论著·智能医学与预测模型

XGBoost-Cox模型可有效预测老年男性急性肾损伤患者死亡风险:一项融合非线性特征工程与可解释机器学习的回顾性队列研究

宋秋月[1],敖强国[2],伍亚舟[1] [1]陆军军医大学(第三军医大学)军事预防医学系军队卫生统计学教研室,重庆;[2]中国人民解放军总医院第二医学中心肾脏病科,北京

[摘要] **目的** 老年急性肾损伤(acute kidney injury, AKI)患者死亡风险高,且随AKI后肾功能变化而改变,而现有预测模型多基于基线变量、可解释性有限。本研究基于老年男性急性肾损伤患者真实世界随访数据,构建融合非线性特征工程和可解释机器学习生存模型的死亡风险预测框架,并评价其区分度、校准度、临床净获益和可解释性。**方法** 本研究为单中心回顾性队列预测模型构建研究,纳入中国人民解放军总医院2008年1月至2025年12月期间发生急性肾损伤7 d后仍存活的老年男性住院患者,以全因死亡为终点结局。2008年1月至2018年12月的患者为模型开发队列,并按7∶3比例采用分层随机抽样方法分为训练集和内部测试集;2019年1月至2025年12月的患者为时间验证队列。收集人口学特征指标,心血管及血流动力学状态指标,肾功能指标,AKI严重程度及肾功能动态变化指标,血液学指标,生化、营养及代谢指标,电解质及矿物质指标,既往病史、合并症及治疗相关因素等指标。采用限制性立方样条(restricted cubic spline, RCS)分析关键连续变量与死亡风险之间的非线性关系;随后结合Boruta算法、随机生存森林(random survival forest, RSF)及临床可解释性进行特征筛选,构建3个递进式预测变量集;分别采用6种生存模型:Cox比例风险模型、样条扩展Cox模型、Elastic-net Cox、RSF、XGBoost-Cox和DeepSurv进行死亡风险预测。采用C-index和时间依赖AUC评价模型区分度,并通过校准曲线和决策曲线分析评价最佳模型的校准度和临床实用性。进一步采用SHAP方法对最佳模型进行全局和个体化解释。**结果** 最终纳入老年男性AKI患者2 335例,其中随访期间死亡1 912例,存活423例。训练集、内部测试集和时间验证集分别为775例、332例和1 228例。B型利钠肽、血尿素氮及AKI后短期估算肾小球滤过率动态变化指标与死亡风险存在非线性关系(P<0.01)。年龄、营养状态、心肾功能负荷、血液学指标、合并症负担、机械通气及AKI后肾功能变化情况是老年AKI患者死亡风险的重要预测因素。纳入AKI后动态肾功能指标的变量集整体预测性能优于仅包含基线变量或AKI发生时变量的模型。测试集中XGBoost-Cox模型获得最高C-index和3年AUC,分别为0.749(95%*CI*:0.722~0.774)、0.846(95%*CI*:0.802~0.891);RSF模型在30 d和90 d预测中表现最好,分别为0.840(95%*CI*:0.785~0.888)、0.829(95%*CI*:0.785~0.868)。时间验证结果中,XGBoost-Cox模型在30 d、90 d、1年和3年死亡风险预测中的AUC分别为0.797(95%*CI*:0.763~0.830)、0.781(95%*CI*:0.755~0.809)、0.802(95%*CI*:0.778~0.824)和0.822(95%*CI*:0.788~0.853),提示模型在时间验证队列中具有较为稳定的泛化性能。XGBoost-Cox模型区分能力较好、3年预测校准总体一致且具有临床净获益,肾功能负荷为其主导预测因素。**结论** 融合非线性特征工程与可解释机器学习生存模型的预测框架能

[基金项目] 国家自然基金面上项目(82574207);国家重点研发计划(2023YFC3605500,2023YFC3605501);重庆市自然科学基金面上项目(CSTB2025NSCQ-GPX0648)

[通信作者] 伍亚舟,E-mail:asiawu@tmmu.edu.cn

2097-0927/Copyright © 2026, 陆军军医大学(第三军医大学)。这篇开放获取文章遵循CC BY许可协议(https://creativecommons.org/licenses/by/4.0/)

引用格式:宋秋月,敖强国,伍亚舟. XGBoost-Cox模型可有效预测老年男性急性肾损伤患者死亡风险:一项融合非线性特征工程与可解释机器学习的回顾性队列研究[J]. 陆军军医大学学报,2026,48(20):2933-2948. DOI:10.16016/j.2097-0927.202607040.

Citation: Song Q Y, Ao Q G, Wu Y Z. XGBoost-Cox model effectively predicts mortality risk in elderly male patients with acute kidney injury: a retrospective cohort study integrating nonlinear feature engineering and interpretable machine learning[J]. J Army Med Univ, 2026, 48(20): 2933-2948. DOI: 10.16016/j.2097-0927.202607040.

够有效识别老年男性AKI患者死亡风险。其中,XGBoost-Cox模型在内部测试与时间验证中区分能力较好,3年死亡风险预测校准总体一致、临床净获益优于全治与不治策略,模型预测依据具有可解释性。

[关键词] 急性肾损伤;老年男性;机器学习;生存分析;死亡风险
[中图法分类号] R195.1;R339.34;R692.07 [文献标志码] A

XGBoost-Cox model effectively predicts mortality risk in elderly male patients with acute kidney injury: a retrospective cohort study integrating nonlinear feature engineering and interpretable machine learning

Song Qiuyue[1], Ao Qiangguo[2], Wu Yazhou[1] [1]Department of Health Statistics, College of Military Preventive Medicine, Army Medical University (Third Military Medical University), Chongqing; [2]Department of Nephrology, Second Medical Center, Chinese PLA General Hospital, Beijing, China

[**Abstract**] **Objective** Elderly patients with acute kidney injury (AKI) face a high mortality risk that varies with renal function trajectories after AKI, whereas existing prediction models rely mainly on baseline variables and offer limited interpretability. Based on real-world follow-up data from elderly male patients with AKI, this study aimed to develop a mortality risk prediction framework that integrates nonlinear feature engineering with interpretable machine-learning survival models, and to evaluate its discrimination, calibration, clinical net benefit, and interpretability. **Methods** A single-center, retrospective cohort study was adopted to construct a prediction model based on older male inpatients who survived at least 7 days after AKI onset at Chinese PLA General Hospital between January 2008 and December 2025, with all-cause mortality as the endpoint outcome. Patients admitted between January 2008 and December 2018 constituted the model development cohort, and were randomly divided into a training set and an internal test set at a 7∶3 ratio; patients admitted between January 2019 and December 2025 constituted the temporal validation cohort. Variables collected included demographic characteristics, cardiovascular and hemodynamic status, renal function indicators, AKI severity and dynamic changes in renal function, hematological indicators, biochemical, nutritional and metabolic indices, electrolytes and minerals, medical history, comorbidities, and treatment-related factors. Restricted cubic splines (RCS) analysis was used to examine the nonlinear relationships between key continuous variables and mortality risk. Feature selection was performed by combining the Boruta algorithm, random survival forest (RSF), and clinical interpretability to construct 3 progressively expanded predictor sets. Six survival models were applied for mortality risk prediction, namely the Cox proportional hazards model, spline-extended Cox model, elastic-net Cox model, RSF, XGBoost-Cox, and DeepSurv. Discrimination was evaluated by the C-index and time-dependent AUC, and the calibration and clinical utility of the best-performing model were assessed by calibration curves and decision curve analysis. SHapley Additive exPlanations (SHAP) were further applied to interpret the best model globally and at the individual level. **Results** A total of 2 335 elderly male patients with AKI were finally included, of whom 1 912 died during follow-up and 423 survived. The training, internal test, and temporal validation sets comprised 775, 332, and 1 228 patients, respectively. BNP, BUN, and short-term post-AKI dynamic eGFR indicators showed nonlinear relationships with mortality risk ($P<0.01$). Age, nutritional status, cardiorenal functional load, hematological indicators, comorbidity burden, mechanical ventilation, and post-AKI changes in renal function were important predictors of mortality risk in elderly patients with AKI. The predictor set incorporating post-AKI dynamic renal function indicators revealed better overall predictive performance than models containing only baseline variables or AKI-onset variables. In the test set, the XGBoost-Cox model achieved the highest C-index and 3-year AUC, which were 0.749 (95%*CI*: 0.722 to 0.774) and 0.846 (95%*CI*: 0.802 to 0.891), respectively; RSF performed best in 30-day and 90-day predictions, with AUC

values of 0.840 (95%*CI*: 0.785 to 0.888) and 0.829 (95%*CI*: 0.785 to 0.868), respectively. In temporal validation, the AUCs of the XGBoost-Cox model for 30-day, 90-day, 1-year, and 3-year mortality risk prediction were 0.797 (95%*CI*: 0.763 to 0.830), 0.781 (95%*CI*: 0.755 to 0.809), 0.802 (95%*CI*: 0.778 to 0.824), and 0.822 (95%*CI*: 0.788 to 0.853), respectively, indicating stable generalizability in the temporal validation cohort. The XGBoost-Cox model showed good discrimination, calibration generally consistent with observed risk at 3 years, and clinical net benefit, with renal functional load as its dominant predictor.

Conclusion The prediction framework integrating nonlinear feature engineering with interpretable machine-learning survival models can effectively identify mortality risk in elderly male patients with AKI. The XGBoost-Cox model showed good discrimination in both internal testing and temporal validation, with generally consistent calibration and net benefit superior to treat-all and treat-none strategies for 3-year mortality risk prediction; SHAP analysis identified renal functional load as its dominant predictor, indicating that the model's predictions are interpretable.

