

基于机器学习的胃间质瘤内镜手术术后出血风险预测模型的构建与验证

刘罗杰¹, 陈健¹, 高福利¹, 冯云赋², 徐晓丹¹

1. 苏州大学附属常熟医院(常熟市第一人民医院)消化内科, 江苏 常熟 215500; 2. 昆山市第一人民医院内镜中心, 江苏 昆山 215300

【摘要】目的:探讨影响胃间质瘤(gGIST)内镜手术术后出血的危险因素,并应用4种不同机器学习算法构建预测模型,以期准确预测gGIST内镜手术术后出血风险。**方法:**收集gGIST患者资料,以8:2的比例将研究对象随机分配到训练队列($n=502$)和验证队列($n=130$)。在训练队列中,应用合成少数类过采样技术(SMOTE)中的变体SMOTE-NC进行过采样。利用梯度提升机(GBM)、深度学习、广义线性模型和分布式随机森林4种机器学习算法构建预测模型。采用最小绝对值收缩和选择算子筛选变量,构建传统逻辑回归(LR)模型。通过计算受试者工作特征曲线下面积(AUC)、灵敏度、特异度、准确度、阳性预测值和阴性预测值评估模型性能。对最优模型进行包括特征重要性、沙普利近似法(SHAP)和力图在内的可解释性分析,并开发一款可实际应用的网络应用程序。**结果:**在632例患者中,78例(12.3%)发生了术后出血。在验证队列中,对比5种预测模型,GBM模型表现最佳,其AUC值为0.889,95%CI为0.829~0.948,优于其他模型。变量重要性分析显示,术者经验、手术时间、术中大出血、肿瘤大小等因素对预测术后出血具有重要影响。SHAP图和力图展示了变量在二分类预测结果中的分布特征,以及各变量对预测结果的影响。**结论:**GBM模型对预测gGIST内镜术后出血具有较好的预测价值。同时,网络应用程序的构建方便了临床使用。

【关键词】胃间质瘤;自动化机器学习;内镜手术;术后出血;预测模型

【中图分类号】R318;R573

【文献标志码】A

【文章编号】1005-202X(2025)04-0550-11

Construction and validation of machine learning-based prediction models for postoperative bleeding following endoscopic resection of gastric gastrointestinal stromal tumor

LIU Luo-jie¹, CHEN Jian¹, GAO Fuli¹, FENG Yunfu², XU Xiaodan¹

1. Department of Gastroenterology, Changshu No.1 People's Hospital/Changshu Hospital Affiliated to Soochow University, Changshu 215500, China; 2. Endoscopy Center, the First People's Hospital of Kunshan, Kunshan 215300, China

Abstract: Objective To explore the risk factors for postoperative bleeding after endoscopic resection of gastric gastrointestinal stromal tumor (gGIST) and to construct prediction models using 4 different machine learning algorithms for accurately predicting postoperative bleeding. **Methods** The clinical data of gGIST patients were collected, and the patients were randomly divided into a training cohort ($n=502$) and a validation cohort ($n=130$) at an 8:2 ratio. Synthetic minority over-sampling technique-nominal continuous was used for oversampling in the training cohort. Four prediction models were constructed using gradient boost machine (GBM), deep learning, generalized linear model and distributed random forest, separately; and in addition, the least absolute shrinkage and selection operator was used to screen variables and construct a traditional Logistic regression model. Model performance was evaluated by calculating the area under the receiver operating characteristic curve (AUC), sensitivity, specificity, accuracy, positive predictive value and negative predictive value. Interpretability analyses, including feature importance, SHapley additive exPlanation and force plot, were performed on the optimal model, and a practically applicable web application was developed. **Results** Among 632 patients, 78 (12.3%) experienced postoperative bleeding. In the validation cohort, GBM model performed best among 5 prediction models, with

【收稿日期】2024-11-15

【基金项目】常熟市科技计划(基础研究计划-医学应用基础研究)(CY202339);苏州市“科教兴卫”青年科技项目(KJXW2023067);常熟市第一人民医院消化内科临床研究机构能力提升项目(SLT2023006)

【作者简介】刘罗杰,博士,主治医师,研究方向:胃肠道间质瘤的临床诊治、机器学习等, E-mail: luojieliu@126.com

【通信作者】徐晓丹,博士,主任医师,研究方向:胃肠道间质瘤的临床诊治、机器学习等, E-mail: xxd20@163.com

an AUC value of 0.889 and a 95%CI of 0.829-0.948, superior to the other 4 models. Variable importance analysis identified surgeon experience, operation time, intraoperative hemorrhage, tumor size as the factors affecting postoperative bleeding prediction. The SHapley additive exPlanation plot and force plot showed the distribution characteristics of variables in the binary classification prediction results and the effect of each variable on the prediction results. **Conclusion** GBM model has high predictive value for postoperative bleeding following endoscopic resection of gGIST, and the construction of the web application facilitates its clinical use.

Keywords: gastric gastrointestinal stromal tumor; automated machine learning; endoscopic surgery; postoperative bleeding; prediction model

前言

胃肠道间质瘤(Gastrointestinal Stromal Tumors, GISTs)是消化道最常见的间叶细胞源性肿瘤,起源于胃肠道Cajal细胞或其前体细胞^[1]。近年来,内镜技术的进步为治疗低风险胃间质瘤(gastric Gastrointestinal Stromal Tumor, gGIST)提供了机会^[2]。多项研究表明,与外科手术相比,内镜下治疗gGIST安全有效,具有术后恢复速度快、住院时间短、手术费用低等优点^[3-5]。但是,与内镜切除胃黏膜病变相比,内镜下治疗gGIST的技术难度更大,术后更易发生出血、穿孔和感染等并发症^[5-7]。既往研究报道,内镜下治疗gGIST发生术后出血的概率为0~8.9%^[8-11]。术后出血不仅会增加患者的医疗费用,还可能对患者的身心健康造成负面影响,影响患者的康复效果。因此,准确预测内镜切除gGIST术后出血风险,并帮助内镜医生识别出高风险患者,早期采取有针对性的干预措施,对提高患者的预后具有重要作用。近年来,国内多项研究显示,手术时间、病灶部位、肿瘤大小、术者经验、基础性疾病史以及是否发生术中大出血等因素与内镜切除gGIST术后出血密切相关^[12-16]。然而,上述研究结论存在较大差异,且部分研究的样本量较小。同时,当前研究尚未实现对内镜切除gGIST发生术后出血的准确预测。而预测模型的建立能够为临床医生提供个体化的风险评估工具,辅助其临床决策。自动化机器学习(Automated Machine Learning, AutoML)是一种融合人工智能和自动化技术的全新方法,它从大量数据中自动学习模型,并采用自动化的方式进行模型的训练和优化,能显著缩短模型构建和优化的时间成本^[17-18]。因此,本研究拟通过多中心的研究进一步探讨影响gGIST内镜手术后出血的危险因素,并应用4种机器学习算法构建预测模型,以期为临床医生提供一个预测术后出血风险的有效工具。

