

基于白蛋白衍生指标建立急性高脂血症型胰腺炎重症化的机器学习预测模型

李明睿¹, 王晓龙^{2*}

¹重庆医科大学附属第二医院重庆市急救医疗中心, 重庆

²重庆医科大学附属第二医院急诊部, 重庆

收稿日期: 2025年2月13日; 录用日期: 2025年3月7日; 发布日期: 2025年3月13日

摘要

背景: 建立基于白蛋白衍生指标的急性高脂血症型胰腺炎(HLAP)重症化的机器学习预测模型。方法: 回顾性收集HLAP患者临床数据, 按7:3分为训练集和验证集。训练集中通过Lasso回归筛选出与入住ICU显著相关的特征变量, 多种机器学习算法构建预测模型, 并在验证集中通过ROC曲线评估模型性能。结果: 最终筛选的特征变量包括APACHE II评分、是否发热、直接胆红素(DB)、是否胰周液体积聚及C反应蛋白/白蛋白(CAR)。逻辑回归模型表现最佳, 验证集中ROC曲线下面积(AUC)为0.757。结论: 基于CAR的预测模型能较准确地评估HLAP患者入住ICU的风险, 为早期识别HLAP患者重症化提供了一种简便实用的工具。

关键词

急性高脂血症型胰腺炎, 机器学习, 白蛋白, 重症化

Establishment of a Machine Learning Prediction Model for the Severity of Acute Hyperlipidemic Pancreatitis Based on Albumin-Derived Indicators

Mingrui Li¹, Xiaolong Wang^{2*}

¹Chongqing Emergency Medical Center, The Second Affiliated Hospital of Chongqing Medical University, Chongqing

²Emergency Department, The Second Affiliated Hospital of Chongqing Medical University, Chongqing

Received: Feb. 13th, 2025; accepted: Mar. 7th, 2025; published: Mar. 13th, 2025

*通讯作者。

文章引用: 李明睿, 王晓龙. 基于白蛋白衍生指标建立急性高脂血症型胰腺炎重症化的机器学习预测模型[J]. 临床医学进展, 2025, 15(3): 1098-1111. DOI: 10.12677/acm.2025.153717

Abstract

Background: Establish a machine learning prediction model for the severity of acute hyperlipidemic pancreatitis (HLAP) based on albumin-derived indicators. **Method:** Retrospective collection of clinical data from HLAP patients, divided into a training set and a validation set in a 7:3 ratio. In the training set, Lasso regression was used to screen out feature variables significantly correlated with ICU admission. Multiple machine learning algorithms were used to construct predictive models, and the model performance was evaluated in the validation set using ROC curves. **Result:** The final selected feature variables include APACHE II score, presence of fever, direct bilirubin (DB), presence of peripancreatic fluid accumulation, and C-reactive protein/albumin (CAR). The logistic regression model performed the best, with an area under the ROC curve (AUC) of 0.757 in the validation set. **Conclusion:** The CAR-based prediction model can accurately assess the risk of HLAP patients being admitted to the ICU, providing a simple and practical tool for the early identification of severe HLAP patients.

Keywords

Acute Hyperlipidemic Pancreatitis, Machine Learning, Albumin, Severe Illness

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 背景

急性胰腺炎(Acute Pancreatitis, AP)是一种胰腺自身消化引起的炎症疾病,起病急,进展快,已成为临床上最常见的急腹症之一[1]。随着社会经济发展及人类饮食结构改变,高脂血症是引起急性胰腺炎常见因素之一,有研究显示,高脂血症型胰腺炎(Hyperlipidemic Acute Pancreatitis, HLAP)在 AP 占比高达 22% [2]。研究显示,与其他类型急性胰腺炎相比,HLAP 具有更严重的病情严重度和更差的预后[2][3]。HLAP 中有高达 50% 的患者会快速进展为重症胰腺炎(Severe Acute Pancreatitis, SAP) [4],其病死率较 AP 增加十倍[5],给患者带来极大的健康损害和经济负担。因此,早期识别急性高脂血症型胰腺炎患者的病情进展具有重要的临床意义。

AP 患者中超过 10% 的患者需要入住 ICU [6], HLAP 入住 ICU 的比例更高[4]。指南推荐有器官功能衰竭和/或全身炎症反应综合征(SIRS)的患者应优先入住监护病床。遗憾的是,目前单纯依靠评分系统(如 APACHE-II、Ranson、BISAP 评分)和影像学均无法准确预测中重度或重症急性胰腺炎的发生和胰腺炎的进展,无法较好地识别出需要入住 ICU 的 HLAP 患者[7]。亟需寻找新的能够反映 HLAP 发病机制并影响其临床结局的新型生物标志物和/或采用先进的机器学习算法充分利用现有数据,以提高 HLAP 入住 ICU 模型的预测准确性。

机器学习是计算机模拟人类大脑进行数据学习的一门科学技术,它能够从大量已知、复杂的数据中学习并总结规律,发现模式及趋势,从而能够对新的数据进行预测或做出决策。在构建医学模型、临床预后、疾病诊断等已取得显著成就[8]-[10]。近年来,基于大数据的 ML 算法在急性胰腺炎领域的应用越来越广泛。

白蛋白(Albumin, ALB)是健康人血浆中最丰富的蛋白质,主要由肝细胞产生,其生理功能包括营养、

抗氧化、抗炎等。目前研究发现, 白蛋白可以作为预测急性胰腺炎持续性器官衰竭的独立危险因素[11]。白蛋白与其他生化指标结合(如白蛋白-球蛋白比值、白蛋白-胆红素比值等)可能提供更复杂的病理生理信息, 有助于精细评估病情严重度和预后。目前已有研究发现白蛋白衍生指标如血清肌酐与白蛋白比值、乳酸与白蛋白比值(LAR)与急性胰腺炎严重程度及预后的相关性研究[12][13]。但尚无在 HLAP 患者中基于白蛋白衍生指标采用机器学习方法建立入住 ICU 模型的相关研究。基于此, 本研究运用机器学习方法, 建立基于白蛋白衍生指标对急性高脂血症型胰腺炎患者入住 ICU 的预测模型。

2. 资料与方法

2.1. 研究对象

选取重庆医科大学附属第二医院 2017 年 1 月至 2022 年 12 月收治的 2311 例 AP 患者, 筛选出 350 例确诊为 HLAP 的患者为研究对象。HLAP 的诊断标准参照亚特兰大标准。本研究方案获得重庆医科大学附属第二医院伦理委员会审批[2024 年科伦审第(28)号]。

