

机器学习和 Logistic 回归模型预测心脏外科术后患者发生急性肾损伤的比较分析

—— 基于 MIMIC-Ⅲ 数据库的资料分析

熊伟¹ 张力凡² 余凯¹ 许果¹ 白上林¹ 刘炫¹

¹ 四川绵阳四〇四医院心脏大血管外科, 四川绵阳 621000; ² 四川大学华西临床医学院, 成都 610041

通信作者: 熊伟, Email: xiongwei1972@163.com

【摘要】 目的 使用机器学习中极限梯度提升(XGBoost)算法构建心脏外科术后患者发生急性肾损伤(AKI)的风险预测模型,探讨心脏外科术后患者发生 AKI 的危险因素和保护因素。**方法** 纳入美国重症监护医学信息数据库Ⅲ(MIMIC-Ⅲ)中全部接受心脏外科手术患者的临床资料,按术后 14 d 内是否发生 AKI 分为 AKI 组和非 AKI 组,并比较两组患者的临床特征。在五折交叉验证的基础上,分别采用 XGBoost 和 Logistic 回归法建立心脏外科术后 AKI 预测模型,并比较两种模型的受试者工作特征曲线下面积(AUC)。采用沙普利加和解释法(SHAP)解释 XGBoost 的输出模型。**结果** 共纳入 6 912 例心脏外科术后患者,其中 5 681 例(82.2%)在术后 14 d 内发生 AKI, 1 231 例(17.8%)未发生 AKI。与非 AKI 组相比,AKI 组患者年龄更大[岁:68.0(59.0, 76.0)比 62.0(52.0, 71.0)],急诊入院及合并肥胖、糖尿病比例更高(52.4% 比 47.8%, 9.0% 比 4.0%, 32.0% 比 22.2%),生命体征指标中呼吸频率(RR)更低[次/min:17.0(14.0, 20.0)比 19.0(15.0, 22.0)],心率(HR)更慢[次/min:80.0(67.0, 89.0)比 82.0(71.5, 93.0)],血压更高[mmHg(1 mmHg≈0.133 kPa):80.0(70.7, 90.0)比 78.0(70.0, 88.0)],实验室指标中血红蛋白(Hb)、血糖、血 K⁺、血肌酐(SCr)更高[Hb(g/L):122.0(109.0, 136.0)比 120.0(106.0, 135.0),血糖(mmol/L):7.3(6.1, 8.9)比 6.8(5.7, 8.5),血 K⁺(mmol/L):4.2(3.9, 4.7)比 4.2(3.8, 4.6),SCr(μmol/L):88.4(70.7, 106.1)比 79.6(70.7, 97.2)],白蛋白(ALB)和三酰甘油(TG)更低[ALB(g/L):38.0(35.0, 41.0)比 39.0(37.0, 42.0),TG(mmol/L):1.4(1.0, 2.0)比 1.5(1.0, 2.2)],且多器官功能障碍综合征(MODS)和脓毒症比例更高(30.6% 比 16.2%, 3.3% 比 1.9%),差异均有统计学意义(均 $P < 0.05$)。Logistic 回归预测模型中的主要影响因素包括血乳酸[Lac;优势比(OR)=1.062, 95% 可信区间(95%CI)为 1.030~1.100, $P=0.050$]、肥胖(OR=2.234, 95%CI 为 1.900~2.640, $P < 0.001$)、男性(OR=0.858, 95%CI 为 0.794~0.928, $P=0.049$)、伴有糖尿病(OR=1.820, 95%CI 为 1.680~1.980, $P < 0.001$)和急诊入院(OR=1.278, 95%CI 为 1.190~1.380, $P < 0.001$)。受试者工作特征曲线(ROC 曲线)分析显示,Logistic 回归模型预测心脏外科术后发生 AKI 的 AUC 为 0.62(95%CI 为 0.61~0.67)。经过网格搜索与五折交叉验证结合优化 XGBoost 模型参数,模型训练效果良好,没有出现拟合或欠拟合。ROC 曲线分析结果显示, XGBoost 模型预测心脏外科术后发生 AKI 的 AUC 为 0.77(95%CI 为 0.75~0.80),明显高于 Logistic 回归预测模型的 AUC($P < 0.01$)。经 SHAP 方法处理后, XGBoost 输出模型中对最终结果最重要的预测因素是年龄和 ALB,其中年龄是危险因素(平均|SHAP 值|为 0.434), ALB 是保护因素(平均|SHAP 值|为 0.221)。**结论** 年龄是心脏外科术后患者发生 AKI 的重要危险因素,而 ALB 则是保护因素。机器学习预测心脏外科术后 AKI 的效能比传统 Logistic 回归更加优秀,能分析变量与结局间更复杂的关系,更精准地个体化预测术后 AKI 的发生风险。

【关键词】 机器学习; 急性肾损伤; 预测模型

基金项目: 四川省绵阳市卫生健康委员会科研项目(202030)

