

## 结直肠癌筛查方法研究进展\*

邱智宇 综述, 李济宾<sup>△</sup> 审校

510060 广州, 中山大学肿瘤防治中心 肝脏外科(邱智宇), 临床研究部(李济宾)

[摘要] 结直肠癌筛查可以有效降低人群的发病率和死亡率, 实现早诊早治, 从而改善患者的预后。目前, 各国主要采用以粪便潜血试验结合肠镜检查为代表的序贯筛查方法, 但这两种方法存在一定的局限。伴随着大数据挖掘技术在肿瘤筛查领域的应用, 近年涌现出许多新的结直肠癌筛查预测模型、筛查技术和潜在标志物, 有望提升结直肠癌的筛查准确性并降低筛查成本。本文综述了近年结直肠癌筛查领域的研究进展及应用情况。

[关键词] 结直肠癌; 筛查策略; 数据挖掘; 预测模型

[中图分类号] R735.3; R730.4 [文献标志码] A doi:10.3969/j.issn.1674-0904.2019.11.013

引文格式: Qiu ZY, Li JB. Research progress of colorectal cancer screening strategies [J]. J Cancer Control Treat, 2019, 32(11): 1024-1030. [邱智宇, 李济宾. 结直肠癌筛查方法研究进展[J]. 肿瘤预防与治疗, 2019, 32(11): 1024-1030.]

### Research Progress of Colorectal Cancer Screening Strategies

Qiu Zhiyu, Li Jibin

Department of Hepatic Surgery, Sun Yat-sen University Cancer Center, Guangzhou 510060, Guangdong, China (Qiu Zhiyu); Department of Clinical Research, Sun Yat-sen University Cancer Center, Guangzhou 510060, Guangdong, China (Li Jibin).

Corresponding author: Li Jibin, E-mail: lijibin3@mail.sysu.edu.cn

This study was supported by Science and Technology Project of Guangzhou Province (NO. 201803040019).

[Abstract] Colorectal cancer (CRC) screening could effectively reduce the incidence and mortality, achieve early diagnosis and treatment, and improve the prognosis of patients. Currently, many countries mainly apply a sequential screening strategy represented by fecal occult blood test in combination with colonoscopy. However, these two methods have inherent limitations in clinical application. With the development of data mining technology in the field of cancer screening, several new CRC risk prediction models, screening techniques and potential biomarkers have emerged in recent years, which would further improve the accuracy of CRC screening and reduce cost. This article reviews the research progress and application related to CRC screening in recent years.

[Key words] Colorectal cancer; Screening strategy; Data mining; Predictive model

结直肠癌是常见的消化系统恶性肿瘤。2018年国际癌症研究组织报道, 全球新发结直肠癌患者约180万例, 死亡人数约86.1万<sup>[1]</sup>。我国是结直肠癌高发国家, 发病率和死亡率分别位居所有癌种第4位和第5位<sup>[2]</sup>。因此, 结直肠癌的筛查与防治已成为我国当前面临的重大公共卫生问题。

近20年来, 得益于结肠镜筛查的普及与避免高危因素的暴露, 发达国家结直肠癌发病率和死亡率呈显著下降趋势<sup>[3]</sup>。但对于发展中国家而言, 由于肠镜筛查成本高、专业医生短缺等原因, 全面开展肠镜筛查并不现实。以我国上海为例, 在结直肠癌初筛阳性的受检者中, 仅有47.01%选择接受进一步的肠镜检查<sup>[4]</sup>。近几年, 医学研究者针对结直肠癌筛查、肿瘤标志物、风险预测模型及数据挖掘方面开展了系列的探索性研究, 本文就现行筛查方法与策略、筛查预测模型、肿瘤标志物组合的研究进展, 以

[收稿日期] 2019-04-22 [修回日期] 2019-09-12

[基金项目] \*广州市科技计划项目(编号:201803040019)

[通讯作者] <sup>△</sup>李济宾, E-mail: lijibin3@mail.sysu.edu.cn

及联合数据挖掘技术的应用前景进行系统综述。

### 1 现行结直肠癌筛查方法

目前,国内外常用的结直肠癌筛查策略以粪便潜血试验(fecal occult blood test, FOBT)为初筛方法,辅以结肠镜或 CT 等检查手段。为减少患者的负担,指南一般推荐采用序贯筛查方法,即在愈创木

脂潜血试验(guaiac FOBT, gFOBT)或粪便免疫化学实验(fecal immunochemical test, FIT)等测试结果阳性的基础上,建议受检查者进行内镜检查。近年来,随着研究的进一步深入,多种基于崭新思路的检验方法陆续出现,部分方法在国外甚至已获批进入临床使用,详见表 1。

