危重症急性肾损伤患者预后预测模型

蒋廷典¹,郭 皓^{2*},刘 阳²,李 贤³

¹承德医学院研究生院,河北 承德 ²河北北方学院研究生院,河北 张家口 ³邯郸市中心医院重症医学科,河北 邯郸

收稿日期: 2025年5月27日; 录用日期: 2025年6月19日; 发布日期: 2025年6月30日

摘要

目的:建立并验证基于机器学习方法的危重症急性肾损伤患者可解释死亡预测模型。方法:本研究提取了美国大型公开重症数据库 MIMIC-IV 的 2008 年至 2019 年、急性肾损伤(AKI)患者诊断 AKI 当天的临床数据,随机将数据分为训练队列和验证队列。提取了 MIMIC-III 数据库 3005 例 AKI 患者数据 作为外部验证集。通过多轮特征选择(低方差过滤、高相关性过滤、互信息筛选、SHAP 值分析和递归特征消除)选择出 12 个最佳特征组合,10 种机器学习方法被用来开发评估住院死亡率的模型。根据其 曲线下面积(AUC)选择最优模型。采用 SHapley Additive exPlanation (SHAP)值来解释最优模型。结果:本研究共计急性肾损伤 3701 例最终纳入患者(中位年龄,65 岁,女性 41.3%)。他们被随机分为 一个培训队列(2591 人,70%)和一个验证队列(1110 人,30%)。10 个机器学习模型中随机森林(RF) 模型具有最好的判别能力,并采用 SHAP 方法解释了模型。最终的模型在内部(AUC = 0.807)和外部 (AUC = 0.720)验证中都能较准确预测 AKI,将有助于临床医生判断重症监护病房 ICU 住院患者的预 后,并进行早期干预。

关键词

重症患者,急性肾损伤,机器学习,预后预测模型,SHAP

Predictive Model for Prognosis of Critically Ill Patients with Acute Kidney Injury

Tingdian Jiang¹, Hao Guo^{2*}, Yang Liu², Xian Li³

¹Graduate School of Chengde Medical University, Chengde Hebei ²Graduate School of Hebei North University, Zhangjiakou Hebei

*通讯作者。

³Department of Critical Care Medicine, Handan Central Hospital, Handan Hebei

Received: May 27th, 2025; accepted: Jun. 19th, 2025; published: Jun. 30th, 2025

Abstract

Objective The aim is to create and confirm a machine learning-based model for predicting mortality in critically ill patients with acute kidney injury (AKI). Methods: Patient clinical records for AKI as of its diagnosis day, spanning 2008 to 2019, were retrieved from the extensive MIMIC-IV public critical care database in the United States. The dataset was arbitrarily segmented into two groups: one for training and another for validation. Furthermore, the MIMIC-III database provided data from 3005 AKI patients, serving as an external validation dataset. A dozen ideal features were chosen after several stages of feature selection, including low variance filtering, high correlation filtering, mutual information screening, SHAP value analysis, and recursive feature elimination. A total of ten machine learning techniques were employed to create models for evaluating mortality rates within hospitals. The selection of the best model was guided by the area beneath the receiver operating characteristic curve (AUC). The optimal model was analyzed using SHAP (SHapley Additive ExPlanation) values. Results: The study eventually incorporated 3701 patients with AKI, averaging 65 years in age and 41.3% female. The subjects were arbitrarily split into two groups: a training group (2591 patients, 70%) and a validation group (1110 patients, 30%). Within the group of 10 machine learning models, the random forest (RF) model stood out as the most effective in discrimination and was analyzed through the SHAP technique. The ultimate model successfully forecasted AKI in both internal (AUC = 0.807) and external (AUC = 0.720) validations. This can aid medical professionals in evaluating the future outlook of patients in the intensive care unit (ICU) and enable prompt medical actions.

