

# 基于全面触发工具与机器学习的贝伐珠单抗严重不良反应预测模型研究<sup>Δ</sup>

符永妃<sup>1,2\*</sup>, 龙 芯<sup>1,2</sup>, 徐宏珍<sup>2</sup>, 唐 健<sup>2</sup>, 利向晴<sup>3</sup>, 龙钰程<sup>2</sup>, 覃 东<sup>1,2#</sup>(1. 桂林医科大学药学院, 广西 桂林 541199; 2. 桂林市人民医院药学部, 广西 桂林 541002; 3. 桂林市人民医院数据中心, 广西 桂林 541002)

中图分类号 R979.1; R969.3; TP181

文献标志码 A

文章编号 1001-0408(2026)04-0497-07

DOI 10.6039/j.issn.1001-0408.2026.04.15



**摘要** **目的** 确定贝伐珠单抗相关不良反应(ADR)的触发器条目,判定并分析相关ADR发生情况,构建该药致严重不良反应(SAR)的预测模型。**方法** 基于全面触发工具(GTT)理念,参考GTT白皮书、药品说明书及相关文献,结合单轮德尔菲法确定贝伐珠单抗相关ADR触发器条目;利用所建条目,基于中国医院药物警戒系统对桂林市人民医院2020年1月至2024年9月相关患者的电子病历进行主动筛查,由药师判定并统计贝伐珠单抗致ADR的发生情况。以所得触发器条目阳性患者的资料为对象(按7:3划分训练集和测试集),通过Boruta算法从训练集相关39项变量中筛选候选特征变量,以是否发生SAR为因变量进行多因素Logistic回归分析;以上述候选特征变量为基础,构建Logistic回归、极端梯度提升、轻量级梯度提升机、随机森林、类别特征提升模型,通过受试者操作特征曲线的曲线下面积(AUC)、召回率等指标评估模型性能,应用Shapley加性解释(SHAP)法分析解释各变量的贡献,并基于最优模型构建列线图。**结果** 最终确定用于贝伐珠单抗相关ADR主动监测的触发器条目38项,含检验指标17项、临床表现13项、干预措施8项。共纳入触发器条目阳性患者483例,其中发生ADR者318例,发生SAR者83例;触发器条目和病例阳性预测值分别为43.57%(708/1 625)和63.84%(318/483)。贝伐珠单抗致ADR涉及7个系统/器官,以血液系统受累最为常见(64.15%)。经Boruta算法共筛选到血钾、红细胞压积、白/球蛋白比值、前白蛋白、既往高血压史、年龄、红细胞计数7个候选变量。多因素Logistic回归分析显示,血钾水平升高(OR=0.234, P=0.002)与贝伐珠单抗致SAR风险降低相关,既往高血压史(OR=2.642, P=0.006)和年龄增加(OR=1.040, P=0.025)与贝伐珠单抗致SAR风险升高有关。Logistic回归模型的AUC、F1值、召回率(0.761、0.447、0.607)均高于其他模型;SHAP评估结果显示,血钾、红细胞压积、年龄等变量的重要性位居前列。**结论** 成功确定38项触发器条目用于贝伐珠单抗相关ADR的主动筛查。血钾水平升高是贝伐珠单抗致SAR的保护因素,而既往高血压史、年龄增加则是危险因素;Logistic回归模型为贝伐珠单抗致SAR的最优预测模型。

**关键词** 贝伐珠单抗;药品不良反应;全面触发工具;机器学习;预测模型

## Predictive model for severe adverse reaction associated with bevacizumab based on the global trigger tool and machine learning

FU Yongfei<sup>1,2</sup>, LONG Xin<sup>1,2</sup>, XU Hongzhen<sup>2</sup>, TANG Jian<sup>2</sup>, LI Xiangqing<sup>3</sup>, LONG Yucheng<sup>2</sup>, QIN Dong<sup>1,2</sup>(1. College of Pharmacy, Guilin Medical University, Guangxi Guilin 541199, China; 2. Dept. of Pharmacy, Guilin People's Hospital, Guangxi Guilin 541002, China; 3. Data Center, Guilin People's Hospital, Guangxi Guilin 541002, China)

**ABSTRACT** **OBJECTIVE** To confirm trigger items for adverse drug reaction (ADR) induced by bevacizumab, to identify and analyze the occurrence of related ADR, and to establish a predictive model for severe adverse reaction (SAR) caused by this drug. **METHODS** Based on the global trigger tool (GTT) theory, and referencing the GTT White Paper, drug package inserts and relevant literature, trigger items for bevacizumab-related ADR were confirmed using a single-round Delphi method. Utilizing these established items, electronic medical records of relevant patients at Guilin People's Hospital from January 2020 to September 2024 were actively screened via the China Hospital Pharmacovigilance System. Pharmacists then identified and tallied the occurrence of bevacizumab-induced ADR. Data from patients with any positive trigger item served as the study subjects (divided into training and test sets at a ratio of 7:3), candidate feature variables were selected from 39 related variables using the Boruta algorithm, and the multivariable Logistic regression analysis was performed with the occurrence of SAR as the dependent variable. Based on these candidate features, Logistic Regression, Extreme Gradient Boosting, Light Gradient

<sup>Δ</sup>基金项目 桂林市科学研究与技术开发计划(No.20210227-10-4, No.20230135-9-7);中国医药卫生事业发展基金会医学科研课题;北京康盟慈善基金会卫生健康发展促进项目-麻醉与重症科研项目(No. MZ-20240922-01);国家卫生健康委医院管理研究所医院药学高质量发展研究课题(智慧药学专项)立项课题(No.NIHAYSZX2543);桂林医科大学校级研究生创新项目(No.GYYK2025016)

