·胰腺疾病· DOI：10.12449/JCH260724

肿瘤相关急性胰腺炎的临床特征及机器学习预测模型构建

杜晨辉[1]，高玉倩[1]，张　硕[2]，何铁英[1]，曹新岭[2,3]

1 新疆医科大学第一附属医院胰腺外科，乌鲁木齐 830054
2 新疆医科大学第一附属医院腹腔镜外科，乌鲁木齐 830054
3 省部共建中亚高发病成因与防治国家重点实验室，乌鲁木齐 830054

通信作者：曹新岭，17590083026@163.com（ORCID：0009-0005-6543-4226）

摘要：目的　分析肿瘤相关急性胰腺炎（TIAP）患者的临床特征，构建并验证TIAP预测模型，以辅助临床早期识别。**方法**　回顾性收集2020年1月—2026年1月新疆医科大学第一附属医院收治的3 051例急性胰腺炎（AP）患者资料，其中TIAP患者72例（TIAP组），为减少类别不平衡，从非TIAP患者中（2 979例）按性别分层随机抽取216例作为常规组，并将TIAP组和常规组共288例患者按7∶3比例随机分为训练集（n=201）和测试集（n=87）。比较两组患者的一般资料及实验室指标差异。不符合正态分布的计量资料两组间比较采用Mann-Whitney U检验；计数资料两组间比较采用χ^2检验。采用递归特征消除和最小绝对收缩与选择算子回归筛选特征变量，分别构建逻辑回归、随机森林、支持向量机、极端梯度提升及轻量梯度提升机模型。通过受试者操作特征曲线、精确率-召回率曲线评估模型性能，采用校准曲线评价模型拟合优度，应用决策曲线分析评价模型的临床实用性，并采用沙普利加性解释对模型进行可解释性分析。**结果**　训练集和测试集中的TIAP患者分别有52例（25.9%）和20例（23.0%）。与常规组相比，TIAP组患者存在胰管扩张比例（χ^2=79.474）更高，年龄（Z=−5.838）更大，白细胞计数（Z=5.630）、淀粉酶（Z=2.606）、血红蛋白（Z=5.038）、中性粒细胞百分比（Z=5.269）水平更低，差异均有统计学意义（P值均<0.05），直接胆红素（Z=0.936）水平有降低趋势，但未表现出统计学差异（P>0.05）。基于上述7个变量构建的5种机器学习模型均具有一定预测效能，其中逻辑回归模型在测试集中表现最佳，曲线下面积为0.893（95%置信区间：0.821～0.953），精确率-召回率曲线的平均精确率为0.654，校准度良好，决策曲线分析显示其具有较好的临床获益。**结论**　基于胰管扩张、年龄、白细胞计数、淀粉酶、血红蛋白、中性粒细胞百分比和直接胆红素构建的TIAP预测模型有较良好的预测性能，可为TIAP的早期诊断提供重要策略指导。

关键词：急性胰腺炎；消化系统肿瘤；机器学习；模型，统计学

基金项目：省部共建中亚高发病成因与防治国家重点实验室基金资助项目（SKL-HIDCA-2023-26）

Clinical features of tumor-induced acute pancreatitis and construction of a machine learning prediction model

Du Chenhui[1], *Gao Yuqian*[1], *Zhang Shuo*[2], *He Tieying*[1], *Cao Xinling*[2,3]

1. Department of Pancreatic Surgery, The First Affiliated Hospital of Xinjiang Medical University, Urumqi 830054, China;
2. Department of Laparoscopic Surgery, The First Affiliated Hospital of Xinjiang Medical University, Urumqi 830054, China;
3. State Key Laboratory of Causes and Prevention of High Morbidity in Central Asia, Urumqi 830054, China

Corresponding author: Cao Xinling, 17590083026@163.com (ORCID: 0009-0005-6543-4226)

Abstract: Objective　To investigate the clinical features of tumor-induced acute pancreatitis (TIAP), to construct and validate a predictive model for TIAP, and to provide help for early identification in clinical practice. **Methods**　A retrospective analysis was performed for the clinical data of 3 051 patients with acute pancreatitis (AP) who were admitted to The First Affiliated Hospital of Xinjiang Medical University from January 2020 to January 2026, among whom there were 72 patients with TIAP. To reduce class imbalance, 216 patients with non-tumor-related AP (2 979 patients) were randomly selected as conventional group using sex-stratified sampling, resulting in a cohort of 288 patients, and this cohort was randomly divided into a training set with 201 patients

and a test set with 87 patients at a ratio of 7∶3. The two groups were compared in terms of general information and laboratory markers. The independent-samples *t* test was used for comparison of normally distributed continuous data between two groups, and the Mann-Whitney *U* test was used for comparison of non-normally distributed continuous data between two groups; the chi-square test was used for comparison of categorical data between two groups. Recursive feature elimination and least absolute shrinkage and selection operator regression were used for screening of characteristic variables, and five machine learning models were constructed, i.e., logistic regression model, random forest model, support vector machine model, extreme gradient boosting model, and light gradient boosting machine model. The receiver operating characteristic curve and the precision-recall curve were used to assess model performance; the calibration curve was used to assess goodness of fit; decision curve analysis was used to evaluate clinical applicability and practicality; Shapley additive explanations were used to assess model interpretability. **Results** In the training set of 201 patients, there were 52 patients (25.9%) in the TIAP group, and in the test set of 87 patients, there were 20 patients (23.0%) in the TIAP group. Compared with the conventional group, the TIAP group had a significantly higher proportion of patients with pancreatic duct dilatation (χ^2=79.474, P<0.05), a significantly higher age (Z=−5.838, P<0.05), and significantly lower levels of white blood cell count (Z=5.630, P<0.05), amylase (Z=2.606, P<0.05), hemoglobin (Z=5.038, P<0.05), and neutrophil percentage (Z=5.269, P<0.05), as well as a lower level of direct bilirubin (Z=0.936, P>0.05). The five machine learning models constructed based on these seven variables had a certain predictive ability, among which the logistic regression model had the best performance in the test set, with an area under the curve of 0.893 (95% confidence interval: 0.821—0.953), an average precision of 0.654 in the precision-recall curve, good calibration, and good clinical benefits based on the decision curve analysis. **Conclusion** The predictive model for TIAP based on pancreatic duct dilatation, age, white blood cell count, amylase, hemoglobin, neutrophil percentage, and direct bilirubin shows good predictive performance and can provide important guidance for the early diagnosis of TIAP.

