

食管癌危险因素及预测模型的研究

李保华¹, 肖 迅², 张 敏¹, 蒋 超¹, 程 鹏¹, 任 铁¹

1. 四川省仪陇县人民医院消化内科, 四川 仪陇 637600; 2. 四川省医学科学院·四川省人民医院(电子科技大学医学院)消化内科, 四川 成都 610072

【摘要】 目的 分析食管癌的相关危险因素并建立预测模型, 为食管癌的机会性筛查提供数据基础, 探讨更优筛查方案。方法 收集仪陇县人民医院 2017 年 1 月至 2021 年 12 月胃镜筛查者数据, 进行多因素 Logistic 回归分析, 将筛选出的影响因素建立预测模型, 利用 ROC 计算曲线下面积(AUC)、特异性及敏感性等指标对预测结果进行评价, 并通过十折交叉验证对模型的稳定性进行验证。结果 性别、年龄、萎缩性胃炎、肿瘤家族史、吸烟、饮酒、腌制酸菜、烫食是食管癌的危险因素($P < 0.05$)。构建的食管癌预测模型受试者工作曲线下面积为 0.783(0.759~0.807); 十折交叉试验后表明模型的稳定好。结论 建立的食管癌人群预测模型具有较好的预测价值, 可用于选择性接受内窥镜检查的早癌精准筛查, 提高筛查效益, 节约经济成本。

【关键词】 机会性筛查; 食管癌; 危险因素; 预测模型

【中图分类号】 R735.1

【文献标志码】 A

【文章编号】 1672-6170(2024)06-0132-06

Research on risk factors and prediction models of esophageal cancer LI Bao-hua¹, XIAO Xun², ZHANG Min¹, JIANG Chao¹, CHENG Peng¹, REN Tie¹ 1. Department of Gastroenterology, Yilong County People's Hospital, Yilong 637600, Sichuan, China; 2. Department of Gastroenterology, Sichuan Academy of Medical Sciences · Sichuan Provincial People's Hospital (Affiliated Hospital of University of Electronic Science and Technology of China), Chengdu 610072, China

【Abstract】 **Objective** To analyze the related risk factors of esophageal cancer and establish a predictive model. This was to provide a data base for the opportunistic screening of esophageal cancer and explore a better screening program. **Methods** Data of patients undergoing gastroscopy screening in our hospital from January 2017 to December 2021 were collected. The data were analyzed by using multivariate logistic regression analysis and a predictive model was established based on selected influencing factors. The area under ROC curve (AUC), the specificity, the sensitivity and other indicators were used to evaluate the prediction results. The stability of the model was verified through 10-fold cross-validation. **Results** Gender, age, atrophic gastritis, family history of cancer, smoking, drinking, pickled sauerkraut, and hot food were risk factors for esophageal cancer ($P < 0.05$). The area under the subject working curve of the esophageal cancer prediction model was 0.783(0.759-0.807). The ten-fold crossover test showed that the model was stable. **Conclusions** The established predictive model of esophageal cancer population has good predictive value. It can be used for selective endoscopic screening of early cancer in order to improve the screening efficiency and save the economic cost.

【Key words】 Opportunistic screening; Esophageal cancer; Risk factors; Prediction model

食管癌是起源于食管黏膜上皮的恶性肿瘤, 是临床常见的恶性肿瘤之一, 其发病率、死亡率居高不下, 在全球范围内, 食管癌的发病率和死亡率均排在前十位。2020 年, 全球食管癌的新发病例数超过 60 万, 死亡人数超过 37 万例。而在我国, 食管癌发病率居所有恶性肿瘤第 8 位, 死亡率位居第 4 位^[1]。积极探索食管癌的相关危险因素并且进行一、二级预防显得尤为重要。据 2022 年四川省早癌汇报会数据, 自我省消化道癌早诊早治项目(农村)机会性筛查项目实施以来, 通过对筛查出的早期患者进行及时的治疗或随访, 有效提高了当地上消化道肿瘤特别是食管癌的早诊率, 取得了一定的社会效益。但是由于内镜为有创检查, 受检者痛苦较大, 筛查依从性较低; 且目前全省食管癌内镜检查

