

# 基于机器学习的心源性休克患者院内死亡的预测研究

周 弛, 廉哲勋\*

青岛大学附属医院心血管内科, 山东 青岛

收稿日期: 2022年10月9日; 录用日期: 2022年11月2日; 发布日期: 2022年11月14日

## 摘 要

目的: 构建心源性休克患者院内死亡预测模型。方法: 利用MIMIC-III数据库, 收集人口特征、实验室检查、合并症等87个指标并进行特征选择后, 使用随机森林、Logistic回归、XGBoost、卷积神经网络算法, 构建预测模型。用敏感性、特异性、曲线下面积、准确性来比较这4种模型的性能。结果: 在这项研究的804名患者中, 有209名患者(26%)出现院内死亡。在预测院内死亡时, 四个模型的接收器工作特征曲线(ROC)的曲线下面积(AUC)在0.757至0.848的范围内。在所有模型中, XGBoost的灵敏度最高(87.3%), 特异性(81%)和准确性最高(84.6%)。结论: 机器学习算法可以准确预测心源性休克患者院内死亡率, 尤其是XGBoost模型。

## 关键词

机器学习, 预测模型, 心源性休克

# Research on Prediction Model of Cardiogenic Shock in-Hospital Death Based on Machine Learning

Chi Zhou, Zhexun Lian\*

Department of Cardiology, The Affiliated Hospital of Qingdao University, Qingdao Shandong

Received: Oct. 9<sup>th</sup>, 2022; accepted: Nov. 2<sup>nd</sup>, 2022; published: Nov. 14<sup>th</sup>, 2022

\*通讯作者 Email: [lianzx566@163.com](mailto:lianzx566@163.com)

## Abstract

**Objective:** To construct a prediction model for in-hospital death in patients with cardiogenic shock. **Methods:** Using the MIMIC-III database, 87 indicators of demographic characteristics, laboratory tests, and comorbidities were collected and then feature selection. Logistic regression, random forest, XGBoost, and convolutional neural network algorithms were employed to build models. Sensitivity, specificity, accuracy, and area under the curve were applied to assess the performance. **Result:** Among 804 patients enrolled in this study, 209 patients (26%) died in hospital. In the prediction of the in-hospital death, the areas under the curve (AUCs) of the receiver operating characteristic curves (ROCs) of the four models ranged from 0.757 to 0.848. Among all models, XGBoost achieved the highest sensitivity (87.3%), specificity (81%) and accuracy (84.6%). **Conclusion:** Machine learning algorithms can accurately predict the in-hospital mortality of patients with cardiogenic shock, especially the XGBoost model.

## Keywords

Machine Learning, Prediction Model, Cardiogenic Shock

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 简介

心源性休克是由于原发性心功能不全而导致的各个器官低灌注的危重状态。引起心源性休克的原因很多,最主要的原因是急性心肌梗死,还包括心律失常、主动脉破裂和心包填塞[1]。虽然心源性休克的发病率低于其他致命疾病,但其住院死亡率高达约 60.1% [2]。尽管在早期预防、及时诊断和积极治疗以挽救心源性休克患者方面取得了进展,但心源性休克仍然是死亡的主要原因之一[3]。

在心源性休克患者管理中,对其进行风险评估是非常重要的,全面评估患者情况,减少院内死亡的发生对于这种复杂患者至关重要。机器学习是一门多领域交叉学科,涉及概率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科,已经在生物信息学、生态学、医学、遗传学、遥感地理学等多领域开展的应用性研究[4],对于数据的审查分析更加科学,能考虑到内在相关性,本研究利用公开数据库,通过机器学习方法,分析影响心源性休克患者的院内死亡相关危险因素。

## 2. 研究对象

重症监护医学信息市场(MIMIC-III)是可公开获得的重症监护数据库[5]。MIMIC-III 数据库的建立得到麻省理工学院和以色列女执事医疗中心的批准。该数据库包含生命体征,药物,实验室测量值以及其他成年患者(18岁或18岁以上)入院的其他重要项目。本研究为回顾性研究,作者获得美国国立卫生研究院保护人类研究参与者考试的批准(ID: 35892903),并且获取访问 MIMIC-III 数据库的权限。使用 PgAdmin 并通过结构化的查询语言提取数据库中资料[6]。