Key words: Pancreatitis; Digestive System Tumors; Machine Learning; Models, Statistical

Research funding: State Key Laboratory of Pathogenesis, Prevention and Treatment of High Incidence Diseases in Central Asia (Co-construction by the Ministry and Province)(SKL-HIDCA-2023-26)

急性胰腺炎(acute pancreatitis,AP)是由胰腺外分泌功能紊乱引发的常见炎症性疾病,是消化系统中最常见导致器官功能衰竭的急腹症之一,其在全球范围内的发病率为32/10万,死亡率为1.45/10万[1]。AP常见的病因包括胆源性、高脂血症及酒精性因素,其他因素可见药物、感染和自身免疫等[2]。在我国,胆石病仍是AP的主要病因[3]。肿瘤是AP的非常规诱因之一,同时AP也可能是胰腺癌、壶腹部癌、转移性癌和淋巴瘤等多种恶性肿瘤的早期临床表现[4-7]。在AP的临床诊疗过程中,肿瘤相关急性胰腺炎(tumor-induced acute pancreatitis,TIAP)极易被漏诊。多数恶性肿瘤具有较差的生物学行为,延误诊断易导致肿瘤进展,严重影响患者的生存预后。

近年来,随着人工智能技术的迅速发展,机器学习(machine learning,ML)[8-9]技术在医学领域中的应用日益深入,在疾病诊断中逐步展现出显著潜力[10-11]。本研究通过可解释性ML算法构建预测模型,旨在提高对TIAP的识别能力,降低肿瘤漏诊率,为肿瘤早期全面诊治提供参考。

1 资料与方法

1.1 研究对象 收集2021年1月—2026年1月新疆医科大学第一附属医院收治的3 015例AP患者的临床资料。其中,TIAP患者共72例,纳入肿瘤组(TIAP组);采用分层随机抽样方法,以性别作为分层因素,从常规病因引起的AP患者中(2 979例)抽取216例患者作为常规组,以确保两组在性别层面的均衡性,增加研究结果的可靠性和可重复性。随后,将TIAP组和常规组共288例AP患者按7∶3的比例随机划分为训练集(n=201)和测试集(n=87)。TIAP组纳入标准:(1)临床诊断为AP;(2)AP明确由肿瘤引起;(3)肿瘤为胰腺、壶腹部或肝外胆道的原发性良性或恶性肿瘤;(4)临床病历资料完整。排除标准:(1)非胰腺或非邻近器官的肿瘤;(2)AP发病前已明确诊断为肿瘤。常规组排除标准:(1)内窥镜逆行胰胆管造影术后所致AP;(2)除肿瘤外由其他罕见因素引起的AP。

1.2 AP的诊断标准与相关定义 AP的诊断依据《中国急性胰腺炎诊治指南(2021)》中的诊断标准[3],符合以

下3项中的2项即可确诊：(1)上腹部持续性疼痛；(2)血清淀粉酶和/或脂肪酶浓度至少高于正常值上限3倍；(3)腹部影像学检查结果显示符合AP的影像学特征改变。AP病情严重程度[3]依据修订版亚特兰大分级和改良计算机体层成像(computed tomography，CT)严重指数评估，分为轻症AP、中重症AP和重症AP。主胰管扩张定义为CT影像上主胰管最大直径>3 mm[12]，胆管扩张定义为CT影像上胆管最大直径>10 mm[13]。

1.3 研究方法 收集人口学特征，包括年龄、性别、病史、吸烟史、饮酒史等；临床资料包括AP严重程度、实验室监测指标和影像学指标等。为避免模型过拟合及多重共线性影响，本研究采用双重筛选策略，基于训练集分别采用递归特征消除和最小绝对收缩与选择算子算法进行特征筛选，以减少单一方法在特征选择时“噪声”的影响，并增强特征选择的稳健性，最终将二者共同筛选出来的特征纳入预测模型。

1.4 模型构建与验证 基于筛选所得变量，应用5种ML算法在训练集队列中构建预测模型，包括极限梯度提升、支持向量机、轻量级梯度提升机、随机森林和逻辑回归。通过多项指标对模型性能进行评估，包括受试者操作特征曲线下面积、精确率-召回率曲线、准确度、敏感度、特异度、阳性预测值、阴性预测值和F1分数。为评估模型的稳健性和泛化能力，在训练集上对模型进行超参数网格搜索和5折交叉验证，并在测试集上进行外部验证，最终选择最佳模型。绘制校准曲线评价模型的一致性，采用临床决策曲线评价模型的临床实用性。鉴于ML模型固有的“黑箱特性”，本研究采用沙普利加性解释(shapley additive explanations，SHAP)对模型进行分析。

1.5 统计学方法 采用R 4.4.3和Python 3.12软件进行数据统计分析。首先进行数据预处理，分析数据中各变量的类型与缺失情况，剔除缺失率≥20%的变量，随后使用多重插补方法对缺失<20%的变量进行插补，并使用链式方程进行5次迭代[14]。不符合正态分布的计量资料以$M(P_{25}\sim P_{75})$表示，两组间比较采用Mann-Whitney U检验。计数资料两组间比较采用χ^2检验。P<0.05为差异有统计学意义。

2 结果

2.1 训练集与测试集临床资料比较 训练集与测试集基线指标对比，差异均无统计学意义(P值均>0.05)(表1)。训练集和测试集TIAP患者分别有52例(25.9%)和20例(23.0%)，表明数据划分具有良好的随机均衡性。