碘染色性率在 20%~30%, 70% 左右的受检者无需进行内镜检查, 筛查的效益成本比仍需提高。因此, 亟需开发一种新的更简便的方法, 快速筛查出高危人群, 为“内镜下碘染色+精准活检”提供初筛的科学依据和医学支持, 让更多的食管癌高危人群得到内镜筛查的机会, 从而提高食管癌高发区早诊早治成本效益和工作效率。本研究基于我院两院区上消化道早癌筛查的数据, 分析食管癌的相关危险因素并建立预测模型, 为食管癌的机会性筛查提供一定的数据基础, 探讨更优筛查方案。

1 资料与方法

1.1 一般资料 2017 年 1 月 1 日至 2021 年 12 月 31 日在仪陇县人民医院参加早癌筛查项目的患者及健康体检人群 6824 例。纳入标准: 所有入选的患者(≥ 40 岁)必须完善完善标准胃镜检查, 含普通光学内镜检查+内镜色素检查, 对可疑病灶进行碘染色+精准活检。排除标准: ①病例资料不完整者; ②

【基金项目】 四川省科技厅科研基金资助项目(编号: 2023YFS0083)

既往有食管癌病史者;③正在接受治疗的食管癌患者;④合并胃癌患者;⑤胃部分切除患者;⑥食管胃结合部恶性肿瘤患者。征得知情同意后,根据纳入和排除标准最终共 5238 例纳入本次研究。其中食管癌患者共 340 例(6.49%)作为试验组,非食管癌组共 4898 例(93.51%)作为对照组。

1.2 方法 根据内镜及活检结果,食管恶性肿瘤(含高级别上皮内瘤变)患者为试验组,非食管恶性肿瘤病例为对照组。收集 2017 年 1 月 1 日至 2021 年 12 月 31 日该院行早癌筛查患者的所有资料,为尽量保证资料的准确性,收集的各项资料均有两位研究人员参与问卷的填写及数据的录入,研究人员需要进行统一培训;如果患者 1 年内曾行多次内镜检查,选择第一次内镜检查结果。所有入组患者均收集以下资料:一般资料(家族史、年龄、性别、文化程度、经济状况、体重指数),合并症(恶性肿瘤、高血压、糖尿病)、既往食管慢性病史(Barrett 食管、慢性食管炎、贲门肠上皮化生、霉菌性食管炎),生活习惯(吸烟、饮酒),饮食习惯(是否喜食肥肉、是否爱食腌制酸菜、是否喜食腌肉饮水来源、是否喜食烫食等),喜食肥肉为每周 ≥ 500 g 脂肪、爱食腌制

酸菜进食酸菜或腌制食物 ≥ 5 次/周、筛查时的症状(反酸、烧心、暖气、吞咽困难、咽部异物感等)。

1.3 统计学方法 应用 SPSS 22.0 统计学软件进行数据分析。计数资料以例数(%)表示,比较采用 χ^2 检验;符合正态分布的计量资料以均数 \pm 标准差表示,两组间比较采用 *t* 检验;非正态分布的计量资料以中位数和四分位数间距表示,组间比较采用 Mann-Whitney *U* 检验;对单因素分析中有统计学意义的指标纳入逐步二元 Logistic 回归分析最终食管癌的独立危险因素。利用 Logistic 回归建立食管癌高危人群筛查模型,ROC 曲线、曲线下面积 AUC、特异度及灵敏度等指标对预测结果进行评价,并通过十折交叉验证筛查模型的稳定性与预测效果。 $P < 0.05$ 为差异有统计学意义。

2 结果

2.1 两组一般情况比较 两组年龄、合并萎缩性胃炎及恶性肿瘤家族史的比例比较,差异有统计学意义($P < 0.05$);两组体重指数(BMI)、文化程度及经济状态,以及高血压、糖尿病、慢性食管炎及胃食管反流病比较,差异均无统计学意义($P > 0.05$)。见表 1。

