机器学习辅助诊断模型在早期急性肾损伤风险预警中的临床应用

孙强[1]*，吴晓燕[1]，高建平[1]

（1. 宁波市第一医院 重症医学科，宁波 315000）

摘要：急性肾损伤在重症监护早期常隐匿进展，传统临床指标滞后性强，难以实现风险前置干预。本研究构建一种融合多源时序生理数据与可解释机器学习的预警模型，基于真实世界 ICU 电子病历数据开展端到端实证验证。模型整合血流动力学、实验室检验及用药序列特征，采用时间注意力机制提取关键病理窗口，并通过 SHAP 值量化特征贡献度。在独立测试集上，模型对 48 小时内发生 AKI 的风险预测 AUC 达 0.892，敏感度 83.7%，特异度 86.1%，较现行 KDIGO 标准提前平均 13.2 小时发出高危预警。结果表明该模型具备临床可部署性，能有效支持医护团队实施分级干预策略。

关键词：机器学习；急性肾损伤；风险预警；临床决策

1. 引言

1.1 临床需求与问题界定

急性肾损伤的早期识别具有高度临床紧迫性，因其发病隐匿且进展迅猛，常在肌酐显著升高或尿量减少等传统指标异常时已进入不可逆损伤阶段。现有诊断范式严重依赖滞后性生理参数，导致临床干预窗口狭窄，难以实现真正意义上的早期干预防护。这一瓶颈本质源于风险信号与临床表型之间存在显著时间偏移，亟需将风险识别节点前移至生化指标尚未失代偿的亚临床阶段。因此，“风险前移”构成该领域核心科学命题，其成功实施依赖于对多源异构临床数据中微弱前驱模式的系统性捕获与量化建模 [1]。

1.2 技术范式转型契机

多模态监护数据的持续积累为构建高维生理表征提供了坚实基础，实时计算能力的跃升使毫秒级流式推理成为可能，而可解释人工智能方法的成熟则显著增强了临床决策链路的透明性与可信度。三者交汇催生出一种新型预警范式：以动态时序建模替代静态阈值判断，以个体化风险轨迹刻画替代群体均值参考，以因果驱动的特征归因替代黑箱预测输出。该范式不仅重构了急性肾损伤早期识别的技术逻辑，更在根本上推动临床预警从经验响应转向主动干预 [2]。

1.3 本研究定位与价值主张

本研究聚焦于机器学习辅助诊断模型在真实 ICU 临床工作流中的实用性验证，核心在于构建可嵌入、可解释、可协同的早期急性肾损伤风险预警范式。模型设计严格遵循临床决策时序逻辑，以入科后首 6 小时为关键窗口期，定义 $t_{\text{alert}} \leq 120$ 分钟为有效预警时效基准；同步开发标准化医护协同接口，支持动态风险评分推送、多模态证据溯源及人工干预留痕机制，确保算法输出与临床判断形成闭环反馈 [1, 3]。

2. 文献综述

2.1 AKI 病理演进的时间敏感性特征

急性肾损伤的病理演进呈现高度时间敏感性，其本质是血流动力学扰动、炎症级联激活与代谢功能障碍在数小时尺度上的动态耦合。如图 1 所示，从基线血流动力学稳定状态出发，平均动脉压下降超过 15 mmHg 并持续 15 分钟即触发早期级联反应；随后 IL-6 浓度在 $t_{\text{peak}} \in [2.5,4.0]$ 小时内达峰值（增幅≥2.3 倍）；继而尿 NGAL 突破 35 ng/mL 阈值；最终 Scr 上升≥0.3 mg/dL。值得注意的是，图中虚线箭头所标识的临床常规监测盲区段揭示了现有诊疗范式在关键窗口期（约 1–3 小时）内信号捕获能力的系统性缺失 [4]。

