

基于常规检验指标构建区分结直肠腺瘤与结直肠癌的预测模型

林俊生¹, 应紫灵², 黄政渊³, 祝先进¹, 曹颖平¹, 卢婷霞¹

A Prediction Model for Colorectal Adenoma and Colorectal Cancer Based on Routine Test
LIN Junsheng¹, YING Ziling², HUANG Zhengyuan³, ZHU Xianjin¹, CAO Yingping¹,
LU Pingxia¹

1. Department of Laboratory Medicine, Fujian Medical University Union Hospital, Fuzhou 350001, China; 2. Department of Laboratory Medicine, School of Medical Technology and Engineering, Fujian Medical University, Fuzhou 350122, China; 3. Department of Emergency Surgery, Fujian Medical University Union Hospital, Fuzhou 350001, China

Corresponding Author: LU Pingxia, E-mail: 1248959330@qq.com

Abstract: Objective To analyze the routine test parameter levels of patients with colorectal adenoma and colorectal cancer, and develop a prediction model. **Methods** A total of 580 patients diagnosed with colorectal adenoma (117 patients) and colorectal cancer (463 patients) were included in the retrospective study. The patients were randomly divided into two groups according to a 7:3 ratio: a training set with 406 cases and a validation set with 174 cases. Logistic regression analysis was used to establish a prediction model, and a nomogram was drawn. The model's discrimination, calibration, and clinical applicability were evaluated using receiver operating characteristic curve (ROC), calibration plot, and decision curve analysis (DCA). **Results** Univariate logistic regression analysis identified 13 potential predictors: age, fecal occult blood test (FOBT), fibrinogen (FIB), thrombin time (TT), albumin (ALB), white blood cell value (WBC), neutrophil count (NEUT#), hematocrit value (HCT), mean corpuscular hemoglobin (MCH), red cell distribution width (RDW), platelet count (PLT), mean platelet volume (MPV), and activated partial thromboplastin time (APTT). Multivariate logistic regression analysis showed MPV, FIB, ALB, FOBT, TT, and HCT were risk factors for colorectal cancer in patients with colorectal adenoma ($P < 0.05$). A nomogram was constructed based on these predictors to build a prediction model. The AUC of the ROC curve was 0.915 for colorectal cancer in the training set and 0.836 in the validation set. Calibration plots demonstrated high prediction accuracy and good model calibration. DCA results indicated the prediction model provided greater net benefit compared with the extreme models at threshold probabilities of approximately 55%-95%. **Conclusion** The developed prediction model exhibits satisfactory discrimination, calibration, and clinical applicability. The model can serve as an auxiliary tool in distinguishing between colorectal adenoma and colorectal cancer in patients.

Key words: Colorectal adenoma; Colorectal cancer; Prediction model; Nomogram

Funding: Fujian Province Joint Funds for the Innovation of Science and Technology (No. 2020Y9088)

Competing interests: The authors declare that they have no competing interests.

摘要: 目的 回顾性分析结直肠腺瘤(CRA)患者与结直肠癌(CRC)患者的常规检验指标并建立预测模型。**方法** 选取580例诊断为CRA患者117例和CRC患者463例,按照7:3随机分为训练集406例、验证集174例。采用Logistic回归法建立模型,并绘制列线图。分别采用受试者工作特征曲线(ROC)、校准图、临床决策曲线(DCA)评估预测模型的区分度、校准度和临床应用的有效性。

结果 单因素Logistic回归分析初步筛选出13个候选变量,包括年龄、大便隐血试验(FOBT)、

纤维蛋白原(FIB)、凝血酶时间(TT)、白蛋白(ALB)、白细胞计数(WBC)、中性粒细胞计数(NEUT#)、红细胞比积(HCT)、平均血红蛋白含量(MCH)、红细胞体积分布宽度(RDW)、血小板计数(PLT)、平均血小板体积(MPV)、活化部分凝血活酶时间(APTT)。多因素Logistic回归显示

收稿日期: 2023-11-07; 修回日期: 2023-12-11

基金项目: 福建省科技创新联合基金项目(2020Y9088)

作者单位: 1. 350001 福州, 福建医科大学附属协和医院检验科; 2. 350122 福州, 福建医科大学医学技术与工程学院医学检验系; 3. 350001 福州, 福建医科大学附属协和医院急诊外科

通信作者: 卢婷霞(1981-), 女, 硕士, 副主任检验师, 主要从事肿瘤及肿瘤耐药的相关研究, E-mail: 1248959330@qq.com, ORCID: 0000-0001-7951-807X

作者简介: 林俊生(1978-), 男, 本科, 主管检验师, 主要从事临床检验工作, ORCID: 0000-0001-9301-6738

MPV、FIB、ALB、FOBT、TT以及HCT是CRA患者患CRC的危险因素 ($P < 0.05$)。构建预测模型, 训练集和验证集发生CRC的ROC曲线下面积 (AUC) 分别为0.915和0.836。校准曲线显示, 模型预测准确率较高, 校正能力良好。DCA结果显示, 当阈值概率为55%~95%时, 训练集与验证集的净收益均大于2个极端模型, 即临床有益。**结论** 本研究构建的预测模型具有较好的区分度、校准度和临床应用的有效性, 可作为区分CRA和CRC患者的辅助工具。

