临床研究 DOI：10.13406/j.cnki.cyxb.003631

基于可解释机器学习的肝硬化患者早期并发食管静脉曲张风险预测研究

尹钰恒[1]，王雨雯[1]，范　捷[2]，杨　春[3]，王　伟[1]
（1.重庆医科大学公共卫生学院，重庆　400016；2.重庆医药高等专科学校公共卫生与应急管理学院，重庆　401331；
3.重庆医科大学附属第一医院感染科，重庆　400016）

【摘　要】**目的：**探讨肝硬化患者并发食管静脉曲张的风险因素，并建立预测模型，以便为预防肝硬化患者早期并发食管静脉曲张提供合理的指导。**方法：**回顾性纳入2006年12月至2021年5月在重庆医渡云大数据平台有电子记录的1 113例肝硬化患者。通过结合递归特征消除算法（recursive feature elimination，RFE）和4种机器学习方法筛选特征，建立5种机器学习预测模型，包括逻辑回归、随机森林、支持向量机（support vector machine，SVM）、决策树和极端梯度提升（eXtreme Gradient Boosting，XGBoost）模型。通过绘制受试者工作特征（receiver operating characteristic，ROC）曲线评估不同模型的性能，并根据表现最优的模型分析肝硬化患者并发食管静脉曲张的风险因素。SHAP图用于解释各风险因素对患者的影响。**结果：**XGBoost模型在预测肝硬化患者并发食管静脉曲张的风险方面表现最佳，ROC曲线下面积为0.872（95%CI=0.813~0.918）。SHAP图表明，影响肝硬化患者并发食管静脉曲张的风险因素包括血小板计数、门静脉内径、胆碱酯酶、白蛋白、谷丙转氨酶、血红蛋白、凝血酶原比值、凝血酶原时间和血清总蛋白等9项临床指标。**结论：**本研究构建的XGBoost预测模型具有较高的预测价值，其筛选出的风险因素对临床防治肝硬化患者早期并发食管静脉曲张具有一定的指导意义。

【关键词】机器学习；肝硬化；食管静脉曲张；风险因素
【中图分类号】R543.6　　【文献标志码】A　　【收稿日期】2024-10-03

Risk prediction of early esophageal varices in patients with liver cirrhosis based on interpretable machine learning

Yin Yuheng[1]，*Wang Yuwen*[1]，*Fan Jie*[2]，*Yang Chun*[3]，*Wang Wei*[1]
（*1. College of Public Health，Chongqing Medical University；2. School of Public Health Emergency Management，Chongqing Medical and Pharmaceutical College；3. Department of Infection Diseases，The First Affiliated Hospital of Chongqing Medical University*）

【**Abstract**】***Objective***：To investigate the risk factors for esophageal varices in patients with liver cirrhosis，to establish a predictive model，and to provide reasonable guidance for the prevention of early esophageal varices in patients with liver cirrhosis. ***Methods***：A retrospective analysis was performed for 1 113 patients with liver cirrhosis who attended the hospitals in Chongqing，China from December 2006 to May 2021. Recursive feature elimination（RFE）and four machine learning methods were used for the screening of features，and five machine learning predictive models were established by logistic regression，random forest，support vector machine（SVM），decision tree，and eXtreme Gradient Boosting（XGBoost）. The receiver operating characteristic（ROC）curve was used to evaluate the performance of each model，and the model with the best performance was used to investigate the risk factors for esophageal varices in patients with liver cirrhosis. SHAP plots were used to explain the impact of each risk factor on patients. ***Results***：The XGBoost model showed the best performance in predicting the risk of esophageal varices in patients with liver cirrhosis，with an area under the ROC curve of 0.872（95%CI=0.813-0.918）. SHAP plots showed that platelet count，diameter of the portal vein，cholinesterase，albumin，alanine aminotransferase，hemoglobin，prothrombin ratio，prothrombin time，and serum total protein were risk factors for esophageal varices in patients with liver cirrhosis. ***Conclusion***：This study shows that the XGBoost predictive model has a relatively high predictive value，and the risk factors obtained by this model have a certain guiding significance for the clinical prevention and treatment of early esophageal varices in patients with liver cirrhosis.

【**Key words**】machine learning；liver cirrhosis；esophageal varices；risk factors

作者介绍：尹钰恒，Email：yyh2290573902@163.com，研究方向：医学大数据分析。
通信作者：王　伟，Email：wwzqbc@cqmu.edu.cn。
基金项目：重庆医科大学未来医学青年创新资助项目（编号：W0150）。
优先出版：https://link.cnki.net/urlid/50.1046.R.20241115.1318.002（2024-11-18）

肝硬化是一种慢性、进行性肝脏疾病，其主要特征是肝组织纤维化和肝功能下降[1]。肝硬化常由慢性肝炎、长期酗酒和非酒精性脂肪性肝病等因素引起[2]。在全球范围内，肝硬化对公共健康造成了严重威胁[3]。随着病情的发展，肝硬化患者常会出现多种并发症，其中食管静脉曲张是一种常见且危险的并发症[4]。肝硬化患者的食管静脉曲张终生患病率为60%~80%，急性食管静脉曲张破裂出血导致的死亡率为15%~55%[5-6]。目前，对肝硬化患者相关风险因素的预防和治疗手段在近几年中降低了食管静脉曲张的发病率[7-8]。因此，为了有效管理和治疗肝硬化的并发症，持续探索肝硬化患者早期并发食管静脉曲张的风险因素以及准确预测食管静脉曲张的发生具有重要意义。