2.2. 纳入与排除标准

纳入标准: ① 符合 HLAP 诊断标准; ② 年龄 ≥ 18 岁。排除标准: ① 合并慢性肝/肾衰竭及妊娠状态; ② 恶性肿瘤; ③ 发病时间超过 24 小时。

2.3. 研究目标

本研究以是否入住 ICU 为主要结局终点。

2.4. 数据收集

基于本研究目的, 收集以下数据: 1) 一般资料, 包括性别、年龄、基础疾病(高血压、糖尿病、脂肪肝); 2) 结局指标(住院天数、ICU 入住率); 3) 入院 24 h 内的实验室指标, 包括空腹血糖、胰腺炎酶学(血清淀粉酶(AMY)、血清脂肪酶(LPS))、生化指标(血脂、肝功能、电解质、肾功能等); 4) 并发症指标(急性胰周液体积聚(Acute Peripancreatic Fluid Collection, APFC)、急性坏死物积聚(Acute Necrotic Collection, ANC)、胰腺假性囊肿、感染性胰腺坏死胸腔积液、腹腔积液、发热(体温 $\geq 37.3^{\circ}\text{C}$)); 5) 白蛋白衍生指标及其计算: PAR: 血小板/白蛋白; CAR: C 反应蛋白/白蛋白; AAR: 碱性磷酸酶/白蛋白; CAR_crea: 肌酐/白蛋白; BAR: 尿素氮/白蛋白。本研究排除了缺失值 $> 20\%$ 的实验室检查指标, 缺失数据使用 mice 包进行插补。

2.5. 模型的开发与验证

基于 6 种常见的机器学习算法构建模型, 分别是支持向量(Support Vector Machines, SVM)、决策树(Decision Tree, DT)、极端梯度提升(eXtreme Gradient Boosting, XGboost)、随机森林(Random Forest, RF)、K 近邻算法(K-Nearest Neighbors)和逻辑回归(Logistic Regression, LR)。首先将数据集按照 7:3 的比例随机划分为训练集和测试集, 训练集用于模型的构建, 测试集用于模型的评估。本研究采用 Lasso 回归进行变量筛选, 利用 10 折交叉验证确定最优正则化参数 λ , 并通过 Glmnet 包进行模型拟合, 将筛选出的变量纳入 SVM、DT、XGboost、RF、KNN、LR 算法中。模型构建过程采用十折交叉验证(Cross-Validation, CV), 即将训练集数据分成 10 等份, 每份数据都在上述机器学习模型中进行训练和测试, 每次选择其中 2 份作为验证集, 剩余 8 份作为训练集, 最终模型评估结果取 10 次训练结果的平均值, 以更准确地评估模型的泛化性能。通过网格搜索(Grid Search)调整模型参数, 不断对模型进行优化。模型性能的评估选用

灵敏度、特异度、准确度、精确度。F1 分数和受试者工作特征(ROC)曲线下面积(AUC)为评价指标。最终绘制 6 个模型的 ROC 曲线, 更加直观地显示各模型的 AUC 值差异。

2.6. 统计学方法

采用 SPSS 27.0 和 R 语言软件进行数据分析。符合正态分布的计量资料以 $\bar{x} \pm s$ 表示, 组间比较采用独立样本 t 检验; 不符合正态分布的计量资料以 M (P25, P75)表示, 组间比较采用 Mann-Whitney U 检验。计数资料以百分比(%)表示, 组间比较采用 χ^2 检验。P < 0.05 为差异具有统计学意义。采用 R 4.3.1 软件进行模型构建, 并计算 ROC 曲线下面积(Area Under Curve, AUC)的灵敏度、特异度等模型评价指标, 绘制 DCA 曲线评估临床获益。

3. 结果

3.1. 基线资料比较

经过逐步筛选, 本研究最终共纳入 HLAP 患者 350 例, 具体流程见图 1。将满足纳入排除标准的患者根据是否入住 ICU 分为入住 ICU 组及未入住 ICU 组, 两组患者的特征情况见表 1。分析两组间基线资料发现, 入住 ICU 患者费用显著高于未入住 ICU 组, 淀粉酶学高于未入住 ICU 组。入住 ICU 患者的甘油三酯水平更高, 白蛋白水平更低, 中性粒细胞计数、C 反应蛋白、降钙素原更高, 更容易合并并发症, 且 APACHE II 评分更高。白蛋白衍生指标中, 入住 ICU 中 PAR、CAR 及 BAR 显著高于未入住 ICU 组, 差异具有统计学意义。

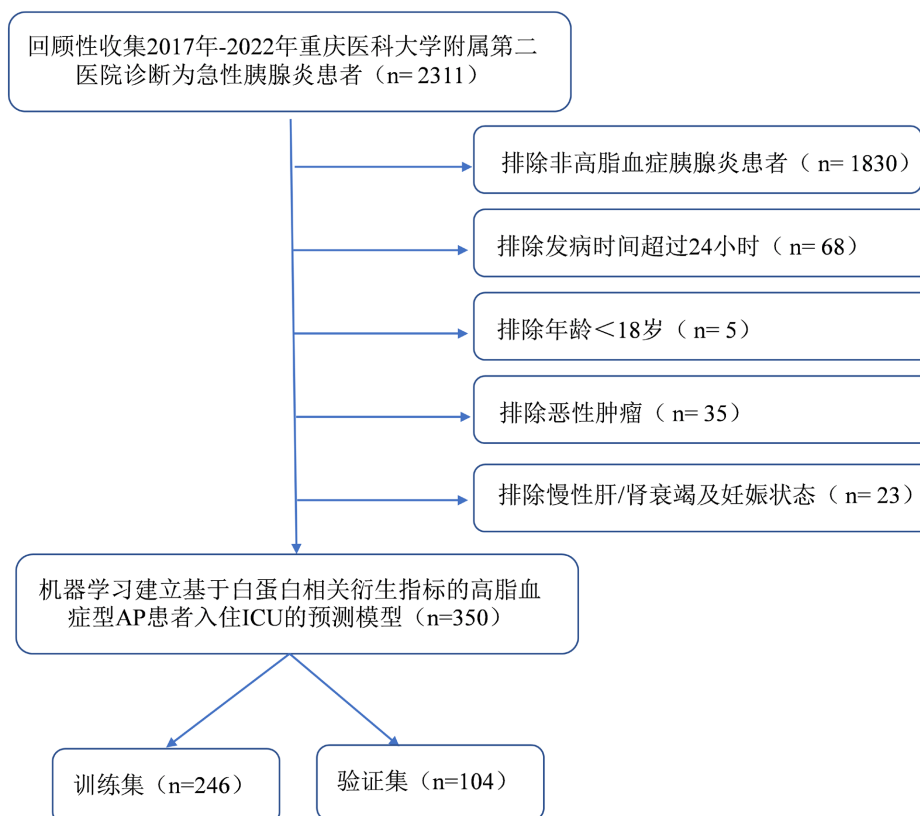


Figure 1. Research flowchart
图 1. 研究流程图

Table 1. Comparison of baseline data between H LAP patients admitted to ICU and those not admitted to ICU
表 1. H LAP 入住 ICU 组与未入住 ICU 组的基线资料比较