### 2.1. 纳入标准

- 1) 年龄位于 18 到 89 岁之间;
- 2) 基线数据缺失率 < 10%。同时,如果数据库期间有超过一项入院

记录, 只会选择第一次入院。根据国际疾病分类第 9 版临床修改诊断代码(ICD-9:785.51)确定诊断为心源性休克的患者。

## 2.2. 研究指标

包括年龄、性别、既往史、实验室指标等。所有指标均为进入 ICU 的 24 内的数值, 包括同一指标的最大值、最小值。删除某一项指标缺失值大于 30%, 最终共纳入 87 项指标进行研究。

## 2.3. 研究方法

所有数据分析及模型建立使用 R (4.0.3)和 Python (3.7.0)进行处理, 对于缺失值使用多重插补法进行填充。符合正态分布的计量资料以均数  $\pm$  标准差( $\bar{x} \pm s$ )表示, 组间比较采用两独立样本 t 检验; 计数资料比较采用卡方检验或 Fisher 确切概率法; 将院内存活和院内死亡组按照 7:3 随机分为训练集和验证集。采用递归特征消除法进行特征筛选, 即根据底层模型的特点, 选出最重要的特征, 把选出来的特征放到一边, 然后在剩余的特征上重复这个过程, 直到遍历了所有的特征。在这个过程中被消除的次序就是特征的排序。将筛选后的特征应用到模型中, 以是否发生院内死亡为结局变量。在训练集中分别建立 Logistic 回归、随机森林、XGBoost 和卷积神经网络模型[7]; 在构建模型过程中, 通过不断调整各项参数, 使每个模型均达到最佳的表现, 用验证集来验证模型的表现。模型的评估采用敏感度、特异度、准确率、曲线下面积来进行衡量。特征选择及模型构建的过程在图 1。使用 Delong test 法计算两个 ROC 曲线之间是否存在的统计意义。使用沙普利值(SHAP value)对最优模型进行可视化解释, 寻找模型中各个变量的贡献程度[8]。

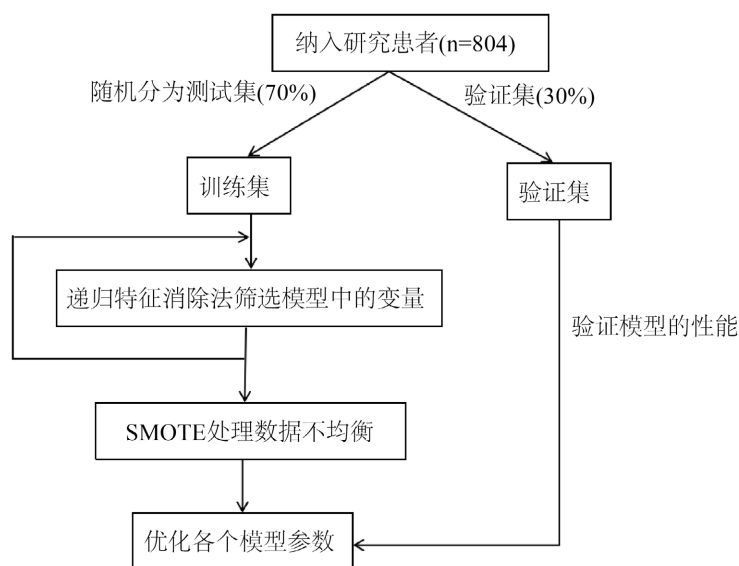


Figure 1. Model building flow chart

图 1. 模型建立流程图

## 3. 结果

### 3.1. 基线资料

在 804 例心源性休克患者中, 563 例(70%)随机进入训练数据, 241 例(30%)进入验证数据。表 1 显示是否发生院内死亡之间的基线变量的区别, 其中两组在许多因素之间具有差异有统计学意义( $P < 0.05$ )。

**Table 1.** Comparison of baseline information of patients with and without in-hospital mortality  
**表 1.** 院内是否发生死亡患者基线资料比较