表 1 两组一般情况比较

| 项目 | | 对照组($n=4895$) | 试验组($n=340$) | 统计量 | <i>P</i> |
|------------------|---------|------------------|------------------|------------------|----------|
| 性别[$n(\%)$] | 女 | 2722(55.57) | 99(29.12) | $\chi^2=89.543$ | <0.001 |
| | 男 | 2176(44.43) | 241(70.88) | | |
| 年龄[$n(\%)$] | 45 岁以下 | 638(13.03) | 4(1.18) | $\chi^2=225.316$ | <0.001 |
| | 45~55 岁 | 1713(34.97) | 36(10.59) | | |
| | 55~65 岁 | 1121(22.89) | 78(22.94) | | |
| | 65 岁以上 | 1426(29.11) | 222(65.29) | | |
| 文化程度[$n(\%)$] | 小学及以下 | 2590(52.88) | 178(52.35) | $\chi^2=1.918$ | 0.383 |
| | 初中、高中 | 2216(45.24) | 159(46.76) | | |
| | 高中以上 | 92(1.88) | 3(0.88) | | |
| 经济状况[$n(\%)$] | 差 | 1133(23.13) | 94(27.65) | $\chi^2=3.659$ | 0.160 |
| | 一般 | 2566(52.39) | 166(48.82) | | |
| | 好 | 1199(24.48) | 80(23.53) | | |
| 体重指数(BMI) | | 20.54 \pm 1.94 | 20.34 \pm 1.94 | $t=1.834$ | 0.0674 |
| 高血压[$n(\%)$] | 无 | 4716(96.28) | 322(94.71) | $\chi^2=2.157$ | 0.142 |
| | 有 | 182(3.72) | 18(5.29) | | |
| 糖尿病[$n(\%)$] | 无 | 4770(97.39) | 333(97.94) | $\chi^2=0.389$ | 0.533 |
| | 有 | 128(2.61) | 7(2.06) | | |
| 慢性食管炎[$n(\%)$] | 无 | 4596(93.83) | 325(95.59) | $\chi^2=1.720$ | 0.190 |
| | 有 | 302(6.17) | 15(4.41) | | |
| 萎缩性胃炎 | 无 | 4600(93.92) | 308(90.59) | $\chi^2=5.964$ | 0.015 |

| 项目 | | 对照组 ($n=4895$) | 试验组 ($n=340$) | 统计量 | P |
|--------------------|---|------------------|-----------------|-----------------|-------|
| 胃食管反流病 [$n(\%)$] | 有 | 298(6.08) | 32(9.41) | $\chi^2=1.937$ | 0.164 |
| | 无 | 4540(92.69) | 322(94.71) | | |
| 肿瘤家族史 [$n(\%)$] | 有 | 358(7.31) | 18(5.29) | $\chi^2=11.688$ | 0.001 |
| | 无 | 4828(98.57) | 327(96.18) | | |
| | 有 | 70(1.43) | 13(3.82) | | |

2.2 两组生活及饮食习惯比较 试验组患者吸烟、酗酒比例、喜欢食腌制酸菜、腌肉、烫食以及进食速度快(平均每餐进食时间低于 10 分钟)的比例明显

高于对照组 ($P<0.05$); 两组饮水、辛辣食物、高温热茶等方面比较差异无统计学意义 ($P>0.05$)。见表 2。

表 2 两组生活及饮食习惯比较 [$n(\%)$]

| 项目 | | 对照组 | 试验组 | χ^2 | P |
|---------|-----|--------------|------------|----------|--------|
| 吸烟 | 无 | 4626(94.45) | 284(83.53) | 64.556 | <0.001 |
| | 有 | 272(5.55) | 56(16.47) | | |
| 酗酒 | 无 | 4567(93.24) | 303(89.12) | 8.28 | 0.004 |
| | 有 | 331(6.76) | 37(10.88) | | |
| 喜食肥肉 | 否 | 4818(98.37) | 331(97.35) | 1.956 | 0.224 |
| | 是 | 80(1.63%) | 9(2.65) | | |
| 腌制酸菜 | 否 | 4674(95.43%) | 313(92.06) | 7.904 | 0.005 |
| | 是 | 224(4.57) | 27(7.94) | | |
| 喜食腌肉 | 否 | 4636(94.65) | 310(91.18) | 7.291 | 0.007 |
| | 是 | 262(5.35) | 30(8.82) | | |
| 饮水 | 井水 | 1919(39.18) | 132(38.82) | 0.017 | 0.897 |
| | 自来水 | 2979(60.82) | 208(61.18) | | |
| 喜烫食 | 否 | 4594(93.79) | 304(89.41) | 10.056 | 0.002 |
| | 是 | 304(6.21) | 36(10.59) | | |
| 喜辛辣食品 | 否 | 4595(93.81%) | 315(92.65) | 0.737 | 0.391 |
| | 是 | 303(6.19%) | 25(7.35) | | |
| 喜高温热茶 | 否 | 4795(97.90) | 329(96.76) | 1.915 | 1.915 |
| | 是 | 103(2.10) | 11(3.24) | | |
| 喜红肉及加工肉 | 否 | 4776(97.51) | 328(96.47) | 1.376 | 0.241 |
| | 是 | 122(2.49) | 12(3.53) | | |
| 喜嚼食槟榔 | 否 | 4828(98.57) | 336(98.82) | 0.146 | 0.703 |
| | 是 | 70(1.43) | 4(1.18) | | |
| 进食速度 | 正常 | 4369(89.20) | 262(77.06) | 45.74 | <0.001 |
| | 快 | 529(10.80) | 78(22.94) | | |
| 新鲜水果 | 喜欢 | 4360(89.02) | 305(89.71) | 0.155 | 0.693 |
| | 不喜欢 | 538(10.98) | 35(10.29) | | |
| 喜食蔬菜 | 否 | 4435(90.55) | 318(93.53) | 3.365 | 0.067 |
| | 是 | 463(9.45) | 22(6.47) | | |