参数	未入住 ICU 组(N = 284)	入住 ICU 组(N = 66)	P 值
一般资料			
年龄(岁)	42.0 (34.0; 51.0)	40.0 (32.0; 46.8)	0.194
男性	206 (72.5%)	46 (69.7%)	0.756
治疗费用(元)	8204 (4767; 13756)	41623 (18837; 59647)	<0.001
既往病史			
脂肪肝	159 (56.0%)	35 (53.0%)	0.766
糖尿病	133 (46.8%)	35 (53.0%)	0.441
高血压	59 (20.8%)	10 (15.2%)	0.388
吸烟	162 (57.0%)	41 (62.1%)	0.539
饮酒	138 (48.6%)	39 (59.1%)	0.161
实验室指标			
胰腺酶学			
淀粉酶(U/L)	183 (93.5; 324)	302 (115; 731)	0.002
脂肪酶(U/L)	640 (174; 1700)	466 (160; 2166)	0.756
血脂			
甘油三酯(mmol/L)	15.1 (9.09; 25.7)	24.1 (13.4; 29.6)	<0.001
总胆固醇(mmol/L)	7.69 (5.88; 10.1)	8.95 (6.30; 11.5)	0.097
高密度脂蛋白(mmol/L)	1.25 (0.73; 2.48)	0.89 (0.55; 2.15)	0.029
低密度脂蛋白(mmol/L)	1.75 (1.26; 2.76)	1.78 (0.96; 3.32)	0.801
载脂蛋白 A (g/L)	1.19 (1.00; 1.37)	1.00 (0.56; 1.22)	<0.001
脂蛋白(mg/L)	35.4 (20.6; 71.1)	24.0 (15.2; 46.5)	0.001
肝功			
白蛋白(g/L)	41.6 (38.9; 44.4)	39.6 (33.8; 43.5)	0.006
ALT (U/L)	23.0 (16.0; 34.0)	24.5 (14.2; 38.8)	0.559
AST (U/L)	22.0 (17.0; 28.0)	27.5 (21.0; 44.0)	<0.001
ALP (U/L)	73.0 (60.0; 87.0)	65.0 (50.8; 91.8)	0.341
γ-GT (U/L)	46.0 (26.0; 82.0)	60.0 (24.2; 112)	0.118
总胆红素(μmol/L)	12.2 (8.87; 16.3)	13.5 (9.45; 21.1)	0.030
直接胆红素(μmol/L)	3.80 (2.80; 5.20)	6.00 (3.52; 9.20)	<0.001
间接胆红素(μmol/L)	8.20 (4.90; 12.2)	7.45 (4.05; 12.2)	0.747
总胆汁酸(mg/L)	1.20 (0.60; 2.60)	1.30 (0.63; 3.17)	0.344

续表

血常规			
红细胞计数($\times 10^{12}/L$)	4.92 (4.48; 5.37)	4.74 (4.30; 5.37)	0.234
血红蛋白(g/L)	155 (141; 168)	157 (138; 173)	0.799
白细胞计数($\times 10^9/L$)	12.6 (9.67; 15.3)	11.8 (8.75; 15.8)	0.594
中性粒细胞比例(%)	82.5 (76.5; 86.4)	85.7 (81.7; 89.1)	<0.001
血小板计数($\times 10^9/L$)	202 (157; 242)	181 (138; 220)	0.037
中性粒细胞计数($\times 10^9/L$)	10.0 (7.63; 13.0)	10.4 (7.37; 13.5)	0.653
单核细胞计数($\times 10^9/L$)	0.06 (0.03; 0.12)	0.05 (0.03; 0.12)	0.843
C 反应蛋白(mg/L)	14.1 (5.00; 52.6)	59.1 (22.1; 197)	<0.001
降钙素原(ng/ml)	0.07 (0.04; 0.15)	0.30 (0.11; 1.39)	<0.001
Na (mmol/L)	136 (132; 139)	135 (130; 138)	0.049
Ca (mmol/L)	2.25 (2.16; 2.35)	2.08 (1.85; 2.19)	<0.001
肾功			
肌酐($\mu\text{mol}/L$)	55.9 (44.6; 67.2)	54.6 (43.7; 82.7)	0.779
尿素氮(mg/dL)	4.51 (3.69; 5.46)	4.74 (3.40; 6.29)	0.273
GFR (mL/min)	116 (105; 129)	121 (98.9; 138)	0.931
血糖(mmol/L)	10.2 (7.20; 13.8)	12.0 (8.20; 15.3)	0.009
并发症			
APFC	48 (16.9%)	28 (42.4%)	<0.001
ANC	9 (3.17%)	17 (25.8%)	<0.001
PPC	16 (5.63%)	2 (3.03%)	0.543
IPN	6 (2.11%)	6 (9.09%)	0.013
SIRS	7 (2.46%)	13 (19.7%)	<0.001
AKI	11 (3.87%)	14 (21.2%)	<0.001
发热	53 (18.7%)	40 (60.6%)	<0.001
白蛋白相关衍生指标			
PAR	0.00 (0.00; 0.00)	0.01 (0.00; 0.04)	<0.001
CAR	0.36 (0.13; 1.43)	1.52 (0.53; 4.67)	<0.001
AAR	1.74 (1.45; 2.10)	1.65 (1.38; 2.38)	0.904
CAR_crea	1.35 (1.06; 1.67)	1.40 (0.97; 2.21)	0.334
BAR	0.11 (0.09; 0.13)	0.13 (0.08; 0.17)	0.042
APACHE II 评分	3.00 (2.00; 5.00)	9.00 (4.00; 10.8)	<0.001

