Scoring Systems to Predict the Severity of Acute Pancreatitis - A Review

Prasanna G*, Thirunavukkarasu VS and Gokul DY

Department of General Surgery, Shri Sathya Sai Medical College & Research Institute, India

*Corresponding author: A Giri Prasanna, Department of General Surgery, Shri Sathya Sai Medical College & Research Institute, India, Email: giriprasanna.adapa@gmail.com

Abstract

Acute pancreatitis is a common and recurrent disorder, where inflammation of the pancreas with inconstant connection of other surrounding tissues or other organ systems. This condition is characterized by long-term pain in the abdomen area, frequent exacerbations of the disease, and insufficiency of the exocrine and/or endocrine. Among the individuals with pancreatitis, identification of patients who are at risk of severe disease and mortality is a crucial step for the purpose of effective management and prevention of mortality. Numerous scoring systems are present to help assess disease severity and prognosis.

Materials and methods: In this review, 22 original research articles and 1 meta-analysis study comparing the different types of scoring systems have been scrutinized and their outcome understood.

Results: While pitfalls were identified in various scoring systems, a generalised outcome from all the studies show that BISAP scoring has good specificity but inadequate sensitivity in predicting disease outcome. APACHE 2 has good predictive value in assessing SAP. EPIC scoring system was found to be a good predictor of organ damage. Modified CTSI also has good predictive value for SAP.

Conclusion: While the results show that APACHE 2, BISAP scoring, Modified CTSI and Ranson’s scoring are the most frequently used and have the best specificity and sensitivity among the various scoring systems, it is of our opinion that similar large studies need be conducted in a south Indian population, in order to better understand the implications of the change in race and food habits of the study population and usefulness of the same scoring systems in ascertaining the prognosis of pancreatitis when compared to the rest of the world.

Keywords: Pancreatitis; Scoring system; APACHE 2; BISAP; Modified CTSI; EPIC

Introduction

Acute pancreatitis (AP) is a common and recurrent disorder, where inflammation of the pancreas with inconstant connection of other surrounding tissues or other organ systems [1]. This condition is characterized by long-term pain in the abdomen area, frequent exacerbations of the disease, and insufficiency of the exocrine and/or endocrine [2]. Among the individuals with pancreatitis, identification of patients who are at risk of severe disease and mortality is a crucial step for the purpose of effective management and prevention of mortality. Numerous scoring systems are present to help assess disease severity and prognosis. However, it is still unclear as to which is the gold standard scoring system. It is the objective of this study to scrutinize available literature and identify an easy to use and reliable scoring system.
Scoring systems to assess the severity of diseases in acute pancreatitis

There are several scoring systems to assess the severity of AP. The few to be named are:

- Bedside index of severity in acute pancreatitis (BISAP)
- Acute physiology and chronic health evaluation - II (APACHE-II)
- Harmless acute pancreatitis score (HAPS)
- Multiple organ system score (MOSS)
- Ranson scores
- Glasgow scores
- Balthazar computer tomography severity index (BCTSI)
- Panc 3 Score
- Computed Tomography Severity Index (CTSI)
- Modified CTSI (MCTSI)
- MR severity index (MRSI)
- Extra-pancreatic inflammation on computed tomography (EPIC)
- Systemic inflammatory response syndrome (SIRS)
- Mixed antagonist response syndrome (MARS)
- Compensatory anti-inflammatory response syndrome (CARS)
- Modified Early Warning Score (MEWS)
- Renal Rim Score

Other biochemical parameters used:

- Procalcitonin (PCT)
- C-Reactive Protein (CRP)
- Interleukin-6 (IL-6)
- Serum Ca2+
- Red cell distribution width (RDW) (Haematological)

Discussion

As elaborated by Wu BU, et al. [3] in 2008 who conducted a cross-sectional observational study for the purpose of identification of factors associated with mortality among individuals who have acute pancreatitis. To predict the in-hospital mortality due to acute pancreatitis, a scoring system based on clinical profiles was developed named “Classification and Regression Tree (CART) analysis”. CART was developed from 2000-2001 on population derived from 212 hospital and validated in 2004-2005 in 177 hospitals by comparing it with APACHE II scoring system. The BISAP Area under the Curve 0.82 and for APACHE II it was 0.83 among the cohort observed for this study. The study concluded that the BISAP scoring system had good predictive validity and is simple to use and good diagnostic accuracy and can be used as a tool to measure the severe AP.

