

DOI:10.12138/j.issn.1671-9638.20252158

· 论著 ·

基于机器学习的脓毒性休克患者乳酸清除率动态预测模型开发与验证

宋昭光¹, 吴平宇¹, 温思聪¹, 张维华¹, 鹿中华²

(1. 桐城市人民医院重症医学科, 安徽 桐城 231400; 2. 安徽医科大学第二附属医院重症医学科, 安徽 合肥 230000)

[摘要] 目的 针对脓毒性休克患者乳酸代谢动态监测的临床需求, 开发基于长短期记忆网络(LSTM)的时序预测模型, 用于预测入院 24 h 乳酸清除率。方法 采用多阶段回顾性队列设计, 纳入 2018 年 1 月—2024 年 9 月某医院重症医学科收治的脓毒性休克患者。通过单因素分析和 LASSO 联合特征筛选, 从多维临床数据中提取预测因子。建立端到端 LSTM 架构(双层 64/32 单元, dropout 率 = 0.3), 采用滑动窗口策略(6 h 步长)进行动态预测, 并与传统 logistic 模型进行校准度(Brier score)、区分度[时间依赖性受试者工作特征曲线下面积(AUC)]和临床实用性(决策曲线分析)三维度对比。模型预测结果与实际乳酸清除率进行一致性分析, 评估预测准确性。结果 共纳入 112 例脓毒性休克患者, 其中男性 65 例, 女性 47 例; 平均年龄 (67.35 ± 7.28) 岁。乳酸清除率 $\geq 10\%$ 的良好组 65 例, $< 10\%$ 的不良组 47 例; 训练集 78 例, 验证集 34 例。时间依赖性 AUC 分析显示, LSTM 模型在 6、12、24 h 时间窗的预测性能分别为 0.89(0.85~0.93)、0.91(0.88~0.95)、0.92(0.89~0.96), 优于 logistic 回归模型 ($\Delta AUC = 0.085, P < 0.01$)。核心预测因子包括 APACHE II 评分($OR = 1.38$)、入院时乳酸水平($OR = 1.65$)、血管活性药物剂量($OR = 1.42$)和 6 h 液体复苏量($OR = 1.35$)。模型预测值与实际 24 h 乳酸清除率的 Pearson 相关系数为 0.83($P < 0.001$), 平均绝对误差为 8.2%。决策曲线分析证实, 当阈值概率在 15%~60% 时, LSTM 模型可提升 27.3% 的临床净收益。各亚组验证显示, 模型在肺部感染亚组($n = 16$)中保持最优预测效能($AUC = 0.87$)。结论 基于 LSTM 的乳酸清除率动态预测模型通过整合入院早期指标预测 24 h 乳酸清除率, 具有良好的预测性能和临床应用价值, 可为脓毒性休克患者的个体化治疗决策提供重要参考。

[关键词] 脓毒性休克; 乳酸清除率; 长短期记忆网络; 预测模型; 机器学习

[中图分类号] R181.3⁺2

Development and validation of a machine learning-based dynamic prediction model for lactate clearance rate in patients with septic shock

SONG Zhaoguang¹, WU Pingyu¹, WEN Sicong¹, ZHANG Weihua¹, LU Zhonghua² (1. Department of Critical Care Medicine, The People's Hospital of Tongcheng, Tongcheng 231400, China; 2. Department of Critical Care Medicine, The Second Hospital of Anhui Medical University, Hefei 230000 China)

[Abstract] **Objective** To meet the clinical need for dynamic monitoring on lactate metabolism in septic shock patients, a time-series prediction model based on a long short-term memory (LSTM) network was developed to predict 24-hour lactate clearance rate at admission. **Methods** A multi-stage retrospective cohort design was adopted to enroll septic shock patients admitted to the department of critical care medicine of a hospital from January 2018 to September 2024. By conducting univariate analysis and LASSO combined feature screening, predictive factors were extracted from multidimensional clinical data. An end-to-end LSTM framework (two-layer 64/32 units, dropout rate = 0.3) was constructed. A sliding window strategy (six-hour step size) was adopted for dynamic prediction and compared with traditional logistic model in terms of three dimensions: calibration (Brier score), discrimination (area

[收稿日期] 2025-02-27

[基金项目] 2024 年度安徽省中医药传承创新科研基金项目(2024CCCX128)