2.2 TIAP组与常规组临床资料比较 单因素分析结果显示(表2)，与常规组相比，TIAP组患者呈现出显著的“老年化、低炎症反应、梗阻性黄疸”特征：TIAP组年龄显著偏大，白细胞计数淀粉酶及中性粒细胞百分比相对较低，而胰管扩张及胆管扩张比例显著升高(P值均<0.05)。值得注意的是，TIAP组的血红蛋白水平显著低于常规组(P<0.05)，提示可能存在肿瘤消耗或慢性失血。

2.3 特征筛选 使用递归特征消除和最小绝对收缩与选择算子算法共同进行特征筛选，最终纳入7个预测变量(图1)：胰管扩张、白细胞计数、年龄、淀粉酶、中性粒细胞百分比、血红蛋白和直接胆红素。上述特征涵盖影像学形态、炎症水平、机体消耗状态及梗阻指标。

2.4 不同ML算法的构建与性能比较 对5种ML模型的性能评价指标进行比较，结果显示逻辑回归模型的性能最为均衡。在测试集中，受试者操作特征曲线分析显示各模型均表现出较好的分类能力，AUC均高于0.85，其中，逻辑回归模型表现最佳，AUC达到0.893；进一步结合精确率-召回率曲线分析发现，逻辑回归模型平均精确率最高(图2)。以逻辑回归模型在测试集中的预测结果为例，在87个样本中共正确分类72例，整体准确度为83%。其中，常规样本中60例被正确识别，7例被误判为肿瘤，对应特异度为90%；肿瘤样本中12例被正确识别，8例被误判为常规，对应敏感度为60.0%(表3)。

2.5 基于最佳模型的解释与临床识别 SHAP值图直观展示了各特征对预测结果的影响，特征重要性汇总图揭示了特征值对预测的正负影响。其中，胰管扩张的SHAP值最高，是判断TIAP的最强正向预测因子；其次，随着年龄增长，TIAP的发病风险呈非线性上升趋势；而炎症指标呈现负相关趋势，即相较于剧烈的胆源性或高脂性胰腺炎，TIAP患者的全身炎症反应程度较轻；血红蛋白低水平提示高风险，与恶性肿瘤引起的慢性贫血的临床特征相符(图3)。校准曲线分析(图4a)评估了预测概率与真实发生率的一致性，尽管测试集中阳性样本量较少(稀疏数据)导致曲线呈现一定波动，但整体趋势表明模型的预测概率与实际发生率呈正相关。尤其是在中等风险区间，预测值与观测值吻合度尚可。这种波动在罕见病因(TIAP占比仅2.40%)的预测研究中属于常见现象，提示未来需扩大样本量以进一步优化模型的校准度。决策曲线分析(图4b)表明，模型在0.05～0.80的风险阈值区间内具有较高的净获益，优于“全部干预”与“无干预”策略，提示其在临床决策中的辅助价值，能支持高风险患者的早期识别。

表1 训练集与测试集临床资料的比较

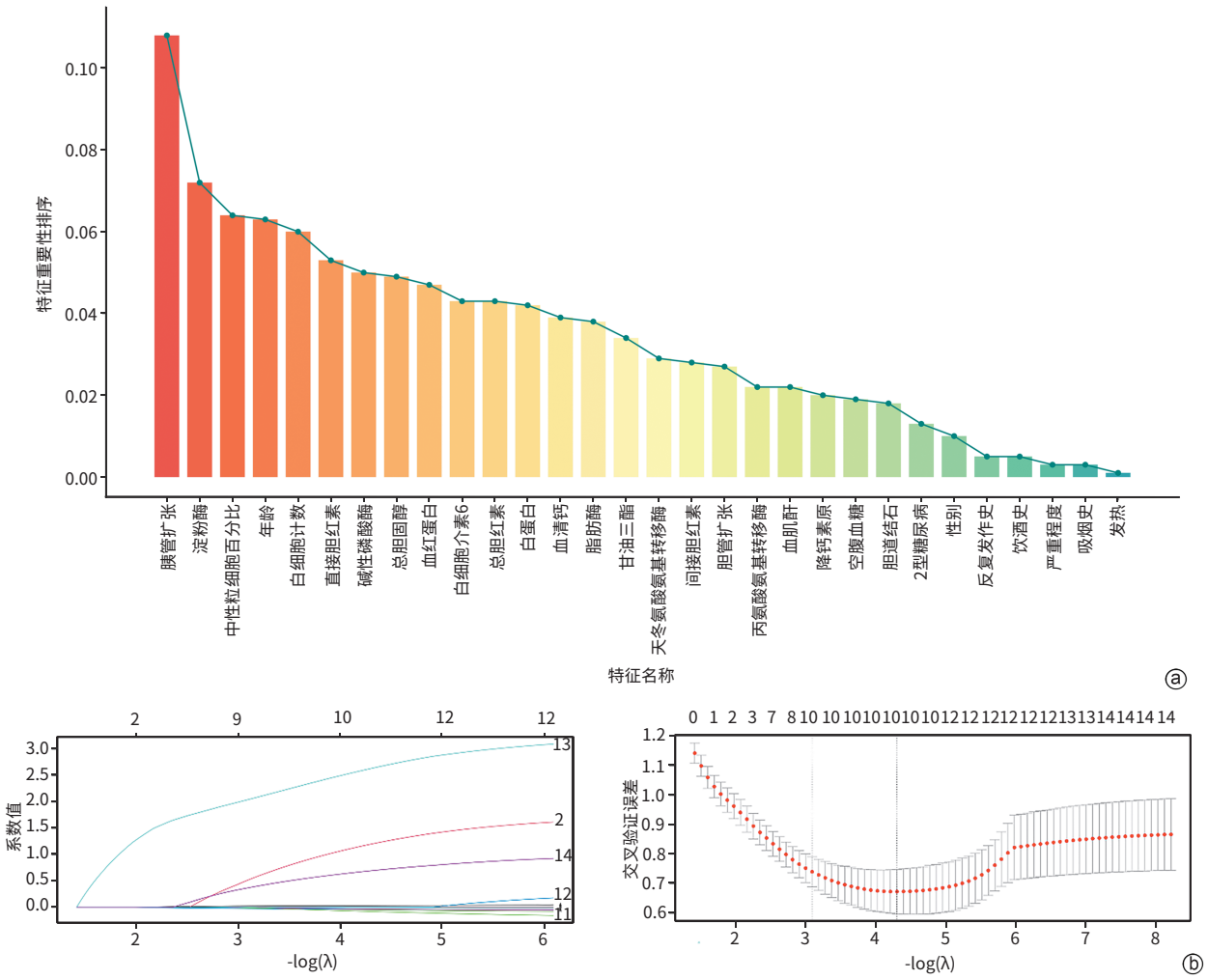
Table 1 Comparison of clinical characteristics between the training and test sets