DOI: 10.3760/cma.j.cn121430-20210223-00279

Comparison of machine learning and Logistic regression model in predicting acute kidney injury after cardiac surgery: data analysis based on MIMIC-Ⅲ database

Xiong Wei¹, Zhang Lifan², She Kai¹, Xu Guo¹, Bai Shanglin¹, Liu Xuan¹

¹Department of Cardiovascular Surgery, Sichuan Mianyang 404 Hospital, Mianyang 621000, Sichuan, China;

²Department of Clinical Medical College, West China Hospital of Sichuan University, Chengdu 610041, Sichuan, China

Corresponding author: Xiong Wei, Email: xiongwei1972@163.com

【Abstract】 Objective To establish an acute kidney injury (AKI) prediction model in patients after cardiac surgery by extreme gradient boosting (XGBoost) machine learning model, and to explore the risk and protective factors for AKI in patients after cardiac surgery. **Methods** All patients who underwent cardiac surgery in Medical Information Mart for Intensive Care-Ⅲ (MIMIC-Ⅲ) database were enrolled, and they were divided into AKI group and non-AKI group according to whether AKI developed within 14 days after cardiac surgery. Their clinical characteristics were

compared. Based on five-fold cross-validation, XGBoost and Logistic regression were used to establish the prediction model of AKI after cardiac surgery. And the area under the receiver operator characteristic curve (AUC) of the models was compared. The output model of XGBoost was interpreted by Shapley additive explanations (SHAP). **Results** A total of 6 912 patients were included, of which 5 681 (82.2%) developed AKI within 14 days after the operation, and 1 231 (17.8%) did not. Compared with the non-AKI group, the main characteristics of AKI group included older age [years: 68.0 (59.0, 76.0) vs. 62.0 (52.0, 71.0)], higher incidence of emergency admission and complicated with obesity and diabetes (52.4% vs. 47.8%, 9.0% vs. 4.0%, 32.0% vs. 22.2%), lower respiratory rate [RR; bpm: times/min: 17.0 (14.0, 20.0) vs. 19.0 (15.0, 22.0)], lower heart rate [HR; bpm: 80.0 (67.0, 89.0) vs. 82.0 (71.5, 93.0)], higher blood pressure [mmHg (1 mmHg $\approx$ 0.133 kPa): 80.0 (70.7, 90.0) vs. 78.0 (70.0, 88.0)], higher hemoglobin (Hb), blood glucose, blood K⁺ level and serum creatinine [SCr; Hb (g/L): 122.0 (109.0, 136.0) vs. 120.0 (106.0, 135.0), blood glucose (mmol/L): 7.3 (6.1, 8.9) vs. 6.8 (5.7, 8.5), blood K⁺ level (mmol/L): 4.2 (3.9, 4.7) vs. 4.2 (3.8, 4.6), SCr (μ mol/L): 88.4 (70.7, 106.1) vs. 79.6 (70.7, 97.2)], lower albumin (ALB) and triacylglycerol [TG; ALB (g/L): 38.0 (35.0, 41.0) vs. 39.0 (37.0, 42.0), TG (mmol/L): 1.4 (1.0, 2.0) vs. 1.5 (1.0, 2.2)] as well as higher incidence of multiple organ dysfunction syndrome (MODS) and sepsis (30.6% vs. 16.2%, 3.3% vs. 1.9%), with significant differences (all $P < 0.05$). In the output model of Logistic regression, important predictors were lactic acid [Lac; odds ratio (OR) = 1.062, 95% confidence interval (95%CI) was 1.030–1.100, $P = 0.005$], obesity (OR = 2.234, 95%CI was 1.900–2.640, $P < 0.001$), male (OR = 0.858, 95%CI was 0.794–0.928, $P = 0.049$), diabetes (OR = 1.820, 95%CI was 1.680–1.980, $P < 0.001$) and emergency admission (OR = 1.278, 95%CI was 1.190–1.380, $P < 0.001$). Receiver operator characteristic curve (ROC curve) analysis showed that the AUC of the Logistic regression model for predicting AKI after cardiac surgery was 0.62 (95%CI was 0.61–0.67). After optimizing the XGBoost model parameters by grid search combined with five-fold cross-validation, the model was trained well with no overfitting or overfitting. ROC analysis showed that the AUC of XGBoost model for predicting AKI after cardiac surgery was 0.77 (95%CI was 0.75–0.80), which was significantly higher than that of Logistic regression model ($P < 0.01$). After SHAP treatment, in the output model of XGBoost, age and ALB were the most important predictors of the final outcome, where age was the risk factor (average |SHAP| value was 0.434), and ALB was the protective factor (average |SHAP| value was 0.221). **Conclusions** Age is an important risk factor for AKI after cardiac surgery, and ALB is a protective factor. The performance of machine learning in predicting cardiac and vascular surgery-associated AKI is better than the traditional Logistic regression. XGBoost can analyze the more complex relationship between variables and outcomes, and can predict the risk of postoperative AKI more accurately and individually.

