

基于多维临床特征与机器学习构建急性胰腺炎复发(RAP)的预测模型及临床验证

苏 俊, 柯庭威, 李海山*

蚌埠医科大学附属合肥市第二人民医院急诊科, 安徽 合肥

收稿日期: 2026年6月17日; 录用日期: 2026年7月11日; 发布日期: 2026年7月21日

摘 要

目的: 探讨急性胰腺炎(AP)患者发生复发性急性胰腺炎(RAP)的独立危险因素, 并结合多维临床特征构建列线图(Nomogram)预测模型。方法: 回顾性分析2023年1月至2025年1月期间蚌埠医科大学附属合肥市第二人民医院收治的380例AP患者。通过LASSO回归与Boruta算法进行特征筛选, 多因素Logistic回归确定预测变量, 并结合血钙分类化、稳健性分析及病因分层分析对关键变量进行解释。结果: 380例患者中位随访时间为18个月(IQR: 12~24个月), RAP发生率为31.8%。多因素分析显示血钙、吸烟史、饮酒史、甘油三酯(TG)、尿酸(UA)、丙氨酸氨基转移酶(ALT)、入院年龄及白蛋白(ALB)是RAP发生的预测因素。血钙分类化及稳健性分析未改变其与RAP风险升高相关的总体方向; 年龄和ALT的负向关联主要结合病因构成及病因去除干预进行解释。验证集AUC为0.822 (95%CI: 0.730~0.903), 校准曲线显示预测概率与实际发生率具有良好一致性。结论: 本研究构建的Nomogram模型有助于临床早期识别RAP高危患者, 辅助个体化随访和二级预防决策。

关键词

急性胰腺炎, 复发, 机器学习, 列线图, 预测模型

Construction and Clinical Validation of a Prediction Model for Recurrent Acute Pancreatitis (RAP) Based on Multidimensional Clinical Features and Machine Learning

Jun Su, Tingwei Ke, Haishan Li*

*通讯作者。

文章引用: 苏俊, 柯庭威, 李海山. 基于多维临床特征与机器学习构建急性胰腺炎复发(RAP)的预测模型及临床验证[J]. 临床医学进展, 2026, 16(7): 1610-1622. DOI: 10.12677/acm.2026.1672682

Department of Emergency, The Second People's Hospital of Hefei Affiliated to Bengbu Medical University, Hefei Anhui

Received: June 17, 2026; accepted: July 11, 2026; published: July 21, 2026

Abstract

Objective: To investigate the predictive factors for recurrent acute pancreatitis (RAP) in patients with acute pancreatitis (AP) and to construct a nomogram prediction model based on multidimensional clinical features. **Methods:** A retrospective analysis was conducted on 380 AP patients admitted to the Second People's Hospital of Hefei, Affiliated to Bengbu Medical University, from January 2023 to January 2025. Feature selection was performed using LASSO regression and the Boruta algorithm, and predictive variables were determined by multivariate logistic regression. Categorized serum calcium analysis, robustness assessment, and etiology-based stratified analyses were further used to interpret key variables. **Results:** During a median follow-up of 18 months (IQR: 12~24 months), the incidence of RAP among the 380 patients was 31.8%. Multivariate analysis showed that serum calcium, smoking history, drinking history, triglycerides (TG), uric acid (UA), alanine aminotransferase (ALT), admission age, and albumin (ALB) were predictive factors for RAP. Categorized and robust analyses of serum calcium showed a consistent direction of association with increased RAP risk, whereas the inverse associations of age and ALT were mainly interpreted in relation to etiological composition and etiological interventions. The AUC of the validation set was 0.822 (95%CI: 0.730~0.903), and the calibration curve showed good consistency between the predicted probability and the actual incidence. **Conclusion:** The nomogram model constructed in this study helps identify high-risk RAP patients and supports individualized follow-up and secondary prevention decision-making.

Keywords

Acute Pancreatitis, Recurrence, Machine Learning, Nomogram, Prediction Model

Copyright © 2026 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. 引言

急性胰腺炎(Acute Pancreatitis, AP)是消化系统最常见的急腹症之一,其发病率在全球范围内呈逐年上升趋势[1]-[3]。尽管大多数AP患者在首次发作后能够完全康复,但仍有约21%的患者会复发,即复发性急性胰腺炎(Recurrent Acute Pancreatitis, RAP) [4]。RAP不仅显著增加了患者的医疗经济负担,降低了生活质量,更是进展为慢性胰腺炎(Chronic Pancreatitis, CP)和胰腺癌的重要危险因素[5]-[7]。研究表明,经历过RAP的患者进展为CP的风险是仅有单次AP发作患者的数倍[6]。因此,早期识别具有高复发风险的AP患者,并针对性地制定个体化的二级预防方案,对于改善患者的长期预后具有重要的临床意义。

目前,关于RAP危险因素的研究已取得一定进展。已知胆道疾病、高甘油三酯血症、长期饮酒以及某些遗传突变是导致AP复发的常见病因[8]-[10]。然而,在临床实践中,许多患者在首次发作时并未表现出明确的病因,或者存在多种潜在危险因素的交互作用,这使得RAP的预测变得极具挑战性。传统的预测方法往往依赖于单一的临床指标或简单的评分系统,难以全面捕捉疾病的复杂性,导致预测效能有

限[11]。近年来,随着医疗大数据的快速发展,机器学习算法在疾病风险预测领域展现出巨大的潜力。与传统统计学方法相比,机器学习能够处理高维、非线性的临床数据,通过多重算法的交叉验证,更精准地筛选出核心预测变量,从而构建出更为稳健和准确的预测模型。

本研究旨在利用机器学习算法,整合 AP 患者出院时的多维临床特征,包括人口学信息、既往病史以及实验室检验指标,深入探讨 RAP 的独立危险因素。通过 LASSO 回归与 Boruta 算法的双重特征筛选,结合多因素 Logistic 回归分析,构建一个可视化、易于临床应用的列线图(Nomogram)预测模型。我们期望该模型能够为临床医生提供一个直观、可靠的风险评估工具,助力 RAP 的高危人群筛查与个体化管理,最终降低 AP 的复发率,改善患者的远期预后。

2. 资料与方法

2.1. 研究对象

本研究采用回顾性队列研究设计,连续收集 2023 年 1 月至 2025 年 1 月期间在蚌埠医科大学附属合肥市第二人民医院消化内科及急诊科收治的急性胰腺炎患者的临床资料。

