

摘要

背景及目的:

结直肠癌（Colorectal Cancer, CRC）是消化系统中最常见的恶性肿瘤之一。2020年的全球癌症统计数据显示 CRC 的发病率在恶性肿瘤中排名第三，超过 190 万例，死亡率居于第二，超过 90 万例。CRC 是影响人们生命健康的主要癌症之一，造成了严重的健康负担，而早期诊断及治疗被普遍认为是降低 CRC 死亡率的最有效途径。在临床中结直肠息肉往往被认为是 CRC 最常见的癌前病变，所以如果能够早期发现并切除结直肠息肉将有望进一步降低 CRC 的发病率，因此研究结直肠息肉发生的危险因素尤为重要。风险预测模型为当下研究的热点，目前对于 CRC 的模型研究较多，但是关于对结直肠息肉的风险预测模型的建立相对较少，故本研究分析结直肠息肉的危险因素，构建并验证结直肠息肉的风险预测模型，从而提高结直肠息肉的筛查率，对预防 CRC 的发生起到指导作用。

方法:

回顾性分析 2021 年 1 月至 2022 年 10 月期间在河南大学淮河医院消化内科住院且行结肠镜检查的患者的临床资料，根据纳入和排除标准构建数据库，共有 1952 例患者符合标准，其中 978 例检测出结直肠息肉，974 例结肠镜检查未见异常。对纳入的 1952 例患者的基本特征进行描述，并对结直肠息肉的发生进行单因素及多因素分析。根据病理类型结果将结直肠息肉组分为腺瘤型息肉、非腺瘤型息肉，然后对结直肠息肉组的基本特征进行描述，并对腺瘤性息肉的发生行单因素及多因素进行分析。关于结直肠息肉风险模型的建立，使用 R 语言将 1952 例数据以 3: 1 完全随机拆分为建模队列（1464 例）与验证队列（488 例），并通过单因素及多因素回归分析对结直肠息肉的危险因素进行筛选，以此为基础建立结直肠息肉风险预测模型，最终以列线图形式呈现临床预测模型，并以区分度、校准度和决策曲线对临床预测模型进行评价。

结果:

(1) 本研究共纳入 1952 例患者，息肉组与无息肉组病例基线特征进行比较，两组间在年龄、性别、TC、TG、LDL-C、TG/HDL、CEA、BMI、腰围、腰高比、吸烟史、饮酒史以及高血压、糖尿病史存在统计学意义 ($P < 0.05$)，而在 HDL-C、TC/HDL、

LDL/HDL、BA、CRP、ALB、TP、尿酸、教育程度无统计学意义 ($P>0.05$)。将所有因素纳入进行单因素分析,结果表明年龄、性别、TC、TG、LDL-C、TG/HDL、BMI、腰围、腰高比、高血压史、糖尿病史、吸烟史、饮酒史均与结直肠息肉的发生有关 ($P<0.05$),而 HDL-C、TC/HDL、LDL/HDL、BA、CRP、ALB、TP、CEA、尿酸、教育程度无统计学意义 ($P>0.05$);将上述单因素中有意义的指标 ($P<0.05$) 纳入多因素中,显示年龄、TG、腰围、饮酒史、高血压史是结直肠息肉的独立危险因素。

(2) 将 978 例患有结直肠息肉的患者,根据病理类型分为腺瘤性息肉组和非腺瘤性息肉组,两组间对比结果显示在年龄、TC、LDL-C、性别、吸烟史、息肉大小、息肉数目、息肉部位方面有统计学意义 ($P<0.05$),而在 TG、HDL-C、TC/HDL、TG/HDL、LDL/HDL、BA、CRP、ALB、TP、CEA、BMI、腰围、腰高比、饮酒史、高血压史、糖尿病史、高尿酸、教育程度无统计学意义 ($P>0.05$)。单因素分析结果可知高龄、男性、TC、LDL-C、吸烟、息肉数目、息肉大小、息肉的部位与腺瘤性息肉的发生有关 ($P<0.05$),而 TG、HDL、TC/HDL、TG/LDL、LDL/HDL、BA、CRP、ALB、TP、CEA、BMI、腰围、腰高比、饮酒史、高血压史、糖尿病史、高尿酸、教育程度与腺瘤性息肉无相关性 ($P>0.05$)。将上述有意义的指标 ($P<0.05$) 纳入多因素分析,显示男性、TC、息肉大小、息肉数目是腺瘤性息肉的独立危险因素。