随着大数据时代的到来，机器学习在肝病领域的应用前景广阔[9-11]。通过对医疗数据的深度挖掘和分析，越来越多机器学习研究在肝硬化并发食管静脉曲张人群中开发和验证[12-13]。然而大多数研究仅使用单一机器学习算法和传统的特征筛选方法[14-15]。这些方法往往受限于样本量、数据收集的全面性以及模型的复杂度，导致预测模型的准确性和泛化能力有限。此外，机器学习算法在构建模型时往往被视为黑箱即模型的内部机制难以被理解和解释，使得直接从模型中提取可解释信息异常困难。Lundberg SM 等[16]提出基于博弈论的 SHapley Additive exPlanations（SHAP）方法，提供了统一的度量方式来刻画每个特征对模型输出的贡献大小。通过揭示模型的决策机制，SHAP 增强了对模型预测结果的可信度，使其在临床应用中更加可靠和易于解释。在既往肝硬化患者并发食管静脉曲张的研究中，往往缺乏对所构建模型的解释，导致其难以在临床上得到进一步实践[17-19]。

因此，本研究的目的是开发和验证 1 个预测模型，用于评估肝硬化患者并发食管静脉曲张的风险因素。随着机器学习的发展，支持向量机（support vector machine，SVM）、随机森林、极端梯度提升（eXtreme Gradient Boosting，XGBoost）等模型在肝病相关疾病的预测上取得了较明显的性能提升[20-22]。这些研究结果为本研究选择模型进行肝硬化并发症的预测提供了坚实的理论基础和实验依据。因此，本研究纳入了逻辑回归、决策树、随机森林、SVM、XGBoost 5 种机器学习模型用于肝硬化患者并发食管静脉曲张的预测研究。通过递归特征消除结合机器学习的方法筛选特征变量，比较并使用 5 种机器学习算法，以克服以往研究的局限性。此外，使用 SHAP 方法对肝硬化患者并发食管静脉曲张的风险因素进一步解释，提升机器学习模型在临床应用中的可行性。研究发现将有助于采取及时且有针对性的干预措施，促进患者康复，为临床医生提供早期干预的指导，并为该领域未来的研究提供重要参考。

1 资料与方法

1.1 一般资料

回顾性收集 2006 年 12 月至 2021 年 5 月在重庆医渡云大数据平台上有电子健康记录的肝硬化患者，共计 1 113 例。纳入标准：①经国际公认的诊断标准（ICD-10 编码）诊断为肝硬化；②有明确的胃镜检查记录。排除标准：①不符合上述纳入标准任意一条者；②肝硬化患者确诊食管静脉曲张的时间间隔<1 周；③数据记录不完整或存在明显错误的患者。

收集电子平台上记录符合标准患者的综合数据，主要包括以下内容：①人口统计学变量（如年龄、性别、吸烟史、饮酒史、冠心病史等）；②实验室检查变量（如丙氨酸氨基转移酶、胆碱酯酶、白蛋白、血红蛋白、门静脉内径等）；③肝硬化评分等级（A、B、C 3 个等级）。数据集中的特征变量缺失值未超过 20%，其余的变量缺失通过随机森林填补法进行填补。

1.2 统计学方法

使用 Python 3.9.12 和 R 4.3.3 完成所有数据的统计分析。计量资料采用平均值 ± 标准差或中位数描述，计数资料用频率描述。使用 t 检验或非参数检验比较连续变量和卡方检验比较分类变量（检验水准 α=0.05）。本研究以 8∶2 的比例随机分为训练集和测试集。采用 4 种机器学习算法结合递归特征消除（recursive feature elimination，RFE）的方法筛选预测因子，以得到最优的特征子集用于预测。为了使模型达到最佳性能，研究使用逻辑回归、决策树、随机森林、SVM、XGBoost 等 5 种机器学习算法对模型进行构建和训练，并通过随机搜索的方法寻找最佳超参数组合。模型的效果评价采用准确率、特异性、召回率、F1 值和受试者工作特征（receiver operating characteristic，ROC）曲线下面积等指数。最后通过 SHAP 方法绘制各特征在 XGBoost 模型中的蜂群图和条形图，更直观地展示各特征对模型预测的贡献度，使模型的可解释性更强。具体技术路线见图 1。