[Key words] Machine learning; Acute kidney injury; Prediction model

Fund program: Scientific Research Project of Mianyang Municipal Health Commission of Sichuan Province (202030)

DOI: 10.3760/cma.j.cn121430-20210223-00279

急性肾损伤 (acute kidney injury, AKI) 是住院患者的常见综合征, 表现为血肌酐 (serum creatinine, SCr) 突然升高和尿量减少^[1]。AKI 的发生与多种因素有关, 如脱水、肾脏毒性药物等^[2]。同时, AKI 也是术后常见并发症, 是手术患者死亡的独立危险因素^[3]。据统计, 术后 AKI 发生率居医院获得性 AKI 第 2 位, 其中心脏外科术后 AKI 发生率较高, 且预后较差^[3]。有研究表明, 心脏外科术后 AKI 的发生率为 20% ~ 60%, AKI 相关病死率也高于其他手术^[4]。因此, 国际上提出了“心血管外科手术相关性急性肾损伤” (cardiac and vascular surgery-associated acute kidney injury, CVS-AKI) 的概念^[5]。这些都表明了心脏外科术后 AKI 的高发生率和高病死率。因此, 早期预测心脏外科术后 AKI 的发生并及时给予适当的干预措施, 对于有效降低患者病死率具有十分重要的意义。

目前, 已有一些研究者提出使用预测模型来预测患者心脏外科术后 AKI 的发生^[6-7], 但是由于样本量较少、模型预测效能不佳等因素, 这些模型的

开发和验证并不理想^[8]。近年来, 机器学习在医学领域发展迅速, 已被广泛应用于图像识别、临床不良事件预测等领域^[9]。本研究基于美国重症监护医学信息数据库 III (Medical Information Mart for Intensive Care- III, MIMIC- III) 中的数据, 拟采用机器学习中极限梯度提升 (extreme gradient boosting, XGBoost) 算法对患者心脏外科术后 AKI 的发生进行预测, 并与传统的 Logistic 回归分析进行比较, 探究导致心脏外科术后 AKI 的危险因素, 以期为中心外科术后患者管理提供新的思路与方法。

1 资料与方法

1.1 数据来源: MIMIC- III 是一个大型综合数据库, 由美国麻省理工学院和哈佛医学院联合开发。本研究中收集了 2001 至 2012 年马萨诸塞州波士顿贝斯以色列女执事医疗中心重症监护病房 (intensive care unit, ICU) 收治患者的相关信息^[10]进行回顾性队列研究。该数据库包括患者的基本信息、生命体征和实验室检测结果等。本研究是对第三方匿名公开数据库的分析, 根据《健康保险可携带性和责任法案》

(Health Insurance Portability and Accountability Act, HIPAA), 使用 MIMIC-III 数据库中的数据进行研究而不构成人体受试者研究。研究者在研究前已获得该数据库使用权限(证书编号: 51586214); 同时本研究符合医学伦理学标准, 获得四川绵阳四〇四医院医学伦理委员会的审批(审批号: 2021-01)。本研究纳入了 MIMIC-III 数据库中所有接受过心脏外科手术的患者, 术后是否发生 AKI 的观察期为患者接受手术后 14 d 内, 根据术后 14 d 内是否发生 AKI 分为 AKI 组与非 AKI 组。

1.1.1 纳入标准: 年龄 ≥ 18 岁; 在 ICU 住院期间接受过心脏外科手术。

1.1.2 排除标准: 在接受心脏外科手术之前出现过 AKI; 临床资料缺失。

1.2 数据提取与处理: 从数据库中获取患者的基本信息, 包括年龄、性别、入院类型、肥胖、是否合并糖尿病; 生命体征, 包括呼吸频率(respiratory rate, RR)、心率(heart rate, HR)、血压、体温; 实验室指标, 包括血红蛋白(hemoglobin, Hb)、白蛋白(albumin, ALB)、血糖、血 K^+ 、血 Na^+ 、动脉血氧分压(arterial partial pressure of oxygen, PaO_2)、pH 值、阴离子隙(anion gap, AG)、SCr、血尿素氮(blood urea nitrogen, BUN)、血乳酸(lactic acid, Lac)、三酰甘油(triacylglycerol, TG); 临床特征及器官支持, 包括机械通气、多器官功能障碍综合征(multiple organ dysfunction syndrome, MODS)、脓毒症、是否应用血管活性药物。所有变量的缺失值通过基于 R 语言 mice package 中的随机森林方法补全。

AKI 的诊断标准参照改善全球肾脏病预后组织(Kidney Disease: Improving Global Outcomes, KDIGO)制定的针对 AKI 的临床指南^[11]。脓毒症的诊断标准采用 2016 年发布的 Sepsis-3^[12]。

1.3 统计学分析: 应用 Python 3.7.9 分析软件进行统计学分析。连续变量均不符合正态分布, 以中位数(四分位数)[$M(Q_L, Q_U)$]表示, 采用 Kruskal-Wallis 秩和检验; 分类变量以频数(百分比)表示, 用 χ^2 检验。 $P < 0.05$ 为差异有统计学意义。