关键词: 结直肠腺瘤; 结直肠癌; 预测模型;

列线图

中图分类号: R446.1; R735.3

开放科学(资源服务)标识码(OSID):



0 引言

结直肠癌 (colorectal carcinoma, CRC) 是全球第三大常见癌症, 癌症死亡率占全球第二位^[1]。CRC的发展大多遵循“腺瘤-癌”序列, 整个发展过程一般需要5~10年, 及时发现并对癌前病变进行干预可以有效阻止CRC的发生。鉴于一级预防对于CRC而言相对困难, 因此降低CRC的发病率和死亡率主要依赖于早期筛查^[2-3]。结直肠腺瘤 (colorectal adenoma, CRA) 是一种起源于结直肠黏膜腺体上皮的良性肿瘤^[4]。据估计, 约95%的CRC是由腺瘤性息肉发展而来的。由于大约40%≥50岁人群有一个或多个腺瘤性息肉, 在癌症转移之前识别这些息肉并切除非常重要^[5]。

目前, 临床常见的CRC筛查手段包括结肠镜检查、直肠镜检查、血清肿瘤标志物, 还有计算机断层扫描 (computed tomography, CT)。然而, 这些方法均存在一些局限性, 如侵入性、高成本、低灵敏度和低特异性^[6-7]。粪便隐血试验 (fecal occult blood test, FOBT) 是一种常用的CRC筛查方法, 已被证实可有效降低CRC的死亡率, 但仍有一定局限性, 包括对人血红蛋白的间接测量, 并且血红蛋白测试的敏感性相对较低^[8]。近年来, 人们开始关注将FOBT与其他标志物相结合进行CRC诊断或构建预测模型, 以提高CRC的检出率和预测癌症发生。Záhorec等^[9]使用血红蛋白 (hemoglobin, Hb)、白蛋白 (albumin, ALB)、纤维蛋白原 (fibrinogen, FIB) 和中性粒细胞与淋巴细胞比值 (neutrophil count/lymphocyte count ratio, NLR) 构建了一个预测模型, 用于区分CRA和CRC。本课题组成员在前期研究中发现, 血小板比容 (platelet hematocrit, PCT) 联合肿瘤标志物癌胚抗原 (carcinoembryonic antigen, CEA) 可提高CRC的检出率^[10]。淋巴细胞 (lymphocyte count, LYM) 是反映机体免疫应答功

能、对抗外界感染和监测体内细胞变异的重要细胞成分。全身炎症反应细胞比率的高阳性预测值在对怀疑患有CRC的患者进行评估时具有一定的意义^[11]。

近年来, 预测模型在疾病诊断中得到广泛应用。许多研究显示, 通过结合多种标志物构建诊断和预测模型比单一标志物更有效。因此, 本研究旨在利用常见检验手段 (包括血常规、FOBT、生化和凝血指标) 构建预测模型, 用于区分CRA和CRC, 旨在对CRC的早期筛查和干预癌前病变提供参考。

1 资料与方法

1.1 资料

选取2019年5月至2021年11月在福建医科大学附属协和医院就诊且符合标准的CRA或CRC的患者580例, 男性349例, 女性231例, 年龄52~68岁。CRA 117例, 男性69例, 女性48例, 年龄51~65岁; CRC 463例, 男性280例, 女性183例, 年龄52~69岁。为建立和验证预测模型将研究对象按照7:3随机分为两部分: 训练集406例, 验证集174例。纳入标准: 经病理确诊为CRA或CRC且入院前未经过任何医学处理的患者; 排除标准: 临床资料不全或丢失的患者, 患有血液系统疾病、慢性肝病、慢性肾病、长期服用激素, 明显合并其他脏器感染性疾病的患者。本研究经医院伦理委员会审核批准 (2021KJCX013)。所有患者均已获得知情同意。

1.2 基本特征和实验室指标收集

收集患者一般临床资料, 包括病理学诊断、性别、年龄。术前的实验室检查指标: 白细胞 (white blood cell value, WBC)、中性粒细胞计数 (neutrophil count, NEUT#)、淋巴细胞计数 (lymphocyte count, LYM#)、单核细胞计数 (monocyte count, MONO#)、血红蛋白 (hemoglobin, Hb)、红细胞比容 (hematocrit value, HCT)、平均红细胞血红蛋白含量 (mean corpuscular hemoglobin, MCH)、红细胞体积分布宽度 (red cell distribution width, RDW)、血小板计数 (platelet count, PLT)、血小板比容 (thrombocytocrit, PCT)、平均血小板体积 (mean platelet volume, MPV)、血小板分布宽度 (platelet distribution width, PDW)、凝血酶原时间 (prothrombin time, PT)、活化部分凝血活酶时间 (activated partial thromboplastin time, APTT)、凝血酶时间 (thrombin time, TT)、纤维

蛋白原（fibrinogen, FIB）、白蛋白（albumin, ALB）、甘油三酯（triglyceride, TG）、总胆固醇（cholesterol, CHOL）和粪便隐血试验（fecal occult blood test, FOBT）。