指标	测试集（n=87）	训练集（n=201）	统计值	P值
肿瘤[例（%）]			χ^2=0.137	0.711
无	67（77.0）	149（74.1）		
有	20（23.0）	52（25.9）		
性别[例（%）]			χ^2<0.001	0.999
女	37（42.5）	87（43.3）		
男	50（57.5）	114（56.7）		
吸烟史[例（%）]			χ^2=0.054	0.817
无	60（69.0）	143（71.1）		
有	27（31.0）	58（28.9）		
饮酒史[例（%）]			χ^2=1.573	0.210
无	57（65.5）	148（73.6）		
有	30（34.5）	53（26.4）		
2型糖尿病[例（%）]			χ^2=0.379	0.538
无	68（78.2）	165（82.1）		
有	19（21.8）	36（17.9）		
胰腺炎反复病史[例（%）]			χ^2=0.103	0.749
无	72（82.8）	171（85.1）		
有	15（17.2）	30（14.9）		
发热[例（%）]			χ^2=0.474	0.491
无	75（86.2）	165（82.1）		
有	12（13.8）	36（17.9）		
严重程度[例（%）]			χ^2=0.474	0.491
轻中度	75（86.2）	165（82.1）		
重度	12（13.8）	36（17.9）		
胆道结石[例（%）]			χ^2<0.001	0.999
无	57（65.5）	131（65.2）		
有	30（34.5）	70（34.8）		
胰管扩张[例（%）]			χ^2=0.264	0.608
无	71（81.6）	157（78.1）		
有	16（18.4）	44（21.9）		
胆管扩张[例（%）]			χ^2=0.379	0.538
无	68（78.2）	165（82.1）		
有	19（21.8）	36（17.9）		
年龄（岁）	50.00（34.50～61.00）	50.00（38.00～61.00）	Z=−0.459	0.647
白细胞计数（×10^9/L）	9.86（6.96～13.57）	9.82（6.47～13.15）	Z=0.393	0.695
中性粒细胞百分比（%）	77.40（67.60～84.45）	78.60（68.40～84.80）	Z=−0.253	0.801
血红蛋白（g/L）	140.00（121.50～155.00）	139.00（125.00～156.00）	Z=−0.466	0.613
淀粉酶（U/L）	182.43（86.16～537.32）	167.26（79.25～499.43）	Z=0.411	0.682
脂肪酶（U/L）	297.76（109.06～1 009.25）	303.86（66.62～1 514.50）	Z=0.196	0.845
血清钙（mmol/L）	2.23（2.10～2.30）	2.21（2.10～2.30）	Z=0.616	0.538
空腹血糖（mmol/L）	6.64（5.33～9.20）	6.70（5.30～9.36）	Z=−0.274	0.784
血肌酐（μmol/L）	58.90（50.34～72.28）	58.26（49.80～73.08）	Z=0.045	0.964
白蛋白（g/L）	41.81（37.73～45.72）	40.37（37.14～44.14）	Z=1.724	0.085
甘油三酯（mmol/L）	1.84（0.99～3.75）	1.65（1.05～5.45）	Z=−0.032	0.975
总胆固醇（mmol/L）	4.56（3.42～5.84）	4.54（3.71～6.04）	Z=−0.677	0.499
总胆红素（μmol/L）	21.88（15.62～39.20）	22.09（15.28～35.78）	Z=0.079	0.937
直接胆红素（μmol/L）	8.29（5.58～13.14）	9.55（6.33～17.79）	Z=−1.051	0.294
间接胆红素（μmol/L）	12.45（8.23～21.84）	12.66（8.29～20.14）	Z=0.136	0.893
天冬氨酸氨基转移酶（U/L）	35.33（24.82～73.37）	37.40（27.01～80.92）	Z=−0.837	0.403
丙氨酸氨基转移酶（U/L）	31.40（21.00～58.23）	33.04（20.00～95.00）	Z=−0.725	0.469
碱性磷酸酶（U/L）	93.44（71.55～136.61）	98.92（74.03～151.13）	Z=−0.510	0.610
降钙素原（ng/L）	0.09（0.05～0.20）	0.11（0.04～0.28）	Z=−0.559	0.576
白细胞介素6（pg/L）	24.90（7.15～65.03）	22.70（7.67～69.64）	Z=−0.301	0.764

表2 肿瘤相关急性胰腺炎组与常规组临床资料的比较

Table 2 Comparison of clinical characteristics between the tumor-induced acute pancreatitis group and the conventional group