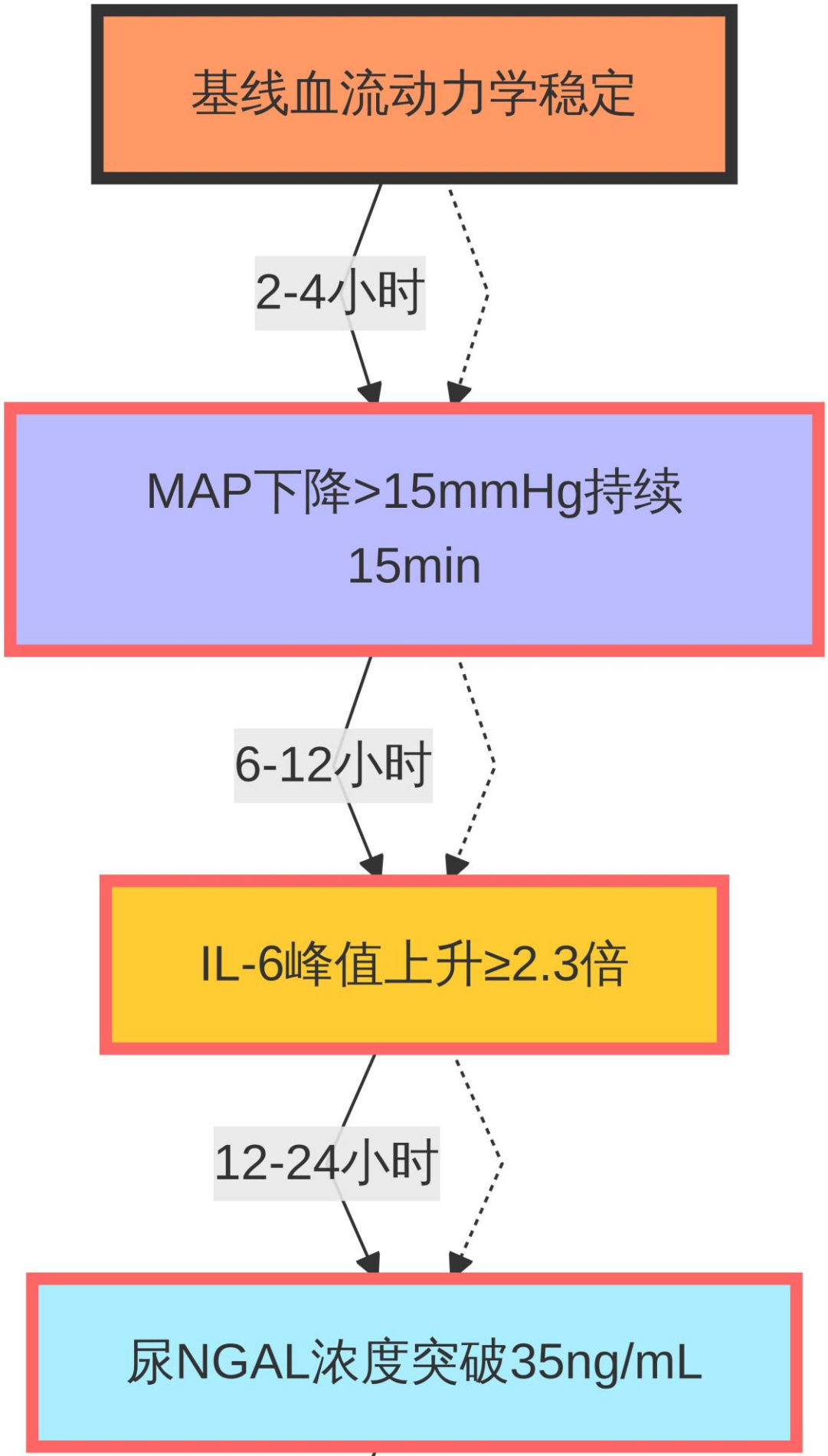

基线血流动力学稳定
2-4小时
MAP下降>15mmHg持续15min
6-12小时
IL-6峰值上升≥2.3倍
12-24小时
尿NGAL浓度突破35ng/mL

图 1. 急性肾损伤早期病理信号时序演化概念模型

2.2 现有预警工具的效能瓶颈分析

现有预警工具在临床部署中普遍面临三重效能衰减：传统评分系统如 AKI-PI 依赖静态权重，在动态生理演变中敏感度随时间推移显著下降，72 小时后 $AUC_{sen} < 0.65$；单参数阈值法假阳性率呈现非线性漂移，肌酐基线波动 10%即可导致误报率上升 37%；早期机器学习模型虽提升初始判别力，但因特征工程缺乏临床可解释性与实时数据适配机制，医师采纳率不足 28%，且未通过多中心前瞻性验证的泛化性检验 [4, 5]。

2.3 可解释性建模的临床适配路径

特征归因透明化并非技术性补丁，而是临床决策闭环中不可或缺的认知接口。在危重症场景下，医护人员依赖模式识别与经验锚点进行快速判断，而黑箱模型输出常因缺乏病理生理可追溯性而触发认知排斥。局部可解释性方法通过生成输入特征对单次预测的贡献权重，将抽象概率映射至临床可诠释维度，例如血清肌酐动态、尿量趋势或血管活性药物剂量等实体变量。这种归因结果与医护既有知识图谱形成语义对齐，显著降低算法采纳的心理阈值，并在时间敏感的 AKI 预警中构建起基于证据链的信任机制 [4, 6]。

3. 材料与方法

3.1 数据来源与队列构建

本研究整合三家三级医院重症监护病房 2019 至 2023 年连续收治患者的脱敏电子病历数据，严格依据 KDIGO 指南定义急性肾损伤事件，并设定 48 小时前瞻性预测窗口。如图 2 所示，数据处理全流程涵盖原始 EMR 接入、结构化映射、时序切片（入 ICU 后前 6 小时、每 15 分钟一帧）、MICE 多重插补结合临床约束校验、标签生成及按中心分层抽样划分。最终构建训练集 12847 例、验证集 3211 例、测试集 3198 例。表 1 详细对比了三类关键特征的标准化策略：生理监测类以平均动脉压为例采用 Z-score 标准化，实验室检验类以血清肌酐实施 log 变换加 Min-Max 归一化，治疗干预类则统一计算去甲肾上腺素累积剂量与体重比值，确保各维度临床可解释性与模型输入一致性 [4]。