表 1 现行结直肠癌筛查方法一览

Table 1. Current Screening Methods for CRC

Screening method	Advantage	Disadvantage	Efficacy <sup>[5]</sup>
High risk factor assessment questionnaire <sup>[6]</sup>	Simple and easy accessibility; Wide coverage	The positive predictive value is low; It is not recommended to use it alone.	Not reported
FOBT(gFBOT and FIT)	Easy to conduct; Suitable for countries with insufficient colonoscopy equipment and low acceptance of endoscopy	Easily affected by diet and other factors, leading to false negative/ positive results	32% lower modality of CRC
Endoscopy			
Colonoscopy <sup>[7]</sup>	Gold standard which can be used to directly observe the whole process of colon and rectum and remove polyps for pathological examinations	Bowel prep is needed; Risk of intestinal perforation; Low acceptance	68% lower modality of CRC
Flexible sigmoidoscopy <sup>[8]</sup>	Easier to conduct, shorter time, lower risk and higher acceptance; Can be used as a substitute to some extent when performing the whole course of colonoscopy is impossible	Can only be used for detecting the distal part of spleen flexure; Lesions in the proximal colon are easily missed; Lower efficacy than colonoscopy	26% - 31% lower modality of CRC
Capsule endoscopy <sup>[9]</sup>	Fewer complications (such as intestinal perforation) than colonoscopy and flexible sigmoidoscopy	Short working time of battery; Can only reach the ileocecum; Higher cost and lower sensitivity than colonoscopy	Not reported
CT colonography <sup>[8]</sup>	Suitable for people who cannot tolerate endoscopy	Lesions cannot be resected; Radioactive exposure; Risk of extra colon lesions, causing additional medical expenses	Similar to endoscopy for polyps ≥ 10 mm; Inferior to endoscopy for polyps < 10 mm
Fecal DNA test (FIT-DNA) <sup>[10]</sup>	Higher sensitivity than FIT	Lower specificity than FIT	Not reported
Circulating DNA test (SEPTIN9) <sup>[11]</sup>	Easy to conduct; Only blood samples are needed	Lower sensitivity than FOBT, FIT and colonoscopy	Not reported

CRC: Colorectal cancer; FOBT: Fecal occult blood test; gFOBT: Guaiac fecal occult blood test; FIT: Fecal immunochemical test.

### 2 结直肠癌筛查策略

结直肠癌筛查是一个完整的鉴别诊断过程,包括筛查对象、筛查年龄、筛查方法和筛查间隔等。除筛查方法外,世界各国在其余几个方面所实施的具体策略亦不近相同。

#### 2.1 筛查对象

以美国的结直肠癌筛查指南(US preventive services task force recommendation statement, USPSTF)为代表,目前达成国际共识的筛查对象是处于一般风险下的无症状人群;对于有结直肠癌或增加结直肠癌患病风险的遗传病(比如 Lynch 综合征或家族性多发结肠息肉)家族史,以及有炎症性肠炎、

腺瘤样息肉、结直肠癌等既往史的高危人群,则需要考虑提前接受筛查<sup>[12]</sup>。

#### 2.2 筛查年龄

对于一般风险人群而言,包括我国在内的绝大多数国家都选择了 50 ~ 74 岁或 75 岁作为筛查年龄。如果将筛查起始年龄提前至 45 岁,可增加每一千人的获得生命年(life-years gained, LYG; 衡量筛查收益的常用指标),且能将结肠镜的筛查间隔从 10 年延长至 15 年,从而部分抵消每一千人的终身结肠镜检查次数(lifetime number of colonoscopies, No. CSY; 衡量筛查负担的常用指标),使 LYG 和 No. CSY 达到更好的平衡。但这种策略所增加的 LYG 其实并不可观,且会增加个人的卫生经济负担以及

内镜检查带来的风险,故没有被 USPSTF 推荐。虽然 76~85 岁人群接受筛查的收益大为降低,但个体若合并其他疾病,可在医生建议下继续接受筛查;对于 85 岁以上人群,尚无足够证据表明他们能从筛查中获益。对于高危人群而言,筛查年龄可一定程度地提前:如果待检查者有 1 位在 60 岁前患结肠癌的一级亲属,或 2 位以上在任意年龄段患结肠癌的一级亲属,起始年龄应提前至 40 岁或亲属患癌年龄的 10 年以前<sup>[12]</sup>。先前未接受过筛查的高危人群在 80 岁后仍能从筛查中获益<sup>[13]</sup>。

