

• 论 著 •

基于机器学习建立胆汁淤积性肝病产气荚膜梭菌感染风险预测模型*

余智航¹, 张雨虹², 张文峰³, 王云英^{2△}

重庆医科大学附属第二医院: 1. 数智中心; 2. 检验科; 3. 肝胆外科, 重庆 400010

摘要:目的 以风险预警为目标构建胆汁淤积性肝病产气荚膜梭菌感染风险预测模型, 用于早期识别高风险患者。方法 收集该院 245 例和重庆市另一三甲医院 35 例胆汁淤积性肝病患者的临床资料进行特征工程筛选指标, 通过 XGBoost 算法构建预测模型, 以受试者工作特征(ROC)曲线下面积(AUC)等评估模型效能。结果 特征重要性分析表明, 降钙素原、C 反应蛋白、中性粒细胞计数和年龄与产气荚膜梭菌感染呈正相关, 嗜酸性粒细胞、红细胞计数、血红蛋白、血细胞比容、血小板计数、血清总蛋白和白蛋白与产气荚膜梭菌感染呈负相关。模型在内部验证、院内验证集和外部验证表现较好, 平均 AUC 分别为 0.860、0.855 和 0.818, 平均灵敏度分别为 0.883、0.847 和 0.840, 平均特异度分别为 0.804、0.764 和 0.696。结论 该风险预测模型能够帮助临床医生较准确地筛查出可能发生产气荚膜梭菌感染的高危胆汁淤积性肝病患者的, 可为临床早期决策提供了有力的参考。

关键词:胆汁淤积性肝病; 产气荚膜梭菌; 机器学习模型; 风险预测

DOI:10.3969/j.issn.1673-4130.2026.06.003 **中图法分类号:**R737.14; R730.43

文章编号:1673-4130(2026)06-0655-07 **文献标志码:**A

A machine learning-based model for predicting Clostridium perfringens infection risk in patients with cholestatic liver disease*

YU Zhihang¹, ZHANG Yuhong², ZHANG Wenfeng³, WANG Yunying^{1△}

1. Digital Intelligence Center; 2. Department of Clinical Laboratory; 3. Department of Hepatobiliary Surgery, the Second Affiliated Hospital of Chongqing Medical University, Chongqing 400010, China

Abstract: Objective To construct a predictive model for Clostridium perfringens infection risk in patients with cholestatic liver disease, aiming at risk warning and early identification of high-risk patients. **Methods** The clinical data of 245 patients with cholestatic liver disease from this hospital and 35 patients from another tertiary hospital in Chongqing were collected for feature engineering to select the screening indicators. A prediction model was constructed using the XGBoost algorithm, and the efficacy of the model was evaluated by metrics such as the area under the receiver operating characteristic (ROC) curve (AUC). **Results** Feature importance analysis revealed that procalcitonin, C-reactive protein, neutrophil count, and age were positively correlated with Clostridium perfringens infection, while eosinophil count, red blood cell count, hemoglobin, hematocrit, platelet count, serum total protein and albumin are negatively correlated with infection. The model demonstrated good performance in internal validation, in-hospital validation, and external validation sets, with mean AUCs of 0.860, 0.855 and 0.818, mean sensitivities of 0.883, 0.847 and 0.840, and mean specificities of 0.804, 0.764 and 0.696, respectively. **Conclusion** This risk prediction model can help clinicians accurately identify cholestatic liver disease patients at high risk of Clostridium perfringens infection, providing a strong reference for early clinical decision-making.

Key words: cholestatic liver disease; Clostridium perfringens; machine learning-based model; risk prediction

产气荚膜梭菌是人类及动物肠道微生物群的组成部分^[1], 目前有关该菌感染的报道主要集中于家禽领域; 在人类中, 产气荚膜梭菌感染多表现为皮肤和

软组织气性坏疽, 而涉及内脏器官的感染多为个案报道^[2], SUZAKI 等^[3]指出, 截至 2023 年全球仅有 100 例报道, HOSHI 等^[4]的研究显示, 2000—2023 年从

* 基金项目: 国家自然科学基金项目(81801963)。

作者简介: 余智航, 男, 博士研究生, 主要从事医疗人工智能研究。△ 通信作者, E-mail: 300818@hospital.cqmu.edu.cn。

64 例肝胆疾病患者中发现并分离出产气荚膜梭菌。然而,本院 2015 年 8 月至 2025 年 8 月从 103 例胆汁淤积性肝病患者的血液及体液中分离出了产气荚膜梭菌,如此高的病例数在我国乃至全球范围内均属罕见。患者感染产气荚膜梭菌后普遍出现严重感染症状,病死率极高。因此,本研究通过回顾性研究,拟基于机器学习方法建立胆汁淤积性肝病患者产气荚膜梭菌感染风险预测模型,分析相关危险因素及其影响关系,为临床早期感染筛查、风险预警、及时干预和降低感染相关病死率提供决策支持与科学依据。