(3) 使用 R 语言将 1952 例数据以 3:1 完全随机拆分为建模队列 (1464 例) 与验证队列 (488 例),结果显示建模队列和验证队列的一般临床资料与实验室检查方面无统计学差异 ($P>0.05$),说明两队列之间基线特征一致,均衡性较好。建模队列中息肉组与非息肉组中的年龄、性别、TC、TG、LDL-C、TG/HDL、BA、BMI、腰围、腰高比、吸烟史、饮酒史、高血压史、糖尿病史等方面均有统计学意义 ($P<0.05$),而在 HDL-C、TC/HDL、LDL/HDL、CRP、ALB、TP、CEA、尿酸、教育程度等方面无统计学意义 ($P>0.05$)。建模队列中单因素分析显示年龄、男性、TC、TG、LDL-C、BA、TG/HDL、BMI、腰围、腰高比、吸烟史、饮酒史、高血压史、糖尿病史均与结直肠息肉的发生有关 ($P<0.05$)。多因素分析显示年龄、TG、腰围以及饮酒史均是结直肠息肉的独立危险因素。将多因素中的危险因素通过 R 语言绘制预测结直肠息肉发生的列线图、校准度评价及决策曲线,结果显示建模队列 ROC 曲线下面积 (AUC) 为 0.706,验证队列 ROC 曲线下面积 (AUC) 为 0.711,表示该模型具有一定的预测作用;两组数据

集的校准曲线均贴近于理想曲线，显示模型具有良好的校准能力；决策曲线显示列线图预测模型具有一定的临床应用价值。

结论：

- (1) 年龄、TG、腰围以及饮酒均是结直肠息肉的独立危险因素；
- (2) 男性、TC、息肉的数目、息肉的大小均是腺瘤性息肉的独立危险因素；
- (3) 建立结直肠息肉的临床预测模型，并且以列线图的形式呈现出来，有助于提高结直肠息肉的检出率。

关键词：结直肠息肉，腺瘤性息肉，危险因素，列线图，临床预测模型

中英文对照表

英文全称	英文缩写	中文全称
Area Under Curve	AUC	ROC 曲线下面积
Bile Acid	BA	胆汁酸
Body Mass Index	BMI	身体指数
Carcinoembryonic antigen	CEA	癌胚抗原
Coli Of Adenomatous Polyposis	CAP	腺瘤性息肉
Colorectal cancer	CRC	结直肠癌
High Density Lipoprotein Cholesterol	HDL-C	高密度脂蛋白胆固醇
Low Density Lipoprotein Cholesterol	LDL-C	低密度脂蛋白胆固醇
Non-Coli Of Adenomatous Polyposis	NCAP	非腺瘤性息肉
Odds Ratio	OR	比值比
Receiver operating characteristic	ROC	受试者工作特征
Total Cholesterol	TC	总胆固醇
Triglyceride	TG	甘油三酯
Uric Acid	UA	尿酸
Waist Circumference	WC	腰围
Waist Height Ratio	WHtR	腰高比

目 录

摘 要.....	I
ABSTRACT	V
中英文对照表.....	IX
目 录.....	XI
引 言.....	1
一. 资料与方法.....	3
1.1 研究对象.....	3
1.2 纳入标准和排除标准.....	3
1.2.1 纳入标准.....	3
1.2.2 排除标准.....	3
1.3 临床资料收集.....	3
1.3.1 临床基础信息.....	3
1.3.2 实验室检查.....	4
1.3.3 相关定义及概念.....	4
1.3.4 分组标准.....	4
1.3.5 电子结肠镜检查.....	5
1.4 统计学方法.....	5
二. 研究结果.....	7
2.1 患者一般临床资料.....	7
2.1.1 所有患者的基线特征比较.....	7
2.1.2 息肉患者基线特征比较.....	8
2.1.3 不同年龄段结直肠息肉的特征比较.....	10
2.1.4 不同性别结直肠息肉的特征比较.....	10
2.1.5 不同 BMI 患者结直肠息肉的特征比较.....	11
2.2 结直肠息肉的单因素及多因素 Logistic 回归分析.....	12

2.3 腺瘤性息肉的单因素及多因素 Logistic 回归分析	13
2.4 结直肠息肉的风险预测模型的建立与验证	14
2.4.1 建模队列与验证队列临床资料比较	14
2.4.2 建模队列一般基线特征	15
2.4.3 建模队列中纳入因素的相关性分析	17
2.4.4 结直肠息肉组的单因素及多因素回归分析	17
2.4.5 回归方程的构建	19
2.4.6 绘制结直肠息肉列线图模型	19
2.4.7 结直肠息肉的模型验证	20
(1) 区分度的评价	20
(2) 校准度评价	20
(3) 决策曲线分析	21
三. 讨论	23
四. 结论	29
参考文献	31
综述	35
参考文献	45
致谢	51