### 2.3 筛查间隔

USPSTF 指出,以下 4 种不同间隔的筛查策略在增加 LYG 方面无显著差别:(1)每年 1 次 FIT;(2)每年 1 次 FIT 联合 10 年 1 次 FS(已有研究表明 FS 联合 FIT 在减少死亡率方面优于单用 FS<sup>[14]</sup>);(3)

Table 2. Common Predictive Models at Home and Abroad

Model	Countries/regions	Factors included	Predictive value (AUC/C statistics)
Ma <sup>[16]</sup>	Japan	Age, BMI, physical activity, smoking, alcohol	C statistics; 0.64
APCS <sup>[17]</sup>	Asian-Pacific countries and regions	Age, gender, family history of CRC, smoking	Not reported
Adelstein <sup>[18]</sup>	Australia	Age, symptom (bloody stool, mucus stool and fatigue)	AUC: 0.83 (CRC); 0.70 (advanced adenoma); 0.67 (non advanced adenoma)
Wang <sup>[19]</sup>	China	16 single nucleotide polymorphism sites	AUC: 0.72
Tao <sup>[20]</sup>	Germany	Age, gender, smoking, history of CRC in first-degree relatives, alcohol, history of polyps, red meat consumption, history of NSAIDs or colonoscopy	AUC: 0.68
Shin <sup>[21]</sup>	Korea	Male: age, height, BMI, fasting blood glucose, serum total cholesterol, family history of cancer, alcohol consumption, meat consumption; Female: age, height, family history of cancer, fasting blood glucose, meat consumption	C statistics; male 0.78; female 0.73
Li <sup>[22]</sup>	China	Age, gender, smoking, alcohol, 7 single nucleotide polymorphism sites	AUC: 0.59

AUC: Area under the curve; BMI: Body mass index; APCS: Asia-Pacific Colorectal Screening; CRC: Colorectal cancer; NSAIDs: Nonsteroidal anti-inflammatory drugs.

研究者期望将优秀的筛查模型用于结肠癌的筛查策略中,从而提高现行筛查策略的卫生经济学效果。有研究指出,结合流行病学危险因素的预测模型能进一步提高“gFOBT/FIT—结肠镜序贯筛查策略”的筛查能力<sup>[23-24]</sup>,但目前已有的模型并未纳入足够的结肠癌预警症状。一项关于预警症状组合的研究表明,阳性预测值最高的组合为贫血与腹壁紧张(>10%),其它表现较优的组合还有腹痛与贫血(6.9%),以及腹壁紧张和体重下降(7.4%)<sup>[25]</sup>,但这些症状并未被纳入现有的预测模型。此外,充分纳入遗传因素,也有望进一步提升现

10 年 1 次结肠镜检查;(4)5 年 1 次 CT 检查。对于新近出现的 FIT-DNA 联合法,指南推荐的间隔为 3 年 1 次,但其筛查效果仍存在争议,尚未明确相较于上述 4 种策略而言是否具有可比性<sup>[12]</sup>。

### 3 结直肠癌风险预测模型

流行病学研究已找到多个与结直肠癌发生相关的因素,除社会人口学因素外,主要可分为遗传因素(结直肠癌及相关疾病的家族史或既往史、基因位点等)和环境因素(饮食习惯、吸烟、基础疾病等)两类<sup>[15]</sup>。在此基础上,国内外的研究者建立了多种包含不同类型因素的风险预测模型。表 2 列举了主要的肠癌风险预测模型,需要注意的是,基于风险预测模型的筛查策略目前仍处在研究阶段,多数模型并未在临床得到验证。

有模型的预测能力。

### 4 肿瘤标志物

肿瘤标志物是另一类备受关注的筛查方法。目前,研究者已在血液、粪便、尿液等标本中发现了多种具有潜在筛查价值的结直肠癌标志物。

#### 4.1 血液标志物

4.1.1 DNA DNA 类肿瘤标志物可以进一步分为三个亚类:(1)细胞游离 DNA (cell free DNA, cfDNA),即由于疾病、治疗等原因从细胞中释放到外周血中的 DNA,通常不包括循环肿瘤 DNA (circulating