\* 第一作者 硕士研究生。研究方向:临床药学。E-mail: amaocie@163.com

# 通信作者 主任药师,硕士生导师。研究方向:药事管理、临床药学。E-mail:68321376@qq.com

Boosting Machine, Random Forest, and Categorical Boosting models were constructed. Model performance was evaluated using metrics including the area under the curve (AUC) of receiver operating characteristic curve and recall rate. The Shapley Additive exPlanations (SHAP) method was applied to analyze and interpret the contribution of each variable. A nomogram was constructed based on the optimal model. **RESULTS** A total of 38 trigger items for active monitoring of bevacizumab-related ADR were determined, comprising 17 laboratory indicators, 13 clinical manifestations, and 8 intervention measures. In total, 483 patients with positive trigger items were included, and 318 patients with bevacizumab-induced ADR were identified, including 83 SARs. The positive predictive values for the trigger items and cases were 43.57% (708/1 625) and 63.84% (318/483), respectively. Bevacizumab-induced ADR involved 7 systems/organs, with the hematological system being the most frequently involved (64.15%). The Boruta algorithm selected 7 variables; serum potassium, hematocrit, albumin-to-globulin ratio, prealbumin, hypertension history, age and red blood cell count. Multivariable Logistic regression showed that elevated serum potassium levels were associated with a decreased risk of bevacizumab-induced SAR (OR=0.234,  $P=0.002$ ), while a history of hypertension (OR=2.642,  $P=0.006$ ) and increased age (OR=1.040,  $P=0.025$ ) were associated with an increased risk. The Logistic Regression model demonstrated superior performance with higher AUC, F1 score and recall rate (0.761, 0.447, 0.607), compared to other models. SHAP evaluation results indicated that variables such as serum potassium, hematocrit, and age ranked highest in importance. **CONCLUSIONS** Totally 38 trigger entries have been successfully identified for active screening of bevacizumab-related ADR. Elevated serum potassium levels are a protective factor against bevacizumab-induced SAR, whereas the hypertension history and increased age are risk factors. The Logistic Regression model is the optimal predictive model.

**KEYWORDS** bevacizumab; adverse drug reaction; global trigger tool; machine learning; predictive model

癌症是导致疾病发生和死亡的重要原因。在全球范围内,我国新发癌症及死亡病例数的占比均较为突出,给卫生系统和患者家庭造成了沉重的经济负担<sup>[1-2]</sup>。近年来,分子靶向治疗发展迅速,贝伐珠单抗(bevacizumab)作为代表性抗血管生成单克隆抗体,可特异性结合血管内皮生长因子A,阻断其与血管内皮生长因子受体2结合,进而抑制肿瘤细胞新生血管形成,有效延缓癌症进展<sup>[3]</sup>。然而,贝伐珠单抗可引发蛋白尿、出血、消化道穿孔等严重不良反应(severe adverse reaction, SAR),其机制主要涉及血管舒张障碍、肾脏滤过屏障受损以及血管修复功能减弱<sup>[4]</sup>;同时,贝伐珠单抗相关SAR还可能导致患者治疗中断或住院时间延长,亟须临床加强识别。

目前,药物不良反应(adverse drug reaction, ADR)监测主要依赖自发上报系统,但该系统存在上报率低、信息不完整、易遗漏等不足<sup>[5]</sup>。全面触发工具(global trigger tool, GTT)作为可主动识别ADR的工具之一<sup>[6]</sup>,已广泛应用于抗真菌药、抗肿瘤药物的相关研究<sup>[6,7]</sup>,但其用于如贝伐珠单抗等具体药物安全性监测的报道相对缺乏。随着电子病历与临床大数据的积累,机器学习技术在ADR监测、安全风险评估方面展现出较高的建模灵活性和良好的预测性能,可在传统统计学分析基础上进一步强化临床对高危患者的识别能力。基于此,本研究首先结合GTT理念确定贝伐珠单抗相关ADR触发器条目,据此筛选触发器条目阳性的患者病历;随后,通过人工审核判定ADR并描述其临床特征;最后,挖掘贝伐珠单抗致SAR的危险/保护因素,构建并优选机器学习模型,将最优模型以列线图形式呈现,以期提高临床对该药相关ADR的早期识别及风险评估能力,为临床优化治疗方案提供参考。

## 1 资料与方法

### 1.1 触发器条目确定

依据GTT白皮书、药品说明书及相关文献<sup>[8-11]</sup>,结合桂林市人民医院(以下简称“我院”)实际情况,本研究初步拟定45项触发器条目,涵盖检验指标、临床表现、干预措施三大类。通过单轮德尔菲法就上述初拟条目向8名专家进行函询,并采用专家打分法计算各条目的重要性均值( $M_j$ )、专家权威系数( $C_j$ )、变异系数( $V_j$ ),以评估专家意见的一致性与可靠性<sup>[12]</sup>。最终以 $V_j > 0.35$ 且 $M_j < 3.5$ 为标准<sup>[12]</sup>,确定贝伐珠单抗相关ADR触发器条目。