1.4 构建 Logistic 回归预测模型: 在排除掉方差膨胀系数(variance inflation factor, VIF) > 10 的高共线性自变量后, 使用基于 Python 语言的 Scikit-learn^[13] 中“Logistic Regression”模块建立 Logistic 回归预测模型, 并绘制受试者工作特征曲线(receiver operator characteristic curve, ROC 曲线), 计算模型的 ROC 曲

线下面积(area under the ROC curve, AUC), 以及各项变量的偏回归系数和优势比(odds ratio, OR)。

1.5 构建机器学习预测模型: XGBoost 是一种基于决策树中梯度增强算法的集成机器学习算法^[14], 是目前最流行的机器学习方法之一。本研究中使用基于 Python 语言的“xgboost package”。模型训练过程中的超参数通过 Scikit-learn 中的网格搜索与五折交叉验证结合的方法来筛查所有参数的组合, 从而确定最优参数, 即将数据分为 5 份, 每次取出 4/5 的数据作为训练集、1/5 的数据作为测试集, 重复该过程 5 次, 比较所有可能的超参数的 5 次 AUC 平均值, 从而确定最优超参数。根据输出模型计算模型的学习曲线及其 AUC。采用沙普利加和解释法(Shapley additive explanations, SHAP)解释模型, 包括各项预测变量的重要性及其对最终预测结果的影响。

2 结果

2.1 一般情况: MIMIC-III 数据库中接受心脏外科手术患者共 7 911 例, 排除术前发生 AKI 者 917 例、年龄 < 18 岁者 82 例, 最终有 6 912 例纳入分析(图 1)。根据 KDIGO 诊断标准, 6 912 例患者中有 5 681 例(82.2%)术后 14 d 内发生 AKI, 其中 AKI I 期 5 291 例(93.1%), AKI II 期 357 例(6.3%), AKI III 期 33 例(0.6%); 1 231 例(17.8%)未发生 AKI。

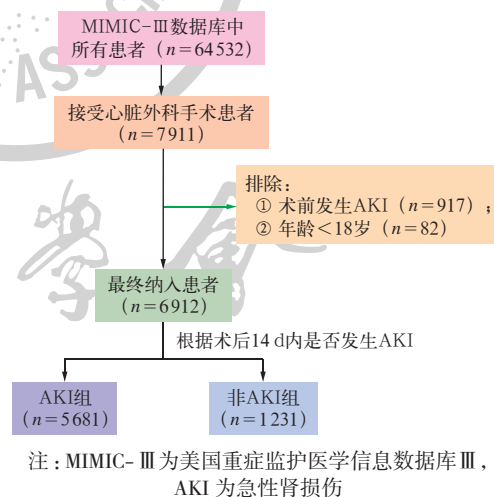


图 1 基于 MIMIC-III 数据库的机器学习和 Logistic 回归模型预测心脏外科术后患者发生 AKI 的比较分析研究对象纳入流程

表 1 显示, 与非 AKI 组相比, AKI 组患者的年龄更大, 急诊入院及合并肥胖、糖尿病的比例更高; 生命体征指标中 RR 更低, HR 更慢, 血压更高; 实验室指标中 Hb、血糖、血 K^+ 、SCr 更高, ALB 和 TG 更低; 且 MODS 和脓毒症的比例更高, 差异均有统计学意义(均 $P < 0.05$)。