	非院内死亡	院内死亡	概率
数量	595	209	
年龄	68.10 ± 12.46	71.94 ± 12.72	<0.001
女性(%)	230 (38.7)	77 (36.8)	0.703
阴离子间隙_最小值	13.79 ± 2.85	15.80 ± 4.42	<0.001
阴离子间隙_最大值	17.69 ± 4.26	20.10 ± 5.28	<0.001
碳酸氢钠_最小值	20.55 ± 4.23	18.68 ± 4.99	<0.001
碳酸氢钠_最大值	24.21 ± 3.79	22.79 ± 4.42	<0.001
胆红素_最小值	0.87 ± 0.79	1.25 ± 1.54	<0.001
胆红素_最大值	1.06 ± 1.30	1.31 ± 1.41	0.016
肌酐_最小值	1.45 ± 1.46	1.78 ± 1.32	0.004
肌酐_最大值	1.82 ± 1.71	2.16 ± 1.48	0.01
血氯_最小值	101.93 ± 5.61	100.55 ± 6.80	0.004
血氯_最大值	106.76 ± 5.95	106.48 ± 6.36	0.571
血细胞比容_最小值	29.93 ± 6.52	29.75 ± 6.41	0.72
血细胞比容_最大值	37.65 ± 5.66	37.19 ± 5.79	0.317
血红蛋白_最小值	10.20 ± 2.27	10.01 ± 2.18	0.289
血红蛋白_最大值	12.60 ± 1.94	12.35 ± 2.02	0.116
乳酸_最小值	1.79 ± 0.97	2.67 ± 2.21	<0.001
乳酸_最大值	3.60 ± 2.67	5.23 ± 4.16	<0.001
血小板_最小值	195.97 ± 85.37	188.51 ± 100.06	0.299
血小板_最大值	256.92 ± 97.30	250.45 ± 109.44	0.424
血钾_最小值	3.68 ± 0.54	3.76 ± 0.62	0.089
血钾_最大值	4.93 ± 0.95	5.06 ± 0.94	0.072
凝血激活酶时间_最小值	36.35 ± 16.86	41.82 ± 22.41	<0.001
凝血激活酶时间_最大值	69.85 ± 42.39	79.20 ± 45.50	0.007
国际标准化比值_最小值	1.41 ± 0.64	1.61 ± 0.74	<0.001
国际标准化比值_最大值	1.90 ± 1.57	2.18 ± 1.60	0.025
凝血酶原时间_最小值	14.99 ± 3.69	16.55 ± 4.61	<0.001
凝血酶原时间_最大值	17.96 ± 8.31	20.85 ± 14.49	<0.001
白细胞_最小值	11.28 ± 4.59	12.68 ± 5.90	<0.001
白细胞_最大值	15.50 ± 6.23	17.32 ± 7.42	0.001
磷酸激酶_最小值	645.25 ± 1174.40	761.40 ± 1415.30	0.245
磷酸激酶_最大值	1376.23 ± 2048.26	1416.47 ± 2187.30	0.81
磷酸肌酸同工酶_最小值	47.62 ± 87.03	48.29 ± 83.56	0.923

## Continued

磷酸肌酸同工酶_最大值	107.14 ± 173.57	102.55 ± 161.51	0.738
谷丙转氨酶_最小值	142.63 ± 440.18	182.61 ± 423.71	0.254
谷丙转氨酶_最大值	166.99 ± 503.64	279.94 ± 723.25	0.014
尿量	2062.43 ± 1300.27	1521.44 ± 1308.78	<0.001
凝血病(%)	103 (17.3)	73 (34.9)	<0.001
慢性肺病(%)	94 (15.8)	39 (18.7)	0.395
充血性心力衰竭(%)	110 (18.5)	49 (23.4)	0.148
糖尿病(%)	45 (7.6)	15 (7.2)	0.976
缺铁性贫血(%)	102 (17.1)	39 (18.7)	0.696
瓣膜病(%)	38 (6.4)	18 (8.6)	0.353
肺循环疾病(%)	15 (2.5)	12 (5.7)	0.045
外周血管疾病(%)	72 (12.1)	25 (12.0)	1
高血压(%)	80 (13.4)	35 (16.7)	0.29
淋巴瘤(%)	5 (0.8)	2 (1.0)	1
肾衰竭(%)	109 (18.3)	47 (22.5)	0.227
肝病(%)	23 (3.9)	17 (8.1)	0.024
艾滋病(%)	0 (0.0)	3 (1.4)	0.023
酗酒(%)	22 (3.7)	8 (3.8)	1
吸毒(%)	11 (1.8)	3 (1.4)	0.932
连续肾替代治疗(%)	33 (5.5)	16 (7.7)	0.353
通气(%)	380 (63.9)	156 (74.6)	0.006
淋巴细胞_最小值	12.33 ± 9.59	10.08 ± 10.58	0.005
淋巴细胞_最大值	13.31 ± 10.53	21.84 ± 148.70	0.164
血钙_最小值	1.05 ± 0.12	1.02 ± 0.13	<0.001
血钙_最大值	1.18 ± 0.15	1.16 ± 0.18	0.329
红细胞分布宽度_最小值	14.49 ± 1.75	15.28 ± 2.12	<0.001
红细胞分布宽度_最大值	14.94 ± 1.95	15.72 ± 2.18	<0.001
主动脉内球囊反搏(%)	235 (39.5)	75 (35.9)	0.401
急性生理学评分 III	51.38 ± 19.09	67.10 ± 22.82	<0.001
最高收缩压	81.83 ± 15.69	79.60 ± 14.92	0.075
最低收缩压	42.07 ± 10.56	38.32 ± 10.01	<0.001
最高血糖	220.77 ± 97.25	244.41 ± 129.14	0.006
最低血糖	109.63 ± 35.92	112.32 ± 41.85	0.374
平均血糖	156.90 ± 44.16	169.89 ± 56.66	0.001
最慢心率	72.43 ± 15.51	75.69 ± 16.60	0.01
平均心率	87.42 ± 15.64	92.71 ± 16.28	<0.001