纳入标准: (1) 符合《中国急性胰腺炎诊治指南(2021)》[8]中 AP 的诊断标准,即具备以下 3 项中的 2 项及以上:① 持续性的上腹部剧烈疼痛,常向背部放射;② 血清淀粉酶或脂肪酶水平升高至正常值上限的 3 倍及以上;③ 腹部影像学(如增强 CT 或 MRI)表现出 AP 的典型特征。(2) 年龄 ≥ 18 岁。(3) 临床资料和实验室检验数据完整。

排除标准: (1) 既往已确诊为慢性胰腺炎或胰腺癌的患者。(2) 合并严重感染、恶性肿瘤或自身免疫性疾病的患者。(3) 妊娠期或哺乳期妇女。(4) 关键临床指标或随访数据缺失的患者。

RAP 的定义: 患者在首次 AP 发作临床症状完全缓解、实验室指标恢复正常且出院至少 30 天后,再次出现符合 AP 诊断标准的急性发作。

根据上述标准,最终纳入 380 例患者进行分析。纳入的 380 例患者按 7:3 的比例随机分为训练集($n = 266$)和验证集($n = 114$)。训练集用于特征筛选和模型构建,验证集用于评估模型的泛化能力和预测性能。

2.2. 临床资料收集

通过医院电子病历系统(EMR)系统地收集患者的临床资料。为更好地反映疾病转归状态并辅助个性化随访决策,所有实验室指标均采集自出院时。收集的变量包括: (1) 人口学特征:性别、入院年龄。(2) 既往病史与生活习惯:糖尿病、高血压、吸烟史、饮酒史、既往胰腺炎史。(3) 初次 AP 病因学分类:结合病史、实验室检查、腹部超声/CT/MRCP 及内镜资料,将病因分为胆源性、高甘油三酯血症性、酒精性、特发性及其他病因。(4) 严重程度分级:依据《中国急性胰腺炎诊治指南(2021)》[8],将患者分为轻症急性胰腺炎(MAP)、中度重症急性胰腺炎(MSAP)和重症急性胰腺炎(SAP)。(5) 关键治疗干预措施:记录胆源性 AP 患者是否接受胆囊切除术或 ERCP,高甘油三酯血症性 AP 患者是否接受降脂治疗及饮食控制,酒精相关 AP 患者是否接受戒酒干预或相关健康教育。(6) 出院时实验室检验指标:甘油三酯(TG)、尿酸(UA)、血钙(Ca)、丙氨酸氨基转移酶(ALT)、白蛋白(ALB)、空腹血糖(GLU)、白细胞计数(WBC)、中性粒细胞与淋巴细胞比值(NLR)、C 反应蛋白(CRP)。

2.3. 统计学分析与模型构建

所有数据的统计分析与可视化均使用 Python (版本 3.11)进行。连续变量若符合正态分布,以均数 $\pm$ 标准差($\bar{x} \pm s$)表示,组间比较采用独立样本 t 检验;若呈偏态分布,则以中位数和四分位数间距[M(P25,

P75)]表示, 组间比较采用 Mann-Whitney U 检验。分类变量以频数和百分比[n (%)]表示, 组间比较采用 χ^2 检验或 Fisher 精确检验。

在模型构建阶段, 首先在训练集中对所有收集的临床变量进行单因素 Logistic 回归分析, 初步筛选出与 RAP 发生相关的潜在危险因素。随后, 采用 LASSO (Least Absolute Shrinkage and Selection Operator) 回归[12]和 Boruta 算法[13]进行双重特征筛选。LASSO 回归通过十折交叉验证确定最优惩罚参数 λ , 筛选出非零系数的特征; Boruta 算法则基于随机森林框架, 通过比较原始特征与影子特征的重要性得分, 进一步确认核心预测变量。

取上述两种机器学习算法筛选结果的并集作为核心变量, 纳入多因素 Logistic 回归模型, 计算各变量的比值比(Odds Ratio, OR)及其 95%置信区间(Confidence Interval, CI), 确定 RAP 的预测因素。针对血钙连续变量 OR 值较高且置信区间较宽的特点, 本研究同步开展血钙分类化和稳健性分析: 一方面按血钙四分位数重新纳入 Logistic 模型, 观察较高血钙水平与 RAP 风险的趋势关系; 另一方面采用异常值缩尾处理和稳健标准误检验结果方向的一致性。对于年龄和 ALT 在多因素模型中的负向关联, 本研究进一步结合初次 AP 病因学分类和关键治疗干预进行病因分层分析, 并设置病因×年龄、病因×ALT 交互项, 以判断其统计关联是否受病因构成和病因去除治疗影响。基于多因素回归结果, 构建预测 RAP 发生风险的列线图(Nomogram)模型。模型性能以 ROC 曲线、AUC、校准曲线和 Brier 得分评价。所有统计检验均为双侧, $P < 0.05$ 为差异有统计学意义。

3. 结果

3.1. 基线特征分析

本研究共纳入 380 例 AP 患者, 其中男性 244 例(64.2%), 女性 136 例(35.8%)。队列的中位随访时间为 18 个月(IQR: 12~24 个月)。在随访期间, 共有 121 例(31.8%)患者发生 RAP, 259 例(68.2%)患者未发生复发(首发组)。两组患者的基线特征比较如表 1 所示。结果显示, RAP 组患者的入院年龄显著低于首发组(37.2 岁 vs 43.8 岁, $P < 0.001$)。在生活习惯方面, RAP 组有吸烟史的比例显著高于首发组(57.0% vs 22.8%, $P < 0.001$), 饮酒史比例也显著升高(44.6% vs 19.3%, $P < 0.001$)。出院时实验室指标方面, RAP 组的甘油三酯(TG)水平显著高于首发组(4.06 mmol/L vs 2.81 mmol/L, $P < 0.001$), 尿酸(UA)水平显著升高(350.5 μ mol/L vs 330.8 μ mol/L, $P = 0.010$), 血钙(Ca)水平也显著升高(2.23 mmol/L vs 2.11 mmol/L, $P < 0.001$)。此外, RAP 组的丙氨酸氨基转移酶(ALT)水平显著低于首发组(27.5 U/L vs 43.8 U/L, $P < 0.001$), 白蛋白(ALB)水平也显著降低(35.4 g/L vs 38.3 g/L, $P < 0.001$)。

Table 1. Comparison of baseline characteristics between the two groups
表 1. 两组患者基线特征比较