tumor DNA, ctDNA)。Gallardo-Gómez 等<sup>[26]</sup>借助微阵列技术比较结直肠癌患者与健康人群的 cfDNA 后,发现二者的甲基化位点存在显著差异; Molparia 等<sup>[27]</sup>则直接从拷贝数变异(copy number variants, CNV)入手,发现 cfDNA 中 >100mb 的 CNV 同样能预测结直肠癌的发生;(2)异常甲基化 DNA,其中最具代表性的便是甲基化 *septin9* (*SEPT9*) 基因,目前已有多款筛查工具(如 Epi's proColon<sup>®</sup>、ColoVantage<sup>®</sup>等)投入临床应用。除 *SEPT9* 外,其他正在研究中的甲基化 DNA 还有 runt 转录相关因子-3、速激肽-1、眼缺失蛋白同源物 4 和生长抑素等,但均停留于试验阶段<sup>[28]</sup>;(3)ctDNA,ctDNA 与 cfDNA 一同被视作“液体活检”技术的代表,当前研究的主要靶点包括 *KRAS*、*APC* 和 *TP53* 等基因<sup>[29]</sup>。2018 年的一项研究表明,甲基化 *APC2* 有望成为中国结直肠癌人群的新标志物<sup>[30]</sup>。但 ctDNA 在临床应用中仍面临着假阳/阴性率高、缺乏标准等问题,其用于结直肠癌筛查的效果具体如何,还有待更深入的研究。

表 3 新近出现的结直肠癌蛋白质标志物组合一览

Table 3. New Protein Marker Combinations of CRC

Biomarker combination	Sensitivity	Specificity	AUC
GDF15, amphiregulin, FasL, Flt3L, TP53 <sup>[33]</sup>	CRC: 56.4%; Advanced adenoma: 22.0%	90%	CRC: 0.82; Advanced adenoma: 0.60
BAG4, IL6ST, VWF, EGFR, CD44 <sup>[34]</sup>	62.3% - 73.3%	90%	0.79 - 0.87
Combination of nucleosomes <sup>[35]</sup>	Clinical stage I: 72%; Clinical stage II: 86%	90%	0.97
Antitumor antibodies; anti-TP53, anti-IMPDH2, anti-MDM2, anti-MAGEA4 <sup>[36]</sup>	26%	90%	Not reported
CEA, GDF15, anti-p53 <sup>[37]</sup>	54%	98%	Not reported

CRC: Colorectal cancer; AUC: Area under the curve; VWF: von Willebrand factor; EGFR: Epidermal growth factor receptor; CEA: Carcinoembryonic antigen.

#### 4.2 粪便标志物

结直肠癌患者的粪便中可检测到从病灶脱落的肿瘤细胞,而它们所携带的突变基因(突变 *KRAS*、*APC*、*TP53*、微卫星不稳定性相关基因等)亦可作为肿瘤筛查的目标靶点<sup>[38]</sup>。表 1 提及的粪便基因检测工具,便是最新一代的多靶点粪便筛查工具 Colo-guard<sup>®</sup>,其检测的基因位点包括突变 *KRAS*、*BMP3* 和 *NDGR4*。此外亦有报道,联合甲基化 *TWIST1* 检测和 FIT 对进展期腺瘤筛查的敏感度可上升至 82.4%<sup>[39]</sup>,表明这项技术的筛查效果还有更进一步的提升空间。

肠道菌群也是影响结直肠癌变的重要因素之一。一项 2017 年的研究表明,微小单胞菌(*Parvimonas micra*)和抗口臭致病菌(*Solobacterium moorei*)这两种细菌与结直肠癌的发生存在显著的相关性。

4.1.2 非编码 RNA 结直肠癌细胞中存在着多种调控癌细胞生物学行为的非编码 RNA(non-coding RNA, ncRNA)。迄今为止,国内外已发现了超过 50 种能作为筛查标志物的 ncRNA 或组合;然而,在较大样本(人数 ≥100)的试验中表现出不逊色于 FIT 的筛查效果(敏感度 ≥79%, 特异度 ≥94%)的 ncRNA 仅有 miR-139-3p 一种<sup>[31]</sup>。最近 Mai 等<sup>[32]</sup>又揭示出,一种与 Piwi 蛋白相互作用的 pi-RNA54365 在结直肠癌组织中呈现出高表达。这种可能促进癌细胞增殖的小分子 RNA 不仅能在患者血清中得以检测,且表达水平与患者的预后呈负相关。

4.1.3 可溶性蛋白 可溶性蛋白种的代表性分子有癌胚抗原、CA19-9、胰岛素样生长因子结合蛋白 2、C 反应蛋白、Dickkopf3、PKM2 和甲基化 FOXE1 等。由于目前并没有找到一个结直肠癌专属的可溶性蛋白分子标志物,故新的研究思路是将多种蛋白质标志物联合用于肿瘤的筛查,继而产生了各类标志物组合(表 3)。