Keywords

Patients with Severe Illnesses, Immediate Renal Damage, Automated Learning, Predictive Forecasting Model, SHAP

Copyright © 2025 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/

CC O Open Access

1. 引言

在 ICU 中,AKI 的发病率和死亡率都处于较高水平[1]-[3]。这不仅严重影响患者的健康,还带来了 沉重的经济负担,而目前医学界尚未找到有效的根治方法[4]。危重症 AKI 的死亡机制十分复杂,涉及多 种因素相互交织。因此,构建一个包含多个相关风险因素的预测模型,或许能为解决这一难题提供更有 效的途径。MIMIC IV 数据库包含 6 万多名患者的临床数据,在医疗大数据中,机器学习方法可以更方便 地处理自变量的多重共线性,可以提高预后预测模型的预测辨别力[5][6]。在医学领域,机器学习技术如 随机森林和梯度提升机已被广泛应用[7][8]。本研究的最佳模型随机森林算法,通过 SHAP 方法,不仅提 供了对整体功能的全局解释,还详细阐述了如何利用个性化输入数据对重症 AKI 患者进行特定预测的局 部解释。最终模型在内部和外部验证中均展现出良好的性能,能够助力临床医生早期识别高危 AKI 患者 并及时进行干预。

2. 资料与方法

2.1. 一般资料

MIMIC 数据库对患者信息进行了脱敏处理。在获得人类研究参与者保护评估(证书编号 55810423)的 授权后,我们使用结构化查询语言(SQL)从该数据库中提取数据。主要研究终点为住院死亡率,数据被分 为训练队列和验证队列,比例 7:3。

急性肾损伤(AKI)的诊断标准依据改善全球肾脏病预后组织(KDIGO)的标准确定,包括:在48小时内血清肌酐增加0.3 mg/dl;血清肌酐较基线升高1.5倍,且该变化发生在前7天内;或者尿量 <0.5 mL/kg/小时,持续6小时。我们利用 Navicat Premium (版本16)中的 SQL 编程工具从 MIMIC-IV 数据库中提取临床数据,数据收集过程遵循了 Deshmukh 等[9]的既定程序。

该研究收集的变量特征包括年龄、性别、体重和身高。合并疾病包括高血压、糖尿病、充血性心力 衰竭、冠心病、急性心肌梗死、脑血管疾病、慢性肺部疾病、慢性肝病、恶性肿瘤。我们收集诊断 AKI24 小时心率、收缩压、舒张压、呼吸频率、体温和 SpO2、中心静脉压、白细胞、红细胞分布宽度、中性粒 细胞计数、血清氯、血清钠、血清钾、乳酸脱氢酶、动脉血乳酸、PCO2、碳酸氢盐、阴离子间隙、血清 纤维蛋白原、凝血酶原时间、活化部分凝血酶时间、D-二聚体、尿酸、肌酸激酶、肌酸激酶同工酶、BNP 和血糖、尿比重、丙氨酸转氨酶、碱性磷酸酶、天冬氨酸转氨酶、总胆红素的最大值,患者住 ICU 期间 最大的肌酐值及最大的 AKI 分期。选择红细胞、红细胞压积、血红蛋白、血小板、淋巴细胞计数、血清 钙、PH、PO2、白蛋白在诊断 AKI24 小时内的最小值。治疗包括在 ICU 住院期间是否机械通气和使用血 管升压药、CRRT 治疗,是否使用肾毒性药(头孢类抗生素、甘露醇)。我们计算了顺序器官衰竭评估(SOFA)、 APACHEII 最大值,斯哥昏迷评分(GCS)的最小值。存活组和死亡组患者在性别、年龄和实验室检查等一 般资料方面比较,差异无统计学意义(P>0.05),具有可比性。

纳入标准: (1) 年龄 ≥18 岁; (2) 符合肾脏疾病:改善全球结局(KDIGO) AKI 诊断标准; (3) ICU 住 院时间超过 6 小时; (4) 对于多次住院的患者,仅包括首次住院的信息。

排除标准为:(1)慢性肾脏病患者。(2)数据严重缺失患者。

2.2. 方法

从 MIMIC-IV 数据库中筛选出 28,484 例首次入住 ICU 的患者数据,排除了 6338 例合并慢性肾脏病、 27 例住院实际时间少于 6 小时的患者,剔除了患者缺失值超过 15%的患者,删除缺失高于 20%的特征变量,最终得到 3701 例可用于分析的患者数据。