1 资料与方法

1.1 一般资料

回顾性收集2010年12月~2022年12月在苏州大学附属第一医院、南京市第一医院、常熟市第一人民

医院、常熟市第二人民医院和昆山市第一人民医院消化内科接受内镜治疗的gGIST患者资料。本研究中所收集的数据来自多中心的电子病历资料库,收集的资料包括患者的性别、年龄、抗凝抗血小板药物史、基础疾病史(包括高血压、冠心病和糖尿病病史)、肿瘤的部位、大小、美国国立卫生研究院(National Institutes of Health, NIH)危险度分级(采用2008年的改良版)^[19]、内镜治疗方式、操作者的手术经验、手术时间、创面闭合方式、术中穿孔、术中大出血的发生情况、术后禁食天数、术后住院天数、R0切除情况、术后出血发生情况。

纳入标准:(1)病理及免疫组化确诊为gGIST;(2)术前血凝、血常规、心电图检查无明显异常;(3)既往无胃部手术史;(4)术前未行分子靶向药物治疗;(5)签署相关知情同意书。排除标准:(1)有淋巴结转移及远处转移;(2)胃内多发病灶同期手术;(3)合并其他部位同期内镜手术;(4)心肺功能较差,无法耐受麻醉及手术;(5)资料不完整;(6)中转外科手术。本研究方案通过了医院伦理委员会的伦理审查(伦理审查批号X2023024)。

1.2 方法

1.2.1 内镜治疗过程 所有患者在术前均常规进行血常规、血生化、血凝、肿瘤标志物、心电图、腹部CT、超声内镜等检查,以排除手术禁忌症并评估肿瘤的具体部位、大小、浸润深度及与邻近器官的关系。同时告知患者及家属相关手术风险及手术的具体过程,签署相关知情同意书。手术过程中,所有患者取左侧卧位,术前行气管插管及全身麻醉。术内在内镜头端安装透明帽,术中使用二氧化碳气泵进行充气,使用注水泵进行冲洗。在这项研究中,进行gGIST内镜治疗的医生具有不同的手术经验。但是,所有参与手术的内镜医生在开展gGIST内镜治疗之前,都已经完成了超过5 000例的胃肠镜检查以及100例以上的内镜下微创手术,具有一定的内镜下治疗经验。本研究共纳入了3种内镜下治疗gGIST的手术方式:包括内镜黏膜下剥离术(Endoscopic Submucosal Dissection, ESD)、内镜下全层切除术(Endoscopic Full-Thickness Resection, EFTR)和经黏膜下隧道内

镜肿瘤挖除术(Submucosal Tunneling Endoscopic Resection, STER),具体操作方法参考文献[9, 20]。

1.2.2 相关定义 (1)手术时间以电子病历系统中麻醉记录单为准(除外麻醉诱导及麻醉苏醒时间)。(2)术中大出血是指在手术中有活动性渗血或喷射性出血且内镜下止血困难,在手术的当日或次日血红蛋白下降 ≥ 2 g/dL,或需要中断手术和(或)输血治疗^[21]。(3)术后出血是指内镜治疗后出现明显呕血或黑便,血红蛋白水平下降 ≥ 2 g/dL,且生命体征发生变化,急诊内镜检查证实为黏膜缺损出血或胃出血^[22]。(4)R0切除是指标本经组织学检查提示侧切缘和垂直切缘无肿瘤残留^[23]。笔者前期通过累积和分析法(Cumulative Sum, CUSUM)发现,内镜切除gGIST的学习曲线大约为50例^[24]。依据该方法,笔者将同一名医生在完成前50例手术时的情况定义为经验不足,当完成超过50例手术时,将其定义为经验丰富。

1.2.3 模型构建 将纳入的患者按照8:2的比例随机分配到训练队列(502例)和验证队列(130例)。然而,由于本研究中术后未出血组与术后出血组的病例数比值大约为7:1(554:78),数据类别存在明显的不平衡,阳性病例的样本量较小。为了解决这一不平衡问题,在训练队列中应用了合成少数类过采样技术(Synthetic Minority Over-sampling Technique, SMOTE)中的SMOTE-NC(SMOTE for Nominal and Continuous features)技术^[25]。SMOTE-NC作为SMOTE的一种变体,能够同时处理连续性和分类变量^[26]。SMOTE可通过合成新的少数类样本来增强数据集的多样性,而不只是简单地复制现有的少数类样本。这有助于提高分类模型对少数类的识别能力,有效地缓解类不平衡问题。

经过SMOTE-NC过采样后,在训练队列中采用最小绝对值收缩和选择算子(Least Absolute Shrinkage and Selection Operator, LASSO)算法筛选变量,并构建传统逻辑回归(Logistic Regression, LR)模型。同时,借助H2O平台的AutoML框架构建多种机器学习模型,主要包括梯度提升机(Gradient Boost Machine, GBM)、基于深度神经网络(Deep Neural Networks, DNN)的深度学习(Deep Learning, DL)、广义线性模型(Generalized Linear Model, GLM)和分布式随机森林(Distributed Random Forest, DRF)^[27]。为了提高模型的性能和泛化能力,在训练队列中进行了十折交叉验证,具体做法为:将数据集分为10个子集,每次使用其中9个子集进行训练,然后用剩下的1个子集进行验证。重复这个过程10次,每次都选择不同的验证子集。最后根据10次验证结果的受试者工作特征(Receiver Operating Characteristic, ROC)曲

线下面积(Area Under Curve, AUC)选择最佳的模型^[28]。研究流程图见图1。

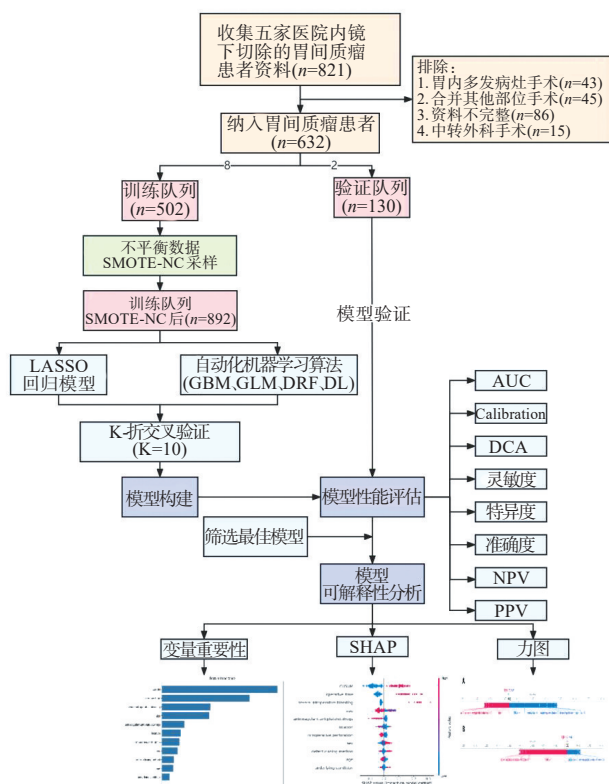


图1 研究流程图

Figure 1 Flowchart of the study

1.2.4 模型性能评估 通过计算AUC值来评估模型性能的优劣。同时,通过绘制校准曲线和临床决策曲线(Decision Curve Analysis, DCA)来判断模型的校准度和临床适用度。此外,为了全面评估模型的综合性能,笔者还计算了各个模型的灵敏度、特异度、准确度、阳性预测值(Positive Predictive Value, PPV)和阴性预测值(Negative Predictive Value, NPV)。

