

基于机器学习构建急性胰腺炎患者并发急性呼吸窘迫综合征风险预测模型

李玉倩¹ 李文哲² 王毅² 王轶希³ 于湘友²

新疆医科大学第一附属医院¹麻醉科,²重症医学科,³脊柱微创与精准骨科,新疆维吾尔自治区乌鲁木齐 830054

通信作者:于湘友,Email:yu2796@163.com

【摘要】目的 通过机器学习,基于临床可及指标,构建急性胰腺炎(AP)患者并发急性呼吸窘迫综合征(ARDS)的风险预测模型,为临床医师提供决策支持。**方法** 基于美国重症监护医学信息数据库Ⅳ(MIMIC Ⅳ) 3.1 数据库开展回顾性队列分析,筛选 AP 患者并根据 ARDS 发生与否进行分组。利用 Logistic 回归筛选特征变量,并基于特征变量子集构建包括 Logistic 回归、梯度提升机(GBM)、自适应提升法(AdaBoost)、随机森林、K-最近邻(KNN)、神经网络、极度梯度提升(XGBoost)和支持向量机(SVM)在内的 8 种预测模型,针对各模型特点进行超参数优化,并综合评价各模型的性能,选择最佳模型进行夏普利加法解释(SHAP)分析。**结果** 共纳入 1 553 例 AP 患者,其中 616 例(39.67%)入住重症监护病房(ICU)后发生 ARDS。Logistic 回归(后退法)筛选出 7 个关键特征变量:急性生理学与慢性健康状况评分Ⅱ(APACHE Ⅱ)评分、序贯器官衰竭评分(SOFA)评分、呼吸频率(RR)、血肌酐(SCr)、血糖、脉搏血氧饱和度(SpO₂)及白蛋白(Alb)为影响 AP 患者并发 ARDS 的独立危险因素[优势比(OR)和 95% 可信区间(95%CI)分别为 1.061(1.030~1.093)、1.185(1.130~1.242)、1.047(1.007~1.090)、0.962(0.949~0.974)、0.587(0.459~0.752)、1.559(1.339~1.814)、1.002(1.000~1.004), P 值分别为 <0.001、<0.001、0.021、<0.001、<0.001、<0.001、0.049]。基于 8 种机器学习算法构建的预测模型中,XGBoost 模型和 AdaBoost 模型表现最佳,受试者工作特征曲线(ROC 曲线)下面积(AUC)及 95%CI 分别为 0.855(0.827~0.883)、0.844(0.815~0.872)。测试集中,XGBoost 模型和 AdaBoost 模型表现稳定,AUC 及 95%CI 分别为 0.838(0.795~0.881)、0.832(0.789~0.876),其预测效能及稳定性优于其他算法。SHAP 分析显示,影响模型性能的主要因素包括:APACHE Ⅱ评分、SOFA 评分、SCr 最大值、Alb 最小值及血糖最大值。**结论** 基于临床常见指标及评分系统构建的 XGBoost 模型和 AdaBoost 模型有良好的预测效能和稳定性,可有效识别并发 ARDS 高风险的 AP 患者,助力临床医生进行早期识别、分层监测和个性化干预。

【关键词】 急性胰腺炎; 急性呼吸窘迫综合征; 机器学习; 预测模型; 精准医疗

基金项目:国家自然科学基金(82460372);新疆维吾尔自治区研究生科研创新项目(XJ2025G160)

DOI: 10.3969/j.issn.1008-9691.2026.01.011

Development and validation of machine learning model for early prediction of acute respiratory distress syndrome in acute pancreatitis patients

Li Yuqian¹, Li Wenzhe², Wang Yi², Wang Yixi³, Yu Xiangyou²

¹Department of Anesthesiology, ²Department of Critical Care Medicine, ³Department of Minimally Invasive Spine and Precision Orthopedics, the First Affiliated Hospital of Xinjiang Medical University, Urumqi 830054, Xinjiang Uyghur Autonomous Region, China

Corresponding author: Yu Xiangyou, Email: yu2796@163.com

【Abstract】 Objective To develop a machine learning based risk prediction model for acute respiratory distress syndrome (ARDS) in patients with acute pancreatitis (AP) using routinely available clinical indicators, thereby providing decision support for clinicians. **Methods** A retrospective cohort study was performed based on the Medical Information Mart for Intensive Care IV (MIMIC IV) 3.1 database. Patients with AP were identified and stratified according to the occurrence of ARDS. Logistic regression analysis was used to screen candidate predictors. Based on the selected feature subset, 8 machine learning models were constructed, including Logistic regression, gradient boosting machine (GBM), adaptive boosting (AdaBoost), random forest, K-nearest neighbors (KNN), neural network, extreme gradient boosting (XGBoost), and support vector machine (SVM). Hyperparameters were optimized according to the characteristics of each algorithm. Model performance was comprehensively evaluated to identify the optimal model, which was subsequently interpreted using Shapley Additive Explanations (SHAP) analysis. **Results** A total of 1 553 patients with AP were included, among whom 616 (39.67%) developed ARDS after intensive care unit (ICU) admission. Logistic regression with backward elimination identified 8 key predictors associated with ARDS in patients with AP: acute physiology and chronic health evaluation II (APACHE II) score, sequential organ failure assessment (SOFA) score, respiratory rate (RR), serum creatinine (SCr), blood glucose, pulse oxygen saturation (SpO₂), and albumin (Alb) [odds ratios (OR) with 95% confidence intervals (95%CI) were 1.061 (1.030–1.093), 1.185 (1.130–1.242), 1.047 (1.007–1.090), 0.962 (0.949–0.974), 0.587 (0.459–0.752), 1.559 (1.339–1.814), and 1.002 (1.000–1.004)], $P < 0.001$, < 0.001 , < 0.001 , < 0.001 , < 0.001 , < 0.001 , < 0.001 , 0.049 respectively. Among the 8 machine learning models, the best performance was achieved

by the XGBoost and AdaBoost models, with areas under the receiver operator characteristic curve (AUC) and 95%CI of 0.855 (0.827–0.883) and 0.844 (0.815–0.872), respectively. In the test set, these models maintained stable performance, with AUC values of 0.838 (95%CI: 0.795–0.881) for XGBoost and 0.832 (95%CI: 0.789–0.876) for AdaBoost, and both outperforming the other algorithms in both predictive accuracy and stability. SHAP analysis indicated that the most influential predictors included APACHE II score, SOFA score, maximum serum creatinine, minimum albumin, and maximum blood glucose. **Conclusions** The XGBoost and AdaBoost models constructed using routinely available clinical indicators and scoring systems demonstrated good predictive performance and stability. These models may effectively identify AP patients at high risk of developing ARDS, facilitating early detection, risk stratification, and personalized clinical intervention.

