | 1.507        | 2.704        |
| PT        | Colorectal adenoma & stage I CRC           | 0.617 | 0.198 | 9.734  | 0.002  | 1.853 | 1.258        | 2.731        |
| TT        | Colorectal adenoma & stage III CRC         | 0.433 | 0.063 | 35.558 | <0.001 | 1.542 | 1.409        | 1.661        |
| TT        | Colorectal adenoma & stage II CRC          | 0.482 | 0.071 | 46.046 | <0.001 | 1.619 | 1.371        | 1.900        |
| TT        | Colorectal adenoma & stage I CRC           | 0.199 | 0.102 | 3.782  | <0.05  | 1.220 | 0.998        | 1.490        |
| CEA       | Colorectal adenoma & stage III CRC         | 0.331 | 0.049 | 46.078 | <0.001 | 1.392 | 1.265        | 1.531        |
| CEA       | Colorectal adenoma & stage II CRC          | 0.327 | 0.047 | 47.848 | <0.001 | 1.386 | 1.264        | 1.521        |
| CEA       | Colorectal adenoma & stage I CRC           | 0.303 | 0.060 | 25.932 | <0.001 | 1.354 | 1.205        | 1.522        |
| A/G       | Colorectal adenoma & stage III CRC         | 0.433 | 0.073 | 35.557 | <0.001 | 1.542 | 1.338        | 1.779        |
| A/G       | Colorectal adenoma & stage II CRC          | 0.482 | 0.071 | 46.046 | <0.001 | 1.619 | 1.371        | 1.900        |
| A/G       | Colorectal adenoma & stage I CRC           | 0.199 | 0.102 | 3.782  | <0.05  | 1.220 | 0.998        | 1.490        |
| FAR       | Colorectal adenoma & stage III CRC         | 0.433 | 0.073 | 35.557 | <0.001 | 1.542 | 1.338        | 1.779        |
| FAR       | Colorectal adenoma & stage II CRC          | 0.482 | 0.071 | 46.046 | <0.001 | 1.619 | 1.371        | 1.900        |
| FAR       | Colorectal adenoma & stage I CRC           | 0.199 | 0.102 | 3.782  | <0.05  | 1.220 | 0.998        | 1.490        |
| HPR       | Colorectal adenoma & stage III CRC         | -2.337 | 0.531 | 19.344 | <0.001 | 0.097 | 0.034        | 0.274        |
| HPR       | Colorectal adenoma & stage II CRC          | -2.830 | 0.529 | 28.567 | <0.001 | 0.059 | 0.021        | 0.167        |
| HPR       | Colorectal adenoma & stage I CRC           | -0.774 | 0.633 | 1.497  | 0.221  | 0.461 | 0.133        | 1.594        |

OR – odds ratio; CI – confidence interval; SE – standard error; β – beta coefficient; Wald – Wald Chi-square; PT – prothrombin time; TT – thrombin time; CEA – carcinoembryonic antigen; A/G – albumin-globulin ratio; FAR – fibrinogen-albumin ratio; HPR – hemoglobin-platelet ratio; CRC – colorectal cancer.
in the diagnosis of CRC was 0.835 (AUC), and the sensitivity was 0.700, while the specificity was 0.830 [22]. The AUC values of the prediction models we proposed in this study using the diagnostic indicators were higher than those of prediction models that are used to diagnose CRC through hematological parameters; in addition, we also provided the reference PPV value. Clinically, the AUC may be different in the case of application of the nomogram to different cohorts, so a higher AUC may ensure the reliability of the prediction [25]. In addition, we did not exclude patients with inflammatory diseases. In our study population, patients with inflammatory diseases accounted for 17% of the total population, indicating that patients with inflammation diseases accounted for a large proportion of the CRC and colorectal adenoma population. Therefore, our CRC diagnosis prediction model used a wider population range than the prediction model that excluded patients with inflammation. A retrospective study in China stated that MPV-to-platelets ratio was used as an effective diagnostic index to distinguish benign from malignant CRC and distinguish early CRC from advanced CRC [19]. Hui et al [21] found that CEA, Hb, lipoprotein (a), and high-density lipoprotein might be powerful and noninvasive diagnostic indicators based on machine-learning approaches. Consistently, we have studied the diagnosis of CRC by hematological indicators, and there is also agreement on diagnostic indicators such as platelet-related parameters and CEA. These 2 studies considered the diagnosis and identification of early and advanced CRC. In addition to CRC stage, we also considered the grade of CRC, and established a related prediction model in participants grouped into CRC that was poorly differentiated, CRC that was moderately and well differentiated, or colorectal adenoma.

There are many diagnostic methods for CRC, including colonoscopy, colorectal barium enema, nuclear magnetic resonance, intracavitary ultrasound, blood tumor markers, fecal occult blood test (FOBT), genetic examination, and others [26]. Among these,

### Table 5. Diagnostic indicators for CRC patients using the multivariate logistic regression model in participants grouped into poorly differentiated CRC, moderately and well-differentiated CRC, and colorectal adenoma.