2.3 影响食管癌发生的多因素分析 以是否患食管癌为因变量,在上述单因素分析中有统计学差异的因素纳入多因素条件 Logistic 回归分析。结果显

示,性别、年龄、萎缩性胃炎、肿瘤家族史、吸烟、饮酒、腌制酸菜、烫食是食管癌的危险因素 ($P<0.05$)。见表 3。

表 3 影响食管癌的多因素 Logistics 分析

| 因素 | 回归系数 | 标准误 | Z | Wald χ^2 | P | OR | 95% CI |
|-------|-------|-------|--------|---------------|--------|-------|---------------|
| 性别 | 0.937 | 0.129 | 7.273 | 52.897 | <0.001 | 2.552 | 1.983 ~ 3.285 |
| 年龄 | 0.932 | 0.075 | 12.492 | 156.043 | <0.001 | 2.539 | 2.193 ~ 2.938 |
| 萎缩性胃炎 | 0.419 | 0.206 | 2.034 | 4.136 | 0.042 | 1.520 | 1.015 ~ 2.276 |
| 肿瘤家族史 | 1.011 | 0.337 | 3.002 | 9.014 | 0.003 | 2.749 | 1.421 ~ 5.320 |
| 吸烟 | 0.760 | 0.171 | 4.449 | 19.796 | <0.001 | 2.139 | 1.530 ~ 2.991 |
| 饮酒 | 0.385 | 0.193 | 1.991 | 3.963 | 0.046 | 1.469 | 1.006 ~ 2.147 |
| 腌制酸菜 | 0.680 | 0.228 | 2.978 | 8.866 | 0.003 | 1.974 | 1.262 ~ 3.089 |
| 烫食 | 0.615 | 0.197 | 3.124 | 9.758 | 0.002 | 1.850 | 1.258 ~ 2.721 |

2.4 模型的建立和验证 基于多因素分析的 8 项独立因素,建立了预测食管癌预测模型,模型公式为: $\ln(p/1-p) = -6.400 + 0.937 \times \text{性别} + 0.932 \times \text{年龄} + 0.419 \times \text{萎缩性胃炎} + 1.011 \times \text{肿瘤家族史} + 0.760 \times \text{吸烟} + 0.385 \times \text{饮酒} + 0.680 \times \text{腌制酸菜} + 0.615 \times \text{烫食}$ (其中 p 代表食管癌为 1 的概率, $1-p$ 代表高级别或癌为 0 的概率)。该模型的敏感性 0.747, 特异性达 0.887, 受试者工作特征曲线下面积为 0.783 (0.759

~ 0.807)。见图 1。

使用十折交叉法对筛查模型进行组内验证,即将队列数据集随机分为 10 份,轮流将其中 9 份作为训练集,另外 1 份,当作测试集,由此来测试模型的内部稳定性与准确性,经十折交叉验证后,曲线下面积是 0.743 (95% CI: 0.645 ~ 0.841),表明食管癌高危人群筛查模型的效果较稳定。见图 2。