表 1 术后 14 d 内是否发生 AKI 两组心脏外科术后患者临床特征比较

变量	全体 (n=6 912)	非 AKI 组 (n=1 231)	AKI 组 (n=5 681)	Z/ χ^2 值	P 值
一般资料					
年龄〔岁, $M(Q_L, Q_U)$ 〕	67.0(58.0, 75.0)	62.0(52.0, 71.0)	68.0(59.0, 76.0)	166.097	<0.001
男性〔例(%)〕	4 679(67.7)	855(69.5)	3 824(67.3)	2.126	0.154
急诊入院〔例(%)〕	3 563(51.5)	588(47.8)	2 975(52.4)	8.576	0.004
肥胖〔例(%)〕	561(8.1)	49(4.0)	512(9.0)	34.902	<0.001
糖尿病〔例(%)〕	2 093(30.3)	273(22.2)	1 820(32.0)	46.581	<0.001
生命体征〔 $M(Q_L, Q_U)$ 〕					
RR(次/min)	17.0(14.0, 21.0)	19.0(15.0, 22.0)	17.0(14.0, 20.0)	72.848	<0.001
HR(次/min)	80.0(68.0, 90.0)	82.0(71.5, 93.0)	80.0(67.0, 89.0)	49.605	<0.001
血压(mmHg)	80.0(70.7, 89.0)	78.0(70.0, 88.0)	80.0(70.7, 90.0)	13.932	<0.001
体温(℃)	36.8(36.5, 37.2)	36.8(36.5, 37.2)	36.8(36.5, 37.2)	3.625	0.057
实验室指标〔 $M(Q_L, Q_U)$ 〕					
Hb(g/L)	122.0(108.0, 136.0)	120.0(106.0, 135.0)	122.0(109.0, 136.0)	8.795	0.003
ALB(g/L)	39.0(36.0, 41.0)	39.0(37.0, 42.0)	38.0(35.0, 41.0)	57.060	<0.001
血糖(mmol/L)	7.2(5.9, 8.8)	6.8(5.7, 8.5)	7.3(6.1, 8.9)	39.767	<0.001
血 K ⁺ (mmol/L)	4.2(3.9, 4.7)	4.2(3.8, 4.6)	4.2(3.9, 4.7)	14.353	<0.001
血 Na ⁺ (mmol/L)	138.0(136.0, 140.0)	138.0(136.0, 140.0)	138.0(136.0, 140.0)	1.956	0.162
PaO ₂ (mmHg)	330.0(235.0, 398.0)	327.0(230.5, 396.0)	331.0(236.0, 398.0)	2.570	0.109
pH 值	7.4(7.4, 7.5)	7.4(7.4, 7.4)	7.4(7.4, 7.5)	0.001	0.977
AG(mmol/L)	13.0(12.0, 15.0)	13.0(12.0, 15.0)	13.0(12.0, 15.0)	0.152	0.697
SCr(μ mol/L)	88.4(70.7, 106.1)	79.6(70.7, 97.2)	88.4(70.7, 106.1)	25.248	<0.001
BUN(mmol/L)	6.4(5.0, 8.5)	6.1(5.0, 8.2)	6.4(5.0, 8.5)	2.774	0.096
Lac(mmol/L)	1.5(1.1, 2.2)	1.5(1.1, 2.2)	1.5(1.1, 2.1)	1.066	0.302
TG(mmol/L)	1.4(1.0, 2.1)	1.5(1.0, 2.2)	1.4(1.0, 2.0)	14.554	<0.001
机械通气〔例(%)〕	342(4.9)	55(4.5)	287(5.1)	0.734	0.433
MODS〔例(%)〕	1 938(28.0)	200(16.2)	1 738(30.6)	103.191	<0.001
脓毒症〔例(%)〕	211(3.1)	23(1.9)	188(3.3)	7.097	0.010
血管活性药物〔例(%)〕	269(3.9)	50(4.1)	219(3.9)	0.116	0.796

注:AKI 为急性肾损伤,RR 为呼吸频率,HR 为心率,Hb 为血红蛋白,ALB 为白蛋白,PaO₂ 为动脉血氧分压,AG 为阴离子隙,SCr 为血肌酐,BUN 为血尿素氮,Lac 为血乳酸,TG 为三酰甘油,MODS 为多器官功能障碍综合征;1 mmHg $\approx$ 0.133 kPa

因两组之间例数差距较大,组间严重不平衡,故采用基于 Python 语言“imbalanced-learn package”中的 AllKNN 方法^[15]进行欠采样。平衡后最终共纳入 4 031 例患者用于构建模型,其中非 AKI 组 1 321 例,AKI 组 2 710 例。

2.2 Logistic 回归预测模型的构建:经过计算,年龄、ALB、AG、血压、血糖、RR、pH 值、血 K⁺、血 Na⁺、HR、体温的 VIF 均>10,提示存在变量间多重共线性,故不纳入模型。Logistic 回归预测模型中主要影响因素包括 Lac、肥胖、男性、伴有糖尿病和急诊入院(表 2)。经计算,Logistic 回归预测模型预测心脏外科术后 AKI 的 AUC 为 0.62〔95% 可信区间(95% confidence interval, 95%CI)为 0.61~0.67;图 2〕。

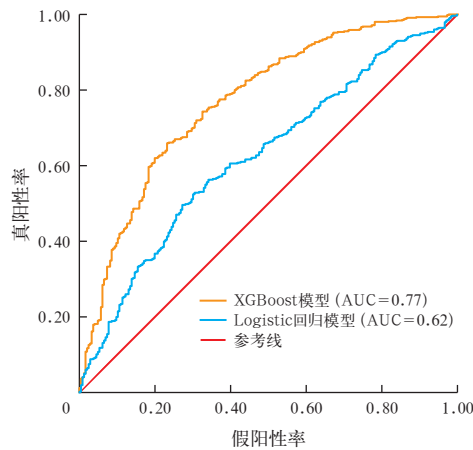
2.3 机器学习模型的构建:XGBoost 的参数在经过网格搜索与五折交叉验证结合优化参数后,设置参数为 eta=0.19, max_depth=3, min_child_weight=3,其余参数保持默认值。在训练过程中,随着从全部病例中抽取样本数增大,训练集的 AUC 逐步下降,

