基于3种机器学习法分析构建急性上消化道出血患者不良临床结局事件预测模型

南阳市中心医院急诊内科(473000) 郭 旭 田 丽 许勤华

【提 要】 目的 基于机器学习探讨血清乳酸/白蛋白比值(lactate/albumin ratio, LAR)、单核细胞/高密度脂蛋白胆固醇比值(monocyte/high-density lipoprotein cholesterol ratio, MHR)及血小板-白蛋白-胆红素指数(plateletary-albumin-bilirubin index, PALBI)与急性上消化道出血患者发生不良临床结局事件的相关性,以期为临床应用提供参考。**方法** 回顾性选取2022年6月至2024年12月南阳市中心医院收治的急性上消化道出血患者205例作为训练集,根据是否发生消化道再次出血的不良临床结局事件分为发生组(*n*=43)和未发生组(*n*=162)。另选我院同期收治的急性上消化道出血患者88例作为验证集,收集患者临床资料并计算LAR、MHR、PALBI,通过单因素分析患者发生不良临床结局事件的影响因素,使用logistic回归、决策分类回归树(decision classification regression tree, CRT)、反向传播神经网络(back-propagation neural network, BPNN)机器学习算法构建急性上消化道出血患者发生不良临床结局事件的预测模型,并采用ROC曲线比较3种方法构建的模型对发生不良临床结局事件的预测价值。**结果** 单因素和多因素logistic回归分析最终结果显示,年龄、肝硬化、LAR、MHR、PALBI均是发生不良临床结局事件的独立影响因素(*P*<0.05)。采用CRT法构建的预测模型显示,LAR、MHR、PALBI、肝硬化均是发生不良临床结局事件的分类因素。BPNN模型结果显示,影响因素重要性排序为PALBI、年龄、MHR、LAR、肝硬化、幽门螺杆菌感染。logistic回归模型、CRT模型、BPNN模型的AUC分别是0.911(95%*CI*为0.860~0.962)、0.854(95%*CI*为0.779~0.930)、0.927(95%*CI*为0.886~0.969),均>0.800,其中logistic回归模型与BPNN模型诊断效能较佳。验证集结果显示logistic回归模型与BPNN模型的诊断效能较决策树模型好。**结论** LAR、MHR及PALBI与急性上消化道出血患者发生不良临床结局事件显著相关,机器学习法构建的模型均具有较好的预测效能,其中logistic回归模型与BPNN模型诊断效能较佳。

【关键词】 急性上消化道出血 不良临床结局事件 机器学习法

【中图分类号】 TP181;R573.2 **【文献标志码】** A **DOI** 10.11783/j.issn.1002-3674.2026.03.026

急性上消化道出血是最为严重的消化道内科疾病,具有较高的致死风险,一旦出血量超过1000毫升,患者可能会迅速出现急性循环衰竭,这将严重威胁到患者的生命安全[1]。在临床上由于每位患者的具体情况不同,即使经过治疗,部分患者仍有可能面临较高的再次出血风险[2]。因此,迅速评估新入院患者的病情,识别那些有较高风险发生不良临床结果的患者,并对他们进行及时干预,是至关重要的。乳酸/白蛋白比值(lactate/albumin ratio, LAR)和单核细胞/高密度脂蛋白胆固醇比值(monocyte/high-density lipoprotein cholesterol ratio, MHR)作为生物标志物,能够预测急诊危重病患者的预后,游平飞等[3]的研究表明,LAR和MHR对急性上消化道出血患者发生不良临床事件具有一定预测作用。同时,另有研究[4]发现血小板-白蛋白-胆红素指数(plateletary-albumin-bilirubin index, PALBI)也能作为预测急性上消化道出血患者预后的一个指标。目前机器学习算法已广泛应用于数据分析和因素探索等领域,其中,一些比较常用的算法包括logistic回归模型、决策分类回归树(decision classification regression tree, CRT)、和反向传播神经网络(back-propagation neural network, BPNN)算法,这些算法不仅操作简单,而且有助于进行更深入的分析研究[5-7]。但在之前的研究中,较少有利用机器学习算法来建立预测急性上消化道出血患者不良临床结局的模型,同时,不同算法在预测效果上的差异也尚未得到清晰的认识。基于此,本研究通过分析LAR、MHR、PALBI与急性上消化道出血患者消化道再次出血不良临床结局事件发生的关联性,并基于logistic回归、CRT、BPNN算法构建急性上消化道出血患者消化道再次出血的不良临床结局事件发生的预测模型,并比较三种模型的预测效能,以期为临床应用提供参考。