注: APFC 为急性胰周液体积聚; ANC 为急性坏死物积聚; PPC 为胰腺假性囊肿; IPN 为感染性胰腺坏死; SIRS 为全身炎症反应综合征; AKI 为急性肾损伤。

3.2. 训练集与测试集的比较

本研究按 7:3 比例将 HLAP 患者随机划分为训练集和测试集, 比较发现训练集与测试集中除了 C 反应蛋白与 CAR 外, 其余参数均无统计学意义(见表 2)。

Table 2. Comparison of general data between the training set and the test set of HLAP patients
表 2. HLAP 患者训练集和测试集一般资料比较

参数	训练集(N = 246)	验证集(N = 104)	P 值
一般资料			
年龄(岁)	41.5 (33.0; 49.8)	42.0 (34.0; 48.2)	0.759
男性	179 (72.8%)	73 (70.2%)	0.719
治疗费用(元)	9780 (5455; 19492)	9925 (5604; 18585)	0.777
既往病史			
脂肪肝	133 (54.1%)	61 (58.7%)	0.502
糖尿病	127 (51.6%)	41 (39.4%)	0.049
高血压	50 (20.3%)	19 (18.3%)	0.768
吸烟	142 (57.7%)	61 (58.7%)	0.966
饮酒	126 (51.2%)	51 (49.0%)	0.798
实验室指标			
胰腺酶学			
淀粉酶(U/L)	186 (89.2; 353)	222 (110; 482)	0.191
脂肪酶(U/L)	633 (172; 1710)	685 (194; 1905)	0.613
血脂			
甘油三酯(mmol/L)	16.1 (9.39; 25.4)	19.2 (10.2; 28.6)	0.091
总胆固醇(mmol/L)	7.81 (5.87; 10.1)	8.43 (6.26; 11.1)	0.209
高密度脂蛋白(mmol/L)	1.11 (0.68; 2.33)	1.49 (0.66; 2.73)	0.547
低密度脂蛋白(mmol/L)	1.70 (1.19; 2.60)	1.84 (1.29; 3.20)	0.073
载脂蛋白 A (g/L)	1.17 (0.94; 1.37)	1.18 (0.90; 1.34)	0.607
肝功			
白蛋白(g/L)	41.5 (38.6; 44.6)	40.9 (37.8; 43.7)	0.324
ALT (U/L)	23.0 (16.0; 34.0)	23.6 (15.0; 37.0)	0.773
AST (U/L)	22.0 (18.0; 28.0)	23.5 (17.8; 32.0)	0.300
ALP (U/L)	75.0 (60.0; 87.0)	68.0 (59.0; 88.0)	0.644
γ-GT (U/L)	48.5 (25.2; 85.5)	44.0 (25.5; 89.7)	0.850
总胆红素(μmol/L)	12.4 (9.05; 17.2)	12.1 (9.17; 17.0)	0.588
直接胆红素(μmol/L)	4.00 (2.90; 5.88)	3.80 (2.80; 5.85)	0.737
间接胆红素(μmol/L)	8.15 (5.00; 12.2)	8.05 (4.20; 12.2)	0.648
总胆汁酸(mg/L)	1.30 (0.60; 2.60)	1.05 (0.60; 2.60)	0.411

续表

血常规			
红细胞计数($\times 10^{12}/L$)	4.87 (4.49; 5.37)	4.91 (4.35; 5.42)	0.613
血红蛋白(g/L)	155 (140; 168)	156 (138; 168)	0.849
白细胞计数($\times 10^9/L$)	12.7 (9.43; 15.6)	11.8 (10.2; 14.9)	0.612
中性粒细胞比例(%)	83.1 (77.9; 86.9)	83.2 (76.8; 87.3)	0.881
血小板计数($\times 10^9/L$)	198 (152; 238)	198 (148; 241)	0.855
中性粒细胞计数($\times 10^9/L$)	10.4 (7.48; 13.3)	9.59 (7.64; 13.0)	0.615
单核细胞计数($\times 10^9/L$)	0.06 (0.03; 0.13)	0.06 (0.03; 0.11)	0.385
C 反应蛋白(mg/L)	33.8 (5.00; 73.8)	8.11 (5.00; 65.6)	0.011
降钙素原(ng/ml)	0.10 (0.05; 0.28)	0.07 (0.04; 0.18)	0.115
Na (mmol/L)	136 (132; 139)	136 (132; 138)	0.431
Ca (mmol/L)	2.24 (2.13; 2.35)	2.22 (2.13; 2.31)	0.245
肾功			
肌酐($\mu\text{mol/L}$)	55.8 (46.0; 67.8)	54.3 (41.7; 68.2)	0.438
尿素氮(mg/dL)	4.62 (3.68; 5.56)	4.41 (3.53; 5.71)	0.430
GFR(mL/min)	117 (105; 130)	117 (100; 132)	0.959
血糖(mmol/L)	10.7 (7.50; 14.4)	10.4 (7.42; 13.1)	0.330
并发症			
APFC	51 (20.7%)	25 (24.0%)	0.587
ANC	16 (6.50%)	10 (9.62%)	0.429
PPC	11 (4.47%)	7 (6.73%)	0.542
IPN	9 (3.66%)	3 (2.88%)	1.000
SIRS	12 (4.88%)	8 (7.69%)	0.433
AKI	18 (7.32%)	7 (6.73%)	1.000
发热	63 (25.6%)	30 (28.8%)	0.621
白蛋白相关衍生指标			
PAR	0.00 (0.00; 0.01)	0.00 (0.00; 0.00)	0.138
CAR	0.84 (0.13; 1.86)	0.20 (0.13; 1.61)	0.034
AAR	1.74 (1.43; 2.15)	1.67 (1.39; 2.09)	0.818
CAR_crea	1.35 (1.07; 1.70)	1.34 (0.99; 1.74)	0.627
BAR	0.11 (0.09; 0.14)	0.11 (0.08; 0.14)	0.488
APACHE II 评分	4.00 (2.00; 7.00)	3.50 (2.00; 6.00)	0.158
入住 ICU	47 (19.1%)	19 (18.3%)	0.973

3.3. 特征变量筛选及模型构建

在训练集队列中, 为避免数据分析存在共线性, 使用 Lasso 回归分析对自变量进行筛选(见图 2(a)、图 2(b)), 共筛选出 5 个自变量, 其系数非零。这些变量包括 APACH II 评分、直接胆红素、ANC、发热以及 CAR。