经过进一步的外部验证,研究者筛选出了 4 种在不同国家结直肠癌患者的粪便标本中稳定存在的微生物基因。这些结果提示,特定菌群及相关的基因有望成为新一类的结直肠癌筛查标志物<sup>[40]</sup>。

#### 4.3 尿液标志物

尿液具有易于采集、较少受到蛋白质与微生物干扰、含有代谢产物等优点,故同样被视作是寻找结直肠癌标志物的理想标本。迄今为止报道的尿液标志物主要包含:(1)尿蛋白,如组织因子、Cyr61、TFF3 等;(2)尿 DNA,如甲基化 *VIM*、甲基化 *WIF1*、甲基化 *NDGR4* 等;(3)代谢产物及其组合,如二乙酰精胺和犬尿氨酸等;(4)前列腺素与活性氧<sup>[41]</sup>。尽管部分尿液标志物已在试验中表现出优异的筛查效果,但目前的临床证据尚不足以支持它们投入临床应用。

## 5 数据挖掘在肿瘤筛查中的潜在价值

近年来,不少思路开阔的研究者将目光投向数据挖掘,期望能将这一项新兴技术应用于临床工作,为临床决策提供更有力的支持。数据挖掘是一种以统计分析为基础,从大量的模糊信息中寻找出有价值的知识的技术,包括人工神经网络、Bayes 判别分

表 4 基于数据挖掘技术建立的结直肠癌标志物组合一览

Table 4. Marker Combinations of Colorectal Cancer Based on Data Mining

Biomarker combination	Data mining method	Specificity	Sensitivity	Positive predictive rate/accuracy
CEA, CA19-9, CA24-2, CY21-1, CA72-4 <sup>[42]</sup>	Neural networks	95%	83%	Positive predictive rate: 95%
CEA, AFP, CA50, CA12-5, CA19-9, CA24-2, CA72-4 <sup>[43]</sup>	Neural networks	96.4%	91.4%	Accuracy: 93.9%
CEA, CA72-4, 4 chosen protein biomarkers <sup>[44]</sup>	Neural networks	96.7%	92.3%	Not reported
CEA, CA19-9, CA72-4, CY21-1 <sup>[45]</sup>	Neural networks	87.01%	80.03%	Accuracy: 81.77%
CEA, CA72-4, CA24-2, CA15-3, HSP60 <sup>[46]</sup>	Machine learning	No data	No data	Accuracy: 62.5% - 91.3%

CEA: Carcinoembryonic antigen; AFP: Alpha-fetoprotein.

类似于蛋白标志物组合,孟潘庆等<sup>[47]</sup>结合神经网络和组织芯片技术建立了一种原位杂交诊断模型(包含 ST13、Bcl-2、Survivin 与 HSF1 mRNA 四种指标)。刘晶等<sup>[48]</sup>则将决策树分类方法运用于激光诱导自体荧光技术,开发出一套可辨别癌变组织的早癌诊断系统。总而言之,数据挖掘技术在结直肠癌筛查方面的应用才刚刚起步。如果将这一技术的应用扩大到对各种筛查方法进行整合与训练,建立一个更完备的筛查模型,有望革新现行的筛查策略。

## 6 讨论

我国的结直肠癌筛查工作尚属起步阶段,但国家对肿瘤筛查高度重视,2005 年启动了包括结直肠癌在内的农村癌症早诊早治项目,涵盖全国的 10 余省份;2012 年启动了国家重大公共卫生服务项目——城市癌症早诊早治项目,涉及我国城市高发的五大肿瘤(肺癌、乳腺癌、肝癌、上消化道癌和结直肠癌),扩展至全国 16 省份的 20 余城市。目前,我国尚未建立起全国性的筛查系统,急需一套易于操作且符合卫生经济学的筛查策略。现行的筛查策略采用高危因素风险评估问卷或 FOBT 作为初筛方法,但前者没有纳入足够的危险因素(包括流行病学因素、预警症状和遗传因素等),筛查效能较低;而后者针对的主要是一般风险人群,难以指引高危人群接受进一步的检查。因此,我国的结直肠癌筛查策略在提高受检者依从性和平衡收益与负担方面仍有较大的改善空间。风险评分、肿瘤标志物的研究进展,以及数据挖掘技术的跨学科应用,为新策略

析和决策树等方法。迄今为止,数据挖掘已在肿瘤诊治这一领域取得了可观的成绩。表 4 汇总了应用数据挖掘技术建立的结直肠癌标志物组合模型。综合来看,标志物组合的筛查效果较单个蛋白标志物表现优异。可以设想,如果在现行的筛查策略中纳入标志物组合,有望进一步提升筛查的准确性。