1.3 统计学方法

采用SPSS 22.0和R4.2.2软件对数据进行统计学分析。符合正态分布的计量资料以均数±标准差（ $\bar{x} \pm s$ ）表示，计数资料以百分比表示；非正态分布的计量资料以中位数和四分位数间距（median [interquartile range, IQR]）表示。采用Box-Tidwell法评估连续型自变量与因变量的logit转换值之间是否存在线性关系，将不存在线性关系的连续型自变量转换为二分类变量。采用方差膨胀因子（variance inflation factor, VIF）诊断自变量之间的多重共线性（ $VIF > 10$ ），并剔除引起共线性的自变量；采用R语言mice包中的多重插补法处理缺失数据（设定随机数种子=2 023），并将剔除共线性变量后的完整数据集以7:3比例随机分为训练集与验证集（设定随机数种子=20 010 123）。两组间比较分别采用两独立样本t检验（正态分布）和非参数秩和检验（非正态分布）；两组样本率的比较采用卡方检验。采用Logistic回归法建立模型，采用单因素Logistic回归分析初步筛选自变量（ $P < 0.1$ 为差异有统计学意义），根据受试者工作特征曲线（ROC）得出的最佳诊断值将筛选出的连续型自变量转换为二分类变量，并进行多因素Logistic回归分析，以最终确定预测CRC风险的危险因素，从而建立预测模型，并绘制可视列线图（Nomogram）。使用ROC曲线、校准图、决策曲线分析（decision curve analysis, DCA）进行模型评估。 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 连续型自变量与因变量的logit转换值之间线性关系的验证

本研究共有43项变量纳入模型分析，包括2个分类变量gender和FOBT；20个连续自变量Age、WBC、NEU、LYM、MON、Hb、HCT、MCH、RDW、PLT、PCT、MPV、PDW、PT、APTT、FIB、TT、ALB、TG和CHOL；20个交互作用项Age*lnAge、WBC*lnWBC、NEUT#*lnNEUT、LYMPH#*lnLYM、MONO#*lnMON、Hb*lnHb、HCT*lnHCT、MCH*lnMCH、RDW-CV*lnRDW-CV、PLT*lnPLT、PCT*lnPCT、MPV*lnMPV、PDW*lnPDW、PT*lnPT、APTT*lnAPTT、FIB*lnFIB、TT*lnTT、ALB*lnALB、TG*lnTG、CHOL*lnCHOL和常数项。采用Box-Tid-

well法评估连续型自变量与因变量的logit转换值之间是否存在线性关系，本研究中除ALB外的连续型自变量与因变量logit转换值之间均存在线性关系，见表1，因此，将ALB转换为二分类变量（ALB赋值 $< 39.3 = 0$ ， $\geq 39.3 = 1$ ）。

表 1 Box-Tidwell法验证连续型自变量与因变量的logit转换值之间是否存在线性关系

Table 1 Linear relationship between continuous independent variables and the logit conversion values of dependent variables by Box-Tidwell method

Variables	P	Variables	P
Age	0.47	Age*lnAge	0.59
WBC	0.64	WBC*lnWBC	0.42
NEUT#	0.97	NEUT#*lnNEUT	0.68
LYMPH#	0.17	LYMPH#*lnLYMPH	0.06
MONO#	<0.01	MONO#*lnMONO#	0.73
Hb	0.42	Hb*lnHb	0.45
HCT	0.20	HCT*lnHCT	0.21
MCH	0.30	MCH*lnMCH	0.34
RDW	0.53	RDW*lnRDW	0.45
PLT	0.88	PLT*lnPLT	0.99
PCT	0.08	PCT*lnPCT	0.29
MPV	0.25	MPV*lnMPV	0.26
PDW	0.06	PDW*lnPDW	0.07
PT	0.97	PT*lnPT	0.96
APTT	0.43	APTT*lnAPTT	0.42
FIB	0.62	FIB*lnFIB	0.87
TT	0.03	TT*lnTT	0.03
ALB	<0.001	ALB*lnALB	<0.001
TG	0.23	TG*lnTG	0.32
CHOL	0.01	CHOL*lnCHOL	0.01
		Constant	0.02

Notes: WBC: white blood cell value; NEUT#: neutrophil count; LYMPH#: lymphocyte count; MONO#: monocyte count; Hb: hemoglobin; HCT: hematocrit value; MCH: mean corpuscular hemoglobin; RDW: red cell distribution width; PLT: platelet count; PCT: thrombocytocrit; MPV: mean platelet volume; PDW: platelet distribution width; PT: prothrombin time; APTT: activated partial thromboplastin time; TT: thrombin time; FIB: fibrinogen; ALB: albumin; TG: triglyceride; CHOL: cholesterol; FOBT: fecal occult blood test.