变量	首发组(n = 259)	RAP 组(n = 121)	统计值	P 值
人口学特征				
性别(男), n (%)	167 (64.5%)	77 (63.6%)	$\chi^2 = 0.002$	0.964
入院年龄(岁)	43.80 (36.80, 52.80)	37.20 (28.70, 47.60)	U = 11454.0	<0.001
既往病史				
糖尿病, n (%)	66 (25.5%)	35 (28.9%)	$\chi^2 = 0.340$	0.560
高血压, n (%)	86 (33.2%)	41 (33.9%)	$\chi^2 = 0.000$	0.989
吸烟史, n (%)	59 (22.8%)	69 (57.0%)	$\chi^2 = 41.777$	<0.001

续表

饮酒史, n (%)	50 (19.3%)	54 (44.6%)	$\chi^2 = 25.346$	<0.001
既往胰腺炎史, n (%)	47 (18.1%)	23 (19.0%)	$\chi^2 = 0.004$	0.952
出院时实验室指标				
甘油三酯(mmol/L)	2.81 (1.73, 4.84)	4.06 (2.34, 6.54)	U = 18956.5	<0.001
尿酸(μ mol/L)	330.80 (265.25, 390.05)	350.50 (298.70, 418.90)	U = 18226.5	0.010
血钙(mmol/L)	2.11 (1.95, 2.26)	2.23 (2.06, 2.40)	U = 20471.5	<0.001
丙氨酸氨基转移酶(U/L)	43.80 (24.40, 75.55)	27.50 (12.60, 43.40)	U = 10395.0	<0.001
白蛋白(g/L)	38.30 (35.00, 41.85)	35.40 (30.50, 39.60)	U = 11056.0	<0.001
空腹血糖(mmol/L)	7.14 (5.39, 9.40)	6.74 (5.43, 8.61)	U = 13410.0	0.272
白细胞计数($\times 10^9$ /L)	7.68 (5.67, 9.90)	8.02 (5.93, 10.39)	U = 16848.0	0.238
NLR	6.80 (4.94, 9.96)	6.47 (4.40, 9.90)	U = 13389.0	0.234
C 反应蛋白(mg/L)	12.60 (4.70, 28.30)	26.85 (15.85, 43.98)	U = 19268.5	<0.001

3.2. 独立危险因素的筛选与多因素分析

3.2.1. 单因素分析

在训练集中,对 16 个潜在的临床和实验室变量进行单因素 Logistic 回归分析。结果显示,入院年龄、吸烟史、饮酒史、甘油三酯、尿酸、血钙、丙氨酸氨基转移酶、白蛋白、C 反应蛋白等变量与 RAP 的发生显著相关($P < 0.05$)。单因素分析结果见表 2,森林图见图 1。

Table 2. Univariate Logistic regression analysis of risk factors for AP recurrence

表 2. AP 复发危险因素单因素 Logistic 回归分析结果

变量	OR 值	95%CI	P 值
入院年龄(岁)	0.956	0.934~0.978	<0.001
性别(男)	1.061	0.614~1.832	0.833
糖尿病	1.423	0.795~2.546	0.235
高血压	0.950	0.537~1.681	0.861
吸烟史	3.841	2.216~6.659	<0.001
饮酒史	3.626	2.050~6.414	<0.001
既往胰腺炎史	1.122	0.576~2.187	0.735
甘油三酯(mmol/L)	1.096	1.015~1.183	0.019
尿酸(μ mol/L)	1.003	1.000~1.007	0.032
血钙(mmol/L)	10.831	3.219~36.443	<0.001
丙氨酸氨基转移酶(U/L)	0.979	0.970~0.989	<0.001
白蛋白(g/L)	0.914	0.871~0.958	<0.001
空腹血糖(mmol/L)	0.928	0.846~1.018	0.113
白细胞计数($\times 10^9$ /L)	1.042	0.972~1.118	0.244
NLR	0.980	0.920~1.043	0.522
C 反应蛋白(mg/L)	1.016	1.007~1.026	<0.001

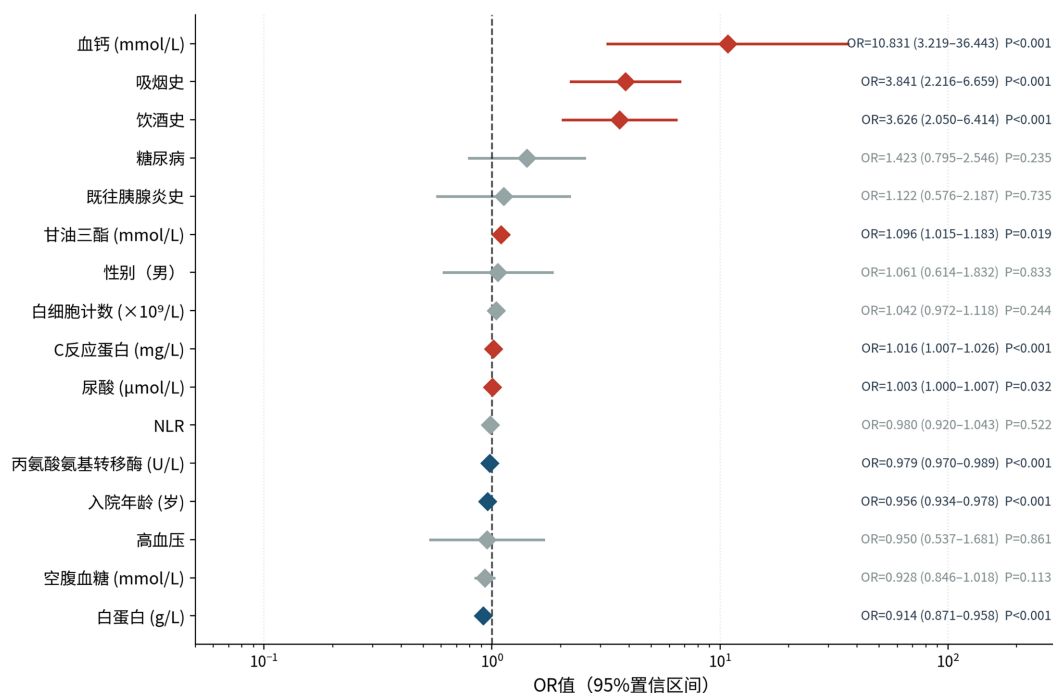


Figure 1. Forest plot of univariate Logistic regression analysis

图 1. 单因素 Logistic 回归分析森林图

3.2.2. LASSO 回归降维

为了进一步筛选核心预测变量并消除多重共线性, 采用 LASSO 回归对单因素分析中具有潜在意义的变量进行降维。通过十折交叉验证, 当惩罚参数 λ 取最优值时, 模型达到最优性能, 此时筛选出非零系数的特征。LASSO 特征筛选路径与交叉验证曲线如图 2 所示[12]。

