BMC Medical Informatics and Decision Making

RESEARCH

Open Access

Automated machine learning for early prediction of systemic inflammatory response syndrome in acute pancreatitis

Rufa Zhang[1], Shiqi Zhu[2], Li Shi[3], Hao Zhang[4], Xiaodan Xu[3], Bo Xiang[1*] and Min Wang[1*]

Abstract

Background Systemic inflammatory response syndrome (SIRS) is a frequent and serious complication of acute pancreatitis (AP), often associated with increased mortality. This study aims to leverage automated machine learning (AutoML) algorithms to create a model for the early and precise prediction of SIRS in AP.

Methods This study retrospectively analyzed patients diagnosed with AP across multiple centers from January 2017 to December 2021. Data from the First Affiliated Hospital of Soochow University and Changshu Hospital were used for training and internal validation, while testing was conducted with data from the Second Affiliated Hospital. Predictive models were constructed and validated using the least absolute shrinkage and selection operator (LASSO) and AutoML. A nomogram was developed based on multivariable logistic regression (LR) analysis, and the performance of the models was assessed through receiver operating characteristic (ROC) curves, calibration curves, and decision curve analysis (DCA). Additionally, the AutoML model's effectiveness and interpretability were assessed through DCA, feature importance, SHapley Additive exPlanation (SHAP) plots, and locally interpretable model-agnostic explanations (LIME).

Results A total of 1,224 patients were included, with 812 in the training cohort, 200 in validation, and 212 in testing. SIRS occurred in 33.7% of the training cohort, 34.0% in validation, and 22.2% in testing. AutoML models outperformed traditional LR, with the deep learning (DL) model achieving an area under the ROC curve of 0.843 in the training set, and 0.848 and 0.867 in validation and testing, respectively.

Conclusion The AutoML model using the DL algorithm is clinically significant for the early prediction of SIRS in AP.

Keywords Automated machine learning, Systemic inflammatory response syndrome, Acute pancreatitis, Predictive models, Artificial intelligence

*Correspondence:
Bo Xiang
xiangbby@foxmail.com
Min Wang
20215255001@stu.suda.edu.com
[1]Department of Gastroenterology, The People's Hospital of Nanchuan, No. 16, Nanda Street, Nanchuan District, Chongqing 408400, China
[2]Department of Gastroenterology, The First Affiliated Hospital of Soochow University, Suzhou, Jiangsu, China
[3]Department of Gastroenterology, Changshu Hospital Affiliated to Soochow University, Changshu No. 1 People's Hospital, Suzhou, Jiangsu, China
[4]Department of Gastroenterology, The Second Affiliated Hospital of Soochow University, Suzhou, Jiangsu, China

Introduction

Acute pancreatitis (AP) is a frequent cause of severe abdominal pain, occurring due to the activation of pancreatic enzymes that cause self-digestion of the pancreas and initiate a chain of inflammatory responses, potentially leading to various complications [1]. While the symptoms of acute pancreatitis (AP) typically resolve on their own in most patients, approximately 20% may develop severe complications, including prolonged organ failure and the onset of systemic inflammatory response syndrome (SIRS) [2]. If SIRS persists, it heightens the risk of persistent organ dysfunction, potentially progressing AP to severe acute pancreatitis (SAP) [3]. SIRS is a key factor in the progression of SAP and is closely linked to its severity, which can lead to multiple organ failure and even death [1, 4, 5]. Therefore, early detection and treatment of SIRS are essential.

In the early stages of AP, the condition starts as a localized sterile inflammation of the pancreas and quickly progresses to SIRS, often accompanied by compensatory anti-inflammatory response syndrome (CARS) [6]. The mechanism of SIRS is complex and is associated with the anti-inflammatory response mediated by Th2/Treg differentiation [7–9]. Studies have shown that gut microbiota translocation, the release of extracellular ATP, and activation of endogenous coagulation pathways are all closely related to inflammatory factors [7, 10, 11]. Additionally, studies have also shown that elevated serum levels of Pentraxin 3 (PTX3) are closely associated with SIRS and ultimately result in fatal outcomes in critically ill patients [12]. Staubli SM et al. indicates that PTX3 performs worse than C-reactive protein (CRP) and Acute Physiology and Chronic Health Evaluation II (APACHE II) in predicting SIRS. All three biomarkers exhibited weak predictive discrimination, with area under the receiver operating characteristic curve (AUROC) values of 0.54, 0.69, and 0.69, respectively. When CRP and PTX3 are used in combination, the AUC value increases to 0.7 [13].

Most of these studies relied on traditional logistic regression (LR) models, with some not fully addressing the problem of multicollinearity among variables. Recently, the growing use of machine learning (ML) in healthcare has attracted attention, with both supervised and unsupervised learning offering powerful algorithms for handling large amounts of clinical data. Unlike traditional LR models, ML provides clear advantages in predicting patient outcomes, complications, and prognosis [14, 15]. Traditional ML algorithms typically include Support Vector Machines (SVM), Artificial Neural Networks (ANN), Bayesian learning, and Random Forests. A newer approach, AutoML, automates the selection of algorithms and hyperparameters to build customized models tailored to specific data. In contrast to traditional ML, AutoML employs techniques like intelligent early stopping, cross-validation, regularization, and hyperparameter optimization, allowing for the development of more accurate models in less time.

In this study, we will utilize the H2O AutoML platform to develop and validate a range of ML models for the early and accurate prediction of SIRS in patients with AP. The performance of these models will be compared with that of the traditional LR method.

Materials & methods

Patients

This study retrospectively gathered clinical data from hospitalized AP patients at Changshu Hospital, the First Affiliated Hospital of Soochow University, and the Second Affiliated Hospital of Soochow University, covering the period from January 2017 to December 2021. All three hospitals are tertiary teaching institutions located in Suzhou, Jiangsu, China. The data from Changshu Hospital and the First Affiliated Hospital were split into a training set and an internal validation set in an 8:2 ratio, while the data from the Second Affiliated Hospital served as an independent test set.

The diagnostic criteria for AP were based on the revised 2012 Atlanta classification [16]. To diagnose acute pancreatitis (AP), a patient must satisfy at least two of the following three criteria: (1) typical abdominal pain; (2) serum amylase levels exceeding three times the normal upper limit; and (3) imaging findings consistent with AP [16]. Adults aged 18 years and older who meet the above criteria will be included in this study. According to the criteria for SIRS, a diagnosis of AP can be made in patients who exhibit two or more of the following conditions: (1) a body temperature greater than 38 °C or less than 36 °C; (2) a heart rate exceeding 90 beats per minute; (3) a respiratory rate greater than 20 breaths per minute or arterial carbon dioxide tension (PaCO2) less than 32 mm Hg; and (4) a white blood cell count greater than 12,000 cells/mm^3, less than 4,000 cells/mm^3, or more than 10% immature (band) forms [17]. During hospitalization, patients who meet the diagnostic criteria for SIRS will be classified into the SIRS group, while those who do not meet the criteria will be classified into the non-SIRS group. Patients with chronic liver disease, chronic kidney disease, hematological disorders, recurrent, chronic, traumatic, or idiopathic pancreatitis, pancreatic cancer, a history of pancreatic surgery, those who have undergone chemoradiotherapy, and pregnant patients will be excluded from the study. All included patients will receive treatment following the established guidelines for managing AP. The ethics committee of Changshu Hospital Affiliated to Soochow University has approved this study.