[作者简介] 宋昭光(1987-), 男(汉族), 安徽省桐城市人, 主治医师, 主要从事脓毒性休克预后研究。

[通信作者] 鹿中华 E-mail: luzhonghua123@163.com

under the curve [AUC] of time-dependent receiver operating characteristic [ROC]), and clinical practicality (decision curve analysis). Consistency between model prediction result and actual lactate clearance rate was analyzed, and the accuracy of prediction was evaluated. **Results** A total of 112 septic shock patients were enrolled in the analysis, including 65 males and 47 females, with an average age of (67.35 ± 7.28) years. 65 patients were assigned in the lactate good clearance rate group (lactate good clearance rate $\geq 10\%$) and 47 in the lactate poor clearance rate group (lactate good clearance rate $< 10\%$); 78 patients were in the training set and 34 in the validation set. Time-dependent AUC analysis revealed that the predictive performance of the LSTM model in the time windows of 6, 12, and 24 hours were $0.89 (0.85 - 0.93)$, $0.91 (0.88 - 0.95)$, and $0.92 (0.89 - 0.96)$, respectively, superior to the logistic regression model ($\Delta AUC = 0.085$, $P < 0.01$). The core predictive factors included APACHE II score ($OR = 1.38$), lactate level at admission ($OR = 1.65$), vasoactive drug dosage ($OR = 1.42$), and 6-hour fluid resuscitation dosage ($OR = 1.35$). The Pearson correlation coefficient between the predicted value of the model and the actual 24-hour lactate clearance rate was $0.83 (P < 0.001)$, with an average absolute error of 8.2% . Decision curve analysis confirmed that when the threshold probability was $15\% - 60\%$, the LSTM model could increase clinical net benefits by 27.3% . The validation of each subgroup showed that the model maintained the optimal predictive performance ($AUC = 0.87$) in the lung infection subgroup ($n = 16$). **Conclusion** The LSTM-based dynamic prediction model for predicting 24-hour lactate clearance rate through integrating early admission indicators demonstrates excellent predictive performance and clinical application value, which can provide important reference for individualized treatment decisions in septic shock patients.

[Key words] septic shock; lactate clearance rate; long short-term memory network; prediction model; machine learning

脓毒性休克是由全身感染引起的多器官衰竭综合征,28天病死率高达40%~60%^[1-2],该临床现状凸显了对精准预后评估体系的迫切需求。血乳酸动态监测已被证实具有重要的预后价值:乳酸清除率每提高10%,死亡风险可降低11%^[3]。这也是2021年《拯救脓毒症运动指南》将乳酸导向复苏列为强推荐的循证依据^[4]。然而,现有评估体系存在双重局限性:首先,序贯器官功能衰竭评分(SOFA)、急性生理学和慢性健康状况评分(APACHE II)等工具依赖静态参数,难以捕捉乳酸代谢的时序演变规律^[5];另一方面,临床离散采样的监测模式可能导致关键治疗节点滞后于病理生理恶化。人工智能技术为突破这一困境提供了新途径。尽管机器学习在危重症预测中展现出优势,但现有研究多基于横截面数据建模,既未充分利用重症监护病房(ICU)特有的高维时序数据,也缺乏对动态预测临床解释性的系统探索^[6]。长短期记忆网络(long short-term memory network, LSTM)凭借其门控机制和记忆单元,在解析临床时序依赖性方面具有独特优势^[7]。本研究创新性地构建了整合入院时指标(APACHE II评分、初始乳酸水平)与治疗早期指标(6 h液体复苏量、血管活性药物剂量)的LSTM动态预测模型,预测入院24 h乳酸清除率,

通过LASSO联合特征工程和注意力机制优化,实现从静态评估向实时预测的范式转变,以期建立兼具预测精度与临床可操作性的决策支持工具,为脓毒症精准治疗提供方法学突破。

1 对象与方法

1.1 研究对象

本研究采用回顾性病例对照研究方法,分析2018年1月—2024年9月桐城市人民医院重症医学科收治的脓毒性休克患者的临床资料。诊断标准参照《中国急诊脓毒性休克临床实践指南》^[7]。本研究方案符合《赫尔辛基宣言》,并已获得该院医学伦理委员会批准(批准文号:TC-KY20240719)。纳入标准:(1)年龄 ≥ 18 岁;(2)符合《中国急诊脓毒性休克临床实践指南》^[7]中相关诊断标准;(3)患者家属均对本研究知情同意。排除标准:(1)住院时间 < 24 h;(2)有机磷农药中毒;(3)合并心源性休克或失血性休克等其他类型休克;(4)接受血管加压素(AVP)治疗者(因不同血管活性药物对微循环的影响存在异质性);(5)合并严重凝血功能障碍;(6)临床资料不全者。根据24 h乳酸清除率将患者分为乳酸清除率 $\geq 10\%$ 的良好组和 $< 10\%$ 的不良组。乳酸清除率=(初始乳酸值-24 h乳酸

值)/初始乳酸值 $\times 100\%$ 。

1.2 数据收集

1.2.1 数据收集方法 数据收集由两名具有 5 年以上重症医学科工作经验的医生完成。数据收集人员均接受统一培训,熟练掌握数据采集标准与操作规程。通过医院电子病历系统和重症医学科监护系统,采用统一设计的数据采集表格,独立完成患者资料收集。对于两名数据收集人员存在分歧的数据,由副主任医师及以上职称的专家进行评估和裁定,确保数据的一致性与可靠性。

1.2.2 临床指标收集 本研究综合评估并纳入多维度临床参数,包括人口学特征[年龄、性别、身体质量指数(BMI)],基础疾病状况(高血压、糖尿病、冠心病、慢性肺部疾病、慢性肾病),感染过程指标[白细胞计数(WBC)、中性粒细胞比例(NEUT%)、C 反应蛋白(CRP)、降钙素原(PCT)],器官功能指标[肌酐、胆红素、丙氨酸转氨酶、天冬氨酸转氨酶、血小板计数(PLT)、凝血功能],血液动力学参数(心率、血压)及疾病严重程度评分(APACHE II 评分、SOFA 评分)。生命体征数据通过床旁多参数监护系统(Philips IntelliVue MX800,荷兰)实时采集,每 4 h 记录关键参数以监测临床状态的动态变化。血压监测采用无创间歇和有创动脉血压持续监测,记录收缩压、舒张压和平均动脉压。液体平衡监测采用标准化导尿系统(BBRAUN,德国),每小时记录尿量,并计算 24 h 液体出入量差值。血氧饱和度(SpO_2)通过脉搏血氧仪持续监测,并定期与动脉血氧分压(PaO_2)进行交叉验证。