2.2 变量间的多重共线性诊断

诊断结果显示，Hb、HCT、PLT和PCT的VIF>10，存在多重共线性。在本研究中，血红蛋白（Hb）与红细胞压积（HCT），血小板计数（PLT）与血小板压积（PCT）在意义上相当，因此我们选择采用逐步回归的方法进行变量筛选。剔除其中引起共线性的2个指标（Hb和PCT），剔除后变量间不存在多重共线性，见表2。

2.3 训练集和验证集患者的基线特征和实验室指标比较

训练集与验证集患者疾病分组、性别、年龄、WBC、NEUT#、LYMPH#、MONO#、HCT、MCH、

表 2 自变量间的共线性诊断

Table 2 Collinearity diagnostics among covariates

Before collinearity diagnostics		After collinearity diagnostics	
Variables	VIF	Variables	VIF
Age	1.17	Age	1.17
Gender	1.39	Gender	1.38
WBC	1.68	WBC	1.68
NEU#	2.14	NEU	2.12
LYM#	1.31	LYM	1.31
MON#	1.46	MON	1.46
Hb	13.56	HCT	2.37
HCT	12.19	MCH	1.76
MCH	1.84	RDW	2.08
RDW	2.16	PLT	1.81
PLT	20.30	MPV	1.89
PCT	17.79	PDW	1.84
MPV	3.10	PT	1.40
PDW	1.88	APTT	1.33
PT	1.40	FIB	1.58
APTT	1.33	TT	1.10
FIB	1.59	ALB	1.27
TT	1.11	TG	1.11
ALB	1.28	CHOL	1.27
TG	1.11	FOBT	1.15
CHOL	1.27		
FOBT	1.16		

Note: VIF: variance inflation factor.

MPV、PLT、PDW、PT、APTT、FIB、TT、ALB、TG、CHOL和FOBT所占比例比较，差别均无统计学意义（ $P>0.05$ ）；RDW所占比例比较，差别有统计学意义（ $P<0.01$ ）。可认为训练集和验证集来自同一总体，均衡性良好，见表3。

2.4 单因素Logistic回归分析与变量类型转化

根据训练集中患者的基线特征和部分实验室指标，使用单因素Logistic回归分析从20个变量中初步筛选出13个候选变量，选择以 $P<0.1$ 表示差别有统计学意义，以确保不漏掉可能具有重要意义的变量，见表4。

2.5 多因素Logistic回归与预测模型构建

以疾病诊断为因变量（赋值情况见表5），以单因素Logistic回归分析筛选的年龄变量及12个二分类预测变量为自变量构建多因素Logistic回归预测模型。根据筛选出的最终变量绘制多因素回归分析森林图（图1）得出，MPV[OR=4.67（1.86~11.72）， $P=0.001$]、FIB[OR=5.54（2.38~12.91）， $P<0.001$]、ALB[OR=5.90（2.07~16.85）， $P=0.001$]、FOBT[OR=

表 3 训练集与验证集患者的基线特征和实验室指标比较

Table 3 Comparison of baseline characteristics and laboratory parameters of patients between training set and validation set

Parameters	Total(n=580)	Training set(n=406)	Validation set(n=174)	P
Groups (n(%))				0.32
CRA	117(20.20)	77(18.97)	40(22.99)	
CRC	463(79.80)	329(81.03)	134(77.01)	
Gender (n(%))				0.28
Female	231(39.82)	168(41.38)	63(36.20)	
Male	349(60.18)	238(58.62)	111(63.80)	
Age (years)	59.48±12.31	59.40±12.37	59.68±12.19	0.80
FOBT (n(%))				0.65
Positive	363(62.59)	257(63.30)	106(60.92)	
Negative	217(37.41)	149(36.70)	68(39.08)	
ALB (g/L, n(%))				0.92
<39.30	107(18.45)	74(18.23)	33(18.97)	
≥39.30	473(81.55)	332(81.77)	141(81.03)	
FIB (g/L)	3.73±0.92	3.73±0.95	3.73±0.85	0.99
TT (s)	16.59±1.23	16.57±1.25	16.63±1.18	0.54
MPV (fl)	10.23±1.06	10.25±1.07	10.18±1.05	0.45
CHOL (mmol/L)	4.67±0.93	4.68±0.93	4.65±0.93	0.68
WBC (10 ⁹ /L)	5.84(4.82-7.02)	5.77(4.77-6.98)	6.00(5.08-7.17)	0.15
NEUT# (10 ⁹ /L)	3.39(2.66-4.36)	3.35(2.56-4.32)	3.49(2.74-4.49)	0.19
LYMPH# (10 ⁹ /L)	1.80(1.47-2.17)	1.78(1.46-2.17)	1.83(1.50-2.18)	0.54
MONO# (10 ⁹ /L)	0.36(0.29-0.46)	0.36(0.29-0.46)	0.36(0.29-0.47)	0.71
HCT (%)	40.20(36.70-43.10)	40.00(36.23-42.80)	40.50(37.60-43.92)	0.08
MCH (pg)	29.90(28.00-31.10)	29.90(27.92-31.10)	29.80(28.22-31.20)	0.74
RDW (%)	12.80(12.30-13.70)	12.90(12.40-13.97)	12.60(12.20-13.28)	<0.01
PLT (10 ⁹ /L)	247.00(209.00-296.00)	248.50(209.00-296.00)	243.00(209.00-294.80)	0.93
PDW (%)	11.40(10.20-12.90)	11.40(10.20-12.70)	11.35(10.12-13.28)	0.60
PT (s)	12.30(11.90-12.70)	12.30(11.90-12.70)	12.30(11.90-12.78)	0.89
APTT (s)	35.70(33.48-38.30)	35.55(33.40-38.27)	36.10(34.00-38.30)	0.26
TG (mmol/L)	1.22(0.93-1.70)	1.21(0.90-1.66)	1.22(0.95-1.81)	0.32

Notes: CRC: colorectal carcinoma; CRA: colorectal adenoma.