| Variables | Model | β     | SE    | Wald  | P     | OR    | 95% CI          |
|-----------|-------|-------|-------|-------|-------|-------|-----------------|
| PT        | Colorectal adenoma & Poorly differentiated CRC | 0.6731 | 0.1365 | 24.3288 | <0.001 | 1.960 | 1.500 - 2.561 |
|           | Colorectal adenoma & Moderately/well-differentiated CRC | 0.8786 | 0.1803 | 23.7555 | <0.001 | 2.408 | 1.691 - 3.428 |
| TT        | Colorectal adenoma & Poorly differentiated CRC | 0.4253 | 0.0748 | 32.3168 | <0.001 | 1.530 | 1.321 - 1.772 |
|           | Colorectal adenoma & Moderately/well-differentiated CRC | 0.5524 | 0.0980 | 31.7703 | <0.001 | 1.737 | 1.434 - 2.105 |
| A/G       | Colorectal adenoma & Poorly differentiated CRC | 0.3245 | 0.0429 | 57.3282 | <0.001 | 1.383 | 1.272 - 1.504 |
|           | Colorectal adenoma & Moderately/well-differentiated CRC | 0.3777 | 0.0602 | 39.3273 | <0.001 | 1.459 | 1.296 - 1.642 |
| PCT       | Colorectal adenoma & Poorly differentiated CRC | 0.2225 | 0.0917 | 5.8865 | 0.015 | 1.249 | 1.044 - 1.495 |
|           | Colorectal adenoma & Moderately/well-differentiated CRC | 0.2493 | 0.1242 | 4.0306 | 0.045 | 1.283 | 1.006 - 1.637 |
| FAR       | Colorectal adenoma & Poorly differentiated CRC | 0.4039 | 0.0642 | 39.5489 | <0.001 | 1.498 | 1.321 - 1.699 |
|           | Colorectal adenoma & Moderately/well-differentiated CRC | 0.5100 | 0.0820 | 38.8682 | <0.001 | 1.665 | 1.418 - 1.956 |
| HPR       | Colorectal adenoma & Poorly differentiated CRC | -4.2277 | 0.9121 | 21.4845 | <0.001 | 0.015 | 0.002 - 0.087 |
|           | Colorectal adenoma & Moderately/well-differentiated CRC | -3.9132 | 1.2829 | 9.3042 | 0.002 | 0.020 | 0.002 - 0.247 |

OR – odds ratio; CI – confidence interval; SE – standard error; β – beta coefficient; Wald – Wald Chi-square; PT – prothrombin time; TT – thrombin time; A/G – albumin-globulin ratio; PCT – plateletcrit (the product of platelet and mean platelet volume); FAR – fibrinogen-albumin ratio; HPR – hemoglobin-platelet ratio; CRC – colorectal cancer.
Figure 2. Nomogram for differentiating colorectal carcinoma (CRC) from colorectal adenoma in the patients grouped into CRC and colorectal adenoma [R v. 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria)]. PLR – platelet-to-lymphocyte ratio; HPR – hemoglobin-to-platelet ratio; CEA – carcinoembryonic antigen; TT – thrombin time; A/G – albumin-to-globulin ratio; FAR – fibrinogen-to-albumin ratio.

Figure 3. An example of the nomogram for differentiating colorectal carcinoma (CRC) from colorectal adenoma in the patients grouped into CRC and colorectal adenoma [R v. 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria)]. PLR – platelet-to-lymphocyte ratio; HPR – hemoglobin-to-platelet ratio; CEA – carcinoembryonic antigen; TT – thrombin time; A/G – albumin-to-globulin ratio; FAR – fibrinogen-to-albumin ratio.
colonscopy is currently the most reliable method for the diagnosis of CRC, but invasiveness, poor compliance, and high cost has limited its application in large-scale screening [26,27]. Computerized tomography (CT), magnetic resonance imaging (MRI), and other imaging examinations do not easily detect small lesions, which limits the early diagnosis of CRC by these methods [28]. FOBT had the advantages of simplicity, non-invasiveness, and economy, but it is prone to false-positive results [29]. It therefore requires multiple and repeated tests to increase its detection rate. In contrast, the method of CRC detection based on inflammation markers, blood biomarkers, and CEA as combined indicators developed in the present study has the advantages of low cost, rapid detection, and acceptability [12,30].

Inflammation plays an important role in the proliferation and survival of malignant cells by changing the homeostasis of tumor tissues and participating in the construction of the tumor microenvironment [31]. Platelets have gradually become involved in the complex interaction between inflammation and cancer [18]. PLR as a typical inflammatory indicator has been previously related to CRC [32]. In a retrospective study evaluating the prognosis of 201 CRC patients treated with curative surgery, overall survival and relapse-free survival were significantly higher in CRC patients in the low-PLR group than in those in the high-PLR group [33]. Also, CRC patients have exhibited significantly higher PLR values than colorectal adenoma patients and healthy controls [34], a finding that was consistent with our results. There have been few predictive models for HPR and FAR to diagnose CRC, but HPR and FAR are related to platelets and related inflammatory factors. In our study, we found that CRC patients have lower HPR and higher FAR levels, compared with colorectal adenoma patients. A/G is a common inflammation-related indicator that has been used previously as a prognostic indicator for CRC patients [35]. The NLR was derived from NEU and LYM of the complete blood count. It has become one of the commonly used inflammation markers due to its cost effectiveness [36]. Several studies have showed NLR to be useful for the diagnosis and prognosis of breast cancer [37], cervical cancer [38,39], non-small cell lung cancer [40], CRC [15] and other malignant tumors. Although we found differences in NLR in the comparison between CRC and colorectal adenoma, no differences were found in the multivariate logistic regression.

CEA was initially considered to be CRC-specific, but later, CEA levels were also detected in other tumors, such as gastric cancer and pancreatic cancer [41]. In addition, CEA could not distinguish between benign and malignant polyps, and CEA is not currently recommended for screening tests according to the guidelines of the European Group on Tumor Markers, European Society of Medical Oncology, and American Society of Clinical Oncology [42,43]. Therefore, we used CEA as one of the combined diagnostic indicators for CRC diagnosis in our study. Similarly, the use of CEA combined with other hematological indicators for CRC diagnosis has also been reported in other studies [15,23]. However, a few limitations were present in our study. First, this study could only provide a reference for early screening of CRC; we did not conduct regular followup of the eligible patients, so we were unable to obtain prognostic status after diagnostic analysis. Second, we only studied several hematological parameters (HPR, FAR, A/G, PLR, CEA, PT, TT, and plateletcrit) for the diagnostic value of CRC and colorectal adenoma, providing a reference for the early detection of CRC. However, other indicators of potential research value, such as genes [44],...
Conclusions

In this study, we developed 3 prediction models for the diagnosis of CRC, stage of CRC, and grade of CRC that used several readily available hematological parameters (HPR, FAR, A/G, PLR, CEA, PT, TT, and plateletcrit) as diagnostic indicators. The prediction models were validated with good discrimination and calibration. Our research provided a reference for the development of auxiliary tools for the early screening of CRC.