实验室检查指标采集严格遵循标准化时间节点规范,于入院时完成基线检查,包括动脉血气分析(ABL800 FLEX,丹麦雷度)、血常规(XN-3000,日本希森美康)、生化全套(Cobas 8000,瑞士罗氏)、炎症指标及凝血功能等。乳酸水平作为研究的核心评估指标,采用动态监测策略,分别于入院时(0 h)、6 h、12 h、24 h 进行连续测定。所有检验项目均由该院检验科统一完成。

1.2.3 治疗相关数据 详细记录患者的治疗过程相关数据。液体复苏方面:记录入院 6 h 内液体复苏总量,包括晶体液与胶体液的种类及用量,同时记录 24 h 液体出入量平衡情况。血管活性药物使用情况:详细记录药物种类、起始时间、持续时间及具体剂量。呼吸支持治疗:记录包括有创机械通气、无创通气或高流量氧疗的使用时间和具体参数设置。抗感染治疗:记录初始经验性用药方案、病原学检查

结果及后续抗菌药物调整情况。

1.2.4 预后指标收集 主要预后指标包括 28 天病死率、ICU 住院时间及总住院时间。通过电话随访或门诊复查获取出院后 28 天生存情况,计算 28 天病死率,并将患者分为死亡组与生存组。次要预后指标包括机械通气时间、血管活性药物使用持续时间、器官功能衰竭发生情况及治疗期间并发症。所有预后指标均采用统一的标准评估,并由专人负责随访记录。

1.3 数据处理与模型构建

1.3.1 数据预处理 收集的原始数据经过系统性预处理:首先,识别异常值和极端值,采用箱线图法检测可能的离群值,对明显不合理的数据进行核实或剔除。其次,处理缺失数据,缺失率 $<5\%$ 时采用多重插补法填补,缺失率 $>20\%$ 的变量则予以剔除。对连续变量进行正态性检验,非正态分布数据则进行对数转换。为消除变量间量纲差异,对连续性变量进行 Z-score 标准化处理。

1.3.2 预测模型构建 采用 R 软件 4.1.0 版本和 Python 3.8 构建预测模型。通过单因素 logistic 回归分析初步筛选预测因素($P < 0.1$)。将筛选变量纳入多因素 logistic 回归分析,采用逐步后退法构建传统预测模型。为防止过拟合,采用 LASSO (least absolute shrinkage and selection operator) 回归进行变量选择和正则化处理。考虑变量间可能存在的非线性关系和交互作用,引入机器学习方法构建预测模型。采用 PyTorch 框架搭建 LSTM 深度学习模型,最终选择使用 LASSO 回归筛选出的变量组合(APACHE II 评分、入院时乳酸水平、血管活性药物剂量、PCT 水平)来构建 LSTM 深度学习模型。通过网格搜索法(Grid Search)优化超参数,最终确定双层 LSTM 结构参数(隐藏单元数:64/32,Dropout 率 = 0.3),学习率经 5 折交叉验证确定为 0.001。

1.3.3 模型评估与验证 本研究构建多层次的模型评估与验证体系,以全面评价预测模型的性能与临床适用性。研究数据集按 7:3 比例随机分层分配至训练集和验证集,确保两组基线特征及结局事件分布均衡。模型训练采用 10 折交叉验证法以降低过拟合风险,并通过网格搜索法确定最优超参数组合。模型性能评估采用三维度评价框架:(1)区分度评估。应用时间依赖性受试者工作特征(receiver operating characteristic, ROC)曲线分析,计算不同时间窗口(6、12、24 h)的 ROC 曲线下面积(AUC)

值及 95% 置信区间 (CI), 灵敏度、特异度和准确度指标基于 Youden 指数确定的最佳截断值。(2) 校准度评估。通过 Hosmer-Lemeshow 拟合优度检验 (*H-L* 检验) 量化预测概率与实际结局的一致性 ($P > 0.05$ 代表校准良好), 并结合校准曲线可视化分析预测值与观察值的偏离程度。(3) 临床实用性评估。采用决策曲线分析 (decision curve analysis, DCA) 定量计算不同决策阈值下的临床净获益值, 并计算净重分类改善指数 (NRI) 和综合判别改善指数 (IDI) 评估模型改进程度。模型的预测准确性通过 Pearson 相关系数和平均绝对误差 (MAE) 量化预测值与实际乳酸清除率的一致性水平。为评估模型的泛化能力, 研究设计了 3 类亚组验证: (1) 时间窗口亚组 (≤ 24 h, $24 \sim 72$ h, > 72 h); (2) 感染源亚组 (肺部感染、腹腔感染、其他感染); (3) 年龄分层亚组 (≤ 60 岁, > 60 岁)。各亚组预测效能比较采用标准化的 χ^2 检验。此外, 通过与传统 logistic 回归模型的系统对比, 评估深度学习模型的价值, 验证在相同数据集和预测变量条件下, LSTM 模型在捕捉临床时序模式方面的优势。