1 资料与方法

1.1 一般资料 选取 2015 年 8 月至 2025 年 8 月重庆医科大学附属第二医院收治的胆汁淤积性肝病患者为研究对象。根据《胆汁淤积性肝病诊治专家共识(2013)》^[5] 标准,将体液或血液中检出产气荚膜梭菌的 103 例患者纳入感染组;同期选取同科室 142 例患者纳入对照组,经核查,其中包含无细菌感染患者 66 例,未感染产气荚膜梭菌但感染其他细菌的患者 76 例;共计 245 例。将 2015 年 8 月至 2023 年 1 月患者纳入实验集,其中包含各感染类型样本 83、51、61 例;将 2023 年 1 月至 2025 年 11 月患者划为院内验证集,其中包含各感染类型样本 20、15、15 例。收集重庆市另一三甲医院产气荚膜梭菌感染病例 14 例,无细菌感染 10 例,其他细菌感染 11 例共 35 例用于外部验证。若同一患者多次感染,仅纳入首次感染资料;临床资料不完整者予以排除。纳入标准:(1)临床资料齐全且自愿参与研究;(2)年龄≥18 岁。排除标准:(1)已参与其他临床试验;(2)有严重认知功能障碍或

精神障碍;(3)妊娠期或哺乳期女性。本研究经本院伦理委员会批准,符合伦理学标准。

1.2 数据收集 从电子病历系统中提取上述各患者临床数据,包括性别、年龄、基础疾病、降钙素原(PCT)、C 反应蛋白(CRP)、白细胞计数(WBC#)、中性粒细胞计数(Neut#)、淋巴细胞计数(Lymph#)、嗜酸性粒细胞百分比(Eos%)、嗜碱性粒细胞百分比(Baso%)、红细胞计数(RBC#)、平均红细胞体积(MCV)、平均血红蛋白浓度(MCHC)、血细胞比容(Hct)、血红蛋白(Hb)、血小板计数(PLT#)、葡萄糖(Glu)、丙氨酸氨基转移酶(ALT)、天门冬氨酸氨基转移酶(AST)、γ-谷氨酰转肽酶(gamma-GGT)、碱性磷酸酶(ALP)、总胆红素(TBIL)、直接胆红素(DBIL)、血清总蛋白(TP)、血清白蛋白(ALB)、球蛋白(GLO)、前白蛋白(PA)、凝血酶原国际标准化比值(PT-INR)、凝血酶原时间(PT)、凝血酶时间(TT)、活化部分凝血活酶时间(APTT)、纤维蛋白原(Fbg)、纤维蛋白原降解产物(FDP)总计 33 个缺失率小于 20% 的指标。

1.3 特征工程及患者基本资料比较 为筛选有效的指标,根据适用情况对感染组和对照组采用非配对 *t* 检验、 χ^2 检验、Mann-Whitney *U* 检验和对数 *t* 检验;若 $P \leq 0.05$,则认为该指标差异有统计学意义。考虑到以上检验捕捉指标关系能力有限,进一步采用随机森林模型对实验集进行预训练,并结合 SHAP 值进行指标的重要性分析,最终选取平均 SHAP 绝对值 ≥ 0.05 或 $P \leq 0.05$ 的指标投入模型的训练。依据以上阈值筛选了以下 21 个指标投入模型的训练,见表 1。

表 1 感染组与对照组一般资料比较(*n* 或 $\bar{x} \pm s$)

项目	感染组(<i>n</i> = 83)	对照组(<i>n</i> = 121)	<i>P</i>		
			<i>t</i> / χ^2 检验	Mann-Whitney <i>U</i>	对数 <i>t</i> 检验
性别			0.801	—	—
男	48	67			
女	35	44			
年龄	63.67 ± 12.79	57.94 ± 14.54	0.004	0.009	0.004
RBC#	3.57 ± 0.67	3.92 ± 0.98	0.004	0.002	0.023
MCV	90.37 ± 8.52	92.15 ± 9.22	0.167	0.104	0.188
MCHC	331.46 ± 38.70	328.02 ± 14.98	0.446	0.511	0.539
Neut#	13.06 ± 11.42	7.09 ± 4.98	<0.001	<0.001	<0.001
Hct	32.24 ± 6.25	36.29 ± 11.88	0.002	0.001	0.010
Eos%	0.36 ± 0.77	1.73 ± 2.48	<0.001	<0.001	<0.001
Hb	106.15 ± 20.99	117.35 ± 28.47	0.002	0.001	0.016
PLT#	148.35 ± 81.62	198.35 ± 113.89	<0.001	0.001	<0.001
ALT	131.47 ± 192.52	70.92 ± 81.18	0.010	0.006	0.004
gamma-GGT	214.29 ± 229.47	255.77 ± 236.97	0.289	0.006	0.005
DBIL	43.73 ± 66.78	41.19 ± 79.55	0.819	0.019	0.069
ALB	34.50 ± 6.14	36.95 ± 6.96	0.015	0.041	0.023
GLO	29.91 ± 6.70	31.19 ± 7.90	0.291	0.153	0.482
TP	64.41 ± 9.54	68.14 ± 10.30	0.014	0.004	0.020