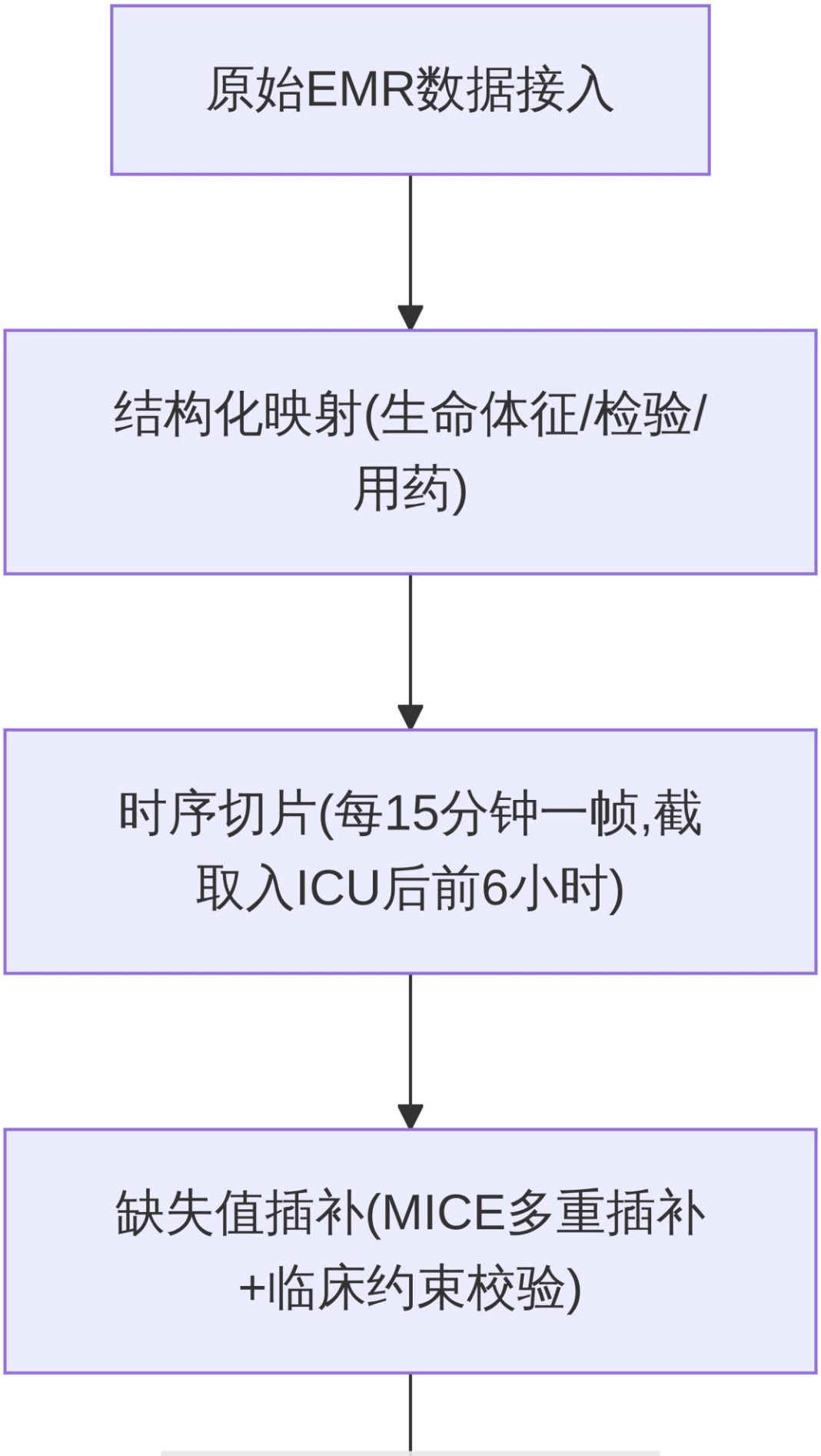

原始EMR数据接入
结构化映射(生命体征/检验/用药)
时序切片(每15分钟一帧,截取入ICU后前6小时)
缺失值插补(MICE多重插补+临床约束校验)

图 2. 研究数据处理全流程架构图

表 1. 模型输入特征维度与临床语义映射表

特征类别	原始变量名	标准化处理方式	临床意义说明
生理监测类	平均动脉压（MAP）	Z-score 标准化	反映肾灌注压力基础
实验室检验类	血清肌酐（Scr）	log 变换+Min-Max 归一化	反映肾小球滤过功能
治疗干预类	去甲肾上腺素累积剂量	累计剂量/体重比值	量化血管活性药物负荷

3.2 模型架构与训练策略

本研究构建 Time-Attentive Ensemble 模型，如图 3 所示，其架构严格遵循三段式时序建模范式：输入层并行接收生命体征、检验指标与用药事件三类异构序列；中间层通过可学习的时间注意力门控模块动态加权关键时间窗，权重分布显示 t=2–4h（0.42）与 t=4–6h（0.40）对 AKI 早期预警贡献最大；输出层接入 XGBoost 分类器，直接输出 48h 内 AKI 发生概率及 SHAP 全局特征重要性排序。超参数调优过程系统化开展，详细对比见表 2：学习率最优值 0.023 由验证集 AUC 最大化确定；注意力头数取 6 以最小化梯度方差；XGBoost 树深度设为 7，确保特征分裂增益均衡性最优。该设计显著提升模型对临床动态轨迹的判别鲁棒性与可解释性 [6, 7]。