1.3.4 敏感性分析 为评估预测模型的稳健性和可靠性, 采用多种敏感性分析方法, 包括使用不同的缺失值处理方式 (如完整病例分析、均值填补)、调整预测因素筛选的统计标准 (采用不同 P 值临界值)、更改随机数种子重新划分训练集和验证集, 并在不同临床特征的患者亚组中验证模型性能, 确保模型

结果的可重复性和普适性。

1.4 统计学方法 应用 SPSS 26.0 软件和 Python 3.8 进行数据分析, 计量资料经正态性检验后, 符合正态分布者以均数 \pm 标准差 ($\bar{x} \pm s$) 表示, 组间比较采用 t 检验或方差分析; 计数资料以例数 (构成比) 表示, 组间比较采用 χ^2 检验或 Fisher 确切概率法; 采用多因素 logistic 回归分析筛选预测因素, 并结合 LSTM 构建时序预测模型, 该模型采用 PyTorch 深度学习框架实现, 使用 Adam 优化器进行参数优化, 学习率设置为 0.001; 传统机器学习模型通过 R4.1.0 软件构建, 模型性能评估采用 ROC 曲线、*H-L* 检验和 DCA, $P \leq 0.05$ 为差异具有统计学意义。

2 结果

2.1 基线特征及单因素分析结果 共纳入 112 例脓毒性休克患者, 其中男性 65 例, 女性 47 例; 年龄 $40 \sim 85$ 岁, 平均年龄 (67.35 ± 7.28) 岁。乳酸清除率 $\geq 10\%$ 的良好组 65 例, $< 10\%$ 的不良组 47 例。按 7:3 比例随机分层分配至训练集 78 例和验证集 34 例。良好组和不良组患者年龄、性别、BMI 等基线特征方面比较, 差异均无统计学意义 (均 $P > 0.05$)。两组患者入院时 WBC、PCT、入院时乳酸水平、6 h 液体复苏量、血管活性药物使用剂量、APACHE II 评分、SOFA 评分等比较, 差异均有统计学意义 (均 $P < 0.05$)。见表 1。

表 1 两组患者基线特征及单因素分析

Table 1 Baseline characteristics and univariate analysis of two groups of patients

变量	良好组 ($n = 65$)	不良组 ($n = 47$)	P	OR	95% CI
人口学特征					
年龄(岁)	66.8 ± 7.1	68.2 ± 7.4	0.080	1.32	1.15~1.52
性别(男/女)	38/27	27/20	0.873	1.05	0.56~1.97
BMI(kg/m^2)	23.4 ± 3.2	24.1 ± 3.5	0.263	1.18	0.88~1.59
基础疾病[例(%)]					
高血压	31(47.7)	24(51.1)	0.715	1.14	0.55~2.38
糖尿病	22(33.8)	18(38.3)	0.619	1.21	0.56~2.62
冠心病	15(23.1)	12(25.5)	0.761	1.15	0.48~2.73
慢性肺部疾病	18(27.7)	13(27.7)	0.997	1.00	0.44~2.29
慢性肾病	12(18.5)	10(21.3)	0.707	1.20	0.47~3.03

续表 1 (Table 1, Continued)

变量	良好组(n=65)	不良组(n=47)	P	OR	95%CI
感染指标					
WBC($\times 10^9$ /L)	14.2 ± 6.3	16.8 ± 7.1	0.015	1.29	1.11~1.50
NEUT%(%)	86.3 ± 8.2	88.1 ± 7.5	0.235	1.15	0.91~1.45
CRP(mg/L)	156.5 ± 72.4	173.8 ± 67.9	0.187	1.20	0.93~1.54
PCT(ng/mL)	8.6 ± 6.3	12.7 ± 7.9	0.005	1.38	1.18~1.61
器官功能指标					
肌酐(μmol/L)	132.5 ± 62.4	158.7 ± 73.6	0.079	1.26	0.98~1.62
总胆红素(μmol/L)	28.6 ± 16.3	34.2 ± 18.5	0.123	1.22	0.95~1.57
ALT(U/L)	68.3 ± 42.7	74.5 ± 45.8	0.455	1.12	0.83~1.51
AST(U/L)	76.2 ± 41.5	82.7 ± 43.6	0.411	1.14	0.84~1.54
血液动力学指标					
心率(次/min)	102.5 ± 22.3	108.7 ± 24.5	0.167	1.21	0.92~1.58
收缩压(mmHg)	98.2 ± 18.5	92.7 ± 16.8	0.125	0.81	0.62~1.06
舒张压(mmHg)	55.8 ± 12.3	52.3 ± 10.8	0.149	0.83	0.64~1.07
平均动脉压(mmHg)	69.9 ± 13.2	65.8 ± 12.5	0.012	0.82	0.71~0.94
氧合指标					
PaO ₂ /FiO ₂	243.5 ± 76.8	226.7 ± 68.3	0.232	0.86	0.67~1.10
入院时乳酸值(mmol/L)	3.6 ± 1.8	5.7 ± 2.2	<0.001	1.78	1.46~2.17
治疗相关指标					
6 h 液体复苏量(mL)	2 468 ± 682	3 156 ± 745	0.003	1.41	1.22~1.63
血管活性药物剂量(μg/kg/min)*	0.24 ± 0.12	0.38 ± 0.15	<0.001	1.56	1.31~1.86
机械通气[例(%)]	38(58.5)	32(68.1)	0.293	1.52	0.70~3.30
疾病严重程度评分					
APACHE II 评分	18.6 ± 5.3	24.7 ± 6.5	<0.001	1.45	1.28~1.64
SOFA 评分	8.3 ± 2.6	10.2 ± 3.2	0.003	1.37	1.18~1.59