Declaration of Figures Authenticity

All figures submitted have been created by the authors who confirm that the images are original with no duplication and have not been previously published in whole or in part.
Table 6. The performance of the prediction model in training and testing sets.

|                          | Patients grouped into CRC and colorectal adenoma | Patients grouped into stage I, II, and III CRC and colorectal adenoma | Patients grouped into poorly differentiated CRC, moderately and well-differentiated CRC, and colorectal adenoma |
|--------------------------|--------------------------------------------------|----------------------------------------------------------|--------------------------------------------------------------------------------------------------|
|                          | Training set  | Testing set | Training set  | Testing set | Training set  | Testing set |
| AUC (95% CI)             | 0.864 (0.834-0.895) | 0.848 (0.800-0.896) | 0.818 (0.799-0.838) | 0.806 (0.775-0.836) | 0.857 (0.838-0.876) | 0.829 (0.797-0.860) |
| Accuracy (95% CI)        | 0.808 (0.777-0.840) | 0.777 (0.726-0.828) | 0.746 (0.729-0.764) | 0.733 (0.706-0.760) | 0.822 (0.805-0.840) | 0.788 (0.759-0.817) |
| Sensitivity (95% CI)     | 0.852 (0.816-0.888) | 0.812 (0.752-0.873) | 0.745 (0.710-0.780) | 0.727 (0.672-0.781) | 0.728 (0.706-0.746) | 0.715 (0.660-0.770) |
| Specificity (95% CI)     | 0.738 (0.681-0.795) | 0.719 (0.629-0.809) | 0.747 (0.727-0.767) | 0.736 (0.704-0.767) | 0.853 (0.832-0.873) | 0.824 (0.791-0.857) |
| PPV (95% CI)             | 0.840 (0.803-0.877) | 0.828 (0.769-0.887) | 0.495 (0.462-0.528) | 0.478 (0.429-0.528) | 0.721 (0.686-0.756) | 0.670 (0.615-0.726) |
| NPV (95% CI)             | 0.754 (0.698-0.811) | 0.697 (0.606-0.787) | 0.898 (0.882-0.913) | 0.890 (0.865-0.914) | 0.877 (0.859-0.896) | 0.853 (0.821-0.884) |

AUC – area under the curve; CI – confidence interval; PPV – positive predictive value; NPV – negative predictive value; CRC – colorectal cancer.

Table 7. Performance of the prediction model in patients with or without inflammatory disease in the testing set.

|                          | Patients grouped into CRC and colorectal adenoma | Patients grouped into stage I, II, and III CRC and colorectal adenoma | Patients grouped into poorly differentiated CRC, moderately and well-differentiated CRC, and colorectal adenoma |
|--------------------------|--------------------------------------------------|----------------------------------------------------------|--------------------------------------------------------------------------------------------------|
|                          | With inflammatory disease                         | Without inflammatory disease                             | With inflammatory disease | Without inflammatory disease | With inflammatory disease | Without inflammatory disease |
| AUC (95% CI)             | 0.879 (0.760-0.997) | 0.835 (0.741-0.928) | 0.860 (0.799-0.921) | 0.814 (0.781-0.846) | 0.851 (0.776-0.926) | 0.827 (0.792-0.861) |
| Accuracy (95% CI)        | 0.854 (0.722-0.939) | 0.764 (0.649-0.856) | 0.740 (0.678-0.802) | 0.720 (0.689-0.750) | 0.833 (0.772-0.894) | 0.790 (0.758-0.822) |
| Sensitivity (95% CI)     | 0.879 (0.767-0.990) | 0.829 (0.714-0.944) | 0.781 (0.718-0.869) | 0.774 (0.717-0.831) | 0.603 (0.603-0.855) | 0.640 (0.603-0.764) |
| Specificity (95% CI)     | 0.800 (0.598-1.002) | 0.677 (0.513-0.842) | 0.694 (0.619-0.770) | 0.702 (0.666-0.738) | 0.885 (0.822-0.949) | 0.834 (0.798-0.870) |
| PPV (95% CI)             | 0.906 (0.805-0.999) | 0.773 (0.649-0.897) | 0.488 (0.383-0.594) | 0.464 (0.412-0.516) | 0.761 (0.638-0.884) | 0.679 (0.617-0.741) |
| NPV (95% CI)             | 0.750 (0.538-0.962) | 0.750 (0.590-0.910) | 0.943 (0.899-0.987) | 0.903 (0.877-0.929) | 0.867 (0.800-0.935) | 0.848 (0.814-0.883) |

AUC – area under the curve; CI – confidence interval; PPV – positive predictive value; NPV – negative predictive value; CRC – colorectal cancer.
### Supplementary Table 1. Proportion of missing variables and sensitivity analysis.