表 2 心脏外科手术患者术后 14 d 内发生 AKI 的 Logistic 回归预测模型分析

因素	偏回归系数	OR 值	95%CI	P 值
SCr	0.081	1.084	1.030~1.140	0.132
Lac	0.060	1.062	1.030~1.100	0.050
机械通气	-0.046	0.956	0.797~1.150	0.802
肥胖	0.806	2.234	1.900~2.640	<0.001
脓毒症	0.352	1.422	1.110~1.830	0.159
TG	-0.000	1.000	0.990~1.000	0.525
血管活性药物	-0.262	0.770	0.633~0.934	0.177
男性	-0.153	0.858	0.794~0.928	0.049
糖尿病	0.599	1.820	1.680~1.980	<0.001
急诊入院	0.245	1.278	1.190~1.380	<0.001
截距	0.230			

注:AKI 为急性肾损伤,SCr 为血肌酐,Lac 为血乳酸,TG 为三酰甘油,OR 为优势比,95%CI 为 95% 可信区间;空白代表无此项

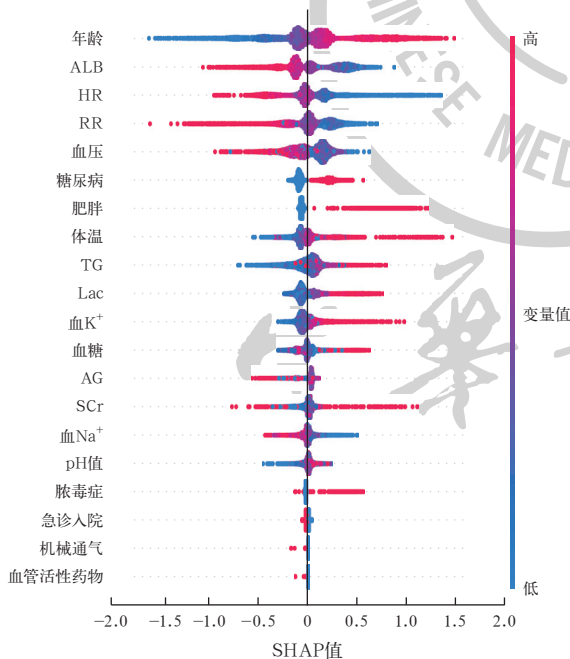
而测试集的 AUC 则逐步上升,提示模型训练效果良好,没有出现拟合或欠拟合。经计算,机器学习模型预测心脏外科术后 AKI 的 AUC 为 0.77(95%CI 为 0.75~0.80;图 2),明显高于 Logistic 回归预测模型的 AUC ($P<0.01$)。



注：XGBoost 为极限梯度提升算法，AKI 为急性肾损伤，ROC 曲线为受试者工作特征曲线，AUC 为 ROC 曲线下面积

图 2 XGBoost 与 Logistic 回归模型预测心脏外科手术后 14 d 内发生 AKI 的 ROC 曲线

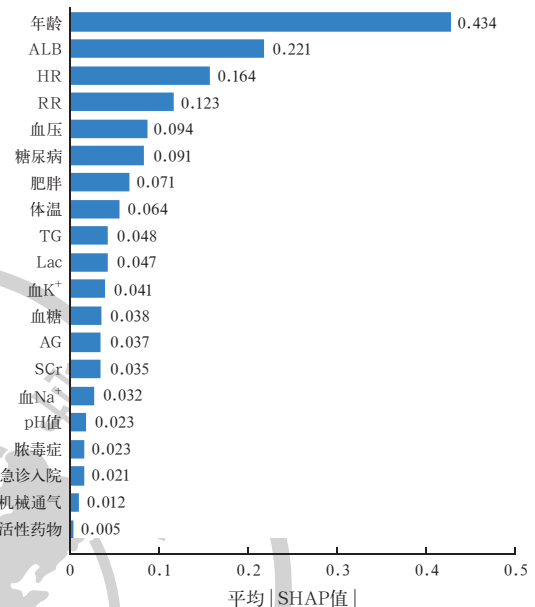
2.4 机器学习模型的解释：SHAP 是一种解释机器学习模型输出的博弈论方法^[16]。在本研究的输出模型中，每位患者的每个预测变量均可转化为对最后结果的贡献，即 SHAP 值，SHAP 值越大，说明患者在心脏外科术后发生 AKI 的风险就越大。结果显示，随着患者年龄增大、ALB 降低，心脏外科术后 AKI 的发生风险逐渐升高（图 3）。



注：AKI 为急性肾损伤，XGBoost 为极限梯度提升算法，SHAP 为沙普利加和解释法，ALB 为白蛋白，HR 为心率，RR 为呼吸频率，TG 为三酰甘油，Lac 为血乳酸，AG 为阴离子隙，SCr 为血肌酐；图中每个点代表一位患者的一个变量值，颜色代表该变量值在同一变量内的相对高低；在输出模型中，每位患者的每个预测变量均可转化为对最后结果的贡献，即 SHAP 值，SHAP 值越大，说明患者在心脏外科术后发生 AKI 的风险就越大