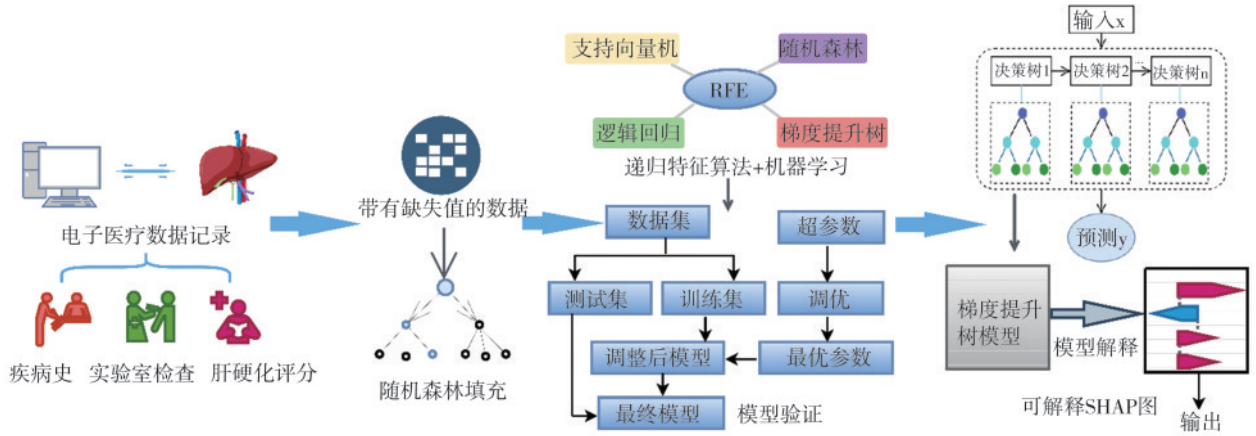


图1　技术路线图

2　结　果

2.1　患者基线资料

本研究共纳入1 113例肝硬化患者，将患者按是否患有食管静脉曲张分为食管静脉曲张组和非食管静脉曲张组。2组间患者的基本特征比较如表1所示。食管静脉曲张患者与非食管静脉曲张患者的平均年龄分别为[(49.91 ± 10.77)岁 vs.(48.76 ± 10.06)岁，t=−1.756，P=0.079]。其中饮酒（F=14.152，P<0.001）、丙氨酸氨基转移酶（F=3.476，P<0.001）、天门冬氨酸氨基转移酶（t=2.418，P=0.016）、前白蛋白（t=9.888，P<0.001）、碱性磷酸酶（t=−3.178，P=0.002）、胆碱酯酶（t=12.897，P<0.001）、白蛋白（t=10.549，P<0.001）、总蛋白（t=5.964，P<0.001）、白细胞计数（t=5.129，P<0.001）、红细胞计数（t=11.694，P<0.001）、淋巴细胞计数（t=8.794，P<0.001）、血红蛋白（t=12.694，P<0.001）、血小板计数（t=11.725，P<0.001）、凝血酶原比值（t=−8.550，P<0.001）、凝血酶原时间（t=1.474，P<0.001）、纤维蛋白原（t=6.600，P<0.001）、活化部分凝血活酶时间（t=−3.760，P<0.001）、活化部分凝血活酶时间比值（t=−3.165，P<0.001）、肝纤维化指数（t=−7，350，P<0.001）、门静脉内径（t=−12.749，P<0.001）、乙型肝炎病毒核心抗体（HBcAb，F=12.845，P<0.001）、肝硬化评分（Child-Pugh分级，F=35.137，P<0.001）在食管静脉曲张组和非食管静脉曲张组之间的差异具有统计学意义。

表1　食管静脉曲张组和非食管静脉曲张组患者基本特征比较（$\bar{x} \pm s$；n，%）

基线资料	食管静脉曲张组（n=699）	非食管静脉曲张组（n=414）	t/χ^2值	P值
年龄（岁）	49.91 ± 10.77	48.76 ± 10.06	−1.756	0.079
性别			0.595	0.440
男	513（73.39）	295（71.26）		
女	186（26.61）	119（28.74）		
吸烟			0.621	0.431
是	292（41.77）	163（39.37）		
否	407（58.23）	251（60.63）		
饮酒			14.152	<0.001
是	270（38.63）	114（27.54）		
否	429（61.37）	300（72.46）		
冠心病			0.054	0.817
是	17（2.43）	11（2.66）		
否	682（97.57）	403（97.34）		
高血压			0.793	0.373
是	52（7.44）	37（8.94）		
否	647（92.56）	377（91.06）		
糖尿病			2.623	0.105
是	92（13.16）	41（9.90）		