### 1.2 病例筛选

本研究采用回顾性研究方法,在中国医院药物警戒系统(China Hospital Pharmacovigilance System, CHPS)中检索2020年1月1日至2024年9月10日于我院住院并使用贝伐珠单抗患者的相关记录。具体操作如下:依据“1.1”项下确定的触发器条目,药师在CHPS的“药品评价系统”模块中完成规则配置及关键字段术语匹配;随后,CHPS基于医嘱、病程及检验信息自动筛查、关联并导出相关病历资料。本研究患者的纳入标准包括:(1)接受贝伐珠单抗治疗的所有住院患者;(2)年龄 $\geq 18$ 岁;(3)触发器条目阳性;(4)住院时间 $\geq 24$  h。本研究患者的排除标准包括:(1)主要住院信息缺失者;(2)经核实未使用贝伐珠单抗者。本研究方案已获我院伦理委员会批准(批准号为2021-033KY)。

### 1.3 ADR判定与分级

针对病历资料,由2名初级职称药师判断是否发生ADR;若存在分歧,则由高级职称药师复核裁决。具体操作如下:依据WHO-乌普萨拉监测中心制定的ADR判定标准,对可疑ADR(触发器条目阳性)与贝伐珠单抗的因果关系(包括“肯定”“很可能”“可能”“不太可能”“待定”“无法评价”)进行评估<sup>[13]</sup>;随后,结合不良事件通用

术语标准5.0版对ADR的严重程度进行分级,将3~5级不良反应定义为SAR(3级为严重但可处理的ADR,4级为危及生命的ADR,5级为导致死亡的ADR)<sup>[14]</sup>;同时,参照现行《药品不良反应报告和监测管理办法》分析ADR(新的、一般的ADR和SAR)的累及系统/器官<sup>[15]</sup>。

#### 1.4 SAR预测模型构建与优选

基于既往研究<sup>[9-11]</sup>,本研究以贝伐珠单抗相关SAR的发生情况为因变量,选取39项指标作为模型自变量,包括——(1)基本信息:性别、年龄、体重、身高;(2)既往病史:高血压史、糖尿病史、脑梗史、过敏史、吸烟史、手术史、饮酒史;(3)治疗暴露:抗肿瘤药物品种数;(4)实验室指标:①血常规相关指标:白细胞计数、红细胞计数、血小板计数、淋巴细胞计数、单核细胞计数、嗜酸性粒细胞绝对值、嗜碱性粒细胞绝对值、中性粒细胞绝对值、红细胞压积、大血小板比率;②电解质与酸碱指标:血钠、血氯、血钾、血钙、碳酸氢盐;③肝肾功能与胆红素分型指标:天冬氨酸转氨酶、丙氨酸转氨酶、谷氨酰转氨酶、碱性磷酸酶、总胆红素、直接胆红素、间接胆红素、尿素、血肌酐;④蛋白与营养指标:总蛋白、前白蛋白、白/球蛋白比值。

将数据集(触发器条目阳性患者的数据资料)按7:3随机划分为训练集和测试集。采用Boruta算法对训练集患者的39项指标进行特征筛选,获得候选特征变量,作为后续建模的统一特征集。以该特征集为自变量、是否发生SAR为因变量在训练集上进行多因素Logistic回归分析,以评估各变量与SAR的独立相关性,并用于识别危险/保护因素。随后,仍以上述特征集为建模范征集,分别构建Logistic回归、极端梯度提升(Extreme Gradient Boosting, XGBoost)、轻量级梯度提升机(Light Gradient Boosting Machine, LightGBM)、随机森林、类别特征提升(Categorical Boosting, CatBoost)预测模型;针对变量类别不平衡问题,在训练集中应用合成少数类过采样技术(Synthetic Minority Oversampling Technique, SMOTE)采样进行过采样处理,同时通过十折交叉验证结合网格搜索的方法优化超参数,以确保模型的泛化能力并避免过拟合。在测试集上以准确率、召回率、F1值、精确度、特异性、受试者操作特征(receiver operating characteristic, ROC)曲线下面积(area under the curve, AUC)为评估指标,从上述5种模型中选取性能最优的模型用于预测SAR的发生风险,并使用Shapley加性解释(Shapley Additive exPlanations, SHAP)法在测试集进行变量可解释性分析。最后,仍以上述特征集为变量构建列线图,对最优模型进行可视化,以便临床进行个体风险评估。

#### 1.5 统计学分析

采用R 4.4.2软件进行数据处理,采用决策链1.0软件建模并进行相关分析。分类变量以频数或率表示,组间比较采用 $\chi^2$ 检验或Fisher精确概率检验;连续变量以 $\bar{x} \pm s$ 、中位数(范围)表示,组间比较采用 $t$ 检验或Mann-Whitney  $U$ 检验。检验水准 $\alpha=0.05$ 。

## 2 结果

### 2.1 基于GTT的触发器条目

本课题组共发放问卷8份,问卷回收率为100%。参与函询的8名专家年龄集中于40~50岁,工作年限以16~30年为主(37.50%),职称以副高级为主(50.00%)。函询结果显示,所有条目的 $M_i$ 均超过3.5分, $C_i$ 为0.75~0.85, $V_j$ 为0.34,表明专家意见一致性、可靠性较高。按 $V_j > 0.35$ 且 $M_i < 3.5$ 的标准筛选,删除“纳差”“便秘”等条目后,最终确定用于预测贝伐珠单抗相关ADR的触发器条目共38项,包含检验指标17项、临床表现13项、干预措施8项,具体见表1。

### 2.2 纳入患者的基本信息

本研究最终纳入患者483例,其中男性247例、女性236例;年龄为(61.82±10.81)岁;住院时间为6(1~46)d,最常见的慢性共病为高血压(137例,28.36%),44.72%的患者有手术史(216例);恶性肿瘤类型以结直肠癌(160例,33.13%)、肺癌(84例,17.39%)为主。

