

85503_Auto_Edited.docx

Name of Journal: *World Journal of Gastroenterology*

Manuscript NO: 85503

Manuscript Type: REVIEW

Acute pancreatitis: A review of diagnosis, severity prediction and prognosis assessment from imaging technology, scoring system and artificial intelligence

Hu JX *et al.* Diagnosis, severity prediction and prognosis assessment of AP

Jian-Xiong Hu, Cheng-Fei Zhao, Shu-Ling Wang, Xiao-Yan Tu, Wei-Bin Huang, Jun-Nian Chen, Ying Xie, Cun-Rong Chen

Abstract

Acute pancreatitis (AP) is a potentially life-threatening inflammatory disease of the pancreas, with clinical management determined by the severity of the disease. Diagnosis, severity prediction, and prognosis assessment of AP typically involve the use of imaging technologies, such as computed tomography (CT), magnetic resonance imaging (MRI) and ultrasound (US), and scoring systems, including Ranson, Acute Physiology and Chronic Health Evaluation II and Bedside Index for Severity in AP scores. CT is considered the gold standard imaging modality for AP due to its high sensitivity and specificity, while MRI and US can provide additional information on biliary obstruction and vascular complications. Scoring systems utilize clinical and laboratory parameters to classify AP patients into mild, moderate, or severe categories, guiding treatment decisions, such as intensive care unit admission, early enteral feeding and antibiotic use. Despite the central role of imaging technologies and scoring systems in AP management, these methods have limitations in terms of accuracy, reproducibility, practicality and economics. Recent advancements of artificial intelligence (AI) provide new opportunities to enhance their performance by analyzing vast amounts of clinical and imaging data. AI algorithms can analyze large amounts of clinical and imaging data, identify scoring system patterns, and predict the clinical course of disease. AI-based models have shown promising results in predicting the severity and mortality of AP, but further validation and standardization are required before widespread clinical application. What's more, we should face up to and pay attention to the correlation between these three technologies, and develop new methods that can accurately, sensitively and specifically be used in the diagnosis, severity prediction and prognosis assessment of AP through complementary advantages.

Key Words: Acute pancreatitis; Imaging technology; Scoring system; Artificial intelligence; Severity prediction; Prognosis assessment

Hu JX, Zhao CF, Wang SL, Tu XY, Huang WB, Chen JN, Xie Y, Chen CR. Acute pancreatitis: A review of diagnosis, severity prediction and prognosis assessment from imaging technology, scoring system and artificial intelligence. *World J Gastroenterol* 2023; In press

Core Tip: In this review, we comprehensively analyze, discuss and summarize the latest progress in the diagnosis, severity prediction and prognosis assessment of acute pancreatitis from the aspects of imaging technology, scoring system and artificial intelligence, so as to provide more guidance and suggestions with clinical value for the diagnosis and treatment of acute pancreatitis.

³ INTRODUCTION

Acute pancreatitis (AP) is an inflammatory disorder resulting from intracellular activation and leakage of improper proteolytic enzymes, including active inflammation and pancreatic injury^[1,2]. AP can result in nausea, vomiting, severe upper abdominal pain, abnormal² release of pancreatic juice, or a systemic inflammatory response syndrome with fever, low blood pressure, and in some cases, failure of one or more organs^[2]. AP¹ is one of the most common causes of hospitalisation from gastrointestinal diseases, with a global incidence rate ranging from 13 to 45 cases per 100000 individuals annually^[2,3]. Globally, the incidence of AP varies, with the North America and Western Pacific regions (as defined by the World Health Organization) experiencing the high rates, surpassing 34 cases per 100000 individuals annually^[4]. The incidence of AP has steadily increased over time in most countries of the Western world^[5]. In the United States, the rate of AP-related hospitalization increased from 65.4 to 81.9 per 100000 adults from 2001 to 2014^[6].

Classification of AP based on severity

AP, often worsened by comorbidities and demographic factors such as obesity, type 2 diabetes, cardiovascular and renal diseases, alcohol use disorder, and age over 45, is classified by severity into three categories: Mild AP (MAP), moderately severe AP (MSAP), and severe AP (SAP)^[7]. SAP can usually result in pancreatic necrosis, systematic inflammation, and multi-organ dysfunction and failure, and its mortality rate, ranging from 20% to 40%, significantly surpasses those of MAP and MSAP^[8]. The volume of extrapancreatic necrosis correlated positively with the complication rate of SAP, potentially serving as an indicator for predicting adverse outcomes in AP^[9]. Predicting SAP with high mortality early on remains a challenge due to the limited accuracy of current predictive tools and the complex clinical features of SAP^[10]. MSAP is characterized by transient organ failure, local complications or exacerbation of comorbid disease, and SAP is defined by persistent organ failure lasting more than 48 h^[11]. MSAP is linked to transient organ failure, while SAP involves persistent organ failure, often necessitating intensive care management^[7]. The early identification of SAP is critical for the stratification and treatment of patients. Additionally, for SAP, it's crucial to avoid interventions that are either excessive and premature or insufficient and delayed; instead, a progressive intervention approach should be implemented at the appropriate time. The critical to play a decisive role is the development of risk stratification tools that meet clinical needs and guide clinicians in terms of resource allocation, patient consultation and clinical audit, as well as the multidisciplinary approaches including evidence-based care that are essential to achieving optimal clinical outcomes^[12]. Therefore, early assessment of the etiology and severity of AP is essential for prompt treatment and close monitoring of severe patients.

Pathophysiology of AP

The pathophysiology of AP involves acinar cell damage, resulting in premature intrapancreatic activation of digestive proteases^[13]. The pathological factors of AP includes calcium (Ca^{2+}) overload, mitochondrial dysfunction, impaired autophagy, endoplasmic reticulum stress, unfolded protein response, intraductal fluid stasis,

genetic mutations (e.g., *PRSS1* or *CTFR* gene), unsaturated fatty acids and exosomes, which mainly lead to inappropriate activation of trypsinogen, infiltration of inflammatory cells and destruction of secretory cells^[14,15]. Ca^{2+} overload is a prevalent mechanism causing cell damage in the body^[15]. Intracellular Ca^{2+} overload and mitochondrial dysfunction, induced by cholecystokinin, excessive alcohol consumption and bile acids, have been identified as key steps in SAP development caused by acinar cell dysfunction^[15]. Mitochondrial dysfunction hinders cell autophagy, leading to increased production of reactive oxygen species and cytokines, which exacerbates pancreatic cell damage^[15]. Mitochondrial injury exacerbates endoplasmic reticulum stress and lysosomal damage, promoting the release and activation of cathepsinogen and trypsinogen, which results in cytoplasmic protein degradation and cell necrosis^[15].

Uncertainty of serum amylase and lipase in diagnosing AP

Common biochemistry markers used in clinical practice include amylase and lipase in serum, but clinicians must be aware of the difference in half-life between the two^[12]. In serum, amylase returns to normal limits within 3 to 5 d, and lipase returns to normal limits within 8 to 14 d^[12]. Elevated serum amylase and/or lipase levels support the clinical suspicion of AP, and the measurement of amylase is more widely used^[16]. However, about 40% of serum amylase is derived from the pancreas, with the rest mainly from the salivary glands^[16]. Therefore, the elevation of serum total amylase is not specific for pancreatitis and some other intra-abdominal diseases should be considered^[16]. For example, Gumaste *et al*^[17] reported that the sensitivity of serum amylase in detecting AP was 72% and the specificity was 99%; In a prospective study including 500 patients with acute abdominal pain, the serum amylase assay had a sensitivity of 85% (with a cutoff value of 300 U/L for the upper reference limit) and a specificity of 91%^[18]; and a prospective study showed that the sensitivity and specificity of total amylase in serum were 45% and 97%, respectively, at the calculated diagnostic threshold of 175 U/L^[19]. In some non-pancreatic diseases, there is also a false elevation of serum amylase. For example, Hu *et al*^[20] reported a case of hyperamylasemia with an

average serum amylase value of 881 U/L, significantly exceeding the reference range of 10-220 U/L. In addition, elevated levels of amylase and lipase, while indicative, are not exclusive to AP and may result from conditions such as bowel obstruction, infarction, cholecystitis, or perforated ulcer^[21]. However, ¹⁸ the sensitivity of serum lipase ranges from 85% to 100%; while some studies reported it was less sensitive than serum amylase, others contended it surpassed amylase in sensitivity^[22].

Current clinical diagnosis of AP

The definition of severity in AP is pivotal for determining ³ the therapeutic approach: patients with MAP typically respond to conservative treatment, while those patients with necrotizing pancreatitis often experience organic dysfunction, necessitating intensive care and regular therapeutic interventions, with a more uncertain prognosis^[1]. Currently, the ⁹ clinical diagnosis of AP necessitates meeting two of the following three criteria: (1) Abdominal pain consistent with AP; (2) serum levels of amylase or lipase exceeding three times the upper normal limit; and (3) cross-sectional abdominal imaging findings consistent with AP^[23]. It is important to note that two of these criteria alone may fail to identify a quarter of AP patients and misdiagnose it in one-tenth^[23].

At present, there is still no single scoring system that can cover all the issues related to the management and evaluation of AP. AP continues to be one of the most intricate digestive disorder in terms of clinical course and outcome, and its inherent variability in each case makes it both challenging and captivating^[24]. Meanwhile, ¹¹ to predict the severity and mortality of AP, clinicians evaluate clinical data, including assessment of organ function, ²⁶ conduct laboratory tests and imaging, and utilize severity-of-the-disease rating systems, such as Ranson, Acute Physiology and Chronic Health Evaluation (APACHE) II, Balthazar's ³⁸ computed tomography severity index (CTSI), modified Mortelet's CTSI (MCTSI), Bedside Index for Severity in AP (BISAP), harmless AP score (HAPS), and first artificial intelligence (AI) model-EASY-APP^[25]. In addition to these, the latest imaging studies and clinical scoring systems for the early diagnosis, prognosis assessment and severity prediction of AP have been extensively studied and reported.

In this review, we provide a detailed discussion and analysis of the latest imaging examinations and some scoring systems applied in this field to afford more valuable guidance to more accurately diagnose, predict and assess AP.

IMAGING TECHNOLOGY

Imaging technology still plays a fundamental role in the initial evaluation, identification of severe cases, prognosis prediction, and decision-making for the treatment and management of AP patients^[1]. An accurate description of imaging findings is crucial in all diseases, particularly in diseases like AP where the appropriate therapy depends on precise diagnosis^[26]. The manifestations of pancreatic diseases are very variable, and imaging plays an important role in the diagnosis and treatment of pancreatic diseases^[27]. Imaging evaluation is still essential to validate the clinical diagnosis, ascertain the etiology, exclude other causes of pain related to elevated levels amylase and/or lipase, and assess the severity and extent of AP^[1].