的制定提供了更广阔的思路。因此,充分利用医学大数据、数据挖掘和人工智能技术,建立更加精确的肿瘤筛查预测模型,预期可以提高结直肠癌筛查的早诊率,获得更高的成本效益。

**作者声明:**本文全部作者对于研究和撰写的论文出现的不端行为承担相应责任;并承诺论文中涉及的原始图片、数据资料等已按照有关规定保存,可接受核查。

**学术不端:**本文在初审、返修及出版前均通过中国知网(CNKI)科技期刊学术不端文献检测系统的学术不端检测。

**同行评议:**经同行专家双盲外审,达到刊发要求。

**利益冲突:**所有作者均声明不存在利益冲突。

**文章版权:**本文出版前已与全体作者签署了论文授权书等协议。

### [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] 孙可欣, 郑荣寿, 张思维, 等. 2015 年中国分地区恶性肿瘤发病和死亡分析[J]. 中国肿瘤, 2019, 28(1): 1-11.
- [3] Siegel R, Desantis C, Jemal A. Colorectal cancer statistics, 2014[J]. CA Cancer J Clin, 2014, 64(2): 104-117.
- [4] 彭慧, 王巧燕, 黄芳, 等. 大肠癌筛查肠镜检查顺应性及影响因素研究[J]. 肿瘤预防与治疗, 2018, 31(5): 340-346.