### 2.3 纳入患者的触发器条目检出情况

在483例患者病历中,共检出触发器条目阳性1625条;其中,经药师判定与贝伐珠单抗致ADR相关的有708条,涉及患者318例,病历整体触发器条目阳性预测值为43.57%(708/1625)。

由纳入患者触发器阳性检出情况(表1)可知,在各类触发器条目中,实验室指标类触发器条目的阳性预测值普遍较高,其中“血钠<136 mmol/L”(95.29%)、“中性粒细胞百分比<40%”(89.89%)、“天冬氨酸转氨酶或丙氨酸转氨酶>2倍ULN或碱性磷酸酶和总胆红素同时升高,其中任一项>2倍ULN”(87.95%)等条目表现出较强的识别效能。临床表现类触发器条目中,“骨髓抑制”的阳性预测值为45.16%,提示其虽为非特异性症状,但仍对贝伐珠单抗相关ADR的识别具有参考价值。干预措施类触发器条目中,“人血白蛋白”“复方甘草酸苷”的阳性预测值分别为28.57%、20.08%,提示部分药物干预可作为贝伐珠单抗相关ADR的潜在安全信号条目。

### 2.4 患者ADR发生情况

#### 2.4.1 ADR类型分布

纳入的483例患者中,318例(65.84%)确定发生了ADR。这318例中,83例患者(26.10%)发生SAR,其主要临床表现包括重度低钾血症(65例,78.31%)、出血相关事件(33例,39.76%)、感染(20例,24.10%)、蛋白尿(7例,8.43%)、严重低蛋白血症(6例,7.23%)、肝胆功能损害(2例,2.41%)、肾功能损害(2例,2.41%)及心力衰竭(1例,1.20%);235例患者(73.90%)发生一般的ADR,主要表现为轻度至中度的实验室检查异常或临床表现,如轻微电解质紊乱、短暂的转氨酶升高等。本研究纳入患者并未发生新的ADR。

#### 2.4.2 ADR累及系统/器官分布

ADR主要累及的7个系统/器官中,血液系统ADR的发生例次最多,而感染类和神经系统ADR的发生例次相对较少。结果见表2。

表1 触发器条目及纳入患者的阳性检出情况

类别	编号	触发器条目	阳性触发 ADR对应的 阳性预测值/		
			频次	阳性触发频次	%
实验室指标	L1	血红蛋白<90 g/L	200	58	29.00
	L2	中性粒细胞百分比<40%	89	80	89.89
	L3	肾小球滤过率<47 mL/(min·1.73 m <sup>2</sup> )	50	33	66.00
	L4	促甲状腺激素<0.33 mU/L或总甲状腺素>224 nmol/L	302	16	5.30
	L5	促甲状腺激素>8.6 mU/L或总甲状腺素<45 nmol/L	406	28	6.90
	L6	天冬氨酸转氨酶或丙氨酸转氨酶>2倍ULN或碱性磷酸酶和总胆红素同时升高,其中任一项>2倍ULN	83	73	87.95
	L7	高敏肌钙蛋白>14 ng/L	38	0	0
	L8	尿酸>420 μmol/L	140	123	87.86
	L9	血小板计数<125×10 <sup>9</sup> L <sup>-1</sup>	275	85	30.91
	L10	尿蛋白阳性	344	84	24.42
	L11	B型钠尿肽前体>100 pg/mL	13	11	84.62
	L12	血钙<2.11 mmol/L	104	62	59.62
	L13	血钙>2.52 mmol/L	47	28	59.57
	L14	血钾<3.2 mmol/L	136	110	80.88
	L15	血钠<136 mmol/L	191	182	95.29
	L16	空腹葡萄糖≥6.11 mmol/L	78	57	73.08
	L17	白细胞计数<3.5×10 <sup>9</sup> L <sup>-1</sup>	403	291	72.21
临床表现	S1	乏力	701	107	15.26
	S2	咳嗽	482	56	11.62
	S3	上呼吸道感染	63	12	19.05
	S4	腹腔感染	7	1	14.29
	S5	肺部感染	130	13	10.00
	S6	便秘	-	-	-
	S7	纳差	-	-	-
	S8	恶心/呕吐	-	-	-
	S9	电解质紊乱	82	29	35.37
	S10	低蛋白血症	143	42	29.37
	S11	手足综合征	-	-	-
	S12	皮炎/瘙痒	-	-	-
	S13	头晕	26	7	26.92
	S14	头痛	106	11	10.38
	S15	腰痛	-	-	-
	S16	关节炎	-	-	-
	S17	骨髓抑制	1 063	480	45.16
S18	贫血	988	139	14.07	
S19	高血压	702	27	3.85	
S20	出血	1 226	41	3.34	
干预措施	A1	亚叶酸钙	156	26	16.67
	A2	人血白蛋白	35	10	28.57
	A3	地榆升白片	256	42	16.41
	A4	甲氧氯普胺	670	80	11.94
	A5	多潘立酮	5	2	40.00
	A6	昂丹司琼	449	79	17.59
	A7	复方甘草酸苷	244	49	20.08
	A8	异甘草酸镁	128	10	7.81