续表1

基线资料	食管静脉曲张组（n=699）	非食管静脉曲张组（n=414）	t/χ^2值	P值
否	607（86.84）	373（90.10）		
丙氨酸氨基转移酶（U/L）	95.69 ± 237.88	155.31 ± 297.12	3.476	<0.001
天门冬氨酸氨基转移酶（U/L）	100.52 ± 209.12	139.08 ± 281.83	2.418	0.016
γ-谷氨酰基转移酶（U/L）	97.95 ± 130.74	94.58 ± 148.04	−0.395	0.693
直接胆红素（μmol/L）	39.52 ± 57.49	38.81 ± 64.03	−0.191	0.849
前白蛋白（mg/L）	117.48 ± 68.99	166.64 ± 86.63	9.844	<0.001
总胆汁酸（μmol/L）	42.16 ± 59.29	35.08 ± 57.06	−1.951	0.051
碱性磷酸酶（U/L）	125.19 ± 77.90	112.61 ± 53.78	−3.178	0.002
胆碱酯酶（kU/L）	4.09 ± 2.09	6.09 ± 2.70	12.892	<0.001
白蛋白（g/L）	36.67 ± 7.36	41.27 ± 6.85	10.549	<0.001
总蛋白（g/L）	71.45 ± 9.78	74.78 ± 8.51	5.964	<0.001
白细胞计数（$\times10^9$/L）	4.50 ± 2.35	5.22 ± 2.06	5.129	<0.001
红细胞计数（$\times10^{12}$/L）	3.95 ± 0.84	4.52 ± 0.75	11.594	<0.001
淋巴细胞计数（$\times10^9$/L）	1.08 ± 0.65	1.49 ± 0.89	8.794	<0.001
血红蛋白（g/L）	119.47 ± 29.41	139.04 ± 21.70	12.694	<0.001
血小板计数（$\times10^9$/L）	79.52 ± 50.43	121.91 ± 62.50	11.725	<0.001
凝血酶时间（s）	18.66 ± 2.69	19.31 ± 8.72	1.474	0.141
凝血酶原比值	1.27 ± 0.25	1.15 ± 0.21	−8.550	<0.001
凝血酶原时间（s）	16.76 ± 3.33	15.16 ± 2.86	−8.437	<0.001
纤维蛋白原（g/L）	2.26 ± 0.82	2.62 ± 0.92	6.600	<0.001
活化部分凝血活酶时间（s）	41.59 ± 6.98	40.03 ± 6.19	−3.760	<0.001
活化部分凝血活酶时间比值	1.28 ± 0.22	1.23 ± 0.19	−3.615	<0.001
肝纤维化指数	8.10 ± 7.23	5.10 ± 6.16	−7.350	<0.001
APRI评分	3.43 ± 6.30	3.14 ± 8.60	−0.630	0.529
门静脉内径（mm）	1.35 ± 0.31	1.15 ± 0.22	−12.749	<.001
HBsAg			0.532	0.466
阳性	673（96.28）	402（97.10）		
阴性	26（3.72）	12（2.90）		
HBeAg			1.560	0.212
阳性	177（25.32）	119（28.74）		
阴性	522（74.68）	295（71.26）		
HBcAb			12.845	<0.001
阳性	195（27.90）	76（18.36）		
阴性	504（72.10）	338（81.64）		
HBsAb			0.295	0.587
阳性	64（9.16）	42（10.14）		
阴性	635（90.84）	372（89.86）		
HBeAb			0.555	0.456
阳性	208（29.76）	132（31.88）		
阴性	491（70.24）	282（68.12）		
Child-pugh分级			35.137	<0.001
A级	385（55.08）	302（72.95）		
B级	244（34.91）	87（21.01）		
C级	70（10.01）	25（6.04）		

2.2 特征变量筛选

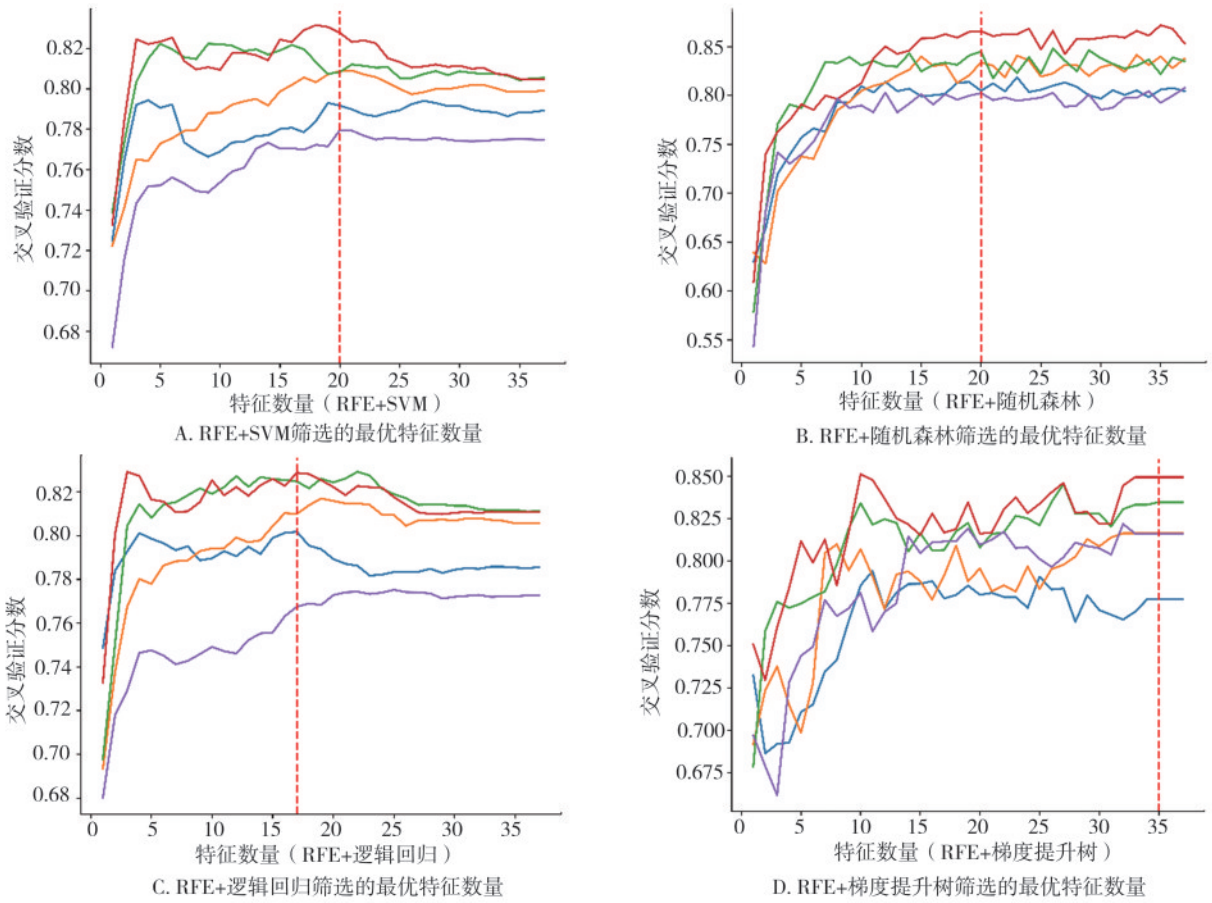
对比2组患者的基线特征时，发现差异具有统计学意义的特征较多。为了从原始数据集得到对肝硬化患者并发食管静脉曲张影响最明显的特征，本研究进行了特征选择。为了避免基线变量中不同量纲之间的影响，本研究对连续型变量进行标准化处理，对离散型变量进行重编码。采用RFE结合4种机器学习的方法进行特征选择，并进行五折交叉验证。随着纳入的特征数量增加，RFE结合4种机器学习模型的准确率变化如图2所示。图2A和图2B表明当特征数量达到20时，RFE+SVM和RFE+随机森林模型的准确率达到最高；图2C表明特征数量达到17时，RFE+逻辑回归模型的准确率最高；图2D表明特征数量达到35时，RFE+XGBoost模型准确率最高。图3A显示了4种方法筛选特征的最优子集，对应的特征集合数量为9个。最终选择的特征集合为血红蛋白、血小板计数、门静脉内径、胆碱酯酶、白蛋白、谷丙转氨酶、凝血酶原比值、凝血酶原时间和血清总蛋白。