- [5] Imperiale TF, Ransohoff DF. Screening for colorectal neoplasia [J]. *N Engl J Med*, 2017, 376(16):1598-1599.
- [6] 李娟, 李其龙, 薛峰, 等. 粪便潜血试验和问卷调查在人群结直肠癌筛查中的应用评价[J]. *中国肿瘤*, 2015, 24(5):385-389.
- [7] Quintero E, Castells A, Bujanda L, et al. Colonoscopy versus fecal immunochemical testing in colorectal-cancer screening[J]. *N Engl J Med*, 2012, 366(8):697-706.
- [8] Bray C, Bell LN, Liang H, et al. Colorectal Cancer Screening [J]. *WMJ*, 2017, 116(1):27-33.
- [9] van Gossum A, Munoz-Navas M, Fernandez-Urien I, et al. Capsule endoscopy versus colonoscopy for the detection of polyps and cancer[J]. *N Engl J Med*, 2009, 361(3):264-270.
- [10] Souverijn JH. Multitarget stool DNA testing for colorectal-cancer screening[J]. *N Engl J Med*, 2014, 370(14):1287-1297.
- [11] Church TR, Wandell M, Lofton-Day C, et al. Prospective evaluation of methylated SEPT9 in plasma for detection of asymptomatic colorectal cancer[J]. *Gut*, 2014, 63(2):317-325.
- [12] Bibbins-Domingo K, Grossman DC, Curry SJ, et al. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement[J]. *JAMA*, 2016, 315(23):2564-2575.
- [13] van Hees F, Saini SD, Lansdorp-Vogelaar I, et al. Personalizing colonoscopy screening for elderly individuals based on screening history, cancer risk, and comorbidity status could increase cost effectiveness[J]. *Gastroenterology*, 2015, 149(6):1425-1437.
- [14] Holme O, Loberg M, Kalager M, et al. Effect of flexible sigmoidoscopy screening on colorectal cancer incidence and mortality: a randomized clinical trial[J]. *JAMA*, 2014, 312(6):606-615.
- [15] 郑晓金, 李凯, 欧凤荣. 结直肠癌相关危险因素研究进展[J]. *实用药物与临床*, 2018, 21(10):1196-1199.
- [16] Ma E, Sasazuki S, Iwasaki M, et al. 10-Year risk of colorectal cancer: development and validation of a prediction model in middle-aged Japanese men[J]. *Cancer Epidemiol*, 2010, 34(5):534-541.
- [17] Yeoh KG, Ho KY, Chiu HM, et al. The Asia-Pacific Colorectal Screening score: a validated tool that stratifies risk for colorectal advanced neoplasia in asymptomatic Asian subjects [J]. *Gut*, 2011, 60(9):1236-1241.
- [18] Adelstein BA, Macaskill P, Turner RM, et al. The value of age and medical history for predicting colorectal cancer and adenomas in people referred for colonoscopy[J]. *BMC Gastroenterol*, 2011, 11(1):97-106.
- [19] Wang HM, Chang TH, Lin FM, et al. A new method for post Genome-Wide Association Study (GWAS) analysis of colorectal cancer in Taiwan[J]. *Gene*, 2013, 518(1):107-113.
- [20] Tao S, Hoffmeister M, Brenner H. Development and validation of a scoring system to identify individuals at high risk for advanced colorectal neoplasms who should undergo colonoscopy screening [J]. *Clin Gastroenterol Hepatol*, 2014, 12(3):478-485.
- [21] Shin A, Joo J, Yang HR, et al. Risk prediction model for colorectal cancer: National Health Insurance Corporation study, Korea [J]. *PLoS One*, 2014, 9(2):e88079.
- [22] 李娇元, 常江, 朱颖, 等. 基于常见遗传变异和传统风险因素的中国南方汉族人群结直肠癌风险预测模型研究[J]. *中华流行病学杂志*, 2015, 36(10):1053-1057.
- [23] Marshall T, Lancashire R, Sharp D, et al. The diagnostic performance of scoring systems to identify symptomatic colorectal cancer compared to current referral guidance[J]. *Gut*, 2011, 60(9):1242-1248.
- [24] Jung YS, Park CH, Kim NH, et al. A combination of clinical risk stratification and fecal immunochemical test is useful for identifying persons with high priority of early colonoscopy[J]. *Dig Liver Dis*, 2018, 50(3):254-259.
- [25] Hamilton W. The CAPER studies: five case-control studies aimed at identifying and quantifying the risk of cancer in symptomatic primary care patients[J]. *Br J Cancer*, 2009, 101(S2):80-86.
- [26] Gallardo-Gómez M, Moran S, Páez de la Cadena M, et al. A new approach to epigenome-wide discovery of non-invasive methylation biomarkers for colorectal cancer screening in circulating cell-free DNA using pooled samples[J]. *Clin Epigenetics*, 2018, 10(1):53-62.
- [27] Molparia B, Oliveira G, Wagner JL, et al. A feasibility study of colorectal cancer diagnosis via circulating tumor DNA derived CNV detection[J]. *PLoS One*, 2018, 13(5):e0196826.
- [28] Ma Z, Williams M, Cheng YY, et al. Roles of methylated DNA biomarkers in patients with colorectal cancer[J]. *Dis Markers*, 2019:2673543.
- [29] Vymetalkova V, Cervena K, Bartu L, et al. Circulating cell-free DNA and colorectal cancer: a systematic review[J]. *Int J Mol Sci*, 2018, 19(11):3356.
- [30] He Y, Sun LY, Wang J, et al. Hypermethylation of APC2 is a predictive epigenetic biomarker for Chinese colorectal cancer[J]. *Dis Markers*, 2018:8619462.
- [31] Wang JJ, Wang X, Song YX, et al. Circulating noncoding RNAs have a promising future acting as novel biomarkers for colorectal cancer[J]. *Dis Markers*, 2019:2587109.
- [32] Mai D, Ding P, Tan L, et al. PIWI-interacting RNA-54265 is oncogenic and a potential therapeutic target in colorectal adenocarcinoma[J]. *Theranostics*, 2018, 8(19):5213-5230.
- [33] Chen H, Qian J, Werner S, et al. Development and validation of a panel of five proteins as blood biomarkers for early detection of colorectal cancer[J]. *Clin Epidemiol*, 2017, 9(1):517-526.
- [34] Rho JH, Ladd JJ, Li CI, et al. Protein and glycomic plasma markers for early detection of adenoma and colon cancer[J]. *Gut*, 2018, 67(3):473-484.
- [35] Rahier JF, Druez A, Faugeras L, et al. Circulating nucleosomes as new blood-based biomarkers for detection of colorectal cancer [J]. *Clin Epigenetics*, 2017, 9(1):53-59.
- [36] Chen H, Werner S, Butt J, et al. Prospective evaluation of 64 serum autoantibodies as biomarkers for early detection of colorectal cancer in a true screening setting[J]. *Oncotarget*, 2016, 7(13):16420-16432.
- [37] Ostendorff HP, Awad A, Braunschweiger KI, et al. Multiplexed VeraCode bead-based serological immunoassay for colorectal cancer [J]. *J Immunol Methods*, 2013, 400-401(1):58-69.
- [38] Carethers JM. Fecal DNA testing for colorectal cancer screening