Imaging modalities for the pancreas encompass plain X-ray, ultrasonography (US), endoscopic ultrasound (EUS), endoscopic retrograde cholangiopancreatography (ERCP), computed tomography (CT), magnetic resonance imaging (MRI), magnetic resonance cholangiopancreatography (MRCP)^[27]. US is usually considered to be the only appropriate modality in the early phase of AP with typical presentations and is used for the detection of gallstones, and CT and MRI are suitable for patients in the early phase of AP with equivocal presentation^[28]. In emergency situations, CT and US are preferred in the imaging modalities of choice due to their advantages of accessibility, speed, and lower cost^[29]. Early detection of CT imaging may influence the diagnosis or treatment in up to 15% of AP patients presenting to the emergency department, particularly in older patients with a history of pancreatitis and biliary interventions; however, abdominal US may offer a more precise screening for biliary etiologies and provide a more informed direction for subsequent treatment^[30]. Based on the fact that US often shows a regular pancreatic structure, the main role of transabdominal US in AP is to identify gallstones and/or choledocholithiasis, seeming

to be useful especially for the evaluation of biliary tract^[31,32]. However, because of the presence and overlap of bowel gas, US is not possible to visualize pancreatic distal abnormality in detection of AP^[31,32]. ¹ In the late phase, typically 48-72 h post-presentation, CT and MRI serve as primary imaging modalities for AP patients, facilitating the assessment of etiology, complications, disease extent, interventions, and subsequent follow-up^[28]. For example, as early as 2007, Stimac *et al*^[33] ²³ reported that non-enhanced MRI was comparable to contrast-enhanced CT (CECT) in the early assessment of AP severity, with both methods demonstrating equal efficacy in predicting local and systemic complications of AP. MRI of the pancreas serves as both a problem-solving tool following CT or US evaluations and an initial imaging examination of choice; furthermore, MRCP is valuable for detecting and evaluating pancreatic ductal anomalies, such as pancreas divisum and annular pancreas^[34].

Abdominal CT

Radiological evaluation, especially by CT, plays a pivotal role in the definition of managing severe cases, particularly in characterizing local complications that impact in the prognosis and dictate the therapeutic approach^[1]. CT is an outstanding non-invasive diagnostic tool for discerning the origins of endocrine and exocrine pancreatic insufficiencies in most patients, and its significance has grown considerably in the diagnosis, treatment, and follow-up of AP patients^[21,27,35]. CT is commonly used to assess the severity of inflammatory process, ascertain the presence and extent of pancreatic necrosis, and identify local complications^[21]. CT with high spatial resolution and rapid acquisition is the preferred for diagnosing AP and associated local complications^[36]. Moreover, CT can clearly display the pancreas and adjacent tissues, and is more precise than US in diagnosing and delineating the extent of pancreatic disease^[36,37]. ⁴⁵ CECT plays a pivotal role in assessing the scope and progression of AP, and stands as the primary imaging modality for initially pinpointing local complications; ³¹ typical cross-sectional imaging features encompass pancreatic enlargement, pancreatic edema, uneven density, peripancreatic fat stranding, and fluid

collection^[15,38]. For example, on CECT, SAP patients typically exhibited larger amounts of peripancreatic retroperitoneal fluid^[39]. Approximately 7 d after the onset of AP, initial CECT playing a significant role in predicting infected pancreatic necrosis underscores the significance for clinicians to contemplate the initial imaging of the pancreas^[40]. What's more, CT is regarded as the gold standard for imaging evaluation of AP, due to its satisfactory effectiveness, outstanding timeliness, and widespread availability^[1]. A lack of clinical response to appropriate conservative treatment within 48-72 h often indicates the necessity for a CT scan to verify the initial diagnosis, assess the severity of the onset, and identify any complications^[41].

In 1990, Balthazar *et al*^[42] developed CTSI by integrating observations of peripancreatic inflammation, phlegmon, and the degree of pancreatic necrosis evident in initial CT examinations. To enhance the accuracy in predicting the prognosis of AP patients, Morteale *et al*^[43] simplified the assessment of fluid collections and the extent of pancreatic necrosis in CTSI, and added features that reflect organ failure and extrapancreatic complications, leading to the development of MCTSI. MCTSI grading of AP was significantly associated with duration of hospitalization, requirements for intensive care unit (ICU), necessity for intervention, and organ failure^[44]. CTSI is an easy-to-calculate and informative tool, and is considered to be a good predictor of mortality and severity of AP^[45]. A prospective study including 50 patients evaluated prognostic correlation and clinical outcome of AP using both Balthazar's CTSI and modified Morteale's CTSI^[27]. In this study, Raghuwanshi *et al*^[27] concluded that the scores derived from the modified Morteale index exhibited a more robust correlation for all outcome parameters in all the patients compared to the Balthazar index. They asserted that CECT served as an outstanding diagnostic tool for staging inflammatory process, identifying pancreatic necrosis, detecting local complications, and grading the severity of AP^[27]. Contrary to expectation, the 2012 Revised Atlanta Classification (RAC) demonstrated greater accuracy than the modified Morteale and Balthazar CTSI in assessing mortality and organ failure among AP patients^[27]. In a study including 178 patients with interstitial edematous pancreatitis (IEP), Song *et al*^[46] indicated that the

initial CECT findings of peripancreatic fluid and heterogeneous enhancement in the pancreatic parenchyma could serve as useful predictors for the progression to necrotizing pancreatitis (NP) in patients initially diagnosed with IEP. However, it was disconcerting that the early CT scan might not conclusively diagnose NP^[46]. Tasu *et al*^[47] demonstrated that a pancreatic enhancement threshold of less than 30 UH on post-contrast CT images during the portal phase provided an accurate and consistent criterion for diagnosing NP. Badat *et al*^[48] highlighted that using the 2012 RAC to categorize pancreatic and peripancreatic collections by CT yielded moderate interobserver agreement, underscoring the potential necessity to either devise a new semiology for characterizing peripancreatic collections by CT or to employ alternative imaging modalities like MRI for more precise analysis of collection contents.

However, the latest relevant clinical research also has encouraging results. A retrospective cross-sectional study enrolled 1924 patients experiencing their first episode of AP from three tertiary referral centers in three different prefecture-level cities of Sichuan Province in China, and revealed a positive rate of 96.7% (1860/1924) for CT findings in AP diagnosis based on CECT^[49]. Among these 1860 AP patients with affirmative CT results, MCTSI exhibited positive correlations with both the 2012 RAC and APACHE II, as evidenced by Spearman's rank correlation coefficients^[49].

Here, there remains a puzzling contradiction, that is, CTSI and MCTSI remain inconsistent in assessing the severity and clinical outcome of AP. Bollen *et al*^[50] determined that there was no notable distinction between CTSI and MCTSI in assessing AP severity, and both CT indexes were more accurate to diagnose AP severity and had a better correlation with the need for intervention and pancreatic infection in comparison with APACHE II. Sahu *et al*^[51] concluded that both CTSI and MCTSI significantly correlated with the clinical outcome of AP and aligned well with RAC grading of severity; MCTSI demonstrated higher sensitivity, albeit with lower specificity than CTSI, in differentiating MAP from MSAP/SAP. Alberti *et al*^[52] determined that CT indexes surpassed APACHE II in assessing the severity in AP, with CTSI holding a slight advantage over MCTSI; additionally, CTSI precisely predicted

pancreatic infections and intervention requirements. Liao *et al*^[53] indicated that both CTSI and MCTSI were significantly associated with clinical prognosis, offering higher accuracy in predicting IPN but less precision in predicting persistent organ failure compared to APACHE II.

Another important factor affecting the effectiveness of CECT in assessing AP severity is the appropriate timing. Dachs *et al*^[54] indicated that early abdominal CT did not offer benefits to afebrile patients experiencing their first episodes of AP. The evidence-based guidelines from International Association of Pancreatology/American Pancreatic Association recommend that the optimal timing for an initial CT assessment should be between 72 and 96 h following the onset of symptoms^[55]. However, until now, the appropriate point in time for when CECT should be performed to provide an accurate assessment for AP has not been well established in clinical practice. For example, in a retrospective study with 309 SAP patients, Huang *et al*^[56] highlighted that the optimal timeframe for CECT evaluation of SAP-associated complications was between 72 h and 1 wk following the onset of SAP, particularly for SAP patients with infection. Their findings revealed that the severity of the disease and its alterations manifested as expanded areas of acute peripancreatic fluid collection (APFC) and increased exudation of pleural effusion within the first 1 wk of SAP onset^[56]. However, the former showed a decrease after 4 wk or more, while the latter reduced after 2 wk or more^[56]. Pocard and Soyer^[57] found that a meticulous review of the current literature failed to offer compelling evidence regarding a specific interval between symptom onset and CT examination, suggesting that the pertinent matter of timely CT examination in AP patients remains inadequately addressed by the existing studies. In this regard, an important and outstanding issue is that the optimal time point for CECT to evaluate SAP patients needs to be determined by larger multicenter clinical studies to improve accuracy of disease diagnosis, avoid unnecessary CECT tests, promote early intervention, and thus improve prognosis.

Chest CT

In AP patients, thoracic complications encompass pleural effusion, pulmonary consolidation, atelectasis, pulmonary embolism, cardiac tamponade, pericardial effusion, elevated diaphragms, mediastinal pseudocysts and acute respiratory distress syndrome (ARDS), the first two of which are common in AP^[58-65]. In AP patients, pleural effusion accounted for 50% on admission, and the emergence of pulmonary consolidation was associated with the onset of respiratory failure^[66]. A retrospective study from three Chinese Acute Pancreatitis Centers showed that 232 out of 465 AP patients had positive pleural effusion, accounting for 49.9%^[67]. In a study including 358 AP patients from seven European centres, more than half of the patients had pleural effusion, with the proportion of 54.4% (195/358), and pleural effusion also appeared mostly bilateral (150/195, 76.9%)^[64]. And it has been reported that AP patients with bilateral pleural effusion had a significantly worse one-year survival^[64]. Bilateral pleural effusion/pulmonary consolidation was suggestive of SAP to a certain extent, and it was considered that measurement of these two parameters has certain clinical value in assessing the severity and prognosis of AP^[68]. Moreover, the early onset of pleural effusion highlights its clinical significance and predicts a poor prognosis in AP^[63].