采用 R 语言中的"mice"软件包对剩余数据中的缺失值进行插补。基于收缩压和舒张压计算出平均 动脉压、身高和体重计算出体重指数。利用方差膨胀因子(VIF)评估特征之间的共线性。存在高共线性的 特征组合包括:身高、体重与体重指数;收缩压、舒张压与平均动脉压;红细胞与血红蛋白,并排除了身 高、体重、收缩压、舒张压以及红细胞,其余 50 个特征用于构建预测模型。

2.3. 外部验证

MIMIC III 患者的外部数据集进行外部验证。纳入和排除标准与 MIMIC IV 数据库提取数据相同。

2.4. 模型开发和比较

本研究从 MIMIC IV 数据库中获取数据,并将其分为两部分:70%的数据用于模型训练,30%的数据 用于内部验证。此外,使用独立的外部数据集进行外部验证。研究中运用了10种机器学习模型来预测危 重急性肾损伤患者的死亡风险,包括自适应增强、人工神经网络、决策树、额外树、梯度增强机、k 近邻、 逻辑回归、随机森林、支持向量机和极限梯度增强。

分类变量以频率(百分比)的形式呈现,符合正态分布的连续变量以均值 ± 标准差表示,不符合正态 分布的连续变量则以中位数(四分位数范围)表示。为了评估模型的性能,采用受试者工作特征曲线下面积 (AUC)、敏感性、特异性、阳性预测值(PPV)、阴性预测值(NPV)、准确性和 F1 评分衡量模型的可靠性和 预测能力。

2.5. 特征选择和模型解释

在本研究中,我们通过多轮特征选择流程(包括低方差过滤、高相关性过滤、互信息筛选、SHAP 值 分析以及递归特征消除)来确定最佳的特征组合。基于 Python 3.12 版本,我们构建了 10 种不同的机器学 习模型,并对其预测性能进行了评估。SHAP 方法解释模型。

2.6. 统计学方法

数据分析采用 Python 3.12 和 R 4.3.2。不服从正态分布的连续变量,采用中位数和四分位数间距来描述其分布特征,Mann-Whitney U 检验或 Kruskal-Wallis H 检验进行不同组别间的比较;分类变量则以频数和百分比的形式呈现,统计分析采用卡方检验或费舍尔精确检验。模型的预测性能通过受试者工作特征曲线下面积(AUC)来评估。当双侧 P 值小于 0.05 时,认为结果具有统计学显著性。

3. 结果

3.1. 患者特征

在 3701 名患者中,中位年龄 65 岁,女性 1532 人(41.3%),417 名为 AKI 1 期,1529 名为 AKI 2 期,1755 名为 AKI 3 期。合并高血压 1065 人(28.7%),合并糖尿病 935 人(25。2%),死亡 999 人(26.9%)。

3.2. 模型开发和性能比较

10 个模型中,随机森林模型(AUC = 0.807)对死亡风险的预测效果最好,其次是额外树模型(AUC = 0.803)和梯度增强机模型(AUC = 0.8029)。表现最好的前 5 个模型的判别性能列于表 1。

机器学习模型	AUC	Sensitivity	Specificity	PPV	NPV	accuracy	F1 score
RF	0.807	0.376	0.942	0.703	0.807	0.703	0.490
ET	0.803	0.339	0.786	0.699	0.858	0.699	0.457
GBM	0.8029	0.403	0.929	0.672	0.929	0.672	0.504
XGboost	0.771	0.380	1.000	0.592	1.000	0.592	0.531
Adaboost	0.759	0.241	0.743	0.586	0.743	0.586	0.461

Table	e 1. Pr	edictive	performa	nce of t	he top :	5 machine	learning	; mode	ls
表1.	前 5	个机器学	之习模型	预测性	能				