注:阳性预测值=ADR对应的阳性触发频次/阳性触发频次×100%;ULN为正常值上限;“-”为已剔除条目。

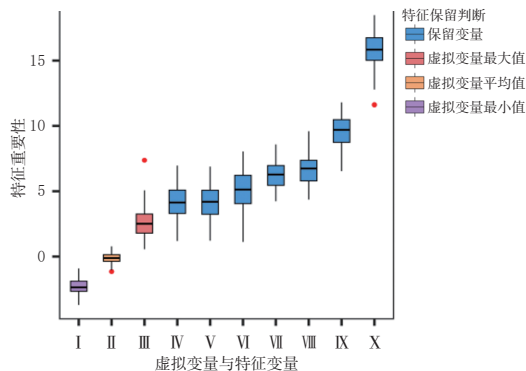
表2 ADR累及系统/器官分布情况

累及系统/器官	临床表现(例次)	总例次	占比/%
血液系统	骨髓抑制(480)、贫血(139)	619	64.15
全身性症状	乏力(107)	107	11.09
代谢及营养系统	低蛋白血症(42)、电解质紊乱(29)	71	7.36
心血管系统	出血(41)、高血压(27)	68	7.05
呼吸系统	咳嗽(56)	56	5.80
感染类	肺部感染(13)、呼吸道感染(12)、腹腔感染(1)	26	2.69
神经系统	头痛(11)、头晕(7)	18	1.87

## 2.5 SAR预测模型

### 2.5.1 统一特征集

本研究通过 Boruta 算法从训练集中共筛选出候选特征变量7个,分别为血钾、红细胞压积、白/球蛋白比值、前白蛋白、既往高血压史、年龄、红细胞计数,将其作为后续危险/保护因素识别和机器学习预测模型及列线图构建的统一特征集。结果见图1。



I: 虚拟变量最小值; II: 虚拟变量平均值; III: 虚拟变量最大值; IV: 红细胞计数; V: 年龄; VI: 既往高血压史; VII: 前白蛋白; VIII: 白/球蛋白比值; IX: 红细胞压积; X: 血钾。

图1 Boruta算法确定的候选特征变量的箱线图

### 2.5.2 危险/保护因素

多因素 Logistic 回归分析结果(表3)显示,血钾水平升高与贝伐珠单抗致 SAR 风险降低相关(OR=0.234, P=0.002),为保护因素(患者血钾升高 1 mmol/L, SAR 风险降低 76.6%);既往高血压史(OR=2.642, P=0.006)和年龄增加(OR=1.040, P=0.025)与贝伐珠单抗致 SAR 风险升高有关,为危险因素(有既往高血压史者的 SAR 风险是无高血压史者的 2.642 倍;患者年龄增加 1 岁, SAR 风险增加 4.0%);而红细胞压积、白/球蛋白比值、前白蛋白、红细胞计数与贝伐珠单抗致 SAR 的发生无关(P>0.05)。

表3 贝伐珠单抗致 SAR 发生风险的多因素 Logistic 回归分析结果

特征变量	回归系数	OR	95% 置信区间	P
血钾	-1.451	0.234	0.093~0.568	0.002
既往高血压史	0.972	2.642	1.328~5.294	0.006
年龄	0.039	1.040	1.006~1.077	0.025
红细胞压积	-0.091	0.913	0.818~1.017	0.099
白/球蛋白比值	-1.093	0.335	0.067~1.548	0.171
前白蛋白	-0.004	0.996	0.989~1.002	0.186
红细胞计数	-0.182	0.833	0.362~1.854	0.661

注:除既往高血压史为二分类变量(有=1,无=0)外,其余变量均为连续变量(取实际测量值);OR:比值比。

### 2.5.3 机器学习模型

本研究以“2.5.2”项下所得的7个候选特征变量为基础<sup>[16]</sup>,分别构建 Logistic 回归、XGBoost、LightGBM、随机森林、CatBoost 预测模型,并在测试集上开展性能评估,结果(图2、表4)显示,各模型的 AUC 为 0.721~0.761,处于可接受水平,提示不同算法构建的模型均具有可接受的区分度,且整体预测表现较为稳定<sup>[17]</sup>;其中, Logistic 回

归模型的AUC、F1值、召回率均高于其他模型,整体性能最优;LightGBM、XGBoost模型次之,二者在准确率、精确度、特异性方面略优于Logistic回归模型,但召回率、F1值稍差。

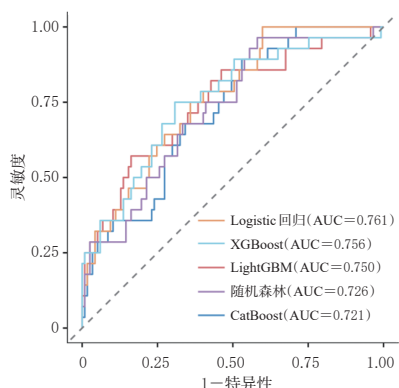


图2 不同预测模型的ROC曲线分析结果

表4 不同预测模型的性能评估结果

模型	准确率	召回率	F1值	精确度	特异性	AUC
Logistic回归	0.710	0.607	0.447	0.354	0.735	0.761
XGBoost	0.800	0.357	0.408	0.476	0.906	0.756
LightGBM	0.807	0.357	0.417	0.500	0.915	0.750
随机森林	0.731	0.393	0.361	0.333	0.812	0.726
CatBoost	0.793	0.357	0.400	0.455	0.897	0.721

此外,SHAP评估结果(图3)显示,血钾、红细胞压积、年龄等变量的重要性位居前列,提示以上变量在贝伐珠单抗致SAR发生风险预测中起关键作用<sup>[18]</sup>。