In a single center study with 309 AP patients, Peng *et al*^[65] explored the predictive significance of semi-quantitative pleural effusion and pulmonary consolidation in determining AP severity using chest CT. In AP patients without organ failure, the values of pleural effusion and pulmonary consolidation were 25.4 ± 23.5 mL and 0.8 ± 1.0 points, respectively, which were lower than the corresponding values of 137.4 ± 116.9 mL and 2.4 ± 1.2 points observed in AP patients with organ failure^[65]. Simultaneously, the values of pleural effusion and pulmonary consolidation in AP patients without death were 39.0 ± 36.0 mL and 1.0 ± 1.1 points, respectively, and were lower than these (144.0 ± 140.3 mL) and (3.0 ± 1.1 points) in the patients who died^[65]. In addition, in predicting SAP, the accuracy of pleural effusion volume (mean value of 41.7 ± 38.0 mL, range of 1-1079 mL) and pulmonary consolidation score (mean value of 1.0 ± 1.2 points, range of 0-5 points) was similar to that of CTSI, APACHE II, and BISAP, and in predicting organ failure, both the parameters had the same accuracy with the three

scores, suggesting that the two parameters could provide prediction of SAP occurrences and organ failure in the early stage^[65]. More importantly, this clinical study may increase the application value of CT due to the important role of these two parameters in predicting AP severity. In a retrospective study from three medical centers, Yan *et al*^[67] reported that the mean volume of pleural effusion was 98.8 ± 113.2 mL in 465 AP patients. The volume of pleural effusion exhibited significant and robust correlations with C-reactive protein (CRP), duration of hospital stay, and such as Ranson, BISAP, Marshall, APACHE II, CTSI, and extra-pancreatic inflammation on CT (EPIC), and displayed considerable accuracy in predicting outcomes like severity, infection, mortality, procedural needs, ICU admission, and organ failure^[67]. Luiken *et al*^[64] categorized the volume of pleural effusion in 195 AP patients as low (48.2%, 94/195), moderate (30.3%, 59/195), and severe (21.5%, 42/195). Their findings suggest that the presence of bilateral and/or moderate to severe amounts of pleural effusion in the early phase of AP could independently predicted SAP^[64].

Thus, the volume of pleural effusion can serve as a dependable radiological biomarker to predict the severity and clinical outcome of AP. In addition, larger and more multi-center prospective studies need to be conducted to promote the clinical application of pleural effusion in the prediction of AP severity. So far, there remains an absence of an established quantitative grading system for pleural effusion, meriting attention in forthcoming clinical research.

Practical problems in the application of CT in AP

Based on advances in predicting and diagnosing AP severity, CECT is considered the diagnostic criterion for assessing AP. However, there is a non-negligible situation where contrast CT is contraindicated in patients with renal dysfunction and in pregnant women, and it is not possible to replicate follow-up studies due to cost and radiation exposure. When uncomplicated AP is diagnosed both clinically and biochemically, CT is superfluous; minimizing its overuse will not only curtail healthcare costs but also diminish radiation exposure to patients^[69]. CT on admission to predict outcome does

not appear to have an advantage compared with the simpler and more readily available clinical scoring systems, so CT on the day of admission to assess severity is not recommended^[70]. Improvement measures aimed at curbing the overuse of early imaging in AP patients may diminish superfluous imaging, elevate quality of care, and curtail wastage^[71].

In addition, CT possesses limitations in assessing the severity of AP, and difficulty distinguishes between necrosis and local effusion in small nonenhanced areas of pancreas^[36]. Without pancreatic parenchymal necrosis, small organized peripancreatic fluid collections might be misconstrued as pseudocysts on CT, leading to an underestimation of extrapancreatic necrosis^[72]. These above disadvantages limit the use of CT in some situations, so there is a need to develop other methods that can be used for the diagnosis and prognosis evaluation of AP. Furthermore, it is recommended that future studies should incorporate reliable non-radiological and laboratory-based categorization tests to enhance the precision in determining and assessing the severity and prognosis of AP, thereby reducing morbidity and mortality associated with post-necrotic inflammation of the pancreas.

MRI

While CT remains the prevalent choice for evaluating AP, MRI has demonstrated greater sensitivity than CT in detecting AP^[34]. MRI, a noninvasive technology boasting high tissue contrast and multiple acquisition sequences, effectively aids in determining the diagnosis, complications, and severity of AP^[36]. When CT yields negative results but there remains a strong clinical suspicion of AP, fat-saturated turbo spin echo T2-weighted or diffusion-weighted imaging sequences can reveal nuanced pancreatic and/or peripancreatic inflammation^[73]. MRI holds a pivotal role in the diagnosis of AP, and is instrumental in assessing and characterizing extrapancreatic necrosis, inflammation, splenomegaly, and tissue involvement—including vascular, transverse-mesocolon, interfascial plane, and gastrointestinal tract—in AP patients^[21,74-80]. And MRI can effectively capture the intra-abdominal inflammatory spread that affects

mesenteric and omental fatty regions, indicative of a pathological manifestation of intra-abdominal fat edema combined with fat necrosis resulting from AP^[81]. MRI is particularly beneficial for imaging of patients with iodine allergies or renal insufficiency, characterizing fluid collections and evaluating abnormalities or disconnections in the pancreatic duct^[38]. As an alternate method for diagnosing AP, MRI shows great potential in clinical applications. MRI offers superior capabilities in diagnosing early extrapancreatic necrosis compared to CT, without the need for radiation, making it suitable for repeated follow-up assessments^[74]. MRI more adeptly identifies the subtlest changes in AP and can delineate the constituents of mild extrapancreatic inflammatory effusions that might be missed on CT^[82]. Fat-saturated T2-weighted MRI offers superior sensitivity in detecting fluid and no liquefied material in extrapancreatic collections compared to CT, while T1-weighted MRI is beneficial for identifying pancreatic or peripancreatic hemorrhage^[74].

MRI in hemorrhage, tissue necrosis, and APFC

Compared to CT, MRI demonstrates superior sensitivity in visualizing hemorrhages, which appear hyperintense on T1WI during the acute phase and maintain their signal intensity longer than on CT^[36]. In necrotizing pancreatitis, MRI offers superior soft tissue contrast compared to CT and excels in visualizing hemorrhage and tissue necrosis^[36]. A retrospective study including 539 AP patients demonstrated that MRI was superior in detecting hemorrhage associated with AP compared to CT, even when CT showed no signs of hemorrhage^[83]. This study revealed that pancreatitis in AP patients with accompanying hemorrhage presented with greater clinical severity, increased susceptibility to organ failure, and prolonged hospital stays, suggesting that early hemorrhage detection on MRI could serve as a novel severity indicator in AP associated with poorer prognosis^[83]. Additionally, due to its enhanced tissue resolution, MRI is poised as the frontline imaging technology for evaluating AP and its complications, notably the identification of hemorrhage^[83]. In a retrospective analysis including 301 AP patients, MRI revealed that 24.9% exhibited at least one peripancreatic

vascular abnormality related to AP, and the incidence of peripancreatic vascular involvement was notably more pronounced in necrotizing pancreatitis compared to edematous pancreatitis^[76]. The common manifestations of early AP on MRI were splenic vein phlebitis and splenic artery involvement/arteritis, and 6.3% of the patients had splenic artery arteritis complicated with hemorrhage in the early phase of AP^[76]. The findings highlighted the efficacy of MRI in delineating the progression of inflammatory processes and associated vascular changes during treatment, and early-stage vascular involvement detected by MRI might serve as a valuable indicator of AP severity^[76].

Since the introduction of abdominal US and CT in the early 1970s, there has been a marked increase in the identification of acute fluid collections in AP, accompanied by a deeper insight into their natural progression and management^[84]. APFC can complicate acute interstitial edematous pancreatitis, manifesting in approximately 30%-50% of such cases^[85]. If APFC was associated with high BISAP (≥ 3) and CRP level (≥ 150 mg/L) after 48 h from admission, or with persistent clinical symptoms reflecting prolonged inflammatory responses, SAP patients with APFC were more likely to develop late complications^[86]. Acute necrotic collections, observed exclusively in necrotic pancreatitis within the first 4 wk of onset, comprise varying amounts of fluid and necrosis, with the latter potentially affecting the pancreatic parenchyma or peripancreatic tissues, or both^[36]. Pancreatic necrosis, characterized by focal, multifocal, or diffuse devitalized tissue within the pancreas, either superficial or deep, is deemed a critical imaging indicator of necrotizing pancreatitis^[81]. A significant correlation exists between the presence of pancreatic necrosis and extrapancreatic fluid collections in relation to the clinical parameters, with an increase in extrapancreatic fluid collections aligning with the escalating severity of AP^[87].

While CT has emerged as the primary non-invasive tool for identifying ⁶ local complications in AP, it struggles to distinguish between APFC and acute necrotic collection in the early phase due to its limited sensitivity in revealing the necrosis debris of peripancreatic tissue^[81]. Given its exceptional resolution for soft tissues, MRI

surpasses CT in delineating pancreatic/peripancreatic fluid collections, especially in quantifying solid debris and fat necrosis, serving as an alternative in cases with CT contraindications^[88]. In MRI findings, hemorrhage in the pancreas and/or surrounding tissues may intermingle with necrosis of these same regions, manifesting as spotted, patchy, or extensive regions of hyperintensity on T1-weighted fat-suppressed images^[81]. In a retrospective study including 70 AP patients, Zhou *et al*^[74] discovered that MRI characteristics of extrapancreatic collections, particularly its extent and amount, could differentiate early extrapancreatic necrosis from peripancreatic fluid collections, suggesting the presence of extrapancreatic necrosis. Moreover, the more extensive the extrapancreatic collections and the broader the scope of extrapancreatic inflammation associated with hemorrhage in AP on MRI, the higher the likelihood of extrapancreatic necrosis^[74].

In a meta-analysis encompassing a total of 566 patients, MRI demonstrated superior accuracy and sensitivity compared to CT for diagnosing AP^[89]. While no study has yet shown that MRI can decrease AP mortality or enhance prognosis, MRI serves as an invaluable diagnostic tool for distinguishing individuals with suspected AP and is regarded as the premier imaging choice for the clinical diagnosis of AP^[89]. Tang *et al*^[82], utilizing MRI and APACHE II, devised a novel model through logistic regression for the early prediction of AP severity, and ascertained that the combined model of extrapancreatic inflammation on MRI (EPIM) and APACHE II excelled in predicting AP severity, surpassing individual parameters. This retrospective analysis including 363 AP patients suggested that merging MRI and APACHE II for gauging AP severity was both viable and more accurate than other scoring mechanisms, potentially facilitating the creation of tailored treatment and management^[82].

MRI severity index in AP

MRI severity index (MRSI), derived from CTSI, evaluates the severity of AP by integrating both peripancreatic inflammation and pancreatic parenchymal necrosis, achieving an effect comparable to that of CTSI in assessing AP severity^[36]. In patients

with pancreatitis, MRSI outperformed APACHE II in assessing local complications, while APACHE II demonstrated superiority in determining systemic complications^[90]. MRSI has emerged as pivotal for the initial assessment, staging, and prognosis of AP. The clinical relevance of MRSI allows for predicting the severity of AP based on initial MRI findings in the early phase, and it holds a significant correlation with APACHE II, incidence of systemic complication, duration of hospital stay, and overall clinical outcome^[81]. In a retrospective study including 337 AP patients, Zhou *et al*^[75] reported that, in the early stages of AP, EPIM based on MRI proved more effective in assessing the severity than EPIC relying on CT. Moreover, the predictive accuracies of EPIM for SAP and organ failure aligned with those of APACHE II and BISAP, surpassing the accuracy of MRSI^[75].