PPV: 阳性预测值, NPV 阴性预测值。PPV: positive predictive value; NPV: negative predictive value。

3.2. 最终模型的识别

最终模型根据 10 个机器学习模型的 AUC 和预测准确率确定。10 种机器学习模型的 ROC 曲线见 图 1。



Figure 1. ROC curves of 10 machine learning models 图 1. 10 种机器学习模型 ROC 曲线

3.3. 最终模型的外部验证



Figure 2. ROC curve for external validation 图 2. 外部验证 ROC 曲线 对于外部验证,最终模型的 AUC 为 0.720,准确率为 0.850,表明最终模型在内部验证和外部验证中都表现出了较好的性能。外部验证 ROC 曲线见图 2。

3.4. 模型的解释



Figure 3. (A) SHAP summary bar chart of the best RF model; (B) SHAP summary lattice chart of the best RF model; (C) SHAP dependency graphThe dependency diagram shows how a single feature affects the output of the prediction model, with each point representing a single patient. A SHAP value for a specific feature greater than zero pushes the decision towards the death category

图 3. (A) RF 最佳模型 SHAP 汇总条形图; (B) RF 最佳模型 SHAP 汇总点阵图; (C) SHAP 依赖关系图依赖图显示单 个特征如何影响预测模型的输出,每个点表示单个患者。超过零的特定特征的 SHAP 值将决策推向死亡类 本研究采用 SHAP 方法解释模型,其提供了模型层面的全局解释和个体层面的局部解释。如图 3A 所示的 SHAP 汇总图,通过平均 SHAP 值来衡量各个特征对模型的贡献程度。此外, SHAP 依赖性图能够揭示单个特征如何影响模型的预测输出。例如,图 3C 展示了 12 个特征的真实值与对应的 SHAP 值之间的关系,其中大于零的 SHAP 值表示该特征对模型的预测产生了正向影响,即提示死亡风险较高。

局部解释专注于分析如何结合特定个体的输入数据来做出预测。以图 4 为例,根据预测模型,该患者的死亡概率为 69%。这种局部解释能够帮助临床医生更好地理解模型是如何针对个体患者做出具体预测的。



Figure 4. Explaining the local model using SHAP method. On the left, the patient characteristic value is shown. The bar chart shows that the SHAP value is positive, indicating that the feature promotes the death of patients, and if it is negative, it reduces the risk of death

图 4. 用 SHAP 方法解释局部模型。左侧为患者特征值,条形图 SHAP 值为正,说明该特征能促进患者死亡,如果为 负,则降低死亡风险

4. 讨论

本研究利用 MIMIC 数据库,运用 10 种机器学习方法建立了危重 AKI 患者死亡风险的预测模型。机器学习方法凭借其处理大数据的能力,能够有效处理高维数据。MIMIC 数据库数据量大、质量高,为预测模型的发展提供了有力支持。

在我们的数据中,所有特征缺失值均小于 20%。R 语言中"mice"包通过链式方程进行多重插补, 生成多个完整数据集,减少偏差,是处理缺失数据的可靠工具。临床预测模型的多轮特征选择通过"粗 筛→精筛→验证"分层优化,逐步剔除低方差、高相关冗余特征,并结合互信息和 SHAP 值捕捉非线性 关联与模型贡献,最终通过递归消除验证稳定性。其优势在于提升模型性能(如 AUC 提高)、增强临床可 解释性(保留关键指标)及高效降维(减少 90%无关变量),同时平衡计算效率与复杂数据适应性,为临床决 策提供精简可靠的预测工具。我们基于随机森林算法构建了一个包含 12 个特征的最终预测模型。这些特 征在 ICU 入院期间能够轻松获取或评估。