【Key words】 Acute pancreatitis; Acute respiratory distress syndrome; Machine learning; Prediction model; Precision medicine

Fund program: National Natural Science Foundation of China (82460372); Xinjiang Uygur Autonomous Region Graduate Student Scientific Research Innovation Project (XJ2025G160)

DOI: 10.3969/j.issn.1008-9691.2026.01.011

急性胰腺炎 (acute pancreatitis, AP) 是一种常见的急性炎症性消化系统疾病,其发病率近年来仍呈持续上升趋势^[1-2]。约 20%~30% 的 AP 患者在早期可继发全身炎症反应综合征 (systemic inflammatory response syndrome, SIRS) 及血管内皮功能障碍,进而导致多器官功能障碍等严重并发症^[3]。Schepers 等^[4]的研究显示,约 38% 的 AP 患者出现器官功能障碍,其中以急性呼吸窘迫综合征 (acute respiratory distress syndrome, ARDS) 最为常见,占比高达 92%,且相关病死率高达 37%。ARDS 是一种由直接或间接肺损伤引发的严重呼吸功能障碍,其病理机制主要为炎症性肺水肿所致的顽固性低氧血症,具有发病急、进展快、病死率高的特点^[5-6]。肺炎、吸入性损伤等是导致 ARDS 的主要直接原因,而脓毒症、严重烧伤和 AP 则是导致 ARDS 最常见的间接原因^[7]。尽管近年来各种疾病的预防和治疗策略取得了进展,但 ARDS 的发病率和病死率未见显著下降,仍是重症医学领域亟待解决的重要难题^[8]。作为 AP 最早出现的器官功能障碍之一,ARDS 的发生机制至今仍未完全阐明。既往有研究表明,AP 相关 ARDS 的发病机制与其他病因引起的 ARDS 在本质上并无显著差异^[9]。因此,如何在临床上早期有效预测和识别 ARDS 的发生风险,并结合现代呼吸支持技术选择适宜的干预策略,成为遏制其发展、改善预后至关重要的研究方向。然而,当前临床实践中仍缺乏精准、有效的预测和预防手段^[10-11]。随着电子健康记录 (electronic health record, EHR)、重症监护数据和大规模临床数据库 [如美国重症监护医学信息数据库 IV (Medical Information Mart for Intensive Care IV, MIMIC IV) 等] 的不断积累与完善,机器学习 (machine learning, ML) 技术在医疗决策及诊疗中的应用与日俱增。相比传统统计模型,

机器学习方法能够处理大规模、多维且复杂的临床数据,更好地识别潜在模式,从而实现对患者未来风险的预测。相关研究显示,相比传统的危重疾病评分系统,机器学习模型在改善患者预后及并发症的预测准确性方面具有明显优势^[12]。鉴于 ARDS 的临床异质性,机器学习通过整合分析包括生命体征、实验室检查和影像学在内的大量临床数据,在重症患者 ARDS 的早期识别方面展现出广阔前景和显著优势。本研究旨在通过机器学习算法,构建 AP 患者并发 ARDS 风险的预测模型。基于 MIMIC IV 数据库中的大规模临床数据,筛选出与 AP 患者 ARDS 发生密切相关的临床特征和实验室指标,构建预测其并发 ARDS 风险的机器学习模型,以期能够助力临床医师早期识别此类高危患者并调整诊疗策略,进而改善其临床预后。

1 资料与方法

1.1 研究数据来源:本研究为一项基于 MIMIC IV 3.1 的回顾性队列研究,构建了 AP 患者并发 ARDS 的风险预测模型。MIMIC IV 数据库由麻省理工学院开发和维护,涵盖了贝斯以色列女执事医疗中心 (Beth Israel Deaconess Medical Center, BIDMC) 重症监护病房 (intensive care unit, ICU) 大量高质量的医疗记录,是目前国际上应用广泛的大规模重症医学公共数据库之一^[13]。本研究团队中有一位成员已获得数据库访问权限 (认证编号: 57264471),并使用 PostgreSQL16.0 和 Navicate Premium17.0 软件完成相关数据提取工作。

1.1.1 纳入标准:根据国际疾病分类 (international classification of diseases, ICD-9、ICD-10) 诊断编码,筛选首次入住 ICU 的 AP 患者作为研究对象^[14]。

1.1.2 排除标准:① 年龄 < 18 岁; ② 非首次住院或非首次入住 ICU; ③ ICU 住院时间 < 24 h; ④ 诊

断为胰腺肿瘤；⑤ 临床信息缺失或实验室检验数据不完整。患者纳入分析见图 1。



1.2 数据筛选与特征提取：本研究提取的变量数据涵盖以下几类：① 人口统计学资料：包括性别、年龄、身高及体质量；② 合并基础疾病情况：包括慢性肾脏疾病、慢性肺部疾病、充血性心力衰竭、风湿性疾病、糖尿病等；③ 疾病严重程度评分；④ 生命体征：包括心率、收缩压、舒张压、平均动脉压 (mean arterial pressure, MAP)、呼吸频率 (respiratory rate, RR)、体温、脉搏血氧饱和度 (pulse oxygen saturation, SpO₂)；⑤ 实验室检查指标：包括血常规、凝血功能、肝肾功能、血气分析等。研究自 AP 患者入住 ICU 起开始随访至转出或死亡，期间筛查所有符合 ARDS 诊断标准的患者。基于入住 ICU 24 h 后监测数据完善患者 ARDS 诊断的判定，结局定义为发生 ARDS^[6]。所有作为预测变量的实验室检测指标和疾病严重程度评分均从患者入住 ICU 后 24 h 内生成的数据中提取。

1.3 统计分析与模型构建：使用 R 4.4.3 版进行统计分析。按 ARDS 发生与否分组，删除缺失值超过 20% 的变量。对计量资料行正态性检验后，正态分布的计量数据以均数 ± 标准差 ($\bar{x} \pm s$) 表示，组间比较采用独立样本 *t* 检验。非正态分布的计量数据以中位数 (四分位间距) [$M(Q_R)$] 表示，两组间的比较采用 Wilcoxon 秩和检验。计数数据以例 (%) 表示，组间比较采用 χ^2 检验。为提升模型变量的临床意