## Continued

最快心率	106.34 ± 19.36	116.18 ± 24.48	<0.001
最大呼吸次数	27.89 ± 6.67	29.61 ± 6.45	0.001
最小呼吸次数	12.76 ± 3.87	12.53 ± 4.05	0.468
平均呼吸次数	19.49 ± 3.78	20.59 ± 4.07	<0.001
最大氧合	99.68 ± 0.72	99.62 ± 1.33	0.404
最小氧合	89.98 ± 10.33	84.74 ± 15.92	<0.001
平均氧合	97.24 ± 2.28	96.15 ± 3.43	<0.001
最大收缩压	140.82 ± 21.90	134.33 ± 23.55	<0.001
平均收缩压	108.21 ± 14.06	101.13 ± 10.24	<0.001
最小收缩压	80.70 ± 14.85	73.34 ± 12.93	<0.001
最高体温	37.60 ± 0.80	37.55 ± 1.06	0.512
最低体温	35.95 ± 0.87	35.72 ± 1.00	0.002
平均体温	36.84 ± 0.69	36.75 ± 0.87	0.139
酸碱度	7.43 ± 0.07	7.42 ± 0.07	0.052
碱剩余	0.36 ± 3.85	-0.51 ± 4.62	0.008
二氧化碳	25.95 ± 4.49	25.02 ± 5.58	0.017