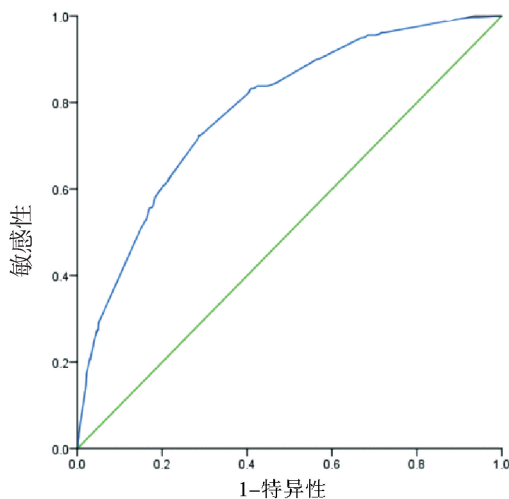


图 1 食管癌预测模型 ROC 曲线

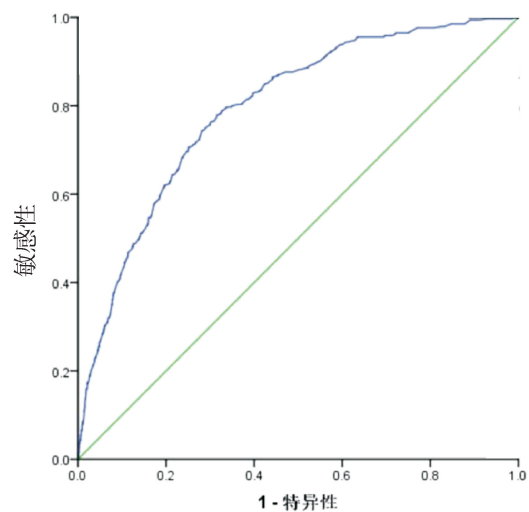


图 2 十折交叉试验后 ROC 曲线

2.5 早期食管癌(含高级别上皮内瘤变)及随访 试验组 340 例阳性病例中早癌患者共 32 例,其中黏膜内癌 4 例,高级别上皮内瘤变 28 例;接受外科手术共 2 例,接受内镜下黏膜剥离术(ESD)患者 26 例,2 例患者外院手术治疗,具体手术方式不详。术后 3 月、6 月及 1 年胃镜随访,均无复发;接受内镜下黏膜剥离术(ESD)患者 2 例术后狭窄,予以内镜下扩张治疗后症状均好转。

3 讨论

因早期食管癌临床症状及体征不典型,绝大多数患者出现明显症状后就诊已处于中晚期,生存率和生存质量较低^[2],因此,食管癌早诊早治对提高生存率,降低死亡率尤为重要^[3]。虽然现阶段国内

外已经有关于食管癌高危人群预测模型的相关研究,但均处于探索阶段,各模型之间的指标、效能均有较大差异^[4-6]。本研究以我院上消化道癌早诊早治项目人群为基础,开发了一个基于既往史、家族史、环境和饮食习惯的因素预测模型,以识别高危个体,为今后机会性筛查提供更好的数据支持。

食管癌的发生目前认为是遗传基因、环境、生活习惯的多发面共同作用的结果^[7]。本研究发现,性别、年龄、萎缩性胃炎、肿瘤家族史、吸烟、饮酒、腌制酸菜、烫食为食管癌发生的独立危险因素。刘浩等研究^[8]发现,南充地区食管癌的发生与年龄、文化程度、吸烟、饮酒、饮食状况、焦虑等密切相关,其结果与本研究一致,但其数据来源于我县

减少,本研究数据更具有说服力。食管癌可发生于任何年龄,食管癌发病率在 40 岁之前处于较低水平,自 40 岁之后快速上升,男、女性发病率均于 80~84 岁年龄组达到高峰^[9]。本研究结果显示,食管癌年龄段大于 65 岁人数最多,占比高达 65.29%,与国家癌症数据中心数据相似。本组资料显示食管癌有明显的性别差异,男女比例为 2.43:1,与其它文献报道相符^[10]。

本研究发现,直系亲属恶性肿瘤病史是食管癌的一个危险因素。Xiao 等^[11]一项关于基于社区的上消化道癌的随机对照试验证实,肿瘤家族史与食管癌前病变的风险增加有关,与本研究结果一致。考虑可能与基因多态性^[12]或相近的饮食结构在其中具有重要作用。

Islami 等^[13]证实,萎缩性胃炎患者罹患食管鳞癌的风险比常人高出 2 倍。此观点在 Nasrollahzadeh 等^[14]的病例对照研究也得到证实,支持萎缩性胃炎是食管鳞癌的危险因素之一,但是未发现近期有关两者之间关系的研究报告。本研究显示,实验组(食管癌组)患者合并萎缩性胃炎的比例明显高于对照组,但是本研究未对胃黏膜萎缩的部位及严重程度进行统计及分析。