续表 1 感染组与对照组一般资料比较 (n 或 $\bar{x} \pm s$)

项目	感染组 (n=83)	对照组 (n=121)	P		
			t/ χ^2 检验	Mann-Whitney U	对数 t 检验
PA	137.12±65.62	148.84±93.23	0.341	0.153	0.649
ALP	234.45±208.31	193.41±199.13	0.188	0.004	0.021
CRP	89.46±63.35	51.09±63.82	<0.001	<0.001	0.017
Fbg	4.50±1.47	3.81±1.46	0.002	<0.001	<0.001
PCT	38.28±68.12	13.91±70.58	0.037	<0.001	<0.001

注:不同菌种的感染过程存在固有生物学异质性,感染相关实验数据呈高离散分布为客观生物学表现,不代表实验操作或数据质量存在异常;—表示无数据。

1.4 建模方法 所有数据操作基于 python 平台, Xgboost 算法来源于 Xgboost(3.0.5)库。模型评估、数据处理主要使用 sklearn(1.7.0)、numpy(1.23.5)、torch(2.7.1)等常见数据库。模型 Shap 值、受试者工作特征(ROC)曲线等可视化来自于 matplotlib(3.7.2)的 pyplot、shap(0.48.0)等包。模型超参数: Xgboost 算法来源于 xgboost 库中 XGBClassifier 模型具体超参数为 n_estimators=200, learning_rate=0.05, max_depth=4, gamma=0.02, reg_alpha=0.1。

1.4.1 模型架构 本研究尝试了采用 XGBoost^[6]、随机森林、LightGBM^[7]、支持向量机(SVM)、Softmax 回归及其混合集成模型构建感染预测模型。本研究并未直接对实验集中的感染组和对照组进行二分类任务训练,而是更精细化建模;模型在训练阶段针对三类标签进行训练:感染组、对照组中的其他细菌感染组及未感染组(分别标记为 2、1、0),而最终实际使用与评估阶段将模型预测为 0 或 1 的样本均归为阴性,预测为 2 的归为阳性。这样保证了模型可以同时获得分辨感染产气荚膜梭菌样本和感染其他细菌样本,以及分辨感染产气荚膜梭菌样本和未感染样本的能力。综合比较各模型性能后,最终选择 XGBoost 作为本研究的主要建模算法。

1.4.2 模型训练与评估 采用曲线下面积(AUC)、准确率、F1 分数、灵敏度、特异度、阴性预测值、阳性预测值和 Brier 分数评估模型性能。为确保模型的泛化性能,内部验证采用重复 K 折交叉验证,其中 K=5,重复 10 次随机划分并求取平均的评估指标值。在每一折训练中采用 Min-Max 归一化,与 KNN 插值法分别对训练集和测试集归一化数据并填补缺失值。

1.4.3 模型最终架构 使用实验集数据进行 5 重重复五折交叉训练,并在每一折训练数据中加入方差 0.01 的高斯噪声;由此共获得 25 个弱分类器,将它们水平聚合为预测模型,作为最终的模型,即预测模型的预测结果是由 25 个弱分类器求平均得到。这样的模型相较于单次训练所有数据所得到的模型,具有更好的稳定性和泛化能力。

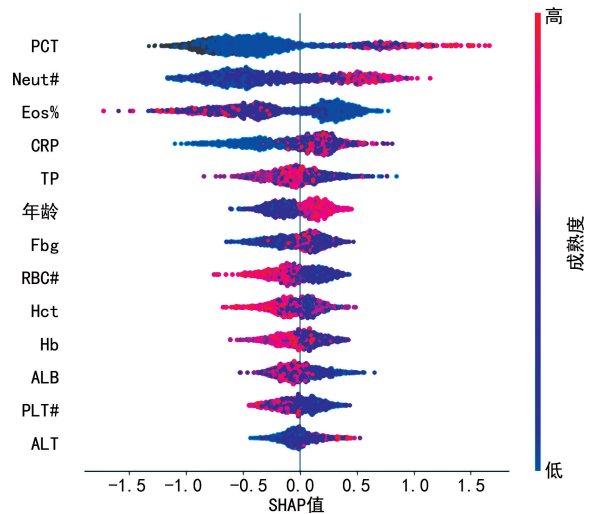
1.4.4 模型验证 通过以上步骤训练得到预测模型,使用院内验证集及外部验证集考察模型的性能,对于这两个集合数据同样采用 Min-Max 归一化,与

KNN 插值法进行预处理。由于两个集合样本数量均较小,得到的评估数值具有偶然性,因此额外采用自助法(Bootstrap)1 000 次验证使得结果稳定性更加可靠。此外按每次随机遮挡 10%,20%验证集特征各重复 100 次模型评估;用于模拟实际场景中数据缺失的情况。