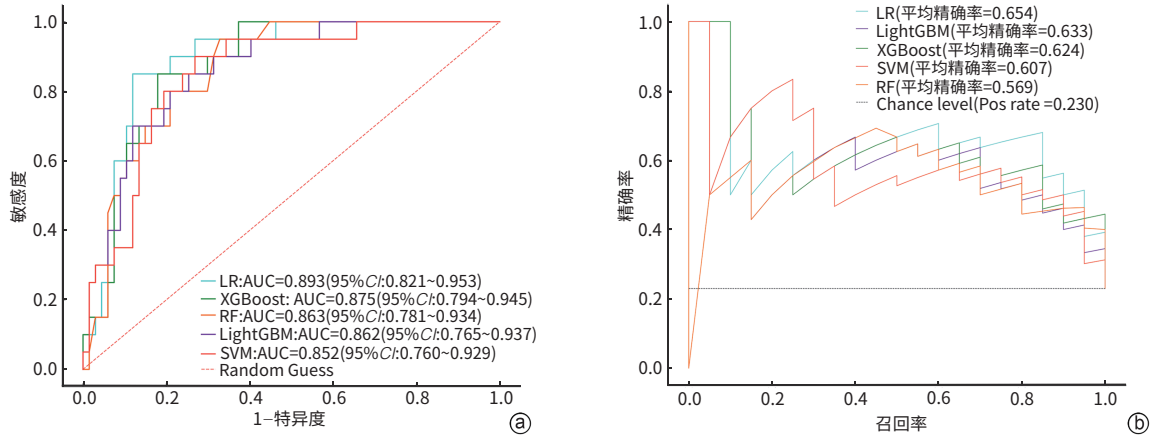
指标	常规组（n=216）	TIAP组（n=72）	统计值	P值
性别[例(%)]			χ^2<0.001	0.999
女	93（43.1）	31（43.1）		
男	123（56.9）	41（56.9）		
吸烟史[例(%)]			χ^2=0.139	0.709
无	154（71.3）	49（68.1）		
有	62（28.7）	23（31.9）		
饮酒史[例(%)]			χ^2=0.683	0.409
无	157（72.7）	48（66.7）		
有	59（27.3）	24（33.3）		
2型糖尿病[例(%)]			χ^2=11.394	0.001
无	185（85.6）	48（66.7）		
有	31（14.4）	24（33.3）		
胰腺炎反复病史[例(%)]			χ^2=12.019	0.001
无	192（88.9）	51（70.8）		
有	24（11.1）	21（29.2）		
发热[例(%)]			χ^2=0.300	0.584
无	178（82.4）	62（86.1）		
有	38（17.6）	10（13.9）		
严重程度[例(%)]			χ^2=2.700	0.100
轻中度	175（81.0）	65（90.3）		
重度	41（19.0）	7（9.7）		
胆道结石[例(%)]			χ^2=3.452	0.063
无	134（62.0）	54（75.0）		
有	82（38.0）	18（25.0）		
胰管扩张[例(%)]			χ^2=79.474	<0.001
无	201（93.1）	32（44.4）		
有	15（6.9）	40（55.6）		
胆管扩张[例(%)]			χ^2=26.077	<0.001
无	190（88.0）	43（59.7）		
有	26（12.0）	29（40.3）		
年龄（岁）	45.00（35.00～56.25）	59.00（51.75～68.00）	Z=−5.838	<0.001
白细胞计数（×10^9/L）	11.09（7.67～13.85）	6.88（5.25～8.94）	Z=5.630	<0.001
中性粒细胞百分比（%）	80.75（72.75～85.43）	69.90（60.83～77.88）	Z=5.269	<0.001
血红蛋白（g/L）	145.00（128.00～161.00）	132.00（115.00～139.50）	Z=5.038	<0.001
淀粉酶（U/L）	182.91（84.47～661.49）	156.09（67.97～269.71）	Z=2.606	0.009
脂肪酶（U/L）	330.16（78.89～1 567.87）	192.43（51.91～891.95）	Z=2.047	0.041
血清钙（mmol/L）	2.21（2.10～2.30）	2.22（2.11～2.30）	Z=−0.011	0.992
空腹血糖（mmol/L）	6.83（5.32～9.94）	6.33（5.20～8.15）	Z=1.299	0.194
血肌酐（μmol/L）	59.61（50.65～73.08）	56.74（48.18～71.06）	Z=0.938	0.349
白蛋白（g/L）	41.87（38.16～45.46）	38.60（33.30～41.07）	Z=4.981	<0.001
甘油三酯（mmol/L）	1.85（1.06～6.96）	1.39（0.98～2.12）	Z=2.975	0.003
总胆固醇（mmol/L）	4.87（3.83～6.30）	3.88（3.26～4.76）	Z=4.681	<0.001
总胆红素（μmol/L）	22.64（16.03～35.51）	18.63（13.68～47.90）	Z=0.662	0.509
直接胆红素（μmol/L）	9.83（6.40～14.89）	7.40（4.91～33.77）	Z=0.936	0.350
间接胆红素（μmol/L）	12.74（8.33～20.06）	12.05（8.04～21.77）	Z=0.137	0.891
天冬氨酸氨基转移酶（U/L）	37.71（26.68～76.54）	33.88（25.30～83.22）	Z=0.600	0.549
丙氨酸氨基转移酶（U/L）	32.64（20.00～95.00）	31.68（20.91～59.34）	Z=0.827	0.409
碱性磷酸酶（U/L）	96.73（73.50～135.01）	103.83（73.46～251.30）	Z=−1.668	0.095
降钙素原（ng/L）	0.11（0.04～0.26）	0.09（0.04～0.21）	Z=0.832	0.405
白细胞介素6（pg/L）	30.38（8.96～80.60）	10.85（4.82～28.16）	Z=3.604	<0.001



注：a，递归特征消除，特征按照重要性进行前后排序，从第9个特征开始，后续特征的重要性逐渐减小；b，最小绝对收缩与选择算子算法的系数路径图与交叉验证图。

图1　特征筛选结果

Figure 1　Feature selection results



注：a，ROC曲线；b，精确率-召回率曲线。ROC曲线，受试者操作特征曲线；AUC，曲线下面积；RF，随机森林；LR，逻辑回归；XGBoost，极限梯度提升；LightGBM，轻量级梯度提升机；SVM，支持向量机；*CI*，置信区间。