2.3 预测模型建立与验证

将筛选出来的血红蛋白、血小板计数、门静脉内径、胆碱酯酶、白蛋白、谷丙转氨酶、凝血酶原比值、凝血酶原时间和血清总蛋白等9个特征变量纳入5种机器学习模型进行训练，并通过随机搜索的方式优化模型的超参数。使用5倍交叉验证评估模型的预测性能，以达到模型的最佳表现。如图3B所示，XGBoost树模型的表现明显优于其他4个模型，ROC曲线下面积为0.872（95%CI=0.813~0.918），准确率为0.803，特异性为0.631，召回率为0.906，F1值为0.851。随机森林、逻辑回归和SVM模型的ROC曲线下面积（area under curve，AUC）均>0.8。5种机器学习模型在测试集上的具体表现见表2。

2.4 SHAP的模型解释

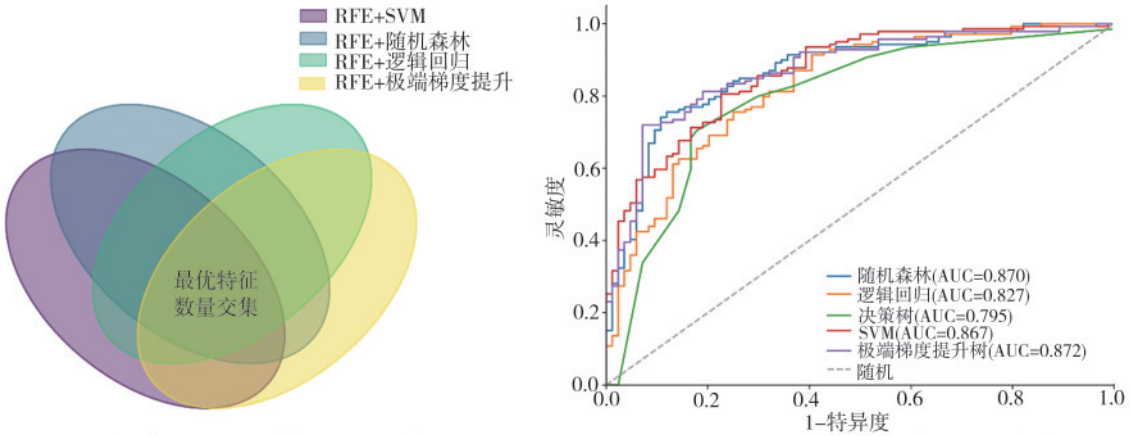
为了使模型具有较强的可解释性，使用SHAP方法展示最优XGBoost模型中各特征的重要性程度。图4A列出了根据平均绝对SHAP值排序的9个风险因素，依次是血小板计数、门静脉内径、胆碱酯酶、白蛋白、谷丙转氨酶、血红蛋白、凝血酶原比值、凝血酶原时间和血清总蛋白。图4B说明了影响因素对肝硬化患者早期并发食管静脉曲张的影响大小。y轴表示风险因素的值，x轴表示风险因素对发生食管静脉曲张的影响。结果表明血小板计数水平较低、门静脉内径较大、胆碱酯酶水平较低、白蛋白和血红蛋白水平较低、谷丙转氨酶水平较低、凝血酶原比值、凝血酶原时间及血清总蛋白水平较高都会增加发生肝硬化患者早期并发食管静脉曲张的风险。