[**Key words**] acute kidney injury; older men; machine learning; survival analysis; mortality risk

Supported by the General Program of National Natural Science Foundation of China (82574207), the National Key Research and Development Program (2023YFC3605500, 2023YFC3605501), and the General Project of Natural Science Foundation of Chongqing (CSTB2025NSCQ-GPX0648).
Corresponding author: Wu Yazhou, E-mail: asiawu@tmmu.edu.cn

急性肾损伤(acute kidney injury, AKI)是老年住院患者常见且预后严重的危重综合征,住院患者发生率为10%~15%,ICU患者发生率超过50%,可显著增加短期死亡风险及远期慢性肾脏病、终末期肾病和心血管事件风险[1-3]。老年患者存在肾功能储备功能减退与合并症多发的特征,AKI发生后病情进展及远期预后存在显著异质性,其中男性老年AKI患者死亡风险水平更高[4]。因此,精准识别老年男性AKI患者的死亡风险,对开展早期风险分层、优化临床治疗决策与长期随访管理均具有重要的临床意义。

既往AKI预后研究多采用Logistic回归或Cox比例风险模型,这类方法具有较好可解释性和可推广性,但通常依赖线性关系、比例风险假设,并常以单因素分析或经验判断进行变量筛选[5]。对于老年AKI这种由多系统、多器官、多维临床因素共同驱动的复杂综合征,单纯依靠AKI分期、血清肌酐、尿量或单一时间点实验室指标,难以充分刻画个体化死亡风险。尤其是年龄、B型利钠肽(B-type natriuretic peptide, BNP)、血尿素氮(blood urea nitrogen, BUN)、电解质、酸碱指标及AKI后肾功能动态变化等连续变量,其与死亡风险之间可能存在阈值效应或非线性关系,若按简单线性形式处理,可能影响模型预测性能和临床解释[6]。近年来临床预测模型研究也强调,模型开发不应仅关注统计显著性或单一算法表现,还需重视变量选择合理性、过拟合控制、外部验证、校准度、临床实用性及透明报告[7]。

限制性立方样条(restricted cubic spline, RCS)、Boruta特征筛选、机器学习生存模型和可解释人工智能为复杂临床预测模型提供了新的技术路径[8-11]。RCS可用于识别连续变量与结局之间的非线性剂量-反应关系;生存模型能够同时利用死亡时间和删失信息,适合真实世界随访数据;随机生存森林(random survival forest, RSF)、XGBoost-Cox和DeepSurv等机器学习生存模型有助于捕捉变量间复杂交互和非线性结构[12-14]。然而,模型性能提升并不等同于临床可用,仍需从区分度、校准度、临床净获益、时间验证及可解释性等方面进行综合评价[15-17]。此外,AKI并非静态事件,AKI发生后48 h至7 d内的短期肾功能变化可能反映肾脏储备、循环状态、炎症负荷和多器官功能状态,具有潜在增量预测价值[1-2]。

基于老年AKI患者真实世界随访数据,本研究构建融合非线性特征工程、机器学习生存模型和SHAP可解释分析的死亡风险预测框架。首先识别关键连续变量与死亡风险之间的非线性关系,并结合Boruta算法、RSF筛选稳定且具有临床意义的预测因子;随后构建基线变量、AKI发生时变量及AKI后短期动态肾功能变量3个递进式变量集,评估动态指标的增量预测价值;进一步比较传统Cox与多种机器学习生存模型在内部测试集和时间验证集中的表现,并通过校准曲线、决策曲线和SHAP分析评价最佳模型的校准度、临床实用性和可解释性,为老年男性AKI患者个体化风险评估和临床决策支持提供依据。

1 资料与方法

1.1 数据来源与研究对象

本研究为单中心回顾性队列预测模型构建研究,

使用医院电子病历系统和实验室信息系统中的脱敏资料,经中国人民解放军总医院医学伦理委员会审核批准(伦审第S2023-725-01号)。2008年1月至2018年12月期间住院并诊断为AKI的老年男性患者作为模型开发队列,共1 107例,并按照7∶3比例采用分层随机抽样方法划分为训练集(775例)和内部测试集(332例);2019年1月至2025年12月期间住院并诊断为AKI的老年男性患者作为独立时间验证队列,共1 228例,用于评估模型随时间推移的泛化能力,研究对象纳排流程见图1。

本研究以KDIGO 2012标准[7]诊断AKI患者,研究主要结局设定为随访期间的全因死亡,观察起始点为患者首次满足AKI诊断标准的时间。由于本研究纳入AKI发生后48 h至7 d的动态肾功能指标,故最终分析仅保留AKI发生后存活超过7 d且随访资料完整的病例,重点评估30 d、90 d、1年及3年的全因死亡风险。

本研究对各时间节点的定义如下:基线定义为患者本次住院入院后、AKI发生前,可获取的最接近入院时间点的临床与实验室检测资料,用于反映患者AKI发生前的基础状态;AKI发生时定义为患者首次符合KDIGO AKI诊断标准的时间点,用于反映AKI事件发生时的急性病情状态;AKI后动态肾功能指标定义为AKI发生后48 h至7 d时间窗口内与估算肾小球滤过率(estimated glomerular filtration rate,eGFR)变化相关的指标,用于反映AKI发生后早期的肾功能恢复轨迹。

1.2 候选变量与数据预处理

1.2.1 人口学特征和心血管及血流动力学状态指标

人口学特征指标:年龄。心血管及血流动力学状态指标:基线收缩压(systolic blood pressure, SBP)、基线舒张压(diastolic blood pressure, DBP)、AKI发生时SBP、AKI发生时DBP,基线BNP和AKI发生时BNP,该类变量用于反映循环状态、血流动力学稳定性及心血管负荷。

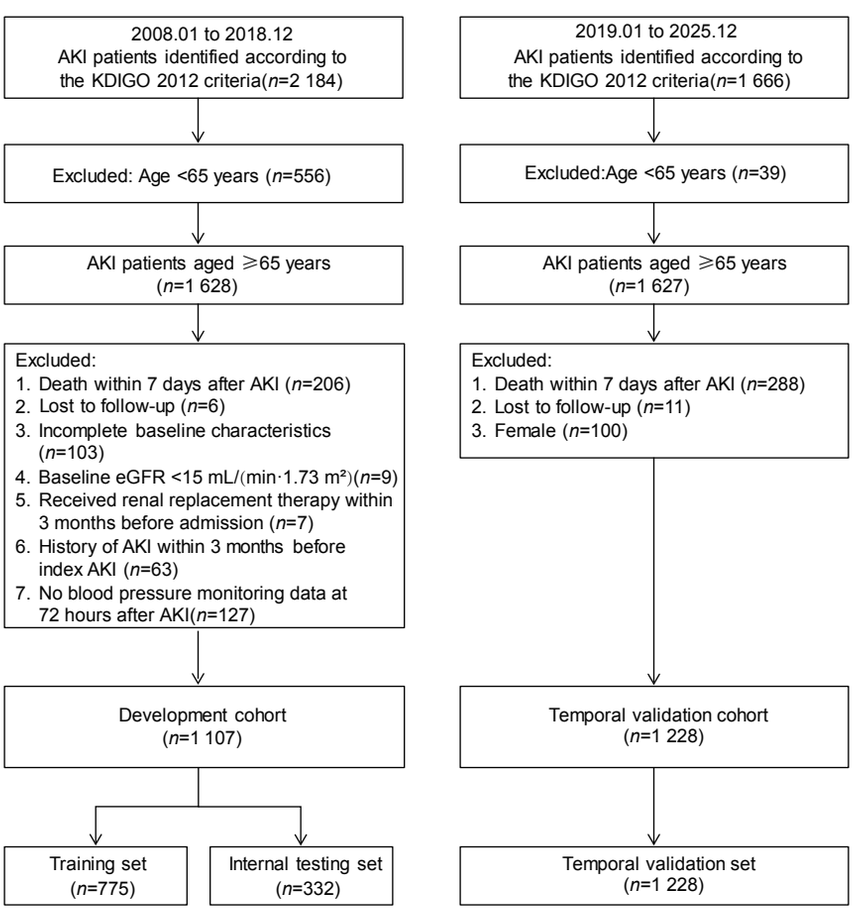


图1 研究对象纳入和排除流程图

Figure 1 Flowchart of participant inclusion and exclusion

1.2.2 肾功能指标 基线血清肌酐(serum creatinine, SCr)、AKI发生时SCr、AKI后48 h内峰值SCr、AKI后48 h内末次SCr、AKI后7 d内峰值SCr、AKI后7 d内末次SCr,基线eGFR、AKI发生时eGFR、AKI后48 h内最高eGFR、AKI后48 h内末次eGFR、AKI后7 d内最高eGFR、AKI后7 d内末次eGFR,基线BUN和AKI发生时BUN。eGFR采用CKD-EPI 2009的公式计算。

1.2.3 AKI严重程度及肾功能动态变化指标 AKI分期,48 h SCr比值、7 d SCr比值、早期SCr升高百分比、第7天SCr超出基线百分比、48 h至7 d SCr恢复百分比,48 h eGFR比值、7 d eGFR比值、早期eGFR下降百分比、第7天eGFR缺损百分比和48 h至7 d eGFR恢复百分比。AKI分期依据KDIGO 2012指南确定。SCr比值和eGFR比值用于量化AKI后肾功能偏离基线的程度,恢复百分比用于反映AKI后早期肾功能改善情况。基于AKI后SCr和eGFR的动态变化进一步构建肾功能动态派生变量计算公式见表1。SCr相关指标基于观察窗口内的SCr峰值计算,表示该时间窗内肌酐升高的最高水平;eGFR相关指标基于观察窗口内的末次测量值计算,表示该时间窗末的滤过功能水平。除恢复百分比外,其余指标数值越大,表示肾功能损伤越明显。恢复百分比数值越大,表示48 h至7 d期间肾功能改善越明显。

1.2.4 血液学指标 基线血红蛋白(hemoglobin, Hb)、基线红细胞(red blood cell, RBC)计数、基线血小板(platelet, PLT)计数、AKI发生时Hb、AKI发生时RBC计数和AKI发生时PLT计数。该类变量用于反映贫血、血液系统状态及凝血相关变化。

1.2.5 生化、营养及代谢指标 基线白蛋白、AKI发生时白蛋白,基线血糖、AKI发生时血糖,基线尿酸和AKI发生时尿酸。该类变量用于评价营养状态、代谢紊乱和疾病负担。