| Characteristic | Data missing ratio (%) | After interpolation | Before interpolation | Statistics | P |
|----------------|------------------------|---------------------|----------------------|------------|---|
| Hb, g/L, Mean±SD | 0.12 | 125.14±24.58 | 125.15±24.59 | t=0.010 | 0.996 |
| RBC, 10¹²/L, Mean±SD | 0.12 | 4.21±0.65 | 4.21±0.65 | t=0.030 | 0.975 |
| WBC, 10⁹/L, M (Q₁, Q₃) | 0.12 | 5.87 (4.77, 7.34) | 5.87 (4.77, 7.34) | Z=0.034 | 0.973 |
| NEUT, 10⁹/L, M (Q₁, Q₃) | 0.12 | 3.67 (2.83, 5.15) | 3.67 (2.83, 5.15) | Z=0.034 | 0.973 |
| LYM, 10⁹/L, M (Q₁, Q₃) | 0.12 | 1.42 (1.07, 1.80) | 1.42 (1.07, 1.80) | Z=0.005 | 0.996 |
| MONO, 10⁹/L, M (Q₁, Q₃) | 0.12 | 0.35 (0.26, 0.46) | 0.35 (0.26, 0.46) | Z=0.032 | 0.975 |
| PLT, 10⁹/L, M (Q₁, Q₃) | 0.12 | 227.00 (181.00, 281.00) | 227.00 (181.00, 281.00) | Z=0.036 | 0.971 |
| MPV, fl, Mean±SD | 1.40 | 10.14±1.49 | 10.13±1.48 | t=0.130 | 0.900 |
| PCT, %, Mean±SD | 1.40 | 0.24±0.08 | 0.24±0.08 | t=0.200 | 0.840 |
| RDW, %, Mean±SD | 0.12 | 13.51±2.39 | 13.51±2.39 | t=0.001 | 0.998 |
| PT, s, Mean±SD | 1.12 | 26.76±4.17 | 26.76±4.17 | t=0.001 | 0.996 |
| APTT, s, Mean±SD | 1.12 | 26.76±4.17 | 26.76±4.17 | t=0.001 | 0.996 |
| TT, s, Mean±SD | 1.12 | 26.76±4.17 | 26.76±4.17 | t=0.001 | 0.996 |
| FIB, g/L, Mean±SD | 1.12 | 3.07±0.84 | 3.07±0.84 | Z=0.030 | 0.973 |
| D-dimer, mg/L, M (Q₁, Q₃) | 1.40 | 0.30 (0.17, 0.60) | 0.30 (0.17, 0.60) | Z=0.049 | 0.961 |
| CEA, ng/mL, M (Q₁, Q₃) | 9.58 | 1.86 (0.90, 3.51) | 1.90 (0.91, 3.58) | Z=0.384 | 0.701 |
| Carbohydrate antigen 199, U/mL, M (Q₁, Q₃) | 10.40 | 10.94 (6.80, 19.42) | 11.11 (6.92, 19.62) | Z=0.279 | 0.780 |
| Carbohydrate antigen 125, U/mL, M (Q₁, Q₃) | 24.53 | 8.50 (5.90, 12.65) | 8.50 (6.30, 12.60) | Z=0.279 | 0.639 |
| GPT, U/L, M (Q₁, Q₃) | 0.58 | 16.00 (11.30, 24.00) | 16.00 (11.60, 24.00) | Z=0.040 | 0.968 |
| GOT, U/L, M (Q₁, Q₃) | 0.58 | 21.00 (17.00, 26.00) | 21.00 (17.00, 26.00) | Z=0.016 | 0.987 |
| Tbil, μmol/L, M (Q₁, Q₃) | 0.58 | 13.00 (9.25, 17.80) | 13.00 (9.20, 17.80) | Z=0.035 | 0.972 |
| ALB+GLB, g/L, Mean±SD | 0.58 | 66.61±7.48 | 66.62±7.49 | t=0.010 | 0.989 |
| ALB, g/L, Mean±SD | 0.58 | 41.63±4.90 | 41.63±4.91 | t=0.030 | 0.977 |
| GLB, g/L, Mean±SD | 0.58 | 24.99±4.56 | 24.99±4.56 | t=0.010 | 0.994 |
| A/G, Mean±SD | 0.58 | 1.71±0.33 | 1.71±0.33 | t=0.020 | 0.982 |
| BUN, mmol/L, M (Q₁, Q₃) | 0.58 | 4.80 (3.80, 5.80) | 4.80 (3.80, 5.80) | Z=0.044 | 0.965 |
| Cr, μmol/L, M (Q₁, Q₃) | 0.58 | 71.00 (61.00, 84.00) | 71.00 (61.00, 84.00) | Z=0.014 | 0.988 |
| UA, μmol/L, M (Q₁, Q₃) | 1.12 | 288.00 (221.00, 359.00) | 288.00 (221.00, 359.00) | Z=0.029 | 0.977 |

SD – standard deviation; M (Q₁, Q₃) – median (interquartile range); Hb – hemoglobin; RBC – red blood cell; WBC – white blood cell; NEUT – neutrophil; LYM – lymphocyte; MONO – monocyte; PLT – platelet; MPV – mean platelet volume; PCT – plateletcrit (the product of PLT and MPV); RDW – red blood cell distribution width; TT – thrombin time; APTT – activated partial thromboplastin time; PT – prothrombin time; FIB – fibrinogen; CEA – carcinoembryonic antigen; GPT – glutamic pyruvic transaminase; GOT – glutamic oxaloacetic transaminase; Tbil – total bilirubin; ALB – albumin; GLB – globulin; A/G – ALB-to-GLB ratio; BUN – blood urea nitrogen; Cr – creatinine; UA – uric acid.
### Supplementary Table 2. Comparison of characteristics among CRC stage I, II, III and colorectal adenoma in the training set.