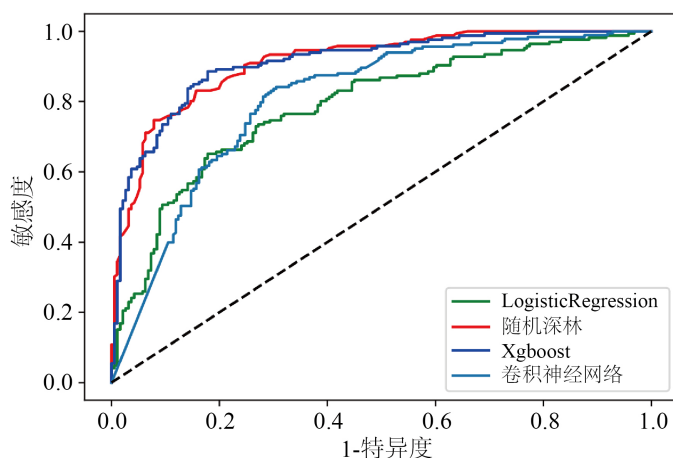
## 3.2. 构建模型的结果

构建的 Logistic 回归模型、随机森林模型、XGBoost 模型及卷积神经网络在验证集中进行内部验证, 结果显示, 各模型的 AUC 均较高, 依次为 0.786、0.826、0.848、0.757。卷积神经网络模型表现一般, 在特异性、敏感性、准确性依次为 0.687、0.852、0.751。XGBoost 模型的表现最优, 其敏感度为 0.81、特异度为 0.873、准确率为 0.846, 各个模型的表现情况如表 2 所示, ROC 曲线在图 2 中展示。经过 Delong test 检验, XGBoost 模型明显优于 Logistic 回归( $P < 0.05$ )。XGBoost 模型中各指标对应的贡献度如图 3 所示, 年龄、白细胞最小值、平均收缩压是最具预测性的特征, 其次是最高体温和红细胞分布宽度最小值。同时, 我们也对 XGBoost 模型中各个特征进行总体可视化。在图 4 中, 每一行中的蓝色和红色点代表了特定变量的低值到高值的参与者, 而 X 轴给出了 SHAP 值, 它给出了对模型的影响: 即它是否倾向于推动事件发生(SHAP 的正值)或事件未发生(SHAP 的负值)。例如: 图 4 中显示年龄是最重要的特征, 当蓝色点主要集中在 SHAP 值小于 0 的区域, 红色点主要集中在 SHAP 值大于 0 的区域, 可见年龄小会降低院内死亡发生发风险, 年龄越大会增加院内死亡发生的风险。

**Table 2.** Comparison of the prediction performance of the models in the test set

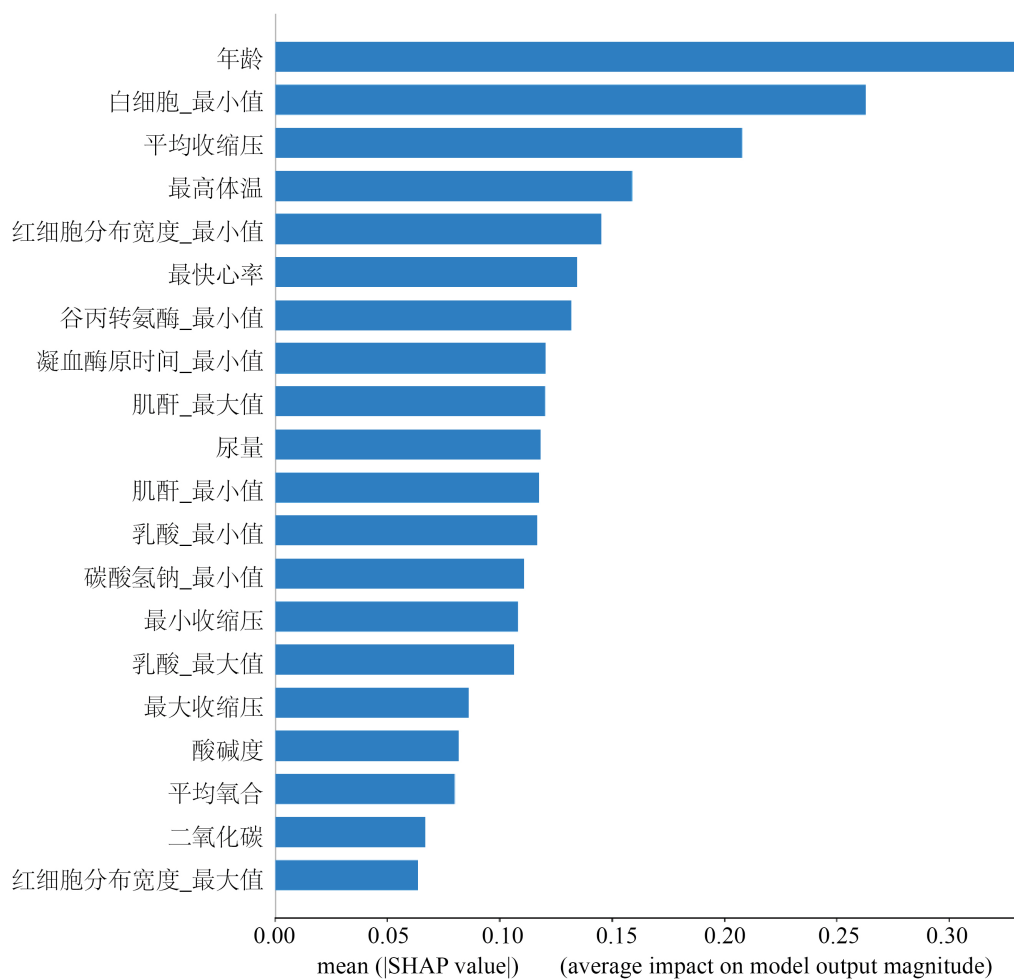
**表 2.** 各模型在测试集中的预测性能

模型	特异性	敏感性	准确率	AUC
Logistic 回归模型	0.699	0.686	0.717	0.786
随机森林模型	0.778	0.867	0.824	0.826
XGBoost 模型	0.810	0.873	0.846	0.848
卷积神经网络模型	0.687	0.852	0.751	0.757