1.2.6 电解质及矿物质指标 基线血钾、血钠、血钙、血氯、血磷,以及AKI发生时血钾、血钠、血钙、血氯和血磷。该类变量用于反映电解质紊乱和矿物质代谢异常。

1.2.7 既往病史、合并症及治疗相关因素 吸烟史、高血压、糖尿病、高胆固醇血症、高尿酸血症、冠心病(coronary heart disease, CHD)、心肌梗死(myocardial infarction, MI)、心房颤动(atrial fibrillation, AF)、慢性心功能不全/慢性心力衰竭(chronic heart failure, CHF)、脑卒中、恶性肿瘤、慢性阻塞性肺疾病(chronic obstructive pulmonary disease, COPD)、慢性肾脏病(chronic kidney disease, CKD)和机械通气(mechanical ventilation, MV)。该类变量用于描述患者基础疾病负担、器官储备状态及治疗支持情况。

本研究在数据提取时严格控制质量。基线描述和组间比较尽量基于原始可用数据进行,不对表2描述性统计结果作额外缺失值填补或异常值替换。进入模型训练阶段时,为保证不同算法能够在同一建模矩阵上运行,若候选变量仍存在少量缺失,则仅在训练集内计算连续变量中位数和分类变量众数进行填补,并将同一填补参数应用于内部测试集和时间验证队列,避免评价数据参与模型开发。连续变量标准化和分类变量哑变量编码根据具体模型算法要求,在相应模型训练流程中完成。

1.3 非线性特征工程

为充分刻画老年男性AKI患者死亡风险的复杂影响因素,本研究将非线性特征工程定义为在临床先验和数据驱动基础上,对候选变量进行动态特征构建、非线性关系识别和Boruta特征筛选的综合分析过程。

首先结合单因素分析、临床意义及后续特征筛选结果,对多个候选连续变量进行探索性RCS分析。考虑论文展示的简洁性和临床解释价值,最终选择6个关键连续指标进行重点展示,包括基线BNP、基线BUN、第7天eGFR缺损百分比、AKI发生时BNP、AKI发生时BUN以及48 h至7 d eGFR恢复百分比。上述变量分别代表心血管负荷、代谢负荷和AKI后短期肾功能动态恢复情况。

根据变量可获得时间构建3个递进式候选变量集。Variable set 1仅纳入基线变量;Variable set 2在Variable set 1基础上进一步加入AKI发生时变量;Variable set 3在Variable set 2基础上增加AKI后48 h至7 d内的动态肾功能指标,以反映AKI后短期肾功能变化过程。

采用Boruta算法[8]进行特征筛选,以死亡状态作为二分类结局,识别与死亡结局稳定相关的候选预测因子。考虑到Boruta算法未直接纳入生存时间和删失信息,本研究进一步采用RSF评估变量对生存结局预测的贡献,选取重要性排名前20的特征。将Boruta确认变量与RSF筛选变量取并集,形成3个变量集对应模型的初始建模变量。为降低变量冗余和多重共线性,进一步结合方差膨胀因子(variance inflation factor, VIF)>5进行变量精简。最终保留的变量用于后续模型训练、内部测试和时间验证。

表1 AKI后肾功能动态衍生变量的定义

Table 1 Definitions of derived variables for post- acute kidney injury (AKI) dynamic renal function changes

衍生变量 Derived variable	定义公式 Definition formula	指标含义 Interpretation
早期损伤严重程度 Early injury severity		
48 h SCr比值 48 h SCr ratio	AKI后48 h内峰值SCr/基线SCr Peak SCr within 48 h after AKI/baseline SCr	数值越大,提示早期SCr升高越明显。 The larger the value, the more pronounced the early increase in SCr.
48 h eGFR比值 48 h eGFR ratio	基线eGFR/AKI后48 h内末次eGFR Baseline eGFR / last eGFR within 48 h after AKI	数值越大,提示早期eGFR下降越明显。 The larger the value, the more pronounced the early decline in eGFR.
早期SCr升高百分比 Percentage of early SCr rise	[(AKI后48 h内峰值SCr-基线SCr)/基线SCr]×100% [(Peak SCr within 48 h after AKI-baseline SCr) / baseline SCr]×100%	数值越大,提示早期肾功能损伤越重。 The larger the value, the more severe the early renal function injury.
早期eGFR下降百分比 Percentage of early eGFR drop	[(基线eGFR-AKI后48 h内末次eGFR)/基线eGFR]×100% [(Baseline eGFR-last eGFR within 48 h after AKI) / baseline eGFR] × 100%	数值越大,提示早期eGFR损失越明显。 The larger the value, the more obvious the early loss of eGFR.
7 d残余肾功能损伤 Residual renal function at 7 d		
7 d SCr比值 7 d SCr ratio	AKI后7 d内峰值SCr/基线SCr Peak SCr within 7 d after AKI / baseline SCr	数值越大,提示7 d内SCr异常升高越明显。 The larger the value, the more pronounced the abnormal SCr elevation within 7 d.
7 d eGFR比值 7 d eGFR ratio	基线eGFR/AKI后7 d内末次eGFR Baseline eGFR / last eGFR within 7 d after AKI	数值越大,提示7 d末eGFR恢复不足或残余损伤越明显。 The larger the value, the greater the insufficient eGFR recovery or residual injury at day 7.
第7天SCr超出基线百分比 7 d SCr excess	[(AKI后7 d内末次SCr-基线SCr)/基线SCr]×100% [(Last SCr within 7 d after AKI-baseline SCr) / baseline SCr]×100%	数值越大,提示7 d内肌酐负荷越高。 The larger the value, the higher the creatinine burden within 7 d.
第7天eGFR缺损百分比 7 d eGFR deficit	[(基线eGFR-AKI后7 d内末次eGFR)/基线eGFR]×100% [(Baseline eGFR-last eGFR within 7 d after AKI) / baseline eGFR] × 100%	数值越大,提示7 d末eGFR水平较基线下降越多。 The larger the value, the greater the decrease in eGFR at day 7 relative to baseline.
48 h至7 d恢复幅度 Recovery magnitude from 48 h to 7 d		
48 h至7 d SCr恢复百分比 Percentage of SCr recovery from 48 h to 7 d	[(AKI后48 h内峰值SCr-AKI后7 d内末次SCr)/AKI后48 h内峰值SCr]×100% [(Peak SCr within 48 h after AKI-last SCr within 7 d after AKI) / peak SCr within 48 h after AKI] × 100%	数值越大,提示Scr下降越明显;若数值较小或为负,提示Scr下降不足或持续升高。 The larger the value, the more obvious the decrease in SCr; a small or negative value indicates insufficient decline or continued elevation of SCr.
48 h至7 d eGFR恢复百分比 Percentage of eGFR recovery from 48 h to 7 d	[(AKI后7 d内末次eGFR-AKI后48 h内末次eGFR)/AKI后48 h内末次eGFR]×100% [(Last eGFR within 7 d after AKI-last eGFR within 48 h after AKI) / last eGFR within 48 h after AKI] × 100%	数值越大,提示eGFR改善越明显;若数值较小或为负,提示eGFR恢复不足或进一步下降。 The larger the value, the more obvious the improvement in eGFR; a small or negative value indicates insufficient recovery or further decline of eGFR.

1.4 生存预测模型构建

本研究采用Cox比例风险模型、样条扩展Cox模型、Elastic-net Cox、RSF、XGBoost-Cox和DeepSurv进行死亡风险建模,比较3个变量集和不同算法的预测性能。最终根据内部测试集和时间队列验证结果,确定表现较优的模型,进一步用于校准分析、决策曲线分析和SHAP可解释性分析。

1.4.1 Cox比例风险模型和样条扩展Cox模型

Cox比例风险模型作为传统统计学基准模型,与样条扩展Cox及机器学习生存模型进行比较。为评估核心

连续变量非线性建模对预测性能的影响,本研究构建样条扩展Cox,统称RCS-Cox。模型先通过RCS绘制关键连续变量与死亡风险的剂量-反应曲线,模型比较阶段则对6个预设关键连续指标引入自然样条项,包括基线BNP、基线BUN、AKI发生时BNP、AKI发生时BUN、第7天eGFR缺损百分比和48 h至7 d eGFR恢复百分比,自由度设为3,其他变量仍以原始形式进入模型。

1.4.2 Elastic-net Cox 构建Elastic-net Cox惩罚回归模型,实现多变量在生存预测中实现收缩并提高模型稳定性。Elastic-net Cox结合L1和L2惩罚项,适用于存在变量相关性和潜在多重共线性的临床预测场景。以Cox部分似然损失函数为优化目标,混合参数α设为0.5;正则化参数λ通过交叉验证确定,交叉验证折数根据训练集中事件数自动设置为3~10折,评价指标采用一致性指数。最终采用lambda.min对应的参数进行预测,并输出个体死亡风险评分。

1.4.3 RSF 构建RSF以捕捉变量与死亡风险之间的非线性关系及高阶交互作用。RSF通过自助抽样构建多棵生存树,并在每个节点随机选择部分候选变量进行分裂,从而降低模型方差并提高预测稳定性。本研究设置树的数量为800,终端节点最小样本量为18,每次分裂候选变量数设为max(2, floor(sqrt(p))),其中p为进入模型的预测变量数;分裂准则采用Log-rank方法,并计算变量重要性。模型内缺失值处理基于变量分布特征的单次插补方法,随机种子设为20260619。

1.4.4 XGBoost-Cox 构建基于梯度提升框架的XGBoost-Cox。为控制模型复杂度并保持分析流程稳定,本研究采用预设参数训练模型:目标函数为survival:cox,评价指标为cox-nloglik,学习率η设为0.03,最大树深度设为2,最小子节点权重设为20,样本采样比例和变量采样比例均设为0.80,L2正则化参数λ设为2.0,L1正则化参数α设为0.5,迭代轮数设为500,线程数设为4。