1.2.5 模型可解释性分析 对性能最佳的模型进行特征重要性、沙普利近似法(SHapley Additive exPlanations, SHAP)和力图等可解释性分析。

1.3 统计学方法

计数资料以例数(%)表示,行卡方检验或Fisher确切概率法。采用Kolmogorov-Smirnov检验对连续变量进行正态性检验,符合正态分布的计量资料以均值 \pm 标准差表示,两组间比较行 t 检验;非正态分布的计量资料以中位数和四分位数间距[M(P25, P75)]表示,两组间比较行Mann-Whitney U检验。采用R(4.1.0版)和Python(3.6版)进行数据分析、绘图及网络应用程序的制作。 $P<0.05$ 为结果有统计学差异。

2 结果

2.1 两组患者基线资料比较

本研究共纳入 632 例患者, 其中 78 例 (12.3%) 发生术后出血。两组患者在肿瘤大小、术者经验、手术时间、术中大出血、术后禁食天数、术后住院天数、R0 切除情况和危险度分级方面比较, 差异有统计学意

义 ($P<0.05$)。但是, 两组患者在性别、年龄、抗凝抗血小板药物史、基础疾病史、肿瘤部位、内镜方式、术中穿孔和创面闭合情况方面比较, 差异均无统计学意义 ($P>0.05$) (表 1)。本研究中, 将总体队列按 8:2 的比例随机分配到训练队列 (502 例) 和验证队列 (130 例), 两队列在术前及术中资料方面比较, 差异均无统计学意义 ($P>0.05$) (表 2)。

表 1 术后出血组与未出血组患者临床资料比较
Table 1 Comparison of clinical data between postoperative bleeding and nonbleeding groups

特征	术后出血组 (n=78)	术后未出血组 (n=554)	χ^2/U 值	P 值
性别[例(%)]			0.013	0.910
男	32(41.0)	231(41.7)		
女	46(59.0)	323(58.3)		
年龄[例(%)]			2.562	0.109
<60 岁	45(57.7)	266(48.0)		
≥60 岁	33(42.3)	288(52.0)		
抗凝抗血小板药物史[例(%)]			0.977	0.323
有	14(17.9)	127(22.9)		
无	64(82.1)	427(77.1)		
基础疾病[例(%)]			1.525	0.217
有	19(24.4)	173(31.2)		
无	59(75.6)	381(68.8)		
肿瘤部位[例(%)]			2.270	0.321
胃上 1/3	58(74.4)	365(65.9)		
胃中 1/3	14(17.9)	127(22.9)		
胃下 1/3	6(7.7)	62(11.2)		
肿瘤大小[例(%)]			13.219	0.001
≥3.0 cm	19(24.4)	63(11.4)		
2.0~3.0 cm	23(29.5)	135(24.4)		
<2.0 cm	36(46.2)	356(64.3)		
手术经验[例(%)]			18.291	<0.001
经验不足	42(53.8)	164(29.6)		
经验丰富	36(46.2)	390(70.4)		
内镜方式[例(%)]			0.296	0.862
ESD	43(55.1)	288(52.0)		
EFTR	34(43.6)	257(46.4)		
STER	1(1.3)	9(1.6)		
创面闭合方式[例(%)]			0.649	0.723
钛夹	60(76.9)	447(80.7)		
尼龙绳	9(11.5)	51(9.2)		
OTSC	9(11.5)	56(10.1)		
手术时间[例(%)]			15.009	<0.001
<90 min	55(70.5)	483(87.2)		
≥90 min	23(29.5)	71(12.8)		
术中大出血[例(%)]			18.767	<0.001
有	7(9.0)	7(1.3)		
无	71(91.0)	547(98.7)		
术中穿孔[例(%)]			0.216	0.642
有	34(43.6)	257(46.4)		
无	44(56.4)	297(53.6)		
术后禁食[d, M(P25, P75)]	3.0(3.0, 5.0)	3.0(2.0, 3.0)	14281	<0.001
术后住院[d, M(P25, P75)]	6.0(6.0, 7.5)	5.0(4.0, 6.0)	11365	<0.001
R0 切除[例(%)]			7.512	0.006
是	65(83.3)	513(92.6)		
否	13(16.7)	41(7.4)		
危险度分级[例(%)]			11.520	0.001
极低/低危险	62(79.5)	508(91.7)		
中高危险	16(20.5)	46(8.3)		

OTSC: 耙状金属夹闭合系统

表 2 训练队列与验证队列临床资料比较[例(%)]

Table 2 Comparison of clinical data between training cohort and validation cohort [cases (%)]

特征	训练队列 (n=502)	验证队列 (n=130)	χ^2	P 值
性别			0.958	0.328
男	204(40.6)	59(45.4)		
女	298(59.4)	71(54.6)		
年龄			0.001	0.996
<60 岁	247(49.2)	64(49.2)		
≥60 岁	255(50.8)	66(50.8)		
抗凝抗血小板药物史			1.399	0.237
有	117(23.3)	24(18.5)		
无	385(76.7)	106(81.5)		
基础疾病			1.388	0.239
有	147(29.3)	45(34.6)		
无	355(70.7)	85(65.4)		
肿瘤部位			5.622	0.060
胃上 1/3	325(64.7)	98(75.4)		
胃中 1/3	121(24.1)	20(15.4)		
胃下 1/3	56(11.2)	12(9.2)		
肿瘤大小			0.509	0.775
≥3.0 cm	67(13.3)	15(11.5)		
2.0~3.0 cm	123(24.5)	35(26.9)		
<2.0 cm	312(62.2)	80(61.5)		
手术经验			0.083	0.773
经验不足	165(32.9)	41(31.5)		
经验丰富	337(67.1)	89(68.5)		
创面闭合方式			1.154	0.562
钛夹	407(81.1)	100(76.9)		
尼龙绳	46(9.2)	14(10.8)		
OTSC	49(9.8)	16(12.3)		
手术时间			2.454	0.117
<90 min	433(86.3)	105(80.8)		
≥90 min	69(13.7)	25(19.2)		
术中大出血			0.561	0.454
有	10(2.0)	4(3.1)		
无	492(98.0)	126(96.9)		
术中穿孔			0.135	0.714
有	233(46.4)	58(44.6)		
无	269(53.6)	72(55.4)		

2.2 训练队列 SMOTE-NC 过采样

本研究中,训练队列中术后出血组和未出血组分别有 65 例和 437 例。术后出血组为少数类样本,经 SMOTE-NC 过采样后该组的病例数为 65×7=455 例,此时术后出血组与术后未出血组的比值为 1.04:1。经过 SMOTE-NC 过采样后,两组在抗凝抗血小板药物史、肿瘤大小、手术经验、手术时间和是否发生术中大出血方面比较,差异具有统计学意义 ($P<0.05$),见表 3。

2.3 LR 模型构建及性能评估

采用 LASSO 回归模型筛选 11 个可能的危险因素。图 2a 展示了 Lasso 回归系数随正则化参数变化的情况。图 2b 展示了 λ 值(对数值)与二项式偏差值之间的关系,当 λ 值(对数值)逐渐增大时,模型的二