| Characteristic                        | Total (n=600) | Colorectal adenoma (n=229) | Stage I (n=48) | Stage II (n=173) | Stage III (n=150) | Statistics | P  |
|---------------------------------------|--------------|-----------------------------|----------------|------------------|-------------------|------------|-----|
| Gender, n (%)                         |              |                             |                |                  |                   |            |     |
| Male                                  | 356 (59.33)  | 142 (62.01)                 | 30 (62.50)     | 99 (57.23)       | 85 (56.67)        | $\chi^2=1.639$ | 0.650 |
| Female                                | 244 (40.67)  | 87 (37.99)                  | 18 (37.50)     | 74 (42.77)       | 65 (43.33)        |            |     |
| Age, years, Mean±SD                  | 56.16±12.43  | 56.78±11.70                 | 55.98±11.68    | 57.14±13.05      | 54.15±12.91       | F=1.871    | 0.133 |
| BMI, kg/m², Mean±SD                  | 22.40±3.42   | 22.93±3.71                  | 22.31±3.34     | 21.90±3.06       | 22.20±3.34        | F=3.289    | 0.020 |
| Diabetes, n (%)                       |              |                             |                |                  |                   | $\chi^2=2.634$ | 0.452 |
| Yes                                   | 575 (95.83)  | 219 (95.63)                 | 157 (97.92)    | 110 (94.22)      | 46 (97.33)        |            |     |
| No                                    | 25 (4.17)    | 10 (4.37)                   | 1 (2.08)       | 10 (5.78)        | 4 (2.67)          |            |     |
| Inflammatory disease, n (%)           |              |                             |                |                  |                   | $\chi^2=2.058$ | 0.560 |
| Yes                                   | 498 (83.00)  | 195 (85.15)                 | 37 (77.08)     | 142 (82.08)      | 124 (82.67)       |            |     |
| No                                    | 102 (17.00)  | 24 (14.85)                  | 11 (22.92)     | 31 (17.29)       | 26 (17.33)        |            |     |
| Hb, g/L, Mean±SD                      | 125.48±24.67 | 137.79±19.88                | 125.35±21.08   | 115.50±25.08     | 118.24±23.99      | F=38.961   | <0.001|
| RBC, 10⁶/µL, Mean±SD                  | 4.24±0.63    | 4.49±0.55                   | 4.08±0.58      | 4.07±0.61        | 4.11±0.67         | F=20.940   | <0.001|
| WBC, 10⁹/L, M (Q₁, Q₃)               | 5.91 (4.82, 7.22) | 5.84 (4.81, 6.68)       | 5.94 (4.73, 7.05) | 6.21 (4.86, 7.72) | 6.00 (4.78, 7.75) | $\chi^2=6.071$ | 0.108 |
| NEUT, 10⁹/L, M (Q₁, Q₃)              | 3.69 (2.89, 5.01) | 3.54 (2.82, 4.50)       | 3.39 (2.94, 4.77) | 4.03 (2.92, 5.60) | 4.01 (2.90, 5.62) | $\chi^2=10.729$ | 0.013 |
| LYM, 10⁹/L, M (Q₁, Q₃)                | 1.43 (1.09, 1.79) | 1.54 (1.24, 1.87)       | 1.46 (1.05, 1.80) | 1.35 (1.08, 1.74) | 1.28 (1.01, 1.67) | $\chi^2=19.112$ | <0.001|
| MONO, 10⁹/L, M (Q₁, Q₃)              | 0.35 (0.26, 0.47) | 0.34 (0.27, 0.44)       | 0.35 (0.26, 0.45) | 0.35 (0.26, 0.45) | 0.36 (0.25, 0.49) | $\chi^2=1.896$ | 0.594 |
| PLT, 10⁹/L, M (Q₁, Q₃)                | 272.00 (181.00, 284.00) | 272.00 (173.00, 253.00) | 272.00 (173.00, 273.00) | 216.00 (186.00, 292.00) | 241.50 (210.00) | $\chi^2=32.943$ | <0.001|
| MPV, fl, Mean±SD                      | 10.20±1.49   | 10.55±1.44                  | 10.32±1.53     | 9.85±1.47        | 10.03±1.45        | F=8.670    | <0.001|
| PCT, %, M (Q₁, Q₃)                    | 0.23 (0.19, 0.28) | 0.22 (0.18, 0.26)       | 0.22 (0.18, 0.26) | 0.24 (0.20, 0.30) | 0.24 (0.19, 0.28) | $\chi^2=20.671$ | <0.001|
| RDW, %, Mean±SD                       | 13.59±2.52   | 13.12±1.61                  | 13.06±1.69     | 14.19±3.26       | 13.77±2.73        | F=7.084    | <0.001|
| PT, s, Mean±SD                        | 11.25±0.98   | 10.80±0.90                  | 11.31±0.97     | 11.55±0.86       | 11.55±1.00        | F=29.317   | <0.001|
| APTT, s, Mean±SD                      | 26.83±4.18   | 26.28±3.40                  | 27.01±5.37     | 27.21±4.27       | 27.17±4.67        | F=2.131    | 0.095 |
| TT, s, Mean±SD                        | 18.27±1.72   | 17.82±1.71                  | 18.38±1.00     | 18.36±1.58       | 18.80±1.89        | F=10.820   | <0.001|
| FIB, g/L, Mean±SD                     | 3.06±0.84    | 2.79±0.71                   | 2.80±0.60      | 3.40±0.88        | 3.17±0.87         | F=22.296   | <0.001|
| D-dimer, mg/L, M (Q₁, Q₃)             | 0.31 (0.17, 0.62) | 0.26 (0.16, 0.50)       | 0.24 (0.17, 0.46) | 0.43 (0.20, 0.84) | 0.36 (0.18, 0.69) | $\chi^2=19.280$ | <0.001|
| CEA, ng/mL, M (Q₁, Q₃)                | 1.82 (0.90, 3.80) | 1.37 (0.74, 2.35)       | 1.25 (0.59, 2.43) | 2.21 (1.06, 4.66) | 3.19 (1.25, 7.82) | $\chi^2=58.969$ | <0.001|
| Carbohydrate antigen 199, U/mL, M (Q₁, Q₃) | 10.84 (6.89, 20.45) | 10.34 (6.50, 17.06) | 9.87 (7.86, 15.02) | 15.16 (6.11, 20.35) | 7.73 (7.32, 32.82) | $\chi^2=19.579$ | <0.001|
| Carbohydrate antigen 125, U/mL, M (Q₁, Q₃) | 8.60 (6.30, 12.50) | 8.60 (6.20, 11.80) | 8.85 (6.45, 15.10) | 8.30 (6.20, 12.70) | 8.50 (6.40, 13.00) | $\chi^2=1.198$ | 0.754 |
Supplementary Table 2 continued. Comparison of characteristics among CRC stage I, II, III and colorectal adenoma in the training set.