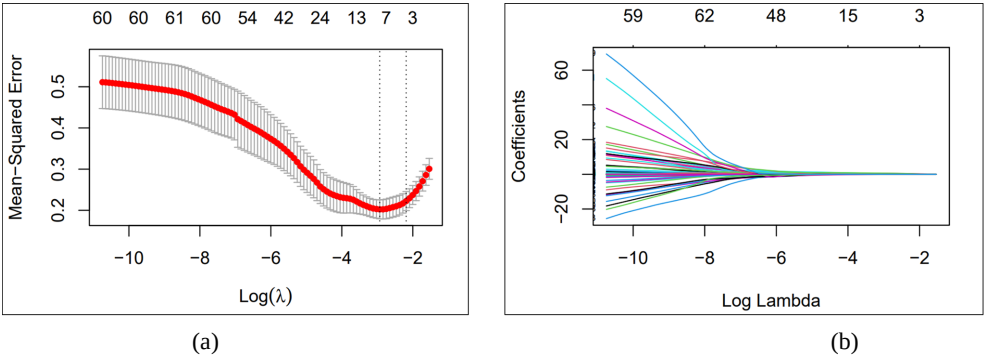
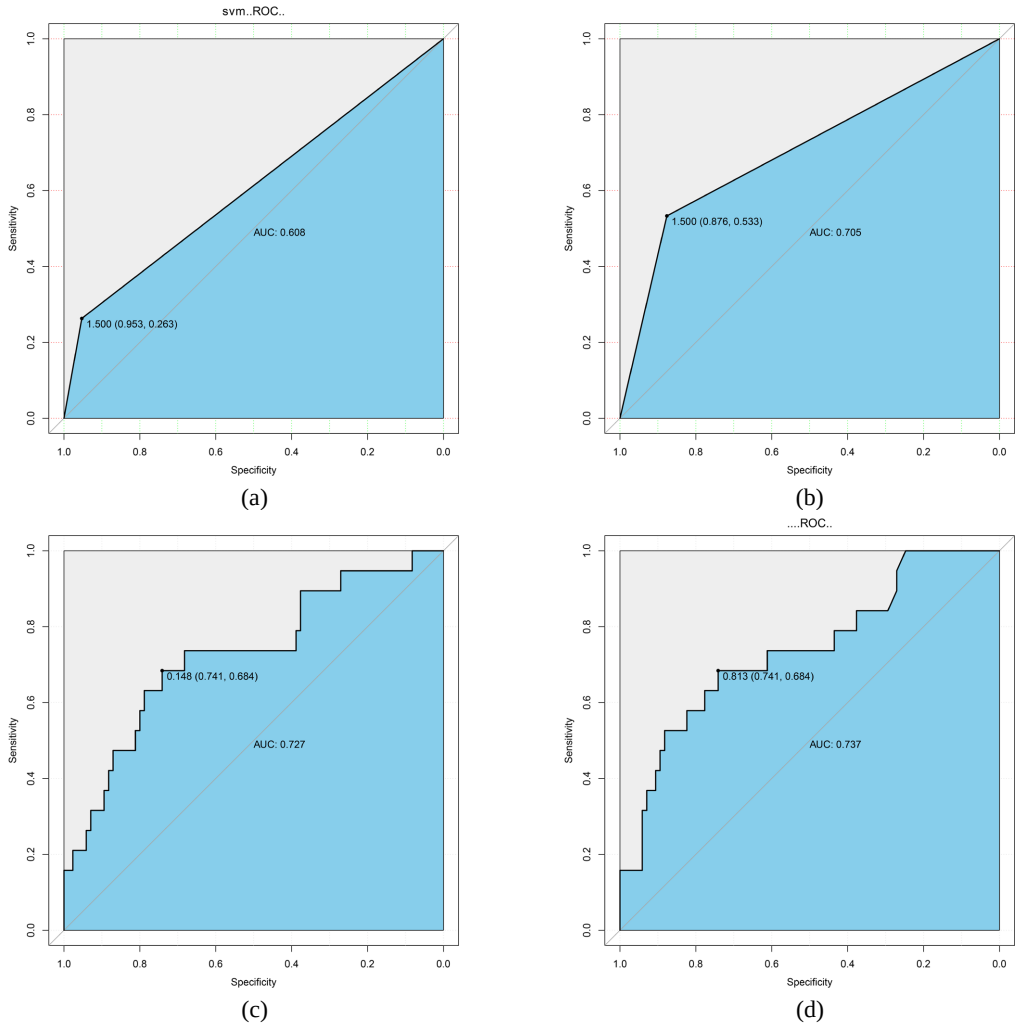


Figure 2. Variable screening through Lasso regression analysis
图 2. Lasso 回归分析筛选变量

3.4. 最优模型解释及可视化

本研究通过多种机器学习建立了急性胰腺炎患者入住 ICU 的预测模型，并评估了它们在验证集中预测急性胰腺炎患者入住 ICU 的准确性与性能。结果显示采用逻辑回归模型 ROC 曲线下面积最大(AUC = 0.7579)，见图 3。不同机器学习模型预测性能指标见表 3。



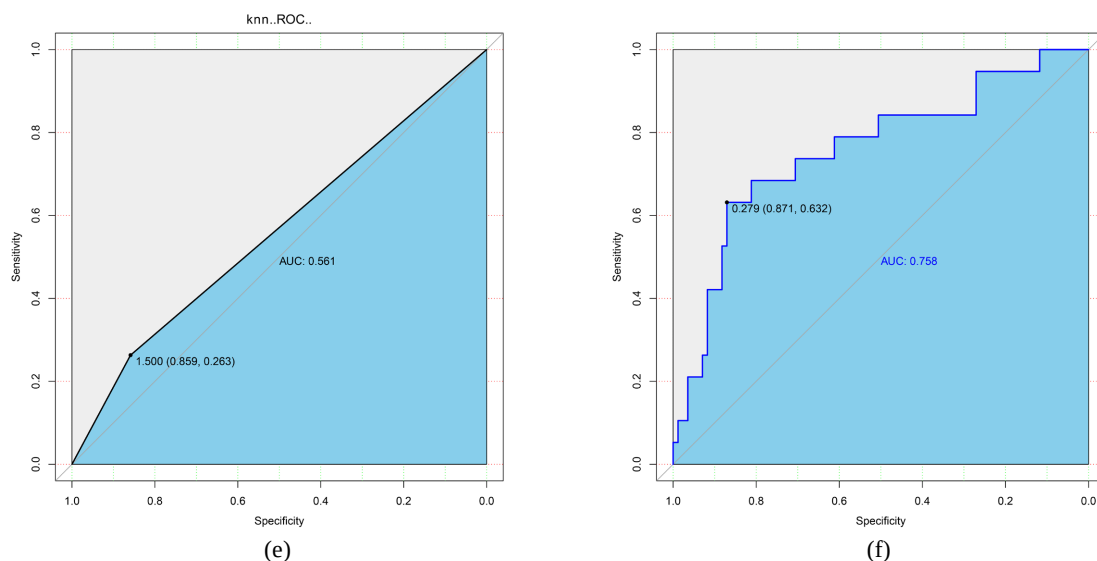


Figure 3. ROC curve of each prediction mode

图 3. 各预测模型的 ROC 曲线

Table 3. Comparison of prediction performance metrics of different machine learning models