1.4.5 DeepSurv 构建DeepSurv作为深度学习生存预测模型。DeepSurv通过多层前馈神经网络学习非线性风险函数,并以Cox部分似然作为损失函数进行模型训练。模型输入变量在训练前进行标准化处理,并剔除无变异特征。网络结构设置为两层全连接隐藏层,节点数分别为32和16,激活函数采用ReLU,并使用批量标准化和dropout降低过拟合风险;dropout比例设为0.15,优化器采用Adam,学习率设为0.001。模型最大训练轮数设为220轮,早停耐心值设为25,验证集比例设为0.15,批量大小根据训练样本量自动设置在16~128,随机种子设为20260619。DeepSurv输出个体风险评分,并与传统统计模型和其他机器学习生存模型进行预测性能比较。

1.5 模型评价

模型性能分别在内部测试集和时间队列验证集中评价。采用全局C-index评估模型对事件发生时序的整体区分效能,采用时间依赖AUC评估模型在30 d、90 d、1年、3年这几个固定随访截点的判别效能[17],并报告点估计值及95%*CI*;绘制时间队列验证集时间依赖ROC曲线以直观比较不同模型的泛化性能。校准度采用校准曲线评价,比较模型预测风险与实际观察风险的一致性;临床实用性采用决策曲线分析评价,通过不同阈值概率下的净获益比较模型相对于treat-all和treat-none策略的潜在临床价值。对于总体表现最佳的模型,进一步采用SHAP方法进行可解释性分析;SHAP结果仅用于解释模型预测依据,不用于因果推断。

1.6 统计学分析

所有分析采用R 4.5.3软件完成。连续变量根据分布情况以$M(P_{25},P_{75})$表示,分类变量以频数(%)表示;组间比较根据变量类型采用两独立样本t检验、Mann-Whitney U检验、χ^2检验或Fisher确切概率。所有统计检验为双侧检验,以P<0.05为差异具有统计学意义。

2 结果

2.1 研究对象基线特征

最终纳入老年男性AKI患者2 335例,其中随访期间死亡1 912例,存活423例。表2为全数据集研究对象基线特征有统计学差异的结果,总体队列中死亡组中位生存时间为190.00(49.00,558.00) d,存活组为1 312.00(914.00,2 066.00) d,2组差异有统计学意义(P<0.001)。死亡组患者年龄显著高于存活组[92.00(87.00,95.00)岁 *vs* 83.00(76.00,91.00)岁,P<0.001]。与存活组相比,死亡组在入院基线及AKI发生时均表现出较低的收缩压和舒张压水平(P<0.001),同时基线及AKI发生时BNP、BUN水平均明显升高(P<0.001),提示死亡患者存在更明显的心肾负荷增加及血流动力学不稳定。基于上述差异,后续进一步针对BNP、BUN及AKI后eGFR动态变化等关键连续指

标,评估其与死亡风险之间的非线性剂量-反应关系。

2.2 BNP、BUN及AKI后eGFR动态指标呈非线性关联,动态肾功能指标贡献增量信息

图2为模型调整年龄、主要合并症及机械通气等协变量后的RCS曲线。结果显示,6个指标与全因死亡风险均存在显著总体关联(P<0.001);除基线BNP的非线性检验P值为0.008外,其余指标的非线性检验P均<0.001,提示上述指标与死亡风险之间普遍存在非线性剂量-反应关系。基线BNP和AKI发生时BNP均与死亡风险呈非线性正相关;基线BUN和AKI发生时BUN同样与死亡风险呈非线性正相关;第7天eGFR缺损与死亡风险呈非线性正相关;48 h至7 d eGFR恢复与死亡风险呈非线性负相关。图2中实线为协变量校正后的HR,阴影区域为95%CI。水平虚线表示HR=1,垂直虚线表示模型的参考值。模型初始候选变量数分别为32、49和68个。Boruta算法分别确认15、16和22个。结合RSF重要性、VIF诊断后,最终分别保留20、23和21个建模变量。

图3展示了Boruta算法和RSF在3个候选变量集中的变量重要性分布。随着AKI发生时变量和动态肾功能指标逐步纳入,重要变量的组成由单纯基线临床特征逐渐扩展至AKI发生时状态及AKI后短期肾功能变化。在Variable set 3中,第7天eGFR缺损百分比、48 h至7 d eGFR恢复百分比、48 h至7 d SCr恢复百分比和7天SCr比值等动态肾功能指标进入重要变量序列,说明AKI后短期肾功能演变可提供区别于基线和AKI发生时变量的增量信息。

2.3 死亡风险生存预测模型构建性能比较:XGBoost-Cox模型总体表现较好

表3展示了不同变量集6种算法的内部测试集性能比较结果。总体来看,随着AKI发生时变量及AKI后动态肾功能指标的逐步纳入,各类模型的预测性能整体呈递增趋势。Variable set 1仅纳入基线变量时,各模型C-index为0.704~0.724;其中XGBoost-Cox表现最佳,C-index为0.724(95%CI:0.691~0.753),其30 d、90 d、1年和3年AUC分别为0.761、0.776、0.815和0.835。

Variable set 2加入AKI发生时变量后,C-index增至

表2 全数据集研究对象的基线特征比较[$M(P_{25},P_{75})$,例(%)]

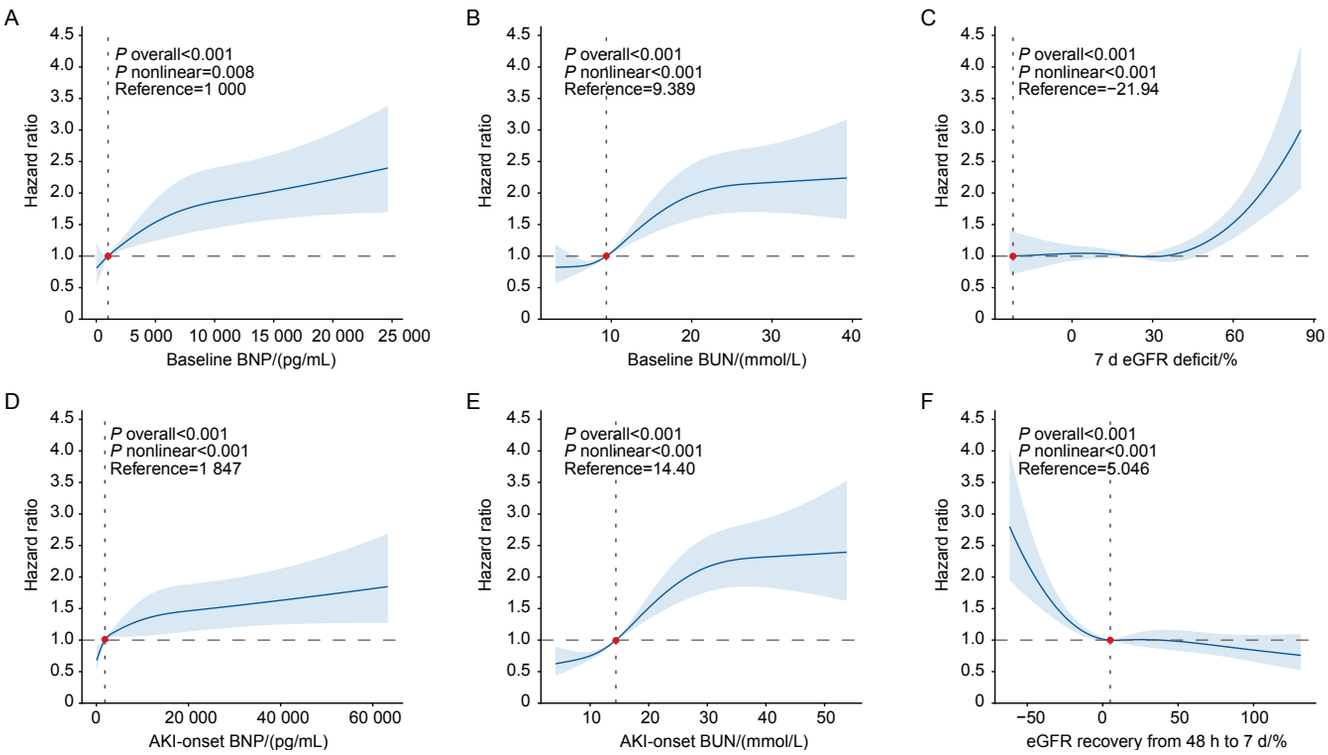
Table 2 Comparison of baseline characteristics of the study population [$M(P_{25},P_{75})$,n(%)]

Variable	Survivor group(n=423)	Non-survivor group (n=1 912)	χ^2/U	P
Demographics				
Survival day/d	1 312(914,2 066)	190(49, 558)	729 013.00	<0.001
Age/years	83(76, 91)	92(87, 95)	213 611.50	<0.001
Cardiovascular and hemodynamic status				
Baseline SBP/mmHg	135.00(124.00, 146.00)	130.00(119.00, 141.00)	455 399.00	<0.001
Baseline DBP/mmHg	71.00(64.00, 79.00)	66.00(60.00, 73.00)	506 882.00	<0.001
AKI-onset SBP/mmHg	127.00(116.00, 140.00)	120.00(108.00, 134.00)	478 260.50	<0.001
AKI-onset DBP/mmHg	68.00(61.00, 76.00)	62.00(55.00, 70.00)	516 037.00	<0.001
Baseline BNP/(pg/mL)	384.20(151.40, 1 174.90)	1 365.00(458.00, 4 040.70)	215 263.00	<0.001
AKI-onset BNP/(pg/mL)	726.00(231.30, 2 727.35)	2 402.00(795.00, 7 104.00)	231 152.50	<0.001
Renal function markers				
Last SCr within 48 h after AKI/(μmol/L)	132.00(96.00, 190.00)	145.00(103.50, 212.00)	372 525.50	0.011
Peak SCr within 7 d after AKI/(μmol/L)	146.00(107.00, 211.00)	164.15(117.00, 256.00)	365 897.00	0.002
Last SCr within 7 d after AKI/(μmol/L)	115.00(83.00, 169.00)	131.00(89.55, 205.00)	360 435.00	<0.001
Baseline eGFR/[mL/(min•1.73 m^2)]	71.40(46.75, 86.53)	63.63(41.31, 81.33)	444 740.50	0.001
Peak eGFR within 48 h after AKI/[mL/(min•1.73 m^2)]	39.04(24.39, 57.07)	33.76(20.82, 49.57)	439 743.00	0.005
Last eGFR within 48 h after AKI/[mL/(min•1.73 m^2)]	42.84(26.01, 62.27)	36.08(22.65, 54.43)	455 918.50	<0.001
Peak eGFR within 7 d after AKI/[mL/(min•1.73 m^2)]	38.60(23.64, 55.93)	30.87(18.05, 47.21)	459 843.50	<0.001
Last eGFR within 7 d after AKI/[mL/(min•1.73 m^2)]	51.00(31.65, 73.21)	40.73(23.58, 65.03)	468 303.00	<0.001
Baseline BUN/(mmol/L)	7.10(5.30, 9.60)	10.20(7.10, 15.30)	247 331.00	<0.001
AKI-onset BUN/(mmol/L)	11.20(7.90, 15.60)	16.10(11.10, 23.50)	251 807.50	<0.001

