How to Evaluate the Severity of Acute Pancreatitis: Back to the Past?

Raffaele Pezzilli

Pancreas Unit, Department of Digestive Diseases and Internal Medicine, Sant’Orsola-Malpighi Hospital, Bologna, Italy

Selecting patients with severe acute pancreatitis as early as possible after the onset of symptoms is crucial for appropriate treatment [1]. The guidelines indicate that the assessment of severity should be carried out as soon as possible by a scoring system, such as Acute Physiology and Chronic Health Evaluation (APACHE) II which is a grade A recommendation [2, 3]. The comment added by the authors of the Italian guidelines is that an APACHE II score greater than 8 is important for determining treatment policy and identifying the need for transfer to a referral unit. In addition, the Italian guidelines also suggest that serum C-reactive protein (CRP) values are useful for severity assessment, although they may not reflect severity within the first 48 h after onset; this is also a grade A recommendation [3]. An open question is: “should we use a scoring system or a single marker?”. The question is not semantic but practical because a single marker is easy to use and clearly comparable between the various hospitals whereas scoring systems may be complex and difficult to use in routine clinical practice. Another open question regarding single markers is: “should we continue to use (CRP) or should we use markers, such as interleukin (IL)-6 or IL-8 which are able to predict the severity of acute pancreatitis earlier than CRP?” [4].

In 2009, Aoun et al. [5] carried out a meta-analysis to assess the accuracy of IL-6 and IL-8 in predicting severe acute pancreatitis. They identified relevant published articles and calculated pooled sensitivities, specificities and likelihood ratios using the random-effect model. The authors included eight reports for IL-6 and analyzed five studies for IL-8. The pooled IL-6 sensitivities ranged from 81.0 to 83.6% and specificities from 75.6 to 85.3% with positive likelihood ratios of 3.43, 4.90 and 4.40 for days 1, 2 and 3, respectively. The IL-8 pooled sensitivities ranged from 65.8 to 70.9% with specificities of 66.5 and 91.3% for days 1 and 2 with positive likelihood ratios of 1.96 and 8.15, respectively. The diagnostic odds ratio for IL-6 was higher than that of IL-8 on the first day of illness. Thus, the authors concluded that IL-6 and IL-8 seem to perform at an acceptable level in predicting severe acute pancreatitis, and we believe that these two markers, or at least IL-6, should be used in the clinical setting in place of CRP. Subsequently, a secondary analysis of three prospective acute pancreatitis cohort studies was carried out using a simpler parameter, such as blood urea nitrogen (BUN) [6]. The meta-analysis and stratified multivariate logistic regression adjusted for age, sex and creatinine levels were calculated to determine the risk of mortality associated with an elevated BUN level at admission and a rise in BUN level at 24 hours. The accuracy of the BUN measurements was determined by the area under the receiver operating characteristic curve; the analysis compared BUN with serum creatinine measurement and the APACHE II score. A total of 1,043 patients with acute pancreatitis were included in the pooled analysis; a BUN level of 20 mg/dL or higher was associated with an odds ratio of 4.6 for mortality. Any rise in BUN level at 24 hours was associated with an odds ratio of 4.3 for mortality. The accuracy of the serial BUN measurement was comparable to that of the APACHE II score. However, we should be careful regarding the use of BUN in patients with known chronic renal insufficiency; however, this single marker may help in evaluating patient response to early resuscitation efforts.

In the last few years, many scoring systems have been proposed other than the APACHE-II, Ranson’s and Glasgow scores to select patients with severe acute pancreatitis, such as the BISAP (BUN, impaired mental status, systemic inflammatory response score (SIRS), age, pleural effusion) [7], HAPS (abdominal tenderness, hematocrit, creatinine) [8], JSS (base excess, PaO₂, BUN, creatinine, LDH, platelet, calcium,
CRP, SIRS, age) [9], Panc 3 (hematocrit, BMI, pleural effusion) [10], POP (age, mean arterial pressure, PaO$_2$, FiO$_2$, arterial pH, BUN, calcium) [11] and SIRS (temperature, heart rate, respiratory rate, white blood cell count) [12]. When taking these various scoring systems into consideration, there is no comparative evaluation of them in a clinical setting. Thus, the recent paper of Mounzer et al. [13] is important because the authors carried out a head-to-head study comparing the accuracy of the various scoring systems in predicting persistent organ failure in two prospective cohorts of patients: one used for training (involving 256 patients) and the other used for validation (involving 397 patients). What are the results? All the existing scoring systems had modest accuracy in the training cohort and in the validation cohort, and the old reliable Glasgow score, never to be forgotten, was the best classifier at admission in both cohorts. Of interest, the serum levels of creatinine and blood urea nitrogen provided similar levels of discrimination in each set of patients. Only a combination of predictive rules is more accurate than each single scoring system, but this is quite unsuitable in a clinical setting. In conclusion, should we go back to using the Glasgow score or can we use a single early marker of severity, such as IL-6? This is the topic to be explored in the near future.

**Conflict of interest**

None

**References**