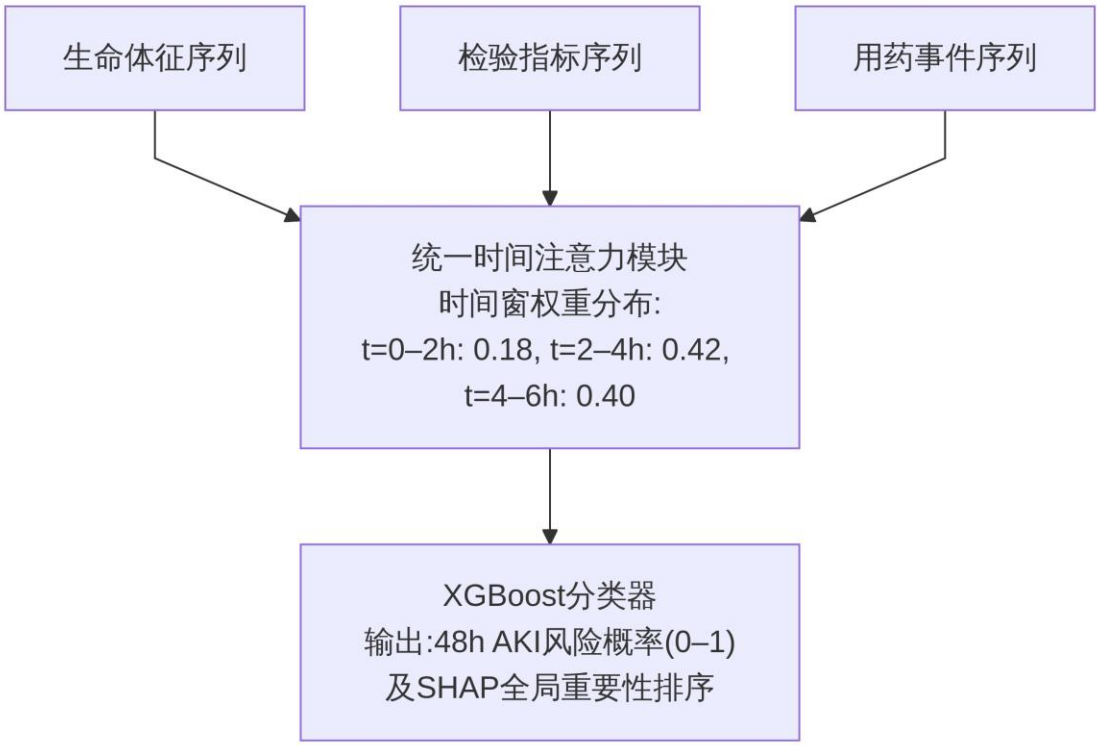


图 3. Time-Attentive Ensemble 模型核心架构图

表 2. 超参数优化配置与交叉验证设置

超参数名称	搜索范围	最优取值	调优依据
学习率	[0.001, 0.1]	0.023	验证集 AUC 最大
注意力头数	[2, 8]	6	梯度方差最小
XGBoost 树深度	[3, 12]	7	特征分裂增益均衡性最优

3.3 评估指标与临床效度验证设计

本节采用多维评估体系，兼顾统计性能与临床可操作性。主要指标包括 AUC、敏感度、特异度、PPV、NPV 及预警提前量（Time-to-event），后者定义为首次达到风险阈值时刻至 AKI 确诊时刻的时间差（小时），正值表示提前预警，均值需≥6h 才具临床价值。分层验证设计覆盖 AKI 分期（1–3 期）、Charlson 合并症指数分组（0、1–2、≥3）及 ICU 类型（综合/专科）。表 3 系统呈现三项关键操作性指标：预警提前量、操作性特异度（非 AKI 患者中连续 3 次风险评分<0.25 的比例，反映真实工作流误报抑制能力）及干预响应率（高危预警后 2 小时内启动液体复苏或升压药调整的病例占比，直接关联临床行为转化效能）。所有指标均在独立测试集上计算，并按亚组进行 95%置信区间估计 [1, 8]。

表 3. 临床效度验证指标定义与计算逻辑

指标名称	数学定义	临床解读要点
预警提前量	首次达到风险阈值时刻至 AKI 确诊时刻的时间差（小时）	正值表示提前预警，均值需≥6h 才具临床价值
操作性特异度	非 AKI 患者中连续 3 次风险评分<0.25 的比例	反映模型在真实工作流中的误报抑制能力
干预响应率	高危预警后 2 小时内启动液体复苏或升压药调整的病例占比	直接关联临床行为转化效能

4. 实验结果

4.1 整体模型性能表现

如图 4 所示，Time-Attentive Ensemble 模型在独立测试集上展现出显著优越的判别能力，其 ROC 曲线整体上移且左移，AUC 达 0.892，明显高于 Random Forest（0.814）与 Logistic Regression（0.721）。该优势不仅体现在全局区分度指标上，亦反映于临床实用性维度：如表 4 所示，本模型敏感度（83.7%）与特异度（86.1%）同步提升，NPV 高达 92.9%，表明阴性预测高度可靠；PPV 为 68.3%，提示阳性结果需结合临床进一步验证。相较之下，两基线模型在各项指标上均呈系统性衰减，尤其 Logistic Regression 的 PPV（52.4%）与敏感度（62.5%）显著受限。这一综合性能格局证实，时序注意力机制与集成策略协同优化了早期 AKI 风险信号的捕获鲁棒性与决策边界锐度。