续表2 全数据集研究对象的基线特征比较[$M(P_{25},P_{75})$,例(%)]

Table 2(continued) Comparison of baseline characteristics of the study population [$M(P_{25},P_{75})$,n(%)]

Variable	Survivor group(n=423)	Non-survivor group (n=1 912)	χ^2/U	P
AKI severity and renal dynamic indices				
AKI stage			15.70	<0.001
Stage 1	316(74.70)	1 395(72.96)		
Stage 2	40(9.46)	302(15.79)		
Stage 3	67(15.84)	215(11.24)		
48 h eGFR ratio/%	1.46(1.21, 1.80)	1.54(1.27, 1.87)	366 028.00	0.002
Early eGFR drop/%	31.70(17.26, 44.36)	35.10(21.15, 46.41)	366 031.50	0.002
7 d SCr ratio/%	1.62(1.52, 1.92)	1.70(1.52, 2.18)	363 351.00	<0.001
7 d eGFR ratio/%	1.28(1.04, 1.65)	1.36(1.10, 1.89)	354 112.00	<0.001
7 d SCr excess/%	42.42(5.65, 62.86)	44.77(12.26, 80.00)	359 110.50	0.001
7 d eGFR deficit/%	21.83(3.54, 39.44)	26.32(8.86, 46.96)	354 111.00	<0.001
SCr recovery from 48 h to 7 d/%	11.27(0.00, 34.12)	9.88(−3.99, 28.47)	458 719.00	<0.001
eGFR recovery from 48 h to 7 d /%	0.00(0.00, 23.24)	2.96(−7.04, 27.03)	430 129.00	0.039
Hematology				
Baseline Hb/(g/L)	125.00(112.00, 137.00)	109.00(96.00, 122.00)	571 714.00	<0.001
Baseline RBC/($\times10^{12}$/L)	3.99(3.58, 4.40)	3.49(3.05, 3.96)	562 296.50	<0.001
Baseline PLT/($\times10^{9}$/L)	177.00(140.00, 219.00)	165.00(119.00, 217.00)	440 431.50	<0.001
AKI-onset Hb/(g/L)	121.00(108.00, 136.00)	106.00(92.00, 120.00)	574 583.50	<0.001
AKI-onset RBC/($\times10^{12}$/L)	3.90(3.48, 4.34)	3.41(2.95, 3.91)	567 795.00	<0.001
AKI-onset PLT/($\times10^{9}$/L)	171.00(137.00, 220.00)	153.00(106.00, 206.00)	473 294.00	<0.001
Biochemistry, nutrition, and metabolism				
Baseline albumin/(g/L)	39.00(35.70, 43.00)	35.10(31.90, 38.40)	566 466.00	<0.001
AKI-onset albumin/(g/L)	37.70(34.50, 41.50)	33.70(30.50, 37.00)	582 276.50	<0.001
Baseline glucose/(mmol/L)	6.06(5.10, 7.47)	6.64(5.40, 8.49)	337 769.00	<0.001
AKI-onset glucose/(mmol/L)	6.56(5.41, 8.37)	7.37(5.78, 9.87)	332 747.50	<0.001
Baseline uric acid/(μmol/L)	299.00(219.00, 363.00)	306.00(223.00, 398.00)	365 624.00	0.003
AKI-onset uric acid/(μmol/L)	353.00(283.00, 421.00)	372.50(288.00, 475.00)	357 935.00	<0.001
Electrolytes and minerals				
Baseline K/(mmol/L)	4.03(3.72, 4.36)	4.10(3.80, 4.44)	372 601.00	0.018
Baseline Na/(mmol/L)	138.00(135.00, 141.00)	138.00(135.00, 143.00)	365 825.50	0.004
AKI-onset Na/(mmol/L)	138.00(134.90, 141.00)	140.00(135.00, 145.00)	326 406.00	<0.001
AKI-onset Ca/(mmol/L)	2.27(2.13, 2.39)	2.23(2.11, 2.37)	432 225.00	0.017
AKI-onset Cl/(mmol/L)	101.60(98.00, 105.10)	102.00(98.00, 107.00)	371 815.50	0.014
Medical history, comorbidities, and treatment				
Smoking history	124(29.31)	721(37.71)	10.57	0.001
Hyperuricemia	130(30.73)	924(48.33)	43.29	<0.001
CHD	213(50.35)	1 362(71.23)	68.78	<0.001
MI	64(15.13)	411(21.50)	8.66	0.003
AF	54(12.77)	532(27.82)	41.78	<0.001
CHF	96(22.70)	943(49.32)	99.42	<0.001
Stroke	86(20.33)	846(44.25)	82.61	<0.001
Malignancy	94(22.22)	657(34.36)	23.40	<0.001
COPD	50(11.82)	547(28.61)	51.30	<0.001
CKD	136(32.15)	880(46.03)	27.12	<0.001
MV	29(6.86)	711(37.19)	147.19	<0.001



A: Baseline BNP; B: Baseline BUN; C: 7 d eGFR deficit; D: AKI-onset BNP; E: AKI-onset BUN; F: eGFR recovery from 48 h to 7 d

图2 6个关键连续预测因子与全因死亡风险的限制性立方样条曲线

Figure 2 Restricted cubic spline curves for 6 key continuous predictors associated with all-cause mortality

0.727~0.741。其中XGBoost-Cox仍表现最佳,C-index为0.741(95%*CI*:0.713~0.766),1年AUC为0.830,3年AUC为0.847。RSF的表现也较好,C-index为0.734(95%*CI*:0.708~0.760),30 d、90 d和1年AUC分别达到0.793、0.806和0.828。上述结果提示,AKI发生时的实验室指标和临床状态能够提升死亡风险预测能力。

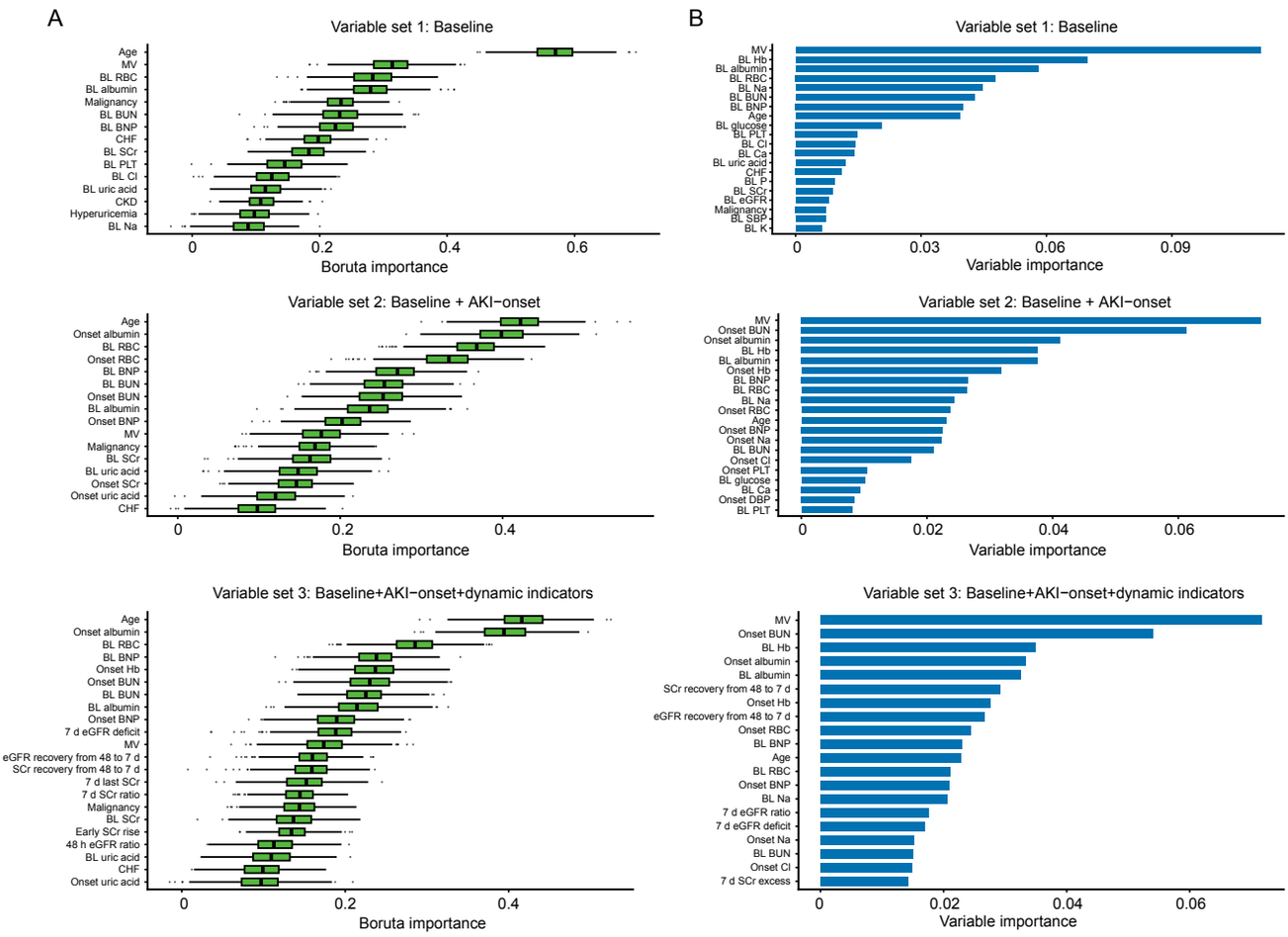
Variable set 3纳入AKI后48 h至7 d的动态肾功能指标后,模型整体性能进一步改善。XGBoost-Cox获得最高C-index,为0.749(95%*CI*:0.722~0.774),其30 d、90 d、1年和3年AUC分别为0.830、0.815、0.837和0.846。RSF在短期预测中表现最突出,30 d AUC为0.840(95%*CI*:0.785~0.888),90 d AUC为0.829(95%*CI*:0.785~0.868),1年AUC为0.837(95%*CI*:0.788~0.878),均为或并列为相应时间点的最高值。Elastic-net Cox和DeepSurv在Variable set 3中也保持较稳定表现,C-index分别为0.735和0.741,但整体未超过XGBoost-Cox和RSF。