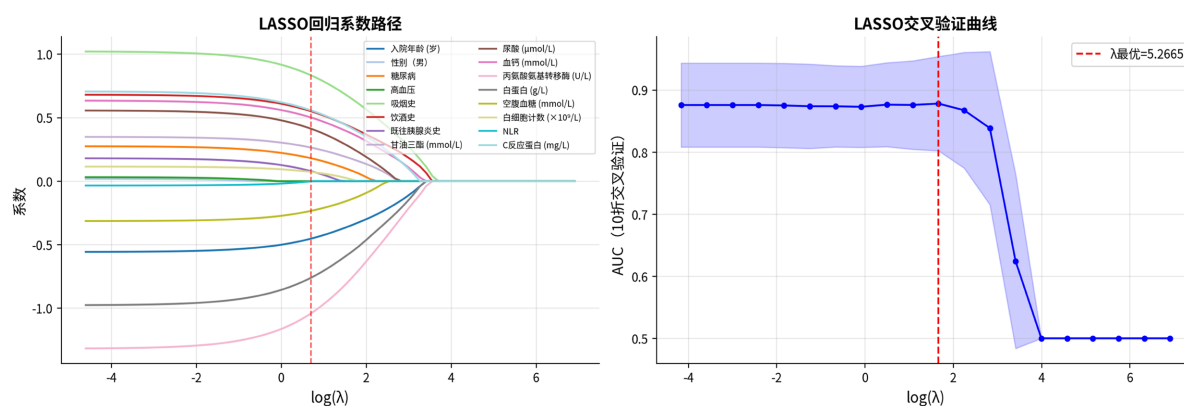


Figure 2. LASSO feature selection path and cross-validation curve

图 2. LASSO 特征筛选路径与交叉验证曲线

3.2.3. Boruta 算法验证

为确保特征筛选的稳健性, 本研究同时采用 Boruta 算法对上述变量进行重要性评估[13]。Boruta 算法基于随机森林框架, 通过比较原始特征与影子特征的重要性得分, 将特征分为“确认特征”、“待定特征”和“拒绝特征”。结果显示, 丙氨酸氨基转移酶、C 反应蛋白、白蛋白、血钙、入院年龄、甘油三酯、

尿酸、吸烟史等变量被确认为对 RAP 预测具有重要贡献的核心特征。Boruta 算法特征重要性排序如图 3 所示。

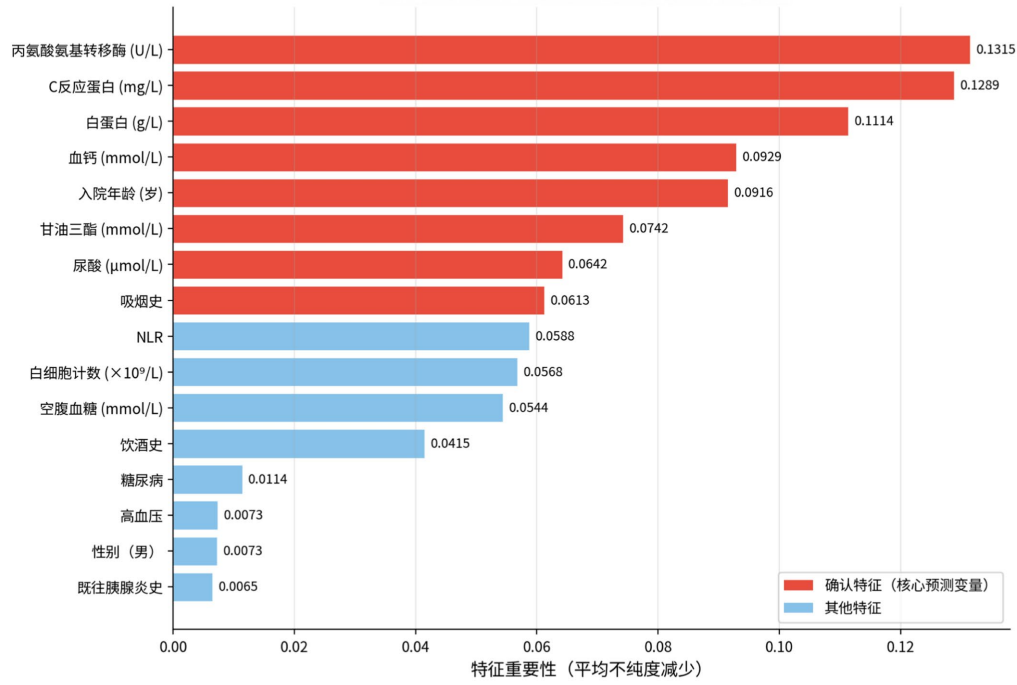


Figure 3. Feature importance ranking by Boruta algorithm
图 3. Boruta 算法特征重要性排序

3.2.4. 多因素 Logistic 回归

综合 LASSO 回归与 Boruta 算法的筛选结果, 将血钙、吸烟史、饮酒史、甘油三酯、尿酸、丙氨酸氨基转移酶、入院年龄、白蛋白这 8 个核心变量纳入多因素 Logistic 回归模型。分析结果表明, 血钙($OR = 13.753$, $95\%CI: 2.729 \sim 69.310$, $P = 0.001$)、吸烟史($OR = 5.388$, $95\%CI: 2.633 \sim 11.026$, $P < 0.001$)、饮酒史($OR = 3.545$, $95\%CI: 1.669 \sim 7.527$, $P < 0.001$)和尿酸($OR = 1.005$, $95\%CI: 1.001 \sim 1.009$, $P = 0.016$)与 RAP 发生风险升高相关; 丙氨酸氨基转移酶($OR = 0.977$, $95\%CI: 0.967 \sim 0.988$, $P < 0.001$)、入院年龄($OR = 0.959$, $95\%CI: 0.932 \sim 0.987$, $P = 0.004$)和白蛋白($OR = 0.879$, $95\%CI: 0.826 \sim 0.936$, $P < 0.001$)与 RAP 发生呈负向关联。血钙 OR 值较高主要与其作为连续变量时的单位尺度有关, 即 OR 表示血钙每增加 1 mmol/L 时 RAP 风险的相对变化; 考虑到临床血钙波动范围较窄, 本文同时结合分类化和稳健性分析解释其结果。年龄和 ALT 的负向关联不直接等同于直接降低复发风险的生物学效应, 而需结合病因学分类和病因去除干预进行解释。虽然甘油三酯($OR = 1.108$, $95\%CI: 0.999 \sim 1.229$, $P = 0.052$)在多因素模型中未达到传统统计学显著性($P < 0.05$), 但考虑到其在单因素分析中的显著性、机器学习算法推荐以及与 RAP 相关的明确病理生理基础 [14]-[17], 本研究仍将其保留在最终预测模型中。多因素分析结果见表 3。