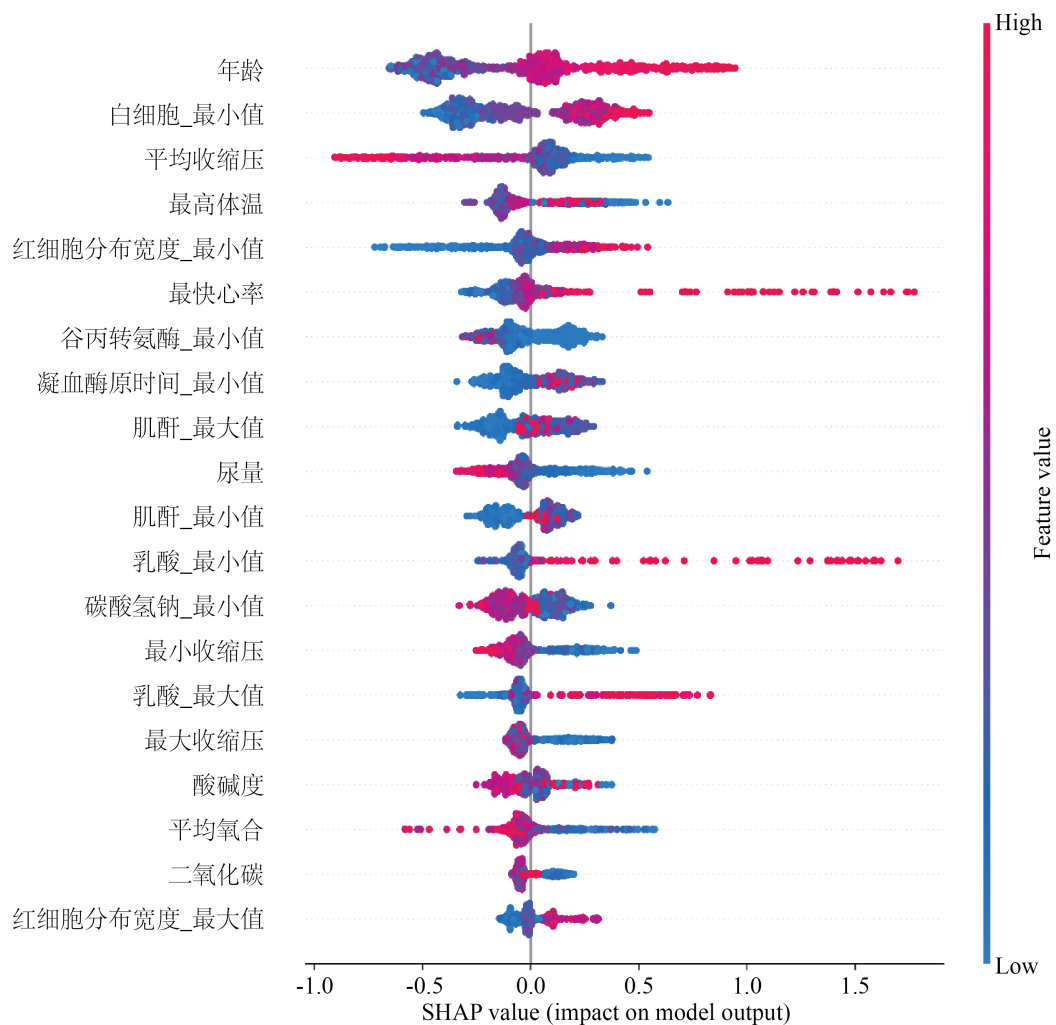
**Figure 2.** Area under the curve as a measure of predictive performance of whether in-hospital death occurs in patients with cardiogenic shock

**图 2.** 曲线下的面积作为衡量心源性休克患者是否发生院内死亡的预测性能



**Figure 3.** Feature importance graph for machine learning models

**图 3.** 机器学习模型的特征重要性



每一行中的蓝色和红色点代表了特定变量的低值到高值的参与者, 而x轴给出了SHAP值, 它给出了对模型的影响: 即它是否倾向于推动事件发生(SHAP的正值)或事件未发生(SHAP的负值)。