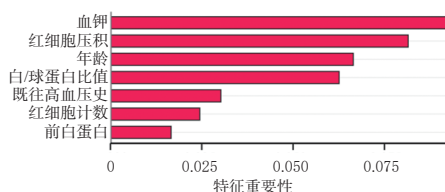


图3 基于SHAP法的候选特征变量的重要性排序图

### 2.5.4 列线图

基于最优的Logistic回归模型,本研究进一步构建了预测贝伐珠单抗致SAR发生风险的列线图,用于直观展示SAR的发生风险。该图(图4)包括红细胞计数、血钾、红细胞压积、前白蛋白、白/球蛋白比值、年龄、既往高血压史7个变量,可通过Logistic回归系数(表3)对各变量赋予0~100的连续积分,再将各变量的积分相加,即得总积分(总积分越高,提示患者发生SAR的风险越高)。

## 3 讨论

### 3.1 触发器条目分析

本研究以GTT理念为指导<sup>[5]</sup>,通过德尔菲法对初始构建的45项触发器条目进行修改和完善,最终形成了包含38项触发器条目的贝伐珠单抗相关ADR主动监测触发器体系,可协助临床发现该药的潜在ADR。函询专家评估结果显示,所有条目的 $M_j$ 均超过3.5分, $C_j$ 不低于0.75,表明专家意见的一致性较高。主动监测实践显示,

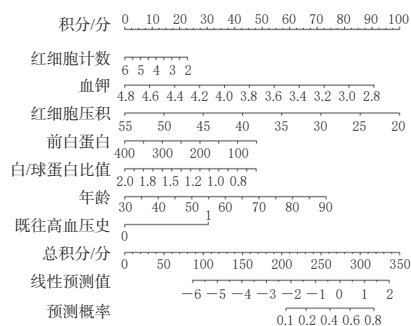


图4 贝伐珠单抗致SAR发生风险预测的列线图

上述触发器条目的阳性预测值也普遍较高,体现了专家意见与临床实践数据的一致性,初步验证了该触发器体系的可行性。

本研究构建的触发器体系包含实验室指标、临床表现、干预措施三大类,其中实验室指标类条目的整体识别效率最高,总阳性预测值为45.57%(1 321/2 899),明显高于临床表现类条目(16.87%,965/5 719)和干预措施类条目15.34%(298/1 943)。具体而言,“血钠<136 mmol/L”(95.29%)、“天冬氨酸转氨酶或丙氨酸转氨酶>2倍ULN或碱性磷酸酶和总胆红素同时升高,其中任一项>2倍ULN”(87.95%)以及骨髓抑制相关指标[“白细胞计数< $3.5 \times 10^9 L^{-1}$ ”(72.21%)、“中性粒细胞百分比<40%”(89.89%)、“血小板计数< $125 \times 10^9 L^{-1}$ ”(30.91%)、“血红蛋白<90 g/L”(29.00%)]在识别潜在ADR时表现突出,且这些指标与贝伐珠单抗的毒性机制高度相关<sup>[3,5]</sup>,提示其在贝伐珠单抗安全预警体系中具有较高的敏感性和较强的生物学解释力。本研究数据显示,贝伐珠单抗致ADR的发生率为65.84%(318/483),致SAR的发生率为17.18%(83/483),与Zhang等<sup>[17]</sup>研究结果相似;此外,相关ADR大多累及血液系统,主要表现包括骨髓抑制、贫血等,与实验室指标类条目阳性预测值较高这一特点相符。

相比之下,临床表现类条目的阳性预测值整体偏低,如乏力、咳嗽、头痛的阳性预测值均低于20%,提示这类非特异性症状在贝伐珠单抗相关ADR识别中的预测价值可能有限。需要注意的是,临床表现类条目在预警体系中主要承担“高灵敏度”的补充作用,更适用于贝伐珠单抗相关ADR的前期筛查和预警。Liu等<sup>[7]</sup>同样基于GTT理念构建了包含37项触发器条目的抗肿瘤药物致ADR主动监测体系,实践显示,血液系统相关ADR占比最高(74.6%),临床表现类触发器条目(如口腔炎、手足综合征、皮疹等)具有较高的阳性预测值(部分指标可达100%)。本研究中,骨髓抑制(45.16%)、电解质紊乱(35.37%)、低蛋白血症(29.37%)亦具有较高的阳性预测值;但由于本研究依据贝伐珠单抗相关ADR进行归因判定,乏力、咳嗽、头痛等非特异临床表现虽触发较多,最终归因比例却较低,故整体阳性预测值偏低,更偏向用于预警初筛。

本研究结果还显示,干预措施类条目的阳性预测值

普遍低于20%，这类条目反映医生的干预行为(如使用护肝、升白细胞等药物)，对ADR的识别依赖于医生是否将药物用于处理疑似ADR，而非用于基础疾病治疗或预防，因此该类条目的阳性预测值普遍较低。

此外，本研究中，阳性预测值低于10%的条目包括“高敏肌钙蛋白 $>14\text{ ng/L}$ ”“促甲状腺激素 $<0.33\text{ mU/L}$ 或总甲状腺素 $>224\text{ nmol/L}$ ”“高血压”“出血”等，尽管在专家函询时获得了较高的重要性评分(3.63~4.38分)，但其临床特异性不强，实际触发效果有限。

### 3.2 SAR影响因素分析

本研究采用Boruta算法从39项临床变量中筛选出7个候选特征变量。多因素Logistic回归分析结果显示，血钾水平较高( $OR=0.234, P=0.002$ )为贝伐珠单抗致SAR发生的保护因素，表明电解质稳态在该药安全风险防控中具有重要意义。Lima等<sup>[19]</sup>基于大样本住院人群的研究表明，血钾水平升高( $\geq 4.9\text{ mmol/L}$ )是固体剂型抗肿瘤药物致ADR发生的独立危险因素之一，与本研究结果有所不同，这可能与纳入人群构成、ADR类型差异有关，但二者均反映出钾离子异常与抗肿瘤药物致ADR高度相关，临床需重点关注。