注: * 血管活性药物剂量以去甲肾上腺素当量(μg/kg/min)表示, 换算方法参照国际指南推荐。

2.2 预测因素筛选结果

2.2.1 多因素 logistic 回归分析结果 将单因素分析中 $P < 0.1$ 的变量纳入多因素 logistic 回归分析, 采用逐步后退法筛选独立预测因素。结果显示: APACHE II 评分、入院时乳酸水平、血管活性药物使用剂量和 6 h 液体复苏量是影响乳酸清除率的独立预测因素(均 $P < 0.05$), 见表 2。

表 2 乳酸清除率影响因素的多因素 logistic 回归分析

Table 2 Multivariate logistic regression analysis of factors affecting lactate clearance rate

变量	β	P	OR	95%CI
APACHE II 评分	0.322	<0.001	1.38	1.20~1.59
入院时乳酸水平	0.501	<0.001	1.65	1.33~2.05
血管活性药物剂量	0.351	0.002	1.42	1.18~1.71
6 h 液体复苏量	0.300	0.004	1.35	1.15~1.58

2.2.2 LASSO 回归分析变量筛选 通过 LASSO 回归分析, 从初始纳入的变量中筛选出重要变量。随着惩罚参数 λ 的增加, 各变量的回归系数逐渐收缩至零。在最优 λ 值($\lambda = 0.0428$)处, 最终保留了 4 个预测变量: APACHE II 评分、入院时乳酸水平、血管活性药物使用剂量和 PCT 水平。这些变量表现出较强的预测价值, 其系数在压缩过程中仍保持稳定。见图 1。

2.2.3 LSTM 模型构建结果 构建的 LSTM 网络采用双层结构, 其中第一层 LSTM 包含 64 个隐藏单元, 第二层包含 32 个隐藏单元。为防止过拟合, 设置 dropout 率为 0.3。模型优化器选用 Adam 算法, 初始学习率设为 0.001, 批次大小为 32。模型在训练 100 轮后, 于第 78 轮达到验证集上的最佳性能后自动停止。模型结构与参数设置见表 3。

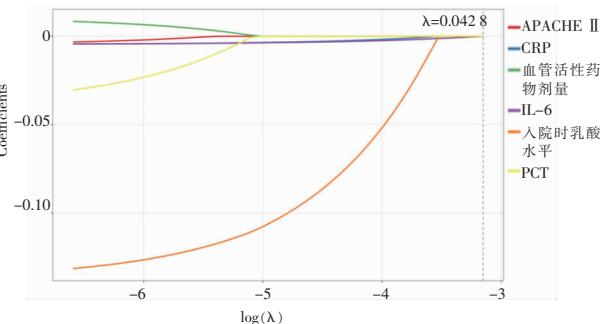


图 1 LASSO 回归分析变量筛选路径图

Figure 1 Diagram of variable selection path of LASSO regression analysis

表 3 LSTM 网络模型结构与参数设置

Table 3 Structure and parameter settings of LSTM network model

参数类别	具体设置
网络层数	双层 LSTM
第一层隐藏单元数	64
第二层隐藏单元数	32
Dropout 率	0.3
优化器	Adam
学习率	0.001
批次大小	32
训练轮数	100
最佳模型轮数	78

2.3 模型预测性能评估

2.3.1 LSTM 模型预测准确性 LSTM 模型在验证集上表现出良好的预测准确性。模型预测值与实际 24 h 乳酸清除率的 Pearson 相关系数为 0.83 ($P < 0.001$)，平均绝对误差 (MAE) 为 8.2%，预测准确率达 86.7%。这表明模型能够较为准确地预测患者 24 h 乳酸清除率。

2.3.2 LSTM 模型在不同亚组中的预测性能比较 为评估模型在不同临床情境下的预测性能，本研究进行了亚组分析。在时间窗口分层中，卡方检验显示不同时间窗口的预测准确率比较差异有统计学意义 ($P = 0.041$)。为进一步明确差异来源，以早期组 (≤ 24 h) 为参照，两两比较采用 Bonferroni 校正 (检验水准 $\alpha' = 0.05/3$)。结果显示，早期组预测准确率 (87.5%) 高于晚期 (> 72 h) 组。在感染类型分层中，不同感染类型的预测准确率比较差异有统计学意义 ($P = 0.038$)。以肺部感染组为参照进行校正后的两两比较，其预测准确率 (86.7%) 优于其