2.4 时间队列验证XGBoost-Cox模型泛化性能最佳

采用时间验证队列评估最终变量集下6种模型在30 d、90 d、1年和3年死亡风险预测中的泛化性能,结果见图4。时间验证集中,各模型在不同预测时间点均表现出较好的区分能力,AUC总体为0.733~0.822。与内部测试集结果一致,XGBoost-Cox在4个时间点均取得最高AUC,分别为0.797(95%*CI*:0.763~0.830)、0.781(95%*CI*:0.755~0.809)、0.802(95%*CI*:0.778~0.824)和0.822(95%*CI*:0.788~0.853)。具体比较显示,RSF在30和90 d预测中次于XGBoost-Cox;1年和3年预测中Elastic-net Cox、DeepSurv及RCS-Cox亦保持较好区分度。总体而言,XGBoost-Cox在短期、中期和长期预测中均表现最稳定。时间验证结果表明,基于最终变量集构建的模型具有较好泛化能力,其中XGBoost-Cox表现最佳,因此作为后续风险分层、校准、决策曲线和SHAP分析的主要模型。

2.5 最佳模型具有临床净获益,肾功能负荷主导3年死亡风险预测

图5展示了最佳模型在3年死亡风险预测中的校准表现、决策曲线分析及SHAP可解释性结果。XGBoost-Cox模型3年预测风险与观察风险总体一致,但高风险区间仍有一定偏离(图5A)。决策曲线显示,在较宽阈值范围内,多数模型净获益高于treat-all和treat-none策略,提示其具有潜在临床应用价值(图5B)。



A: Results of the Boruta feature selection with green boxes denoting confirmed variables; B: Variable importance ranking from the RSF, only the top-ranked variables within each set are shown. BL indicates the baseline measurement of the indicator, and Onset indicates the measurement of the indicator at AKI onset

图3 3个候选预测变量集中的Boruta算法和RSF变量重要性分布

Figure 3 Variable importance distributions from the Boruta algorithm and random survival forest (RSF) across 3 candidate predictor sets

基于XGBoost-Cox的SHAP分析显示,AKI发生时BUN、机械通气、AKI发生时白蛋白、恶性肿瘤、基线白蛋白、基线BNP、基线血红蛋白、年龄、AKI发生时BNP、AKI发生时氯离子、基线BUN、AKI后48 h至7 d eGFR恢复比值、基线尿酸、7 d eGFR缺损和AKI发生时钠离子等变量是影响3年死亡风险预测的主要因素(图5C)。上述结果提示肾功能负荷、营养状态和合并症负担在老年AKI患者死亡风险预测中具有重要作用。SHAP summary plot进一步展示了各变量取值与模型预测方向之间的关系(图5D)。

个体化SHAP解释显示,高风险个体的预测风险为0.934,高于总体平均预测水平(0.843)。该个体风险升高主要来自AKI发生时BUN明显升高(23.8 mmol/L, SHAP值+0.46)与AKI发生时BNP明显升高(5 040 pg/mL, +0.16)。降低死亡风险的特征中幅度最大的是较低水平的基线BNP(195 pg/mL, −0.24)、未机械通气(−0.16)与无恶性肿瘤(−0.15)。低风险个体的预测风险为0.141,明显低于总体平均水平,其风险降低的前5个特征主要是AKI发生时白蛋白较高(41.5 g/L, −0.53)、年龄相对较低(74岁, −0.50)、基线白蛋白较高(46.3 g/L, −0.32)、AKI发生时BUN较低(8.2 mmol/L, −0.31)、基线BNP较低(140 pg/mL, −0.25),见图5E、F。上述结果表明,SHAP方法能够从全局和个体层面解释模型预测依据,有助于提高机器学习生存模型在老年AKI死亡风险预测中的透明度和临床可解释性。

表3 不同变量集6种算法的内部测试集性能比较

Table 3 Performance comparison of 6 survival prediction models

Model	C-index	AUC-30 d	AUC-90 d	AUC-1 year	AUC-3 years
Variable set 1					
Cox	0.708(0.676—0.737)	0.738(0.661—0.803)	0.750(0.691—0.798)	0.791(0.744—0.836)	0.828(0.777—0.873)
RCS-Cox	0.704(0.671—0.734)	0.716(0.634—0.790)	0.741(0.681—0.792)	0.793(0.744—0.838)	0.828(0.780—0.873)
Elastic-net Cox	0.711(0.680—0.740)	0.747(0.672—0.809)	0.754(0.696—0.802)	0.793(0.745—0.838)	0.836(0.789—0.882)
RSF	0.707(0.678—0.735)	0.764(0.694—0.829)	0.768(0.713—0.815)	0.800(0.752—0.844)	0.813(0.767—0.863)
XGBoost-Cox	0.724(0.691—0.753)	0.761(0.690—0.826)	0.776(0.718—0.822)	0.815(0.770—0.858)	0.835(0.788—0.881)
DeepSurv	0.718(0.688—0.745)	0.769(0.703—0.830)	0.773(0.717—0.822)	0.802(0.752—0.848)	0.837(0.790—0.880)
Variable set 2					
Cox	0.727(0.699—0.754)	0.768(0.699—0.828)	0.783(0.728—0.829)	0.809(0.762—0.854)	0.842(0.792—0.887)
RCS-Cox	0.727(0.698—0.754)	0.755(0.678—0.821)	0.782(0.726—0.830)	0.812(0.764—0.856)	0.845(0.798—0.888)
Elastic-net Cox	0.728(0.701—0.755)	0.771(0.701—0.829)	0.784(0.730—0.829)	0.812(0.764—0.855)	0.843(0.793—0.888)
RSF	0.734(0.708—0.760)	0.793(0.727—0.850)	0.806(0.756—0.845)	0.828(0.783—0.870)	0.839(0.791—0.884)
XGBoost-Cox	0.741(0.713—0.766)	0.791(0.721—0.851)	0.801(0.749—0.844)	0.830(0.785—0.873)	0.847(0.798—0.890)
DeepSurv	0.729(0.701—0.755)	0.754(0.679—0.816)	0.783(0.731—0.827)	0.821(0.774—0.866)	0.844(0.794—0.891)
Variable set 3					
Cox	0.734(0.706—0.762)	0.796(0.731—0.857)	0.798(0.743—0.844)	0.816(0.772—0.858)	0.836(0.787—0.881)
RCS-Cox	0.734(0.705—0.762)	0.785(0.711—0.850)	0.797(0.745—0.845)	0.817(0.774—0.861)	0.837(0.792—0.883)
Elastic-net Cox	0.735(0.708—0.763)	0.805(0.741—0.860)	0.798(0.745—0.844)	0.819(0.775—0.860)	0.837(0.790—0.884)
RSF	0.744(0.717—0.771)	0.840(0.785—0.888)	0.829(0.785—0.868)	0.837(0.788—0.878)	0.831(0.780—0.877)
XGBoost-Cox	0.749(0.722—0.774)	0.830(0.767—0.886)	0.815(0.767—0.858)	0.837(0.794—0.878)	0.846(0.802—0.891)
DeepSurv	0.741(0.711—0.770)	0.821(0.752—0.882)	0.799(0.747—0.848)	0.819(0.773—0.862)	0.838(0.790—0.882)