在模型中,血清阴离子间隙是影响危重 AKI 患者死亡风险的最重要因素。已有研究[10]证实,高阴 离子间隙(AG)与脓毒症[11]等疾病的严重程度或不良预后呈正相关。其升高常与乳酸酸中毒、酮症酸中毒 或尿毒症相关。酸中毒通过激活肾小管凋亡通路,抑制肾小管再生能力,延缓肾功能恢复,还可抑制心 肌收缩力、降低血管反应性,加重休克并增强促炎因子(如 IL-6、TNF-*a*)释放,增加患者死亡风险。此外, 年龄也是重症患者死亡率的重要影响因素,随着年龄的增长,重症患者的死亡率显著上升[12]。随着年龄 增加,肾单位逐渐减少,肾动脉硬化导致肾脏血流调节能力减弱,且肾小管上皮细胞再生能力降低,AKI 后肾功能恢复延迟,肾小球滤过率下降更易发生肾毒性药物相关 AKI,且高龄患者固有免疫与适应性免 疫减退,增加患者死亡风险。高钠血症被证明是 ICU 住院死亡的独立危险因素,而低钠血症可能导致神 经系统症状、昏迷甚至死亡[13]。低尿量同样是 AKI 患者死亡的危险因素,它可能引发高钾血症、代谢 性酸中毒和多器官功能衰竭等并发症。凝血酶原时间主要反映凝血因子的合成状态,是评估重症患者凝 血功能的重要指标,间接影响患者的预后[14]。

高 BMI (肥胖)患者脂肪细胞分泌促炎因子(如 IL-6、瘦素),加剧全身炎症反应,游离脂肪酸(FFA)堆积导致肾小球系膜细胞凋亡,加重蛋白尿和肾纤维化,增加患者死亡风险。血清白蛋白水平是反映营养状态的重要指标。白蛋白可结合自由基(如活性氧 ROS),抑制脂质过氧化,保护肾小管上皮细胞,并可结合内毒素及促炎因子(如 IL-6、TNF-a),减轻全身炎症反应。低白蛋白减少对自由基的清除能力,导致线粒体功能障碍和肾小管上皮细胞凋亡。血清白蛋白每下降 10 g/L,AKI 患者死亡风险增加 40%。为了阐释机器学习模型的"黑箱"问题,我们还采用了 SHAP 解释模型。

本研究构建的预测模型未区分病因。然而,不同病因对重症监护室急性肾损伤患者的预后可能存在 影响。这一不足可能会限制模型在特定病因患者群体中的预测精准度,未来研究可考虑将病因纳入模型 构建因素,以进一步优化模型性能。

然而,本研究也存在一些局限性。第一,模型中缺失了一些重要的特征,例如中心静脉压和心肌酶 等。第二,我们在未考虑急性肾损伤病因的情况下建立了 AKI 死亡风险预测模型。第三,该模型基于美 国人群的数据库构建和验证,其在全球人群中的适用性尚不明确,需要进一步评估其泛化能力。第四, ICU 住院期间诊断 AKI 后的 24 小时数据可能存在选择偏倚。

综上,我们成功开发了一个具有临床应用潜力的预测模型。该模型最终纳入的 12 个特征在临床实践 中易于获取。最终的随机森林模型在内部和外部验证中均表现出良好的预测能力。未来需要开展更多的 随机对照研究,以确定基于该预测模型的及时治疗措施是否能够改善重症 AKI 患者的预后。

参考文献

- [1] Chawla, L.S., Amdur, R.L., Shaw, A.D., Faselis, C., Palant, C.E. and Kimmel, P.L. (2014) Association between AKI and Long-Term Renal and Cardiovascular Outcomes in United States Veterans. *Clinical Journal of the American Society of Nephrology*, 9, 448-456. <u>https://doi.org/10.2215/cjn.02440213</u>
- [2] Bouchard, J., Soroko, S.B., Chertow, G.M., Himmelfarb, J., Ikizler, T.A., Paganini, E.P., *et al.* (2009) Fluid Accumulation, Survival and Recovery of Kidney Function in Critically III Patients with Acute Kidney Injury. *Kidney International*, 76, 422-427. <u>https://doi.org/10.1038/ki.2009.159</u>
- [3] Coca, S.G., Yusuf, B., Shlipak, M.G., Garg, A.X. and Parikh, C.R. (2009) Long-Term Risk of Mortality and Other Adverse Outcomes after Acute Kidney Injury: A Systematic Review and Meta-Analysis. *American Journal of Kidney*