Overall, MRI serves as an excellent instrument for identifying and distinguishing prevalent local complications subsequent to AP. MRI offers diagnostic and prognostic value on par with CT, though it presents certain limitations in clinical practice. The scans necessitate greater cooperation of patient, including prolonged immobility and apnea, and are more time-consuming and costly^[1]. Additionally, MRI has the limitation of a restricted field of view, preventing it from capturing extensive regions of the chest and pelvic cavity simultaneously, as CT can^[81].

US

Based on its quick, simple, repeatable, radiation-free, bedside applicability, US is the first-line imaging method in most medical centers to confirm the diagnosis of AP and exclude other causes of acute abdomen. In the early period, the advantage of US is its capability to assess gallbladder and biliary tract, detecting gallstones and identifying bile duct dilatation^[21]. However, US may show normal pancreas in MAP patients, and is not able to differentiate the diagnosis between interstitial and necrotizing pancreatitis because of not allowing the assessment of parenchymal perfusion^[21]. Endoscopic US (EUS) can identify choledocholithiasis and hidden pancreatic tumors that remain elusive on CT or MRI in recurrent AP patients, endoscopic US-guided fine needle

puncture biopsy can distinguish focal pancreatitis from pancreatic tumor, and color Doppler US can be used to assess vascular complications such as false arterial aneurysms or portal vein thrombosis^[21]. Xu reported that EUS outperformed CT in accurately categorizing symptomatic peripancreatic fluid collections and emerged as a preferred imaging modality for detecting solid necrotic debris^[91]. EUS-guided lumen-apposing metal stents for pancreatic fluid collections were feasible, effective with preferable technical and clinical success rates^[92].

In a retrospective analysis with a cohort of 6069 patient, Froes *et al*^[93] evaluated the impact of abdominal ultrasound (AUS) on the length of service (LOS) for patients hospitalized for AP who lacked radiographic evidence of AP on CT of the abdomen and pelvis (CTAP). Additionally, they further assessed how AUS affected the probability of subsequent interventions, such as ERCP or cholecystectomy^[93]. In patients with AP, undergoing AUS within 48 h resulted in a reduced LOS by 1.099 d; those who underwent AUS were 1.126 times more likely to proceed with subsequent ERCP compared to those who only had CTAP; patients receiving AUS after CTAP had a 2.711 times higher likelihood of undergoing subsequent cholecystectomy^[93]. In this cohort of patients admitted for AP, conducting AUS within 48 h after negative CTAP correlated with reduced LOS; moreover, patients undergoing AUS were not only more inclined to undergo ERCP but also exhibited a higher likelihood of undergoing cholecystectomy^[93].

In a study with a total of 196 patients, Cai *et al*^[94] investigate the diagnostic accuracy of US and contrast-enhanced US (CEUS) for AP. They demonstrated that CEUS outperformed US in diagnosing AP and SAP and produced excellent results in the staging of AP severity^[94]. In this study, compared to results from CECT, the diagnostic rates for pancreatic swelling using US and CEUS were 121% (148/122) and 91% (111/122), respectively, while for peripancreatic fluid collection, they were 84.8% (151/178) and 96.6% (172/178), respectively^[94]. The findings confirmed that CEUS surpassed US in specificity when visualizing pancreatic parenchyma edema, pancreatic border-capsula, collection fluid of peripancreas, and peripancreatic necrosis; this discrepancy between US and CEUS might arise from the ability of CEUS to visualize

vessels upon contrast agent injection^[94]. The conclusion drawn was that CEUS serves as a trustworthy method for diagnosing and monitoring AP and SAP, potentially acting as an alternative to CECT^[94].

Summary

The application of imaging in patients with AP is an essential aspect of modern clinical management. While there are challenges associated with their use, continuous research, technological advances, and thoughtful implementation of guidelines can optimize their role in patient management.

Imaging technologies for diagnosing and managing AP have made great strides, but inappropriate imaging tests can increase economic costs to the health system, subject patients to excess radiation, and elevate complication rates without benefiting patients. The choice of appropriate imaging modality for AP depends exactly on available time, technique, and clinical situation of the patient. Although imaging examination is widely used and carefully evaluated during the diagnosis process of AP, it remains unclear when imaging should be performed, especially given the economic costs associated with imaging and the financial burden on patients. In terms of the economic and financial implications of diagnostic imaging for AP patients, early imaging may not be advisable for those presenting with characteristic clinical symptoms and pronounced laboratory results. However, when clinical manifestations are unclear, early imaging examination is often used to identify suspected AP, discover potential etiology, diagnose complications, assess severity, implement risk stratification, and guide treatment. For AP patients, imaging technologies remain pivotal in initial diagnosis, identification of severe cases, assessment of prognosis, and decision of therapeutic management.

Radiomics is a data science technique that extracts a large number of quantitative features from medical images using advanced algorithms. These features capture subtle differences in the texture, shape, and intensity of image regions, which may be difficult for human observers to discern. By extracting these features, radiomics can transform

images into high-dimensional data that can be analyzed and mined using machine learning and other data science techniques. This allows for more objective and precise diagnosis, treatment planning, and prognosis evaluation in AP. Therefore, radiomics has the potential to revolutionize medical imaging and improve patient outcomes in the 21st century.

CLINICAL SCORING SYSTEMS

Over the decades, many clinical scoring systems have been developed and applied, and their efficacy and accuracy have been compared. Clinically, an ideal scoring system should be responsive, simple and reliable, and universally applicable across diverse patient populations and geographical areas, maintaining its relevance over time. Such clinical scoring systems are imperative to predict complications, severity, mortality, and ICU admission requirement in AP patients^[95]. As we all know, numerous “traditional” multifactorial clinical scoring systems, such as APACHE II, Ranson, Glasgow, Systemic Inflammatory Response Syndrome (SIRS), HAPS, Japanese Severity Score (JSS), CTSI, Sequential Organ Failure Assessment (SOFA) and BISAP, provide insights into systemic complications to some extent and possess commendable predictive capabilities for severity and mortality of disease^[12,82,96-98]. Based on the 2012 RAC, these scoring systems primarily stratified the severity of AP into MAP, MSAP, and SAP^[99]. Development of the original APACHE severity-of-illness classification system began in 1978, and APACHE II was derived from the results of a simplified effort based on the 12 most commonly used physiological measures included in the original APACHE system^[100]. APACHE II, initially designed for intensive care applications, necessitates the aggregation of numerous ¹³ parameters, some of which might not be pertinent to the prognosis in AP, while it overlooks key indicators such as pancreatic injury and significant regional complications^[101,102]. Ranson was first reported to be used to assess the severity of AP in 1974, and has been used for nearly 50 years^[103]. Ranson is relatively accurate in classifying the severity of AP patients; however, its limitation is the 48-h duration required for completion, thereby missing a crucial early therapeutic

opportunity^[102]. The main limitation of Glasgow, much like Ranson's, is the need for a 48-h duration to finalize the calculation^[96]. However, based on the local characteristics of CT examination, CTSI mainly emphasizes local complications but falls short in representing the systemic inflammatory response^[42]. In addition, for SAP, MCTSI demonstrates prognostic value for short-term mortality, while CTSI effectively predicts the necessity for intervention^[104]. SOFA, similar to APACHE II, is a detailed scoring system that takes into account acute and chronic illness, signs, and laboratory values in patients^[12].

Comparison of different scoring systems used in AP

For a more thorough understanding of the various attributes inherent in distinct scoring systems, we will embark on a comprehensive discussion and detailed analysis of the utilization of commonly employed clinical scoring systems within the context of AP in the following sections.

In two independent, prospectively enrolled cohorts [training ($n = 256$) and validation ($n = 397$)] of AP patients, Mounzer *et al*^[105] compared the accuracy of the scoring systems including APACHE II, BISAP, Glasgow, HAPS, JSS, Ranson and SIRS in predicting persistent organ failure. In this study, they discovered that these scoring systems exhibited moderate accuracy, with area under the curve (AUC) at admission ranging from 0.62-0.84 in the training cohort and 0.57-0.74 in the validation cohort; notably, Glasgow emerged as the superior classifier at admission in both cohorts^[105]. In a retrospective study including 161 patients, statistically significant cutoff values in predicting SAP were APACHE II ≥ 8 , Ranson ≥ 3 , BISAP ≥ 2 , CTSI ≥ 3 , and CRP₂₄ ≥ 21.4 mg/dL, and APACHE II had the highest accuracy in predicting SAP^[106]. Confusingly, different studies have shown that these scoring systems vary widely in accuracy, sensitivity, and specificity for the desired purpose of prediction, as follows. In a retrospective study including 326 patients diagnosed with hyperlipidemic AP (HLAP), the predictive abilities of APACHE II, BISAP, Ranson and MCTSI were compared for assessing MSAP and SAP, local complications, and HLAP mortality^[107]. The results

showed that the four scoring systems have their own advantages and characteristics. For example, Ranson lacked a distinct advantage in predicting severity and prognosis of HLAP compared to other three scoring systems; APACHE II excelled in predicting HLAP severity but fell short in predicting local complications; MCTSI demonstrated exceptional prowess in predicting local complications, yet was less adept in predicting severity and mortality; BISAP offered a commendable accuracy in evaluating the severity, local complications, and mortality of HLAP, yet there remains room for refining its precision in future assessments^[107].