引言

最新的全球癌症统计报告^[1]显示，结直肠癌（Colorectal Cancer, CRC）发病率在恶性肿瘤中排名第三，死亡率居于第二；而且 CRC 被视为社会经济发展的标志，发病率往往随着人类发展指数的增加而上升^[2]。CRC 的发展主要通过“腺瘤-腺癌”、“锯齿状病变-癌症”和“炎症-癌症”等三种途径^[3]，其中在临床中最常见的是 Morson 学者提出“腺瘤-腺癌”途径学说^[4]。尽管息肉和癌症之间有着众所周知的联系，但据估计，息肉至少需要 10 年的时间才能长成癌症^[5]。因此在结直肠息肉进展为 CRC 之前，早期对结直肠息肉的识别并切除，有望减少 CRC 的发生率。

结直肠息肉是结直肠粘膜表面隆起并突起到肠腔内的病变的总称；根据胃肠道肿瘤病理分类的标准，将结直肠息肉分为肿瘤性息肉和非肿瘤性息肉^[6]。肿瘤性息肉与非肿瘤性的主要区别特征在于肿瘤性息肉有发生恶性转化的潜在风险，肿瘤性息肉约占所有结肠息肉的 2/3，是 CRC 最常见的癌前病变。目前随着人民生活水平的提高以及我国消化内镜技术与诊疗快速发展，结肠镜的检查得到不断的改善和提高，使得结直肠息肉的检出率也随之增加，故研究结直肠息肉发生的危险因素是当下的热点，所以如果能在临床中明确结直肠息肉的发生的危险因素，那么能够大幅度预防 CRC 的发生。其中列线图模型是一种直观的预测风险模型的评分量表，它使用多个变量为疾病的发作或进展创建可视化的个体预后，具有直观且易于解释的优点。

风险预测模型为当下研究的热点，目前对于 CRC 的模型研究较多^[7-9]，其中张琴通过回顾性研究北京大学附属医院 2011 至 2015 年行结肠镜体检的无症状患者的临床资料，多因素结果表明性别、年龄、BMI、吸烟、糖尿病是 CRC 的独立危险因素，ROC 曲线下面积为 0.783，并通过验证组验证得出风险预测模型具有一定的临床使用价值。但是目前国内外文献关于结直肠息肉的风险预测模型却很少，其中中国学者邢洁^[10]通过回顾性分析 2013 年至 2014 年在首都医科大学就诊患者的基本情况、个人病史、性别、年龄、BMI、肠道清洁度等信息，与原发高血压患者结直肠息肉发生的相关危险因素进行统计分析，多因素回归分析显示年龄、饮酒是高血压患者发生结直肠息肉的独立危险因素，并对其建立预测模型显示 ROC 曲线下面积为 0.684。由于结直肠息肉风险预测模型的研究较少，因此本研究拟通过收集、归纳和处理结直肠息肉患者的一般临床资料和生化参

数，建立用于结直肠息肉疾病预测的列线图风险模型及评价其预测价值，利用现有指标对临床数据进行统计分析，依据临床数据对结直肠息肉疾病进行前瞻性预测，并对相关人群进行分层，从而实现高危人群的筛查和针对性的结肠镜检查。

一. 资料与方法

1.1 研究对象

选自 2021 年 1 月至 2022 年 10 月在淮河医院消化内科住院且行结肠镜检查的 1952 例患者为研究对象，系统回顾这些患者临床资料，根据纳入标准和排除标准，经过严格筛选发现结肠镜检查可见结直肠息肉的 978 例患者为病例组，同期结肠镜检查未见异常的 974 例患者为对照组。这些患者签署结肠镜检查及治疗的知情同意书，且试验符合淮河医院伦理学要求。

1.2 纳入标准和排除标准

1.2.1 纳入标准

- (1) 临床资料完整的患者。
- (2) 年龄 ≥ 18 岁。
- (3) 肠道准备波士顿量表 (Boston) 评分 > 6 分。
- (4) 行结直肠息肉切除术者或者电子结肠镜检查未见异常者。

1.2.2 排除标准

- (1) 临床资料不完整的患者。
- (2) 电子结肠镜检查未到达回盲部。
- (3) 电子结肠镜检查提示为结直肠癌患者。
- (4) 年龄 < 18 岁。
- (5) 既往或本次结肠镜检查除结直肠息肉以外的其他肠道疾病患者。
- (6) 肠道准备不充分，Boston 评分 < 6 分，影响结肠镜进境观察的患者。

1.3 临床资料收集

1.3.1 临床基础信息

从医院数据库中收集消化内科病区受试者的临床基础信息，具体信息如下：年龄、

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/865121224112012010>