义可解释性，采用 Logistic 回归分析筛选关键特征，并基于特征变量子集建立包括 Logistic 回归、梯度提升机 (gradient boosting machine, GBM)、自适应提升法 (adaptive boosting, AdaBoost)、随机森林、K-最近邻 (K-nearest neighbor, KNN)、神经网络、极度梯度提升 (extreme gradient boosting, XGBoost) 及支持向量机 (support vector machine, SVM) 在内的 8 种预测模型，基于各模型特点进行超参数优化。以受试者工作特征曲线 (receiver operator characteristic curve, ROC 曲线)、决策曲线 (decision curve analysis, DCA)、校准曲线及准确率、敏感度、特异度、精确率、F1 分数评价模型性能，选择最优模型并进行夏普利加法解释 (Shapley Additive Explanations, SHAP) 分析。*P* < 0.05 提示差异有统计学意义。

2 结果

2.1 纳入患者的基线资料 (表 1 ~ 2)：本研究共纳入 1 553 例 AP 患者，其中 616 例 (39.67%) 入住 ICU 后并发 ARDS。与非 ARDS 组比较，ARDS 组体质量、查尔森合并症指数、序贯器官功能衰竭评分 (sequential organ failure assessment, SOFA)、急性生理学与慢性健康状况评分 II (acute physiology and chronic health evaluation II, APACHE II)、Logistic 器官功能障碍评分 (Logistic organ dysfunction score, LODS)、基础疾病 (包括慢性肾脏疾病、慢性肺疾病、充血性心力衰竭)、生命体征 (包括心率最大值、RR 最大值、体温最大值、SpO₂ 最小值)、总住院时间、ICU 住院时间、使用机械通气比例、28 d 病死率、白细胞计数 (white blood cell count, WBC)、中性粒细胞计数 (neutrophil count, NEU) 最大值、血尿素氮 (blood urea nitrogen, BUN)、血肌酐 (serum creatinine, SCr)、血乳酸 (lactic acid, Lac)、血糖、总胆红素 (total bilirubin, TBil)、国际标准化比值 (international normalized ratio, INR)、凝血酶原时间 (prothrombin time, PT)、活化部分凝血活酶时间 (activated partial thromboplastin time, APTT) 最大值均明显升高，而收缩压最小值、舒张压最小值、MAP 最小值、白蛋白 (albumin, Alb) 最小值、pH 值和动脉血氧分压 (arterial partial pressure of oxygen, PaO₂)、剩余碱最小值均明显降低 (均 *P* < 0.05)。两组性别、年龄、身高、种族、风湿性疾病、糖尿病、血小板计数 (platelet count, PLT) 最小值、动脉血二氧化碳分压 (arterial partial pressure of carbon dioxide, PaCO₂) 最大值比较差异均无统计学意义 (均 *P* > 0.05)。

表 1 基于机器学习构建 AP 患者并发 ARDS 风险预测模型纳入患者的基线资料

指标	全体 (n=1 553)	非 ARDS 组 (n=937)	ARDS 组 (n=616)	χ^2/t Z 值	P 值	指标	全体 (n=1 553)	非 ARDS 组 (n=937)	ARDS 组 (n=616)	χ^2/t t 值	P 值
性别 [例(%)]				1.536	0.138	合并基础疾病 [例(%)]					
男性	859 (55.3)	533 (56.9)	326 (52.9)			慢性肾脏疾病	253 (16.3)	138 (14.7)	115 (18.7)	-2.059	0.047
女性	694 (44.7)	404 (43.1)	290 (47.1)			慢性肺部疾病	313 (20.2)	172 (18.4)	141 (22.9)	-2.180	0.035
年龄 (岁, $\bar{x} \pm s$)	60.5 ± 17.4	60.6 ± 18.1	60.3 ± 16.3	0.301	0.759	充血性心力衰竭	296 (19.1)	159 (17.0)	137 (22.2)	-2.591	0.012
身高 (cm, $\bar{x} \pm s$)	169.0 ± 10.3	169.0 ± 0.4	169.0 ± 10.0	0.041	0.967	风湿性疾病	52 (3.3)	29 (3.1)	23 (3.7)	-0.672	0.589
体质量 [kg, $M(Q_R)$]	80.5 (28.6)	79.0 (25.0)	82.7 (26.7)	-3.986	<0.001	糖尿病	412 (26.5)	256 (27.3)	156 (25.3)	0.876	0.416
种族 [例(%)]				0.514	0.310	生命体征 ($\bar{x} \pm s$)					
白种人	1 074 (69.2)	637 (68.0)	437 (70.9)			心率最大值 (次/min)	111.0 ± 22.5	109.0 ± 21.6	114.0 ± 23.5	-4.458	<0.001
黑种人	199 (12.8)	132 (14.1)	67 (10.9)			收缩压最小值 (mmHg)	95.0 ± 14.9	97.1 ± 14.8	91.8 ± 14.5	6.998	<0.001
黄种人	72 (4.6)	42 (4.5)	30 (4.9)			舒张压最小值 (mmHg)	50.4 ± 10.6	51.3 ± 10.8	49.1 ± 10.2	3.996	<0.001
其他人种	208 (13.4)	126 (13.4)	82 (13.3)			MAP 最小值 (mmHg)	62.7 ± 11.7	63.6 ± 11.7	61.2 ± 11.5	4.012	<0.001
查尔森合并症指数 (分, $\bar{x} \pm s$)	3.89 ± 2.55	3.73 ± 2.53	4.13 ± 2.57	-3.023	<0.001	RR 最大值 (次/min)	28.80 ± 5.47	28.30 ± 5.46	29.50 ± 5.41	-4.506	<0.001
SOFA 评分 (分, $\bar{x} \pm s$)	4.98 ± 3.80	4.09 ± 3.18	6.34 ± 4.23	-11.898	<0.001	体温最高值 (°C)	37.4 ± 0.6	37.4 ± 0.5	37.5 ± 0.6	-4.986	<0.001
APACHE II 评分 (分, $\bar{x} \pm s$)	16.00 ± 6.24	14.60 ± 5.73	18.10 ± 6.38	-11.447	<0.001	SpO ₂ 最小值	0.920 ± 0.031	0.924 ± 0.029	0.913 ± 0.033	6.266	<0.001
LODS 评分 (分, $\bar{x} \pm s$)	4.42 ± 3.05	3.57 ± 2.60	5.72 ± 3.23	-14.441	<0.001	使用机械通气 [例(%)]	572 (36.8)	204 (21.8)	368 (59.7)	-16.433	<0.001
						总住院时间 (d, $\bar{x} \pm s$)	13.1 ± 10.5	10.9 ± 9.1	16.4 ± 11.5	-10.617	<0.001
						ICU 住院时间 (d, $\bar{x} \pm s$)	3.79 ± 4.15	2.70 ± 2.78	5.46 ± 5.21	-13.550	<0.001
						28 d 病死率 [% (例)]	9.9 (153)	6.6 (62)	14.8 (91)	-5.321	<0.001