Data collection

Demographic information, clinical data, and details of comorbidities were extracted from electronic medical records. Relevant laboratory tests, including complete blood counts, coagulation tests, and serum biochemical markers, were collected within 24 h of admission. A total of 34 variables were analyzed, with more details provided in Supplementary Table S1. Missing data were assumed to be random and were imputed using a random forest algorithm through the "mice" package in R software [18]. The study flowchart is shown in Fig. 1.

Statistical analysis

In this study, we used the Shapiro-Wilk test to assess whether continuous variables followed a normal distribution. Continuous variables were expressed as mean ± standard deviation (SD) if they followed a normal distribution, and as median (interquartile range) if not. Categorical variables were presented as frequencies. For group comparisons, categorical variables were analyzed using Pearson's Chi-square test or Fisher's exact test, while continuous variables were compared using the non-parametric Mann-Whitney U test. A two-sided p-value

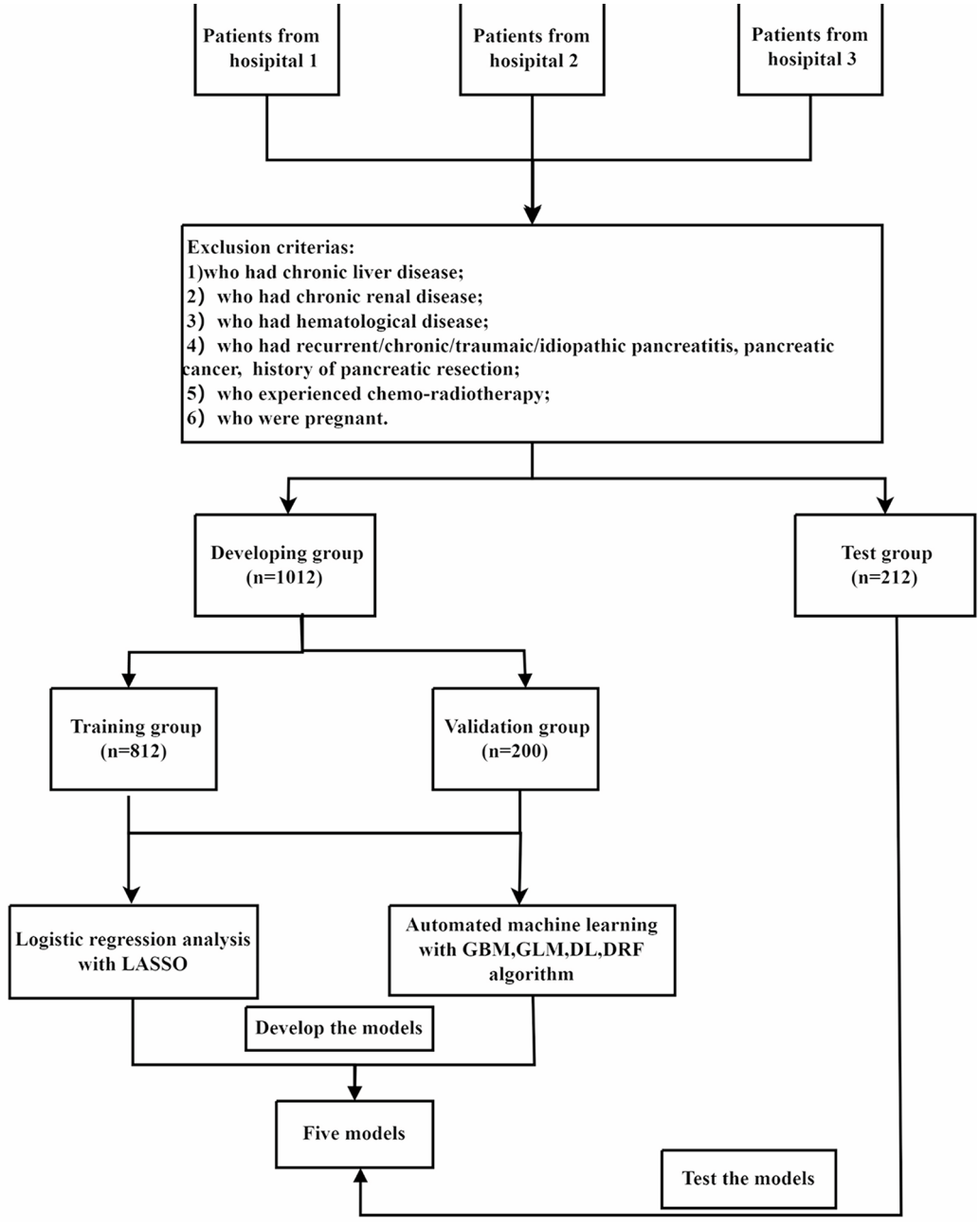

Fig. 1 The flowchart of this study

of less than 0.05 was considered statistically significant. All statistical analyses were performed using R software (version 4.2.1) with the following packages: H2O (version 3.36.0.2), tableone (version 0.12.0), tidyverse (version 1.3.0), tidyquant (version 1.0.2), and lime (version 0.5.1).

Logistic regression algorithms and automated machine learning algorithms

To address multicollinearity among variables, the LASSO regression model was applied using the "λmin + 1se" criterion for univariate analysis. Independent risk factors were identified through binary LR with a backward stepwise method, and a nomogram was constructed. The model's predictive ability was evaluated using ROC curves, calibration curves, and decision curve analysis (DCA).

For the ML component, this study utilized the H2O package (www.h2o.ai) for AutoML, which automatically selects suitable algorithms and combines them into various ensemble models. These included the default Random Forest (DRF), a random grid of Gradient Boosting Machines (GBM), Extremely Randomized Forest (XRF), a random grid of Deep Learning (DLs), and a fixed grid of Generalized Linear Models (GLMs). Hyperparameters were optimized through a grid search with 5-fold cross-validation on the training set, and the area under the curve (AUC) was used to evaluate the performance of different hyperparameter combinations. AutoML visualizations included feature importance, SHapley Additive exPlanation (SHAP) plots, and Local Interpretable Model-agnostic Explanations (LIME). Feature importance highlights each feature's contribution to the ML model's predictions. SHAP provides insights into how specific features influence individual predictions, offering clear, interpretable explanations. LIME interprets predictions by fitting a locally interpretable model around a specific instance.