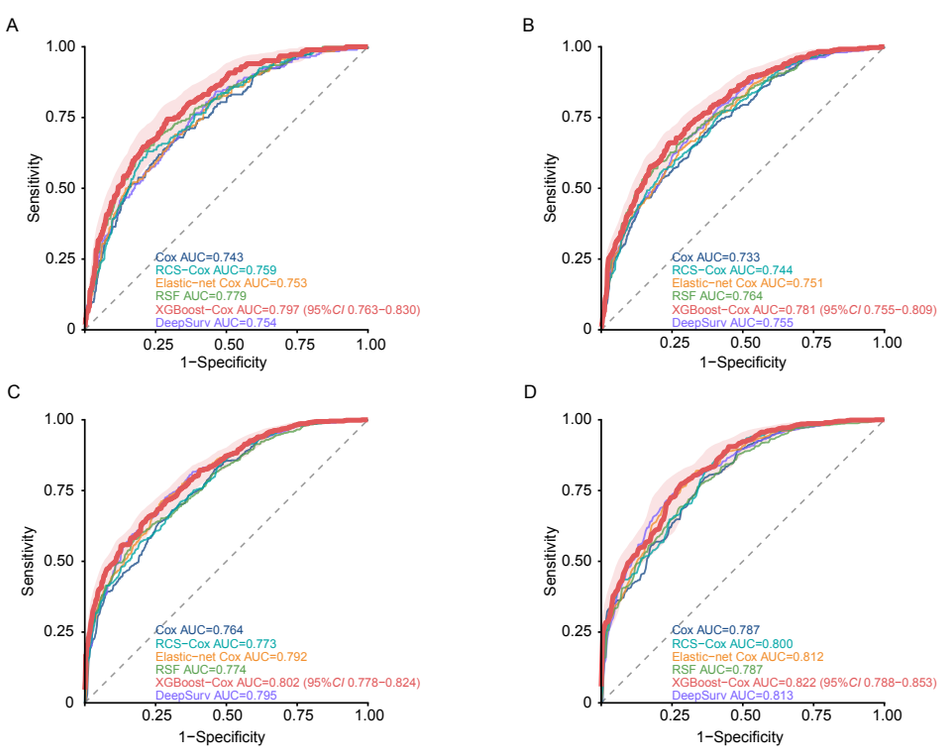
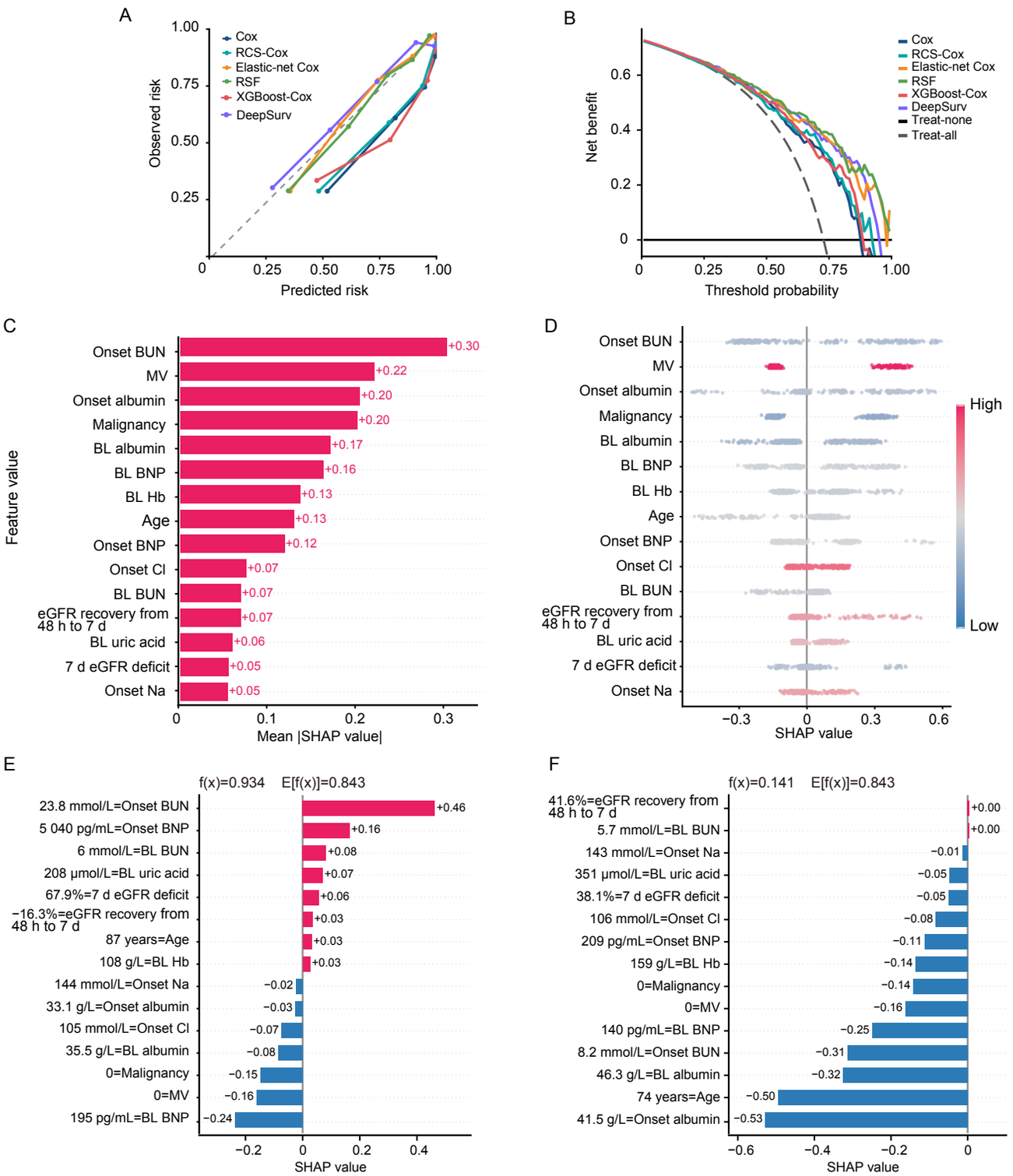


图4 时间验证队列中预测30 d(A)、90 d(B)、1年(C)和3年(D)全因死亡的时间依赖ROC曲线

Figure 4 Time-dependent ROC curves for predicting all-cause mortality at 30 (A) and 90 d (B), and 1 (C) and 3 years (D) in the temporal validation cohort



A: Calibration curves of 6 models for 3-year mortality prediction with the gray dashed line representing perfect calibration; B: Decision curve analysis of 6 models for 3-year mortality prediction; C: SHAP feature importance ranking of the XGBoost-Cox model based on mean absolute SHAP values; D: SHAP summary plot of the XGBoost-Cox model in the test set; E and F: Local SHAP explanations for representative high-risk and low-risk individuals with red bars indicating variables increasing the predicted risk and blue bars indicating variables decreasing the predicted risk. BL indicates the baseline measurement of the indicator, and Onset indicates the measurement of the indicator at AKI onset

图5 XGBoost-Cox模型的校准、决策曲线及SHAP可解释性分析

Figure 5 Calibration, decision curve analysis, and SHAP-based interpretability analysis of the XGBoost-Cox model

3 讨论

本研究构建并验证了融合非线性特征工程与可解释机器学习生存模型的老年男性AKI死亡风险预测框架。主要发现为:BNP、BUN及AKI后短期eGFR动态变化与死亡风险存在非线性关系;死亡风险受肾功能负荷、营养状态、心血管功能、血液学指标、合并症和器官支持治疗等多维因素共同影响,研究中BUN、白蛋白、机械通气、年龄及恶性肿瘤等变量的重要性与既往发现基本一致[18];纳入AKI后动态肾功能指标可改善预测性能;XGBoost-Cox在内部测试和时间验证中总体表现最佳,并具有较好的校准度、临床净获益和可解释性。

3.1 非线性特征工程:筛选BNP、BUN及AKI后eGFR动态变化为死亡风险的关键非线性预测因子

既往AKI预后研究多依赖Cox回归、Logistic回归或LASSO等方法进行变量筛选和风险建模。这类方法具有较好的可解释性,但通常更强调变量的线性主效应,对连续变量的阈值效应、平台效应及非线性风险变化识别不足。本研究在特征工程阶段引入RCS进行非线性关系识别,并结合Boruta算法、RSF变量重要性和临床先验筛选预测因子,使变量选择既考虑统计稳定性,也兼顾临床可解释性。研究发现BNP、BUN及AKI后短期eGFR动态变化与死亡风险的非线性关系提示,心肾交互、代谢负荷和肾功能恢复过程均与预后密切相关。具体来看,基线BNP和AKI发生时BNP均与死亡风险呈非线性正相关。随着BNP水平升高,死亡风险总体呈上升趋势,且在高值区间置信区间增宽,提示极端高值范围内风险估计的不确定性增加。基线BUN和AKI发生时BUN同样与死亡风险呈非线性正相关,死亡风险随BUN水平升高而增加,并在较高水平后趋于平台。肾功能动态变化指标方面,第7天eGFR缺损与死亡风险呈非线性正相关,尤其在缺损程度较高时风险上升更为明显;48 h至7 d eGFR恢复百分比与死亡风险呈非线性负相关,eGFR恢复不足或持续下降者死亡风险较高,而随着eGFR恢复程度增加,死亡风险逐渐降低。因此,临床上可结合AKI发生后48 h至7 d eGFR恢复情况、第7天肾功能缺损程度,动态识别肾功能持续受损或恢复不良的高危患者,加强肾功能监测、病因及可逆危险因素评估,并及时采取个体化管理措施。表2中48 h至7 d eGFR恢复百分比的未调整分布与RCS结果存在一定差异,可能受检测时间窗和复测频率影响,因此本研究未将其单因素组间差异作为主要解释依据,而是结合RCS、多变量模型综合判断其预后价值。上述结果证实,BNP、BUN及AKI后eGFR动态变化不仅与死亡风险显著相关,而且其风险效应并非简单线性,支持后续预测模型保留这些连续变量并采用能够捕捉非线性关系的建模方法。

3.2 多变量集生存预测模型比较证实:纳入AKI后动态肾功能指标的XGBoost-Cox模型预测性能最优

与传统二分类模型相比,生存模型可同时利用死亡时间和删失信息,更适配真实世界随访数据。本研究系统比较6种生存预测模型,结果显示:纳入AKI后动态肾功能指标后,机器学习生存模型整体表现优于传统Cox模型;其中XGBoost-Cox在内部测试和时间验证中表现稳定,在3个变量集中均获得最高C-index,总体区分能力最稳定;RSF在Variable set 3中对30 d、90 d和1年死亡风险的时间依赖性AUC最高或接近最高,对短期和中期死亡风险预测能力较强。XGBoost-Cox可在Cox部分似然框架下学习变量的非线性关系和交互作用;Boruta和RSF比单纯LASSO类筛选方法更适合识别非线性贡献较强、在线性模型中可能被低估的变量。RCS-Cox可处理部分连续变量的非线性关系,但本研究仅对预设关键变量引入样条项,性能提升有限,提示单纯对少数变量做样条扩展,不足以完全刻画老年AKI患者复杂的风险结构。