In a prospective study including 50 AP patients, Harshit Kumar and Singh Griwan^[108] assessed the accuracy of APACHE II, BISAP, Ranson and MCTSI in predicting the severity of AP, referencing the 2012 RAC. In this study, MCTSI demonstrated the highest AUC values for predicting SAP (0.919), pancreatic necrosis (0.993), organ failure (0.893), and ICU admission (0.993); meanwhile, APACHE II ranked second in accuracy for predicting SAP (0.834) and organ failure (0.831)^[108]. The findings indicated that APACHE II demonstrated a high sensitivity in predicting pancreatic necrosis (93.33%), organ failure (92.86%) and ICU admission (92.31%), while also maintaining a substantial negative predictive value (NPV) for predicting pancreatic necrosis (96.15%), organ failure (96.15%) and ICU admission (95.83%)^[108]. Biberici Keskin *et al*^[109] retrospectively investigated 690 patients who had been admitted due to AP by five scoring systems including HAPS, Ranson, BISAP, Glasgow, and JSS. In this study, NPV of each score was notably superior to their respective positive predictive value (PPV)^[109]. Of the five scoring systems, JSS exhibited the highest value of AUC across all end-points (0.80 for in-hospital major adverse events, 0.94 for in-hospital mortality, 0.91 for 30-d mortality); nevertheless, none of the five scoring systems effectively predicted 30-d readmission^[109]. Li *et al*^[110] conducted a retrospective assessment of four scoring systems (Ranson, BISAP, Glasgow, and APACHE II) to predict AP outcomes in 918 patients, categorizing them into two age groups: the elderly (≥ 60 years old) and the younger (< 60 years old). In this study, they drew several following conclusions: BISAP effectively predicted the severity, pancreatic necrosis, and mortality in elderly AP

patients; APACHE II was more suitable for assessing severity in younger patients; both Ranson and Glasgow were generally applicable for evaluating most AP patients; however, Ranson demonstrated heightened efficacy in assessing severity among younger patients^[110]. In this study, the criterion of predicting SAP was different between the elderly and the younger (the elderly: Ranson ≥ 4 , Glasgow ≥ 3 , APACHE II ≥ 9 , BISAP ≥ 3 ; the younger: Ranson ≥ 3 , Glasgow ≥ 2 , APACHE II ≥ 8 , BISAP ≥ 2), suggesting that the scoring cutoffs for the elderly were consistently one point higher than those for the younger^[110]. The variation in the cutoff value for predicting SAP enhanced the specificity of the four scoring systems, albeit with a marginal reduction in their sensitivity to SAP^[110]. In a retrospective analysis including 653 AP patients, Teng *et al*^[111] investigated and compared the characteristics of six scores in predicting SAP, ICU admission, and mortality, including Ranson, Glasgow, APACHE II, BISAP, HAPS, and SOFA. In predicting SAP, SOFA exhibited the lowest sensitivity at 13.6% but boasted the highest specificity at 99.7%; conversely, Ranson maintained the highest sensitivity at 92.6% but had one of the lowest specificities at 51.9%, with only HAPS registering a slightly lower specificity at 49.7%^[111]. In predicting ICU admission, APACHE II and Ranson displayed a sensitivity at 100.0%, BISAP demonstrated the lowest sensitivity at 25.0% and a specificity at 93.4%, and SOFA demonstrated the highest specificity at 99.2%^[111]. In predicting mortality, APACHE II and Ranson displayed a sensitivity at 100.0%, BISAP showcased the lowest sensitivity at 25.0%, and SOFA had the highest specificity at 98.9%, similar to ICU admission^[111]. All scores had high and comparable NPVs in the prediction of SAP, ICU admission, and mortality in AP patients^[111]. In this study, they concluded that SOFA and 48-h Ranson outperformed other clinical scorings (Glasgow, APACHE II, BISAP, HAPS) in predicting severity, ICU admission, and mortality^[111]. In a prospective observational study including 164 patients, Venkatesh *et al*^[112] reported that, based on ROC curves, Ranson at admission demonstrated superior diagnostic accuracy in predicting severity, organ failure and mortality, and outperformed the other three scores (APACHE II, BISAP, and modified Glasgow) in predicting AP severity. In addition, this study revealed that while BISAP might be

calculated within 24 h of admission, both APACHE II and modified Glasgow demonstrated superior diagnostic accuracy, with APACHE II exhibiting the strongest association with mortality in SAP patients^[112]. Asfuroğlu Kalkan *et al*^[113] retrospectively analyzed 1150 AP patients, and reported that these scoring systems including BISAP, Ranson, HAPS, APACHE II and Glasgow were capable of predicting mortality. However, APACHE II predicted mortality with a sensitivity of 90% and specificity of 92%^[113].

Drawing on the insights gleaned from the aforementioned body of literature, we have meticulously synthesized a detailed appraisal of the application of various clinical scoring systems in prognosticating severity, local complications, organ failure, and mortality rates associated with AP. These summarizations are comprehensively depicted in Table 1. Various scoring systems exhibited diverse levels of sensitivity, specificity, and accuracy in forecasting the severity, local complications, organ failure, and associated mortality. Further, it is noteworthy that numerous studies have indicated the existence of substantial differences among these scoring systems, highlighting their lack of uniform standards and, in some instances, a concerning degree of inconsistency in their projections. Given the variability in accuracy among diverse scoring systems for predicting the severity, local complications, organ failure, and mortality associated with AP, there is a plausible need for further refinement and design optimization of each scoring system to enhance the precision of these predictions. Moreover, another potential area of research could be the amalgamation of multiple existing scoring systems to boost the predictive accuracy for AP through a more comprehensive scoring approach.

To enhance a more comprehensive understanding of the clinical utility of prevalent scoring systems, such as BISAP, SOFA, and qSOFA, in predicting AP outcomes, we conducted independent discussions and analyses on the latest advancements of these tools to provide invaluable reference and guidance for their practical application in clinical settings.

BISAP

In a large population-based study, Wu *et al*^[101] identified five variables for prediction of in-hospital mortality by Classification and Regression Tree analysis to derive a prognostic scoring system (BISAP) including blood urea nitrogen (BUN) (> 25 mg/dL), age (> 60 years), SIRS, pleural effusion, and impaired mental status. BUN emerged as the most efficient primary discriminative variable, age and SIRS further distinguished between high- and low-risk cases, and mental status and pleural effusion further refined the categorization of intermediate-risk patients^[101]. Introduced in 2008, BISAP, with its advantages of simplicity and precision, had been employed for the early identification of AP patients with an elevated risk of in-hospital mortality^[101]. BISAP is adept at identifying AP patients at heightened risk of mortality, representing the advancement of intermediate markers of severity within 24 h of onset, and its risk stratification ability could hold potential for enhancing clinical care and streamlining enrollment in clinical trials^[114]. BISAP was considered to be as good as APACHE II in predicting severity, death, and especially organ failure in AP, and outperformed Ranson, CTSI, CRP, hematocrit (HCT) and body mass index (BMI), with a score of 2 being a statistically significant cutoff value^[115]. BISAP is a streamlined scoring system designed to predict the severity of AP, and is instrumental in early risk stratification of AP. A prospective study of 87 patients experiencing their first episode of AP revealed that BISAP (≥ 2) demonstrated comparability to both APACHE II (≥ 8) and MCTSI (≥ 8) in metrics of accuracy, sensitivity, specificity, and NPV^[116]. In a systematic review and meta-analysis, a pooled analysis from 12 prospective cohorts showcased the exemplary performance of BISAP in predicting SAP across diverse patient populations and disease severity^[117]. Furthermore, the performance of BISAP was notably superior when severe pancreatitis was characterized by the persistence of organ failure for 48 h or more^[117]. A European cohort study indicated that BISAP effectively predicted SAP, mortality, and ICU admission, making it invaluable for triaging patients toward ICU care^[118].

Chen *et al*^[119] assessed the accuracy of BISAP in predicting the severity and prognoses of AP in Chinese patients. In this study, they retrospectively analyzed clinical data from

497 AP patients, comparing BISAP with APACHE II, Ranson, and CTSI regarding their predictive capacities for the severity of AP and the occurrence of mortality, pancreatic necrosis and organ failure in SAP patients^[119]. They highlighted that BISAP outperformed traditional scoring systems in terms of simplicity and speed, and maintained a performance comparable to other scoring systems in predicting both SAP and its associated prognoses^[119]. Zhang *et al*^[120] evaluated the efficacy of BISAP, APACHE II and Ranson in predicting the severity, mortality and pancreatic necrosis of AP based on the 2012 RAC at a tertiary care center in China. From their study involving 155 patients, they determined that BISAP might serve as a reliable tool for risk stratification and prognostic assessment in Chinese AP patients^[120]. Gao *et al*^[121] conducted a meta-analysis to systematically assess the accuracy of BISAP in predicting mortality and SAP, and affirmed that BISAP served as a dependable tool for identifying AP patients at elevated risk for adverse outcomes. While BISAP demonstrated superior specificity compared to Ranson and APACHE II, it exhibited a slightly diminished sensitivity for both mortality and SAP^[121]. An Indian study with 119 AP patients showed that BISAP was an accurate means of risk stratification, and patients with BISAP ≥ 4 invariably developed SAP or pancreatic necrosis and had high mortality^[122]. The available studies collectively demonstrated that BISAP performs very well in predicting SAP, and the simplicity and accuracy of the calculation make BISAP a valuable tool for clinical care of AP patients. Additionally, before confidently advocating for the adoption of BISAP, its integration into clinical practice should be evaluated to determine its potential to enhance outcomes in AP.

BISAP, an easily computed clinical prediction scale, leverages data from initial assessment of patients and routine laboratory results, demonstrating excellent performance in predicting SAP. Of particular concern is that BISAP is less cumbersome to calculate, more economical at cost, which makes it an ideal scoring system. It is considered that BISAP should be popularized at primary and secondary care institutions for severity classification and risk stratification of early AP, so that SAP

patients can be referred to higher-level medical centers for more reasonable clinical intervention.

However, as more and more clinical studies have been conducted, BISAP has shown inconsistent predictive power and results in predicting SAP, as reported in the next studies. A prospective study included 51 patients showed that BISAP was inferior to APACHE II in predicting the severity of AP, especially for SAP^[123]. In a meta-analysis including 1972 subjects, Yang and Li evaluated the diagnostic performance of BISAP in predicting SAP^[124]. They concluded that, despite its high specificity, BISAP was not the optimal standalone method for assessing AP severity due to its low sensitivity^[124]. In a prospective study including 50 AP patients, the accuracy of BISAP in predicting SAP stand at 84%, surpassing that of serum PCT (≥ 3.29 ng/mL) at 76%, which was on par with APACHE II; moreover, in logistic regression analysis, BISAP demonstrated greater statistical significance than serum PCT^[125]. They determined that BISAP outperformed serum PCT, APACHE II, Glasgow, and BCTSI in accurately predicting AP severity, positioning it as a promising tool for gauging the clinical progression of AP^[125]. Hagjer *et al*^[126] evaluated the usefulness of BISAP and PCT for AP prediction in a prospective observational study including 60 patients. Based on this study, in predicting severity, mortality, and organ failure, they finally concluded that BISAP was as effective as APACHE II and surpassed Ranson, CTSI, CRP, HCT and BMI in evaluating AP patients, and PCT was a good independent prognostic marker and was comparable with BISAP and APACHE II in accuracy^[126].

A multicenter validation study is essential to corroborate these findings and further elucidate the role of BISAP in AP. Meanwhile, further well-designed prospective studies are warranted to investigate the conditions under which BISAP can be used to more accurately, sensitively, specifically assess severity and prognosis in AP.

Combination of BISAP and other diagnostic indicators

In a retrospective analysis including 114 cases, the severity and mortality of AP escalated with the increase of BISAP, and BISAP exhibited a positive correlation with

CRP, D-dimer and serum glucose and negatively correlated with serum Ca^{2+} ^[127]. Based on the positive correlation between CRP and APACHE-II, Ranson, BISAP and CTSI, when CRP was included into BISAP, the AUC of predicting SAP and death were 0.873 and 0.909, respectively, showing that the combination of BISAP and CRP had better predictive value for severity and death of AP^[127]. In a study including 117 SAP patients, Wu *et al*^[128] reported that combining BISAP with miR-155 yielded a superior AUC compared to individual predictions, suggesting that this combination could enhance the clinical predictive accuracy for AP severity.