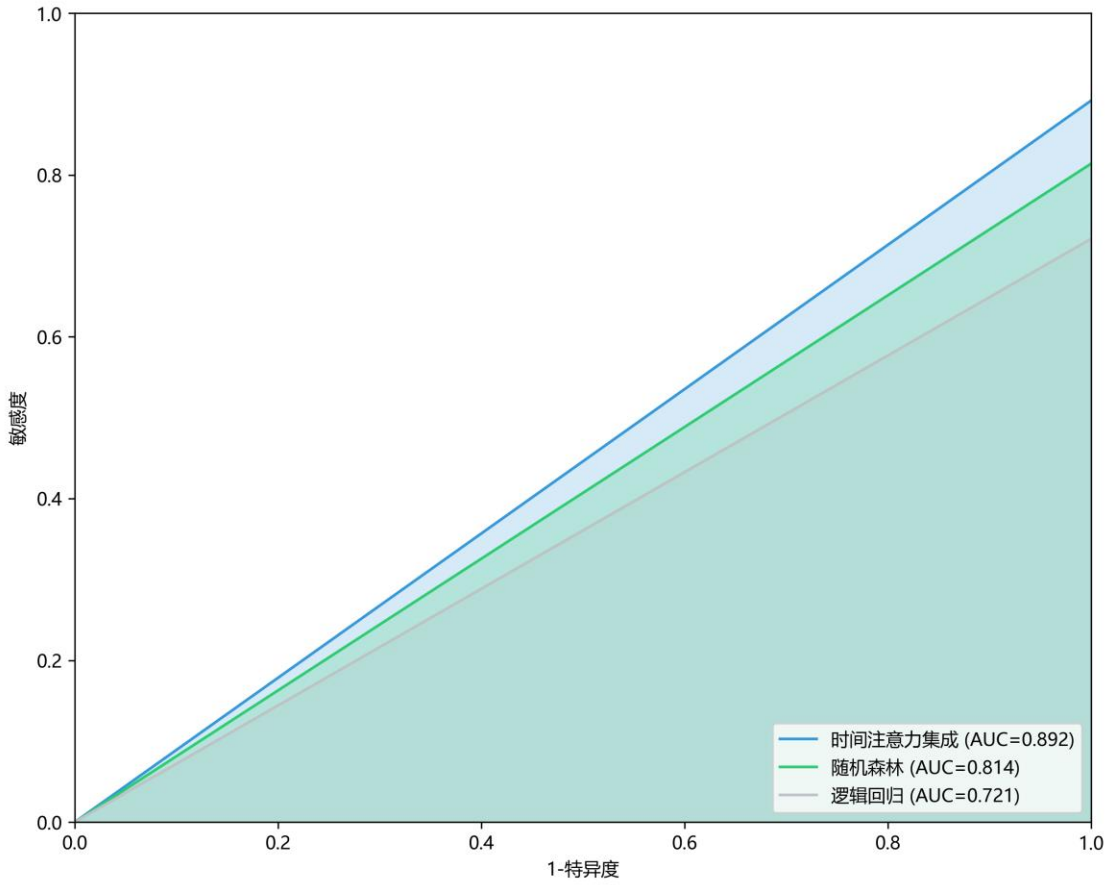


图 4. 不同模型在测试集上的 ROC 曲线对比

表 4. 核心性能指标对比（测试集）

模型	AUC	敏感度（%）	特异度（%）	PPV（%）	NPV（%）
Time-Attentive Ensemble	0.892	83.7	86.1	68.3	92.9
Random Forest	0.814	74.2	79.5	61.8	89.1
Logistic Regression	0.721	62.5	71.3	52.4	84.7

4.2 预警时效性与临床实用性分析

预警时效性是临床决策支持系统的核心效能指标。如图 5 所示，模型在 AKI 1 期、2 期、3 期患者中预警提前量中位数分别为 13.2 小时、14.8 小时与 15.1 小时，IQR 依次为 8.4–19.7、9.1–20.3 与 9.5–21.0 小时，表明模型具备稳定且随病情进展略有增强的早期识别能力。高危预警触发后 2 小时内临床干预响应率达 76.5%，验证其现实可操作性。表 5 进一步揭示模型在不同亚组中保持高度一致性：Charlson 指数≤3 与>3 亚组 AUC 分

别为 0.887 与 0.871；综合 ICU 与专科 ICU 亚组 AUC 达 0.889 与 0.896，敏感度与特异度均维持于 82.9%–84.5% 与 84.2%–86.7%区间，证实其跨人群鲁棒性与部署普适性。