多因素Logistic回归分析结果显示，既往有高血压病史( $OR=2.642, P=0.006$ )、年龄增加( $OR=1.040, P=0.025$ )是贝伐珠单抗致SAR发生的危险因素，提示相关SAR的发生可能与患者基础心血管状态存在关联。已有研究指出，在使用贝伐珠单抗治疗的转移性结直肠癌患者中，既往高血压史( $OR=3.30$ )及年龄 $\geq 60$ 岁( $OR=2.04$ )是此类患者高血压等ADR发生的独立危险因素<sup>[20]</sup>；另有研究指出，与中青年患者相比，老年患者(尤其是合并高血压病史等心血管高危因素的老年患者)在接受贝伐珠单抗等抗血管内皮生长因子药物治疗后，更易出现高血压、蛋白尿等ADR，临床需进行更为严格的评估和监护<sup>[21]</sup>。贝伐珠单抗可通过抑制血管内皮生长因子信号通路而引发血管内皮功能障碍、微血管稀疏和血流动力学失衡<sup>[22]</sup>，这可能是该药心血管毒性发生的潜在机制。

### 3.3 预测模型的构建及评估

本研究基于经Boruta算法筛选得到的7个候选特征变量，并基于此分别构建了5种机器学习模型以预测贝伐珠单抗相关SAR的发生风险。其中，Logistic回归作为广义线性模型，可通过逻辑函数将输出映射至0~1区间，适用于二元分类变量<sup>[23]</sup>；XGBoost和LightGBM基于梯度提升树算法，能够在大规模、高维度或稀疏特征数据中有效捕捉复杂的非线性关系及高阶交互作用<sup>[24]</sup>；随机森林模型是基于决策树的集成学习算法，其运行逻辑是通过构建多棵基于随机样本和随机特征的决策树，提高模型的准确性和鲁棒性<sup>[24]</sup>；CatBoost则具有较高的精确度和可扩展性，尤其适用于包含多类别变量的大规模、高维度数据集<sup>[25]</sup>。

本研究通过综合评估测试集的AUC、F1值、召回率

等关键指标发现，Logistic回归模型的各项评估指标相对均衡，整体表现最佳，在高灵敏度场景下更具实用性；XGBoost和LightGBM虽在AUC上与Logistic回归模型较为接近，但在召回率、F1值等综合指标方面并未表现出明显优势。笔者分析认为，Logistic回归模型优于其余几种机器学习模型的原因可能包括以下几点：(1)本研究中，SAR阳性事件数相对有限(83/483)，限制了复杂机器学习模型在避免过拟合前提下充分挖掘非线性关系的能力；(2)本研究纳入的预测变量与结果(SAR是否发生)之间的关系多呈线性或可加性，削弱了非线性模型的优势，更符合Logistic回归的建模假设；(3)相较于其余模型，Logistic回归模型的可解释性有助于特征筛选和临床验证，从而确保了纳入的预测因子具有临床相关性<sup>[23]</sup>。

### 3.4 列线图的临床应用价值

本研究将最优模型(Logistic回归模型)以列线图形式呈现，可整合多项实验室指标与一般临床特征，便于临床计算出个体化的SAR发生率，并在贝伐珠单抗治疗前及治疗过程中开展动态风险分层。总体而言，对于总积分较高的患者，临床应警惕其潜在增加的用药风险，及时关注与SAR相关的电解质异常、血液学变化及出血相关事件等信号，并结合患者基础状况(如合并高血压等)合理调整随访及监测频率；对于总积分较低的患者，其短期安全性尚可，但临床仍需按常规流程管理。需要强调的是，列线图仅作为临床风险估计及分层管理的参考工具，并不能替代临床综合判断，其外推应用仍需进一步的外部验证。

### 3.5 本研究局限性

本研究存在如下不足：(1)本研究为单中心研究，样本量相对有限，且ADR的判定依赖人工审核，难以完全避免主观误差。(2)模型变量主要取自常规临床记录，累积剂量、具体联合用药等潜在风险因子因数据库限制而未实现系统化提取；此外，本研究以“抗肿瘤药物品种数”作为替代指标，虽能在一定程度上反映联合用药情况，但无法精确反映剂量-反应关系及药物间的相互作用。(3)本研究贝伐珠单抗致SAR中，重度低钾血症和出血相关事件较为常见，其余类型事件数量有限，故本研究所建预测模型的外推性尚需在多中心、大样本的研究中进一步验证。

## 4 结语

基于GTT理念与德尔菲法筛选的触发器条目，可显著提高贝伐珠单抗相关ADR的识别效率，具有较高的临床应用价值，该方法可作为医疗机构开展药物警戒与安全管理的有效工具，为ADR主动监测提供实践参考。所有候选模型中，Logistic回归模型表现最优，可为临床使用贝伐珠单抗的风险评估和干预策略制定提供可行路径及证据支撑。未来可结合自动化数据采集与算法优化，探索智能化药物安全监测平台的建设。