Mention must be of Papachristou GI, et al. [4] who conducted a prospective study in 2010 to compare BISAP and SIRS scoring systems with the available “traditional” systems i.e. Ranson’s system, APACHE-II, CTSI in predict severe AP, necrosis, and mortality in AP patients. In the total 185 patients included in the study, forty (22%) participants developed organ failure and were classified as severe AP, 36 (19%), had necrosis and mortality rate was 3.8%. Diagnostic accuracy for BISAP score was 81%, Ranson’s score was 94%, APACHE-II was 72%, and CTSI was 84% respectively. The study concluded that the BISAP score was a precise instrument to predict risk factors in AP patients with added advantage of ease of use and early predictability.

Gompertz M, et al. [5] published a retrospective review in 2012 in 128 patients with acute pancreatitis. The BISAP, APACHE II and Balthazar scoring system was used to check the severity. The Area under the curve for BISAP score was 0.977. The sensitivity was 71.4%, specificity was 99.1% and the positive and negative predictive value was 83.3% and 98.3 A statistically significant correlation was found was among BISAP scores and duration of hospital stay. The study concluded that BISAP system of scoring was an important tool for predicting SAP. Moreover, all the parameters were on the first day of hospital. The sensitivity and specificity was better in BISAP compared to APACHE II and Balthazar systems.
Bezmarevic M, et al. [6] conducted a prospective study in 2012 in fifty one patients with pancreatitis patients out of which 29 had severe AP, to find the best scoring systems in predicting the SAP within 24 h of hospital admission. PCT, CRP, BISAP score and APACHE II score was used as measurements. Sensitivity was found to be 74% and specificity was found to be 59% for BISAP score. Sensitivity was found to be 89% and specificity was found to be 69% for APACHE II, CRP had 75% sensitivity and 86% specificity and PCT had 86% sensitivity and 63% specificity, respectively. The study findings concluded that PCT significantly predicts the disease outcome compared to other systems (p < 0.001). APACHE II score was found to be a stronger predictor of the disease severity than BISAP score.

Kim BG, et al. [7] in 2013 conducted a comparative study between BISAP and serum procalcitonin in predicting SAP. Fifty patients were included in the study. The scoring systems compared were BISAP score with serum PCT, Ranson’s score, and APACHE-II, Glasgow, and BCTSI scores. The diagnostic accuracy of BISAP score i.e. >/= 2 was 84% at predicting SAP while it was only 76% for the serum PCT. This was similar to the APACHE-II score. In univariate analysis, BISAP had better statistical significance than serum PCT. Based on the study findings, the authors concluded that, BISAP was a more accurate tool for predicting AP severity than the serum PCT, APACHE-II, Glasgow, and BCTSI scores.

Khanna AK, et al. [8] conducted a study in 2013 on comparison of multiple scoring systems such as Ranson, Glasgow, APACHE-II, CTSI scores, IL-6, CRP, and procalcitonin in for the purpose of assessing the disease severity, risk of progression, pancreatic necrosis and death among those with in acute pancreatitis. Of the total 72 subjects, 31 encountered organ failure and severe acute pancreatitis, 17 participants suffered from pancreatic necrosis, and 12.5% participants expired. The study concluded that for the purpose of assessing disease severity, the best parameters are CRP and IL-6. On the other hand, for the prediction of mortality due to acute pancreatitis, best scoring systems are APACHE-II and Ranson score.

Zhang J, et al. [9] conducted a study from 2010 to 2013 in China to compare BISAP scores and Atlanta classification in for the purpose of assessing the disease severity of acute pancreatitis. From 155 patients included in the study, 16.7% had severe AP, and 3.2% participants deceased. It was found that with increase in BISAP score, there were statistically significant increase in severity, necrosis and deaths. The AUC for BISAP was 79.3%, APACHE II was 83.6% and by Ranson score was 90.3%. The area under the curve for BISAP in the predictive validity, necrosis was 0.834 compared to 0.801 for APACHE II and 0.840 for Ranson score and for mortality it was least for BISAP. The study concluded that the BISAP scoring may be of great utility for the purpose of classifying patients of acute pancreatitis based on risk and predicting the clinical course and prognosis.