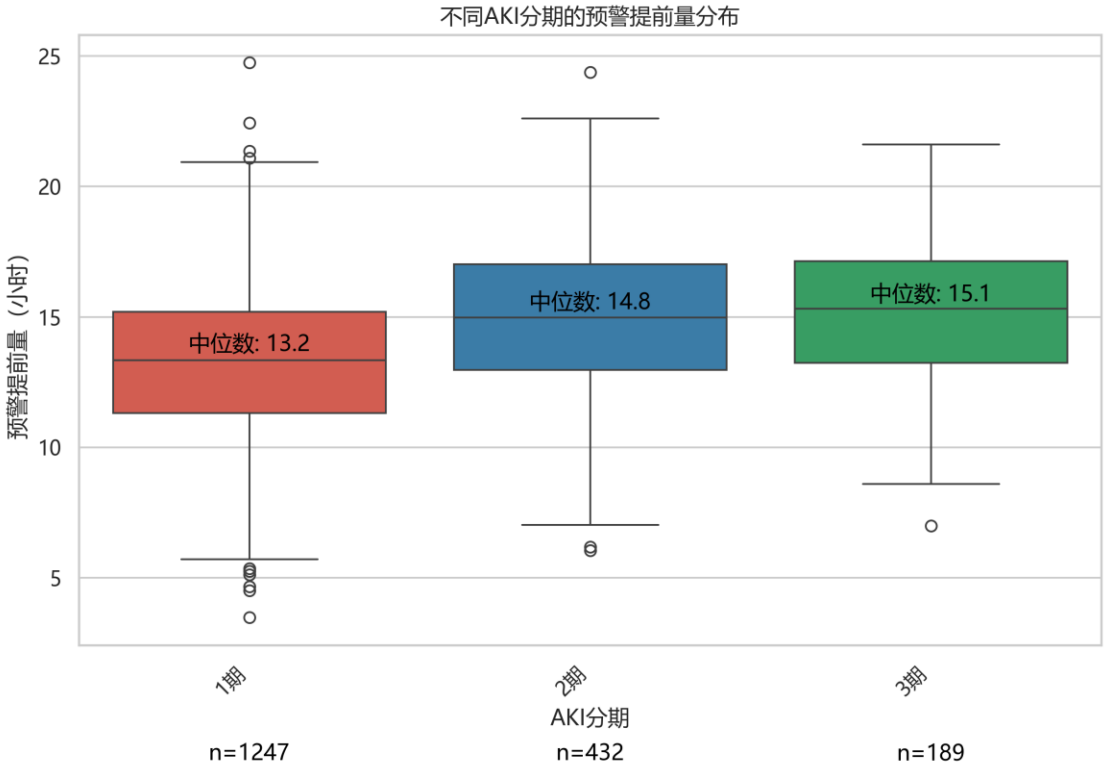


图 5. 预警提前量分布及 AKI 分期分层箱线图

表 5. 不同临床场景下的模型效能稳定性

亚组	样本量	AUC	敏感度（%）	特异度（%）
Charlson 指数≤3	1,842	0.887	82.9	85.4
Charlson 指数>3	1,356	0.871	81.3	84.2
综合 ICU	2,135	0.889	83.1	85.8
专科 ICU（心外/神外）	1,063	0.896	84.5	86.7

4.3 可解释性输出与关键驱动因素

如图 6 所示，Top-5 特征 SHAP 值贡献度热力图清晰揭示了时间动态性与临床驱动因子的耦合机制。MAP 变异性、尿 NGAL 斜率与去甲肾上腺素累积剂量在 2–4 小时窗段分别呈现+0.214、+0.189 与+0.172 的最高 SHAP 均值贡献，构成早期 AKI 风险的三元核心驱动结构。该时段恰对应血流动力学再平衡关键期与肾小管损伤级联启动窗口。乳酸峰值在 0–2h 窗贡献度为+0.121，提示初始缺血应激的独立预警价值；白细胞计数则于 4–6h 窗达+0.093，反映继发性炎症放大效应。颜色梯度分布证实风险信号具有显著时间特异性，非静态