**Figure 4.** The first 20 clinical variables affecting the model in the Xgboost model

**图 4.** Xgboost模型中前 20 个影响模型临床变量

## 4. 讨论

心源性休克在是患者心脏输出功能衰竭的极期表现, 具有病死率高的特点, 因此, 在心源性休克早期评估中, 识别高危患者, 积极采取各种措施具有重要意义。既往关于心源性休克的风险研究局限于危险因素、预后探究[9], 仅少数基于危险因素的预测模型研究[3]。机器学习是指机器通过分析大量数据来进行学习, 其属于人工智能的一部分, 重视寻找数据中的关联。与传统的统计学不同的是, 机器学习更关注模型的预测效果, 即该事件是否发生。Shin 发现机器学习与传统统计模型在预测心衰患者再次入院和死亡的风险中具有更好的判别力[10]。D'Ascenzo 利用基于机器学习的方法来识别急性冠状动脉综合征后全因死亡, 心肌梗塞和大出血的发生率[11]。

在本研究中, 我们使用 4 种流行的机器学习算法构建心源性休克患者院内死亡模型, 经 Delong test 检验显示, 机器学习算法表现明显优于传统的 Logistic 回归模型, 其中综合预测效能最优的为 XGBoost 模型, AUC 为 0.848。



相对于深度学习中出现黑盒理论, 利用 SHAP 法能反映出每一个样本中的特征的影响力, 而且还表现影响的正负性。本研究中的 XGBoost 提示, 年龄、炎症指标、体温是模型中影响院内死亡的重要因素。年龄是判断许多疾病预后的经典指标。高龄患者预后在绝大多数疾病中相对较差, 年龄同时也是心源性休克患者预测预后的独立的危险因素。研究显示在使用 ECMO、机器循环通气患者中年龄与患者住院死亡率的增加呈线性相关[12]。炎症指标在心源性休克患者中扮演重要角色。休克患者常常出现严重的全身性炎症表现, 实验室检查中可发现典型炎症介质水平, 如 C 反应蛋白, 白介素-6 升高[13] [14]。外周灌注降低引起内皮细胞的损伤, 各种机制导致炎症因子的聚集, 随着缺血时间的延长, 最终引起各个器官功能障碍。红细胞分布宽度近几年也被认为是一个新的炎症指标来预测患者预后。越来越多的研究发现红细胞分布宽度在心脑血管疾病中有着良好的预测价值。Kim 研究指出在严重脓毒症或败血性休克患者住院期间红细胞分布宽度较基线明显增加与不良的临床预后相关[15]。对于心源性休克患者, 红细胞分布宽度的升高与患者全因死亡风险增加相关[16]。

本研究使用的 XGBoost 模型, 是于 2016 年发布的一种新算法, 但其在数据科学界获得了广泛的欢迎[17]。因为它是一种高效的、可扩展的分类和回归的模型, 可以处理稀疏, 复杂和非线性的数据类型, 这与临床上多变的数据相一致。

综上所述, 从模型的特异性、敏感性方面来看, 本研究搭建机器学习心源性休克院内死亡率预测模型是成功的, 其具有良好的稳定性, 可以提前预警高危心源性休克患者, 根据预测结果可以实行积极干预。模型中使用的数据均为临床常见的指标, 便于获得, 具有很强的可操作性和推广性。

## 5. 研究的局限性

本文利用了公开数据库进行分析, 数据详实, 但仍需要进一步外部验证模型的准确性, 但考虑到在研究开始随机分组, 分为训练集、验证集, 所获得的模型具有一定的可靠性, 但仍需要进一步的研究。