3.2.5. 血钙分类化、稳健性及病因分层分析

为降低连续变量单位尺度对血钙 OR 值的影响, 本研究进一步将血钙按四分位数进行分类化处理, 并结合异常值缩尾处理和稳健标准误进行敏感性分析。结果显示, 较高血钙水平与 RAP 风险升高保持同向关系, 未出现方向性逆转, 提示血钙与 RAP 复发风险的关联并非单纯由连续变量尺度或个别异常值造成。病因分层和交互作用分析显示, 入院年龄及 ALT 与 RAP 的负向关联在不同病因亚组中存在差异,

Table 3. Multivariate Logistic regression analysis of risk factors for AP recurrence
表 3. AP 复发危险因素多因素 Logistic 回归分析结果

变量	OR 值	95%CI (下限~上限)	P 值
血钙(mmol/L, 连续变量)	13.753	2.729~69.310	0.001
吸烟史	5.388	2.633~11.026	<0.001
饮酒史	3.545	1.669~7.527	<0.001
甘油三酯(mmol/L)	1.108	0.999~1.229	0.052
尿酸(μmol/L)	1.005	1.001~1.009	0.016
丙氨酸氨基转移酶(U/L)	0.977	0.967~0.988	<0.001
入院年龄(岁)	0.959	0.932~0.987	0.004
白蛋白(g/L)	0.879	0.826~0.936	<0.001

注：血钙以连续变量形式进入主模型，OR 值表示血钙每增加 1 mmol/L 时 RAP 风险的相对变化；分类化和稳健性分析用于辅助判断该关联方向的稳定性。

其中胆源性 AP 患者接受胆囊切除术或 ERCP 等病因去除干预后复发风险下降，可部分解释 ALT 升高与较低复发风险之间的关联；而高甘油三酯血症性、酒精性及特发性 AP 患者更常见于较年轻人群，其复发风险更容易受到持续代谢异常和生活方式因素影响。因此，年龄和 ALT 在本模型中更适合作为反映病因背景和治疗路径差异的预测变量，而不应简单解释为独立保护性因素。

3.3. 列线图预测模型的构建

基于多因素 Logistic 回归分析确定的 8 个核心预测变量(血钙、吸烟史、饮酒史、甘油三酯、尿酸、丙氨酸氨基转移酶、入院年龄、白蛋白)，构建了预测 RAP 发生风险的列线图(Nomogram)模型。在该列线图中，每个变量对应一个特定的评分(Points)，将所有变量的评分相加得到总分(Total Points)，总分向下对应即可得出患者发生 RAP 的预测概率。临床应用时，血钙、年龄和 ALT 的评分结果需结合血钙分类化结果、初次 AP 病因学分类、严重程度分级及关键治疗干预进行综合判断。列线图模型如图 4 所示。

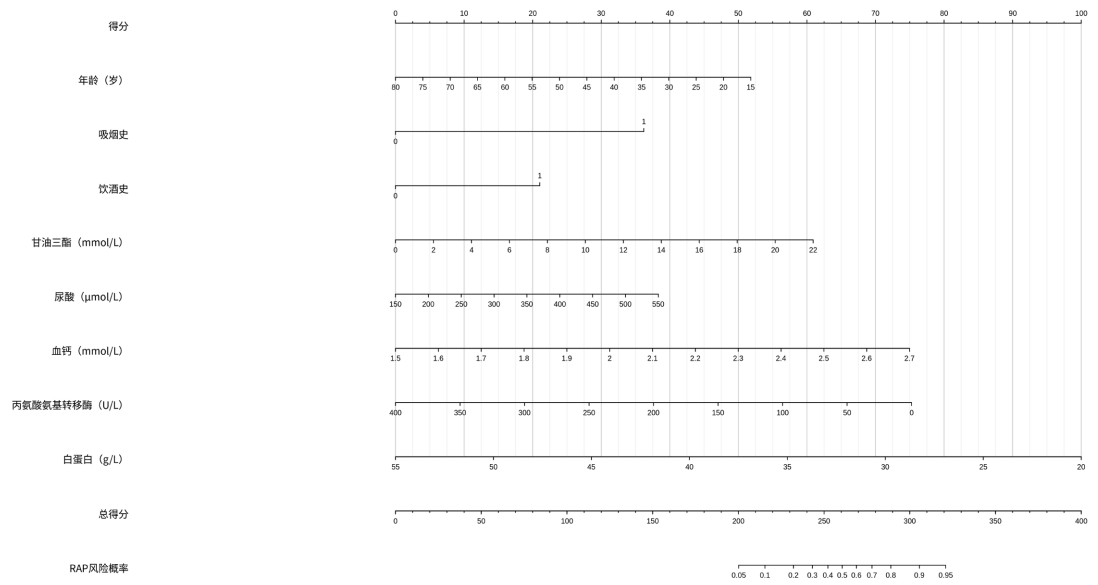


Figure 4. Nomogram for predicting RAP risk
图 4. RAP 风险预测列线图(Nomogram)

3.4. 预测模型的性能评估与验证

3.4.1. 模型区分度评估

采用受试者工作特征(ROC)曲线评估列线图模型的区分度。在训练集中,该模型的曲线下面积(AUC)为 0.876 (95%CI: 0.830~0.918);在验证集中, AUC 达到 0.822 (95%CI: 0.730~0.903),表明该模型在区分 RAP 患者与非 RAP 患者方面具有良好的性能。训练集与验证集的 ROC 曲线如图 5 所示。