图2　5种机器学习模型在测试集上的ROC曲线和精确率-召回率曲线

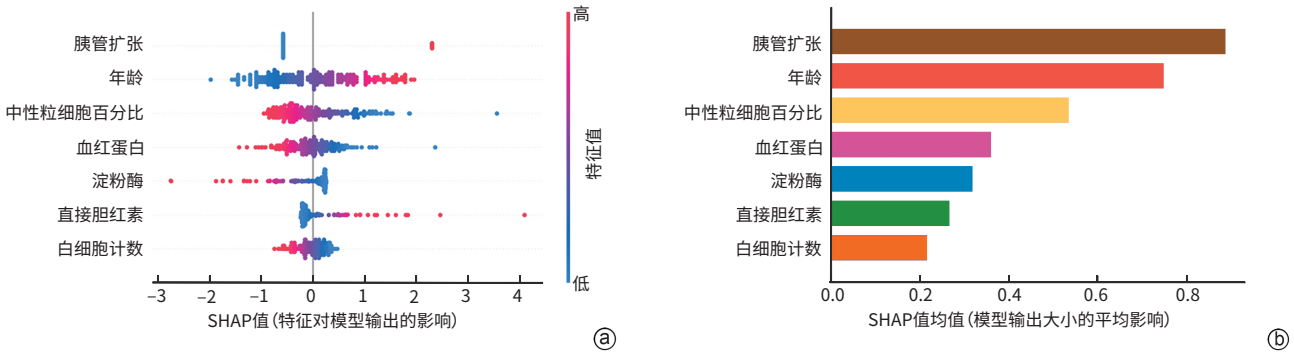
Figure 2　ROC curves and precision-recall curves for five machine learning models on the test set

表3　5种机器学习算法性能的比较

Table 3　Performance comparison of five machine learning algorithms

项目	准确度	敏感度	特异度	阳性预测值	阴性预测值	F1分数	Kappa分数
训练集							
RF	1.00	1.00	1.00	1.00	1.00	1.00	1.00
LR	0.86	0.62	0.95	0.80	0.88	0.70	0.61
XGBoost	1.00	0.98	1.00	1.00	0.99	0.99	0.99
LightGBM	1.00	1.00	1.00	1.00	1.00	1.00	1.00
SVM	0.80	0.27	0.98	0.82	0.79	0.41	0.32
测试集							
RF	0.80	0.60	0.87	0.57	0.88	0.59	0.46
LR	0.83	0.60	0.90	0.63	0.88	0.62	0.50
XGBoost	0.82	0.65	0.87	0.59	0.89	0.62	0.50
LightGBM	0.83	0.65	0.88	0.62	0.89	0.63	0.52
SVM	0.78	0.30	0.93	0.55	0.82	0.39	0.27

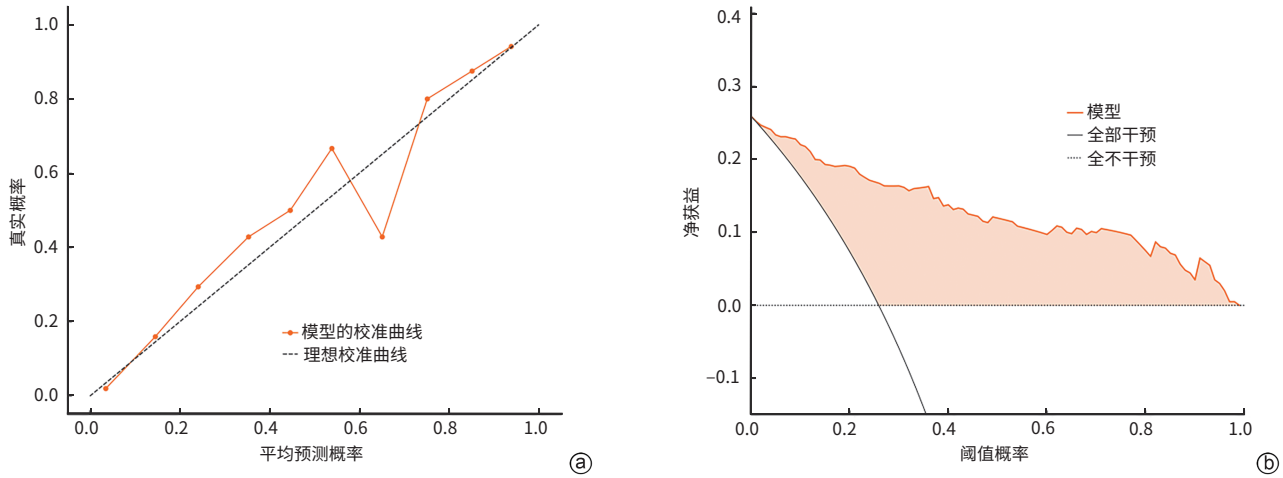
注：RF，随机森林；LR，逻辑回归；XGBoost，极限梯度提升；LightGBM，轻量级梯度提升机；SVM，支持向量机。