A. RFE+SVM筛选的最优特征数量　B. RFE+随机森林筛选的最优特征数量

C. RFE+逻辑回归筛选的最优特征数量　D. RFE+梯度提升树筛选的最优特征数量

图2 RFE结合4种机器学习的五折交叉验证

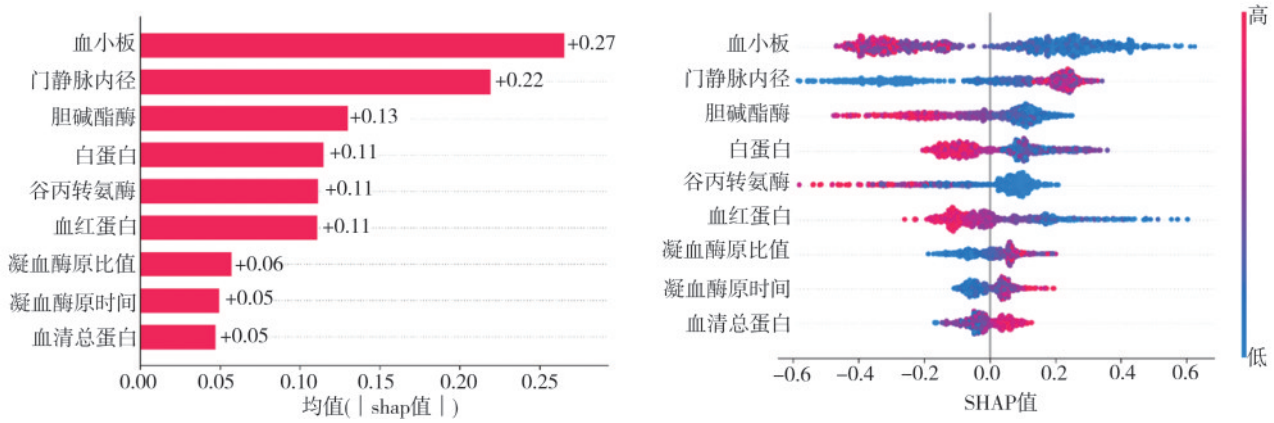


A. RFE结合4种机器学习筛选的特征交集　　B. 5种机器学习模型在测试集上的ROC曲线

图3　特征筛选交集和测试集上的模型验证

表2　5种机器学习模型在测试集上的性能表现

预测模型	准确率	特异性	召回率	F1	AUC(95%CI)
随机森林	0.803	0.667	0.885	0.848	0.866(0.819~0.907)
逻辑回归	0.762	0.631	0.842	0.815	0.827(0.773~0.873)
决策树	0.753	0.631	0.827	0.807	0.795(0.738~0.864)
SVM	0.785	0.667	0.856	0.832	0.867(0.817~0.906)
梯度提升树	0.803	0.631	0.906	0.851	0.872(0.813~0.918)



A. SHAP条形图　　B. SHAP蜂群图

图4　可解释SHAP图

3　讨　论

本研究旨在探讨肝硬化患者早期并发食管静脉曲张的风险因素，通过RFE筛选特征变量，并构建机器学习预测模型。结合SHAP的方法来分析食管静脉曲张相关风险因素的相对重要性，进一步提高模型的可解释性。目前已有一些用于评估肝硬化患者并发食管静脉曲张的风险因素的研究。例如Dong TS等[23]收集238例肝硬化病人的临床信息，利用随机森林算法开发了EVendo评分。EVendo评分用于识别肝硬化患者并发食管静脉曲张的AUC值为0.820。Yan YJ等[24]和王伟等[25]分别招募了218例和330例肝硬化患者，利用图像特征方法来识别肝硬化患者中轻度食管静脉曲张。基于图像特征的模型在内部验证数据集上AUC值分别为0.732和0.782。相对于之前的研究，本研究在样本量、特征选择的方法、模型的可解释性上具有一定的优势。

本次研究一共纳入1 113例肝硬化患者，有利于更准确地反映患者的特征和提高研究结果的可靠性。在特征选择方面，RFE不需要对特征的分布做出提前假设，能够减少模型的复杂性和降低过拟合的风险。因此，研究中采用了RFE结合4种机器学习方法，利用最终取得的交集来建立预测模型。在建立的5种机器学习模型里，XGBoost模型的性能最佳，AUC为0.872。在模型解释性方面，采用的SHAP图有利于增强机器学习模型的可解释性。

SHAP值结果表明肝硬化患者并发食管静脉曲张的风险因素包括血小板计数、门静脉内径、胆碱酯酶、白蛋白、谷丙转氨酶、血红蛋白、凝血酶原比值、凝血酶原时间和血清总蛋白。在纳入的风险因素中，血小板计数对模型预测的贡献最大，可能表示血小板计数对肝硬化患者并发食管静脉曲张有明显的影响。目前已有的研究和共识都表明血小板计数减少会增加肝硬化患者并发食管静脉曲张的风险以及不良结局[26-28]。肝硬化患者中的门静脉内径与门静脉高压症相关，而门静脉高压是导致食管静脉曲张的重要原因[29]。本研究中表明门静脉内径增大对肝硬化患者并发食管静脉曲张有重要影响，与相关的研究结果一致[30]。胆碱酯酶是一种出色的肝硬化生物标志物，低水平的胆碱酯酶可以作为晚期肝病的可靠预后标志[31-32]。本研究结果SHAP图表明胆碱酯酶在肝硬化患者并发食管静脉曲张的风险因素重要性中排第3位，揭示了胆碱酯酶有望作为肝硬化患者并发食管静脉曲张的重要预测因子。此外，本研究指出血红蛋白、白蛋白、凝血酶原比值、凝血酶原时间与肝硬化患者并发食管静脉曲张的风险密切相关，与郑文凤等[33]的研究结果一致。分析由于肝脏是合成白蛋白的主要器官，肝功能受损时会导致白蛋白合成量减少，造成低蛋白血症。凝血酶原比值和凝血酶原时间是反映肝脏的合成功能的指标，凝血酶原比值增大和凝血酶原时间延长都表明凝血因子的合成出现障碍。因此，血红蛋白和白蛋白水平较低、凝血酶原比值和时间会增加肝硬化患者发生食管静脉曲张的风险，早期应更密切地监测肝硬化患者的血液指标变化。