An earlier prospective study by Senapati D, et al. [10] in 2014 who conducted study to determine the accuracy of the BISAP score in the prediction of disease severity among patients who have acute pancreatitis. 246 patients were included in the study. 15.8% patients developed organ failure. 6.9% patients had persistent organ failure, 5.2% patients in the study died. It was also found that a BISAP score of >/= 3 had 92% sensitivity, 76% specificity of, 17% positive predictive value, and 99% negative predictive value. The study concluded that BISAP score was simple and accurate method for early detection of SAP and mortality.

Sharma V, et al. [11] conducted a cross-sectional study for the purpose of comparison of various scores such as CTSI, modified CTSI, BISAP scoring, EPIC scoring among patients who have acute pancreatitis. They aimed to compare the risk of multi-organ failure, effect of various interventions and the risk of deaths. They included a total of 105 patients in the study. Among the study population, 8 individuals expired. 71 had POF. Only 16 needed some form of intervention. The area under curve for predicting POF was 0.75 which was significantly higher compared to other scores. Similarly, the area under curve for the need for radiologic intervention for BISAP score was 0.64 which was also significantly higher compared to other scoring systems. The area under curve for the prediction of in-hospital mortality for BISAP was 90% which was much higher and highly significant compared to other scores. Thus BISAP was found to have the most significant utility for the purpose of prediction of mortality and POF among those with acute pancreatitis.

Zheng J, et al. [12] in 2015 conducted a retrospective study to explore the early evaluations of BISAP and C-reactive protein in predicting the severity of acute pancreatitis. 114 participants were enrolled in the study. The study findings suggested that with increase in BISAP scores, there was increase in severity and mortality of acute pancreatitis (P < 0.01). BISAP scores were positively correlated with CRP, D-dimer and serum glucose and negatively correlated with serum calcium. The area under the curve of predicting SAP was 0.873 for and mortality was 0.909. The study concluded that BISAP score plus CRP had a good predictive in predicting SAP and death.

Vinnik YS, et al. [13] in 2015 conducted a retrospective study to evaluate the diagnostic accuracy of Tolstoy-Krasnogorov score, Ranson, BISAP, Glasgow, MODS2, APACHE II and CTSI, at estimating the severity of acute pancreatitis. Results showed the best scoring system used was valuables score for estimation of acute pancreatitis severity was BISAP.
with sensitivity of 98.10%. The study concluded that most effective scores for estimation of acute pancreatitis severity are Score of estimation by Tolstoy-Krasnogorov, Ranson, Glasgow and BISAP Scoring systems MODS 2, APACHE I high specificity and positive predictive value.

Koziel D, et al. [14] in 2015 conducted a study to determine prognostic value of various scales of acute pancreatitis in determining the severity and mortality on 1014 patients. BISAP, Panc 3 scores and Ranson scales were calculated within 24 h of hospitalization. The 81.1% had mild AP 12% had moderate AP and 6.9% had severe AP and 3.7% expired. (APACHE) II was found to be the most accurate scale with AUC of 0.724 followed by BISAP with AUC 0.693 in determining the disease severity. With respect to the prognosis for mortality, AUC for APACHE II was 0.726 and BISAP was 0.707, not significantly different.

Shabbir S, et al. [15] conducted a cross-sectional study in 2015 to determine the accuracy of BISAP score with Ranson’s score in determining the severity of illness, prognosis and risk of deaths among those who have acute pancreatitis in Islamabad. Of the total 80 participants, almost one-third (31.25%) were identified as having severe acute pancreatitis and only three participants (3.75%) had features which were suggestive of necrosis. The mortality among the study participants was 5%. Nearly one-fifth (18.75%) of the participants had a BISAP score of more than or equal to 3 while 31.25% had Ranson’s score more than or equal to 3. The study concluded that BISAP score was simple and an accurate tool for classification of severity and was as effective as the Ranson’s score among patients with acute pancreatitis in determining the frequency of severity and subsequently mortality.