表 4 两组患者治疗前研究指标的单因素Logistic回归分析
Table 4 Univariate Logistic regression analysis of indicators in two groups before treatment

Indicators	CRC (n=329)	CRA (n=77)	OR (95%CI)	P
Gender (n(%))			1.08(0.65-1.78)	0.77
Female	135(41.03)	33(42.86)		
Male	194(58.97)	44(57.14)		
Age (years)	60.09±12.27	56.45±12.43	1.02(1.00-1.04)	0.02
FOBT (%)			14.13(7.64-28.02)	<0.001
Positive	244(74.16)	13(16.88)		
Negative	85(25.84)	64(83.12)		
FIB (g/L)	3.89±0.94	3.06±0.67	4.61(3.00-7.42)	<0.001
TT(s)	16.40±1.25	17.25±1.00	0.48(0.36-0.62)	<0.001
ALB (g/L)			2.13(1.18-3.76)	0.01
<39.30	52(15.80)	22(28.57)		
≥39.30	277(84.20)	55(71.43)		
WBC (10 ⁹ /L)	5.89(4.83-7.17)	5.29(4.50-6.23)	1.28(1.10-1.52)	<0.001
NEUT# (10 ⁹ /L)	3.46(2.65-4.45)	2.92(2.33-3.57)	1.52(1.22-1.93)	<0.001
LYMPH# (10 ⁹ /L)	1.78(1.46-2.18)	1.78(1.43-2.08)	1.06(0.69-1.67)	0.80
MONO# (10 ⁹ /L)	0.35(0.28-0.45)	0.39(0.30-0.47)	0.66(0.20-2.52)	0.49
HCT (%)	39.90(35.70-42.40)	41.00(37.65-43.40)	0.93(0.88-0.97)	<0.001
MCH (pg)	29.80(27.80-31.00)	30.40(29.20-31.50)	0.89(0.81-0.96)	<0.001
PDW (%)	11.40(10.20-12.70)	11.50(10.30-12.90)	0.93(0.83-1.05)	0.29
RDW (%)	13.00(12.40-14.15)	12.70(12.30-13.20)	1.25(1.08-1.52)	<0.001
PLT (10 ⁹ /L)	256.00(213.00-305.00)	229.00(189.50-255.50)	1.01(1.00-1.01)	<0.001
MPV (fl)	10.31±1.10	10.01±0.87	1.32(1.03-1.71)	0.03
PT (s)	12.30(11.90-12.70)	12.30(11.95-12.90)	0.90(0.70-1.19)	0.43
APTT (s)	35.40(33.20-38.10)	36.30(34.50-38.95)	0.94(0.89-1.00)	0.04
TG (mmol/L)	1.19(0.91-1.64)	1.33(0.85-1.87)	0.94(0.75-1.23)	0.62
CHOL (mmol/L)	4.66±0.90	4.81±1.06	0.84(0.65-1.10)	0.21

表 5 多因素Logistic回归分析赋值表
Table 5 Assignment table of multivariate Logistic regression analysis

Variables	Assignment
WBC (10 ⁹ /L)	<6.42=0, ≥6.42=1
NEUT# (10 ⁹ /L)	<3.28=0, ≥3.28=1
HCT (%)	<35.7=0, ≥35.7=1
MCH (pg)	<29.2=0, ≥29.2=1
RDW (%)	<13.6=0, ≥13.6=1
MPV (fl)	<34.2=0, ≥34.2=1
PLT (10 ⁹ /L)	<271=0, ≥271=1
APTT (s)	<34.2=0, ≥34.2=1
FIB (g/L)	<3.01=0, ≥3.01=1
TT (s)	<16.5=0, ≥16.5=1

11.19 (5.05~24.80), $P<0.001$]、TT[OR=7.38 (3.07~17.70), $P<0.001$]的高值以及HCT[OR=0.22 (0.05~0.92), $P=0.039$]的低值是CRA患者患CRC的危险因素 ($P<0.05$)。MPV≥34.2 fl是CRA患者患CRC的危险因素。根据预测变量绘制列线图, 见图2, 得出CRA患者发生CRC的风险率。

2.6 预测模型性能评估

该预测模型的表现指标评估主要基于模型的区分度和校准度。该预测模型在训练集和验证集的ROC曲线下面积 (area under curve, AUC) 分别为

0.915和0.836, 见图3, 表明预测模型具有良好的判别能力, 能够区分CRA和CRC患者。训练集和验证集的校准曲线显示, 该模型的预测风险与实际发病风险一致程度较高, 预测模型的校正能力良好, 见图4。

2.7 预测模型临床应用评估

利用DCA评估预测模型临床应用的有效性, CRA发生CRC风险列线图的DCA, 见图5。结果显示, 当阈值概率在55%~95%区间内时, 预测模型的净收益均显著大于2个极端模型 (全干预或者全不干预), 即可认为模型的预测价值优于2个极端情况。