表 3. 不同机器学习模型预测性能指标对比

Models		Testing Dataset						
	Predict survived	Predict dead	Accuracy [95% CI]	Positive predictive value [95% CI]	Negative predictive value [95% CI]	Specificity [95% CI]	Sensitivity [95% CI]	AUC [95% CI]
SVM								
survived	81	14	0.8269	0.8526	0.5556	0.2632	0.9529	0.608
dead	4	5						
决策树								
survived	78	11	0.8269	0.8764	0.5333	0.4211	0.9176	0.7049
dead	7	8						
XGboost								
survived	79	13	0.8173	0.8587	0.5	0.3158	0.9294	0.7269
dead	6	6						
随机森林								
survived	78	12	0.8173	0.8667	0.5	0.3684	0.9176	0.7375
dead	7	7						
KNN								
survived	70	13	0.7308	0.8434	0.2857	0.3158	0.8235	0.561
dead	15	6						
逻辑回归								
survived	78	7	0.8077	0.9176	0.3158	0.4615	0.8571	0.7579
dead	13	6						

3.5. 列线图

本研究发现逻辑回归曲线下面积最大, 以此结果建立预测急性高脂血症型胰腺炎入住 ICU 的列线图 (图 4)。构建列线图的变量分别为 APACH II 评分、直接胆红素、ANC、发热以及 CAR, 各变量在列线图 中的权重由其回归系数决定, 其中 CAR 对总评分的贡献较高, 反映其在预测 ICU 风险中的重要性。通过简单的变量测量与分值累加, 可快速评估患者的 ICU 入住概率, 为临床干预提供支持。

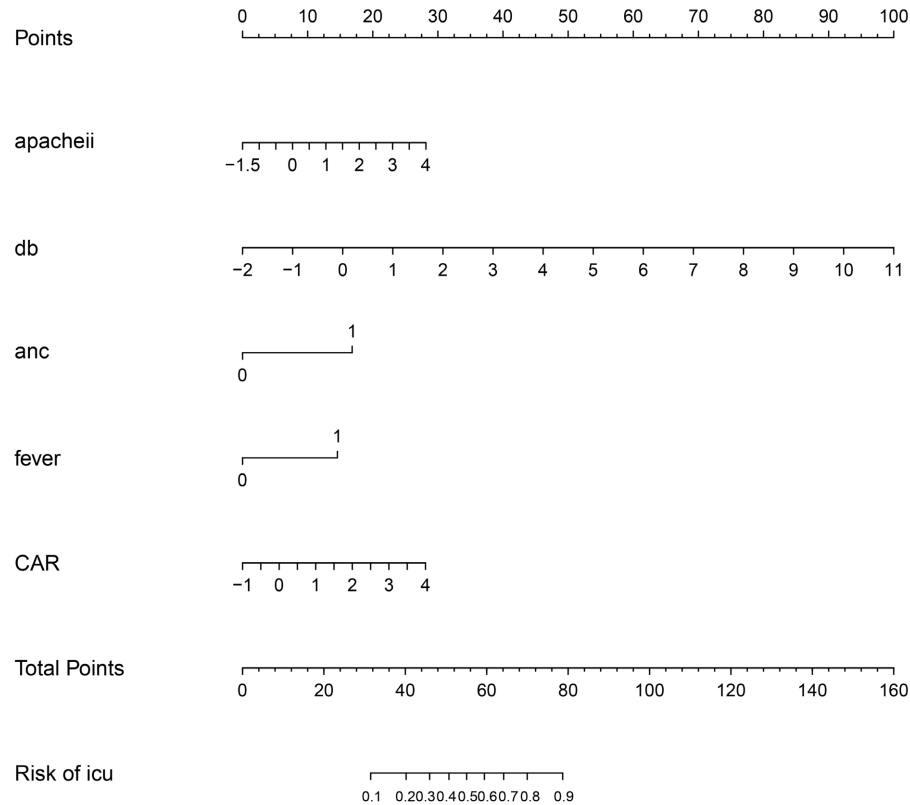


Figure 4. Nomogram
图 4. 列线图

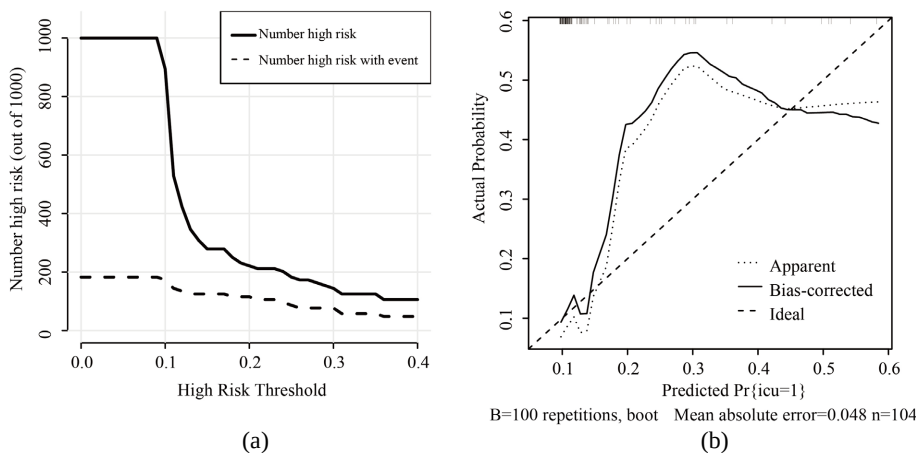


Figure 5. Decision curve analysis
图 5. 决策曲线分析

3.6. DCA 曲线和校准曲线

决策曲线分析(图 5(a))显示, 该预测模型在阈值概率范围 0.1 至 0.3 之间具有较高的净获益。表明本模型在平衡灵敏度与特异度方面具有良好的表现, 适用于 ICU 资源有限的临床场景。校准曲线(图 5(b))显示, 模型在低风险范围(<0.2)的预测概率与实际 ICU 入住概率接近理想状态。然而, 在中高风险范围(>0.2), 预测值略高于实际发生率, 提示模型在高风险患者中的预测可能存在一定程度的高估。通过偏差校正后的模型校准表现显著改善, 与理想曲线更加接近, 验证了模型预测结果的可靠性。