注: 1 mmHg ≈ 0.133 kPa

表 2 基于机器学习构建 AP 患者并发 ARDS 风险预测模型纳入患者入住 ICU 后 24 h 内实验室检查指标 ($\bar{x} \pm s / M(Q_R)$)

指标	全体 (n=1 553)	非 ARDS 组 (n=937)	ARDS 组 (n=616)	t 值	P 值	指标	全体 (n=1 553)	非 ARDS 组 (n=937)	ARDS 组 (n=616)	t / Z 值	P 值
WBC 最大值 ($\times 10^9/L$)	14.50 ± 6.35	13.90 ± 6.19	15.50 ± 6.47	-4.803	<0.001	PaCO ₂ 最大值 (mmHg)	43.50 ± 8.13	43.20 ± 8.10	44.00 ± 8.16	-1.903	0.057
LYM 最小值 ($\times 10^9/L$)	1.07 ± 0.61	1.09 ± 0.62	1.03 ± 0.61	2.148	0.032	剩余碱最小值 (mmol/L)	-3.17 ± 3.94	-2.77 ± 3.73	-3.78 ± 4.16	4.983	<0.001
NEU 最大值 ($\times 10^9/L$)	11.20 ± 5.51	10.80 ± 5.46	11.80 ± 5.54	-3.380	<0.001	Lac 最大值 (mmol/L)	2.79 ± 2.45	2.67 ± 2.24	2.96 ± 2.73	-2.217	<0.001
Hb 最小值 (g/L)	103.0 ± 23.5	104.0 ± 23.1	101.0 ± 24.0	2.510	0.012	血糖最大值 (mmol/L)	9.21 ± 3.69	8.88 ± 3.52	9.71 ± 3.87	-4.389	<0.001
PLT 最小值 ($\times 10^9/L$)	193 ± 115	195 ± 113	190 ± 117	0.689	0.491	血糖最小值 (mmol/L)	6.38 ± 1.63	6.27 ± 1.58	6.49 ± 1.69	-2.937	0.004
BUN 最大值 (mmol/L)	8.83 ± 5.70	8.15 ± 5.48	9.83 ± 5.70	-5.771	<0.001	TBil 最大值 (μmol/L)	17.1 (32.5)	17.1 (29.1)	18.8 (32.5)	-2.316	0.024
SCr 最大值 (μmol/L)	123.76 ± 84.86	113.15 ± 76.11	140.56 ± 94.59	-6.229	<0.001	INR 最大值	1.43 ± 0.42	1.40 ± 0.40	1.48 ± 0.44	-3.374	<0.001
Alb 最小值 (g/L)	30.50 ± 5.86	31.40 ± 5.85	29.20 ± 5.60	7.669	<0.001	PT 最大值 (s)	15.8 ± 4.4	15.6 ± 4.5	16.1 ± 4.4	-2.225	0.026
pH 值最小值	7.32 ± 0.11	7.34 ± 0.10	7.31 ± 0.11	5.311	<0.001	APTT 最大值 (s)	36.1 ± 14.7	35.1 ± 13.7	37.6 ± 16.0	-3.343	<0.001
PaO ₂ 最小值 (mmHg)	98.0 ± 36.3	102.0 ± 37.4	92.7 ± 33.7	4.718	<0.001						

注: 1 mmHg ≈ 0.133 kPa

2.2 Logistic 回归分析筛选特征变量 (表 3; 图 2): 将单因素分析 $P < 0.05$ 的变量纳入 Logistic 回归模型 (后退法) 进行特征筛选, 最终确定 7 个关键特征变量 APACHE II 评分、SOFA 评分、RR、SCr 及血糖水平升高, SpO₂ 及 Alb 水平降低均为 AP 患者并发 ARDS 的独立危险因素 (均 $P < 0.05$)。

表 3 多因素 Logistic 回归分析影响 AP 患者并发 ARDS 的危险因素

变量	β 系数	s_e	χ^2 值	P 值	OR (95%CI)
APACHE II 评分	0.059	0.015	15.735	<0.001	1.061 (1.030 ~ 1.093)
SOFA 评分	0.170	0.024	49.403	<0.001	1.185 (1.130 ~ 1.242)
RR 最大值	0.046	0.020	5.313	0.021	1.047 (1.007 ~ 1.090)
SpO ₂ 最小值	-0.039	0.006	36.490	<0.001	0.962 (0.949 ~ 0.974)
Alb 最小值	-0.533	0.126	17.874	<0.001	0.587 (0.459 ~ 0.752)
SCr 最大值	0.444	0.077	32.809	<0.001	1.559 (1.339 ~ 1.814)
血糖最大值	0.002	0.001	3.651	0.049	1.002 (1.000 ~ 1.004)