项式偏差值先减小后增大。当二项式偏差值达到最小值时(左侧虚线),共有 10 个特征项被纳入模型。当 λ 取值为“ λ_{1se} ”时,模型仅包含 6 个特征项。最终经过多因素分析,得到了一个简化的预测模型,该模型包含 4 个特征项:手术时间、术者的经验、术中穿孔和术中大出血。使用这 4 个特征构建了 LR 模型,并通过 Nomogram 图对模型结果进行了可视化展示(图 3)。

验证队列中,LR 模型的 AUC 值为 0.851,95%CI 为 0.782~0.920(图 4a)。图 4b 展示了在经过 500 次自助法重复抽样后得到的校准曲线,其平均绝对误差为 0.026,表明 LR 模型的校准度总体可接受。图 4c 展示了 LR 模型的临床决策曲线,可以发现在术后出血的预测阈值介于 20%~80% 时,通过早期干预可能会带来 1%~35% 的净获益。

表3 经SMOTE-NC过采样后训练队列中两组患者临床资料比较[例(%)]

Table 3 Comparison of clinical data between two groups in the training cohort after SMOTE-NC oversampling [cases (%)]

特征	术后出血组(n=455)	术后未出血组(n=437)	χ^2	P值
性别			0.517	0.472
男	193(42.4)	175(40.0)		
女	262(57.6)	262(60.0)		
年龄			1.151	0.283
<60岁	235(51.6)	210(48.1)		
≥60岁	220(48.4)	227(51.9)		
抗凝抗血小板药物史			4.135	0.042
有	84(18.5)	105(24.0)		
无	371(81.5)	332(76.0)		
基础疾病			1.615	0.204
有	121(26.6)	133(30.4)		
无	334(73.4)	304(69.6)		
肿瘤部位			0.847	0.655
胃上1/3	294(64.6)	278(63.6)		
胃中1/3	103(22.6)	109(24.9)		
胃下1/3	58(12.7)	50(11.4)		
肿瘤大小			20.269	<0.001
≥3.0 cm	88(19.3)	53(12.1)		
2.0~3.0 cm	139(30.5)	101(23.1)		
<2.0 cm	228(50.1)	283(64.8)		
手术经验			49.241	<0.001
经验不足	243(53.4)	132(30.2)		
经验丰富	212(46.6)	305(69.8)		
创面闭合方式			2.313	0.315
钛夹	356(78.2)	358(81.9)		
尼龙绳	44(9.7)	39(8.9)		
OTSC	55(12.1)	40(9.2)		
手术时间			52.607	<0.001
<90 min	309(67.9)	385(88.1)		
≥90 min	146(32.1)	52(11.9)		
术中大出血			51.482	<0.001
有	61(13.4)	4(0.9)		
无	394(86.6)	433(99.1)		
术中穿孔			0.180	0.671
有	207(45.5)	205(46.9)		
无	248(54.5)	232(53.1)		

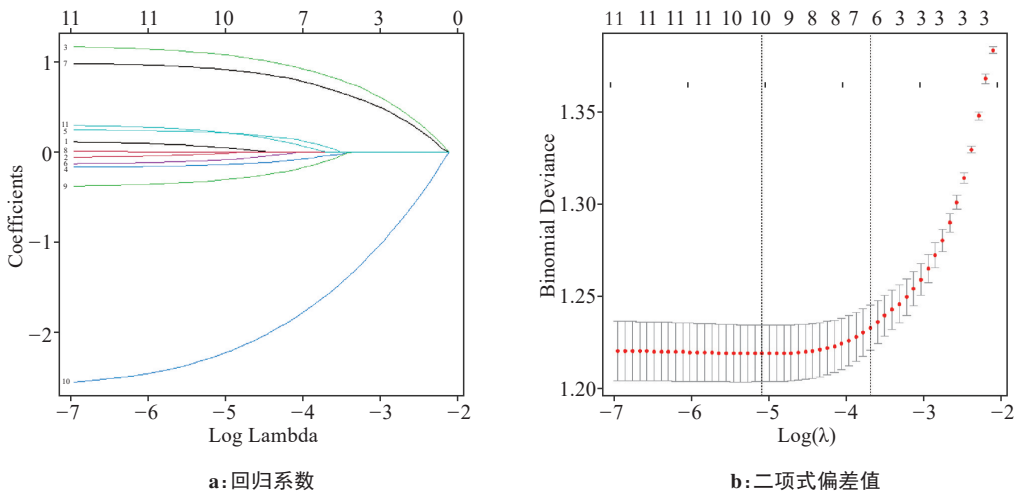


图2 LASSO回归变量筛选

Figure 2 Variable screening in LASSO regression

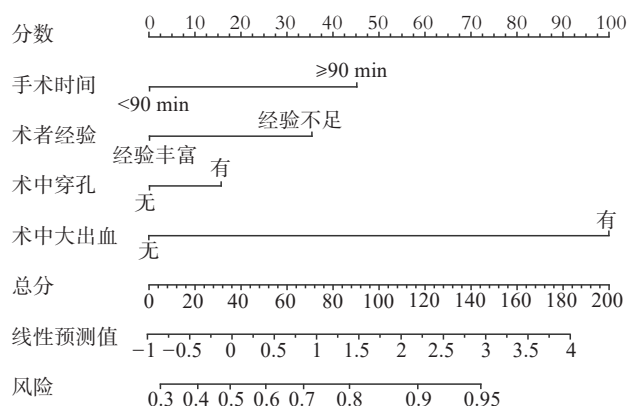


图3 Nomogram图预测gGIST内镜手术术后出血风险
Figure 3 Nomogram for predicting the risk of postoperative bleeding after endoscopic resection of gGIST

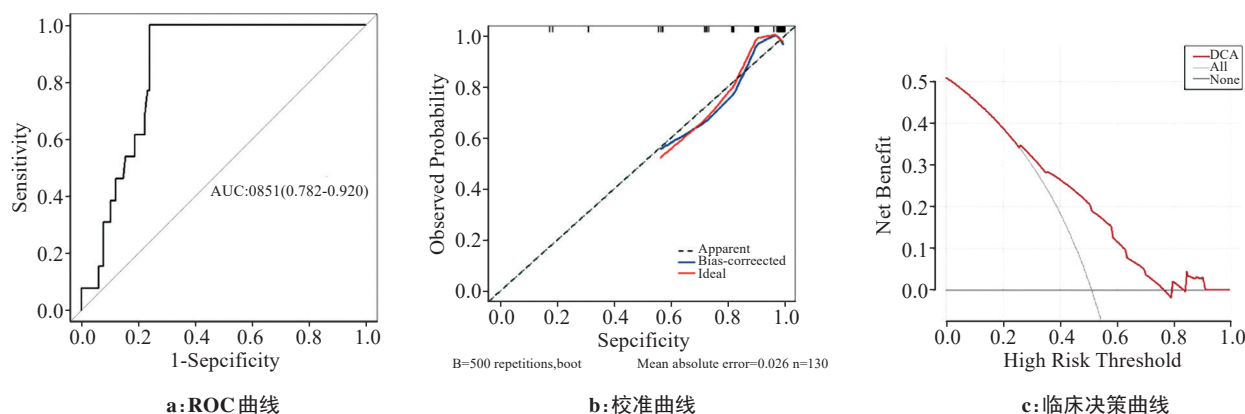


图4 验证队列中LR模型性能评估曲线
Figure 4 Performance evaluation curve of LR model in the validation cohort