1.4.5 模型对比 为了体现三分类训练模型执行二分类任务这种精细化建模方法的优越性,对比以下 4 种建模的性能:(1)使用本文上述方法建模;(2)使用 Softmax 回归算法进行三分类训练,再用于目标二分类任务;(3)直接使用实验集中的感染组和对照组进行二分类训练 Xgboost 模型;(4)直接使用实验集中的感染组和对照组进行二分类训练逻辑回归模型(注:逻辑回归可视为 Softmax 回归的二分类版本)。

2 结果

2.1 指标分析 由于本研究选取的指标较多,因此将表 1 中经检验 $P \leq 0.05$ 的指标纳入分析。这些指标在模型 50 次训练的 SHAP 点图见图 1,该图可以反映各指标对模型输出的影响程度:PCT、Neut#、CRP、Age、Fbg 和 ALT 与产气荚膜梭菌感染呈正相关,而 Eos%、RBC#、Hct、Hb、PLT#、TP 和 ALB 与产气荚膜梭菌感染呈负相关。



注:SHAP 值表示该特征对预测的影响程度;红色表示该点对应的样本该指标值较高,蓝色表示对应的样本该指标值较低。

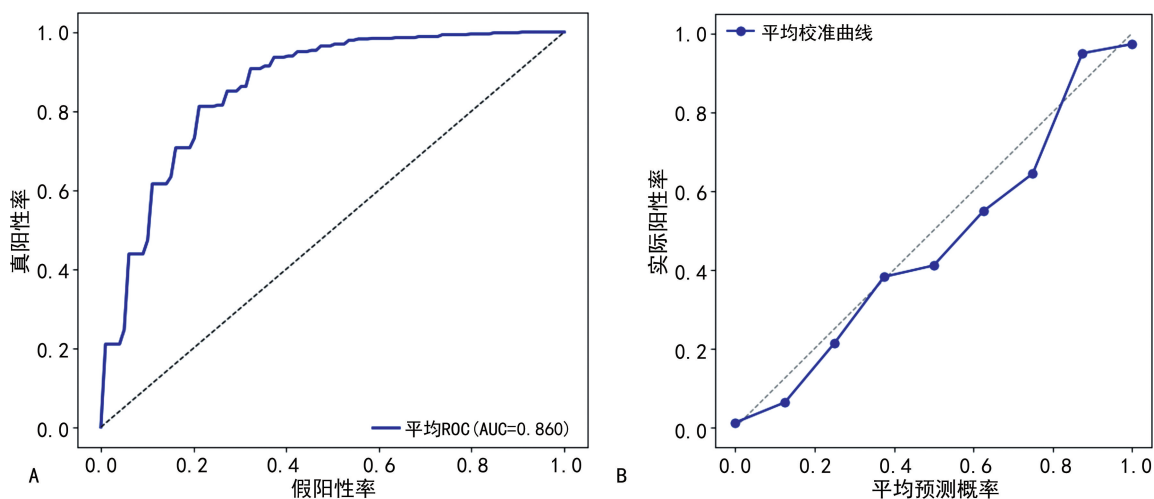
图 1 模型 SHAP 点图

2.2 模型内部评估 模型内部验证的平均 ROC 曲线和校准曲线见图 2,模型内部验证各评估指标对应

值如下: AUC 为 0.860 ± 0.054 , 准确率为 0.836 ± 0.047 , F1 分数为 0.811 ± 0.045 , 灵敏度为 0.883 ± 0.092 , 特异度为 0.804 ± 0.107 , 阴性预测值为 0.921 ± 0.054 , 阳性预测值为 0.765 ± 0.092 , Brier 分数为 0.032 ± 0.028 。从精确度看, 模型拥有良好且较稳定的 AUC、准确率、F1 分数、灵敏度和阴性预测值, 说明模型有较好的整体判别能力及阴性样本的识别能力; 特异度和阳性预测值在可接受且较稳定的范围; 这种差异表明模型倾向于筛查型模型, 满足本研究的建模需求。ROC 曲线整体形状平滑、无剧烈变化, 反映模型较为稳定。从校准性看, 校准曲线接近对角线且 Brier 分数在 95% CI 下小于 0.06, 反映模