Cho JH, et al. [16] conducted a study in prospective study to find the predictive validity of many scoring systems for the purpose of predicting the severity of illness among the individuals with acute pancreatitis between 2011 and 2012 was retrospectively analyzed. Ranson, APACHE-II, BISAP scores, and CTSI scores were determined for all participants. 161 patients were included in the study and 13% were categorized as SAP, and the risk of mortality was 1.9% among the study participants. The severe AP was associated with higher scores with respect to Ranson score ≥ 3 and the AUC was 0.69, BISAP score ≥ 2 and the AUC was 0.74, APACHE-II score ≥ 8 and the AUC was 0.78, CTSI score ≥ 3 and the AUC was 0.69, and CRP (24) score ≥ 21.4 and the AUC was 0.68. It was seen that APACHE-II had highest diagnostic accuracy for prediction of SAP. However, there was no significant difference when pairwise comparison was made. The study concluded that all the scoring systems included in the study had the same predictive accuracy.

An earlier retrospective study conducted by Qiu L, et al. [17] in 2015 between 2008 and 2014 to determine various conventional scoring systems such as BISAP, Ranson, CTSI, SIRS scores in prognosis of Acute Pancreatitis among patients with hyperlipidemia as the etiology. Among the study population enrolled, 14.2% had Hyperlipidemic AP, 22.8% had SAP. The AUC for BISAP was 0.905, Ranson was 0.938, SIRS was 0.812, and CTSI was 0.834 in predicting SAP. The study concluded that BISAP, Ranson, SIRS, and CTSI all had good accuracy in predicting the prognosis of HLAP.

An earlier meta analyzed study by Gao W, et al. [18] showed the accuracy of the BISAP score in predicting death and severe acute pancreatitis in 2015. A total of 12 cohorts from 10 different studies were taken into consideration. The BISAP score of ≥3 for mortality demonstrated a pooled sensitivity of 56% and a specificity of 91%. Regarding the outcome of SAP, the pooled sensitivity was 51% and the specificity was 91%. Hence it is a useful rule-out test, for ruling out patients who are at a lesser risk of severe illness. The pooled positive LR was 7.23 and the pooled negative likelihood ratios was determined 0.56. The study concluded that when Ranson criteria and APACHE score were compared with BISAP score; both the tools showed higher sensitivity but had lower specificity for SAP and mortality. The BISAP score had a very good specificity, while having a comparatively lower sensitivity for the risk of death and SAP.

Yang L, et al. [19] conducted a retrospective study from 2006 to 2015 to compare four different scoring systems - BISAP, Ranson, MCTSI, and APACHE II - in predicting SAP in a total of 326 participants. The results demonstrated that all the above-mentioned systems are of value for the purpose of prediction of disease severity, local complications, and the risk of death. APACHE II was found to have the highest predictive value in assessing the severity of hyperlipidemic acute pancreatitis. Still it had difficulties in predicting the risk of local complications. MCTSI performed excellently in predicting the risk of pancreatic complications, but it struggled to predict the severity of acute pancreatitis and the risk of in-hospital death. On the other hand, the BISAP score had excellent accuracy with respect to the disease severity of acute pancreatitis, local complications, and the risk of mortality due to hyperlipidemic acute pancreatitis. But there is scope for improvement of the accuracy of BISAP score for HLAP in the future.

Yang LX, et al. [20] conducted a retrospective study from 2006 to 2015 to explore the four measures, BISAP, Ranson score, MCTSI and APACHE in assessment of severity and prognosis of hyperlipidemic acute pancreatitis in a total of 326 patients. As for the severity, the sensitivity of APACHE was 57% and AUC was 0.814, this was higher compared
An earlier study by Ye JF, et al. [21] (2017) conducted to see consistency of BISAP, Modified Early Warning Score (MEWS), serum Ca2+, and red cell distribution width (RDW) in predicting acute pancreatitis severity. A total population of 302 patients was included in the study of, 209 had mild acute pancreatitis and 93 suffered from severe acute pancreatitis. Univariate analysis, was done which showed that the BISAP, MEWS and serum Ca2+ had good predictive capacity to detect the severity of AP (P-value<0.001), whereas RDW was not associated with good prediction. The study concluded that BISAP and serum Ca2+ have high predictive value for the severity of AP.