2.5 最优模型可解释性分析

2.5.1 变量重要性分析 在模型性能评估阶段,笔者发现 GBM 模型在各项评估指标上的表现均最佳。变量重要性分析显示,术者经验是影响术后出血最重要的因素,其次分别为手术时间、术中大出血、肿瘤大小、抗凝抗血小板药物史、肿瘤部位、术中穿孔、性别、创面闭合方式、年龄和基础疾病(图7)。

2.5.2 SHAP 分析 SHAP 分析图展示了 11 个特征项对预测结果的影响程度。笔者观察到术者经验、手术时间和术中大出血是对预测结果影响最大的 3 个特征。SHAP 图揭示,当术者经验不足、手术时间超过 90 min 以及发生术中大出血时,其 SHAP 值均大于 0,这表明当出现上述情况时发生术后出血的可能性较大(图8)。

2.5.3 力图分析 图9展示了基于 GBM 模型对验证队列中随机选取的两个病例进行的力图分析。图9a

2.4 机器学习模型构建及性能评估

基于 H2O 平台的 AutoML 算法,笔者成功构建了 46 种不同的模型,其中包括 24 种 GBM 模型,1 种 GLM 模型,20 种 DL 模型以及 1 种 DRF 模型。在验证队列中,GBM 模型展现出了较好的分类性能,其 AUC 值为 0.889, 95%CI 为 0.829~0.948 (图 5a)。GLM 模型的 AUC 值为 0.886, 95%CI 为 0.775~0.996 (图 5b)。DRF 模型的 AUC 值为 0.797, 95%CI 为 0.691~0.902 (图 5c)。DL 模型的 AUC 值为 0.871, 95%CI 为 0.804~0.938 (图 5d)。此外,图 6 展示了这 4 种机器学习模型在验证队列中的临床决策曲线,结果显示最大净收益接近 40%。表 4 详细展示了在验证队列中 5 种不同预测模型的综合性能表现。GBM 模型的 AUC、灵敏度、特异度、准确度、PPV 和 NPV 均高于其他预测模型。

中的病例实际诊断为阴性,模型预测发生术后出血的风险为 38%,预测值较低,与实际诊断相符。图 9b 中的病例预测术后出血的概率为 62%,其实际诊断为阳性,与实际相符,再次验证了模型预测的准确性。

2.6 本地网络应用程序

基于 GBM 模型,笔者设计了一款本地网络应用程序。利用这款程序,用户可以从 11 个特征项中任意选择 2~11 个特征项,用于预测患者术后出血的风险。通过扫描图 10 中的二维码可了解该应用程序的具体操作方法。

3 讨论

手术治疗仍是当前治疗 gGIST 的首选方式^[29]。随着内镜技术的发展,手术方式已经从传统的开腹手术向腹腔镜手术和内镜手术演变。由于内镜手术具有创伤小、费用低、术后恢复快等优势,目前已经

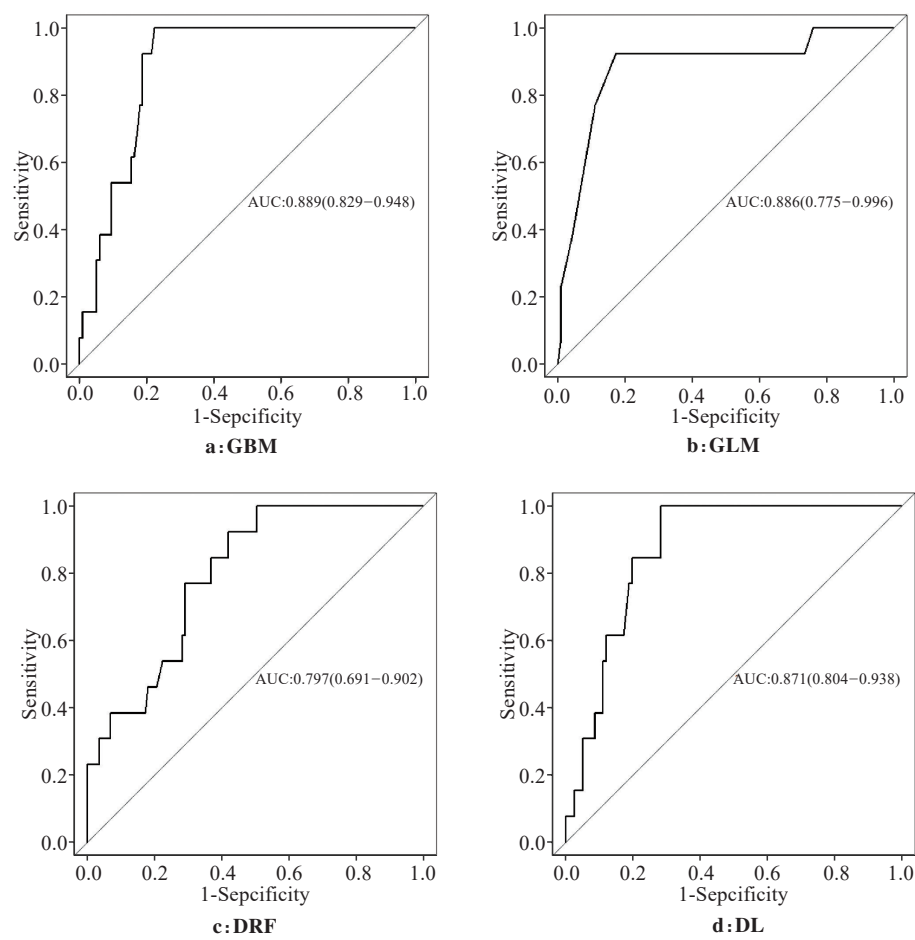


图5 验证队列中不同机器学习模型的 ROC 曲线

Figure 5 ROC curves for different machine learning models in the validation cohort