Diseases, 53, 961-973. https://doi.org/10.1053/j.ajkd.2008.11.034

- [4] Hoste, E.A.J., Bagshaw, S.M., Bellomo, R., Cely, C.M., Colman, R., Cruz, D.N., et al. (2015) Epidemiology of Acute Kidney Injury in Critically Ill Patients: The Multinational AKI-EPI Study. Intensive Care Medicine, 41, 1411-1423. https://doi.org/10.1007/s00134-015-3934-7
- [5] Chen, V., Li, J., Kim, J.S., Plumb, G. and Talwalkar, A. (2022) Interpretable Machine Learning. *Communications of the ACM*, **65**, 43-50. <u>https://doi.org/10.1145/3546036</u>
- Song, X., Liu, X., Liu, F. and Wang, C. (2021) Comparison of Machine Learning and Logistic Regression Models in Predicting Acute Kidney Injury: A Systematic Review and Meta-Analysis. *International Journal of Medical Informatics*, 151, Article 104484. <u>https://doi.org/10.1016/j.ijmedinf.2021.104484</u>
- [7] Yue, S., Li, S., Huang, X., Liu, J., Hou, X., Zhao, Y., et al. (2022) Machine Learning for the Prediction of Acute Kidney Injury in Patients with Sepsis. *Journal of Translational Medicine*, 20, Article No. 215. <u>https://doi.org/10.1186/s12967-022-03364-0</u>
- [8] Katz, S., Suijker, J., Hardt, C., Madsen, M.B., Vries, A.M., Pijpe, A., *et al.* (2022) Decision Support System and Outcome Prediction in a Cohort of Patients with Necrotizing Soft-Tissue Infections. *International Journal of Medical Informatics*, 167, Article 104878. <u>https://doi.org/10.1016/j.ijmedinf.2022.104878</u>
- [9] Deshmukh, F. and Merchant, S.S. (2020) Explainable Machine Learning Model for Predicting GI Bleed Mortality in the Intensive Care Unit. *American Journal of Gastroenterology*, **115**, 1657-1668. https://doi.org/10.14309/ajg.0000000000632
- [10] Cheng, B., Li, D., Gong, Y., Ying, B. and Wang, B. (2020) Serum Anion Gap Predicts All-Cause Mortality in Critically Ill Patients with Acute Kidney Injury: Analysis of the MIMIC-III Database. *Disease Markers*, 2020, Article 6501272. https://doi.org/10.1155/2020/6501272
- [11] Akinosoglou, K., Schinas, G., Almyroudi, M.P., Gogos, C. and Dimopoulos, G. (2023) The Impact of Age on Intensive Care. Ageing Research Reviews, 84, Article 101832. <u>https://doi.org/10.1016/j.arr.2022.101832</u>
- [12] Mohr, N.M., Vakkalanka, J.P., Faine, B.A., Skow, B., Harland, K.K., Dick-Perez, R., et al. (2018) Serum Anion Gap Predicts Lactate Poorly, but May Be Used to Identify Sepsis Patients at Risk for Death: A Cohort Study. Journal of Critical Care, 44, 223-228. <u>https://doi.org/10.1016/j.jcrc.2017.10.043</u>
- [13] Grim, C.C.A., Termorshuizen, F., Bosman, R.J., Cremer, O.L., Meinders, A.J., Nijsten, M.W.N., *et al.* (2021) Association between an Increase in Serum Sodium and in-Hospital Mortality in Critically III Patients. *Critical Care Medicine*, 49, 2070-2079. <u>https://doi.org/10.1097/ccm.000000000005173</u>
- [14] Arshad, A., Ahmed, W., Rehman, N., Naseem, Z. and Ghos, Z. (2024) Tackling a Deadly Global Phenomenon: Sepsis Induced Coagulopathy: A Narrative Review. *Journal of the Pakistan Medical Association*, 74, 959-966. <u>https://doi.org/10.47391/jpma.10194</u>