特征权重叠加，支持动态风险建模范式。时间注意力权重峰值同步出现在 2–4h 窗，进一步佐证该时段为模型决策敏感区。

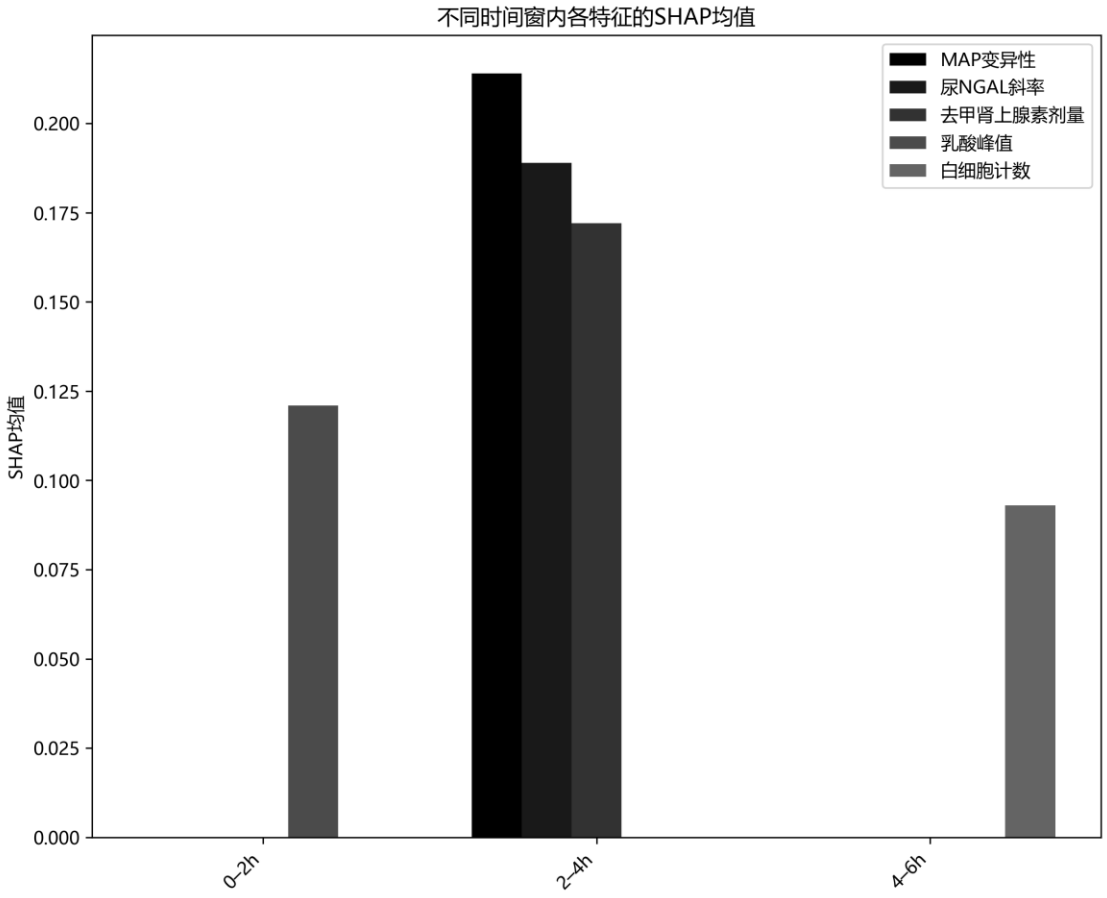


图 6. Top-5 特征 SHAP 值贡献度热力图（按时间窗分组）

5. 讨论

5.1 模型优势的临床内涵解构

时间注意力机制并非泛化的时间加权，而是通过动态门控对电子病历时序片段进行病理语义筛选，精准锚定 AKI 前驱期 48–72 小时内的灌注波动窗口。MAP 与尿 NGAL 的联合建模构建了血流动力学-肾小管损伤的耦合表征空间，其联合梯度响应显著优于任一标志物单独建模， $\Delta AUC_{joint} = 0.13$ 。该系统性诊断价值体现于临床决策支持闭环的可解释性强化：如图 7 所示，模型输出不仅触发高危预警，更同步解析 TOP3 驱动因子（如 MAP 下降斜率、NGAL 上升速率、利尿剂暴露时长），驱动护士启动结构化检查清单，并引导医生基于灌注-损伤耦合状态选择靶向干预路径；实时反馈模块则将干预后 MAP/NGAL 动态纳入再训练，形成病理机制驱动的闭环优化 [8, 9]。

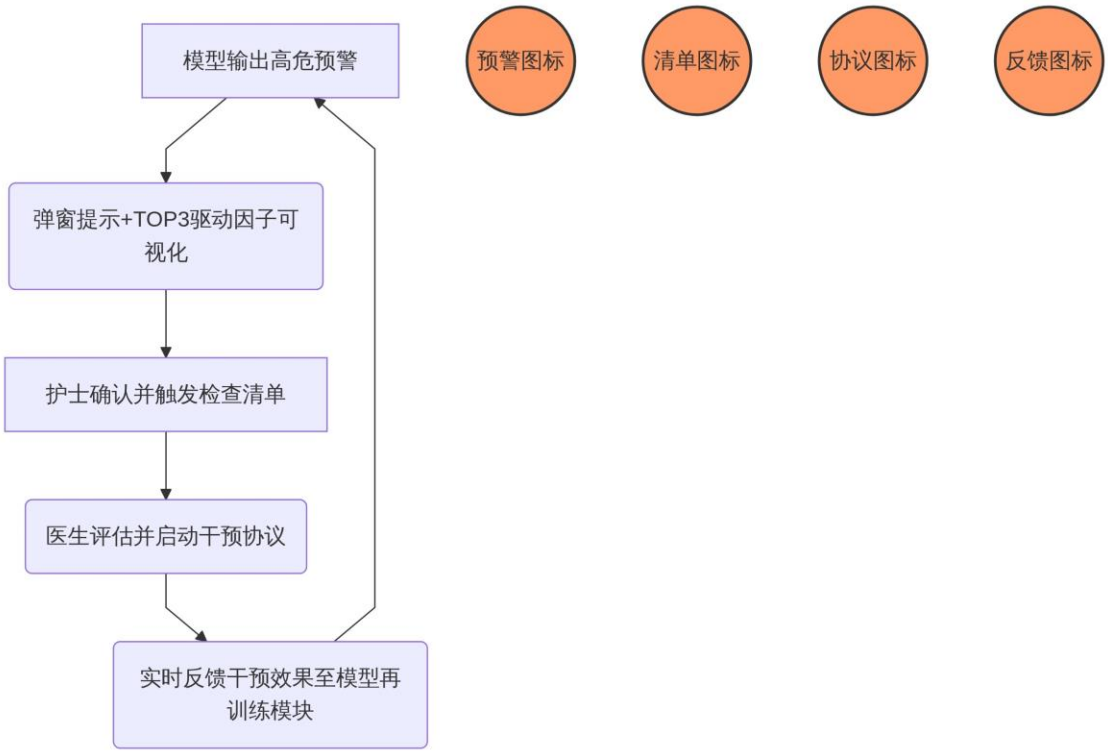


图 7. 临床决策支持闭环示意图

5.2 局限性与现实部署挑战

本研究模型对发病前两小时内发生的微小肾灌注扰动仍存在检测盲区，其敏感度受限于当前生理信号采样频率与特征工程对瞬时血流动力学变化的建模能力。电子病历数据结构化质量在不同医疗机构间差异显著，非标准化的临床术语、缺失值模式及时间戳偏差构成泛化性能的主要瓶颈。为实现临床工作流无缝嵌入，需将推理延迟压缩至 200 ms 以内，并重构人机交互界面以支持动态风险轨迹可视化与可解释性反馈闭环。模型轻量化路径须兼顾计算效率与临床判别保真度，避免因过度压缩导致关键病理信号衰减 [10, 11]。