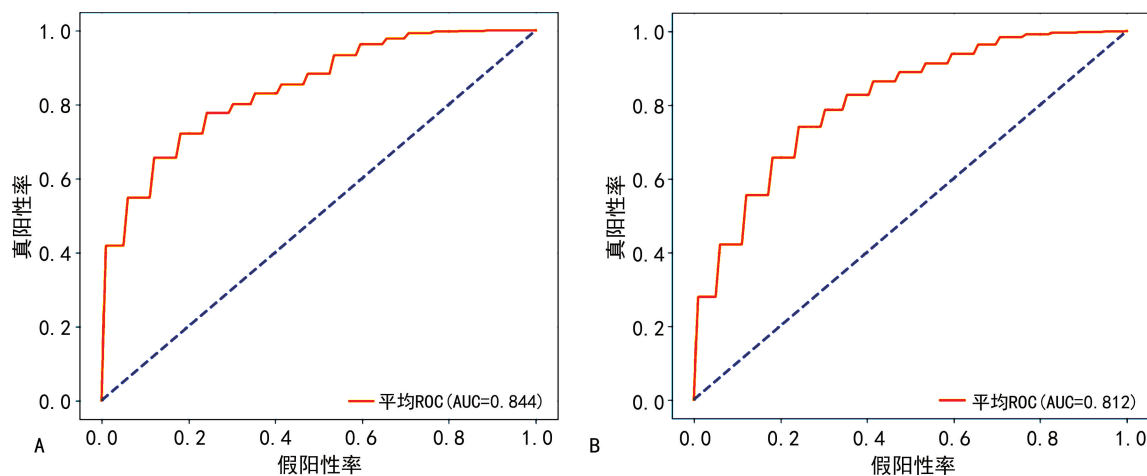
型的预测概率与真实风险匹配良好。总体而言, 该模型在内部验证中表现出良好可靠的判别能力、预警能力与较高的校准度。

2.3 模型在院内验证集的评估 模型随机 10% 和 20% 特征遮挡下验证的平均 ROC 曲线见图 3; 模型在直接验证、自助法、随机特征遮挡 10% 和 20% 情况下各评估指标对应值见表 2。自助法结果与内部验证结果接近, 表明模型没有过拟合。在随机特征遮挡 10% 和 20% 情况下, 模型的评估数值依旧处于可以接受的范围; 这表明模型具有一定的鲁棒性, 在特征少量缺失的情形依旧保持一定性能。



注: A 为 ROC 曲线; B 为校准曲线。

图 2 模型平均 ROC 曲线和校准曲线



注: A 为在 10% 特征遮挡下模型平均 ROC 曲线; B 为在 20% 特征遮挡下模型平均 ROC 曲线。

图 3 院内验证集评估

表 2 模型在院内验证集上性能指标

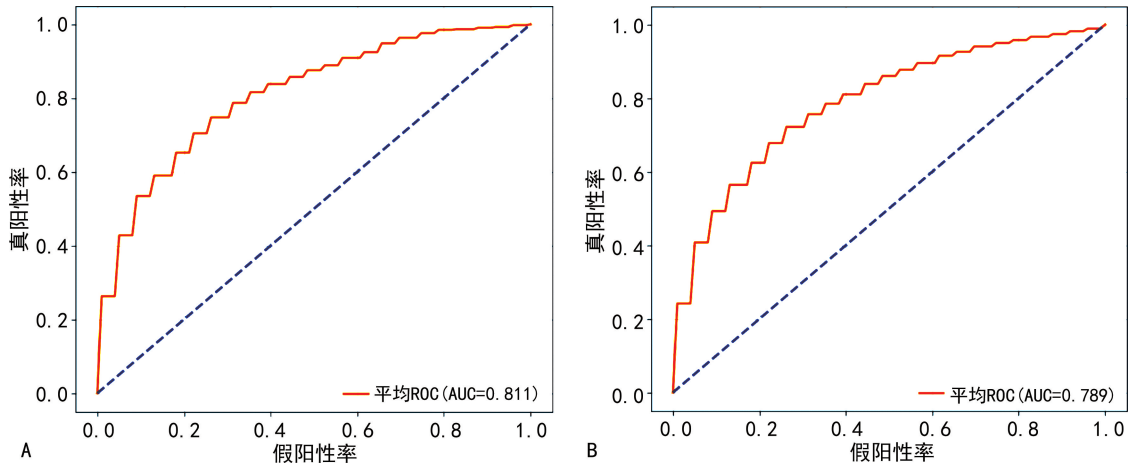
指标	直接验证	自助法 (Bootstrap)	随机特征遮挡 (10%)	随机特征遮挡 (20%)
AUC	0.834	0.855	0.844	0.812
准确率	0.740	0.789	0.770	0.734
F1 分数	0.727	0.778	0.752	0.713

续表 2 模型在院内验证集上性能指标

指标	直接验证	自助法 (Bootstrap)	随机特征遮挡 (10%)	随机特征遮挡 (20%)
灵敏度	0.780	0.847	0.821	0.769
特异度	0.700	0.764	0.744	0.707
阴性预测值	0.800	0.875	0.833	0.798
阳性预测值	0.680	0.740	0.706	0.667

2.4 模型外部验证 模型随机 10% 和 20% 特征遮挡下验证的平均 ROC 曲线见图 4, 模型在直接验证、自助法、随机特征遮挡 10% 和 20% 情况下各评估指标对应值见表 3。模型性能处于可以接受的范围, 与