- [J]. Annu Rev Med, 2019, 71(7):1-11.
- [39] Suehiro Y, Zhang Y, Hashimoto S, et al. Highly sensitive faecal DNA testing of TWIST1 methylation in combination with faecal immunochemical test for haemoglobin is a promising marker for detection of colorectal neoplasia[J]. Ann Clin Biochem, 2018, 55(1): 59-68.
- [40] Yu J, Feng Q, Wong SH, et al. Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer[J]. Gut, 2017, 66(1):70-78.
- [41] Iwasaki H, Shimura T, Kataoka H. Current status of urinary diagnostic biomarkers for colorectal cancer[J]. Clin Chim Acta, 2019, 498(11):76-83.
- [42] 余捷凯, 杨美琴, 姜铁军, 等. 血清肿瘤标志物优化组合人工神经网络模型在大肠癌诊断中的应用[J]. 浙江大学学报(医学版), 2004, 33(5):36-39.
- [43] 潘爱萍, 李铿, 曹俊涛. 肿瘤标志物人工神经网络模型对结肠直肠癌的诊断价值[J]. 结直肠肛门外科, 2009, 15(6):380-382.
- [44] 谭顺厚, 胡丽敏. 3 种人工神经网络模型预测大肠癌的初步研究[J]. 检验医学与临床, 2011, 8(8):941-942, 944.
- [45] 江永平, 张勇, 蒋宁, 等. 肿瘤标志物联合人工神经网络对大肠癌的预警[J]. 中国卫生检验杂志, 2015, 25(3):371-373, 377.
- [46] 郗群, 毛文虹. 基于机器学习的结直肠癌血清标志物筛选及早期诊断模型评估[J]. 解放军预防医学杂志, 2016, 34(6):879-883, 889.
- [47] 孟潘庆, 贾玉生, 郑树, 等. 组织芯片技术与人工智能神经网络在大肠肿瘤诊断中的应用[J]. 中华肿瘤防治杂志, 2007, 14(17):1324-1327.
- [48] 刘晶, 廖志芳, 樊晓平, 等. 基于决策树分类技术的大肠早癌诊断系统研究[J]. 中国医学工程, 2005, 13(5):462-465.



· 读者 · 作者 · 编者 ·

## 本刊对统计学处理的有关要求

1. 科研设计: 应交代科研方法的名称和主要做法。如调查设计(分为前瞻性、回顾性或横断面调查研究); 实验设计(应交代具体的设计类型, 如自身配对设计、成组设计、交叉设计、析因设计、正交设计等); 临床试验设计(应交代属于第几期临床试验, 采用了何种盲法措施等)。主要做法应围绕 4 个基本原则(随机、对照、重复、均衡)概要说明, 尤其要交代如何控制重要非试验因素的干扰和影响。

2. 资料的表达与描述: 用  $\bar{x} \pm s$  表达近似服从正态分布的定量资料, 用  $M(P_{25} \sim P_{75})$  表达呈偏态分布的定量资料; 用统计表时, 要合理安排纵横表目(三线表), 并将数据的含义表达清楚; 用统计图时, 所用统计图的类型应与资料性质相匹配, 并使数轴上刻度值的标法符合数学原则; 用相对数时, 分母不宜小于 20, 要注意区分百分率与百分比。

3. 统计分析方法的选择: 对于定量资料, 应根据所采用的设计类型、资料所具备的条件和分析目的, 选用合适的统计分析方法, 不应盲目套用  $t$  检验和单因素方差分析; 对于定性资料, 应根据所采用的设计类型、定性变量的性质和频数所具备的条件以及分析目的, 选用合适的统计分析方法, 不应盲目套用  $\chi^2$  检验。对于回归分析, 应结合专业知识和散点图, 选用合适的回归类型, 不应盲目套用简单线性回归分析, 对具有重复实验数据的回归分析资料, 不应简单化处理; 对于多因素、多指标资料, 要在一元分析的基础上, 尽可能运用多元统计分析方法, 以便对因素之间的交互作用和多指标之间的内在联系进行全面、合理的解释和评价。

4. 统计结果的解释和表达: 当  $P < 0.05$  (或  $P < 0.01$ ) 时, 应说明对比组之间的差异有统计学意义, 而不应说对比组之间具有显著性(或非显著性)的差别; 应写明所用统计分析方法的具体名称(如: 成组设计资料的  $t$  检验、两因素析因设计资料的方差分析、多个均数之间两两比较的  $q$  检验等), 应尽可能给出统计量的具体值(如  $t = 3.454$ ,  $\chi^2 = 4.682$ ,  $F = 6.791$  等)和具体的  $P$  值(如  $P = 0.023$ ); 当涉及到总体参数(如总体均数、总体率等)时, 在给出显著性检验结果的同时, 再给出 95% 置信区间。