Results

Baseline characteristics

The clinical characteristics and baseline data of 1,224 patients are detailed in Table 1. Based on the diagnostic criteria, 389 patients (31.8%) were identified as having developed SIRS. For model validation, patients from Changshu Hospital and the First Affiliated Hospital of Soochow University were randomly allocated into a training set ($n = 812$, 80%) and a validation set ($n = 200$, 20%). Furthermore, patients from the Second Affiliated Hospital of Soochow University served as an independent test set ($n = 212$). Within the training, internal validation, and test sets, 274 patients (33.7%), 68 patients (34.0%), and 47 patients (22.2%) respectively developed SIRS.

Development of prediction model

Univariate analysis was performed using the LASSO regression model with the "λmin + 1se (0.059)" criterion and 5-fold cross-validation to address multicollinearity among variables (Supplementary Figure S1). For multivariate analysis, backward stepwise LR identified three independent risk factors out of 34 variables, which were used to create a nomogram (Fig. 2). And, the results of the multivariate LR can be found in Supplementary Table 2. Calibration curves for the training, validation, and test datasets (Fig. 3) showed mean absolute errors of 0.049, 0.023, and 0.044, respectively, highlighting the LASSO model's strong predictive accuracy compared to actual outcomes. DCA for the test set indicated that the LASSO model provided additional benefits when predicting SIRS probabilities between 10% and 90%, with a net benefit ranging from 1 to 17%. For instance, if a clinician estimated a 40% probability of SIRS in a patient, early intervention could offer a 6% additional benefit, equating to detecting 6 cases of SIRS without unnecessary treatments out of 100 patients. This approach outperforms the "treat none" strategy (represented by the horizontal line in Supplementary Figure S2), which assumes no true positives or false positives. Net benefit analysis confirms that the LASSO model improves patient outcomes regardless of individual preferences. The test set's ROC curve (Supplementary Figure S3) achieved an AUC of 0.856, as detailed in Table 2.

The study constructed 53 models using four ML algorithms: Deep Learning (DL), Gradient Boosting Machine (GBM), Generalized Linear Model (GLM), and Distributed Random Forest (DRF). Stacked ensemble models were omitted due to challenges in interpretation. Among these, the DL model demonstrated the best performance, achieving the highest AUC of 0.867 in the test cohort and effectively managing imbalanced data. Figure 4 highlights that in the GBM model, neutrophil count (N), white blood cell count (WBC), and C-reactive protein (CRP) were the most critical features, followed by lactate dehydrogenase (LDH), serum calcium (Ca^{2+}), alkaline phosphatase (ALP), lymphocyte ratio (Lr), activated partial thromboplastin time (APTT), gamma-glutamyl transferase (GGT), and international normalized ratio (INR). N and CRP were common significant variables in both the GBM and LASSO models. The SHAP plot in Fig. 5 ranks the top ten important features in the GBM model: WBC, CRP, N, LDH, Ca^{2+}, Lr, ALP, GGT, prothrombin time (PT), and total bilirubin (TB). Variables closer to a value of 1 were more strongly associated with a higher likelihood of SIRS progression. For example, the red portion of CRP, concentrated on the right of axis = 0, suggests that elevated CRP levels in patients with AP are linked to an increased risk of SIRS. The LIME plot of the GBM model demonstrates the contributions of key variables to

Table 1 Baseline characteristics of patients in training, validation and test groups