Early diagnosis and timely assessment of the severity are critical because early aggressive treatment reduces morbidity and mortality of AP. However, an ideal multifactor scoring system for early assessment of AP severity has not been determined. Based on an analysis of the available data and evidence, we recommend that BISAP as a multifactor scoring system combined with characteristic biochemical markers present at 48 h, in order to achieve optimal early assessment of AP severity.

SOFA

In October 1994, the ESICM convened a consensus meeting in Paris to establish the SOFA score, aiming to quantitatively and objectively describe the degree of organ dysfunction/failure over time in patient groups or even in individual cases^[129]. Although SOFA is primarily designed for patients with sepsis, it was deemed necessary to expand its application beyond this specific patient group^[129]. At present, SOFA is widely utilized in ICU to evaluate, prognosticate and assess patients; since its validation, it has been applied in diverse medical settings, including trauma, surgical, cardiac, and neurological ICUs^[130]. Minne *et al*^[131] conducted a systematic review on the utility of SOFA-based models for predicting the risk of mortality in ICU patients, and recommended an integration of a traditional model derived from data within the initial 24 h post-ICU admission with sequential SOFA. SOFA could be easily integrated into contemporary cardiac ICU through an electronic algorithm, and the day 1 SOFA demonstrated strong predictive capability for short-term mortality among a broad

spectrum of patients in cardiac ICU^[132]. Among the critical care systems, SOFA has distinct benefits, including its simplicity in computation, incorporation of therapeutic needs, and facilitating comparisons of AP with other critical care diseases^[24].

Adam *et al*^[133] retrospectively evaluated the efficacy of APACHE II, SOFA and modified Ranson in predicting mortality among 43 SAP patients as well as other factors influencing mortality in patients admitted to ICU, and concluded that SOFA was superior to Ranson and APACHE II in determining prognosis. In this study, SOFA had a significant correlation with mortality, and all patients with SOFA ≥ 11 at any point during ICU stay exhibited a heightened mortality risk, with a sensitivity of 80% and a specificity of 79%^[133]. Tee *et al*^[134] retrospectively obtained serial measurements of Ranson, APACHE II, and SOFA in 159 patients with SAP, assessing the efficacy of serial measurement using these three scoring systems. In this study, besides acquiring Ranson and APACHE II on admission and at 48 h, they took serial weekly measurements of SOFA, including data from admission, 48 h, as well as days 7, 14, and 21^[134]. The three scoring systems reliably predicted both overall and ICU mortality; however, the SOFA on day 7 exhibited the largest AUC, with any increase or lack of change in SOFA on day 7 of hospitalization correlating with elevated mortality^[134]. They concluded that both APACHE II and SOFA were sensitive in predicting mortality for AP, serial SOFA proved reliable for guiding clinical decisions and day 7 of hospitalization was a reasonable time for SOFA reassessment to predict late mortality in SAP^[134]. A retrospective study enrolling 146 AP patients demonstrated that an increase in SOFA independently heightened the likelihood of adverse outcome during hospitalization for AP patients, and SOFA > 5 was highly predictive of in-hospital mortality compared to other scores^[135]. Utilizing a straightforward tool like SOFA, validated in intensive care settings, could enhance the stratification of in-hospital mortality risk and clinical deterioration among AP patients admitted to medical wards. Teng *et al*^[111] reported that both SOFA and 48-h Ranson effectively predicted the severity, ICU admission, and mortality associated with AP, with SOFA showing particularly favorable results.

qSOFA

46

The qSOFA includes respiratory rate (breaths per minute), systolic blood pressure (mm Hg), and Glasgow Coma Scale score^[136]. In a 17-year observation study including 1059 patients, the ROC curve analysis revealed that the AUC values of APACHE II, SOFA, qSOFA scores in predicting the prognosis of infected patients stood at 0.713, 0.744 and 0.662, respectively^[137]. In this study, Qin *et al*^[137] posited that qSOFA, due to its advantages of rapid acquisition, was anticipated to serve as an efficient tool for assessing the prognosis of ICU patients with infections. Given its extraordinary simplicity, qSOFA would be an appropriate score particularly for the initial patient evaluation in the emergency, and was considered to be a rapidly available prognostic score in AP with limited prognostic validity^[138]. In a cohort study including 203 patients, Rasch *et al*^[138] reported that qSOFA could predict ICU admission and multiple organ dysfunction syndrome in AP. In a retrospective cohort study involving 161 patients with the diagnosis of alcohol-induced AP, a qSOFA score of 2 or higher both upon and 48 h post-admission exhibited a specificity of 94% or greater and sensitivity of 33% or higher for assessing pancreatitis severity and determining the necessity for intensive care admission, intubation, or vasopressor^[139]. In a three-year cohort study from USA, Hallac *et al*^[140] evaluated the ability of qSOFA and SIRS in predicting extended hospital stays among patients presenting with AP to the emergency department and hospital ward. A qSOFA of 2 or higher was linked to a diagnosis of significant AP with a specificity of 99% and a sensitivity of 4%; in contrast, a SIRS score of 2 displayed a specificity of 61% and a sensitivity of 80% in detecting patients with significant AP^[140]. Based on their findings, they inferred that relying solely on qSOFA for triaging AP patients could lead to under recognition and potential undertreatment^[140].

HAPS

HAPS was calculated rapidly from the following three parameters: presence or absence of rebound tenderness or guarding, hematocrit (> 43 mg/dL for men or > 39.6 mg/dL

25

for women) and serum creatinine (> 2 mg/dL)^[105,141]. Oskarsson *et al*^[142] reported that HAPS predicting nonsevere AP progression had a specificity of 96.3% and a corresponding PPV of 98.7% in 531 patients experiencing either a first-time or a recurrent attack of AP, emphasizing HAPS as a highly specific scoring algorithm predicting nonsevere AP progression^[142]. In a prospective pilot study with 103 AP patients from Indian, the sensitivity, specificity, PPV, NPV and AUC of HAPS as a predictor of non-severe disease were 76.3%, 85.7%, 93.8%, 56.6% and 0.848, respectively^[143]. In a study including 703 AP patients from China, the sensitivity, specificity, PPV, NPV and AUC of HAPS on admission for predicting MAP was 48.2%, 97.7%, 95.6%, 64.1% and 0.749, respectively^[97]. These aforementioned two studies validated the utility of HAPS at admission in predicting non-severe AP in India and MAP in China, respectively. Maisonneuve *et al*^[144] evaluated PPV of HAPS by performing a meta-analysis of 20 reports covering 6374 patients. They concluded that HAPS accurately identified patients with nonsevere AP who wouldn't necessitate ICU care, enabling the pinpointing of patients suitable for brief general ward stays or home-based care^[144].

HAPS may offer significant advantages in the triage of AP patients when compared to other scoring systems, underscoring its potential utility in optimizing patient classification and guiding treatment strategies. In a study including 60 patients with the first attack of AP, Gupta *et al*^[145] reported that the sensitivity, specificity, PPV, NPV and AUC of HAPS predicting SAP were 90.91%, 59.81%, 33.33%, 96.67% and 0.75, respectively. The high NPV indicated that HAPS could very accurately identify, within the first hour of admission, patients who had a mild course of disease, did not require intensive management and were not at risk of dying from the disease^[145]. Based on this result, they argued that the patient typically tended to experience a milder course of illness if the evaluation of HAPS yielded a negative result^[145]. Conversely, in instances where the score was positive, the patient's clinical progression could unfold in any direction, demonstrating the uncertainty associated with such an outcome^[145]. In this same study, the sensitivity, specificity, PPV, NPV and AUC of BISAP in the prediction

of SAP were 63.64%, 100%, 100%, 92.45% and 0.82, respectively^[145]. In comparison to BISAP in this study, HAPS demonstrated a heightened sensitivity towards processes predicting mortality and severity, and played a pivotal role in determining whether patients necessitated costly imaging procedures, thereby potentially enabling significant hospital cost savings^[145]. In a study with 116 patients, Al-Qahtani *et al*^[146] compared HAPS with Ranson in predicting the severity of AP, and concluded that HAPS was effective in rapidly identifying patient likely to experience a non-severe course of the disease.

Of significant importance is the fact that assessment of HAPS can be accomplished within the first hour of a patient visit, offering a distinct advantage in terms of time efficiency. In contrast, while Ranson might offer superior accuracy, it necessitates a full 48 h to reach completion, highlighting a potential trade-off between speed and precision in these scoring systems. Considering that the substantial majority of individuals diagnosed with AP typically exhibit a milder form of the disease, the capacity to accurately distinguish these patients of MAP is of utmost significance. Drawing upon the aforementioned analysis and discussion, HAPS appears to be a commendable choice for assisting physicians in evaluating the severity of AP. Furthermore, HAPS could potentially be perceived as a gold standard for facilitating both the early identification and cost-effective management of this disease. In addition, due to the readily accessible parameters required for its computation, HAPS can be effectively utilized in a wide range of healthcare facilities, including those located in developing countries. This ease of implementation makes HAPS an inclusive and practical tool for global health contexts.

Other recent clinical scoring systems

Hong *et al*^[147] developed a prognostic score termed SABP, encompassing systemic inflammatory response syndrome, serum albumin, blood urea nitrogen and pleural effusion. The SABP score could serve as an instrumental tool to categorize patients at risk of developing SAP as per the latest revised Atlanta criteria and its application on

admission may enhance clinical care and refine management approaches for AP^[147]. He *et al*^[148] retrospectively analyzed a clinical data of 469 patients with AP, and selected seven prognostic indicators to establish ²unweighted predictive score (unwScore) and weighted predictive score (wScore) for MSAP and SAP. The early multi-indicator prediction models for MSAP and SAP demonstrated robust predictive efficacy, offering a meaningful clinical benchmark for diagnosis and treatment^[148].

In a retrospective analysis encompassing a total of 1,295 AP patients, Feng *et al*^[149] developed an independent predictive tool, known as a nomogram, to predict the likelihood of sepsis occurrence in this patient population. In this study, the predictive ¹performance and clinical utility of the newly established nomogram surpassed those of other scoring systems such as SIRS, BISAP, SOFA, and qSOFA^[149]. The innovative risk-prediction system could precisely estimate the likelihood of sepsis in AP patients, assisting clinicians in formulating personalized treatment strategies for the patients; by doing so, it not only alleviated the disease burden of the patients but also facilitated the reasonable distribution of medical resources, which was a crucial aspect of tertiary prevention^[149]. The nomogram incorporated all the independent prognostic factors, including body temperature, phosphate, Ca²⁺, sodium (Na⁺), lactate, albumin, platelet count, urinary output, ¹mean blood pressure (MBP), Glasgow Coma Scale (GCS), and ¹Charlson Comorbidity Index (CCI)^[149]. These diverse elements collectively contributed to its predictive strength.