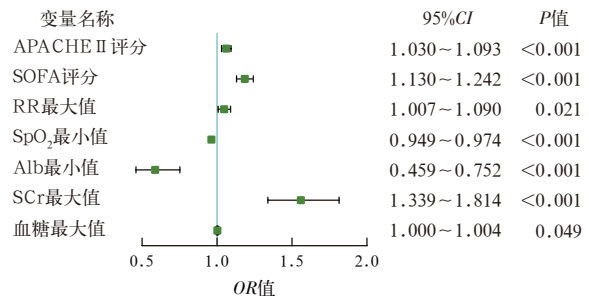


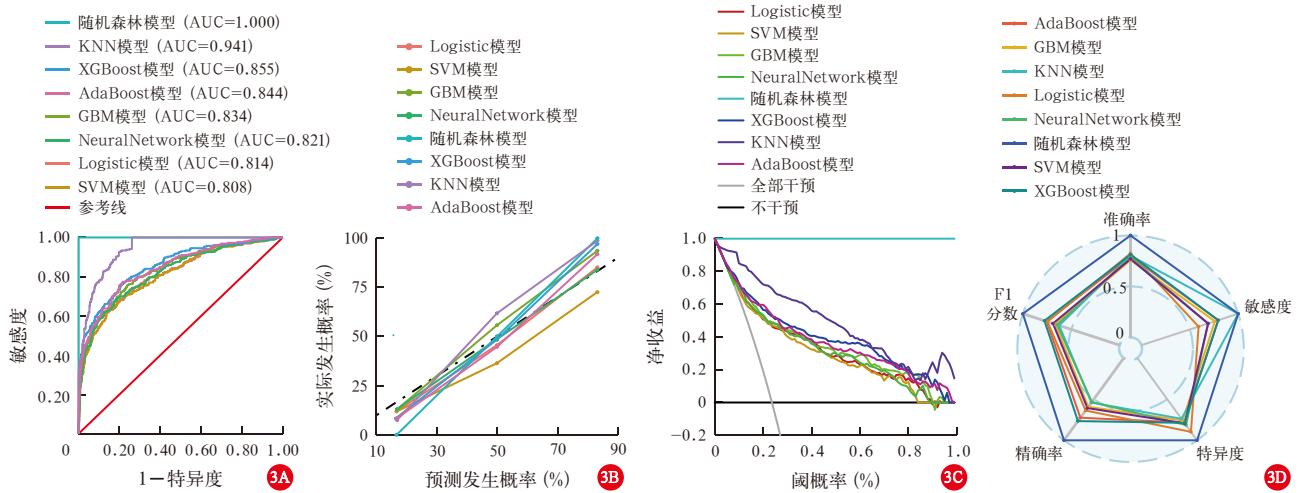
图 2 基于 Logistic 回归分析方法筛选的特征变量

2.3 模型构建 (图 3): 使用 R 软件按 7 : 3 比例将患者随机分为训练集和测试集。将多因素 Logistic 回归分析筛选出的 7 个独立危险因素纳入 8 种机器学习算法, 基于训练集构建预测 AP 患者 ARDS 发生风险模型。8 种模型的 ROC 曲线分析显示,

尽管经过超参数调优,但随机森林和KNN仍存在显著过拟合风险,XGBoost模型和AdaBoost模型具有良好的预测表现,ROC曲线下面积(area under the curve, AUC)及95%可信区间(95% confidence interval, 95%CI)分别为0.855(0.827~0.883)、0.844(0.815~0.872)。8个模型的校准曲线如图3B所示,为了解各模型的预测可靠性提供了重要依据。XGBoost模型和AdaBoost模型的预测概率与观察结果之间一致性较好。关于临床实用性,DCA曲线分析中XGBoost模型和AdaBoost模型在80%~90%净获益区间范围内显示出良好的临床效能。图3D展示了通过多项指标综合评价模型的预测性能。

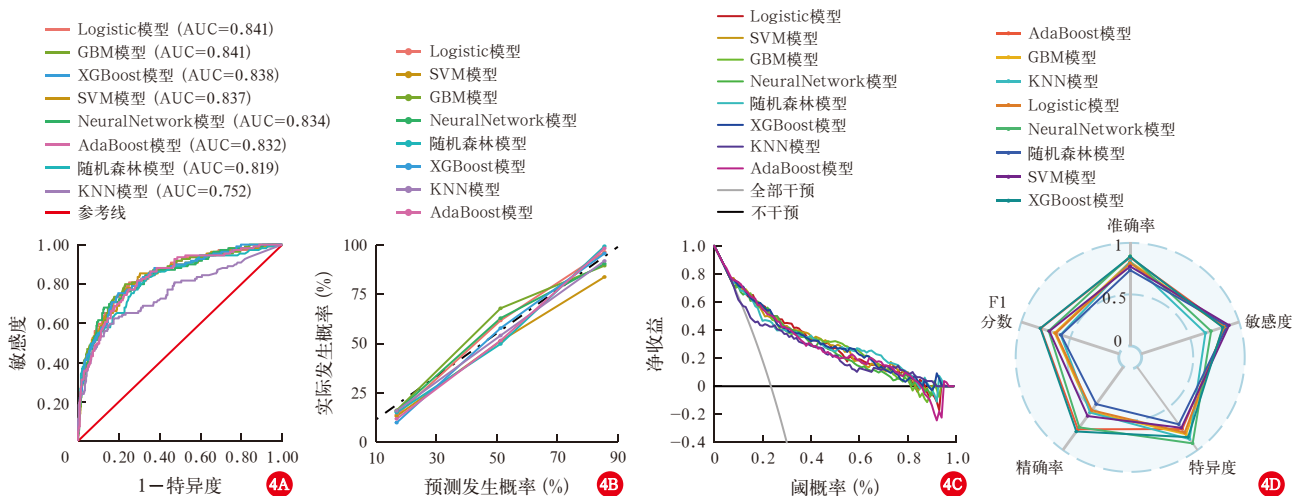
2.4 模型的评价(图4):在测试集中,XGBoost模型和AdaBoost模型表现稳定,AUC及95%CI分别为0.838(0.795~0.881)、0.832(0.789~0.876),预测效能及稳定性优于其他算法。在综合比较多项评价指标后,XGBoost模型和AdaBoost模型被确定为预测AP患者并发ARDS的最佳预测模型。

2.5 模型可视化(图5):通过SHAP算法分析了XGBoost模型和AdaBoost模型的可解释性。特征重要性条形图(图5A~5B)和蜂群图(图5C~5D)分别显示了特征对模型的贡献程度。如针对APACHE II评分,SHAP均值越大,变量值越高,表明该评分升高会显著增加罹患ARDS的风险。最终筛选出



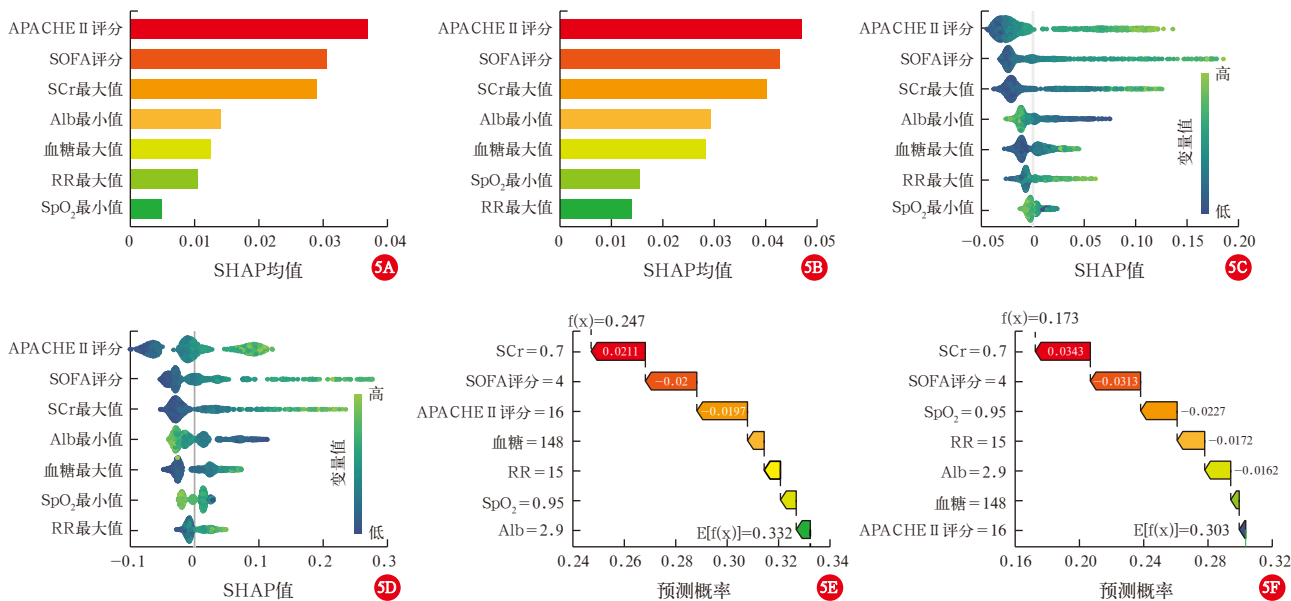
注:A为训练集的ROC曲线,B为训练集的决策曲线,C为训练集的校准曲线,D为归一化雷达图展示各模型在训练集中的准确率、敏感性、特异性、精确率和F1分数