注：a，SHAP特征重要性汇总图，展示每个样本的SHAP值和特征值分布；b，平均SHAP值柱状图，展示整体特征重要性排序。SHAP，沙普利加性解释。

图3　SHAP值的模型解释图

Figure 3　Model interpretation diagram of shapley additive explanations values



注：a，校准曲线；b，决策曲线。

图4　校准曲线与决策曲线分析

Figure 4　Calibration curve and decision curve analysis

3 讨论

本研究结果显示，胰管扩张、白细胞计数、年龄、淀粉酶、直接胆红素、血红蛋白和中性粒细胞百分比与TIAP的发生密切相关。临床上，尽管多数肿瘤并不会诱发AP，但5%～14%的良性或恶性胰胆管肿瘤患者以AP为首发表现或并发AP[15]。目前，关于TIAP的具体发病机制尚未完全阐明，既往文献研究主要归纳为机械性梗阻、血管侵犯、酶原激活和肿瘤所致高钙血症等途径[16]。不同于常规病因胰腺炎“急性胰酶异常激活-腺泡细胞损伤-炎症瀑布反应”的经典路径，TIAP更常体现为“肿瘤导致的胰管或壶腹部梗阻-胰液引流障碍-导管压力改变与局灶性炎症”，并叠加肿瘤微环境、免疫异常及胆道梗阻等因素，使其炎症表现和实验室谱系常呈现“非典型化”特征，这是其主要发病机制[17-19]。结合本研究结果，TIAP患者血清淀粉酶水平显著低于常规病因AP患者，白细胞计数及中性粒细胞百分比亦相对较低，该结果从机制层面支持TIAP并非以广泛腺泡细胞坏死和胰酶大量外溢为主要病理基础。在我国常规病因胰腺炎中，胆源性胰腺炎往往占较大比例[3]。胆道结石在急性期可造成壶腹部短暂性梗阻，从而引起胆汁反流和胆道压力升高，表现为一过性或中度胆红素升高。这一现象提示，胆红素升高并非TIAP的必备特征，其诊断价值需结合病因构成与疾病阶段综合评估。同时也揭示了TIAP以胰管局部狭窄或慢性梗阻为主，而非急性胆胰共同梗阻的发病模式。

在病因谱与影像特征方面，大样本队列研究显示，胰腺癌、壶腹周围癌及肿瘤性胰腺囊肿是TIAP最主要的病因，占比超过80%；该研究进一步指出，胰管与胆管的同时扩张（即“双管征”）可显著提高TIAP的诊断特异度，且胰管扩张在胰腺癌患者中更为常见[20]。本研究中，TIAP组多数患者肿瘤位于胰腺区域，约55.6%的患者伴有显著的胰管扩张。这一结果与上述文献报道高度吻合，有力佐证了本研究模型将“胰管扩张”筛选为关键预测因子的病理学基础。

临床实践中，提示AP患者存在潜在肿瘤的因素在很大程度上尚未被充分认识。CT在诊断和评估AP严重程度时被广泛应用，但在鉴别病因诊断方面常不具有较高的诊断价值[21]。弥漫性胰腺炎症渗出易掩盖局部肿块，因此Cho等[22]建议，对于病因不明的AP患者，应在出院后3个月内常规行CT检查，以发现潜在肿瘤的可能。本研究中，TIAP组患者较常规胰腺炎患者年龄偏大，且部分患者3个月内有AP反复发作病史。肿瘤患者更易合并慢性炎症性贫血、营养不良、隐匿出血或骨髓抑制等，使其血红蛋白水平与常规病因AP群体出现显著差异。低血红蛋白与高龄作为重要特征，反映了恶性肿瘤的慢性消耗状态及流行病学规律，当老年患者出现不明原因的AP伴贫血时，应高度怀疑合并肿瘤的可能性。这一发现提示临床医生，对于症状及炎症指标“看似轻微”但影像学存在异常的老年AP患者，应保持高度警惕，避免因症状不典型而导致漏诊[23]。

本研究应用ML算法为TIAP的早期识别及风险预测提供了初步证据，具有一定的临床指导意义。然而，本研究仍存在以下局限性：（1）为单中心回顾性研究，无法避免回顾性研究带来的选择偏倚；（2）仅证实了肿瘤与常规病因引起AP之间的差异，尚未深入分析不同肿瘤部位、病理类型与患者预后的相关性。因此，后续仍需开展多中心、大样本的前瞻性研究，以进一步验证本研究的最佳模型（逻辑回归模型）的稳定性与泛化能力，提升其在临床实践中的应用价值。

伦理学声明：本研究方案于2025年12月4日经由新疆医科大学第一附属医院审批通过，批号：20250422-10。

利益冲突声明：本文不存在任何利益冲突。

作者贡献声明：杜晨辉负责数据分析及文章撰写；高玉倩和张硕负责数据获取与分析；何铁英负责审阅文章；曹新岭负责修改和审阅文章。

参考文献：

[1] Li T Y, Qin C, Zhao B B, et al. Global and regional burden of pancreatitis: Epidemiological trends, risk factors, and projections to 2050 from the global burden of disease study 2021[J]. BMC Gastroenterol, 2024, 24(1): 398. DOI: 10.1186/s12876-024-03481-8.

[2] 曲元旭, 曹锋, 丁乙轩, 等. 不同病因急性胰腺炎的临床特征及预后情况[J]. 中华消化外科杂志, 2025, 24(5): 623-629. DOI: 10.3760/cma.j.cn115610-20250305-00086.
Qu Yuanxu, Cao Feng, Ding Yixuan, et al. Clinical characteristics and prognosis of acute pancreatitis with different etiologies[J]. Chin J Dig Surg, 2025, 24(5): 623-629. DOI: 10.3760/cma.j.cn115610-20250305-00086.

[3] 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.
Chinese Pancreatic Surgery Association, Chinese Society of Surgery, Chinese Medical Association. Guidelines for diagnosis and treatment of acute pancreatitis in China(2021)[J]. Chin J Surg, 2021, 59(7): 578-587. DOI: 10.3760/cma.j.cn112139-20210416-00172.

[4] Feng Q, Li C, Zhang S, et al. Recurrence and survival after surgery for pancreatic cancer with or without acute pancreatitis[J]. World J Gastroenterol, 2019, 25(39): 6006-6015. DOI: 10.3748/wjg.v25.i39.6006.