图 3 心脏外科手术后 14 d 内发生 AKI 的 XGBoost 预测模型中所有变量的 SHAP 分析

通过 SHAP 值的绝对值(|SHAP 值|)来衡量各项变量在最终结果中的重要性，结果显示，输出模型中最重要的预测因素是年龄和 ALB，其中年龄是危险因素，ALB 是保护因素（图 4）。



注：AKI 为急性肾损伤，XGBoost 为极限梯度提升算法，SHAP 为沙普利加和解释法，ALB 为白蛋白，HR 为心率，RR 为呼吸频率，TG 为三酰甘油，Lac 为血乳酸，AG 为阴离子隙，SCr 为血肌酐；在输出模型中，每位患者的每个预测变量均可转化为对最后结果的贡献，即 SHAP 值，以 SHAP 值的绝对值(|SHAP 值|)来衡量各项变量在最终结果中的重要性，|SHAP 值|越大，说明该变量对于预测结果就越重要

图 4 心脏外科手术后 14 d 内发生 AKI 的 XGBoost 预测模型中所有变量的重要性分析

3 讨论

本研究回顾分析了 MIMIC-III 数据库中所有接受心脏外科手术患者的临床资料，对术后 AKI 的发生进行了预测研究，比较了机器学习与传统 Logistic 回归的预测效能。预测模型及危险因素的研究对于接受心脏外科手术患者的风险评估和治疗策略的选择具有重要意义。本研究通过 XGBoost 模型发现，年龄和 ALB 对于心脏术后发生 AKI 有重要预测价值。

已有研究表明，年龄、SCr、ALB 等对于 AKI 的发生具有重要预测价值^[17-18]。在本研究 XGBoost 模型中，年龄和 ALB 是最重要的预测因素，而 SCr 则没有如预期对结局产生重要影响。一方面，这可能与本研究收集的是患者的术前资料，而其他研究多为术后 SCr 有关。另一方面，利用 SHAP 方法分析后，SCr 对预后的影响可能是存在的。随着 SCr 升高其 SHAP 值逐渐增加，直至在大约 176.8 $\mu\text{mol/L}$ 时达到最高值。但在整体水平上，SCr 对结果影响并不大，其 |SHAP 值| 小于年龄、ALB 等其他因素。

本研究表明,机器学习比传统 Logistic 回归模型具有更好的预测效能。一方面,传统的线性模型可能无法很好地解释某些变量与结局之间复杂的非线性关系。另一方面,机器学习中常用的 XGBoost 算法可以在迭代的过程中聚焦于预测错误的结果,这意味着当迭代次数足够时模型的表现会大大提升。其他研究结果也支持机器学习技术在预测方面比传统的回归模型更好^[19-20]。

XGBoost 是一种在医学研究中尚未得到广泛应用的新技术。它已经成功地应用于解决一些复杂的问题,如预测 AKI 患者的容量反应性^[21]。XGBoost 模型的效能明显优于传统的线性模型也在其他研究中得到了验证^[22]。同时, XGBoost 技术可通过交叉验证和正则化来防止过拟合或欠拟合,还可以自动处理缺失值,这对于临床实践中随时进行预测具有重要意义。应用 XGBoost 模型可以在术前任意时刻给出对术后 AKI 发生风险的相对准确的预测。此外,我们可以利用 SHAP 方法直观地了解模型的工作原理及如何简化模型以便于临床工作人员在术前对患者术后 AKI 发生风险作出大致判断。

本研究存在一定局限性:首先,本研究是一项回顾性研究,需要进一步开展前瞻性研究来验证结论。其次,本研究中总体人群存在年龄偏大、术后 AKI 发生率过高等问题,即便通过欠采样、标准化等方式来减小其影响,也仍然可能会影响预测模型在其他人群中的应用。