图3 训练集中各机器学习模型的预测性能



注:A为测试集的ROC曲线,B为测试集的决策曲线,C为测试集的校准曲线,D为归一化雷达图展示各模型在测试集的准确率、敏感性、特异性、精确率和F1分数

图4 测试集中各机器学习模型的预测性能



注: A 为条形图显示 XGBoost 模型中 7 个重要变量的平均 SHAP 值, B 为条形图显示 AdaBoost 模型中 7 个重要变量的平均 SHAP 值, C 为蜂群图可视化 XGBoost 模型中 7 个关键变量的 SHAP 值分布, D 为蜂群图可视化 AdaBoost 模型中 7 个关键变量的 SHAP 值分布, E 为 SHAP 瀑布图展示某病例的 XGBoost 模型临床应用, F 为 SHAP 瀑布图展示某病例的 AdaBoost 模型临床应用

图 5 XGBoost 模型和 AdaBoost 模型的 SHAP 分析结果

影响 AP 患者并发 ARDS 风险的前 5 个因素是: APACHE II 评分、SOFA 评分、SCr 最大值、Alb 最小值及血糖最大值。此外,瀑布图通过临床实例解析 XGBoost 模型和 AdaBoost 模型的预测价值(图 5E ~ 5F)。

3 讨论

AP 的临床管理取决于疾病的严重程度,其诊断、病情严重程度、并发症预测、预后评估通常基于影像学技术(如计算机断层扫描、磁共振成像和超声)以及评分系统(包括 Ranson 评分等)^[15]。尽管这些方法在临床管理中占据核心地位,但其在准确性、可重复性、临床实用性及经济成本方面存在一定的局限性^[16]。作为人工智能的重要分支之一,机器学习可通过分析大规模数据集识别潜在模式,具有自学习、自适应、高容错性和高效性等优势^[17]。在医学数据分析中,机器学习能有效处理复杂的高维数据集,识别潜在的疾病风险因素,构建精准的预测模型,协助临床医生进行疾病诊断、治疗决策及预后评估^[18]。

本研究基于真实世界的临床数据,运用 8 种机器学习算法构建了 AP 并发 ARDS 风险的预测模型,并通过 SHAP 算法对最优模型进行了解释性分析。结果显示, XGBoost 模型和 AdaBoost 模型均表现出良好的预测性能,能在实际临床应用中提供显著的

净获益。两个模型仅纳入了与 AP 患者并发 ARDS 密切相关的 7 个变量,包括 APACHE II 评分、SOFA 评分、SCr 最大值、Alb 最小值、血糖最大值、RR 最大值和 SpO₂ 最小值,这些均为患者早期的静态数据,具有良好的临床易获取性,有利于模型的推广与实施。此外, SHAP 分析揭示了每个变量对模型预测结果的具体影响,直观展示了各临床指标对患者预后的积极或消极影响,增强了模型的可解释性。因此,本研究构建的预测模型不仅能有效识别 AP 患者并发 ARDS 的关键风险因素,还克服了传统机器学习模型“黑箱”的缺陷,提高了临床医生运用此模型进行临床决策的信心和可行性^[19]。根据 SHAP 分析结果, APACHE II 评分和 SOFA 评分是模型中最重要的预测特征变量,而这两项评分也是 ICU 内最常用的疾病严重程度评分系统^[20]。有研究表明, APACHE II 评分的升高与 AP 疾病严重程度存在显著的正相关性^[21]。此外,有研究显示, SOFA 评分在预测 AP 患者院内病死率、多器官功能障碍发生率及 ICU 入住需求方面均优于其他指标,因此能有效进行早期病情分层和预后预测^[22]。APACHE II 评分与 SOFA 评分联合应用可快速识别 AP 的高危患者,并对治疗效果与风险变化进行动态评估^[23]。同时,这两项评分也是识别继发 ARDS 高风险患者的重要工具。一项荟萃分析表明, APACHE II 评分、SOFA

评分、肺源性脓毒症、吸烟、胰腺炎和 C-反应蛋白是脓症患者发生 ARDS 的重要预测指标^[24]。

急性肾损伤 (acute kidney injury, AKI) 是 AP 的严重并发症之一,不同研究报告其发病率差异较大,约为 7.9%~42.0%,而入住 ICU 患者的 AKI 发生率可高达 70% 以上^[25]。AKI 的发生不仅反映肾功能下降,也可能通过液体潴留、炎症反应等机制促进肺损伤与 ARDS 的发生发展。因此入院时较高的基线 SCr 水平,或继发 AKI 时 SCr 的升高,均与 ARDS 的发生和死亡风险显著相关^[26]。尽管大多数研究更关注于 AKI 的发生及整体预后,但已有研究证实 SCr 升高可作为 ARDS 的预测指标^[27]。此外,也有研究表明,SCr 有助于区分低炎症与高炎症 ARDS 亚型,并与其预后密切相关^[28]。

Alb 是维持血浆胶体渗透压的关键,低白蛋白血症 (≤ 30 g/L) 时血浆胶体渗透压下降导致肺泡毛细血管通透性增强,是 ARDS 发生的重要机制之一。此外,Alb 在体内具有抗氧化、抗炎、保护血管内皮屏障及物质载运等功能,低白蛋白血症时上述生理作用受损可能加剧炎症相关损伤^[29]。一项国际多中心前瞻性研究显示,AP 患者入院时低白蛋白血症的患病率为 19.0%,住院期间升高至 35.7%,且低白蛋白血症与疾病严重程度、病死率及器官衰竭风险呈剂量依赖性增加,并与住院时间延长相关,是疾病严重程度和病死率的独立危险因素^[30]。另一项研究评估了 AP 患者入院后 24 h 内 Alb 水平与预后结局及病死率的相关性,结果显示,低白蛋白血症与持续性全身炎症反应综合征及胸腔积液发生相关,在预测持续性器官功能衰竭和病死率方面具有重要临床意义^[31]。