食管癌的发病与饮食因素、生活习惯密切相关,已经得到广泛的共识。在我国食管癌高发区食物中存在严重的真菌污染,而且多为不同菌株的混合污染。其作用机制包括产生促癌毒素或促进食物中亚硝酸胺的合成并与其协同致癌^[15]。本研究显示,喜食腌菜也是食管癌发生的独立危险因素,与张瑞等的研究^[16]结果一致,考虑可能腌菜或腌菜中富含亚硝酸盐有关^[17]。

基于上述 8 项独立因素,建立了预测食管癌预测模型,该模型的敏感性 0.747,特异性达 0.887,受试者工作特征曲线下面积为 0.783 (0.759~0.807),表示模型区分度较好,能较好地识别高危和非高危人群。经十折交叉验证后,曲下面积是 0.743, (95% CI:0.645~0.841),表明食管癌高危人群筛查模型的效果较稳定。

本研究发现,食管癌的发生与日常饮食、生活习惯有密切关系。爱吸烟、饮酒、喜食腌制酸菜及烫食的人患食管癌的风险较普通人群明显升高。此外,随着年龄的增长,食管癌的发生率也呈逐步上升趋势。直系亲属肿瘤家族史也是不可忽略的因素。因此,对于今后的食管癌筛查工作中,特别是机会性筛查中,对于有吸烟、饮酒、喜食腌制酸菜及烫食生活及饮食习惯的人群,特别是老年患者,且有肿瘤家族史这部分人群,最好进行精准筛查,如染色、

放大等手段,必要时短期内复查,以免减少漏诊。

早期食管癌传统的治疗方法是外科根治手术,但随着消化内镜微创治疗技术的不断发展与成熟,其与外科手术具有相当的远期疗效和生存率,且具有微创、恢复快、并发症少和术后生活质量高等优点^[18]。夹线法或隧道法提高了内镜下治疗的可行性,内镜下治疗有着高整块切除率和低局部复发率,已得到广泛认可。有研究发现内镜下治疗组和食管切除组之间的死亡率、食管鳞状细胞癌死亡率或无复发生存率没有显著差异。而且,食管切除术的总体不良事件发生率高于内镜下治疗组。本次研究期间发现的早期食管癌患者绝大多数接受的是内镜手术治疗,通过密切随访,均未复发的征像。术后并发症方面,只有 2 例患者出现术后狭窄,考虑与全周病变有关^[19],经过后期的处理,均得以有效解决。

本研究也存在一定的局限性。本研究模型以仪陇县 40 岁以上人群为研究对象,是否适用于其他地区、其他人群亟待外部验证。其次,该模型采用的是回顾性问卷调查,在获取既往信息时,可能存在回忆偏倚;此外,研究中的预测模型基于人群病例对照研究建立,需要对前瞻性收集结局事件和对暴露因素随访动态观察。接下来我们将联合市内、县域内多家机构进行区域内多中心研究,减少上述引起引起的偏倚。

【参考文献】

- [1] Ferlay JEM, Lam F, Colombet M, et al. Global cancer observatory: cancer today [EB/OL]. [2022-01-20]. <https://gco.iarc.fr/today>.
- [2] 赫捷,陈万青,李兆申,等. 中国食管癌筛查与早诊早治指南(2022,北京)[J]. 中华消化外科杂志,2022,21(6):677-700.
- [3] He Z, Liu Z, Liu M, et al. Efficacy of endoscopic screening for esophageal cancer in China (ESECC): design and preliminary results of a population-based randomised controlled trial[J]. Gut, 2019, 68(2): 198-206.
- [4] 李贺,陈万青. 食管癌发病风险评估模型的研究进展[J]. 中国肿瘤,2019,28(12):883-890.
- [5] Xie S H, Lagergren J. A model for predicting individuals' absolute risk of esophageal adenocarcinoma: moving toward tailored screening and prevention[J]. International Journal of Cancer, 2016, 138(12): 2813-2819.
- [6] Kunzmann AT, Thrift AP, Cardwell CR, et al. Model for identifying individuals at risk for esophageal adenocarcinoma[J]. Clinical Gastroenterology and Hepatology,2018,16(8): 1229-1236.
- [7] 谷变利,马丽霞,石林林,等. 食管鳞癌的病因学研究进展[J]. 食管疾病,2023,5(1):47-51.
- [8] 刘浩,冉茂均,刘奕初,等. 四川省南充市食管癌发病因素的调查研究[J]. 大家健康:学术版,2013,7(20):24-25.