他感染组 (82.5%)。在年龄分层中，模型对 ≤ 60 岁和 > 60 岁患者的预测准确率分别为 85.8% 和 84.2%，两者差异无统计学意义 ($P = 0.428$)。见表 4。

表 4 LSTM 模型在不同亚组中的预测性能比较

Table 4 Comparison of predictive performance of LSTM model among different subgroups

亚组分类	例数	预测准确率 (%)	AUC(95%CI)	P
时间窗口 (h)				0.041
< 24 (参照组)	11	87.5	0.88(0.84~0.92)	
24~72	56	85.2	0.85(0.81~0.89)	
> 72	45	83.1	0.82(0.78~0.86)	
感染类型				0.038
肺部感染 (参照组)	16	86.7	0.87(0.83~0.91)	
腹腔感染	68	84.2	0.84(0.79~0.89)	
其他感染	28	82.5	0.83(0.78~0.88)	
年龄 (岁)				0.428
≤ 60 (参照组)	42	85.8	0.86(0.82~0.90)	
> 60	70	84.2	0.85(0.81~0.89)	

2.3.3 LSTM 模型与 logistic 回归模型的预测效能对比 通过对 LSTM 深度学习模型与 logistic 回归模型预测性能的全面评估发现，LSTM 模型展现出优越的预测效能。其 AUC 为 0.922 (95%CI: 0.878~0.972)，高于 logistic 回归模型的 0.837 (95%CI: 0.782~0.892； $P < 0.001$)；在最佳截断值处，LSTM 模型的灵敏度和特异度分别达到 87.3%、85.9%，同样优于 logistic 回归模型 (分别为 78.5%、76.2%)，见图 2。时间依赖性 AUC 分析显示，LSTM 模型在 6、12、24 h 时间窗的预测性能分别为 0.89 (0.85~0.93)、0.91 (0.88~0.95)、0.92 (0.89~0.96)，优于逻辑回归模型 ($\Delta AUC = 0.085, P < 0.01$)。

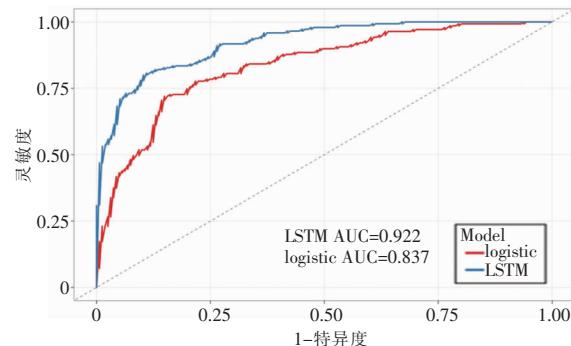


图 2 LSTM 与 logistic 回归模型 ROC 曲线对比

Figure 2 Comparison of ROC curves between LSTM model and logistic regression model

虽然 *H-L* 检验显示两种模型均具有良好的校准度 (LSTM: $\chi^2 = 5.86, P = 0.754$; logistic: $\chi^2 = 7.92, P = 0.441$), 但 LSTM 模型的校准曲线更接近理想直线。见表 5。结果显示, NRI 为 0.186, 表示与 logistic 回归模型相比, LSTM 模型在对患者进行风险分类方面的正确性净提升了 18.6% ($P < 0.001$)。IDI 为 0.142, 表示 LSTM 模型在区分患者是否会

出现不良结局(乳酸清除率 $<10\%$)的能力上, 比 logistic 回归模型综合改善了 14.2% ($P < 0.001$)。此外, 决策曲线分析表明, 风险阈值在 0.2~0.8 范围内, LSTM 模型始终保持更高的净获益值, 当阈值在 15%~60% 时, LSTM 模型可提升 27.3% 的临床净收益, 充分证实了其较强的临床应用价值。见图 3。

表 5 LSTM 模型与 logistic 回归模型预测效能比较

Table 5 Comparison of predictive performance between LSTM model and logistic regression model

模型	AUC(95%CI)	灵敏度(%)	特异度(%)	准确度(%)
LSTM 模型	0.92(0.88~0.97)	87.3	85.9	86.7
logistic 回归模型	0.84(0.78~0.89)	78.5	76.2	77.4
P	<0.001	<0.001	<0.001	<0.001

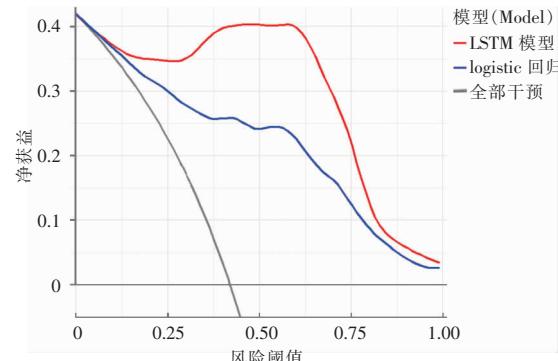


图 3 决策曲线分析图

Figure 3 Decision curve analysis

3 讨论

乳酸清除率是评估脓毒性休克患者组织灌注的关键指标, 其动态变化较单次测量具有更高的预测价值^[8]。乳酸清除率与患者预后密切相关, 研究表明, 乳酸清除率每提高 10%, 死亡风险可降低 11%^[3]。本团队前期通过传统 logistic 回归分析模型发现, AVP 分泌调节异常是脓毒性休克患者不良预后的独立危险因素, 且联合 APACHE II 评分、B 型脑钠肽水平可有效预测患者 28 天预后^[9]。本研究创新性地构建了基于 LSTM 的时序预测模型, 利用入院早期指标预测 24 h 乳酸清除率, 可为临床早期决策提供客观依据。