资料与方法

1. 样本量计算

参照王家良[8]的样本量计算方法,样本应为因素的10~20倍,本研究共纳入17个因素,并考虑10%样本流失率,因此样本范围为187~374例,结合我院实际患者数量,最终纳入205例患者。

2. 一般资料

选取2022年6月至2024年12月南阳市中心医院收治的急性上消化道出血患者205例,作为训练集,进行回顾性队列研究,根据是否发生消化道再次出血的不良临床结局事件分为发生组(*n*=43)和未发生组(*n*=162),另选我院同时期收治的急性上消化道出血患者88例作为验证集,纳入标准:①符合《急性上消化道出血急诊诊治流程专家共识》[9]中的诊断标准;②年满18周岁;③拥有完整的临床资料记录。排除标准:①因

食物或药物等引起的呕血和黑便，经内镜检查确认非急性上消化道出血患者；②严重的心、肾功能障碍；③患有血液系统疾病的患者；④有恶性肿瘤并发的患者。本研究已通过医院伦理委员会审批。

3. 资料收集

根据医院电子病历系统统计患者的各项临床资料，包括一般资料、合并症、出血次数、症状（呕血、晕厥、黑便、心力衰竭）、凝血功能，并根据实验室指标计算LAR、MHR和PALBI。其中LAR=血清乳酸（mmol/L）/白蛋白（g/dL），MHR=单核细胞计数绝对值（10^9/L）/高密度脂蛋白胆固醇（mmol/L），PALBI=2.02×$\log_{10}$[总胆红素（μmoL/L）]−0.37×$\log_{10}$[总胆红素（μmoL/L）]^2−0.04×白蛋白（g/L）−3.48×$\log_{10}$[血小板（1000/μL）]+1.01×$\log_{10}$[血小板（1000/μL）]^2。

4. 统计分析方法

采用SPSS 26.0软件进行统计学分析及模型构建，计量资料采用（$\bar{x}\pm s$）描述，组间比较采用两样本均数t检验；计数资料采用[例（%）]表示，组间比较采用卡方检验。采用单因素分析急性上消化道出血患者发生不良临床结局事件的影响因素，以P<0.05为差异有统计学意义。使用训练集构建logistic回归、CRT模型、BPNN模型，并采用ROC曲线评估模型的区分度和预测效能，比较三种模型AUC、灵敏度和特异度。使用验证集进行模型验证。

结　果

1. 急性上消化道出血患者发生不良临床结局事件的单因素分析

单因素分析结果显示，年龄、肝硬化、幽门螺杆菌感染、LAR、MHR、PALBI在两组间的差异均有统计学意义（P<0.05），见表1。

表1　急性上消化道出血患者发生不良临床结局事件的单因素分析

项目	发生组（n=43）	未发生组（n=162）	t/χ^2值	P值
年龄（岁，$\bar{x}\pm s$）	66.28±7.26	62.81±6.51	3.032	0.003
肝硬化[n（%）]			4.867	0.027
有	14（32.56）	28（17.28）		
无	29（67.44）	134（82.72）		
幽门螺杆菌感染[n（%）]			5.014	0.025
有	19（44.19）	43（26.54）		
无	24（55.81）	119（73.46）		
LAR（$\bar{x}\pm s$）	0.69±0.14	0.58±0.11	5.488	<0.001
MHR（$\bar{x}\pm s$）	0.61±0.17	0.46±0.13	6.281	<0.001
PALBI（$\bar{x}\pm s$）	−1.52±0.46	−2.21±0.54	7.669	<0.001