3 讨论

CRC是一种异质性疾病, 由遗传和环境因素相互作用引起, 随遗传和表观遗传变化的逐渐积累而发展, 最终导致正常结肠黏膜转变为浸润性癌症。大多数CRC病例是由癌前腺瘤和锯齿状息肉引起^[12]。内镜筛查可以降低CRC的发病率和死亡率。对于已行息肉切除的患者, 结肠镜检查的监测可以预防CRC的发生^[13]。然而, 对于有锯齿状息肉的个体, 现有的前体病变监测指南的适用性差异很大, 且缺乏足够的支持证据。一些高风险患者可能没有

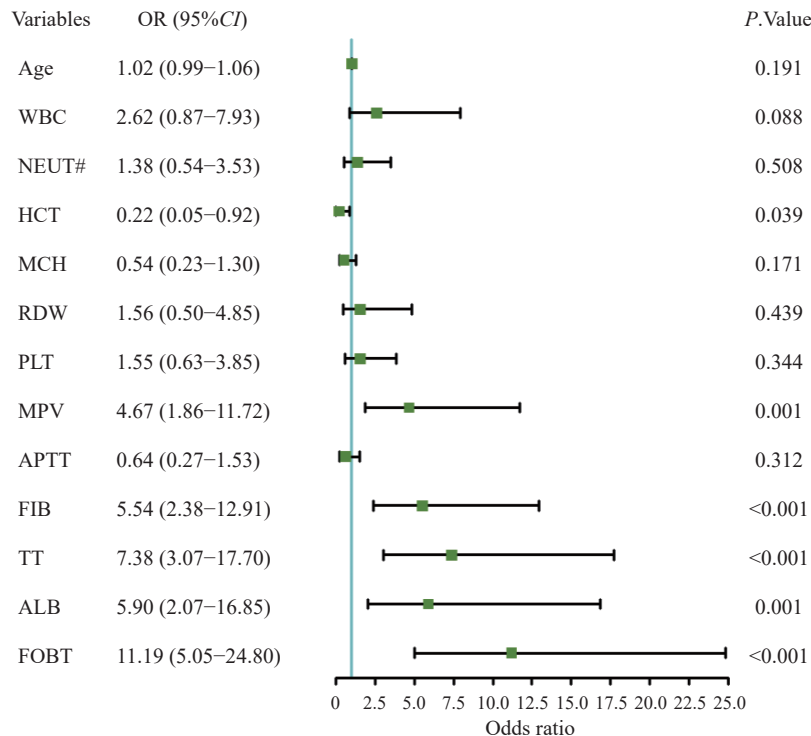


图1 多因素回归分析森林图

Figure 1 Forest plot of multivariate regression analysis

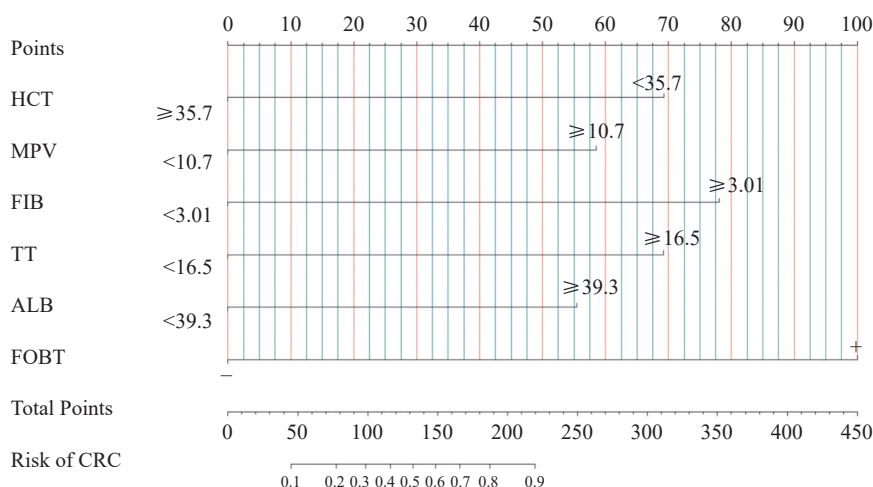


图2 结直肠腺瘤患者发生结直肠癌的预测模型列线图

Figure 2 Nomogram predicting colorectal cancer in patients with colorectal adenoma

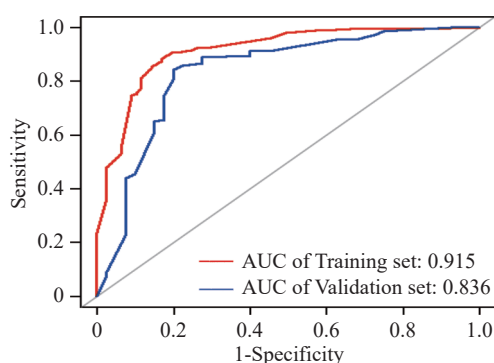


图3 预测模型的ROC曲线

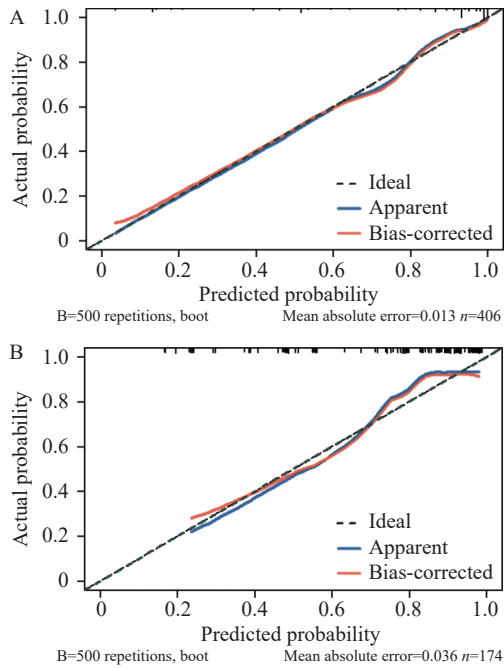
Figure 3 ROC curve of the prediction model