通过多因素分析, 本研究确定了影响乳酸清除率的关键预测因素。结果显示, APACHE II 评分作为独立预测因素, 反映了患者入院时的病情严重

程度, 与 Tian 等^[10]的研究结果一致。APACHE II 评分是 ICU 最权威、应用最为广泛的危重病情评估系统, 通过评估患者病情及预测病死率, 为制定诊疗方案、优化医疗资源配置及确定治疗时机提供客观依据^[11-12]。入院时乳酸水平反映组织灌注与细胞代谢状态, Khodashahi 等^[13]研究证实初始乳酸水平 $>4 \text{ mmol/L}$ 时, 乳酸清除率下降风险显著增加。此外, 血管活性药物使用剂量与 6 h 液体复苏量的预测价值, 凸显了早期血流动力学管理的重要性, 这与 Dalimonte 等^[14]的系统评价结果一致。

在模型构建过程中, 本研究采用 LASSO 回归进行变量选择优化, 显著降低了过拟合风险。与传统 logistic 回归相比, 优化后的 LSTM 模型预测效能显著提高 ($AUC = 0.922$)。该结果优于 Yee 等^[15]报道的传统静态评分系统, 主要得益于模型对器官功能障碍动态演变过程的时序特征捕捉能力。

本研究比较了乳酸清除率预测与病死率预测的关系和临床意义。相比直接预测病死率的模型, 乳酸清除率预测模型具有以下独特价值: (1)作为过程指标, 能够更早反映治疗反应; (2)具有明确的病理生理机制, 较单纯病死率预测更具解释性; (3)干预策略更加明确, 有助于指导临床个体化治疗方案的调整。该发现与 Hernández 等^[2]基于乳酸清除率引导治疗的临床试验结果相符。

模型验证采用严格的内部验证与多层次的亚组验证, 时间依赖性 AUC 分析显示, 模型在 6、12、24 h 时间窗均表现出良好的预测性能, 尤其是在入院早期 ($\leq 24 \text{ h}$) 预测准确率最高 (87.5%), 这对于临床快速评估和早期干预具有重要价值。同时, 模型在

不同感染源亚组中表现出稳定的预测效能,显示出良好的临床泛化能力。此外,本研究系统考虑了各项指标的生理学意义。通过 LASSO 回归筛选的核心预测因子不仅与 Bataille 等^[16]的多中心研究结果相符,更重要的是反映了疾病进展的不同维度。其中,APACHE II 评分反映整体病情,乳酸水平提示组织灌注状态,血管活性药物使用反映血流动力学稳定性,PCT 水平则指示感染程度。Roy 等^[17]的研究进一步证实这些指标与微循环功能和细胞代谢的密切关联。在模型性能评估方面,本研究的 LSTM 模型相较其他机器学习方法展现出明显优势。相比 Kim 等^[18]的随机森林算法和 Misra 等^[19]的 XG-Boost 模型,本模型实现了更高的预测准确性。Landi 等^[20]的机制研究解释了这一优势,认为 LSTM 的门控机制在识别临床指标长短期相关性方面具有独特作用。最后,在临床实用性方面,本研究特别注重模型的可解释性。通过保持临床指标的原始计量单位并结合注意力机制,模型的可解释性显著优于 Aushev 等^[21]报道的传统深度神经网络。Hah 等^[22]研究证实,这种设计显著提升了临床医生对 AI 辅助决策的接受度,为模型的实际应用奠定了基础。

本研究构建的基于 LSTM 的乳酸清除率动态预测模型展现出优异性能。该模型通过深度学习算法有效整合时序特征,为脓毒性休克患者的预后评估提供了新的方法学思路和临床决策支持工具。本研究存在以下局限性:单中心回顾性设计可能导致选择偏倚,需通过多中心前瞻性研究验证;样本量有限可能影响模型泛化能力,需扩大数据规模;未纳入微循环监测指标(如舌下微循环流速)可能限制预测精度。尽管 LSTM 模型表现优异,但其临床应用需解决实时数据整合问题(如电子病历系统与床旁设备的无缝对接),后续可开发嵌入式临床决策支持模块。

利益冲突:所有作者均声明不存在利益冲突。

参 考 文 献

[1] Gao ZW, Li C, Chen H, et al. Association between diastolic blood pressure during the first 24 h and 28-day mortality in patients with septic shock: a retrospective observational study [J]. Eur J Med Res, 2023, 28(1): 329.

[2] Hernández G, Ospina-Tascón GA, Damiani LP, et al. Effect of a resuscitation strategy targeting peripheral perfusion status vs serum lactate levels on 28-day mortality among patients with septic shock: the ANDROMEDA-SHOCK randomized clinical trial [J]. JAMA, 2019, 321(7): 654–664.