## 参考文献

- [1] Levy, B., Bastien, O., Karim, B., *et al.* (2015) Experts' Recommendations for the Management of Adult Patients with Cardiogenic Shock. *Annals of Intensive Care*, **5**, Article No. 52. <https://doi.org/10.1186/s13613-015-0052-1>
- [2] Zannad, F., Mebazaa, A., Juillière, Y., *et al.* (2006) Clinical Profile, Contemporary Management and One-Year Mortality in Patients with Severe Acute Heart Failure Syndromes: The EFICA Study. *European Journal of Heart Failure*, **8**, 697-705. <https://doi.org/10.1016/j.ejheart.2006.01.001>
- [3] Pöss, J., Köster, J., Fuernau, G., *et al.* (2017) Risk Stratification for Patients in Cardiogenic Shock after Acute Myocardial Infarction. *Journal of the American College of Cardiology*, **69**, 1913-1920. <https://doi.org/10.1016/j.jacc.2017.02.027>
- [4] 李昊朋. 基于机器学习方法的智能机器人探究[J]. 通讯世界, 2019, 26(4): 241-242.
- [5] Johnson, A.E., Pollard, T.J., Shen, L., *et al.* (2016) MIMIC-III, a Freely Accessible Critical Care Database. *Scientific Data*, **3**, Article ID: 160035. <https://doi.org/10.1038/sdata.2016.35>
- [6] Johnson, A.E., Stone, D.J., Celi, L.A. and Pollard, T.J. (2018) The MIMIC Code Repository: Enabling Reproducibility in Critical Care Research. *Journal of the American Medical Informatics Association*, **25**, 32-39. <https://doi.org/10.1093/jamia/ocx084>
- [7] 周昀锴. 机器学习及其相关算法简介[J]. 科技传播, 2019, 11(6): 153-154, 165.
- [8] Lundberg, S.M. and Lee, S.-I. (2017) A Unified Approach to Interpreting Model Predictions. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Curran Associates Inc., Long Beach, 4768-4777.
- [9] Abdin, A., Pöss, J., Fuernau, G., *et al.* (2018) Prognostic Impact of Baseline Glucose Levels in Acute Myocardial Infarction Complicated by Cardiogenic Shock—A Sub-Study of the IABP-SHOCK II-Trial [Corrected]. *Clinical Research in Cardiology*, **107**, 517-523. <https://doi.org/10.1007/s00392-018-1213-7>
- [10] Shin, S., Austin, P.C., Ross, H.J., *et al.* (2021) Machine Learning vs. Conventional Statistical Models for Predicting Heart Failure Readmission and Mortality. *ESC Heart Fail*, **8**, 106-115. <https://doi.org/10.1002/ehf2.13073>

- 
- [11] D'Ascenzo, F., De Filippo, O., Gallone, G., *et al.* (2021) Machine Learning-Based Prediction of Adverse Events Following an Acute Coronary Syndrome (PRAISE): A Modelling Study of Pooled Datasets. *The Lancet*, **397**, 199-207. [https://doi.org/10.1016/S0140-6736\(20\)32519-8](https://doi.org/10.1016/S0140-6736(20)32519-8)
- [12] Chung, M., Zhao, Y., Strom, J.B., Shen, C. and Yeh, R.W. (2019) Extracorporeal Membrane Oxygenation Use in Cardiogenic Shock: Impact of Age on In-Hospital Mortality, Length of Stay, and Costs. *Critical Care Medicine*, **47**, e214-e221. <https://doi.org/10.1097/CCM.0000000000003631>
- [13] Geppert, A., Dorninger, A., Delle-Karth, G., Zorn, G., Heinz, G. and Huber, K. (2006) Plasma Concentrations of Interleukin-6, Organ Failure, Vasopressor Support, and Successful Coronary Revascularization in Predicting 30-Day Mortality of Patients with Cardiogenic Shock Complicating Acute Myocardial Infarction. *Critical Care Medicine*, **34**, 2035-2042. <https://doi.org/10.1097/01.CCM.0000228919.33620.D9>
- [14] Thérout, P., Armstrong, P.W., Mahaffey, K.W., *et al.* (2005) Prognostic Significance of Blood Markers of Inflammation in Patients with ST-Segment Elevation Myocardial Infarction Undergoing Primary Angioplasty and Effects of Pexelizumab, a C5 Inhibitor: A Sub-Study of the COMMA Trial. *European Heart Journal*, **26**, 1964-1970. <https://doi.org/10.1093/eurheartj/ehi292>
- [15] Kim, C.H., Park, J.T., Kim, E.J., *et al.* (2013) An Increase in Red Blood Cell Distribution Width from Baseline Predicts Mortality in Patients with Severe Sepsis or Septic Shock. *Critical Care*, **17**, R282. <https://doi.org/10.1186/cc13145>
- [16] Wang, B., Aihemaiti, G., Cheng, B. and Li, X. (2019) Red Blood Cell Distribution Width Is Associated with All-Cause Mortality in Critically Ill Patients with Cardiogenic Shock. *Medical Science Monitor*, **25**, 7005-7015. <https://doi.org/10.12659/MSM.917436>
- [17] Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, San Francisco, 785-794. <https://doi.org/10.1145/2939672.2939785>