Variables	Group	The developing, dataset ($n=1012$)		*p*-value	The test dataset ($n=212$)		*p*-value
		Non-SIRS($n=670$)	SIRS($n=342$)		Non-SIRS($n=165$)	SIRS($n=47$)	
Sex (%)	Male	374 (55.8)	220 (64.3)	0.011	107 (64.8)	35 (74.5)	0.289
	Female	296 (44.2)	122 (35.7)		58 (35.2)	12 (25.5)	
Age(mean [IQR])		53.50 [40.00, 67.00]	46.50 [34.00, 60.75]	< 0.001	49.00 [37.00, 65.00]	39.00 [33.50, 65.50]	0.138
Etiology (%)	Biliary	308 (46.0)	136 (39.8)	< 0.001	78 (47.3)	15 (31.9)	0.003
	Hyperlipidemia	101 (15.1)	101 (29.5)		23 (13.9)	18 (38.3)	
	Alcoholic	32 (4.8)	21 (6.1)		18 (10.9)	3 (6.4)	
	Others	229 (34.2)	84 (24.6)		46 (27.9)	11 (23.4)	
Smoke (%)	No	591 (88.2)	284 (83.0)	0.05	135 (81.8)	35 (74.5)	0.364
	Yes	79 (11.6)	58 (17.0)		30 (18.2)	12 (25.5)	
Hypertension (%)	No	447 (66.7)	221 (64.6)	0.551	122 (73.9)	32 (68.1)	0.543
	Yes	223 (33.3)	121 (35.4)		43 (26.1)	15 (31.9)	
Diabetes (%)	No	580 (86.6)	295 (86.3)	0.969	144 (87.3)	34 (72.3)	0.025
	Yes	90 (13.4)	47 (13.7)		21 (12.7)	13 (27.7)	
SBP (median [IQR])		131.00 [120.00, 143.00]	134.00 [121.00, 144.75]	0.197	130.00 [117.00, 140.00]	128.00 [122.00, 136.50]	0.892
DBP (median [IQR])		79.00 [71.00, 85.00]	80.00 [72.00, 88.00]	0.012	75.00 [70.00, 87.00]	80.00 [70.00, 89.00]	0.106
MAP (median [IQR])		96.67 [89.33, 104.33]	98.33 [89.75, 106.67]	0.051	92.67 [86.00, 102.67]	96.00 [88.00, 102.83]	0.233
PLT (*10^9/L) (median [IQR])		190.00 [152.25, 234.00]	204.00 [162.00, 247.50]	0.005	210.00 [170.00, 248.00]	217.00 [173.00, 270.00]	0.262
WBC (*10^9/L) (median [IQR])		10.73 [8.26, 13.80]	15.40 [12.74, 18.84]	< 0.001	11.10 [8.40, 13.40]	15.60 [12.55, 18.35]	< 0.001
N (*10^9/L) (median [IQR])		8.79 [6.20, 11.67]	13.37 [10.77, 16.58]	< 0.001	9.00 [6.40, 11.60]	13.30 [11.05, 15.90]	< 0.001
L (*10^9/L) (median [IQR])		1.27 [0.80, 1.90]	1.10 [0.70, 1.60]	0.001	1.30 [0.80, 1.80]	1.00 [0.60, 1.65]	0.041
HCT (L/L) (median [IQR])		0.42 [0.39, 0.46]	0.44 [0.40, 0.47]	< 0.001	0.43 [0.40, 0.48]	0.44 [0.41, 0.48]	0.267
RDW (%) (median [IQR])		12.80 [12.40, 13.40]	12.90 [12.40, 13.40]	0.843	12.80 [12.40, 13.30]	12.80 [12.30, 13.05]	0.644
Lr (%) (median [IQR])		12.35 [7.20, 18.67]	6.90 [4.40, 10.97]	< 0.001	12.60 [8.10, 17.80]	6.90 [3.90, 11.45]	< 0.001
Cr (μmol/L) (median [IQR])		64.00 [54.52, 76.25]	62.00 [49.10, 75.57]	0.011	65.00 [55.00, 75.00]	65.00 [51.50, 75.50]	0.741
TB (μmol/L) (median [IQR])		20.70 [14.53, 32.10]	22.45 [15.33, 32.12]	0.335	17.40 [12.10, 27.30]	17.30 [12.65, 23.80]	0.673
DB (μmol/L) (median [IQR])		7.10 [4.60, 13.80]	7.60 [4.12, 12.80]	0.642	8.00 [5.50, 13.50]	8.30 [5.05, 12.25]	0.583
Urea (mmol/L) (median [IQR])		4.80 [3.80, 6.20]	5.20 [3.90, 7.00]	0.005	4.20 [3.40, 5.80]	4.30 [3.55, 7.25]	0.159
LDH (U/L) (median [IQR])		211.00 [171.93, 287.00]	260.80 [209.72, 390.75]	< 0.001	193.00 [163.00, 236.00]	271.00 [174.00, 351.00]	< 0.001
Ca^{2+}(mmol/L) (median [IQR])		2.18 [2.08, 2.29]	2.08 [1.95, 2.23]	< 0.001	2.11 [2.01, 2.21]	2.06 [1.92, 2.16]	0.035
TG (mmol/L) (median [IQR])		1.31 [0.84, 2.52]	1.94 [1.08, 5.93]	< 0.001	1.01 [0.64, 2.31]	2.28 [0.95, 8.34]	0.001
GLU (mmol/L) (median [IQR])		6.95 [5.70, 8.92]	7.72 [6.23, 10.38]	< 0.001	6.64 [5.43, 8.82]	10.46 [7.14, 13.70]	< 0.001
ALT (U/L) (median [IQR])		41.90 [19.00, 144.50]	30.30 [16.83, 78.07]	0.001	43.00 [19.00, 162.00]	21.00 [13.00, 66.00]	0.018
AST (U/L) (median [IQR])		33.20 [19.18, 100.50]	26.90 [17.22, 53.73]	< 0.001	29.00 [17.00, 77.00]	23.00 [16.00, 48.50]	0.088
GGT (U/L) (median [IQR])		89.55 [33.62, 281.55]	78.20 [38.23, 203.95]	0.437	102.00 [40.00, 273.00]	103.00 [42.50, 197.00]	0.824
ALP (U/L) (median [IQR])		92.10 [69.90, 141.75]	78.00 [59.50, 106.90]	< 0.001	85.00 [67.00, 135.00]	83.00 [64.00, 104.50]	0.325
ALB (g/L) (median [IQR])		37.20 [33.90, 40.60]	35.00 [31.80, 39.00]	< 0.001	38.90 [35.40, 41.10]	36.60 [33.45, 39.50]	0.005
K^+(U/L) (median [IQR])		3.99 [3.72, 4.30]	3.98 [3.73, 4.28]	0.772	4.11 [3.84, 4.39]	4.24 [4.00, 4.51]	0.066
PT (s) (median [IQR])		13.00 [12.20, 14.00]	13.50 [12.50, 14.70]	< 0.001	13.70 [13.20, 14.50]	14.00 [13.40, 14.75]	0.173

Table 1 (continued)

Variables	Group	The developing, dataset ($n=1012$)		*p*-value	The test dataset ($n=212$)		*p*-value
		Non-SIRS($n=670$)	SIRS($n=342$)		Non-SIRS($n=165$)	SIRS($n=47$)	
INR (median [IQR])		1.07 [1.01, 1.15]	1.11 [1.02, 1.22]	< 0.001	1.06 [1.02, 1.14]	1.10 [1.02, 1.16]	0.212
APTT (s) (median [IQR])		31.40 [28.20, 35.70]	33.20 [28.72, 38.00]	0.002	37.00 [34.20, 41.30]	37.30 [34.15, 41.10]	0.97
CRP (mg/L) (median [IQR])		13.62 [2.50, 75.72]	109.72 [16.58, 208.45]	< 0.001	66.50 [15.10, 127.10]	165.30 [113.95, 238.45]	< 0.001

SBP, Systolic blood pressure; DBP, Diastolic blood pressure; MAP, Mean artery pressure; N, Neutrophil count; L, Lymphocyte count; Lr, Percentage of lymphocytes; Cr, Creatinine; TB, Total bilirubin; DB, Direct bilirubin; TG, Total triglycerides; GLU, Glucose; ALB, Albumin

SIRS progression, with Supplementary Figure S4 showing that both cases 1 and 2 were predicted to have a high risk of SIRS. In both cases, Ca^{2+} was the most influential feature, while N contributed the least to the risk of SIRS progression.

DCA for the test set showed that AutoML models predicting SIRS probabilities between 10% and 100% could offer additional benefits ranging from 1 to 20%. For instance, if a clinician assessed a patient's SIRS likelihood at 20%, early intervention could improve outcomes by at least 15%. This approach significantly outperforms the "treat none" strategy (represented by the horizontal line in Supplementary Figure S5), which results in no true positives or false positives.

Comparisons models developed by LR and automl

In the test set, the AUC values of the five models were as follows: the DL model had the highest AUC at 0.867, followed by the LASSO model at 0.856, the GLM at 0.853, the GBM at 0.833, and the DRF at 0.830. Among these models, the DL model achieved an AUC and accuracy both exceeding 0.80, demonstrating better predictive performance compared to the other models. Details are listed in Table 2.

Discussion

AP can be classified into three stages based on its pathological progression: the acute response phase, systemic infection phase, and residual infection phase. During the acute response phase, the self-digestion of pancreatic enzymes triggers the activation of inflammatory factors, accompanied by the activation of inflammatory cells, the release of chemokines, adhesion molecules, reactive oxygen species, platelet-activating factor (PAF), and endothelin. This causes local inflammation to escalate into SIRS [3, 5]. If SIRS persists, endotoxins and phospholipase A2, through systemic circulation and the mesenteric lymphatic pathway, sustain and amplify the inflammatory cascade, potentially leading to more severe SIRS and even multiple organ dysfunction syndrome (MODS) [3]. Research by Tan et al. and Sharma et al. further found that AP patients with SIRS are more prone to developing infected pancreatic necrosis (IPN), and the duration of SIRS is comparable to APACHE II and CT Severity Index (CTSI) scores in predicting IPN, MODS, and mortality [19, 20]. Studies have shown that early onset of SIRS in AP is usually associated with a higher mortality rate [21]. Therefore, early and accurate assessment of AP with SIRS is crucial for improving prognosis.