既往研究多聚焦ICU-AKI或脓毒症相关AKI的短期/院内死亡预测,如梁晶晶等[19]通过LightGBM预测SA-AKI 28 d死亡测试集AUC为0.817,Liu等[20]构建的ICU-AKI死亡风险stacking模型AUC为0.91。本研究不以急性期早期死亡预警为目标,限定于AKI发病7 d后仍存活的老年男性患者的中长期风险分层,采用生存模型处理随访时间和删失信息,纳入AKI后eGFR动态恢复指标,改进之处在于适用场景、动态再评估和模型可解释性。本研究最终选择XGBoost-Cox作为解释模型,开展校准曲线、决策曲线和SHAP分析:校准结果提示模型预测风险与观察风险一致性较好,决策曲线提示模型在一定阈值范围内存在潜在临床净获益,SHAP结果显示模型核心依赖变量具备明确临床含义[18]。同时,SHAP解释反映模型层面的关联贡献而非因果效应[21],但可帮助临床医生理解模型风险判定,提高机器学习模型的透明度和可接受性。

3.3 老年AKI死亡风险由多维驱动因素

研究结果显示,老年AKI死亡风险由肾功能损伤、心血管负荷、营养炎症状态、基础疾病和治疗支持等因素共同影响。SHAP分析提示,AKI发生时BUN、

机械通气、AKI发生时白蛋白、恶性肿瘤、基线白蛋白、基线BNP、基线血红蛋白和年龄等变量是影响模型预测的重要因素。这些变量分别反映肾功能负荷、营养状态、合并症负担、心血管功能、贫血及年龄相关脆弱性。该结果说明,老年AKI患者的预后不能仅依赖年龄、血清肌酐或尿量等单一指标判断,而应综合评估患者的全身状态和器官储备。

与单一时间点指标相比,AKI后短期肾功能变化更能反映疾病演变过程。对于老年患者而言,AKI后eGFR能否恢复,不仅取决于肾脏本身损伤程度,还与循环状态、感染炎症负荷、营养状态、用药暴露及多器官功能储备有关。因此,将动态肾功能指标纳入预测模型,有助于从"发生AKI时的风险评估"进一步过渡到"AKI后恢复过程中的风险再评估",更符合真实临床决策场景[1,22-23]。

本研究的创新性主要体现在以下方面。第一,以非线性关系识别和稳定变量筛选为核心开展特征工程,结合RCS、Boruta算法和RSF提高变量构建及筛选的稳健性。第二,基于变量可获得时间构建递进式变量集,分别评估基线变量、AKI发生时变量和AKI后短期动态肾功能指标的增量预测价值,并将动态变量集定位为AKI后7 d仍存活患者的条件风险再评估场景。第三,系统比较传统统计模型、惩罚回归模型、树模型、梯度提升模型和深度学习生存模型,并结合时间外部验证、校准曲线、决策曲线和SHAP分析评价模型表现及可解释性,形成兼顾预测性能、临床实用性和透明度的智能医学分析框架[24]。本研究也存在一定局限性。本研究为单中心回顾性设计,AKI后动态肾功能指标依赖临床复测记录,可能受到住院时间、病情严重程度和检测频率影响;同时时间验证队列中有11例患者因缺乏有效随访信息被排除,占初始研究人群的0.66%(11/1 666),虽失访比例较低,但现有资料无法证实其为随机失访,仍不能完全排除失访相关选择偏倚。此外,本研究仅纳入男性患者,研究时段内符合纳入标准的女性病例样本量有限,模型结论暂不能推广至女性AKI人群,后续需纳入足量女性样本开展独立验证。未来研究需在前瞻性队列中采用统一复测时间点进一步验证。

综上,融合非线性特征工程与可解释机器学习生存模型的预测框架能够较好识别老年男性AKI患者死亡风险。XGBoost-Cox模型在区分度、时间验证表现和可解释性方面具有一定优势。需要注意的是,本模型仅用于AKI发病7 d后仍存活的老年男性患者的中长期死亡风险分层,不适用于AKI发病1周内危重死亡高危人群。AKI后短期肾功能恢复指标虽具有潜在预测价值,但仍需通过多中心外部验证以及前瞻性研究,进一步提升模型在老年男性AKI个体化风险评估和临床决策支持中的应用价值。

利益冲突 本文所有作者声明不存在利益冲突

参考文献:

[1] Ostermann M, Lumlertgul N, Jeong R, *et al*. Acute kidney injury [J]. Lancet, 2025, 405 (10474): 241-256. DOI: 10.1016/S0140-6736(24)02385-7.

[2] Ronco C, Bellomo R, Kellum J A. Acute kidney injury [J]. Lancet, 2019, 394 (10212): 1949-1964. DOI: 10.1016/S0140-6736(19)32563-2.

[3] Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group. KDIGO clinical practice guideline for acute kidney injury [R]. 2012:1-138.

[4] Grant C H, Dahiya A, Palechuk T, *et al*. Health inequalities and outcomes following acute kidney injury: a systematic review & meta-analyses of observational studies [J]. BMC Nephrol, 2025, 26 (1): 494. DOI: 10.1186/s12882-025-04391-x.

[5] Harrell F E Jr. Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis [M]. Cham: Springer International Publishing, 2015.

[6] Wynants L, van Smeden M, McLernon D J, *et al*. Three myths about risk thresholds for prediction models [J]. BMC Med, 2019, 17(1): 192. DOI:10.1186/s12916-019-1425-3.

[7] Collins G S, Moons K G M, Dhiman P, *et al*. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods [J]. BMJ, 2024, 385: e078378. DOI:10.1136/bmj-2023-078378.

[8] Kursa M B, Rudnicki W R. Feature selection with the boruta package [J]. J Stat Softw, 2010, 036 (i11): 1-13. DOI: 10.18637/jss.v036.i11.

[9] 伍亚舟, 陈锡程, 易东. 人工智能在临床领域的研究进展及前景展望[J]. 第三军医大学学报, 2022, 44(1): 89-102. DOI:10.16016/j.1000-5404.202112092.
Wu Y Z, Chen X C, Yi D. Advances and perspective of artificial intelligence in clinical area [J]. J Third Mil Med Univ, 2022, 44 (1): 89-102. DOI: 10.16016/j. 1000-5404. 202112092.

[10] Haug C J, Drazen J M. Artificial intelligence and machine learning in clinical medicine, 2023 [J]. N Engl J Med, 2023, 388(13): 1201-1208. DOI:10.1056/NEJMra2302038.

[11] Rajpurkar P, Chen E, Banerjee O, *et al*. AI in health and medicine[J]. Nat Med, 2022, 28(1): 31-38. DOI:10.1038/

s41591-021-01614-0.

[12] Ishwaran H, Kogalur U B, Blackstone E H, *et al*. Random survival forests[J]. Ann Appl Stat, 2008, 2(3): 841-860. DOI:10.1214/08-aoas169.

[13] Chen T Q, Guestrin C. XGBoost: a scalable tree boosting system [C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016: 785-794. DOI:10.1145/2939672.2939785.

[14] Katzman J L, Shaham U, Cloninger A, *et al*. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network[J]. BMC Med Res Methodol, 2018, 18(1): 24. DOI: 10.1186/s12874-018-0482-1.

[15] Van Calster B, McLernon D J, van Smeden M, *et al*. Calibration: the Achilles heel of predictive analytics[J]. BMC Med, 2019, 17(1): 230. DOI: 10.1186/s12916-019-1466-7.

[16] Vickers A J, Elkin E B. Decision curve analysis: a novel method for evaluating prediction models[J]. Med Decis Making, 2006, 26(6): 565-574. DOI:10.1177/0272989X06295361.

[17] Heagerty P J, Lumley T, Pepe M S. Time-dependent ROC curves for censored survival data and a diagnostic marker [J]. Biometrics, 2000, 56(2): 337-344. DOI: 10.1111/j.0006-341x.2000.00337.x.

[18] Jiao Y R, Wu Z, Zhang Y B, *et al*. Explainable machine learning-based 28-day mortality prediction model for elderly patients with acute kidney injury[J]. BMC Nephrol, 2026, 27(1): 157. DOI:10.1186/s12882-026-04792-6.

[19] 梁晶晶, 肖回宇, 屈壕, 等. LightGBM预测脓毒症相关急性肾损伤患者28天死亡风险性能最优: 基于MIMIC-Ⅳ与eICU数据库构建可解释性机器学习模型及验证对比[J]. 陆军军医大学学报, 2026, 48(15): 2129-2138. DOI: 10.16016/j.2097-0927.202605041.
Liang J J, Xiao H Y, Qu H, *et al*. LightGBM demonstrates optimal performance in predicting 28-day mortality risk in patients with sepsis-associated acute kidney injury: development and validation of an interpretable machine learning model based on MIMIC-Ⅳ and eICU databases[J]. J Army Med Univ, 2026, 48(15): 2129-2138. DOI: 10.16016/j.2097-0927.202605041.

[20] Liu M Q, Fan Z P, Gao Y, *et al*. A two-tier feature selection method for predicting mortality risk in ICU patients with acute kidney injury[J]. Sci Rep, 2024, 14(1): 16794. DOI:10.1038/s41598-024-63793-3.

[21] Lundberg S M, 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. DOI: 10.1038/s42256-019-0138-9.

[22] Zarbock A, Nadim M K, Pickkers P, *et al*. Sepsis-associated acute kidney injury: consensus report of the 28th Acute Disease Quality Initiative workgroup[J]. Nat Rev Nephrol, 2023, 19(6): 401-417. DOI: 10.1038/s41581-023-00683-3.

[23] Chawla L S, Bellomo R, Bihorac A, *et al*. Acute kidney disease and renal recovery: consensus report of the Acute Disease Quality Initiative (ADQI) 16 Workgroup[J]. Nat Rev Nephrol, 2017, 13(4): 241-257. DOI: 10.1038/nrneph.2017.2.

[24] Steyerberg E W, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation[J]. Eur Heart J, 2014, 35(29): 1925-1931. DOI:10.1093/eurheartj/ehu207.

(**收稿**:2026-07-09;**修回**:2026-09-10)

(编辑 邓强庭)