| Characteristic | Total (n=600) | Colorectal adenoma (n=229) | Stage I (n=48) | Stage II (n=173) | Stage III (n=150) | Statistics | P |
|----------------|--------------|----------------------------|----------------|------------------|-------------------|------------|---|
| GPT, U/L, M (Q₁, Q₃) | 16.00 (11.00, 24.00) | 18.90 (14.00, 26.00) | 14.00 (11.50, 22.00) | 14.00 (10.00, 22.00) | 16.00 (11.00, 22.00) | X²=28.445 | <0.001 |
| GOT, U/L, M (Q₁, Q₃) | 21.00 (17.00, 26.00) | 23.00 (19.00, 27.00) | 19.00 (16.00, 25.00) | 19.00 (16.00, 25.00) | 20.50 (16.00, 26.00) | X²=23.232 | <0.001 |
| TBil, μmol/L, M (Q₁, Q₃) | 13.00 (9.30, 17.80) | 15.30 (11.40, 19.70) | 8.40 (8.10, 17.45) | 11.50 (8.30, 15.40) | 11.30 (8.50, 17.00) | X²=43.204 | <0.001 |
| ALB+GLB, g/L, Mean±SD | 66.72±7.47 | 69.57±6.86 | 65.23±6.20 | 65.36±7.36 | 64.42±7.53 | F=20.237 | <0.001 |
| ALB, g/L, Mean±SD | 41.70±4.78 | 42.38±4.21 | 42.09±4.37 | 41.32±5.07 | 40.97±5.26 | F=3.225 | 0.022 |
| GLB, g/L, Mean±SD | 25.02±4.60 | 27.19±4.80 | 23.14±3.47 | 24.03±3.90 | 23.45±4.07 | F=3.257 | <0.001 |
| A/G, Mean±SD | 17.13±3.23 | 16.04±3.09 | 18.55±3.07 | 17.55±3.04 | 17.86±3.22 | F=16.516 | <0.001 |
| BUN, mmol/L, M (Q₁, Q₃) | 4.80 (3.75, 5.80) | 5.20 (4.30, 6.20) | 4.15 (3.30, 5.30) | 4.70 (3.50, 5.80) | 4.40 (3.20, 5.40) | X²=42.558 | <0.001 |
| Cr, μmol/L, M (Q₁, Q₃) | 71.00 (61.00, 84.00) | 73.00 (62.00, 86.00) | 68.00 (62.00, 76.00) | 72.00 (61.00, 86.00) | 70.00 (59.00, 82.00) | X²=5.528 | 0.137 |
| UA, μmol/L, M (Q₁, Q₃) | 286.00 (220.00, 360.00) | 325.00 (258.00, 391.00) | 268.00 (215.00, 333.00) | 258.00 (207.00, 336.00) | 260.00 (200.00, 323.00) | X²=47.807 | <0.001 |
| NLR, M (Q₁, Q₃) | 2.39 (1.81, 3.95) | 2.72 (1.72, 3.08) | 1.76 (1.35, 3.41) | 1.93 (1.31, 3.41) | 2.95 (2.03, 4.67) | X²=25.636 | <0.001 |
| PLR, M (Q₁, Q₃) | 156.44 (116.62, 229.85) | 136.30 (106.14, 174.32) | 141.73 (111.07, 234.94) | 187.38 (130.72, 264.84) | 180.18 (128.30, 269.23) | X²=48.147 | <0.001 |
| MLR, M (Q₁, Q₃) | 0.24 (0.18, 0.35) | 0.21 (0.18, 0.30) | 0.24 (0.17, 0.34) | 0.25 (0.17, 0.39) | 0.27 (0.19, 0.38) | X²=11.564 | 0.009 |
| MPV/PLT, M (Q₁, Q₃) | 4.85 (3.37, 5.96) | 4.88 (3.83, 6.20) | 4.64 (3.55, 6.43) | 4.81 (3.83, 6.53) | 4.39 (3.28, 5.83) | X²=37.396 | <0.001 |
| FAR, M (Q₁, Q₃) | 6.92 (5.87, 8.55) | 6.23 (5.43, 7.47) | 6.30 (5.66, 7.51) | 7.78 (6.58, 9.72) | 7.47 (6.26, 9.39) | X²=70.445 | <0.001 |
| PNI, Mean±SD | 417.01±47.83 | 423.84±42.09 | 420.92±43.66 | 413.22±50.71 | 409.71±52.64 | F=3.225 | 0.022 |
| SII, M (Q₁, Q₃) | 581.29 (377.24, 1034.25) | 472.22 (339.95, 698.73) | 473.36 (368.35, 832.25) | 689.45 (459.50, 1385.74) | 707.59 (382.92, 1357.32) | X²=43.260 | <0.001 |
| dNLR, M (Q₁, Q₃) | 1.84 (1.38, 2.61) | 1.73 (1.35, 2.15) | 1.74 (1.30, 2.32) | 1.98 (1.45, 2.97) | 2.14 (1.46, 3.14) | X²=20.077 | <0.001 |
| HPR, M (Q₁, Q₃) | 0.57 (0.42, 0.75) | 0.67 (0.53, 0.81) | 0.60 (0.41, 0.81) | 0.46 (0.33, 0.62) | 0.52 (0.33, 0.70) | X²=72.312 | <0.001 |

SD – standard deviation; M (Q₁, Q₃) – median (interquartile range); BMI – body mass index; Hb – hemoglobin; RBC – red blood cell; WBC – white blood cell; NEUT – neutrophil; LYM – lymphocyte; MONO – monocyte; PLT – platelet; MPV – mean platelet volume; PCT – plateletcrit (the product of PLT and MPV); RDW – red blood cell distribution width; PT – prothrombin time; APTT – activated partial thromboplastin time; TT – thrombin time; FIB – fibrinogen; CEA – carcinoembryonic antigen; GPT – glutamic pyruvic transaminase; GOT – glutamic oxaloacetic transaminase; TBil – total bilirubin; ALB – albumin; GLB – globulin; A/G – ALB-to-GLB ratio; BUN – blood urea nitrogen; Cr – creatinine; UA – uric acid; NLR – NEUT-to-LYM ratio; PLR – PLT-to-LYM ratio; MLR – MONO-to-LYM ratio; FAR – FIB-to-ALB ratio; PNI – 10×ALB+0.005×LYM; SII – systemic immune-inflammatory index; dNLR – NEUT/(WBC-NEUT); HPR – Hb-to-PLT ratio.
### Supplementary Table 3. Comparison of characteristics among CRC (poorly differentiated, moderately differentiated and well-differentiated) and colorectal adenoma in the training set.