在院内验证集上测试结果性能差距不大, 表明模型学习到区分产气荚膜梭菌的关键规律, 具备一定的泛化能力。



注: A 在 10% 特征遮挡下模型平均 ROC 曲线 B 在 20% 特征遮挡下模型平均 ROC 曲线。

图 4 外部验证

表 3 模型外部验证性能指标

指标	直接验证	自助法 (Bootstrap)	随机特征遮挡 (10%)	随机特征遮挡 (20%)
AUC	0.816	0.818	0.811	0.789
准确率	0.771	0.755	0.739	0.719
F1 分数	0.793	0.779	0.754	0.733
灵敏度	0.840	0.842	0.816	0.796
特异度	0.696	0.661	0.659	0.637
阴性预测值	0.800	0.801	0.787	0.764
阳性预测值	0.750	0.733	0.723	0.704

续表 4 各模型性能对比

指标	Xgboost (三分类)	Xgboost (二分类)	Softmax (三分类)	逻辑回归
特异度	0.661	0.620	0.562	0.495
阴性预测值	0.801	0.690	0.734	0.630
阳性预测值	0.733	0.680	0.673	0.608

2.5 各建模性能对比 各模型在外部验证集, 使用自助法 1 000 次的平均性能指标见表 4。由表 4 可见, Xgboost 在此任务中性能优于 Softmax 类的线性模型, 这说明产气荚膜梭菌的感染的判别特征难以通过线性结构充分提取, 而树模型能够更有效地捕捉非线性关系。采用三分类训练的模型表现均优于直接进行二分训练的同种模型。这一结果是可以预见的, 因为若将训练集中无感染组和其他细菌感染组混为一类, 模型可能难以学到区分产气荚膜梭菌感染和其他感染之间的关键特征。

表 4 各模型性能对比

指标	Xgboost (三分类)	Xgboost (二分类)	Softmax (三分类)	逻辑回归
AUC	0.818	0.761	0.725	0.617
准确率	0.755	0.705	0.701	0.610
F1 分数	0.779	0.720	0.718	0.620
灵敏度	0.842	0.730	0.773	0.655

3 讨论

胆汁淤积性肝病产气荚膜梭菌引起的感染之前的报道很少, 特别是在胆汁淤积性肝病患者中大量分离更是罕见, 机体肝胆系统感染产气荚膜梭菌后, 会引起坏疽性胆囊胆管炎和肝脓肿, 而且无明显的临床症状特征并迅速发展为休克, 导致临床医生无法及时识别感染^[8]。HOSHI 等^[4]认为治疗胆汁淤积性肝病感染时, 在拿到微生物结果之前需重视产气荚膜梭菌的感染。但在实际临床工作中产气荚膜梭菌感染的诊断却很困难, 主要原因有: (1) 临床医生对此菌的感染危害性认识不足; (2) 实验室检测微生物能力有限; (3) 没有合适的指南或诊断模型参考。因此有必要建立胆汁淤积性肝病患者产气荚膜梭菌感染风险预测模型, 分析出感染危险因素及其相关性, 同时提高实验室的微生物检测能力, 为临床早期感染筛查、风险预警、及时干预和降低感染相关病死率提供科学依据与决策支持。

模型的 SHAP 点图显示: PCT、Neut #、CRP、Age、Fbg 和 ALT 与产气荚膜梭菌感染呈正相关, 而与 Eos%、RBC #、Hct、Hb、PLT #、TP、ALB 与感染呈负相关。模型在内部交叉验证中取得 0.860 ±

0.054 的 AUC, 表明模型在区分是否感染产气荚膜梭菌具有较强能力, 较高的灵敏度(0.883±0.092)和阴性预测值(0.921±0.054), 这提示模型在排除风险方面性能较优。模型的特异度与阳性预测值在可以接受的范围, 分别低于灵敏度和阴性预测值, 这种倾向性说明模型“宁可多报阳性, 也不能少漏”, 适合临床早期感染筛查或风险预警。Brier 分数为 0.033±0.028 说明模型的概率预测接近真实标签, 可靠性较高。各指标的标准差(SD)均在 0.03~0.10, 说明模型表现较为稳定, 没有过度依赖某一折的数据。院内验证集上结果与内部验证接近, 表明模型未出现明显过拟合; 外部验证性能尚好, 反映模型具有良好的泛化能力; 随机遮挡验证结果表明, 模型在数据少量缺失情况下依旧保持一定性能。

模型显示, PCT、Neut#、CRP 与产气荚膜梭菌感染呈正相关, 感染时 PCT、Neut#、CRP 明显升高, 而 Eos% 与产气荚膜梭菌感染概率呈负相关, 其机制可能与急性细菌感染引发的免疫应激反应有关; 当机体遭受感染时, 免疫系统分泌大量内源性糖皮质激素, 抑制骨髓中嗜酸性粒细胞释放至外周血, 同时诱导外周血中已存在嗜酸性粒细胞的程序性死亡, 使其滞留于组织中, 从而导致外周血嗜酸性粒细胞数量显著下降^[9]。