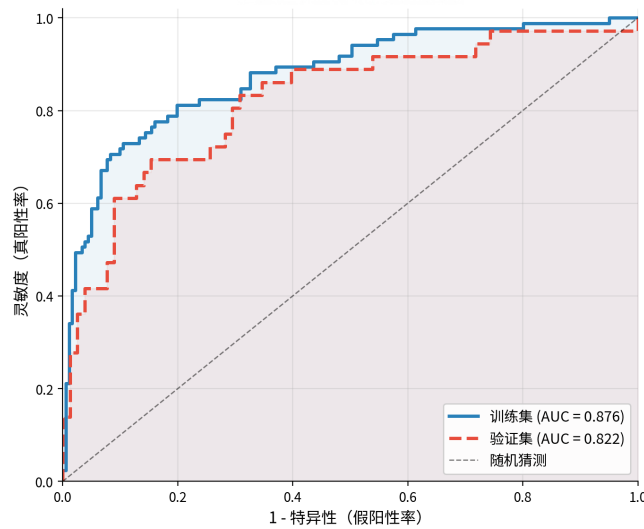


Figure 5. ROC curves of training set and validation set
图 5. 训练集与验证集 ROC 曲线

3.4.2. 模型校准度评估

通过绘制校准曲线(Calibration curve)评估模型预测概率与实际发生率之间的一致性。结果显示,无论是在训练集还是验证集中,校准曲线均紧密贴合理想的 45 度对角线,表明模型的预测概率与实际观察到的 RAP 发生率高度一致。训练集的 Brier 得分为 0.1136,验证集的 Brier 得分为 0.1562,进一步证实了模型具有优良的校准度。训练集与验证集的校准曲线如图 6 所示。

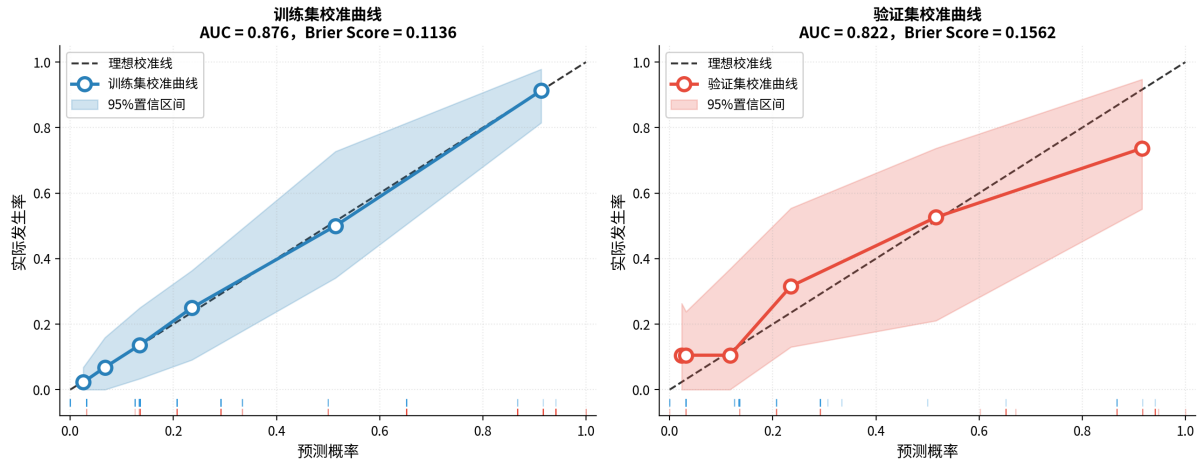


Figure 6. Calibration curves of training set and validation set
图 6. 训练集与验证集校准曲线

Table 4. Performance evaluation of the prediction model**表 4.** 预测模型性能评估

评估指标	训练集(n = 266)	验证集(n = 114)
AUC (95%CI)	0.876 (0.830~0.918)	0.822 (0.730~0.903)
Brier Score	0.1136	0.1562
最佳截断值(Youden 指数)	0.434	0.434
灵敏度	72.9%	61.1%
特异性	89.5%	88.5%
阳性预测值(PPV)	76.5%	71.0%
阴性预测值(NPV)	87.6%	83.1%

AUC: 曲线下面积(Area Under the Curve); Brier Score 越小表示校准度越好(完美校准 = 0)。

4. 讨论

本研究通过回顾性分析 380 例 AP 患者的临床资料, 利用 LASSO 回归与 Boruta 算法双重机器学习特征筛选策略, 结合多因素 Logistic 回归, 构建了一个包含血钙、吸烟史、饮酒史、甘油三酯、尿酸、丙氨酸氨基转移酶、入院年龄和白蛋白 8 个核心变量的 RAP 预测模型。该模型在验证集中表现出良好的区分度(AUC=0.822)和校准度, 为临床早期识别 RAP 高危患者提供了直观、可靠的工具。与单纯报告回归结果相比, 本研究进一步结合血钙分类化、稳健性分析和病因分层结果对关键变量进行解释, 使模型结果更符合 AP 复发的临床异质性。训练集与验证集的预测模型性能评估如表 4 所示。

4.1. 出院时实验室指标的临床价值

本研究创新性地采用出院时实验室指标构建预测模型, 具有以下优势: 首先, 出院时指标反映了急性炎症消退后的代谢状态, 更能体现患者的基础病理生理特征(如持续高甘油三酯血症、高尿酸血症), 这些是 RAP 复发的根本驱动因素。其次, 出院时指标可直接用于出院评估, 帮助临床医生在患者离院前完成风险分层, 制定个体化随访计划(如高危患者缩短随访间隔、强化代谢管理)。第三, 避免了入院时急性炎症反应的干扰(如 ALT、WBC 急性期升高), 可能掩盖患者的真实代谢状态。最后, 出院时常规检测这些指标无需额外检测, 操作简便, 易于推广。

4.2. 高甘油三酯血症与 RAP 的关联

本研究结果显示, 甘油三酯(TG)水平升高是 RAP 发生的重要危险因素。这一发现与既往多项关于高甘油三酯血症性急性胰腺炎(HTG-AP)高复发率的研究结果高度一致[14]-[17]。高水平的 TG 在胰腺局部被胰脂肪酶水解, 产生大量游离脂肪酸(FFA)。过量的 FFA 不仅具有直接的细胞毒性, 能破坏胰腺腺泡细胞和毛细血管内皮细胞, 还会引发强烈的全身炎症反应和微循环障碍, 导致胰腺组织缺血坏死[14]。此外, TG 水平的波动往往与患者的饮食习惯、脂代谢异常及遗传易感性密切相关, 这些因素在首次 AP 发作后若未得到有效控制, 极易导致疾病的反复发作[16]。因此, 对于合并高甘油三酯血症的 AP 患者, 出院后长期的降脂治疗和严格的饮食管理是预防 RAP 的关键。