本研究围绕肝硬化患者早期并发食管静脉曲张的相关风险因素进行分析，构建了5种机器学习模型，对预测肝硬化患者并发食管静脉曲张的最佳模型进行了解释性分析。该模型可解释机器学习模型对于识别肝硬化人群中早期食管静脉曲张病例以及辅助临床医生做出适当的治疗决策具有重要意义。目前的研究还存在一些局限性。首先，本研究是1项回顾性研究，因此不可避免地存在一定的偏倚。此外，本研究尚未对所构建的预测模型进行外部验证，这使得模型的泛化能力仍然存在不确定性，亟需进一步的验证。未来，本研究计划收集足够的外部验证数据集，以进一步优化该模型。

综上所述，XGBoost模型预测肝硬化患者早期并发食管静脉曲张效能最优，揭示了其相关的风险因素，可为肝硬化患者早期食管静脉曲张的临床决策提供参考。

参 考 文 献

[1] Ginès P, Krag A, Abraldes JG, et al. Liver cirrhosis[J]. Lancet, 2021, 398(10308): 1359-1376.

[2] Huang DQ, Terrault NA, Tacke F, et al. Global epidemiology of cirrhosis-aetiology, trends and predictions[J]. Nat Rev Gastroenterol Hepatol, 2023, 20(6): 388-398.

[3] Huang DQ, Mathurin P, Cortez-Pinto H, et al. Global epidemiology of alcohol-associated cirrhosis and HCC: trends, projections and risk factors[J]. Nat Rev Gastroenterol Hepatol, 2023, 20(1): 37-49.

[4] Tapper EB, Parikh ND. Diagnosis and management of cirrhosis and its complications: a review[J]. JAMA, 2023, 329(18): 1589-1602.

[5] Garcia-Tsao G, Abraldes JG, Berzigotti A, et al. Portal hypertensive bleeding in cirrhosis: risk stratification, diagnosis, and management: 2016 practice guidance by the American Association for the study of liver diseases[J]. Hepatology, 2017, 65(1): 310-335.

[6] Rigo GP, Merighi A, Chahin NJ, et al. A prospective study of the ability of three endoscopic classifications to predict hemorrhage from esophageal varices[J]. Gastrointest Endosc, 1992, 38(4): 425-429.

[7] Garbuzenko DV, Arefyev NO. Primary prevention of bleeding from esophageal varices in patients with liver cirrhosis: an update and review of the literature[J]. J Evid Based Med, 2020, 13(4): 313-324.

[8] Marasco G, Dajti E, Ravaioli F, et al. Spleen stiffness measurement for assessing the response to β-blockers therapy for high-risk esophageal varices patients[J]. Hepatol Int, 2020, 14(5): 850-857.

[9] 蔡偌欣，马亚楠，闻德亮. 机器学习在非酒精性脂肪肝预测中的应用[J]. 中国卫生统计，2024，41(2)：316-318.
Cai RX, Ma YN, Wen DL. Application of machine learning in prediction of nonalcoholic fatty liver disease[J]. Chin J Health Stat, 2024, 41(2): 316-318.