RBC、Hb 和 Hct 及 PLT# 与产气荚膜梭菌感染呈负相关, 产气荚膜梭菌分泌的 α -毒素具有独特的磷脂酶 C 活性, 可水解红细胞膜磷脂, 导致红细胞溶解及数量下降、Hb 和 Hct 降低^[10-12], 如果这些指标持续下降亦提示病情可能进一步加重。此外, PLT# 与产气荚膜梭菌感染呈显著负相关, 其机制可能与感染诱导的全身促凝反应有关, 产气荚膜梭菌感染可引起大量促凝物质释放, 导致全身微血管内广泛形成由血小板和纤维蛋白构成的微血栓, 使血小板被大量消耗。同时, α -毒素还能上调细胞膜上 CD11b 的表达, 促进血小板与白细胞聚集^[13]。上述双重机制共同导致血液中血小板水平急剧下降。

年龄与产气荚膜梭菌感染呈正相关, 这与 DI BELLA 等^[14]报道的结果一致。随着年龄增长, 机体免疫功能逐渐下降, 从而显著增加了病原微生物感染的可能性。TP 和 ALB 与产气荚膜梭菌感染呈负相关, TP 降低可导致肠壁水肿, 使肠道黏膜细胞间连接松弛, 从而增加肠壁通透性, 促进细菌移位。本研究纳入的患者均为胆汁淤积性肝病患者, 胆道梗阻致肠道胆汁缺乏, 导致肝胆汁酸及循环胆汁酸积聚。肠道胆汁的缺乏使其失去抗炎、抑菌、营养上皮细胞及维持上皮紧密连接的生理作用, 进而引发肠道菌群过度生长、黏膜萎缩、通透性增加及肠道运动功能障碍^[15]。此外, 胆汁淤积还可损害肝胰壶腹括约肌功能, 使小肠中定植的产气荚膜梭菌经胆肠反流或通过淋巴及

血行途径进入肝胆系统, 引发感染^[16-17]。因此, 肠道胆汁酸缺乏与 TP 降低的协同作用可能共同增加产气荚膜梭菌感染风险。

综上所述, 胆汁淤积性肝病患者如为高龄人群, TP 较低, 在此基础上发生产气荚膜梭菌感染的概率大大增加。当患者疑似感染时, 首先关注感染相关生物标志物 PCT、Neut# 和 CRP 的升高情况和 Eos% 是否减少。产气荚膜梭菌的毒素对红细胞独特的溶解作用, 造成 RBC 减少、Hb 和 Hct 降低, 所以要特别关注与红细胞相关指标的动态变化, 这些指标持续降低亦提示病情可能进一步加重。产气荚膜梭菌毒素作用使 PLT 减少, 纤维蛋白升高。同时, 应加强实验室检测能力, 对患者分泌物特别是体液样本进行需氧与厌氧双重培养, 并结合新型检测技术, 以提高产气荚膜梭菌的检出率。值得注意的是, 产气荚膜梭菌对头孢类抗菌药物具有天然耐药性, 若能在早期准确识别感染并及时采用针对性抗菌药物及必要的外科干预, 可显著降低其引起菌血症及脓毒症的风险, 减少病死率并提高治愈率。

本研究存在一定局限性, 首先, 因产气荚膜梭菌感染较为罕见, 调查的样本量有限, 可能导致模型训练还不够充分, 效力有限; 其次, 外部验证样本较少, 即使在多次自助法验证下, 验证效力依然有限。因此, 希望有更多的其他医院参与对预测模型有效性的评估。此外, 本研究团队未来将尝试基于多模态数据构建预测模型, 以进一步提高模型的准确性和临床适用性。

参考文献

- [1] KIU R, HALL L J. An update on the human and animal enteric pathogen *Clostridium perfringens*[J]. *Emerg Microbes Infect*, 2018, 7(1):141.
- [2] OSÓRIO C, SILVA D, TELES L, et al. An unusual fatal outcome of laparoscopic cholecystectomy: a case report [J]. *Cureus*, 2023, 15(1):e34365.
- [3] SUZAKI A, HAYAKAWA S. Clinical and microbiological features of fulminant haemolysis caused by *Clostridium perfringens* bacteraemia; unknown pathogenesis[J]. *Microorganisms*, 2023, 11(4):824.
- [4] HOSHI Y, TAKESHIMA K, MATSUOKA S, et al. Survival in a case of emphysematous cholecystitis with sepsis caused by *Clostridium perfringens*[J]. *Cureus*, 2023, 15(11):e49705.
- [5] 胆汁淤积性肝病诊断治疗专家委员会. 胆汁淤积性肝病诊断治疗专家共识 2013[J]. *中国肝脏病杂志*, 2013, 1(5):53-63.
- [6] CHEN T, GUESTRIN C. XGBoost: a scalable tree boosting system [C]//KRISHNAPURAM B, SHAH M, SMOLA A J, et al. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco, California, USA: ACM, 2016: 785-794.