2. 模型输入变量的选择

以是否发生不良临床结局事件为因变量，以单因素分析有差异的因素为自变量，将6个变量分别纳入3种模型作为自变量。

3. 多因素logistic回归分析

多因素logistic回归分析变量不良临床结局事件（发生=1；未发生=0），根据王家良[8]书中可知，自变量形式可为连续变量或分类变量，因此年龄、LAR、MHR、PALBI（原值输入），肝硬化、幽门螺杆菌感染（有=1；无=0），赋值后，结果显示，年龄、肝硬化、LAR、MHR、PALBI是急性上消化道出血患者发生不良临床结局事件的独立因素（P<0.05），见表2。

表2　急性上消化道出血患者发生不良临床结局事件的多因素logistic回归分析

相关危险因素	回归系数（b）	标准误	Wald χ^2值	P值	OR	95%CI	
						下限	上限
年龄	0.095	0.035	7.228	0.007	1.099	1.026	1.178
肝硬化	1.225	0.573	4.574	0.032	3.405	1.108	10.468
LAR	4.010	2.048	8.049	0.005	33.654	6.027	104.721
MHR	3.324	1.617	10.844	0.001	20.148	8.629	58.991
PALBI	2.346	0.501	21.928	<0.001	10.439	3.911	27.863
常量	−10.165	3.007	11.424	0.001			

4. CRT模型构建

将单因素分析中有统计学意义的6个变量纳入决策树模型，父节点和子节点中最小个案数分别为10和5，检验水准为0.05，基于CRT法构建的分类树模型共4层，11个结点，6个终结点。筛选出的因素包括PALBI、MHR、LAR、幽门螺杆菌感染。各相关危险因素发生不良临床事件占比见图1。

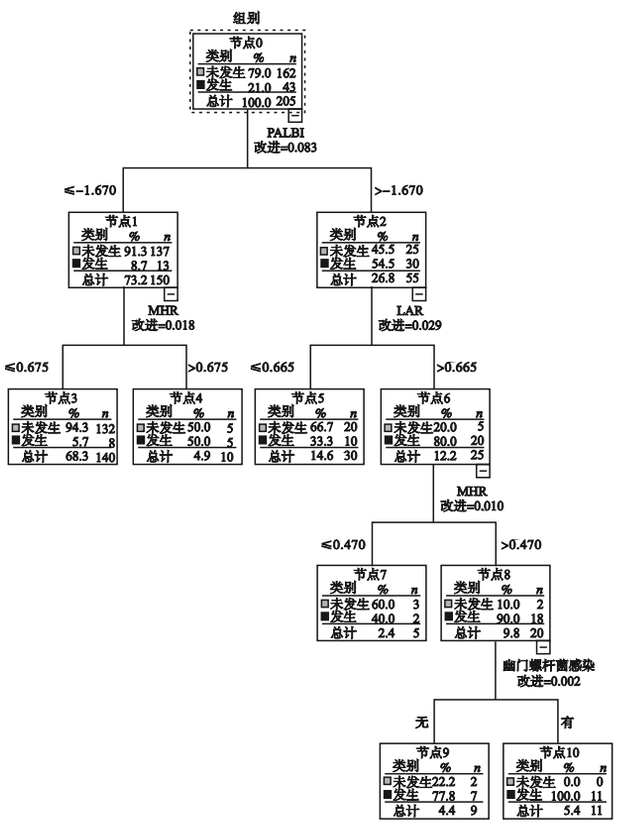


图1　CRT模型

5. BPNN模型构建

采用BPNN构建急性上消化道出血患者发生不良临床结局事件的预测模型，以205例患者中152例为建模集，53例为验证集做神经网络模型拟合，设置隐藏层数最小1，最大50，其中包含H(1∶1)、H(1∶2)、H(1∶3)、H(1∶4)4个节点，隐藏层激活函数为双曲正切，输出层激活函数为Softmax。各变量对模型的重要性正态化后，依次为PALBI、年龄、MHR、LAR、肝硬化、幽门螺杆菌感染。见图2、表3。