| Characteristic        | Total (n=600) | Groups                                      | Groups                                      | Groups                                      | Statistics | P      |
|-----------------------|---------------|---------------------------------------------|---------------------------------------------|---------------------------------------------|------------|--------|
|                       |               | Colorectal adenoma (n=229)                 | Poorly differentiated (n=65)                 | Moderately well-differentiated (n=306)      |            |        |
| Gender, n (%)         |               | Male                                        | Female                                      |                                             |            |        |
|                       |               | 356 (59.33)                                 | 244 (40.67)                                 |                                              |            |        |
|                       |               | 142 (62.01)                                 | 87 (37.99)                                  |                                              |            |        |
|                       |               | 35 (53.85)                                  | 50 (76.92)                                  |                                              |            |        |
|                       |               | 179 (58.50)                                 | 127 (41.50)                                 |                                              |            |        |
|                       |               |                                             |                                             |                                             |            |        |
| Age, years, Mean±SD   | 56.16±12.43   | 56.78±11.70                                 | 57.00±13.51                                 | 55.52±12.73                                 | F=0.840    | 0.432  |
|                       |               |                                             |                                             |                                             |            |        |
| BMI, kg/m², Mean±SD   | 22.40±3.42    | 22.93±3.71                                  | 22.24±3.03                                 | 22.04±3.23                                  | F=4.58     | 0.011  |
|                       |               |                                             |                                             |                                             |            |        |
| Diabetes, n (%)       |               |                                             |                                             |                                             |            |        |
|                       |               |                                             |                                             |                                              |            |        |
|                       |               | No                                          | Yes                                         |                                              |            |        |
|                       |               | 575 (95.83)                                 | 25 (4.17)                                   |                                             |            |        |
|                       |               | 219 (95.63)                                 | 10 (4.37)                                   |                                             |            |        |
|                       |               | 60 (92.31)                                  | 5 (7.69)                                    |                                             |            |        |
|                       |               | 296 (96.73)                                 | 10 (3.27)                                   |                                             |            |        |
|                       |               |                                             |                                             |                                             |            |        |
| Inflammatory disease, n (%) |      |                                             |                                             |                                              |            |        |
|                       |               |                                             |                                             |                                              |            |        |
|                       |               | No                                          | Yes                                         |                                              |            |        |
|                       |               | 498 (83.00)                                 | 102 (17.00)                                 |                                             |            |        |
|                       |               | 195 (85.15)                                 | 34 (14.85)                                  |                                             |            |        |
|                       |               | 60 (92.31)                                  | 5 (7.69)                                    |                                             |            |        |
|                       |               | 243 (79.41)                                 | 73 (20.59)                                  |                                             |            |        |
|                       |               |                                             |                                             |                                             |            |        |
| Hb, g/L, Mean±SD      | 125.48±24.67  | 137.79±19.88                                | 118.89±22.00                               | 118.09±24.79                                | F=54.365   | <0.001 |
|                       |               |                                             |                                             |                                             |            |        |
| RBC, 10⁹/L, Mean±SD   | 4.24±0.63     | 4.49±0.55                                   | 4.03±0.65                                  | 4.10±0.63                                   | F=31.548   | <0.001 |
|                       |               |                                             |                                             |                                             |            |        |
| WBC, 10⁹/L, M (Q₁, Q₃) | 5.91     | (4.82, 7.22)                                | 6.84                                       | (4.88, 7.79)                                | F=4.856    | 0.088  |
|                       |               |                                             |                                             |                                             |            |        |
| NEUT, 10⁹/L, M (Q₁, Q₃) | 3.69     | (2.89, 5.01)                                | 3.54                                       | (2.82, 4.50)                                | F=10.878   | 0.004  |
|                       |               |                                             |                                             |                                             |            |        |
| LYM, 10⁹/L, M (Q₁, Q₃) | 1.43     | (1.09, 1.79)                                | 1.54                                       | (1.24, 1.87)                                | F=22.893   | <0.001 |
|                       |               |                                             |                                             |                                             |            |        |
| MONO, 10⁹/L, M (Q₁, Q₃) | 0.35     | (0.26, 0.47)                                | 0.34                                       | (0.27, 0.44)                                | F=3.859    | 0.145  |
|                       |               |                                             |                                             |                                             |            |        |
| PLT, 10⁹/L, M (Q₁, Q₃) | 229.00   | (181.00, 84.00)                             | 212.00                                     | (173.00, 253.00)                            | F=27.175   | <0.001 |
|                       |               |                                             |                                             |                                             |            |        |
| MPV, fl, Mean±SD      | 10.20±1.49    | 10.55±1.44                                  | 9.96±1.42                                  | 9.99±1.49                                   | F=10.820   | <0.001 |
|                       |               |                                             |                                             |                                             |            |        |
| PCT, %, M (Q₁, Q₃)    | 0.23          | (0.19, 0.28)                                | 0.22                                       | (0.18, 0.27)                                | F=19.367   | <0.001 |
|                       |               |                                             |                                             |                                             |            |        |
| RDW, %, Mean±SD       | 13.59±2.52    | 13.12±1.61                                  | 13.52±2.34                                 | 13.95±3.01                                  | F=7.266    | <0.001 |
|                       |               |                                             |                                             |                                             |            |        |
| PT, s, Mean±SD        | 11.25±0.98    | 10.80±0.90                                  | 11.71±1.23                                 | 11.48±0.85                                  | F=44.481   | <0.001 |
|                       |               |                                             |                                             |                                             |            |        |
| APTT, s, Mean±SD      | 26.83±4.18    | 26.28±3.40                                  | 27.85±5.32                                 | 27.02±4.40                                  | F=4.244    | 0.015  |
|                       |               |                                             |                                             |                                             |            |        |
| TT, s, Mean±SD        | 18.27±1.72    | 17.82±1.71                                  | 18.79±2.41                                 | 18.49±1.46                                  | F=13.933   | <0.001 |
|                       |               |                                             |                                             |                                             |            |        |
| FIB, g/L, Mean±SD     | 3.06±0.84     | 2.79±0.71                                   | 3.36±0.83                                  | 3.20±0.87                                   | F=22.828   | <0.001 |
|                       |               |                                             |                                             |                                             |            |        |
| D-dimer, mg/L, M (Q₁, Q₃) | 0.31     | (0.17, 0.62)                                | 0.26                                       | (0.16, 0.50)                                | F=11.271   | 0.004  |
|                       |               |                                             |                                             |                                             |            |        |
| CEA, ng/mL, M (Q₁, Q₃) | 1.82     | (0.90, 3.80)                                | 1.37                                       | (0.74, 2.35)                                | F=35.221   | <0.001 |
|                       |               |                                             |                                             |                                             |            |        |
| Carbohydrate antigen 199, U/mL, M (Q₁, Q₃) | 6.89, 20.45 | (6.50, 17.06)                              | 12.38                                     | (6.95, 22.43)                               | F=5.860    | 0.053  |
|                       |               |                                             |                                             |                                             |            |        |
| Carbohydrate antigen 125, U/mL, M (Q₁, Q₃) | 8.60   | (6.30, 12.50)                               | 8.60                                       | (6.10, 13.80)                               | F=9.025    | 0.630  |
### Supplementary Table 3 continued. Comparison of characteristics among CRC (poorly differentiated, moderately differentiated and well-differentiated) and colorectal adenoma in the training set.