• 论 著 •

前哨淋巴结阴性乳腺癌患者血清 STC-1、MMPs、VEGFs 水平与术后复发风险的关系*

谌彪¹, 马静², 王伟¹, 康鹏^{1△}

1. 乌鲁木齐市第一人民医院普外科, 新疆乌鲁木齐 830011; 2. 新疆医科大学附属肿瘤医院乳甲外科, 新疆乌鲁木齐 830011

摘要:目的 探讨血清斯钙素 1(STC-1)、基质金属蛋白酶(MMPs)及血管内皮生长因子(VEGFs)对前哨淋巴结阴性乳腺癌患者术后复发风险的预测价值。方法 选取 2021 年 4 月至 2023 年 5 月在乌鲁木齐市第一人民医院接受手术治疗的 275 例早期乳腺癌患者为研究对象。随访 2 年,根据术后是否复发分为复发组与未复发组。对比两组患者临床特征、STC-1、MMPs 及 VEGFs 水平的差异。通过多因素 Logistic 回归分析筛选与前哨淋巴结阴性乳腺癌患者术后复发相关的影响因素,受试者工作特征(ROC)曲线评价 STC-1、MMPs、VEGFs 水平预测患者术后复发风险的效能。结果 入选的前哨淋巴结阴性乳腺癌患者均完成随访,其中术后局部复发 13 例,区域复发 5 例,孤立性复发 10 例,总复发率为 10.1%(28/275)。复发组和未复发组患者肿瘤最大径、组织学分级、手术方式比较,差异均有统计学意义($P < 0.05$),复发组患者血清 STC-1、MMP-2、MMP-9、VEGF-A、VEGF-C 水平均高于未复发组($P < 0.05$)。调整协变量后,多因素 Logistic 回归分析结果显示,血清 STC-1($OR = 1.044, 95\% CI: 1.010 \sim 1.089$)、MMP-9($OR = 1.019, 95\% CI: 1.003 \sim 1.035$)、VEGF-A($OR = 1.016, 95\% CI: 1.007 \sim 1.025$)均是前哨淋巴结阴性乳腺癌患者术后复发风险的独立危险因素($P < 0.05$)。ROC 曲线分析结果显示,STC-1、MMP-9、VEGF-A 联合预测前哨淋巴结阴性乳腺癌患者术后复发风险的曲线下面积(AUC)及 95%CI 为 0.924(0.875~0.972),其预测效能高于各因素独立预测($Z = 4.341, 3.262, 3.081$,均 $P < 0.05$)。结论 术后复发的前哨淋巴结阴性乳腺癌患者血清 STC-1、MMP-2、MMP-9、VEGF-A、VEGF-C 水平显著高于未复发患者,血清 STC-1、MMP-9、VEGF-A 水平较高均是患者术后复发的独立危险因素,对术后复发风险具有较高预测效能。

关键词:前哨淋巴结阴性乳腺癌; 复发; 斯钙素 1; 基质金属蛋白酶; 血管内皮生长因子

DOI:10.3969/j.issn.1673-4130.2026.06.004

中图法分类号:R446.1;R737.9

文章编号:1673-4130(2026)06-0661-07

文献标志码:A

Relationship between the serum levels of STC-1, MMPs and VEGFs and the risk of postoperative recurrence in patients with sentinel lymph node-negative breast cancer*

CHEN Biao¹, MA Jing², WANG Wei¹, KANG Peng^{1△}

1. Department of General Surgery, Urumqi First People's Hospital, Urumqi, Xinjiang 830011, China; 2. Department of Breast and Thyroid Surgery, Affiliated Tumor Hospital of Xinjiang Medical University, Urumqi, Xinjiang 830011, China

Abstract: Objective To explore the predictive value of serum stanniocalcin-1 (STC-1), matrix metalloproteinases (MMPs), and vascular endothelial growth factors (VEGFs) for the risk of postoperative recurrence in patients with sentinel lymph node-negative breast cancer. **Methods** A total of 275 patients with early-stage breast cancer who underwent surgical treatment at Urumqi First People's Hospital from April 2021 to May 2023 were selected as the research subjects. The patients were followed up for 2 years and divided into the recurrence group and the non-recurrence group based on whether they had recurrence after surgery. The differences in clinical characteristics, STC-1, MMPs, and VEGFs levels between the two groups were compared. Through multivariate Logistic regression analysis, the influencing factors related to postoperative recurrence in patients with sentinel lymph node-negative breast cancer were screened. The receiver operating characteristic (ROC) curve was used to evaluate the efficacy of STC-1, MMPs, and VEGFs levels in predicting the risk of postoperative recurrence in patients. **Results** All the selected patients with negative sentinel lymph

* 基金项目:新疆维吾尔自治区自然科学基金项目(2021D01C246)。

作者简介:谌彪,男,副主任医师,主要从事乳腺、甲状腺疾病研究。△ 通信作者,E-mail:kangpeng7308@163.com。