In this study, we developed and validated a range of models for the early detection of SIRS using both AutoML and LR methods. Compared to traditional univariate and multivariate analyses, AutoML significantly reduced the time required while improving accuracy,

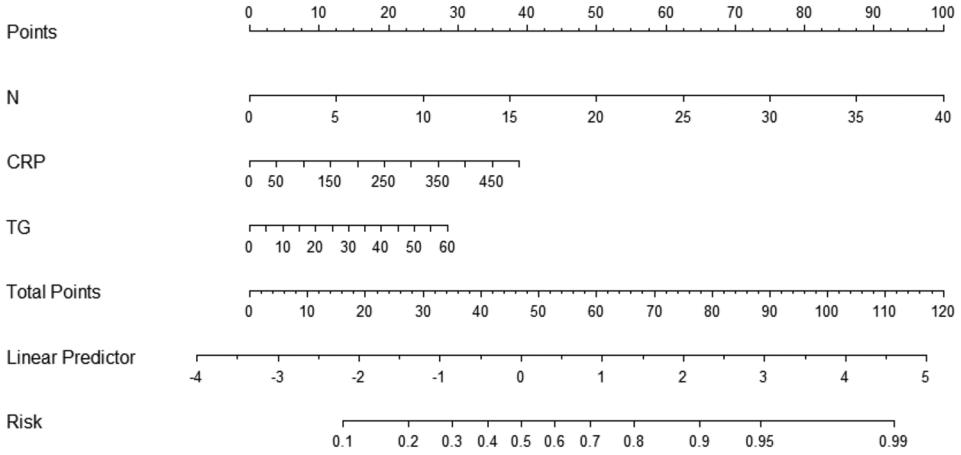

Fig. 2 Nomogram of the LASSO model for the early prediction of SIRS

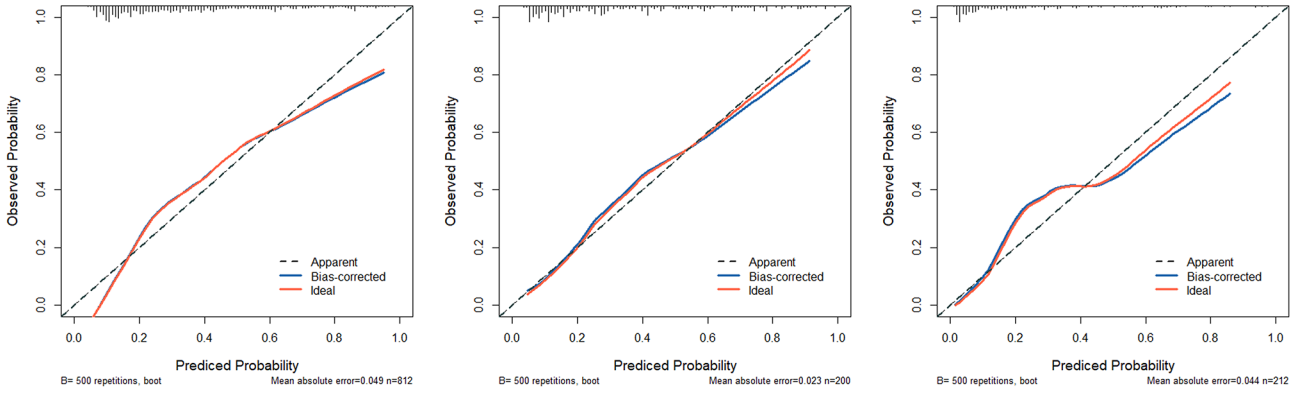

Fig. 3 Calibration curve of the LASSO model in the training, validation, and test set

Table 2 Comparison of LR and automl models for early prediction of SIRS in the test cohort

AUC		Sensitivity	Specificity	Accuracy	PPV	NPV	LR+	LR−
AutoML								
GBM	0.833	0.936	0.612	0.684	0.407	0.971	2.414	0.104
DRF	0.830	0.979	0.564	0.656	0.390	0.989	2.243	0.038
GLM	0.853	0.957	0.624	0.698	0.421	0.981	2.548	0.068
DL	0.867	0.787	0.812	0.807	0.544	0.931	4.190	0.262
Logistic regression								
LASSO	0.856	0.915	0.697	0.745	0.462	0.966	3.019	0.122

LR, Logistic regression; AutoML, Automated machine learning; SIRS, Systemic inflammatory response syndrome; PPV, Positive predictive value; NPV, Negative predictive value; LR+, Positive likelihood ration; LR−, Negative likelihood ratio

which greatly enhanced operational efficiency. Additionally, the ensemble model integrated multiple machine learning algorithms and used several classifiers to combine prediction outcomes, resulting in improved overall performance [22]. We applied four AutoML algorithms (GBM, DRF, GLM, and DL) for the early prediction of SIRS, with all models outperforming standard algorithmic techniques. The DL model achieved the highest AUC in the test set, a critical indicator for evaluating model effectiveness. Since our goal was to quickly identify AP patients at risk of developing SIRS, sensitivity and accuracy were also crucial assessment metrics. Both the DRF and GLM models demonstrated sensitivities above 0.950, while the DL model achieved an accuracy of 0.807. Therefore, in this study, the DL model was found to be the most successful.

WBC and N are important indicators of inflammatory response and are widely used in the diagnosis and prognostic assessment of inflammatory diseases [23, 24]. When the body experiences severe infection or an inflammatory response, both WBC and N levels typically rise significantly, reflecting the immune system's reaction