得到足够的监测，而低风险受试者可能会接受过度监测^[13]。因此，对于有结直肠息肉的患者，在资源匮乏或无法进行结肠镜检查的情况下，需要探索

一种便捷、经济且非侵入性的检查方法，给临床医生在早期筛查和诊断上提供更多的支持。

已知的传统CRA和CRC的危险因素包括年龄、男性、家族史、肥胖和缺乏身体活动以及红肉摄入量^[14]。FOBT是一种常见的非侵入性CRC筛查方法，已证实可以降低CRC相关的死亡率，但并不能降低CRC的发病率^[15]。本研究确认了FOBT作为CRA患者发展为CRC的独立危险因素，可见FOBT的阳性结果与CRC的发生存在显著关联（ $OR=11.19$ ， $P<0.001$ ），说明FOBT在识别CRC高风险患者方面具有一定的预测能力。

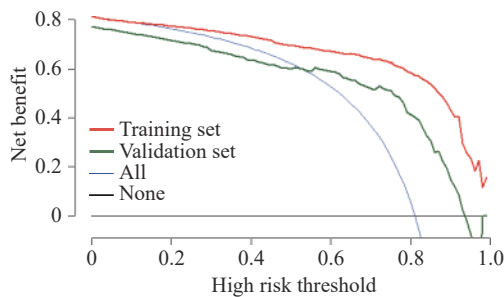
血栓栓塞是CRC、乳腺癌等癌症患者可能面临的一个严重并发症^[16]。红细胞的定量和定性变化以及红细胞与止血系统的细胞和分子成分的相互作用



A: training set; B: validation set.

图4 预测模型的校准曲线

Figure 4 Calibration curves of the prediction model



The horizontal line represents the net benefit without intervention, while the diagonal line represents the net benefit with intervention.

图5 预测模型的DCA分析

Figure 5 Decision curve analysis of the prediction model

会影响出血和血栓形成。其中，影响止血和血栓形成的红细胞变化包括HCT和血液成分的定性变化，例如变形性、聚集、黏附蛋白和磷脂酰丝氨酸的表达、细胞外微泡的释放和溶血^[17]。FIB是最丰富的凝血因子，血凝块形成的最后阶段是将可溶性FIB转化为不溶性纤维蛋白，不溶性纤维蛋白是用于止血或阻塞血管的血凝块的聚合物支架。FIB是一种黏弹性聚合物，纤维蛋白支架的结构和机械性能决定了其在止血中的有效性以及血栓并发症的发展和结果^[18]。越来越多的证据表明，术前FIB与前白蛋白比值（fibrinogen to prealbumin ratio, FPR）、FIB与ALB比值（fibrinogen to albumin ratio, FAR）或ALB与FIB比值（albumin to fibrinogen ratio, AFR）等预测因子正在成为综合指标，ALB、前白蛋白和

FIB在CRC的进展中起着关键性调节作用^[19]。在共同凝血途径中，凝血酶可以使FIB转变为纤维蛋白，因此可用TT来反映凝血功能。本研究也显示，HCT、FIB、TT和ALB均为CRA患CRC的危险因素。

炎症蛋白激活PLT，而PLT与癌症进展和发展直接相关。MPV是关键的血小板参数之一，作为血小板活化的标志物，MPV近几十年来备受关注。许多研究评估了MPV与各种恶性肿瘤的关联，对癌症MPV改变的深入研究可以揭示该参数在癌症诊断、治疗反应和预后中的潜在用途^[20]。目前，血常规指标较少用于CRC的辅助诊断与预测。本研究显示， $MPV \geq 34.2$ fl是CRA患者患CRC的危险因素 $[OR=4.67 (1.86 \sim 11.72), P=0.001]$ 。

列线图是一种广泛用于临床实践的预测工具，通过预测个体的临床结果、呈现相关风险变量和风险评分来改善医患沟通，并促进个性化治疗^[21]。本研究纳入HCT、MPV、FIB、TT、ALB和FOBT等6个危险因素构建列线图模型。该预测模型具有较好的区分度、校准度和临床应用的有效性，具备一定的预测能力，对CRC的早期筛查和干预癌前病变具有一定的指导意义。

尽管本研究对常规检验指标与CRA和CRC之间关系做出了一个初步的探索工具，但仍然存在一些不足之处。首先，该研究是一项单中心回顾性的研究，仍需要多中心研究来评估模型的真实性能。其次，本研究的列线图预测模型目前仅经过了内部验证，需要进行外部验证来进一步验证其准确性和可靠性。最后，该研究模型仅纳入了常规检验指标，未考虑其他与疾病相关的因素，如临床表现、病理检查结果等。未来的研究可以将这些因素纳入模型，以提高预测模型的准确性和应用价值。