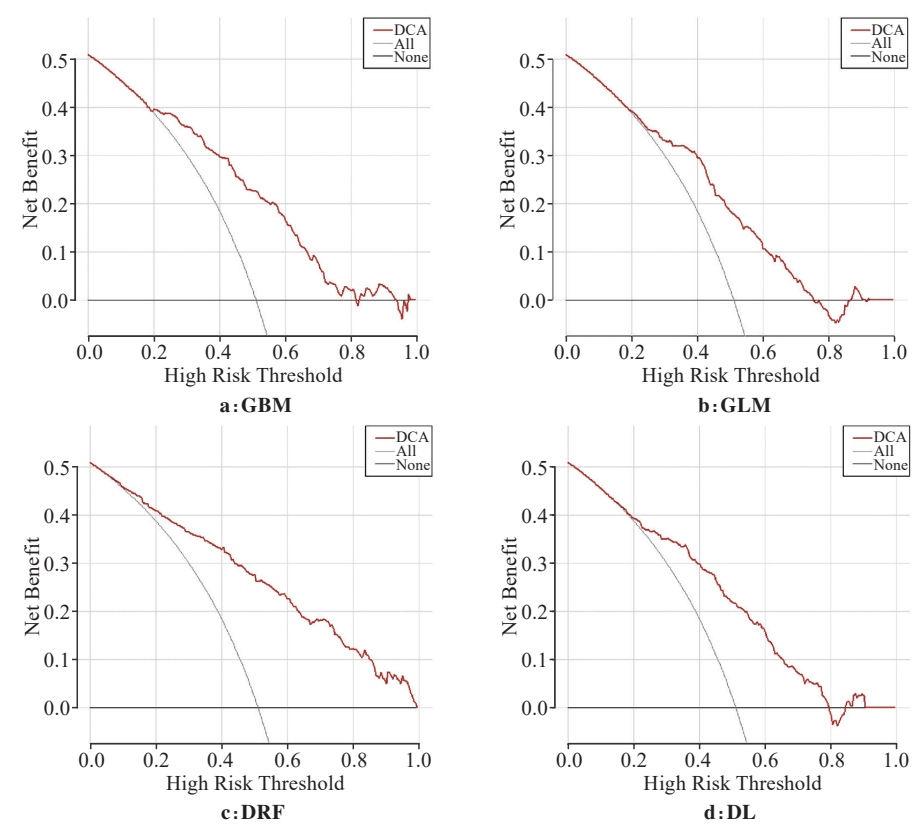


图6 验证队列中不同机器学习模型的临床决策曲线

Figure 6 Decision curve analysis for different machine learning models in the validation cohort

表 4 验证队列中 5 种预测模型性能表现

Table 4 Performance evaluation of 5 prediction models in the validation cohort

模型	AUC	灵敏度	特异度	准确度	PPV	NPV
GBM	0.889	1.000	0.829	0.846	0.394	1.000
GLM	0.886	0.923	0.829	0.838	0.375	0.990
DRF	0.797	0.923	0.581	0.615	0.197	0.986
DL	0.871	1.000	0.718	0.746	0.283	1.000
LR	0.851	1.000	0.761	0.785	0.371	1.000

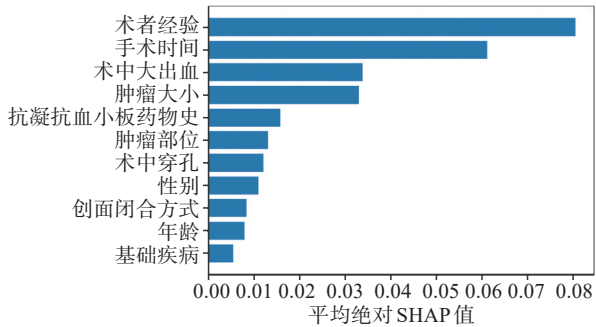


图 7 训练队列中 GBM 模型变量重要性排名

Figure 7 Ranking of variable importance for GBM model in the training cohort

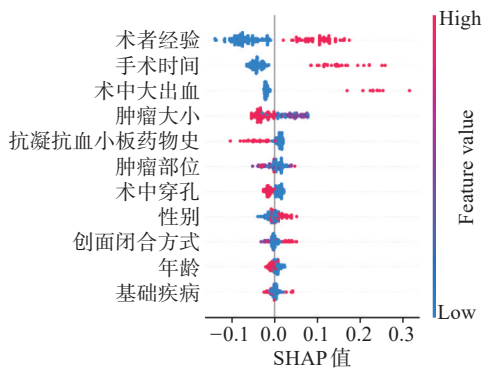


图 8 训练队列中 GBM 模型 SHAP 特征分析

Figure 8 SHAP feature analysis for GBM model in the training cohort

成为治疗 gGIST 的首选手术方式^[3-5]。但是,内镜下治疗 gGIST 存在术后出血的风险^[8-11]。本研究中,632 例 gGIST 患者中共有 78 例(12.3%)发生了术后出血。张媛等^[15]的研究报道显示,ESD 技术治疗 gGIST 术后出血的发生率为 8.89%,稍低于本文报道。这可能与两项研究中瘤体直径大小、术者手术经验、内镜术式组成、手术时间长短等因素不同有关。

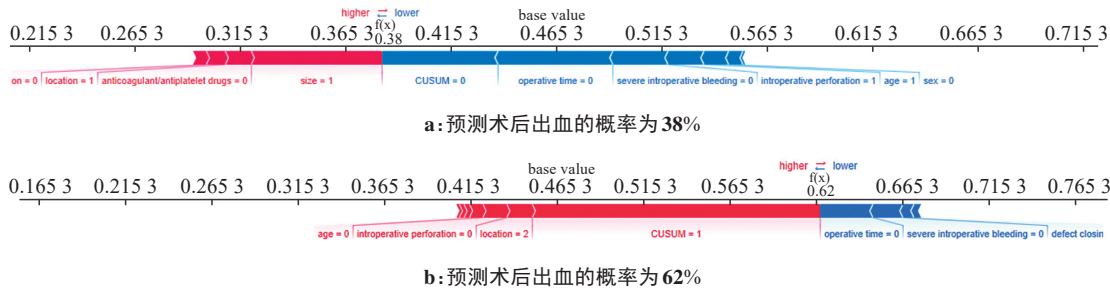


图 9 验证队列中 GBM 模型的力图分析

Figure 9 Force plot analysis of GBM model in the validation cohort



图 10 基于 GBM 模型的本地网络应用程序操作视频

Figure 10 Operational video of a local web application based on GBM model

本研究中,术后出血组与术后未出血组的病例数比例大约为 1:7,两组之间存在明显的数

据不平衡。这种不平衡可能会导致模型偏向于多数类,从而降低对少数类的预测能力。这一偏差可能促使模型优先提升总体准确率,而忽视对阳性样本的识别。王梅英等^[30]在预测肿瘤患者化疗后发生下呼吸道感染的研究中,建模组中发生下呼吸道感染的病例数为 99 例,而未发生下呼吸道感染的病例数为 1 569 例,两组的比值约为 1:16,数据极不平衡。基于该数据构建的预测模型 AUC 值为 0.780。而通过 SMOTE 过采样技术处理后,模型的 AUC 值为 0.949。此外,在乳腺癌生存预测^[31]、肝纤维化严重程度鉴定^[32]、肥胖程度评估^[33]等研究中,都应用了 SMOTE-NC 技术进行过采样处理,并且取得较为理想的结果。因此,本研究中笔者采用 SMOTE-NC 技术对术后出血组的病例进行了过采样处理。

GBM模型的变量重要性和SHAP分析显示,术者经验、手术时间和术中大出血是影响术后出血的重要因素。其中,术者经验是影响术后出血最关键的因素。与本研究结论一致,李云等^[14]在研究STER技术治疗gGIST术后出血的危险因素时发现,术者经验不足,操作不熟练是导致术后出血的独立危险因素。具有丰富经验的内镜医生通常拥有更加熟练的操作技能和卓越的决策能力。这使得他们能够在手术过程中即时识别并妥善处理潜在风险。Toyokawa等^[34]认为,术者经验不足和操作不熟练可能导致手术时间延长,这对年纪较大、基础情况较差的患者影响尤为显著,可能增加术后并发症的风险。因此,在进行gGIST手术时,若术者经验较为有限,应在瘤体切除后仔细检查创面,并即时处理暴露的血管。同时,术后应密切观察患者的病情变化,以便即时发现并处理可能出现的并发症。