Summary

Score systems, utilizing 4-25 factors, have been developed to predict severity, yet they frequently rely on multiple parameters not measurable daily and often necessitate over 24 h to finalize, leading to critical time loss^[150]. While these scores can predict failure or severity of specific organs, their reliance on dichotomous parameters leads to information loss, limiting their practical application in clinical settings^[150]. Based on the current literatures, here are the identified problems and potential solutions for applying

clinical scoring systems to the diagnosis, severity prediction, and prognosis assessment of AP.

Inconsistency: Different scoring systems like Ranson, Glasgow, BISAP, and CTSI may yield inconsistent results, leading to confusion in clinical decision-making.

Solution: Research to validate and compare different scoring systems can help identify the most accurate and reliable ones. Standardizing the use of a particular scoring system across healthcare settings can reduce inconsistency.

Complexity: Some scoring systems are complex and require multiple parameters, making them time-consuming to calculate. This complexity can hinder their practical application in urgent care settings.

Solution: Creating simplified and user-friendly scoring systems that maintain accuracy can make them more practical for clinicians to use, especially in urgent care settings.

Lack of sensitivity and specificity: Some scoring systems may lack sensitivity or specificity in predicting the severity and prognosis of AP, leading to inaccurate assessments.

Solution: Combining scoring systems with comprehensive clinical assessment can lead to more accurate care. This solution is more of a clinical recommendation rather than a documented research finding.

Lack of personalization: Scoring systems are often based on population-level data and may not account for characteristics of individual patient, leading to generalized predictions that may not be applicable to all patients.

Solution: Considering patient-specific factors, such as comorbidities, lifestyle, and preferences, in conjunction with scoring systems, can lead to more personalized and effective care.

Over-reliance on scoring systems: Sole reliance on scoring systems without considering clinical judgment and other patient-specific factors may lead to suboptimal care.

Solution: Providing education and training to healthcare professionals on how to effectively use scoring systems, including their limitations, can enhance their application in clinical practice.

In conclusion, while clinical scoring systems are valuable tools in managing AP, they present challenges that are recognized both in clinical practice and in the research literature. The solutions outlined above, grounded in current research and clinical wisdom, can enhance the effectiveness of these systems in providing accurate and personalized treatment for patients with AP.

AI

In the era of AI, machine learning algorithms have been devised to accurately predict the severity, complications, recurrence, mortality, and even the optimal timing of surgery for AP patients. However, the quality of research evaluating the accuracy of AI is still low and lacks studies comparing AI with these commonly used clinical scores. Therefore, more research is needed before we can routinely use AI in our daily clinical practice. Prior to this, the easy-to-calculate and applicable scoring systems seems to be the most reasonable choice.

Recently, AI applications, utilizing machine learning, have been progressively integrated into the medical field, demonstrating superior performance in predicting complications compared to logistic regression analysis^[151]. AI-based machine learning is booming and creating a technological revolution, especially in the healthcare industry^[152]. Machine learning, a subset of AI, employs statistical methods to train algorithms for predictions, enabling a computer system to self-learn and enhance its performance based on experience^[150]. Machine learning has garnered significant attention and recognition from clinicians, driven by advancements in statistical theory and computer technology^[153]. Machine learning adeptly discerns intricate relationships

between diseases and variables, categorizes variables based on specific criteria, predicts outcomes from foundational features, and recognizes objects with analogous patterns^[152]. Innovative machine learning technologies have been extensively employed in predictive models for a spectrum of diseases, consistently demonstrating superior performance over traditional logistic regression or Cox regression analyses^[153].

In this age of technological advancement, AI stands as a pinnacle of innovation, proficiently discerning the intricate non-linear relationships between numerous biochemical parameters and their associated disease outcomes^[150]. For example, a retrospective study demonstrated that, when juxtaposed with the traditional logistic regression model, machine learning models (extreme gradient boosting and random forest) utilizing readily accessible features upon admission exhibited superior performance in predicting acute kidney injury (AKI) among AP patients^[151]. Leveraging such machine learning algorithms in predictive models could enable clinicians to foresee AKI at an early stage, potentially mitigating further renal damage^[151]. Based on an international cohort of 1,184 patients and a validation cohort of 3543 patients, Kui *et al*^[154] devised a user-friendly web application named EASY-APP, which employs multiple continuous variables accessible at admission. The EASY prediction score serves as an effective tool for pinpointing patients at elevated risk for severe AP within hours of hospitalization, and the web application was made available to clinicians, enhancing the utility and precision of the model^[154]. Zhou *et al*^[155] demonstrated that the extreme gradient boosting model (XGBoost) algorithm possesses the capability to precisely predict the severity of AP, offering clinicians valuable assistance in identifying severe AP at an early stage. In a prospective cohort study integrating necrosis prediction with AI, the XGBoost machine learning algorithm was employed to analyze the data from 2,387 AP patients^[156]. This model in the predictive capability rivals those existing clinical scoring systems and its performance is anticipated to enhance with continued use^[156]. In the United States, Thapa *et al*^[7] applied machine learning algorithms to predict which AP patients need SAP treatment, and developed three models using logistic regression, neural networks, and XGBoost. In this study, machine

learning models were trained and tested to utilize data from 61894 patients, with the XGBoost model surpassing the performance of both ¹ logistic regression and neural network-based models^[7]. Furthermore, the XGBoost model achieved a superior AUROC compared both HAPS and BISAP in identifying patients likely to be diagnosed with SAP^[7]. They concluded that machine learning has the potential to refine the precision of AP risk stratification methods, facilitating prompter treatment and intervention initiation^[7].

In a large retrospective study enrolling 5460 patients, Yuan *et al*^[157] ² developed and validated a novel machine learning tool, APCU, leveraging clinical, laboratory, and radiologic data to predict ICU admission among AP patients. They showed that the APCU effectively categorized ² AP patients into high-risk and low-risk groups, demonstrating a superior discriminative capability compared to other risk scores like Ranson, APACHE II, SIRS, NEWS in predicting ICU admission for AP patients and specific subgroups within 48 h of hospitalization^[157]. Notably, this study marked the inaugural application of a machine learning algorithm for the predictions of ² ICU admission in AP patients within 48 h of hospitalization, relying on widely accessible clinical, laboratory, and radiologic data^[157].

In a retrospective analysis involving 648 AP patients, Hong *et al*^[158] developed random forest (RF) and logistic regression (LR) models using a training sample; the RF model, notable for its interpretability, showcased the most superior discriminative performance in predicting SAP. In a retrospective study involving 631 AP patients, Luo *et al*^[159] developed a machine learning model, culminating in a nomogram designed for the early identification of SAP during the progression of AP. Their findings indicated that the RF model delivered optimal predictive performance, with the nomogram offering a visual scoring model suitable for clinical application^[159]. Such models have the potential to act as functional tools, enabling personalized treatment choices and enhancing clinical results by stratifying AP patients prior to treatment^[159]. In a study with a total of 1,012 patients, Yin *et al*^[160] ¹² developed a series of effective models for early prediction of SAP based on automated machine learning (AutoML) platform, and these

models outperformed the existing scoring systems, which might offer insights into AutoML applications in future medical studies. The AutoML model based on the GBM algorithm for early prediction of SAP showed evident clinical practicability^[160].

In a recent retrospective study involving a cohort of 460 AP patients, to predict ARDS in these patients at admission, Zhang and Pang^[161] constructed and optimized four machine learning models, including Support Vector Machine (SVM), Ensembles of Decision Trees (EDTs), Bayesian Classifier (BC) and nomogram models, based on 31 features with significant differences between the groups with and without ARDS. Among the four models, the BC algorithm exhibited superior predictive performance with the highest AUC (0.891), surpassing SVM (0.870), EDTs (0.813) and the nomogram (0.874) in the test set^[161]. Concurrently, the EDT algorithm achieved the highest accuracy at 0.891, precision at 0.800 and F1 score at 0.615, but registered the lowest FDR at 0.200 and the second-highest NPV at 0.902^[161]. In terms of predictive performance for ARDS as a complication of AP, they concluded that BC was the superior predictive model in the test set and EDTs exhibited promising potential for predicting large samples^[161].

Summary

The application of AI in the diagnosis, severity prediction, and prognosis assessment of AP represents an exciting development in the field of medicine. However, based on these current researches, we recognize several limitations and potential challenges that must be addressed to fully leverage the capabilities of AI in this context.

36 Data quality and availability: AI algorithms require high-quality, comprehensive, and diverse data to build robust and accurate models. In the context of AP, such data sets may not be readily available, especially for rare subtypes of the disease or patient populations with specific comorbidities. Furthermore, incomplete or inconsistent data can lead to biased or flawed results.

Interpretability: AI models, especially those employing complex algorithms like deep learning, often operate as 'black boxes', providing outputs without clear, understandable reasons for their decisions. This can limit their acceptance in the clinical setting, as healthcare professionals typically prefer to understand the reasoning behind a diagnosis or prediction.

Standardization: AI algorithms are typically designed and validated on specific datasets. Their generalizability to other populations or healthcare settings, especially those that are vastly different from the original context, is not guaranteed. This lack of standardization can lead to inconsistent results when the models are used in different settings.

Generalizability: Models trained on a specific set of data may not perform well when applied to different datasets, especially if there are demographic or geographical differences. For example, an AI model trained on data from a high-income country might not work as well in a low-income setting due to differences in healthcare infrastructure, disease prevalence, and patient characteristics.

Regulation: The use of patient data to develop and apply AI models raises significant concerns around data privacy, consent, and security. It's crucial that these concerns are addressed to ensure ethical usage and maintain public trust. For instance, who is responsible if an AI system makes an incorrect diagnosis or prognosis? How is patient data privacy ensured?

Implementation: The successful implementation of AI in healthcare settings requires clinicians to have a certain level of understanding and trust in the technology. This can be challenging due to varying levels of digital literacy among healthcare providers, and resistance to change.

Given these challenges, ongoing research is critical to improve the reliability, interpretability, and generalizability of AI tools in healthcare, and to address the ethical, legal, and workflow integration issues associated with their use. It's important that as we move forward, these tools are developed and used in a manner that complements the expertise of healthcare professionals, rather than seeking to replace it.

CONCLUSION

Early aggressive treatment of AP has been proven to reduce the incidence and mortality rates, therefore early diagnosis and severity assessment of AP are extremely necessary, and there is a particular need for ideal technological approaches to early evaluate and predict the progression of AP.