4. 讨论

急性高脂性胰腺炎病情进展迅速, 部分患者需入住重症监护室(ICU)进行强化治疗。因此, 准确预测患者是否需要 ICU 管理, 对于优化医疗资源分配、指导临床决策具有重要意义。本研究基于白蛋白衍生指标构建预测模型, 发现基于 CAR 的临床预测模型能较好地预测 HLAP 患者是否需要入住 ICU。本研究结果提供一种简便、实用的工具, 帮助临床医师快速识别 HLAP 的高危患者。

白蛋白是反映机体营养和炎症状态的重要指标, 其衍生参数, 如 CAR、PAR、AAR、肌酐/白蛋白, 结合了白蛋白与炎症、凝血及代谢相关因子的比值, 更全面地反映了患者的全身炎症反应和器官功能状态, 能有效预测患者预后[14][15]。既往研究发现肌酐/白蛋白是 AP 患者短期和长期全因死亡率的独立预测因子[12][16]。在本研究中, 通过 Lasso 回归筛选发现, CAR 与 AHLF 患者入住 ICU 的风险显著相关, 这一结果提示 CAR 能够作为一种简便且具有临床应用价值的指标, 为预测病情严重程度提供了重要参考。

CAR 结合了急性炎症反应的敏感指标 C 反应蛋白(CRP)与营养和炎症状态的标志白蛋白, 是评估全身炎症反应综合指数的有效工具。CRP 水平的升高提示急性炎症的存在, 而白蛋白的降低则与疾病严重程度及预后密切相关。CAR 在整合这两个参数后, 能够更精准地反映机体炎症负荷和营养耗竭状态, 从而具有较高的预测效能。目前已有研究证实 CAR 在脓毒症、急性呼吸窘迫综合征(ARDS)、急性心肌梗死等危重症中具有良好的预测效能[17][18]。同时, 在急性胰腺炎中能有效预测急性肾损伤的发生[19], 本研究发现 CAR 在预测 HLAP 患者入住 ICU 需求中的有效性进一步验证了其在重症医学领域的广泛适用性。

相比传统统计方法, 机器学习能够在复杂数据中识别非线性关系并有效处理多维特征。本研究应用了多种机器学习方法, 包括支持向量机(SVM)、决策树(DT)、极端梯度提升(XGboost)、随机森林(RF)、K 近邻算法(KNN)和逻辑回归(LR)。最终发现, 基于 CAR 逻辑回归模型的 ROC 面积最大, 提示其在预测 AHLF 患者入住 ICU 方面具有良好的性能。逻辑回归的优势在于模型简单、可解释性强, 更适合实际临床应用。

机器学习在急性胰腺炎中已有相关研究, Balázs Kui 等运用机器学习方法构建早期预测 SAP 的模型, 结果发现 XGBoost 机器学习算法最佳, 六个最具影响的特征分别是呼吸频率、体温、腹肌紧张、性别、年龄和血糖水平[20]。Zhang 等发现利用机器学习构建急性胰腺炎发生急性肾损伤的预测模型, 结果发现 XGBoost 方法构建的预测模型 ROC 最大, 贡献度最大的前五位参数依次为 SOFA 评分、APACHE II 评分、血小板计数、尿素氮和尿酸[21]。本研究中根据 Lasso 回归筛选的变量为 APACHE II 评分、发热、直接胆红素、ANC 和 C 反应蛋白/白蛋白。采用逻辑回归构建的预测模型在验证集中的 AUC 为 0.757, 能有效预测 HLAP 患者是否需要入住 ICU, 具有一定的临床应用价值。

本研究存在以下局限性: 1) 本研究为单中心、小样本量研究, 可能限制模型的外部推广性和稳定性; 2) 本研究纳入的变量可能存在未全面涵盖 HLAP 病情进展的潜在重要因素, 从而影响研究结果; 3) 本研究仅进行训练集建模, 未行外部验证, 同时基于静态指标, 未考虑患者病情随时间变化的数据。因此,

未来需通过多中心大样本的研究来验证模型的普适性; 引入更多变量, 特别是动态变化数据且尝试更复杂的机器学习方法优化模型性能, 以提升对 HLAP 患者的精准预测和治疗价值。

5. 结论

本研究采用逻辑回归构建的基于 C 反应蛋白/白蛋白的机器学习模型能有效预测 HLAP 患者是否需要入住 ICU, 其预测性能与临床应用的解释性较好。该模型可帮助临床医师快速识别病情进展迅速的 HLAP 高危患者, 优化 ICU 资源分配, 提高临床决策的科学性, 有助于改善患者预后。