| Characteristic | Total (n=600) | Colorectal adenoma (n=229) | Poorly differentiated (n=65) | Moderately and well-differentiated (n=306) | Statistics | P  |
|---------------|--------------|-----------------------------|----------------------------|------------------------------------------|------------|----|
| GPT, U/L, M (Q1, Q3) | 16.00 (11.00, 24.00) | 18.90 (14.00, 26.00) | 14.00 (10.00, 19.00) | 15.00 (11.00, 23.00) | χ²=30.629 | <0.001 |
| GOT, U/L, M (Q1, Q3) | 21.00 (17.00, 26.00) | 23.00 (19.00, 27.00) | 20.00 (17.00, 25.00) | 20.00 (16.00, 25.00) | χ²=22.080 | <0.001 |
| Tbil, μmol/L, M (Q1, Q3) | 13.00 (9.30, 17.80) | 15.30 (11.40, 19.70) | 10.90 (8.50, 14.50) | 8.00 (8.30, 16.10) | χ²=40.947 | <0.001 |
| ALB+GLB, g/L, Mean±SD | 66.72±7.47 | 69.57±6.86 | 64.71±7.69 | 65.01±7.21 | F=29.651 | <0.001 |
| ALB, g/L, Mean±SD | 41.70±4.78 | 42.38±4.21 | 41.23±5.35 | 41.29±5.01 | F=3.811 | 0.023 |
| GLB, g/L, Mean±SD | 25.02±4.60 | 27.19±4.80 | 23.40±4.15 | 23.72±3.87 | F=47.653 | <0.001 |
| A/G, Mean±SD | 17.13±3.23 | 16.04±3.09 | 17.96±3.20 | 17.77±3.12 | F=22.772 | <0.001 |
| BUN, mmol/L, M (Q1, Q3) | 4.80 (3.75, 5.80) | 4.60 (3.40, 5.40) | 4.60 (3.40, 5.70) | 5.20 (3.40, 5.70) | χ²=36.821 | <0.001 |
| Cr, μmol/L, M (Q1, Q3) | 71.00 (61.00, 84.00) | 73.00 (62.00, 86.00) | 74.00 (63.00, 83.00) | 69.50 (60.00, 84.00) | χ²=3.604 | 0.165 |
| UA, μmol/L, M (Q1, Q3) | 286.00 (220.00, 360.00) | 325.00 (258.00, 391.00) | 271.00 (225.00, 364.00) | 258.00 (205.00, 324.00) | χ²=48.833 | <0.001 |
| NLR, M (Q1, Q3) | 2.55 (1.81, 3.95) | 2.22 (1.72, 3.08) | 3.72 (2.36, 5.65) | 2.68 (1.84, 4.28) | χ²=30.638 | <0.001 |
| PLR, M (Q1, Q3) | 156.44 (116.62, 229.85) | 136.30 (106.14, 174.32) | 187.41 (137.18, 252.66) | 169.92 (124.12, 269.23) | χ²=42.616 | <0.001 |
| M LR, M (Q1, Q3) | 0.24 (0.18, 0.35) | 0.21 (0.18, 0.30) | 0.25 (0.19, 0.40) | 0.26 (0.18, 0.38) | χ²=9.869 | 0.007 |
| MPV/PLT, M (Q1, Q3) | 4.39 (3.37, 5.96) | 4.85 (3.83, 6.20) | 4.25 (3.25, 6.14) | 3.99 (2.93, 5.65) | χ²=29.131 | <0.001 |
| FAR, M (Q1, Q3) | 6.92 (5.87, 8.55) | 6.23 (5.43, 7.47) | 8.10 (6.60, 9.37) | 7.40 (6.22, 9.16) | χ²=54.759 | <0.001 |
| PNI, Mean±SD | 417.01±47.83 | 423.84±42.09 | 412.33±53.45 | 412.90±50.13 | F=3.811 | 0.023 |
| SII, M (Q1, Q3) | 581.29 (377.24, 1034.25) | 472.22 (339.95, 698.73) | 510.66 (334.54, 636.54) | 636.51 (399.65, 1308.46) | χ²=39.079 | <0.001 |
| dNLR, M (Q1, Q3) | 1.38 (1.38, 2.01) | 1.73 (1.35, 2.15) | 2.46 (1.71, 3.60) | 1.93 (1.39, 2.85) | χ²=26.104 | <0.001 |
| HPR, M (Q1, Q3) | 0.57 (0.42, 0.75) | 0.67 (0.53, 0.81) | 0.52 (0.40, 0.75) | 0.50 (0.33, 0.69) | χ²=61.731 | <0.001 |

SD – standard deviation; M (Q1, Q3) – median (interquartile range); BMI – body mass index; Hb – hemoglobin; RBC – red blood cell; WBC – white blood cell; NEUT – neutrophil; LYM – lymphocyte; MONO – monocyte; PLT – platelet; MPV – mean platelet volume; PCT – plateletcrit (the product of PLT and MPV); RDW – red blood cell distribution width; PT – prothrombin time; APTT – activated partial thromboplastin time; TT – thrombin time; FIB – fibrinogen; CEA – carcinoembryonic antigen; GPT – glutamic pyruvic transaminase; GOT – glutamic oxaloacetic transaminase; TBI – total bilirubin; ALB – albumin; GLB – globulin; A/G = ALB-to-GLB ratio; BUN – blood urea nitrogen; Cr – creatinine; UA – uric acid; NLR – NEUT-to-LYM ratio; PLR – PLT-to-LYM ratio; M LR – MONO-to-LYM ratio; FAR – FIB-to-ALB ratio; PNI – 10×ALB+0.005×LYM; SII – systemic immune-inflammatory index; dNLR – NEUT/(WBC-NEUT); HPR – Hb-to-PLT ratio.