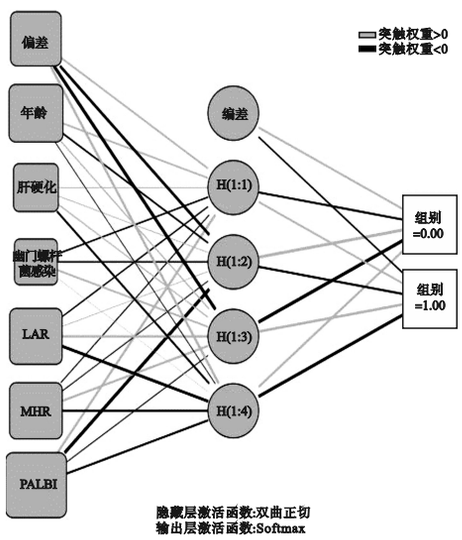


图2 BPNN模型

表3 影响因素重要性排序

项目	重要性	正态化自变量重要性(%)
PALBI	0.317	100.00
年龄	0.207	65.30
MHR	0.188	59.20
LAR	0.185	58.30
肝硬化	0.066	20.90
幽门螺杆菌感染	0.037	11.60

6. 三种模型预测性能比较

以三种预测模型计算得到预测变量为测试变量，急性上消化道出血患者不良临床结局事件发生情况为状态变量绘制ROC曲线，见图3。logistic回归模型AUC值为0.911(95%*CI*为0.860~0.962)，灵敏度为0.744，特异度为0.957；CRT模型AUC值为0.854(95%*CI*为0.779~0.930)，灵敏度为0.814，特异度为0.815；BPNN模型AUC值为0.927(95%*CI*为0.886~0.969)，灵敏度为0.884，特异度为0.827。非参数DeLong检验结果显示，logistic回归模型和BPNN模型的AUC均大于CRT模型，logistic回归模型与BPNN模型的AUC比较差异无统计学意义(*P*>0.05)，见表4。另选88例患者作为验证集，logistic回归模型AUC值为0.909(95%*CI*为0.860~0.958)，灵敏度为0.889，特异度为0.929；CRT模型AUC值为0.796(95%*CI*为0.731~0.861)，灵敏度为0.828，特异度为0.815；BPNN模型AUC值为0.939(95%*CI*为0.900~0.978)，灵敏度为0.814，特异度为0.932，见图4。

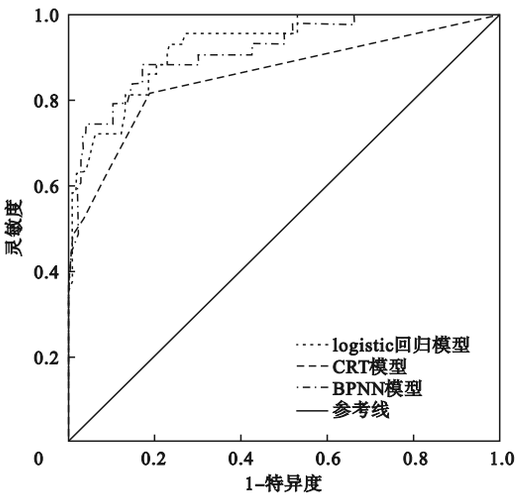


图3 训练集logistic回归模型、决策树CRT模型、BPNN模型的ROC曲线

表4 3种模型ROC曲线下面积的成对比较

检验结果对	*Z*值	*P*值	AUC差异	95%*CI*	
				下限	下限
A−B	1.979	0.048	0.057	0.245	0.113
A−C	−1.449	0.147	−0.016	0.216	0.006
B−C	−2.586	0.010	−0.073	0.236	−0.018