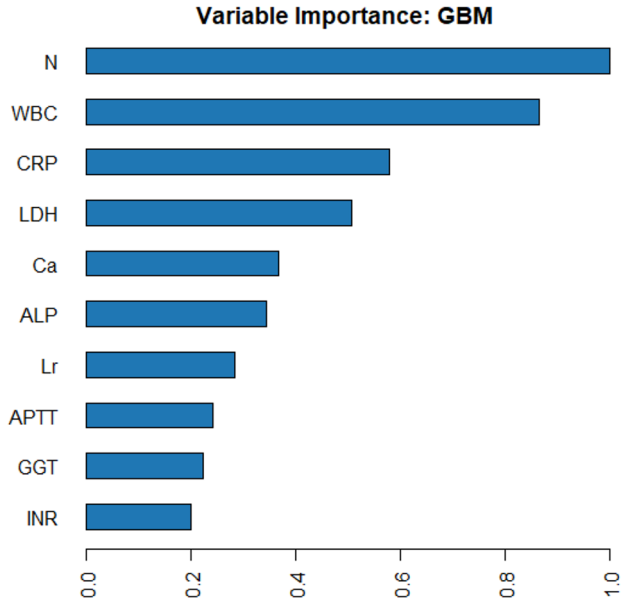

Fig. 4 Variable importance of the GBM model in the training set

to pathogens or injury [24, 25]. Therefore, it is not surprising that both are considered predictor variables for SIRS. CRP is a common acute phase reactant closely related to inflammatory responses. It is synthesized and secreted by the liver and is essential for maintaining normal immune function. Under normal circumstances, CRP levels remain within a standard range, but they rise rapidly when the body is subjected to an inflammatory cell invasion. CRP has excellent stability and detectability, making its level changes highly sensitive indicators of the degree of inflammatory response and cellular or tissue damage [13]. Relevant studies indicate that changes in CRP levels can occur significantly earlier than the onset of clinical symptoms [24]. Some research also suggests that the CRP level at admission serves as a predictive factor for the severity of acute pancreatitis [26].

Kolodecik TR et al. found that measuring LDH activity within 12 h of onset can serve as a biomarker for the early prediction of acute pancreatitis (AP) prognosis, with a sensitivity of 63.6% and specificity of 89.6% [27]. Additionally, LDH is also a key variable in the RANSON

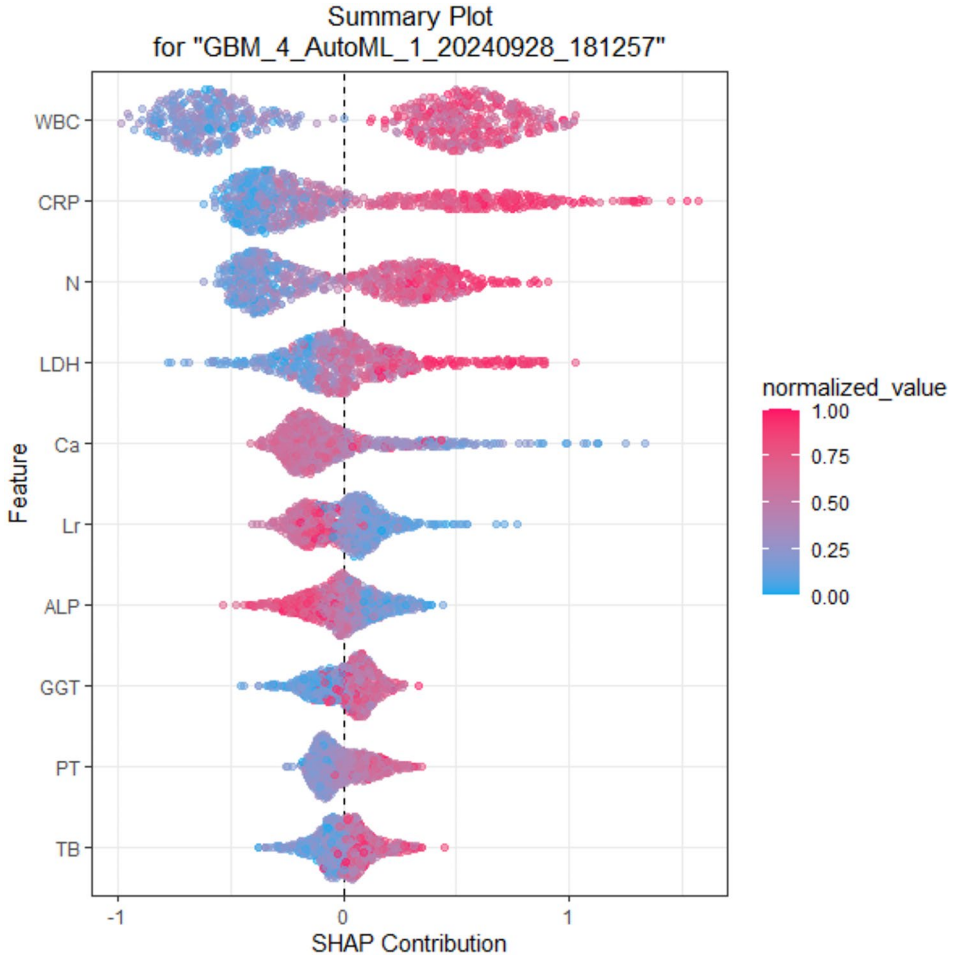

Fig. 5 SHAP of the GBM model in the training set

score [28]. In an in vitro study of lymphocytes from burned rats, it was found that the transformation activity of lymphocytes significantly increased, accompanied by elevated levels of interleukins (IL-1, IL-2, IL-6), which was consistent with changes in cell proliferation [29]. Furthermore, the study indicated that in addition to macrophages and neutrophils, lymphocytes also expressed Toll-like receptor 4 (TLR4) and nuclear factor kappa B (NF-κB), both of which provide the basis for initiating and promoting the inflammatory response [30]. TLR4, one of the earliest discovered TLR subtypes, is a type I transmembrane protein composed of extracellular, I transmembrane, and intracellular regions. It is the only confirmed receptor and signaling molecule within the TLR family that recognizes lipopolysaccharide (LPS), a key endotoxin component, thus playing an initiating role in the inflammatory response [31]. When LPS binds to TLR4 on lymphocytes, it activates the main signaling pathway, the NF-κB pathway [32]. This pathway triggers and promotes the transcription of various genes related to the inflammatory response, encoding a range of cytokines and inflammatory mediators, including TNF-α, intercellular adhesion molecule-1, IL-1, IL-2, IL-6, and prostaglandins, thereby driving the inflammatory process. If these cytokines and inflammatory mediators are excessively expressed and released, they can lead to a cascade of inflammatory mediators, rapidly escalating local inflammation into SIRS.

When the host's immune system is unable to combat pathogenic factors, it may lead to hepatocyte damage, obstructed bile secretion, and disrupted bilirubin metabolism, resulting in the appearance of jaundice. Research indicates that the dysfunction of stem cells during the development of sepsis may be associated with the release of inflammatory cytokines, which is often closely related to the progression of SIRS. SIRS is a systemic inflammatory response of the body to infection or other pathogenic factors, characterized by the excessive release of inflammatory cytokines. GGT, ALP, and TB are widely distributed in the biliary system and are primarily excreted through the bile ducts, making them important serum markers for assessing biliary obstruction [33]. In our study, patients with higher levels of these three markers had an increased risk of developing SIRS, which may be related to prolonged biliary obstruction and poor bacterial growth in the bile. Animal studies have also shown that the concentration of the GGT marker is closely associated with the risk of death from SIRS [34].