另外,血糖水平升高是 AP 患者的常见表现,与疾病严重程度、并发症发生率及病死率密切相关。一项国际前瞻性队列研究显示,住院期间血糖水平与 AP 严重程度呈剂量依赖关系^[11]。AP 引发的炎症因子释放及氧化应激可导致微循环功能障碍和胰腺外器官损伤,而高血糖则可能加剧炎症和氧化损伤,进一步增加肺泡毛细血管的通透性并诱发 ARDS。此外,三酰甘油葡萄糖指数 (triglyceride-glucose index, TyG),作为反映胰岛素抵抗的指标,与住院期间急性呼吸衰竭 (包括 ARDS) 的发生显著相关,提示血糖及其他营养代谢指标在评估 AP 患者呼吸功能障碍风险时亦应予以重视^[32]。

RR 及 SpO₂ 是反映肺通气功能和氧合状态的

直接指标。研究表明,RR 升高是 ARDS 早期常见体征之一^[33]。一项基于急诊 ICU 数据库的轨迹建模研究显示,持续 RR 升高的 ARDS 患者 28 d 病死率明显高于 RR 较低者,提示 RR 加快可作为 ARDS 高风险的预警信号^[34]。此外,浅快呼吸指数 (rapid shallow breathing index, RSBI) 通过 RR 与潮气量的比值,反映了肺顺应性下降和通气效率的恶化程度,在 ARDS 进程中具有重要的参考价值。研究还表明,SpO₂/吸入氧浓度 (fractional of inspired oxygen, FiO₂) 比值可作为 PaO₂/FiO₂ 比值的可靠替代指标,尤其在无法获取动脉血气的环境中更具实用性^[35]。2023 ARDS 全球新定义专家共识建议在诊断 ARDS 时纳入 SpO₂/FiO₂ 比值^[6, 36]。RR 与 SpO₂ 的监测为 ARDS 的早期识别和风险评估提供了简便、经济且高度敏感的手段。将这两项指标与临床评分系统或机器学习模型结合,有望进一步提升 ARDS 的预测准确性,优化个体化治疗策略。

综合所述,基于 ICU 的常见评分与指标构建预测模型能有效评估 AP 患者早期发生 ARDS 的风险,但本研究仍存诸多局限性:第一,本研究受限于单中心、回顾性设计,且未纳入炎症因子和影像学特征等潜在危险因素,模型的泛化能力仍有待进一步验证;第二,AP 的病因多样,其异质性可能会影响模型指标的贡献度,且本研究中仅通过 Logistic 回归进行变量筛选,可能在一定程度上限制了模型对复杂非线性关系的捕捉能力,未来值得针对其亚型进一步验证模型的预测有效性;第三,机器学习辅助诊断虽能提升效率,但模型可解释性仍待加强,模型在临床实践中的部署应用方案亦亟待解决。

4 结论

本研究基于临床常见指标及评分通过机器学习构建并验证了 AP 患者并发 ARDS 风险的预测模型,结果显示, XGBoost 模型和 AdaBoost 模型有良好的预测效能和稳定性,可助力于早期准确识别并发 ARDS 的高风险患者,辅助临床医师进行分层监测与个性化干预。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 中华医学会外科学分会胰腺外科学组. 中国急性胰腺炎诊治指南 (2021) [J]. 中华外科杂志, 2021, 59 (7): 578-587. DOI: 10.3760/cma.j.cn112139-20210416-00172.
- [2] Tenner S, Vege SS, Sheth SG, et al. American college of gastroenterology guidelines: management of acute pancreatitis [J]. Am J Gastroenterol, 2024, 119 (3): 419-437. DOI: 10.14309/ajg.0000000000002645.