参考文献

- [1] Gliem, N., Ammer-Herrmann, C., Ellenrieder, V. and Neesse, A. (2020) Management of Severe Acute Pancreatitis: An Update. *Digestion*, **102**, 503-507. <https://doi.org/10.1159/000506830>
- [2] Jin, M., Bai, X., Chen, X., Zhang, H., Lu, B., Li, Y., et al. (2019) A 16-Year Trend of Etiology in Acute Pancreatitis: The Increasing Proportion of Hypertriglyceridemia-Associated Acute Pancreatitis and Its Adverse Effect on Prognosis. *Journal of Clinical Lipidology*, **13**, 947-953.e1. <https://doi.org/10.1016/j.jacl.2019.09.005>
- [3] Vipperla, K., Somerville, C., Furlan, A., Koutroumpakis, E., Saul, M., Chennat, J., et al. (2017) Clinical Profile and Natural Course in a Large Cohort of Patients with Hypertriglyceridemia and Pancreatitis. *Journal of Clinical Gastroenterology*, **51**, 77-85. <https://doi.org/10.1097/mcg.0000000000000579>
- [4] Zhou, W., Liu, Q., Wang, Z., Yao, L., Chen, J. and Yang, X. (2024) Analysis of the Clinical Profile and Treatment Efficiency of Hyperlipidemic Acute Pancreatitis. *Lipids in Health and Disease*, **23**, Article No. 70. <https://doi.org/10.1186/s12944-024-02057-5>
- [5] Rashid, N., Sharma, P.P., Scott, R.D., Lin, K.J. and Toth, P.P. (2017) All-Cause and Acute Pancreatitis Health Care Costs in Patients with Severe Hypertriglyceridemia. *Pancreas*, **46**, 57-63. <https://doi.org/10.1097/mpa.0000000000000704>
- [6] Finkenstedt, A., Jaber, S. and Joannidis, M. (2023) Ten Tips to Manage Severe Acute Pancreatitis in an Intensive Care Unit. *Intensive Care Medicine*, **49**, 1127-1130. <https://doi.org/10.1007/s00134-023-07121-9>
- [7] Tenner, S., Vege, S.S., Sheth, S.G., Sauer, B., Yang, A., Conwell, D.L., et al. (2023) American College of Gastroenterology Guidelines: Management of Acute Pancreatitis. *American Journal of Gastroenterology*, **119**, 419-437. <https://doi.org/10.14309/ajg.0000000000002645>
- [8] Han, T., Cheng, T., Liao, Y., He, Y., Liu, B., Lai, Q., et al. (2022) Development and Validation of a Novel Prognostic Score Based on Thrombotic and Inflammatory Biomarkers for Predicting 28-Day Adverse Outcomes in Patients with Acute Pancreatitis. *Journal of Inflammation Research*, **15**, 395-408. <https://doi.org/10.2147/jir.s344446>
- [9] Xu, F., Chen, X., Li, C., Liu, J., Qiu, Q., He, M., et al. (2021) Prediction of Multiple Organ Failure Complicated by Moderately Severe or Severe Acute Pancreatitis Based on Machine Learning: A Multicenter Cohort Study. *Mediators of Inflammation*, **2021**, Article ID: 5525118. <https://doi.org/10.1155/2021/5525118>
- [10] Ye, J., Zhao, Y., Ju, J. and Wang, W. (2017) Building and Verifying a Severity Prediction Model of Acute Pancreatitis (AP) Based on BISAP, MEWS and Routine Test Indexes. *Clinics and Research in Hepatology and Gastroenterology*, **41**, 585-591. <https://doi.org/10.1016/j.clinre.2016.11.013>
- [11] Li, S., Zhang, Y., Li, M., Xie, C. and Wu, H. (2017) Serum Albumin, a Good Indicator of Persistent Organ Failure in Acute Pancreatitis. *BMC Gastroenterology*, **17**, Article No. 59. <https://doi.org/10.1186/s12876-017-0615-8>
- [12] Wang, J., Li, H., Luo, H., Shi, R., Chen, S., Hu, J., et al. (2024) Association between Serum Creatinine to Albumin Ratio and Short- and Long-Term All-Cause Mortality in Patients with Acute Pancreatitis Admitted to the Intensive Care Unit: A Retrospective Analysis Based on the MIMIC-IV Database. *Frontiers in Immunology*, **15**, Article 1373371. <https://doi.org/10.3389/fimmu.2024.1373371>
- [13] Liu, Q., Zheng, H., Wu, M., Wang, Q., Yan, S., Wang, M., et al. (2022) Association between Lactate-to-Albumin Ratio and 28-Days All-Cause Mortality in Patients with Acute Pancreatitis: A Retrospective Analysis of the MIMIC-IV Database. *Frontiers in Immunology*, **13**, Article 1076121. <https://doi.org/10.3389/fimmu.2022.1076121>
- [14] Liu, S., Zhao, K., Shao, C., Xu, L., Cui, X. and Wang, Y. (2024) Association between Alkaline Phosphatase to Albumin Ratio and Mortality among Patients with Sepsis. *Scientific Reports*, **14**, Article No. 3170. <https://doi.org/10.1038/s41598-024-53384-7>
- [15] Ye, D., Jiang, W. and Gu, D. (2025) Association between Platelet-Albumin-Bilirubin Grade and the 30-Day Mortality in Patients with Acute Respiratory Distress Syndrome: Evidence from the MIMIC-IV Database. *Balkan Medical Journal*, **42**, 66-74. <https://doi.org/10.4274/balkanmedj.galenos.2024.2024-8-7>

-
- [16] Yang, L., Cao, S., Chen, M., Zhang, J., He, C. and Wang, W. (2024) Association of Serum Albumin-to-Creatinine Ratio with In-Hospital Mortality in Patients with Severe Acute Pancreatitis: A Retrospective Study. *BMC Gastroenterology*, **24**, Article No. 401. <https://doi.org/10.1186/s12876-024-03493-4>
- [17] Oh, T.K., Song, I. and Lee, J.H. (2018) Clinical Usefulness of C-Reactive Protein to Albumin Ratio in Predicting 30-Day Mortality in Critically Ill Patients: A Retrospective Analysis. *Scientific Reports*, **8**, Article No. 14977. <https://doi.org/10.1038/s41598-018-33361-7>
- [18] Yu, Y., Wu, W., Dong, Y. and Li, J. (2021) C-reactive Protein-to-Albumin Ratio Predicts Sepsis and Prognosis in Patients with Severe Burn Injury. *Mediators of Inflammation*, **2021**, Article ID: 6621101. <https://doi.org/10.1155/2021/6621101>
- [19] Wu, W., Zhang, Y., Pan, Y., He, Z., Tan, Y., Wang, D., *et al.* (2024) Predictive Value of C-Reactive Protein/Albumin Ratio for Acute Kidney Injury in Patients with Acute Pancreatitis. *Journal of Inflammation Research*, **17**, 5495-5507. <https://doi.org/10.2147/jir.s473466>
- [20] Kui, B., Pintér, J., Molontay, R., Nagy, M., Farkas, N., Gede, N., *et al.* (2022) EASY-APP: An Artificial Intelligence Model and Application for Early and Easy Prediction of Severity in Acute Pancreatitis. *Clinical and Translational Medicine*, **12**, e842. <https://doi.org/10.1002/ctm2.842>
- [21] Zhang, W., Chang, Y., Cheng, C., Zhao, X., Tang, X., Lu, F., *et al.* (2024) A Machine Learning Model for Predicting Acute Kidney Injury Secondary to Severe Acute Pancreatitis. *Chinese Medical Journal*, **137**, 619-621. <https://doi.org/10.1097/cm9.0000000000003027>