手术时间也是影响gGIST内镜手术术后出血的危险因素之一。本研究中,78例发生术后出血的患者中,29.5%(23/78)的病例手术时间超过90 min,而在未发生术后出血的554例患者中,这一比例仅为12.8%(71/554)。这表明手术时间越长,术后出血的风险越高。同时,芦迅达等^[35]和胡康等^[36]的研究也发现,手术时间是影响早期胃癌ESD术后出血的危险因素之一。手术时间的延长可能会导致胃黏膜的损伤及炎症反应的增加,从而增加术后出血的风险。此外,长时间手术通常意味着手术过程更复杂,术中电凝电切的次数更多,增加了术后出血的可能性。因此,笔者建议术前应充分评估手术难度,并制定精确的治疗方案以缩短手术时间。对于手术时间较长的病例,术后应加强生命体征监测和恢复情况跟踪,并即时进行相应处理。同时,应用GBM模型提前预测术后出血风险,并采取早期干预措施,以降低术后出血发生率。

Suzuki等^[37]的研究显示,发生术中大出血是ESD术后出血的一个重要危险因素。Lu等^[38]的研究同样发现,ESD术后出血与发生术中出血有关。与上述研究结果一致,本文研究发现,术中大出血是gGIST内镜手术术后出血的重要危险因素。本研究中,78例发生术后出血的患者中,有9.0%(7/78)的病例发生了术中大出血,而554例未发生术后出血的患者中仅1.3%(7/554)的病例发生了术中大出血,两组间比较差异有统计学意义($P<0.001$)。大量出血可能导致内镜视野不清,出血点难以定位或接近,尤其在操作技术不熟练时,这大大增加了止血的难度。因此,手术过程中应细致观察黏膜下层和瘤体周边的血管,对粗大血管提前予以热活检钳电凝处理。对于内镜止

血困难的病例,应及时求助经验丰富的医生,并在必要时中转外科手术。

本研究也存在不足:(1)本研究为回顾性研究,可能存在一定的选择偏倚;(2)本研究对3种内镜技术进行了综合分析,不同内镜技术的操作步骤存在差异,发生术后出血的风险可能也不同,未来可以针对这3种技术展开更为专门化的研究。然而,据笔者所知,本研究是目前探讨内镜切除gGIST术后出血风险中,纳入样本量最大的研究,这显著增强了统计结果的稳定性和可靠性。同时,笔者应用了先进的机器学习算法来构建预测模型,这些算法能够处理复杂的非线性关系并自动提取数据中的关键特征,从而提高了预测的准确性和模型的可靠性。此外,本文模型得到了良好的预测结果,这进一步验证了模型的有效性和实用性。最后,本文研究综合考虑了多种可能影响内镜切除gGIST术后出血的因素,提供了更全面的分析。

综上所述,本研究基于不同机器学习算法分析影响gGIST内镜手术术后出血的相关危险因素,并构建了多种预测模型。笔者发现,术者经验、手术时间和术中大出血是影响gGIST内镜手术术后出血的重要危险因素。应用SMOTE-NC过采样及机器学习算法构建的模型中,GBM模型的性能优于其他预测模型。此外,网络应用程序的开发方便了模型在临床实践中的应用。

【参考文献】

- [1] Nishida T, Goto O, Raut CP, et al. Diagnostic and treatment strategy for small gastrointestinal stromal tumors[J]. *Cancer*, 2016, 122(20): 3110-3118.
- [2] Chen HM, Li BW, Li LY, et al. Current status of endoscopic resection of gastric subepithelial tumors[J]. *Am J Gastroenterol*, 2019, 114(5): 718-725.
- [3] Kelly CM, Gutierrez Sainz L, Chi P. The management of metastatic GIST: current standard and investigational therapeutics[J]. *J Hematol Oncol*, 2021, 14(1): 2.
- [4] Meng X, Hu YF, Mao XL, et al. Clinical outcomes of endoscopic versus laparoscopic resection for senile patients with gastric gastrointestinal stromal tumours (2 to 4 cm) originating from the muscularis propria layer[J]. *Prz Gastroenterol*, 2022, 17(3): 196-202.
- [5] Liu S, Zhou XX, Yao YX, et al. Resection of the gastric submucosal tumor (G-SMT) originating from the muscularis propria layer: comparison of efficacy, patients' tolerability, and clinical outcomes between endoscopic full-thickness resection and surgical resection[J]. *Surg Endosc*, 2020, 34(9): 4053-4064.
- [6] An W, Sun PB, Gao J, et al. Endoscopic submucosal dissection for gastric gastrointestinal stromal tumors: a retrospective cohort study[J]. *Surg Endosc*, 2017, 31(11): 4522-4531.
- [7] Kanehira E, Kamei A, Umezawa A, et al. Long-term outcomes of percutaneous endoscopic intragastric surgery in the treatment of gastrointestinal stromal tumors at the esophagogastric junction[J]. *Surg Endosc*, 2016, 30(5): 2036-2042.
- [8] Zhang Q, Gao LQ, Han ZL, et al. Effectiveness and safety of endoscopic resection for gastric GISTs: a systematic review[J]. *Minim Invasive Ther Allied Technol*, 2018, 27(3): 127-137.