注：A表示logistic回归模型；B表示CRT模型；C表示BPNN模型。

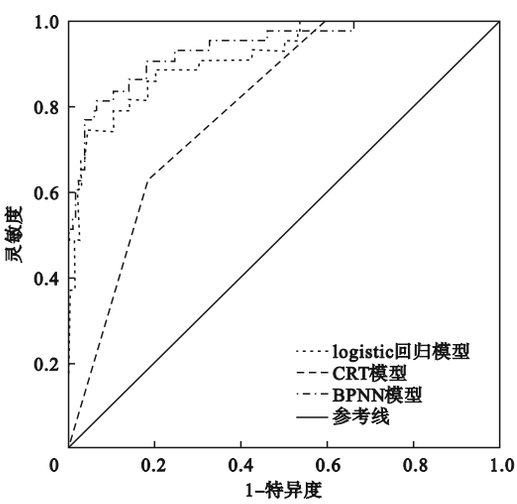


图4 验证集logistic回归模型、决策树CRT模型、BPNN模型的ROC曲线

讨 论

本研究采用了3种机器学习算法来构建预测模型。其中，logistic回归结果显示，年龄、肝硬化、LAR、MHR、PALBI是急性上消化道出血患者发生不良临床结局事件的独立因素。高龄患者机体功能显著减退，其组织自我修复的能力较弱，使得这些患者再次出血的风险相对较高[10]。肝硬化患者由于肝功能减退，可能会引起血管内压力升高和血浆蛋白水平降低，从而

增加了患者再次出血的风险[11]。血清乳酸的正常范围约在0.5~2.2mmol/L，白蛋白的正常范围约是3.5~5.0g/dL，本研究中两组LAR值的差异较小，但差异具有统计学意义，提示LAR可能是一个有用的预测因子。已有研究[12]指出，LAR是预测重症患者病死率的一个独立危险因素，其升高与器官衰竭密切相关，本研究结果与文献结果相符。再出血的患者可能会出现组织缺氧和器官功能障碍等严重情况，这些情况会导致体内乳酸水平的升高，同时在机体严重消耗状态下，白蛋白的分解也会增加，进而导致白蛋白水平的下降[13]，因此发生不良临床结局事件患者的LAR升高。MHR是一种能够反映体内炎症反应严重程度的指标[14]。当MHR水平升高时，这表明患者体内的炎症反应更为活跃，且病情可能进展得更快[15]，因此，MHR水平的升高也可能预示着患者发生不良临床结局事件的风险更大。PALBI指数能反映门静脉高压的严重程度，门静脉高压是导致食管和胃底静脉曲张破裂出血的一个重要危险因素，同时也是引起消化道溃疡的一个原因，因此，PALBI指数能够客观地反映急性上消化道出血患者病情程度[16]。

本研究构建的CRT模型共4层，11个结点，6个终结点。分类所用因素包括PALBI、MHR、LAR、幽门螺杆菌感染，与logistic回归分析相比，增加了幽门螺杆菌感染，表明合并幽门螺杆菌感染也增加不良临床结局事件发生风险。其可能机制在于幽门螺杆菌通过鞭毛穿透胃黏膜表面的黏液层，并利用尿素酶的作用将周围环境变为碱性，从而损伤胃黏膜，同时其还能引发细胞释放多种炎症因子，激发炎症反应，进一步加剧胃黏膜的损害，从而提高上消化道出血的风险[17]。因此，对于急性上消化道出血的患者，如果合并幽门螺杆菌感染，他们的再出血风险较高，预后也较差，这与高伟华等[18]的研究结果一致。

本研究BPNN模型结果显示，影响因素重要性排序为PALBI、年龄、MHR、LAR、肝硬化、幽门螺杆菌感染。与多因素logistic回归分析和SVM模型相比，PALBI、MHR、LAR同样是比较重要的影响因素，提示临床护理人员应密切关注。

3种机器学习算法构建的模型的AUC均>0.800，预测准确性良好，非参数DeLong检验结果显示logistics回归模型和BPNN模型的AUC均大于CRT模型，logistic回归模型与BPNN模型的AUC比较差异无统计学意义，表明logistic回归模型与BPNN模型对急性上消化道出血患者发生不良临床结局事件的预测效能好，能识别高风险患者。