## 参考文献

- [1] BRAY F, LAVERSANNE M, SUNG H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [J]. *CA Cancer J Clin*, 2024, 74(3):229-263.
- [2] CHEN S M, CAO Z, PRETTNER K, et al. Estimates and projections of the global economic cost of 29 cancers in 204 countries and territories from 2020 to 2050[J]. *JAMA Oncol*, 2023, 9(4):465-472.
- [3] ZHANG M Y, ZHU J, BAO Y, et al. Bevacizumab in ovarian cancer therapy: current advances, clinical challenges, and emerging strategies[J]. *Front Bioeng Biotechnol*, 2025, 13:1589841.
- [4] MAO C L, SEOW K M, CHEN K H. The utilization of bevacizumab in patients with advanced ovarian cancer: a systematic review of the mechanisms and effects[J]. *Int J Mol Sci*, 2022, 23(13):6911.
- [5] MENG X, WU Y Z, LIU Z X, et al. Active monitoring of antifungal adverse events in hospitalized patients based on global trigger tool method[J]. *Front Pharmacol*, 2024, 15:1322587.
- [6] WU S, YIN Q N, WU L Y, et al. Establishing a trigger tool based on global trigger tools to identify adverse drug events in obstetric inpatients in China[J]. *BMC Health Serv Res*, 2024, 24(1):72.
- [7] LIU Y, LIU X J, XIA B B, et al. The application of global trigger tool in monitoring antineoplastic adverse drug events: a retrospective study[J]. *Front Oncol*, 2024, 14:1230514.
- [8] GARCIA J, HURWITZ H I, SANDLER A B, et al. Bevacizumab (avastin®) in cancer treatment: a review of 15 years of clinical experience and future outlook[J]. *Cancer Treat Rev*, 2020, 86:102017.
- [9] PATEL J N, JIANG C, OWZAR K, et al. Pharmacogenetic and clinical risk factors for bevacizumab-related gastrointestinal hemorrhage in prostate cancer patients treated on CALGB 9040: alliance [J]. *Pharmacogenomics J*, 2024, 24(2):6.
- [10] 徐宏珍, 陈嘉, 廖倩倩, 等. 基于CHPS系统主动监测对贝伐珠单抗不良反应发生特点及相关性回顾性分析[J]. *中文科技期刊数据库(引文版)医药卫生*, 2024(5):65-68.
- [11] 丁年羊, 李莉, 方攀奇, 等. 我院贝伐珠单抗生物类似药与原研药相关不良反应回顾性分析[J]. *中国药房*, 2024, 35(4):472-475.
- [12] 杨江, 王明航, 李建生, 等. 基于改良德尔菲法的支气管哮喘急性发作风险预警因素调查表及条目筛选研究[J]. *中国全科医学*, 2022, 25(35):4425-4432.
- [13] SUN J, DENG X Y, CHEN X P, et al. Incidence of adverse drug reactions in COVID-19 patients in China: an active monitoring study by hospital pharmacovigilance system[J]. *Clin Pharmacol Ther*, 2020, 108(4):791-797.
- [14] ZHANG Y F, ZHANG H Y, XU H Q, et al. Efficacy and safety of hepatic arterial infusion chemotherapy combined with lenvatinib and PD-1 inhibitors for advanced hepatocellular carcinoma with macrovascular invasion[J]. *World J Surg Oncol*, 2024, 22(1):122.
- [15] 卫生部. 药品不良反应报告和监测管理办法[EB/OL]. (2011-05-04) [2025-07-31]. [http://www.gov.cn/zhengce/2021-06/29/content\\_5723552.htm](http://www.gov.cn/zhengce/2021-06/29/content_5723552.htm).
- [16] MOONS K G M, ALTMAN D G, REITSMA J B, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration[J]. *Ann Intern Med*, 2015, 162(1):W1-W73.
- [17] ZHANG N, PAN L Y, CHEN W Y, et al. A risk-factor model for antineoplastic drug-induced serious adverse events in cancer inpatients: a retrospective study based on the global trigger tool and machine learning[J]. *Front Pharmacol*, 2022, 13:896104.
- [18] LUNDBERG S M, LEE S I. A unified approach to interpreting model predictions[J]. *NeurIPS*, 2017, 30:4765-4774.
- [19] LIMA S I V C, MARTINS R R, SALDANHA V, et al. Development and validation of a clinical instrument to predict risk of an adverse drug reactions in hospitalized patients[J]. *PLoS One*, 2020, 15(12):e0243714.
- [20] ZHENG Z L, ZHAO Y H, XIE J W, et al. Clinical risk factors of bevacizumab-related hypertension in patients with metastatic colorectal cancer: a retrospective study[J]. *Front Pharmacol*, 2024, 15:1463026.
- [21] ROUSSEAU A, GÉRAUD A, GEISS R, et al. Safety of solid oncology drugs in older patients: a narrative review [J]. *ESMO Open*, 2024, 9(11):103965.
- [22] MOHAMMED T, SINGH M, TIU J G, et al. Etiology and management of hypertension in patients with cancer[J]. *Cardiooncology*, 2021, 7(1):14.
- [23] KIM J M, JUNG H, KWON H E, et al. Predicting prognostic factors in kidney transplantation using a machine learning approach to enhance outcome predictions: a retrospective cohort study[J]. *Int J Surg*, 2024, 110(11):7159-7168.
- [24] GENG Z, YANG C Y, ZHAO Z Y, et al. Development and validation of a machine learning-based predictive model for assessing the 90-day prognostic outcome of patients with spontaneous intracerebral hemorrhage[J]. *J Transl Med*, 2024, 22(1):236.
- [25] LIU X Z, XIE Z L, ZHANG Y, et al. Machine learning for predicting in-hospital mortality in elderly patients with heart failure combined with hypertension: a multicenter retrospective study[J]. *Cardiovasc Diabetol*, 2024, 23(1):407.

(收稿日期:2025-08-05 修回日期:2026-01-05)  
(编辑:张元媛)