利益冲突声明：

所有作者均声明不存在利益冲突。

参考文献：

- [1] Bray F, Ferlay J, Soerjomataram I, *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA Cancer J Clin*, 2018, 68(6): 394-424.
- [2] Brenner H, Kloor M, Pox CP. Colorectal cancer[J]. *Lancet*, 2014, 383(9927): 1490-1502.
- [3] Ladabaum U, Dominitz JA, Kahi C, *et al.* Strategies for colorectal cancer screening[J]. *Gastroenterology*, 2020, 158(2): 418-432.
- [4] Wang B, Wang X, Tseng Y, *et al.* Distinguishing colorectal adenoma from hyperplastic polyp by WNT2 expression[J]. *J Clin*

- Lab Anal, 2021, 35(10): e23961.
- [5] Sawicki T, Ruskowska M. A Review of colorectal cancer in terms of epidemiology, risk factors, development, symptoms and diagnosis[J]. *Cancers (Basel)*, 2021, 13(9): 2025.
- [6] Uraoka T, Hosoe N, Yahagi N. Colonoscopy: Is it as effective as an advanced diagnostic tool for colorectal cancer screening?[J]. *Expert Rev Gastroenterol Hepatol*, 2015, 9(2): 129-132.
- [7] Dekker E, Rex DK. Advances in CRC Prevention: Screening and surveillance[J]. *Gastroenterology*, 2018, 154(7): 1970-1984.
- [8] Robertson DJ, Selby K. Fecal Immunochemical test: The world's colorectal cancer screening test[J]. *Gastrointest Endosc Clin N Am*, 2020, 30(3): 511-526.
- [9] Záhorec R, Marek V, Waczulikova I, *et al.* Predictive model using hemoglobin, albumin, fibrinogen, and neutrophil-to-lymphocyte ratio to distinguish patients with colorectal cancer from those with benign adenoma[J]. *Neoplasma*, 2021, 68(6): 1292-1300.
- [10] Zhu X, Cao Y, Lu P, *et al.* Evaluation of platelet indices as diagnostic biomarkers for colorectal cancer[J]. *Sci Rep*, 2018, 8(1): 11814.
- [11] Hernandez-Ainsa M, Velamazán R, Lanas A, *et al.* Blood-cell-based inflammatory markers as a useful tool for early diagnosis in colorectal cancer[J]. *Front Med (Lausanne)*, 2022, 9: 843074.
- [12] Binefa G, Rodríguez-Moranta F, Teule A, *et al.* Colorectal cancer: From prevention to personalized medicine[J]. *World J Gastroenterol*, 2014, 20(22): 6786-6808.
- [13] He X, Hang D, Wu K, *et al.* Long-term risk of colorectal cancer after removal of conventional adenomas and serrated polyps[J]. *Gastroenterology*, 2020, 158(4): 852-861. e4.
- [14] Sninsky JA, Shore BM, Lupu GV, *et al.* Risk factors for colorectal polyps and cancer[J]. *Gastrointest Endosc Clin N Am*, 2022, 32(2): 195-213.
- [15] Jodal HC, Helsing LM, Anderson JC. Colorectal cancer screening with fecal testing, sigmoidoscopy or colonoscopy: A systematic review and network meta-analysis[J]. *BMJ Open*, 2019, 9(10): e032773.
- [16] Gervaso L, Dave H, Khorana AA. Venous and arterial thromboembolism in patients with cancer: JACC: CardioOncology State-of-the-art review[J]. *JACC CardioOncol*, 2021, 3(2): 173-190.
- [17] Weisel JW, Litvinov RI. Red blood cells: The forgotten player in hemostasis and thrombosis[J]. *J Thromb Haemost*, 2019, 17(2): 271-282.
- [18] Litvinov RI, Pieters M, de Lange-Loots Z, *et al.* Fibrinogen and fibrin[J]. *Subcell Biochem*, 2021, 96: 471-501.
- [19] Chen C, Liu Y, Han P, *et al.* Research progress of preoperative FPR, FAR or AFR in patients with colorectal cancer[J]. *Cancer Manag Res*, 2021, 13: 1791-1801.
- [20] Detopoulou P, Panoutsopoulos GI, Mantoglou M, *et al.* Relation of mean platelet volume (MPV) with cancer: A systematic review with a focus on disease outcome on twelve types of cancer[J]. *Curr Oncol*, 2023, 30(3): 3391-3420.
- [21] Wang X, Lu J, Song Z, *et al.* From past to future: Bibliometric analysis of global research productivity on nomogram (2000-2021)[J]. *Front Public Health*, 2022, 10: 997713.

[编辑: 尤婷婷; 校对: 安凤]

作者贡献:

林俊生: 数据收集分析、论文撰写

应紫灵: 数据收集分析、研究方法设计

黄政渊: 数据收集、论文修改

祝先进: 论文构思、统计指导

曹颖平: 论文审阅

卢娉霞: 论文审阅修改、基金支持