综上所述,对于预测心脏外科术后患者是否发生 AKI, XGBoost 是一种较好的方法,根据其输出结果,最重要的危险因素是年龄,最重要的保护因素则是 ALB。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] Singbartl K, Kellum JA. AKI in the ICU: definition, epidemiology, risk stratification, and outcomes [J]. *Kidney Int*, 2012, 81 (9): 819-825. DOI: 10.1038/ki.2011.339.
- [2] Hoste EA, De Corte W. Epidemiology of AKI in the ICU [J]. *Acta Clin Belg*, 2007, 62 Suppl 2: 314-317. DOI: 10.1179/acb.2007.071.
- [3] Thiele RH, Isbell JM, Rosner MH. AKI associated with cardiac surgery [J]. *Clin J Am Soc Nephrol*, 2015, 10 (3): 500-514. DOI: 10.2215/CJN.07830814.
- [4] Wang Y, Bellomo R. Cardiac surgery-associated acute kidney injury: risk factors, pathophysiology and treatment [J]. *Nat Rev Nephrol*, 2017, 13 (11): 697-711. DOI: 10.1038/nrneph.2017.119.
- [5] Nadim MK, Forni LG, Bihorac A, et al. Cardiac and vascular surgery-associated acute kidney injury: the 20th International Consensus Conference of the ADQI (Acute Disease Quality Initiative) Group [J]. *J Am Heart Assoc*, 2018, 7 (11): e008834. DOI: 10.1161/JAHA.118.008834.
- [6] 李贵森, 吴昌为. 人工智能在急性肾损伤诊断中的价值和应用 [J]. *中国血液净化*, 2019, 18 (11): 769-771. DOI: 10.3969/j.issn.1671-4091.2019.11.010.
- [7] 凌国兴, 罗程, 黎玉贵, 等. 心脏手术后急性肾损伤的研究进展 [J]. *医学综述*, 2018, 24 (6): 1103-1108. DOI: 10.3969/j.issn.1006-2084.2018.06.013.
- [8] Park S, Lee H. Acute kidney injury prediction models: current concepts and future strategies [J]. *Curr Opin Nephrol Hypertens*, 2019, 28 (6): 552-559. DOI: 10.1097/MNH.0000000000000536.
- [9] Deo RC. Machine learning in medicine [J]. *Circulation*, 2015, 132 (20): 1920-1930. DOI: 10.1161/CIRCULATIONAHA.115.001593.
- [10] Johnson AE, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database [J]. *Sci Data*, 2016, 3: 160035. DOI: 10.1038/sdata.2016.35.
- [11] Khwaja A. KDIGO clinical practice guidelines for acute kidney injury [J]. *Nephron Clin Pract*, 2012, 120 (4): e179-184. DOI: 10.1159/000339789.
- [12] Rhodes A, Evans LE, Alhazzani W, et al. Surviving Sepsis Campaign: international guidelines for management of sepsis and septic shock: 2016 [J]. *Intensive Care Med*, 2017, 43 (3): 304-377. DOI: 10.1007/s00134-017-4683-6.
- [13] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python [J]. *J Mach Learn Res*, 2011, 12 (85): 2825-2830.
- [14] Chen TQ, Guestrin C. XGBoost: a scalable tree boosting system [C]// Krishnapuram B, Shah M. KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, 2016. New York: Association for Computing Machinery, 2016: 785-794. DOI: 10.1145/2939672.2939785.
- [15] Tomek I. An experiment with the edited nearest-neighbor rule [J]. *IEEE Trans Syst Man Cybern B Cybern*, 1976, 6 (6): 448-452. DOI: 10.1109/TSMC.1976.4309523.
- [16] Lundberg SM, Lee SI. A unified approach to interpreting model predictions [C]// von Luxburg U, Guyon I, Bengio S, et al. NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, 2017. New York: Curran Associates Inc., 2017: 4765-4774. DOI: 10.48550/arXiv.1705.07874.
- [17] Kate RJ, Pearce N, Mazumdar D, et al. A continual prediction model for inpatient acute kidney injury [J]. *Comput Biol Med*, 2020, 116: 103580. DOI: 10.1016/j.compbiomed.2019.103580.
- [18] Merchant ML, Brier ME, Slaughter MS, et al. Biomarker enhanced risk prediction for development of AKI after cardiac surgery [J]. *BMC Nephrol*, 2018, 19 (1): 102. DOI: 10.1186/s12882-018-0902-9.
- [19] 张素珍, 唐素娟, 戎珊, 等. 基于机器学习的重症监护病房脓毒性休克患者早期发生急性肾损伤风险的预测模型构建 [J]. *中华危重病急救医学*, 2022, 34 (3): 255-259. DOI: 10.3760/cma.j.cn121430-20211126-01790.
- [20] 彭川, 熊辉. 机器学习在急诊科诊疗中的应用研究进展 [J]. *中国中西医结合急救杂志*, 2022, 29 (2): 253-256. DOI: 10.3969/j.issn.1008-9691.2022.02.028.
- [21] Zhang ZH, Ho KM, Hong YC. Machine learning for the prediction of volume responsiveness in patients with oliguric acute kidney injury in critical care [J]. *Crit Care*, 2019, 23 (1): 112. DOI: 10.1186/s13054-019-2411-z.
- [22] 魏子健, 李静, 李雪岩, 等. 基于机器学习的院内消化道致命性再出血预测与指标筛选 [J]. *中华危重病急救医学*, 2019, 31 (3): 359-362. DOI: 10.3760/cma.j.issn.2095-4352.2019.03.020.

(收稿日期: 2021-02-23)

关于经过广告审批后的广告中存在不规范医学名词术语未予更改的声明

依照广告审批的相关规定,按照广告厂家的要求,本刊刊登的新活素、血必净及佳维体广告图片和内容均按照广告审查批准文件的原件刊出,故广告内容中“适应症”“禁忌症”未按标准医学名词术语修改为“适应证”“禁忌证”,“其它”未修改为“其他”,“成份”未修改为“成分”,时间单位仍用汉字表示,剂量单位“ml”未修改为“mL”,标示数值范围的标点符号“-”未修改为“~”。特此声明!