综上所述，LAR、MHR及PALBI与急性上消化道出血患者发生不良临床结局事件显著相关，机器学习法构建的模型均具有较好的预测效能，其中以logistic回归模型与BPNN模型诊断效能较佳，可进一步推广应用以验证预测模型的效能。但本研究亦存在一定缺陷，研究患者均来自本院，无法代表所有的急性上消化道出血患者，后续可联合多地区医院对急性上消化道出血患者发生不良临床结局事件的预测进行细化研究。

参 考 文 献

[1] 许清华,孙鹏飞. 急性上消化道出血患者应用泮托拉唑和生长抑素联合急诊治疗的效果与结局观察[J]. 中华养生保健,2024,42(13):190-193.
[2] 胡萍,周敏,熊小芳,等. 难治性急性非静脉曲张性上消化道出血消化内镜止血后再出血状况及集束化护理预防研究[J]. 基层医学论坛,2024,28(22):79-81.
[3] 游平飞,金红旭. LAR联合MHR对急性上消化道出血患者预后的预测价值[J]. 中国医科大学学报,2024,53(7):640-645.
[4] 吴永美,卞熠,俞远怀,等. 年龄休克指数、PALBI评分及BUN/Cr比值对急性消化道出血输血治疗患者预后的预测价值[J]. 分子诊断与治疗杂志,2023,15(12):2079-2082.
[5] 张宁,陈雁,张红梅,等. 2型糖尿病患者轻度认知功能障碍发生风险的logistic回归和决策树预测模型的构建[J]. 中国实用护理杂志,2023,39(32):2514-2520.
[6] 崔海艳,景涛,邹新亮,等. 急性心肌梗死患者PCI后服药依从性较差的主要影响因素及其BP神经网络模型构建[J]. 实用心脑肺血管病杂志,2024,32(6):45-50.
[7] 方积乾. 医学统计手册[M]. 北京:中国统计出版社,2018.
[8] 王家良. 临床流行病学:临床科研设计,测量与评价[M]. 上海:上海科学技术出版社,2014.
[9] 中国医师协会急诊医师分会,中华医学会急诊医学分会,全军急救医学专业委员会,等. 急性上消化道出血急诊诊治流程专家共识[J]. 中国急救医学,2021,41(1):1-10.
[10] 马阳阳,郭凯华. 上消化道出血患者术后早期再出血风险预测模型的构建及效果验证[J]. 护理实践与研究,2023,20(21):3271-3276.
[11] 柴振伟. 急性上消化道出血患者的临床特征及不良预后的影响因素[J]. 河南医学研究,2022,31(19):3560-3563.
[12] 周银瑞,年士艳,冯磊,等. 乳酸、白蛋白及其比值在ICU高致死率疾病中的应用进展[J]. 分子诊断与治疗杂志,2024,16(4):783-786.
[13] 王肖,李玲玲. PGE2、IL-6及乳酸水平对肝硬化合并消化道出血的诊断及预后预测价值[J]. 医学理论与实践,2024,37(5):836-839.
[14] 臧传丽,朴玉华,朴红梅. 单核细胞计数与高密度脂蛋白比值及全身免疫炎症指数预测急性肺栓塞病情严重程度及预后的临床意义[J]. 延边大学医学学报,2023,46(4):276-281.
[15] 刘伟,庄雷,李妙男. MHR、NLR与急性缺血性脑卒中静脉溶栓患者预后的关系[J]. 中国现代医学杂志,2024,34(3):20-25.
[16] 常琳琳,代子艳,周亚柏,等. OASIS评分联合血小板-白蛋白-胆红素指数与急性上消化道出血病情的相关性及对预后评估价值[J]. 河北医学,2024,30(3):441-448.
[17] 盛锦义,姜景平,汪福昌. 老年消化道出血临床特征及与HP感染、非甾体抗炎药物使用的关系[J]. 中国老年学杂志,2022,42(24):5978-5980.
[18] 高伟华,赵杰,李晓敏. 急性上消化道出血预后相关危险因素分析[J]. 中外医疗,2024,43(17):58-60.

（责任编辑：林美艺）