Valverde-Lopez F, et al. [22] conducted a prospective study from 2018 to 2012 aimed at assessing ability of certain biomarkers and few scoring systems in acute pancreatitis from 269 patients. 6.3% were classified as having SAP and the mortality was 3%. It was found that BISAP had the highest predictive validity (90%) in terms of predicting the risk of severity of acute pancreatitis. Furthermore, it also had high accuracy for predicting the risk of in-hospital deaths (97%) and the rate of admission in intensive care (89%). The blood urea nitrogen performed after 48 hrs had an area under curve of 96%.

Mention must be of study by Chen C, et al. [23] who compared EPIC score, Balthazar score, bedside index of severity in acute pancreatitis (BISAP), and systemic inflammatory response syndrome (SIRS) score in 208 patients to determine severity and organ failure at early stages of AP. 22.6% participants developed organ failure, and 2.4% had infection and had to undergo surgery. Two patients expired. The area under the curve for EPIC score, BISAP and SIRS were 0.724, 0.773, and 0.801 and here was no significant difference between the groups. The study concluded that the EPIC scoring system was a good instrument in predicting of organ failure. However, it didn’t differentiate severity and number of failed organs in early phase AP.

An earlier prospective study by Kaushik M, et al. [24] (2017) to find the correlation between BISAP score and Marshall score to among those who have acute pancreatitis for the purpose of assessing predictive value in organ failure. The outcome was seen within 24hrs and at 28 days. 50 patients were assessed, and it was found that the BISAP score measured within the period of 24 hours of getting admitted to the hospital had good prediction at detecting severity and complication (P<0.001). When outcome was assessed after 4 weeks, a little more than half of the study population recovered (54%), one-third (36%) experienced complications and one-tenth (10%) died. The study concluded that BISAP score was a reliable and consistent measure for predicting the severity and the risk of organ failure and for classifying patients with Acute Pancreatitis (AP) <24 hours from admission.

Harshit Kumar A, et al. [1] conducted a prospective study to compare the APACHE II, BISAP, Ranson’s score and CTSI for the purpose of predicting the severity of illness in the patients with acute pancreatitis. The study population included 50 patients. The prevalence of severe acute pancreatitis among the study population was 28%, 30% had pancreatic necrosis. Among the study population, 28% had progression to persistent organ failure and 28% needed ICU admission. Modified CTSI had the highest area under curves for prediction of severe acute pancreatitis (91.9%), pancreatic necrosis (99.3%), organ failure (89.3%) and ICU admission (99.3%). APACHE II had the second highest values for area under curve for predicting severe acute pancreatitis (83.4%) and organ failure (83.1%). The sensitivity of APACHE II for prediction of organ failure, pancreatic necrosis and ICU admission were 92.86%, 93.33% and 92.31% respectively.

Vasudevan S, et al. [25] conducted a study for the purpose of comparison of the various blood parameters and scores measured on the admission day in predicting the outcome. They included a total of 343 patients. Among them, severe Acute pancreatitis was seen in nearly half the patients (49.6%), 28% of the patients developed infection and necrosis of the pancreas while the rate of mortality was 18%. The study concluded that both BISAP and APACHE II are similar with respect to the prediction of outcome in acute pancreatitis. But the BISAP score utilized the same cut off and predicted all 3 outcomes and hence is a more useful and standard scoring system, compared to the APACHE II.

**Conclusion**

The results show that APACHE 2, BISAP scoring, Modified CTSI and Ranson’s scoring are the most frequently used and have the best specificity and sensitivity among the various scoring systems. However, it was observed that studies conducted on Indian population, especially south Indian population, are relatively scarce. It is of our opinion that similar large studies need to be conducted in a south Indian population, in order to better understand the implications of the change in race and food habits of the study population and usefulness of the same scoring systems in ascertaining the prognosis of pancreatitis when compared to the rest of the population.
the world. The key lacunae identified in the literature are relative scarcity of the studies comparing the predictive validity of BISAP scoring and variable levels of sensitivity and specificity reported across different studies.

**References**