[5] Pitlick M M, Abeykoon J P, Dao L, et al. Non-Hodgkin lymphoma

presenting as acute pancreatitis: A rare occurrence[J]. Clin Case Rep, 2018, 7(1): 107-109. DOI: 10.1002/ccr3.1885.
[6] Tanaka H, Nakazawa T, Yoshida M, et al. Metastasis-induced acute pancreatitis in a patient with small cell carcinoma of the lungs[J]. JOP, 2009, 10(5): 557-561.
[7] Dixon J M, Chapman R W, Berry A R. Carcinoid tumour of the *Ampulla* of Vater presenting as acute pancreatitis[J]. Gut, 1987, 28(10): 1296-1297. DOI: 10.1136/gut.28.10.1296.
[8] Cuocolo R, Caruso M, Perillo T, et al. Machine learning in oncology: A clinical appraisal[J]. Cancer Lett, 2020, 481: 55-62. DOI: 10.1016/j.canlet.2020.03.032.
[9] Chung H, Ko Y, Lee I S, et al. Prognostic artificial intelligence model to predict 5 year survival at 1 year after gastric cancer surgery based on nutrition and body morphometry[J]. J Cachexia Sarcopenia Muscle, 2023, 14(2): 847-859. DOI: 10.1002/jcsm.13176.
[10] 殷民月, 朱锦舟, 刘璐, 等. 急性胰腺炎机器学习模型的研究进展[J]. 临床肝胆病杂志, 2023, 39(12): 2978-2984. DOI: 10.3969/j.issn.1001-5256.2023.12.034.
Yin Minyue, Zhu Jinzhou, Liu Lu, et al. Research advances in machine learning models for acute pancreatitis[J]. J Clin Hepatol, 2023, 39(12): 2978-2984. DOI: 10.3969/j.issn.1001-5256.2023.12.034.
[11] 梁谭, 陈子祥, 陈江明, 等. 人工智能在胰腺癌全程化管理中的应用及展望[J]. 中华消化外科杂志, 2024, 23(11):1459-1464. DOI: 10.3760/cma.j.cn115610-20241028-00471.
Liang Tan, Chen Zixiang, Chen Jiangming, et al. Application and prospect of artificial intelligence in the whole-process management of pancreatic cancer[J]. Chin J Dig Surg, 2024, 23(11): 1459-1464. DOI: 10.3760/cma.j.cn115610-20241028-00471.
[12] Ko S E, Choi I Y, Cha S H, et al. Clinical and radiologic characteristics of pancreatic head carcinoma without main pancreatic duct dilatation: Using dual-phase contrast-enhanced CT scan[J]. Clin Imag, 2016, 40(3): 548-552. DOI: 10.1016/j.clinimag.2015.10.008.
[13] Rizvi A, Sethi A, Poneros J, et al. Does incidentally detected common bile duct dilation need evaluation?[J]. Cleve Clin J Med, 2022, 89(6): 315-319. DOI: 10.3949/ccjm.89a.21029.
[14] Ren L J, Wang T, Sekhari Seklouli A, et al. A review on missing values for main challenges and methods[J]. Inf Syst, 2023, 119: 102268. DOI: 10.1016/j.is.2023.102268.
[15] Tenner S, Vege S S, Sheth S G, 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.
[16] Mujica V R, Barkin J S, Go V L. Acute pancreatitis secondary to pancreatic carcinoma[J]. Pancreas, 2000, 21(4): 329-332. DOI: 10.1097/00006676-200011000-00001.
[17] Li J, Zhou R, Zhang J, et al. Calcium signaling of pancreatic acinar cells in the pathogenesis of pancreatitis[J]. World J Gastroenterol, 2014, 20(43): 16146-16152. DOI: 10.3748/wjg.v20.i43.16146.
[18] Canakis A, Baron T H. The role of underlying pancreatic ductal obstruction and risk of post-ERCP pancreatitis[J]. Gastrointest Endosc, 2025, 101(3): 694. DOI: 10.1016/j.gie.2024.09.033.
[19] Shi J Q, Nie S, Shi Z, et al. Clinical characteristics of pancreatic cancer presenting with acute pancreatitis: A propensity-matched cohort study[J]. Surg Endosc, 2025, 39(5): 3307-3316. DOI: 10.1007/s00464-025-11654-z.
[20] Xiong Y Y, Zhao Y, Han X L, et al. Clinical characteristics and outcome of tumor-associated acute pancreatitis: A single-center cohort study[J]. Ann Transl Med, 2021, 9(8): 639. DOI: 10.21037/atm-20-7196.
[21] 任夏, 叶叶, 刘罗杰, 等. 基于炎症因子、肺超声和CT评分系统的急性胰腺炎预后不良列线图预测模型的构建[J]. 临床肝胆病杂志, 2025, 41(4): 713-721. DOI: 10.12449/JCH250417.
Ren Xia, Ye Ye, Liu Luojie, et al. Establishment of a nomogram prediction model for poor prognosis of acute pancreatitis based on inflammatory factors, lung ultrasound, and CT scores[J]. J Clin Hepatol, 2025, 41(4): 713-721. DOI: 10.12449/JCH250417.
[22] Cho J H, Choi J S, Hwang E T, et al. Usefulness of scheduled follow-up CT in discharged patients with acute pancreatitis[J]. Pancreatology, 2015, 15(6): 642-646. DOI: 10.1016/j.pan.2015.09.004.
[23] 张慧敏, 金梦, 钱家鸣, 等. 胰腺癌合并急性胰腺炎和单纯急性胰腺炎诊断分析[J]. 中国实用内科杂志, 2017, 37(2): 145-147. DOI: 10.19538/j.nk2017020114.
Zhang Huimin, Jin Meng, Qian Jiaming, et al. Differential diagnosis between pancreatic cancer with acute pancreatitis and simple acute pancreatitis cancer[J]. Chin J Pract Intern Med, 2017, 37(2): 145-147. DOI: 10.19538/j.nk2017020114.

收稿日期：2025-12-22；**录用日期**：2026-02-24
本文编辑：王莹

引证本文：杜晨辉, 高玉倩, 张硕, 等. 肿瘤相关急性胰腺炎的临床特征及机器学习预测模型构建[J]. 临床肝胆病杂志, 2026, 42(7): 1661-1669.
Du Chenhui, Gao Yuqian, Zhang Shuo, et al. Clinical features of tumor-induced acute pancreatitis and construction of a machine learning prediction model[J]. J Clin Hepatol, 2026, 42(7): 1661-1669.