- [3] Iyer H, Elhence A, Mittal S, et al. Pulmonary complications of acute pancreatitis [J]. *Expert Rev Respir Med*, 2020, 14 (2): 209–217. DOI: 10.1080/17476348.2020.1698951.
- [4] Schepers NJ, Bakker OJ, Besselink MG, et al. Impact of characteristics of organ failure and infected necrosis on mortality in necrotising pancreatitis [J]. *Gut*, 2019, 68 (6): 1044–1051. DOI: 10.1136/gutjnl-2017-314657.
- [5] ARDS Definition Task Force. Acute respiratory distress syndrome: the Berlin definition [J]. *JAMA*, 2012, 307 (23): 2526–2533. DOI: 10.1001/jama.2012.5669.
- [6] Matthay MA, Arabi Y, Arroliga AC, et al. A new global definition of acute respiratory distress syndrome [J]. *Am J Respir Crit Care Med*, 2024, 209 (1): 37–47. DOI: 10.1164/rccm.202303-0558WS.
- [7] 李文哲, 杜欣欣, 于湘友. 急性呼吸窘迫综合征药物治疗的窘境 [J]. *中国中西医结合急救杂志*, 2025, 32 (1): 122–126. DOI: 10.3969/j.issn.1008-9691.2025.01.024.
- [8] Gorman EA, O’Kane CM, McAuley DF. Acute respiratory distress syndrome in adults: diagnosis, outcomes, long-term sequelae, and management [J]. *Lancet*, 2022, 400 (10358): 1157–1170. DOI: 10.1016/S0140-6736(22)01439-8.
- [9] Lankisch PG, Rahlf G, Koop H. Pulmonary complications in fatal acute hemorrhagic pancreatitis [J]. *Dig Dis Sci*, 1983, 28 (2): 110–116. DOI: 10.1007/BF01315139.
- [10] Yang P, Sjöding MW. Acute respiratory distress syndrome: definition, diagnosis, and routine management [J]. *Crit Care Clin*, 2024, 40 (2): 309–327. DOI: 10.1016/j.ccc.2023.12.003.
- [11] Shah J, Rana SS. Acute respiratory distress syndrome in acute pancreatitis [J]. *Indian J Gastroenterol*, 2020, 39 (2): 123–132. DOI: 10.1007/s12664-020-01016-z.
- [12] Zhou Y, Mei S, Wang J, et al. Development and validation of a deep learning-based framework for automated lung CT segmentation and acute respiratory distress syndrome prediction: a multicenter cohort study [J]. *EclinicalMedicine*, 2024, 75: 102772. DOI: 10.1016/j.eclinm.2024.102772.
- [13] Johnson A, Bulgarelli L, Shen L, et al. MIMIC-IV, a freely accessible electronic health record dataset [J]. *Sci Data*, 2023, 10 (1): 1. DOI: 10.1038/s41597-022-01899-x.
- [14] Leppäniemi A, Tolonen M, Tarasconi A, et al. 2019 WSES guidelines for the management of severe acute pancreatitis [J]. *World J Emerg Surg*, 2019, 14: 27. DOI: 10.1186/s13017-019-0247-0.
- [15] 韩宝华, 杨文, 王慧, 等. 基于 CT 评分和炎症因子的重症急性胰腺炎预后模型构建及评价 [J]. *中华危重病急救医学*, 2023, 35 (1): 82–87. DOI: 10.3760/cma.j.cn121430-20220411-00351.
- [16] Hu JX, Zhao CF, Wang SL, et al. Acute pancreatitis: a review of diagnosis, severity prediction and prognosis assessment from imaging technology, scoring system and artificial intelligence [J]. *World J Gastroenterol*, 2023, 29 (37): 5268–5291. DOI: 10.3748/wjg.v29.i37.5268.
- [17] Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery [J]. *Lancet Oncol*, 2019, 20 (5): e262–e273. DOI: 10.1016/S1470-2045(19)30149-4.
- [18] 高欣, 林嘉希, 吴爱荣, 等. 基于 XGBoost 算法的机器学习模型在早期预测重症急性胰腺炎中的应用 [J]. *中华危重病急救医学*, 2023, 35 (4): 421–426. DOI: 10.3760/cma.j.cn121430-20221019-00930.
- [19] Sanchez-Pinto LN, Luo Y, Churpek MM. Big data and data science in critical care [J]. *Chest*, 2018, 154 (5): 1239–1248. DOI: 10.1016/j.chest.2018.04.037.
- [20] 周文君, 张频捷, 余维丽, 等. 重症急性胰腺炎治愈出院患者长期生存质量现状调查及影响因素分析 [J]. *中华危重病急救医学*, 2025, 37 (2): 146–152. DOI: 10.3760/cma.j.cn121430-20240626-00540.
- [21] Chauhan R, Saxena N, Kapur N, et al. Comparison of modified Glasgow-Imrie, Ranson, and APACHE II scoring systems in predicting the severity of acute pancreatitis [J]. *Pol Przegl Chir*, 2022, 95 (1): 6–12. DOI: 10.5604/01.3001.0015.8384.
- [22] Cazacu SM, Parscoveanu M, Cartu D, et al. NLR48 is better than CRP, and mCTSI, and similar to BISAP and SOFA scores for mortality prediction in acute pancreatitis: a comparison of 6 scores [J]. *J Inflamm Res*, 2023, 16: 4793–4804. DOI: 10.2147/JIR.S432408.
- [23] Metri A, Bush N, Singh VK. Predicting the severity of acute pancreatitis: current approaches and future directions [J]. *Surg Open Sci*, 2024, 19: 109–117. DOI: 10.1016/j.sopen.2024.03.012.
- [24] Mayow AH, Ahmad F, Afzal MS, et al. A systematic review and Meta-analysis of independent predictors for acute respiratory distress syndrome in patients presenting with sepsis [J]. *Cureus*, 2023, 15 (4): e37055. DOI: 10.7759/cureus.37055.
- [25] Li F, Wang Z, Bian R, et al. Predicting the risk of acute kidney injury in patients with acute pancreatitis complicated by sepsis using a stacked ensemble machine learning model: a retrospective study based on the MIMIC database [J]. *BMJ Open*, 2025, 15 (2): e087427. DOI: 10.1136/bmjopen-2024-087427.
- [26] Antonucci E, Garcia B, Chen D, et al. Incidence of acute kidney injury and attributive mortality in acute respiratory distress syndrome randomized trials [J]. *Intensive Care Med*, 2024, 50 (8): 1240–1250. DOI: 10.1007/s00134-024-07485-6.
- [27] Lin H, Ren Y, Cui J, et al. Nomogram risk prediction model for acute respiratory distress syndrome following acute kidney injury [J]. *Front Med (Lausanne)*, 2025, 12: 1563425. DOI: 10.3389/fmed.2025.1563425.
- [28] Spadaro S, Park M, Turrini C, et al. Biomarkers for acute respiratory distress syndrome and prospects for personalised medicine [J]. *J Inflamm (Lond)*, 2019, 16: 1. DOI: 10.1186/s12950-018-0202-y.
- [29] Mendes RS, Pelosi P, Schultz MJ, et al. Fluids in ARDS: more pro than cons [J]. *Intensive Care Med Exp*, 2020, 8 (Suppl 1): 32. DOI: 10.1186/s40635-020-00319-x.
- [30] Ocskay K, Vinkó Z, Németh D, et al. Hypoalbuminemia affects one third of acute pancreatitis patients and is independently associated with severity and mortality [J]. *Sci Rep*, 2021, 11 (1): 24158. DOI: 10.1038/s41598-021-03449-8.
- [31] Amri F, Rahaoui M, Aissaoui H, et al. Is serum albumin a pivotal biomarker in anticipating acute pancreatitis outcomes? [J]. *BMC Gastroenterol*, 2024, 24 (1): 234. DOI: 10.1186/s12876-024-03314-8.
- [32] Lv J, Zhou Y, Tao C, et al. Association between the triglyceride glucose index and the risk of acute respiratory failure in patients with acute pancreatitis [J]. *BMC Gastroenterol*, 2025, 25 (1): 182. DOI: 10.1186/s12876-025-03771-9.
- [33] Brochard L, Slutsky A, Pesenti A. Mechanical ventilation to minimize progression of lung injury in acute respiratory failure [J]. *Am J Respir Crit Care Med*, 2017, 195 (4): 438–442. DOI: 10.1164/rccm.201605-1081CP.
- [34] Li M, Liu F, Yang Y, et al. Identifying vital sign trajectories to predict 28-day mortality of critically ill elderly patients with acute respiratory distress syndrome [J]. *Respir Res*, 2024, 25 (1): 8. DOI: 10.1186/s12931-023-02643-8.
- [35] National Heart, Lung, and Blood Institute PETAL Clinical Trials Network. Early neuromuscular blockade in the acute respiratory distress syndrome [J]. *N Engl J Med*, 2019, 380 (21): 1997–2008. DOI: 10.1056/NEJMoa1901686.
- [36] 袁雪燕, 刘玲, 邱海波. 2023 急性呼吸窘迫综合征全球新标准: 进步与局限 [J]. *中华医学杂志*, 2024, 104 (15): 1216–1220. DOI: 10.3760/cma.j.cn112137-20231016-00770.

(收稿日期: 2025-07-30)

(责任编辑: 邸美仙)