[10] 黄良江，毛德文，郑景辉，等. 人工智能在肝性脑病风险预测模

型中的应用进展[J]. 实用医学杂志，2024，40(3)：289-294.
Huang LJ，Mao DW，Zheng JH，et al. Application progress of artificial intelligence in risk prediction model of hepatic encephalopathy[J]. J Pract Med，2024，40(3)：289-294.
[11] Le Berre C，Sandborn WJ，Aridhi S，et al. Application of artificial intelligence to gastroenterology and hepatology[J]. Gastroenterology，2020，158(1)：76-94.
[12] Agarwal S，Sharma S，Kumar M，et al. Development of a machine learning model to predict bleed in esophageal varices in compensated advanced chronic liver disease：a proof of concept[J]. J Gastroenterol Hepatol，2021，36(10)：2935-2942.
[13] Bayani A，Hosseini A，Asadi F，et al. Identifying predictors of varices grading in patients with cirrhosis using ensemble learning[J]. Clin Chem Lab Med，2022，60(12)：1938-1945.
[14] 林子淮，李欣明，全显跃. 基于机器学习 CT 影像组学特征联合血清学特征模型预测高出血风险食管静脉曲张 [J]. 中国 CT 和 MRI 杂志，2023，21(6)：93-96.
Lin ZH，Li XM，Quan XY. Predicting high bleeding risk esophageal varices based on machine learning CT imaging features combined with a serological profile model[J]. Chinese Journal of CT and MRI，2023，21(6)：93-96.
[15] Malik S，Tenorio BG，Moond V，et al. Systematic review of machine learning models in predicting the risk of bleed/grade of esophageal varices in patients with liver cirrhosis：a comprehensive methodological analysis[J]. J Gastroenterol Hepatol，2024.
[16] Lundberg SM，Erion G，Chen H，et al. From local explanations to global understanding with explainable AI for trees[J]. Nat Mach Intell，2020，2(1)：56-67.
[17] Şİmşek C，Tekİn E，Sahİn H，et al. Artificial intelligence to predict esophageal varices in patients with cirrhosis[J]. Acibadem Univ Saglik Bilimleri Dergisi，2021，12(3)：625-629.
[18] Abd El-Salam SM，Ezz MM，Hashem S，et al. Performance of machine learning approaches on prediction of esophageal varices for Egyptian chronic hepatitis C patients[J]. Inform Med Unlocked，2019，17：100267.
[19] 王 伟，李文杰，李佳晨. 基于社交信息扩散的网络传染病传播防控研究[J]. 重庆医科大学学报，2023，48(12)：1393-1400.
Wang W，Li WJ，Li JC. Prevention and control strategy for network infectious disease spreading based on social information diffusion[J]. J Chongqing Med Univ，2023，48(12)：1393-1400.
[20] Sato M，Morimoto K，Kajihara S，et al. Machine-learning approach for the development of a novel predictive model for the diagnosis of hepatocellular carcinoma[J]. Sci Rep，2019，9(1)：7704.
[21] Wang J，Wang ZQ，Chen MK，et al. An interpretable artificial intelligence system for detecting risk factors of gastroesophageal variceal bleeding[J]. NPJ Digit Med，2022，5(1)：183.
[22] Wu CC，Yeh WC，Hsu WD，et al. Prediction of fatty liver disease using machine learning algorithms[J]. Comput Methods Programs Biomed，2019，170：23-29.
[23] Dong TS，Kalani A，Aby ES，et al. Machine learning-based development and validation of a scoring system for screening high-risk esophageal varices[J]. Clin Gastroenterol Hepatol，2019，17(9)：1894-1901. e1.
[24] Yan YJ，Li Y，Fan CL，et al. A novel machine learning-based radiomic model for diagnosing high bleeding risk esophageal varices in cirrhotic patients[J]. Hepatol Int，2022，16(2)：423-432.
[25] 王 伟，李文杰，李佳晨. 基于社交信息扩散的网络传染病传播防控研究门[J]. 重庆医科大学学报，2023，48(12)：1393-1400.
Wang W，Li W，Li JC. Prevention and control strategy fornetwork infectious disease spreading based on socialinformation diffusionU]. JChongging Med Univ，2023，48(12)：1393-1400.
[26] 张俊文，段 红. 肝硬化患者血小板功能的研究[J]. 重庆医科大学学报，2000，25(2)：186-187，193.
Zhang JW，Duan H. Study on the platelet functions in patients with liver cirrhosis[J]. J Chongqing Med Univ，2000，25(2)：186-187，193.
[27] Schepis F，Cammà C，Niceforo D，et al. Which patients with cirrhosis should undergo endoscopic screening for esophageal varices detection?[J]. Hepatology，2001，33(2)：333-338.
[28] Zaman A，Becker T，Lapidus J，et al. Risk factors for the presence of varices in cirrhotic patients without a history of variceal hemorrhage [J]. Arch Intern Med，2001，161(21)：2564-2570.
[29] Wang CY，Huang YF，Liu CC，et al. Diagnosis of clinically significant portal hypertension using CT- and MRI-based vascular model[J]. Radiology，2023，307(2)：e221648.
[30] 杨舒馨影，李 涛. 脑出血患者循环血中 HBP、ALT、AST、WBC、NEU、PLT 表达与感染及预后的相关性[J]. 重庆医科大学学报，2024，49(10)：1119-1124.
Yang SXY，Li T. Correlation of HBP，ALT，AST，WBC，NEU，and PLT Expression in Circulating Blood with Infection and Prognosis in Patients with Intracerebral Hemorrhage[J]. J Chongging Med Univ，2024，49(10)：1119-1124.
[31] 罗 亚，黄维嘉. 肝病患者血清胆碱酯酶测定及意义[J]. 重庆医科大学学报，2001，26(2)：199-200.
Luo Y，Huang WJ. Measurement of serum cholinesterase in patients with hepatopathy[J]. J Chongqing Med Univ，2001，26(2)：199-200.
[32] Ramachandran J，Sajith KG，Priya S，et al. Serum cholinesterase is an excellent biomarker of liver cirrhosis[J]. Trop Gastroenterol，2014，35(2)：15-20.
[33] 郑文凤，殷琳琳，李山林，等. 肝硬化食管静脉曲张破裂出血的危险因素分析[J]. 医药论坛杂志，2024，45(10)：1050-1053.
Zheng WF，Yin LL，Li SL，et al. Analysis of risk factors for esophageal variceal bleeding in patients with liver cirrhosis[J]. Journal of Medical Forum，2024，45，(10)：1050-1053.

（责任编辑：周一青）

本文引用格式：
尹钰恒，王雨雯，范 捷，等. 基于可解释机器学习的肝硬化患者早期并发食管静脉曲张风险预测研究[J]. 重庆医科大学学报，2025，50(3)：389-396.