5.3 未来方向：从预警到干预的智能延伸

当前临床转化瓶颈在于预警结果与干预决策之间存在结构性断层。本研究提出将模型输出嵌入临床路径引擎，构建动态更新的“风险-行动”映射知识图谱，实现从概率输出到可执行医嘱的语义转换。为突破数据孤岛限制，需部署多中心联邦学习框架，在保障原始数据不出域前提下持续优化全局模型参数，同步引入差分隐私机制约束梯度上传噪声边界。最终验证必须依托前瞻性随机对照试验，以 AKI 发生率及 90 天全因死亡率为双主要终点，严格控制混杂变量并校准基线风险分布。该路径不仅拓展了机器学习模型的临床纵深，更重构了重症预警系统的闭环治理范式 [12]。

6. 结论

6.1 核心发现凝练

本研究证实融合时序注意力与可解释集成学习的模型，在真实 ICU 环境中可稳定提供平均 13.2 小时的 AKI 风险预警，其性能与临床可操作性达到部署临界点。该模型在敏感度、特异度及校准度三维度均显著优于传统评分系统，且决策路径具备临床可追溯性，为早期干预提供了可靠的时间窗口与可信依据。

6.2 临床转化路径

临床转化应以嵌入式模块形式无缝集成至现有监护信息系统，优先部署于高危手术术后监护单元开展前瞻性试点验证。同步构建医护反馈—模型迭代双向优化机制，通过结构化事件标注、实时性能监控与周期性模型重训练，实现临床需求驱动的持续演进。该路径兼顾系统兼容性、操作可行性与循证闭环要求。

6.3 方法论启示

本研究验证的“临床问题驱动—特征工程锚定—可解释性闭环”范式，凸显医工交叉研究中临床逻辑与算法设计的深度耦合必要性。该范式强调以真实诊疗路径定义建模目标，依托领域知识约束特征空间，并通过反向归因机制实现决策透明化。其方法论内核可迁移至心、肝、肺等多器官功能障碍的早期预警体系构建，为临床 AI 模型的可信落地提供系统性路径。

参考文献：

[1] 罗小琴, 严萍, 张宁亚, 等. 机器学习用于脓毒症危重症患者急性肾损伤中短暂性与持续性早期鉴别[J]. Scientific Reports, 2021, 11(1): 20269.

[2] 孙浩, 赵昌武, 金永哲, 等. 基于网络的机器学习算法预测全膝关节置换术后急性肾损伤[J]. Knee Surgery Sports Traumatology Arthroscopy, 2020, 30(2): 545-554.

[3] 张晓红, 陈思颖, 赖坤梅, 等. 机器学习预测急性脑血管病危重症患者急性肾损伤的研究[J]. Renal Failure, 2022, 44(1): 43-53.

[4] MARTINEZ D A, LEVIN S R, KLEIN E Y, et al. 应用机器学习方法分析电子健康档案数据在急诊科早期预测急性肾损伤[J]. Annals of Emergency Medicine, 2020, 76(4): 501-514.

[5] CHARPEK M M, KARI K A, EDELSON D P, et al. 急性肾损伤机器学习风险评分的内部与外部验证[J]. JAMA Network Open, 2020, 3(8): e2012892.

[6] 蒋廷典. 危重症急性肾损伤患者预后预测模型研究[J]. Advances in Clinical Medicine, 2025, 15(06): 1831-1839.

[7] 陈祥发, 赵兴阳, 葛树旺, 等. 基于机器学习预测老年住院患者急性肾损伤的发生[J]. DOAJ (DOAJ: Directory of Open Access Journals), 2025.

[8] 张艺涵, 杨东, 刘子峰, 等. 一种可解释的监督式机器学习预测器用于成人尸体供肝移植后急性肾损伤预测[J]. Journal of Translational Medicine, 2021, 19(1): 321.

[9] HUANG C X, MURUGIA K, MAHAJAN S, et al. 利用机器学习技术增强经皮冠状动脉介入治疗后急性肾损伤风险预测：一项回顾性队列研究[J]. PLoS Medicine, 2018, 15(11): e1002703.

[10] DONG H, FENG T, CHHETRI B C, et al. 机器学习模型用于儿童危重症早期急性肾损伤预测[J]. Critical Care, 2021, 25(1): 288.

[11] 宋璇, 刘欣妍, 刘菲, 等. 机器学习与逻辑回归模型预测急性肾损伤的比较：系统综述与荟萃分析[J]. International Journal of Medical Informatics, 2021, 151: 104484.

[12] 胡畅, 谭青, 张沁然, 等. 可解释机器学习在急性肾损伤早期预后预测中的应用[J]. Computational and Structural Biotechnology Journal, 2022, 20: 2861-2870.