[3] Jouffroy R, Léguillier T, Gilbert B, et al. Prehospital lactate clearance is associated with reduced mortality in patients with septic shock [J]. Am J Emerg Med, 2021, 46: 367–373.

[4] Nazer L, Abusara A, Aloran B, et al. Patient diversity and author representation in clinical studies supporting the surviving sepsis campaign guidelines for management of sepsis and septic shock 2021: a systematic review of citations [J]. BMC Infect Dis, 2023, 23(1): 751.

[5] Macdonald SPJ, Arends G, Fatovich DM, et al. Comparison of PIRO, SOFA, and MEDS scores for predicting mortality in emergency department patients with severe sepsis and septic shock [J]. Acad Emerg Med, 2014, 21(11): 1257–1263.

[6] Hua YX, Zhao ZF, Li RP, et al. Deep learning with long short-term memory for time series prediction [J]. IEEE Commun Mag, 2019, 57(6): 114–119.

[7] 中国医师协会急诊医师分会. 中国急诊感染性休克临床实践指南 [J]. 中国急救医学, 2016, 36(3): 193–206. Emergency Physician Branch of Chinese Medical Association. Chinese emergency septic shock clinical practice guidelines [J]. Chinese Journal of Critical Care Medicine, 2016, 36(3): 193–206.

[8] Takahashi N, Nakada TA, Walley KR, et al. Significance of lactate clearance in septic shock patients with high bilirubin levels [J]. Sci Rep, 2021, 11(1): 6313.

[9] 张维华, 宋昭光, 温思聪, 等. 感染性休克患者 AVP-Time 变化及 AVP 分泌调节异常对预后的预测价值和临床意义 [J]. 中国急救复苏与灾害医学杂志, 2024, 19(3): 322–326. Zhang WH, Song ZG, Wen SC, et al. Prognostic value and clinical significance of AVP-time changes and abnormal regulation of AVP secretion in patients with septic shock [J]. China Journal of Emergency Resuscitation and Disaster Medicine, 2024, 19(3): 322–326.

[10] Tian Y, Yao Y, Zhou J, et al. Dynamic APACHE II score to predict the outcome of intensive care unit patients [J]. Front Med (Lausanne), 2022, 8: 744907.

[11] Rani D, Kumari V, Kumari S. Prognostic value of APACHE II score, sofa score and biomarkers in patients of sepsis and septic shock – a comparative study [J]. Int J Acad Med Pharm, 2023, 5(4): 649–654.

[12] Bloria SD, Chauhan R, Sarna R, et al. Comparison of APACHE II and APACHE IV score as predictors of mortality in patients with septic shock in intensive care unit: a prospective observational study [J]. J Anaesthesiol Clin Pharmacol, 2023, 39(3): 355–359.

[13] Khodashahi R, Sarjamee S. Early lactate area scores and serial blood lactate levels as prognostic markers for patients with septic shock: a systematic review [J]. Infect Dis (Lond), 2020, 52(7): 451–463.

[14] Dalimonte MA, DeGrado JR, Anger KE. Vasoactive agents for adult septic shock: an update and review [J]. J Pharm

Pract, 2020, 33(4): 523–532.

[15] Yee CR, Narain NR, Akmaev VR, et al. A data-driven approach to predicting septic shock in the intensive care unit[J]. Biomed Inform Insights, 2019, 11: 1178222619885147.

[16] Bataille B, de Selle J, Moussot PE, et al. Machine learning methods to improve bedside fluid responsiveness prediction in severe sepsis or septic shock: an observational study[J]. Br J Anaesth, 2021, 126(4): 826–834.

[17] Roy TK, Secomb TW. Effects of impaired microvascular flow regulation on metabolism-perfusion matching and organ function[J]. Microcirculation, 2021, 28(3): e12673.

[18] Kim J, Chang H, Kim D, et al. Machine learning for prediction of septic shock at initial triage in emergency department [J]. J Crit Care, 2020, 55: 163–170.

[19] Misra D, Avula V, Wolk DM, et al. Early detection of septic shock onset using interpretable machine learners[J]. J Clin Med, 2021, 10(2): 301.

[20] Landi F, Baraldi L, Cornia M, et al. Working memory connections for LSTM[J]. Neural Netw, 2021, 144: 334–341.

[21] Aushev A, Ripoll VR, Vellido A, et al. Feature selection for

the accurate prediction of septic and cardiogenic shock ICU mortality in the acute phase[J]. PLoS One, 2018, 13(11): e0199089.

[22] Hah H, Goldin DS. How clinicians perceive artificial intelligence-assisted technologies in diagnostic decision making: mixed methods approach[J]. J Med Internet Res, 2021, 23(12): e33540.

(本文编辑:陈玉华)

本文引用格式:宋昭光,吴平宇,温思聪,等.基于机器学习的脓毒症休克患者乳酸清除率动态预测模型开发与验证[J].中国感染控制杂志,2025,24(8):1097–1105. DOI:10.12138/j.issn.1671-9638.20252158.

Cite this article as: SONG Zhaoguang, WU Pingyu, WEN Sicong, et al. Development and validation of a machine learning-based dynamic prediction model for lactate clearance rate in patients with septic shock[J]. Chin J Infect Control, 2025, 24(8): 1097–1105. DOI: 10.12138/j.issn.1671–9638.20252158.