In recent years, there has been heightened interest in leveraging imaging technologies, scoring systems and AI to improve the diagnosis, severity prediction, and prognosis evaluation of AP. Different imaging modalities, such as CT, MRI, and US, are used to assess the severity and extent of pancreatic inflammation and detect any complications that may arise. Several ²scoring systems have been developed to assess the severity of AP and predict the risk of complications, such as Ranson, APACHE II, BISAP, SOFA and HAPS. These scoring systems take into account ⁴⁰various clinical and laboratory parameters, such as age, blood pressure, serum glucose, and white blood cell count, to provide a numerical score that reflects the severity of the disease. AI is a rapidly developing field that has the potential to revolutionize the diagnosis and management of AP. AI algorithms can be trained to analyze large datasets of imaging and clinical data to predict the severity and prognosis of AP. AI algorithms have been developed to analyze CT scans of patients with AP to predict the risk of complications such as pancreatic necrosis, abscess, or pseudocyst. The algorithms can detect subtle changes in the pancreas that may be missed by human radiologists and can provide more accurate and timely predictions of the risk of complications.

The integration of imaging technologies, scoring systems, and AI in the diagnosis, severity prediction, and prognosis assessment of AP has several advantages, including:

(1) More accurate diagnosis: Imaging technologies and AI algorithms can provide more accurate diagnoses, reducing the risk of misdiagnosis and unnecessary treatment; (2) Improved risk assessment: Scoring systems and AI algorithms can provide more accurate risk assessments, which can help healthcare providers make more informed treatment decisions; (3) Personalized treatment: The combination of imaging technologies, scoring systems, and AI can provide a more personalized approach to treatment, taking into account each patient's unique circumstances; and (4) Improved patient outcomes: The earlier and more accurate diagnosis, as well as the more personalized treatment options, can lead to improved patient outcomes and reduced healthcare costs.

Despite these advantages, there are several challenges that need to be addressed when integrating imaging technologies, scoring systems, and AI in the management of AP. These challenges include the need for standardized imaging protocols and scoring systems, the need for large datasets of imaging and clinical data to train AI algorithms, and ethical and legal challenges associated with the use of AI in healthcare. In conclusion, the integration of imaging technologies, scoring systems, and AI has the potential to revolutionize the diagnosis, severity prediction, and prognosis assessment of AP.

13%

SIMILARITY INDEX

PRIMARY SOURCES

- | | | |
|----------|--|---------------------------|
| 1 | www.researchgate.net
<small>Internet</small> | 497 words — 4% |
| <hr/> | | |
| 2 | www.ncbi.nlm.nih.gov
<small>Internet</small> | 213 words — 2% |
| <hr/> | | |
| 3 | www.scielo.br
<small>Internet</small> | 68 words — < 1% |
| <hr/> | | |
| 4 | atm.amegroups.com
<small>Internet</small> | 63 words — < 1% |
| <hr/> | | |
| 5 | www.revistas.usp.br
<small>Internet</small> | 56 words — < 1% |
| <hr/> | | |
| 6 | Bo Xiao, Hai-Bo Xu, Zhi-Qiong Jiang, Ju Zhang, Xiao-Ming Zhang. "Current concepts for the diagnosis of acute pancreatitis by multiparametric magnetic resonance imaging", Quantitative Imaging in Medicine and Surgery, 2019
<small>Crossref</small> | 52 words — < 1% |
| <hr/> | | |
| 7 | Subhash Chandra, Arvind Murali, Reena Bansal, Dipti Agarwal, Adrian Holm. "The Bedside Index for Severity in Acute Pancreatitis: a systematic review of prospective studies to determine predictive performance", Journal of Community Hospital Internal Medicine Perspectives, 2017 | 49 words — < 1% |

-
- 8 helda.helsinki.fi 48 words — < 1%
Internet
-
- 9 www.wjgnet.com 45 words — < 1%
Internet
-
- 10 ijchmr.com 41 words — < 1%
Internet
-
- 11 Julia Walkowska, Nicol Zielinska, R. Shane Tubbs, Michał Podgórski, Justyna Dłubek-Ruxer, Łukasz Olewnik. "Diagnosis and Treatment of Acute Pancreatitis", Diagnostics, 2022 38 words — < 1%
Crossref
-
- 12 www.frontiersin.org 36 words — < 1%
Internet
-
- 13 B U Wu, R S Johannes, X Sun, Y Tabak, D L Conwell, P A Banks. "The early prediction of mortality in acute pancreatitis: a large population-based study", Gut, 2008 30 words — < 1%
Crossref
-
- 14 Kai Siang Chan, Vishal G Shelat. "Diagnosis, severity stratification and management of adult acute pancreatitis—current evidence and controversies", World Journal of Gastrointestinal Surgery, 2022 29 words — < 1%
Crossref
-
- 15 Chiyoung Park, Hyehyeon Kim, Saehee Kim, Chulhee Kim. "Enzyme Responsive Nanocontainers with Cyclodextrin Gatekeepers and Synergistic Effects in Release of Guests", Journal of the American Chemical Society, 2009 27 words — < 1%
Crossref

-
- 16 bmcgastroenterol.biomedcentral.com 26 words — < 1%
Internet
-
- 17 Ting Zhou, Meng-yue Tang, Yan Deng, Jia-long Wu, Huan Sun, Yong Chen, Tian-wu Chen, Xiaoming Zhang. "MR Imaging for Early Extrapancreatic Necrosis in Acute Pancreatitis", Academic Radiology, 2019 22 words — < 1%
Crossref
-
- 18 C. S. Pitchumoni. "Chapter 55-1 Acute Pancreatitis", Springer Science and Business Media LLC, 2020 21 words — < 1%
Crossref
-
- 19 docksci.com 21 words — < 1%
Internet
-
- 20 encyclopedia.pub 21 words — < 1%
Internet
-
- 21 translational-medicine.biomedcentral.com 21 words — < 1%
Internet
-
- 22 Biswanath Sahu, Pooja Abbey, Rama Anand, Ashok Kumar, Shaili Tomer, Ekta Malik. "Severity assessment of acute pancreatitis using CT severity index and modified CT severity index: Correlation with clinical outcomes and severity grading as per the Revised Atlanta Classification", Indian Journal of Radiology and Imaging, 2021 20 words — < 1%
Crossref
-
- 23 Davor Štimac. "The Role of Nonenhanced Magnetic Resonance Imaging in the Early Assessment of Acute Pancreatitis", The American Journal of Gastroenterology, 5/2007 18 words — < 1%
Crossref

24 Thomas L. Bollen, Hjalmar C. van Santvoort, Marc G.H. Besselink, Wouter H. van Es, Hein G. Gooszen, Maarten S. van Leeuwen. "Update on Acute Pancreatitis: Ultrasound, Computed Tomography, and Magnetic Resonance Imaging Features", Seminars in Ultrasound, CT and MRI, 2007 18 words — < 1%
Crossref

25 link.springer.com 18 words — < 1%
Internet

26 Jia Zhang, Muhammad Shahbaz, Ruliang Fang, Benjia Liang et al. "Comparison of the BISAP scores for predicting the severity of acute pancreatitis in Chinese patients according to the latest Atlanta classification", Journal of Hepato-Biliary-Pancreatic Sciences, 2014 16 words — < 1%
Crossref

27 Ting Zhou, Yong Chen, Jia-Long Wu, Yan Deng, Ju Zhang, Huan Sun, Chuan Lan, Xiao-ming Zhang. "Extrapancreatic Inflammation on Magnetic Resonance Imaging for the Early Prediction of Acute Pancreatitis Severity", Pancreas, 2020 15 words — < 1%
Crossref

28 Yu-San Tee, Hsin-Yueh Fang, I.-Ming Kuo, Yann-Sheng Lin, Song-Fong Huang, Ming-Chin Yu. "Serial evaluation of the SOFA score is reliable for predicting mortality in acute severe pancreatitis", Medicine, 2018 15 words — < 1%
Crossref

29 Shanshan Yu, Yangyang Xiong, Jun Xu, Xianquan Liang, Yangyang Fu, Danyu Liu, Xuezhong Yu, Dong Wu. "Identification of Dysfunctional Gut Microbiota Through Rectal Swab in Patients with Different Severity of Acute Pancreatitis", Digestive Diseases and Sciences, 2020 14 words — < 1%
Crossref

-
- 30 Sumitra Hagjer, Nitesh Kumar. "Evaluation of the BISAP scoring system in prognostication of acute pancreatitis – A prospective observational study", *International Journal of Surgery*, 2018
Crossref 14 words — < 1%
-
- 31 doi.org
Internet 13 words — < 1%
-
- 32 edepositireland.ie
Internet 13 words — < 1%
-
- 33 "Evidence-Based Emergency Imaging", Springer Science and Business Media LLC, 2018
Crossref 12 words — < 1%
-
- 34 Georgios I Papachristou. "Comparison of BISAP, Ranson's, APACHE-II, and CTSI Scores in Predicting Organ Failure, Complications, and Mortality in Acute Pancreatitis", *The American Journal of Gastroenterology*, 10/27/2009
Crossref 12 words — < 1%
-
- 35 Ling-Ji Song, Bo Xiao. "Medical imaging for pancreatic diseases: Prediction of severe acute pancreatitis complicated with acute respiratory distress syndrome", *World Journal of Gastroenterology*, 2022
Crossref 12 words — < 1%
-
- 36 appmaster.io
Internet 12 words — < 1%
-
- 37 qims.amegroups.com
Internet 12 words — < 1%
-
- 38 Sreejith Vasudevan, Pooja Goswami, Ujjwal Sonika, Bhaskar Thakur, Vishnubhatla Sreenivas, 11 words — < 1%

Anoop Saraya. "Comparison of Various Scoring Systems and Biochemical Markers in Predicting the Outcome in Acute Pancreatitis", Pancreas, 2018

Crossref

-
- 39 worldwidescience.org 11 words — < 1%
Internet
-
- 40 www.science.gov 11 words — < 1%
Internet
-
- 41 "Poster Presentations", Journal of Gastroenterology and Hepatology, 2012. 10 words — < 1%
Crossref
-
- 42 Byung Geun Kim, Myung Hwan Noh, Choong Heon Ryu, Hwa Seong Nam et al. "A comparison of the BISAP score and serum procalcitonin for predicting the severity of acute pancreatitis", The Korean Journal of Internal Medicine, 2013 10 words — < 1%
Crossref
-
- 43 Piero Alberti, Elizabeth Pando, Rodrigo Mata, Laura Vidal et al. "Evaluation of the modified CT severity index (MCTSI) and CT severity index (CTSI) in predicting severity and clinical outcomes in acute pancreatitis", Journal of Digestive Diseases, 2020 10 words — < 1%
Crossref
-
- 44 Yi Wang, Kaixiang Liu, Xisheng Xie, Bin Song. "Potential role of imaging for assessing acute pancreatitis-induced acute kidney injury", The British Journal of Radiology, 2020 10 words — < 1%
Crossref
-
- 45 Ziqiang Zhao, Jie Shen, Deqing Zhang, Jiaqing Shen. "The prognostic role of peripheral 10 words — < 1%

lymphocyte subsets in patients with acute pancreatitis", The American Journal of the Medical Sciences, 2018

Crossref

46	publications.aap.org	10 words — < 1%
	Internet	
47	www.bmj.com	10 words — < 1%
	Internet	

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	< 10 WORDS
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	< 10 WORDS