4.3. 尿酸与血钙在 RAP 预测中的价值

本研究发现, 血清尿酸(UA)水平升高同样与 RAP 风险增加相关(OR = 1.005, P = 0.016)。近年来, 越来越多的证据表明, 高尿酸血症不仅是痛风和心血管疾病的危险因素, 还参与了 AP 的病理生理过程[18]。

高水平尿酸可能通过氧化应激、局部炎症反应及 NLRP3 炎症小体相关通路促进胰腺损伤,从而增加复发风险[18]。血钙在本研究多因素模型中显示出较高 OR 值($OR = 13.753$, $95\%CI: 2.729 \sim 69.310$, $P = 0.001$)。该结果需结合变量尺度进行理解:当前 OR 值基于血钙每增加 1 mmol/L 计算,而临床中血钙通常集中在较窄范围内, 1 mmol/L 代表较大的生理跨度,因而可能表现为较高的 OR 值和较宽的置信区间。分类化及稳健性分析显示,较高血钙水平与 RAP 风险升高的总体方向保持一致。机制上,高钙状态可能通过促进胰管内钙盐沉积、增强腺泡细胞内钙超载和激活胰蛋白酶原,增加胰腺炎复发风险[19]。因此,血钙可作为 RAP 风险评估中的重要代谢指标,但其效应量解释应同时参考连续变量模型和分类化模型。

4.4. 年龄与肝功能指标负向关联的病因学解释

本研究中入院年龄($OR = 0.959$, $P = 0.004$)和丙氨酸氨基转移酶(ALT, $OR = 0.977$, $P < 0.001$)在多因素模型中与 RAP 发生呈负向关联。该结果不宜被简单表述为年龄或 ALT 本身具有“保护作用”,而应从 AP 病因构成和病因去除干预角度进行解释。年轻患者中,酒精性、高甘油三酯血症性、遗传性或特发性 AP 比例相对较高,这些病因常与长期生活方式、脂代谢异常或遗传易感性相关,若首次发作后未能有效控制,复发风险可能持续存在[20]。相反,老年患者中胆源性 AP 更为常见,而胆源性 AP 患者在首次发作后接受胆囊切除术或 ERCP 等干预后,病因得以去除,复发风险可明显下降[21] [22]。ALT 升高通常提示胆道梗阻或胆源性 AP 倾向,因此 ALT 与较低复发风险之间的关联更可能是胆源性 AP 经病因处理后复发率下降的间接反映,而非 ALT 本身对 RAP 具有保护性作用。

本研究的病因分层和交互作用分析进一步支持了上述解释:年龄和 ALT 的效应方向及强度在不同病因亚组中并不完全一致,提示其统计学关联受到胆源性、酒精性、高甘油三酯血症性和特发性 AP 构成差异的影响。因而,在临床应用列线图时,对于年龄较轻且 ALT 较低的患者,应重点追问饮酒、血脂控制、既往代谢异常及特发性复发倾向;对于 ALT 较高且符合胆源性 AP 特征的患者,则应结合胆囊切除术、ERCP 等病因去除措施完成复发风险再评估。

4.5. 模型的临床应用与局限性

本研究构建的预测模型将复杂的回归方程转化为直观的评估工具,临床医生只需根据患者出院时的常规指标,即可快速估计其发生 RAP 的概率。这对于优化医疗资源配置、制定个体化随访计划具有重要实用价值。例如,对于模型预测为高危的患者,应加强出院后的健康教育,密切监测血脂、尿酸和血钙等代谢指标,并结合病因学分类采取针对性干预:胆源性 AP 患者应关注胆囊切除术或 ERCP 等病因去除措施的落实,高甘油三酯血症性 AP 患者应强化降脂治疗和饮食管理,酒精相关 AP 患者则需加强戒酒干预和长期随访。对于年龄较轻、ALT 较低而伴随吸烟、饮酒或 TG 升高的患者,不应因年龄或 ALT 在模型中的负向关联而降低警惕,而应结合其病因背景进行更积极的复发预防。

然而,本研究也存在一定的局限性。首先,作为一项单中心回顾性研究,样本量相对有限,且可能存在选择偏倚。其次,本研究虽然补充追溯了初次 AP 病因学分类、严重程度分级及关键治疗干预措施,并将其用于关键变量解释、分层和交互作用分析,但回顾性资料对于院外干预时机、长期降脂治疗依从性、戒酒和饮食控制持续性的记录仍不如前瞻性随访完整,可能存在信息偏倚。再次,血钙作为连续变量进入主模型时出现 OR 值较高且 $95\%CI$ 较宽,提示效应量解释需要结合分类化模型和稳健性分析,而不能仅依据连续变量 OR 进行临床推断。此外,本研究的随访时间相对有限,部分远期复发病例可能未被记录,从而低估真实复发率。未来仍需开展多中心、大样本、前瞻性队列研究,进一步验证和优化该预测模型。

综上所述,本研究基于多维临床特征与机器学习算法,构建了预测急性胰腺炎患者复发风险的模型。

该模型包含血钙、吸烟史、饮酒史、甘油三酯、尿酸、丙氨酸氨基转移酶、入院年龄和白蛋白 8 个核心预测变量,具有良好的区分度(验证集 AUC = 0.822)和校准度(Brier Score = 0.1562)。血钙在主模型中表现出较高 OR 值,结合分类化和稳健性分析后,其与 RAP 风险升高的总体方向保持一致;年龄和 ALT 的负向关联则更适合从病因构成和病因去除干预角度解释,而不宜简单归纳为保护性因素。该模型不仅有助于临床医生早期识别 RAP 高危人群,还可结合初次 AP 病因学分类、严重程度分级及关键治疗措施,为制定个体化二级预防策略和随访计划提供依据。