There is a close interaction between the inflammatory response and the coagulation mechanism. The inflammatory response can activate the coagulation system, while the coagulation mechanism also plays a regulatory role in the inflammatory response. During the development of sepsis, pathogens produce endotoxins and exotoxins, which stimulate the release of a large number of inflammatory cytokines and mediators, activating the coagulation system and leading to a hyper-coagulable state in the blood. At the same time, this response also inhibits the fibrinolytic and anticoagulant systems to varying degrees. The formation of microthrombi in the microvasculature can result in microcirculatory disturbances [35]. Therefore, it is not surprising that PT is used as a predictive indicator for SIRS. In the systemic inflammatory response of sepsis, inflammatory factors such as interleukin-1β and interleukin-6 can bind to calcium-sensing receptors (CaSR) present in the parathyroid gland and renal tubules, leading to a decrease in parathyroid hormone (PTH), 1,25-dihydroxyvitamin D, and blood calcium levels [36]. Vitamin D is closely related to calcium regulation, and Watkins has noted a high incidence of vitamin D deficiency in patients with sepsis. The SIRS can impair the expression of vitamin D receptors in immune cells and reduce the levels of vitamin D binding protein. Supplementing with vitamin D may have anti-inflammatory effects and could help improve sepsis and the coagulopathy associated with disseminated intravascular coagulation (DIC) [37].

This study has the advantage of utilizing AutoML to construct a series of models that are more accurate and sensitive in the early prediction of SIRS compared to traditional algorithms. The predictive factors in these models are based on routine detection indicators for AP, and our study is multicenter, facilitating the models' application and promotion. However, there are some limitations to this study. First, novel biomarkers highlighted in recent research were not included, as they have not yet been widely implemented in clinical practice. Second, since this is a retrospective study, further prospective research is required to validate our findings. Finally, the sample size of this study is relatively small, and larger sample-size studies are needed to confirm our conclusions.

Conclusions

We developed and validated a series of models using the AutoML platform for the early prediction of SIRS in patients with AP. These models outperformed scoring systems constructed with traditional algorithms and may provide direction for the application of AutoML in future medical research. Additionally, the performance of the DL model is better than that of the other models.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12911-025-02997-7.

Supplementary Figure S1: Penalty chart of predictive factors for SIRS based on LASSO regression analysis

Supplementary Figure S2: Decision curve analysis of the LASSO model in test set

Supplementary Figure S3: ROC curves of all proposed models (GLM, GBM, DRF, DL, and LASSO models) in the test set

Supplementary Figure S4: LIME of the GBM model in the test set

Supplementary Figure S5: Decision curve analysis plots of the 4 models in the test set, indicating net benefits of around 1-20%

Supplementary Table S1: Attributes and variables extracted from electronic medical records

Supplementary Table S2: Multifactorial logistic regression analysis of acute pancreatitis (AP) complicated by systemic inflammatory response syndrome (SIRS)

Acknowledgements
Not applicable.

Author contributions
R.F.Z. was responsible for data curation, methodology, software development, and visualization, contributing equally to the writing of the original draft and review & editing. S.Q. Z. led the formal analysis and made equal contributions to methodology and resources. L.S. provided support in data curation and writing the original draft. H. Z. supported formal analysis, resources, software, and validation. X.D. X. led the funding acquisition and provided support in resources and supervision. B.X. contributed equally in supervision, writing the original draft, and review & editing. M.W. led in resources and supervision, contributing equally to writing the original draft and review & editing.

Funding
Suzhou Key Disease Project (LCZX202334).

Data availability
The data that support the findings of this study are available on request from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
This study has received ethical approval from the Affiliated Changshu Hospital of Soochow University, and informed consent has been obtained from all patients. Our study complies with the relevant provisions of the Declaration of Helsinki.

Consent for publication
Not applicable.

Clinical trial number
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 14 November 2024 / Accepted: 7 April 2025
Published online: 17 April 2025