- [9] Andalib I, Yeoun D, Reddy R, et al. Endoscopic resection of gastric gastrointestinal stromal tumors originating from the muscularis propria layer in North America: methods and feasibility data[J]. *Surg Endosc*, 2018, 32(4): 1787-1792.
- [10] Zhang Y, Mao XL, Zhou XB, et al. Long-term outcomes of endoscopic resection for small (≤ 4.0 cm) gastric gastrointestinal stromal tumors originating from the muscularis propria layer[J]. *World J Gastroenterol*, 2018, 24(27): 3030-3037.
- [11] Cai XL, Li XY, Liang C, et al. Endoscopic or laparoscopic resection for small gastrointestinal stromal tumors: a cumulative meta-analysis[J]. *Chin Med J (Engl)*, 2020, 133(22): 2731-2742.
- [12] 黄剑平, 张会禄, 刘懿, 等. 胃间质瘤内镜黏膜下剥离术后出血危险因素的研究[J]. *中国肿瘤临床与康复*, 2019, 26(6): 656-658. Huang JP, Zhang HL, Liu Y, et al. Risk factors for postoperative hemorrhage in patients undergoing endoscopic submucosal dissection for gastric stromal tumors[J]. *Chinese Journal of Clinical Oncology and Rehabilitation*, 2019, 26(6): 656-658.
- [13] 陈冬梅, 王红建. 内镜黏膜下剥离术治疗胃间质瘤出现并发症的可能危险因素分析[J]. *临床消化病杂志*, 2021, 33(6): 406-410. Chen DM, Wang HJ. Analysis of risk factors for complications in endoscopic submucosal dissection treatment of gastric stromal tumors[J]. *Chinese Journal of Clinical Gastroenterology*, 2021, 33(6): 406-410.
- [14] 李云, 刘超, 刘琳, 等. 胃间质瘤经黏膜下隧道内镜肿瘤切除术术后出血的危险因素分析[J]. *局解手术学杂志*, 2021, 30(6): 502-505. Li Y, Liu C, Liu L, et al. Risk factors for postoperative bleeding after submucosal tunneling endoscopic resection of gastric stromal tumor[J]. *Journal of Regional Anatomy and Operative Surgery*, 2021, 30(6): 502-505.
- [15] 张媛, 林金灿, 栗华. 内镜黏膜下剥离术治疗胃间质瘤的临床效果及术后出血的危险因素分析[J]. *中国内镜杂志*, 2021, 27(8): 33-38. Zhang Y, Lin JC, Li H. Clinical efficacy of endoscopic submucosal dissection in the treatment of gastric stromal tumor and risk factors of postoperative hemorrhage[J]. *China Journal of Endoscopy*, 2021, 27(8): 33-38.
- [16] Liu LJ, Ye Y, Wang QH, et al. Risk factors for postoperative complications in endoscopic resection of gastric gastrointestinal stromal tumors: a multi-center analysis[J]. *Surg Endosc*, 2023, 37(9): 6844-6851.
- [17] Leite D, Martins A Jr, Rativa D, et al. An automated machine learning approach for real-time fault detection and diagnosis[J]. *Sensors (Basel)*, 2022, 22(16): 6138.
- [18] Puri M. Automated machine learning diagnostic support system as a computational biomarker for detecting drug-induced liver injury patterns in whole slide liver pathology images[J]. *Assay Drug Dev Technol*, 2020, 18(1): 1-10.
- [19] Joensuu H. Risk stratification of patients diagnosed with gastrointestinal stromal tumor[J]. *Hum Pathol*, 2008, 39(10): 1411-1419.
- [20] Du C, Linghu EQ. Submucosal tunneling endoscopic resection for the treatment of gastrointestinal submucosal tumors originating from the muscularis propria layer[J]. *J Gastrointest Surg*, 2017, 21(12): 2100-2109.
- [21] Saito I, Tsuji Y, Sakaguchi Y, et al. Complications related to gastric endoscopic submucosal dissection and their managements[J]. *Clin Endosc*, 2014, 47(5): 398-403.
- [22] Goto O, Fujishiro M, Oda I, et al. A multicenter survey of the management after gastric endoscopic submucosal dissection related to postoperative bleeding[J]. *Dig Dis Sci*, 2012, 57(2): 435-439.
- [23] 中华医学会消化内镜学分会, 中国抗癌协会肿瘤内镜专业委员会. 中国早期胃癌筛查及内镜诊治共识意见(2014年, 长沙)[J]. *中华消化内镜杂志*, 2014, 31(7): 361-377. Chinese Society of Digestive Endoscopy, The Society of Oncological Endoscopy of Chinese Anti-Cancer Association. Consensus on the screening, endoscopic diagnosis and treatment of early gastric cancer in China[J]. *Chinese Journal of Digestive Endoscopy*, 2014, 31(7): 361-377.
- [24] Sun Y, Liu LJ, Shi DT, et al. Assessment of a real-world learning curve for the endoscopic resection of gastric gastrointestinal stromal tumors[J]. *J Int Med Res*, 2023, 51(8): 3000605231194448.
- [25] Gök EC, Olgun MO. SMOTE-NC and gradient boosting imputation based random forest classifier for predicting severity level of Covid-19 patients with blood samples[J]. *Neural Comput Appl*, 2021, 33(22): 15693-15707.
- [26] Chawla NV, Bowyer KW, Hall LO, et al. SMOTE: synthetic minority over-sampling technique[J]. *J Artif Intell Res*, 2002, 16(1): 321-357.
- [27] Chen Y, Liu XJ, Gao L, et al. Using the H2O automatic machine learning algorithms to identify predictors of Web-Based medical record nonuse among patients in a data-rich environment: mixed methods study[J]. *JMIR Med Inform*, 2023, 11: e41576.
- [28] van der Gaag M, Hoffman T, Remijns M, et al. The five-factor model of the Positive and Negative Syndrome Scale II: a ten-fold cross-validation of a revised model[J]. *Schizophr Res*, 2006, 85(1/3): 280-287.
- [29] Akahoshi K, Oya M, Koga T, et al. Current clinical management of gastrointestinal stromal tumor[J]. *World J Gastroenterol*, 2018, 24(26): 2806-2817.
- [30] 王梅英, 杨敏, 刘佳微, 等. 基于SMOTE算法的化疗肿瘤患者下呼吸道感染预警模型构建[J]. *中国感染控制杂志*, 2021, 20(12): 1094-1101. Wang MY, Yang M, Liu JW, et al. Construction of early warning model of lower respiratory tract infection in chemotherapy tumor patients based on SMOTE algorithm[J]. *Chinese Journal of Infection Control*, 2021, 20(12): 1094-1101.
- [31] Yuan H, Xu HZ. Deep multi-modal fusion network with gated unit for breast cancer survival prediction[J]. *Comput Methods Biomech Biomed Engin*, 2024, 27(7): 883-896.
- [32] Hassoun S, Bruckmann C, Ciardullo S, et al. Setting up of a machine learning algorithm for the identification of severe liver fibrosis profile in the general US population cohort[J]. *Int J Med Inform*, 2023, 170: 104932.
- [33] Gozukara Bag HG, Yagin FH, Gormez Y, et al. Estimation of obesity levels through the proposed predictive approach based on physical activity and nutritional habits[J]. *Diagnostics (Basel)*, 2023, 13(18): 2949.
- [34] Toyokawa T, Inaba T, Omote S, et al. Risk factors for perforation and delayed bleeding associated with endoscopic submucosal dissection for early gastric neoplasms: analysis of 1123 lesions[J]. *J Gastroenterol Hepatol*, 2012, 27(5): 907-912.
- [35] 芦迅达, 冯志杰, 纪晨光, 等. 早期胃癌内镜黏膜下剥离术后出血的危险因素分析[J]. *临床荟萃*, 2021, 36(12): 1087-1091. Lu XD, Feng ZJ, Ji CG, et al. Analysis on risk factors of postoperative hemorrhage of endoscopic submucosal dissection in early gastric cancer[J]. *Clinical Focus*, 2021, 36(12): 1087-1091.
- [36] 胡康, 周永强, 邹百仑, 等. 早期胃癌患者内镜黏膜下剥离术后出血风险评估及预测模型的建立[J]. *中国医刊*, 2023, 58(2): 156-160. Hu K, Zhou YQ, Zou BC, et al. Assessment and prediction model of bleeding risk of patients with early gastric cancer after endoscopic submucosal dissection[J]. *Chinese Journal of Medicine*, 2023, 58(2): 156-160.
- [37] Suzuki S, Chino A, Kishihara T, et al. Risk factors for bleeding after endoscopic submucosal dissection of colorectal neoplasms[J]. *World J Gastroenterol*, 2014, 20(7): 1839-1845.
- [38] Lu XY, Ma J. Risk factors of delayed post-polypectomy bleeding after treatment of gastrointestinal polyps with snare-assisted endoscopic sub-mucosal dissection[J]. *J Minim Access Surg*, 2023, 19(2): 272-277.

(编辑: 薛泽玲)