声 明

本研究已获得医院伦理委员会的批准(2021-科研-015),并豁免了患者的知情同意。

基金项目

蚌埠医科大学自然科学重点项目(编号: 2024byzd400)。

参考文献

- [1] Boxhoorn, L., Voermans, R.P., Bouwense, S.A., Bruno, M.J., Verdonk, R.C., Boermeester, M.A., *et al.* (2020) Acute Pancreatitis. *The Lancet*, **396**, 726-734. [https://doi.org/10.1016/s0140-6736\(20\)31310-6](https://doi.org/10.1016/s0140-6736(20)31310-6)
- [2] Iannuzzi, J.P., King, J.A., Leong, J.H., Quan, J., Windsor, J.W., Tanyingoh, D., *et al.* (2022) Global Incidence of Acute Pancreatitis Is Increasing over Time: A Systematic Review and Meta-Analysis. *Gastroenterology*, **162**, 122-134. <https://doi.org/10.1053/j.gastro.2021.09.043>
- [3] Petrov, M.S. and Yadav, D. (2019) Global Epidemiology and Holistic Prevention of Pancreatitis. *Nature Reviews Gastroenterology & Hepatology*, **16**, 175-184. <https://doi.org/10.1038/s41575-018-0087-5>
- [4] Hajibandeh, S., Jurdon, R., Heaton, E., Hajibandeh, S. and O'Reilly, D. (2023) The Risk of Recurrent Pancreatitis after First Episode of Acute Pancreatitis in Relation to Etiology and Severity of Disease: A Systematic Review, Meta-Analysis and Meta-Regression Analysis. *Journal of Gastroenterology and Hepatology*, **38**, 1718-1733. <https://doi.org/10.1111/jgh.16264>
- [5] Sankaran, S.J., Xiao, A.Y., Wu, L.M., Windsor, J.A., Forsmark, C.E. and Petrov, M.S. (2015) Frequency of Progression from Acute to Chronic Pancreatitis and Risk Factors: A Meta-Analysis. *Gastroenterology*, **149**, 1490-1500.E1. <https://doi.org/10.1053/j.gastro.2015.07.066>
- [6] Ahmed Ali, U., Issa, Y., Hagenaars, J.C., Bakker, O.J., van Goor, H., Nieuwenhuijs, V.B., *et al.* (2016) Risk of Recurrent Pancreatitis and Progression to Chronic Pancreatitis after a First Episode of Acute Pancreatitis. *Clinical Gastroenterology and Hepatology*, **14**, 738-746. <https://doi.org/10.1016/j.cgh.2015.12.040>
- [7] Szatmary, P., Grammatikopoulos, T., Cai, W., Huang, W., Mukherjee, R., Halloran, C., *et al.* (2022) Acute Pancreatitis: Diagnosis and Treatment. *Drugs*, **82**, 1251-1276. <https://doi.org/10.1007/s40265-022-01766-4>
- [8] 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]. 中华消化外科杂志, 2021, 20(7): 730-739.
- [9] Pacella, D., De Simone, A., Pisanu, A., Pellino, G., Selvaggi, L., Murzi, V., *et al.* (2025) A Systematic Review of the Predictive Factors for the Recurrence of Acute Pancreatitis. *World Journal of Emergency Surgery*, **20**, Article No. 32. <https://doi.org/10.1186/s13017-025-00601-x>
- [10] Chen, Y., Huang, S., Luo, B., Jiang, J., Ren, W., Zou, K., *et al.* (2024) Prediction and Evaluation of a Nomogram Model for Recurrent Acute Pancreatitis. *European Journal of Gastroenterology & Hepatology*, **36**, 554-562. <https://doi.org/10.1097/meg.0000000000002732>
- [11] Yang, A.L. and McNabb-Baltar, J. (2020) Hypertriglyceridemia and Acute Pancreatitis. *Pancreatology*, **20**, 795-800. <https://doi.org/10.1016/j.pan.2020.06.005>
- [12] Tibshirani, R. (1996) Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **58**, 267-288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- [13] Kursu, M.B. and Rudnicki, W.R. (2010) Feature Selection with the Boruta Package. *Journal of Statistical Software*, **36**, 1-13.
- [14] Mosztbacher, D., Hanák, L., Farkas, N., Szentesi, A., Mikó, A., Bajor, J., *et al.* (2020) Hypertriglyceridemia-Induced Acute Pancreatitis: A Prospective, Multicenter, International Cohort Analysis of 716 Acute Pancreatitis Cases. *Pancreatology*, **20**, 608-616. <https://doi.org/10.1016/j.pan.2020.03.018>

-
- [15] Guan, L., Ding, L., Wan, J., Xia, L., He, W., Xiong, H., *et al.* (2023) Serum Triglyceride Levels Are Associated with Recurrence in Patients with Acute Hypertriglyceridemic Pancreatitis. *Frontiers in Medicine*, **10**, Article ID: 1079637. <https://doi.org/10.3389/fmed.2023.1079637>
- [16] Kota, S., Kota, S., Jammula, S., Krishna, S.V.S. and Modi, K. (2012) Hypertriglyceridemia-Induced Recurrent Acute Pancreatitis: A Case-Based Review. *Indian Journal of Endocrinology and Metabolism*, **16**, 141-143. <https://doi.org/10.4103/2230-8210.91211>
- [17] Song, K., Guo, C., Li, C. and Ding, N. (2021) Risk Factors of Recurrence of Acute Pancreatitis: A Retrospective Research. *Turkish Journal of Gastroenterology*, **32**, 971-978. <https://doi.org/10.5152/tjg.2020.20775>
- [18] Li, L., Wu, H., Yin, H. and Cao, H. (2024) Serum Uric Acid Concentration in Acute Pancreatitis Is Significantly Higher than Healthy Population. *Scientific Reports*, **14**, Article No. 18279. <https://doi.org/10.1038/s41598-024-69425-0>
- [19] Baronia, A.K. and Ahmed, A. (2016) Hypocalcemia in Acute Pancreatitis Revisited. *Indian Journal of Critical Care Medicine*, **20**, 173-177. <https://doi.org/10.4103/0972-5229.178182>
- [20] Sun, Y., Jin, J., Zhu, A., Hu, H., Lu, Y., Zeng, Y., *et al.* (2022) Risk Factors for Recurrent Pancreatitis after First Episode of Acute Pancreatitis. *International Journal of General Medicine*, **15**, 1319-1328. <https://doi.org/10.2147/ijgm.s344863>
- [21] Umans, D.S., Hallensleben, N.D., Verdonk, R.C., Bouwense, S.A.W., Fockens, P., Santvoort, H.C., *et al.* (2020) Recurrence of Idiopathic Acute Pancreatitis after Cholecystectomy: Systematic Review and Meta-Analysis. *Journal of British Surgery*, **107**, 191-199. <https://doi.org/10.1002/bjs.11429>
- [22] Gagyí, E., Teutsch, B., Veres, D.S., Pálkás, D., Vörhendi, N., Ocskay, K., *et al.* (2024) Incidence of Recurrent and Chronic Pancreatitis after Acute Pancreatitis: A Systematic Review and Meta-Analysis. *Therapeutic Advances in Gastroenterology*, **17**, 1-16.