References

1. Prajapati R, Manay P, Sugumar K, Rahandale V, Satoskar R. Acute pancreatitis: predictors of mortality, pancreatic necrosis and intervention. Turk J Surg. 2021;37(1):13–21.
2. Xu F, Chen X, Li C, Liu J, Qiu Q, He M, Xiao J, Liu Z, Ji B, Chen D et al. Prediction of Multiple Organ Failure Complicated by Moderately Severe or Severe Acute Pancreatitis Based on Machine Learning: A Multicenter Cohort Study. Mediators Inflamm. 2021:5525118.
3. Komara NL, Paragomi P, Greer PJ, Wilson AS, Breze C, Papachristou GI, Whitcomb DC. Severe acute pancreatitis: capillary permeability model linking systemic inflammation to multiorgan failure. Am J Physiol Gastrointest Liver Physiol. 2020;319(5):G573–83.
4. Habtezion A, Gukovskaya AS, Pandol SJ. Acute pancreatitis: A multifaceted set of organelle and cellular interactions. Gastroenterology. 2019;156(7):1941–50.
5. Boxhoorn L, Voermans RP, Bouwense SA, Bruno MJ, Verdonk RC, Boermeester MA, van Santvoort HC, Besselink MG. Acute pancreatitis. Lancet (London England). 2020;396(10252):726–34.
6. Sendler M, van den Brandt C, Glaubitz J, Wilden A, Golchert J, Weiss FU, Homuth G, De Freitas Chama LL, Mishra N, Mahajan UM, et al. NLRP3 inflammasome regulates development of systemic inflammatory response and compensatory Anti-Inflammatory response syndromes in mice with acute pancreatitis. Gastroenterology. 2020;158(1):253–e269214.
7. Yang R, Tenhunen J, Tonnessen TI. HMGB1 and histones play a significant role in inducing systemic inflammation and multiple organ dysfunctions in severe acute pancreatitis. Int J Inflam. 2017;2017:1817564.
8. Chen Z, Dong WH, Wu Q, Wang J. Two-layer regulation of TRAF6 mediated by both TLR4/NF-kB signaling and miR-589-5p increases Proinflammatory cytokines in the pathology of severe acute pancreatitis. Am J Transl Res. 2020;12(6):2379–95.
9. Iyer S, Bawa EP, Tarique M, Dudeja V. Know Thy Enemy-Understanding the role of inflammation in severe acute pancreatitis. Gastroenterology. 2020;158(1):46–8.
10. Dixit A, Cheema H, George J, Iyer S, Dudeja V, Dawra R, Saluja AK. Extracellular release of ATP promotes systemic inflammation during acute pancreatitis. Am J Physiol Gastrointest Liver Physiol. 2019;317(4):G463–75.
11. Dumnicka P, Maduzia D, Ceranowicz P, Olszanecki R, Drożdż R, Kuśnierz-Cabala B. The interplay between inflammation, coagulation and endothelial injury in the early phase of acute pancreatitis: clinical implications. Int J Mol Sci. 2017;18(2).
12. Liu S, Qu X, Liu F, Wang C. Pentraxin 3 as a prognostic biomarker in patients with systemic inflammation or infection. Mediators Inflamm. 2014;2014:421429.
13. Staubli SM, Schäfer J, Rosenthal R, Zeindler J, Oertli D, Nebiker CA. The role of CRP and pentraxin 3 in the prediction of systemic inflammatory response syndrome and death in acute pancreatitis. Sci Rep. 2019;9(1):18340.
14. Arefan D, Mohamed AA, Berg WA, Zuley ML, Sumkin JH, Wu S. Deep learning modeling using normal mammograms for predicting breast cancer risk. Med Phys. 2020;47(1):110–8.
15. Le S, Hoffman J, Barton C, Fitzgerald JC, Allen A, Pellegrini E, Calvert J, Das R. Pediatric severe sepsis prediction using machine learning. Front Pead. 2019;7:413.
16. Banks PA, Bollen TL, Dervenis C, Gooszen HG, Johnson CD, Sarr MG, Tsiotos GG, Vege SS. Classification of acute pancreatitis–2012: revision of the Atlanta classification and definitions by international consensus. Gut. 2013;62(1):102–11.
17. Mounzer R, Langmead CJ, Wu BU, Evans AC, Bishehsari F, Muddana V, Singh VK, Slivka A, Whitcomb DC, Yadav D, et al. Comparison of existing clinical scoring systems to predict persistent organ failure in patients with acute pancreatitis. Gastroenterology. 2012;142(7):1476–82. quiz e1415-1476.
18. Blazek K, van Zwieten A, Saglimbene V, Teixeira-Pinto A. A practical guide to multiple imputation of missing data in nephrology. Kidney Int. 2021;99(1):68–74.
19. Tan C, Yang L, Shi F, Hu J, Zhang X, Wang Y, Deng Z, Li J, Yuan H, Shi T, et al. Early systemic inflammatory response syndrome duration predicts infected pancreatic necrosis. J Gastrointest Surg. 2020;24(3):590–7.
20. Sharma D, Jakkampudi A, Reddy R, Reddy PB, Patil A, Murthy HVV, Rao GV, Reddy DN, Talukdar R. Association of systemic inflammatory and Anti-inflammatory responses with adverse outcomes in acute pancreatitis: preliminary results of an ongoing study. Dig Dis Sci. 2017;62(12):3468–78.
21. Machicado JD, Gougol A, Tan X, Gao X, Paragomi P, Pothoulakis I, Talukdar R, Kochhar R, Goenka MK, Gulla A, et al. Mortality in acute pancreatitis with persistent organ failure is determined by the number, type, and sequence of organ systems affected. United Eur Gastroenterol J. 2021;9(2):139–49.
22. Goh KH, Wang L, Yeow AYK, Poh H, Li K, Yeow JJL, Tan GYH. Artificial intelligence in sepsis early prediction and diagnosis using unstructured data in healthcare. Nat Commun. 2021;12(1):711.
23. Thomas MR, Storey RF. The role of platelets in inflammation. Thromb Haemost. 2015;114(3):449–58.
24. Liang Y, Zhao X, Meng F. Procalcitonin, C-Reactive protein, and neutrophil ratio contribute to the diagnosis and prognosis of severe acute pancreatitis. Iran J Public Health. 2019;48(12):2177–86.
25. Shahim P, Linemann T, Inekci D, Karsdal MA, Blennow K, Tegner Y, Zetterberg H, Henriksen K. Serum Tau fragments predict return to play in concussed professional ice hockey players. J Neurotrauma. 2016;33(22):1995–9.

26. Valverde-López F, Matas-Cobos AM, Alegría-Motte C, Jiménez-Rosales R, Úbeda-Muñoz M, Redondo-Cerezo E. BISAP, RANSON, lactate and others biomarkers in prediction of severe acute pancreatitis in a European cohort. J Gastroenterol Hepatol. 2017;32(9):1649–56.
27. Kolodecik TR, Reed AM, Date K, Shugrue CA, Patel V, Chung SL, Desir GV, Gorelick FS. The serum protein Renalase reduces injury in experimental pancreatitis. J Biol Chem. 2017;292(51):21047–59.
28. Qu C, Gao L, Yu XQ, Wei M, Fang GQ, He J, Cao LX, Ke L, Tong ZH, Li WQ. Machine Learning Models of Acute Kidney Injury Prediction in Acute Pancreatitis Patients. Gastroenterol Res Pract. 2020:3431290.
29. Kataranovski M, Kucuk J, Lilić D, Drasković-Pavlović B, Colić M, Pejnović N, Dujić A. Increased activity of lymph node cells in thermal injury. Reg Immunol. 1992;4(4):197–203.
30. Gururajan M, Jacob J, Pulendran B. Toll-like receptor expression and responsiveness of distinct murine Splenic and mucosal B-cell subsets. PLoS ONE. 2007;2(9):e863.
31. Säemann MD, Weichhart T, Zeyda M, Staffler G, Schunn M, Stuhlmeier KM, Sobanov Y, Stulnig TM, Akira S, von Gabain A, et al. Tamm-Horsfall glycoprotein links innate immune cell activation with adaptive immunity via a Toll-like receptor-4-dependent mechanism. J Clin Invest. 2005;115(2):468–75.
32. Janssens S, Beyaert R. Role of Toll-like receptors in pathogen recognition. Clin Microbiol Rev. 2003;16(4):637–46.
33. Cheng L, Niu J, Cheng Y, Liu J, Shi M, Huang S, Ding X, Li S. Risk factors for systemic inflammatory response syndrome after percutaneous transhepatic cholangioscopic lithotripsy. J Inflamm Res. 2024;17:2575–87.
34. Jaramillo C, Renaud DL, Arroyo LG, Kenney DG, Gamsjaeger L, Gomez DE. Serum haptoglobin concentration and liver enzyme activity as indicators of systemic inflammatory response syndrome and survival of sick calves. J Vet Intern Med. 2022;36(2):812–9.
35. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, Bellomo R, Bernard GR, Chiche JD, Coopersmith CM, et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3). JAMA. 2016;315(8):801–10.
36. Hendy GN, Canaff L. Calcium-sensing receptor, Proinflammatory cytokines and calcium homeostasis. Semin Cell Dev Biol. 2016;49:37–43.
37. Watkins RR, Yamshchikov AV, Lemonovich TL, Salata RA. The role of vitamin D deficiency in sepsis and potential therapeutic implications. J Infect. 2011;63(5):321–6.

Publisher's note