*JOP. J Pancreas (Online) 2006; 7(2):245-248.*

LETTER

# Assessment of the Severity of Acute Pancreatitis. The Usefulness of ROC Analysis in Comparative Studies of Clinical and Imaging Prognostic Indices

**Paraskevas S Brestas[1], Urania G Dafni[2]**

[1]Department of Radiology, General Hospital of Athens. [2]Division of Public Health, Laboratory of Biostatistics, Department of Nursing, University of Athens. Athens, Greece

Dear Sir:

We read the original article regarding assessment of the predictive value of the CT severity index (CTSI) for the severity of acute pancreatitis by Gürleyic *et al*. with great interest [1]. The purpose of this study was to make a comparative assessment of the accuracy of the CTSI proposed by Baltazar *et al*. [2], the APACHE II score and serum CRP concentrations in predicting the severity of acute pancreatitis. The cut-off values of 3 for the CTSI and 7 for the APACHE II were used in the present study, based on the results of studies which did not have exactly the same discrimination endpoint [2, 3]. The discrimination endpoint in this study is defined as the ability to separate those patients who had mild pancreatitis from those who had severe pancreatitis, according to the classification criteria of the 1992 Atlanta International Symposium [4].

Receiver operating characteristic (ROC) curves could have been used to determine the most appropriate cut-off point for the selected discrimination endpoint which corresponds to the best possible trade-off between sensitivity and specificity which were estimated in the present sample. Moreover, the area under the ROC curve is a reliable measure of overall predictive discrimination and a previously described method for comparison of the areas under the ROC curves, derived from the same cases, could also have been used [5, 6].

The additional information provided by ROC curves in studies of prognostic indices of acute pancreatitis severity derives from the complete illustration of the relationship between sensitivity and specificity for a certain discrimination endpoint (severe *vs*. mild pancreatitis in this case). This might have been useful because: a) the clinical impact of the two types of misclassification (failure to correctly identify a case of severe pancreatitis or failure to correctly identify a case of mild pancreatitis) is not the same, and b) it is necessary to realize new prospective comparative studies for assessing the clinical impact of promising imaging techniques, such as MRI or contrast-enhanced US in the near future [7, 8].

Moreover, it should be noted that even though the Atlanta classification system provides a reliable basis for experimental studies for the clinical management of acute pancreatitis, it is not considered to be a perfect system since intermediate forms of the disease do occur [9]. If an imperfect gold standard is used, the estimated accuracy of the tests may suffer ("imperfect gold standard bias"). Another type of bias affecting ROC analysis is "verification bias" and takes place if some of the patients with test results do not have verified disease status or if the decision to

verify a patient is influenced by the test results. Calculation of the accuracy of a diagnostic test using standard definitions unavoidably includes the risk of some kind of bias under certain circumstances. ROC analysis offers the possibility of bias-correction methods [10] and methods of non-parametric estimation of ROC curves have been also suggested recently in the case in which the gold standard is not binary or in the absence of a gold standard, [11, 12]. Thus, ROC analysis should be the preferred method for the assessment of the predictive value of imaging techniques.

Received December 14th, 2005

**Keywords** Inflammation; Pancreas; Pancreatitis, Acute Necrotizing; ROC Curve

**Abbreviations** CTSI: CT severity index

**Correspondence**
Paraskevas S Brestas
Ag. Marinis 21, Melissia
15127 Athens
Greece
Phone: +30-210.804.6453/4120, +30-697.272.9154
Fax: +30-210.804.4120
E-mail: pbrestas@yahoo.gr

**References**

1. Gurleyik G, Emir S, Kilicoglu G, Arman A, Saglam A. Computed tomography severity index, APACHE II score, and serum CRP concentration for predicting the severity of acute pancreatitis. JOP. J Pancreas (Online) 2005; 6:562-7. [PMID 16286706]

2. Balthazar EJ, Robinson DL, Megibow AJ, Ranson JH. Acute pancreatitis: value of CT in establishing prognosis. Radiology 1990; 174:331-6. [PMID 2296641]

3. Larvin M, McMahon MJ. APACHE-II score for assessment and monitoring of acute pancreatitis. Lancet 1989; 2:201-5. [PMID 2568529]

4. Bradley EL 3rd. A clinically based classification system for acute pancreatitis. Summary of the International Symposium on Acute Pancreatitis, Atlanta, Ga, September 11 through 13, 1992. Arch Surg 1993; 128:586-90. [PMID 8489394]

5. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982; 143:29-36. [PMID 7063747]

6. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983; 148:839-43. [PMID 6878708]

7. Pezzilli R, Fantini L. The imaging assessment of the severity of acute pancreatitis may change in the near future. JOP. J Pancreas (Online) 2005; 6:467-9. [PMID 16186671]

8. Brocchi E, Piscaglia F, Bonora M, Celli N, Venturi A, Fantini L, et al. Echo-enhanced ultrasonography: is it the future gold standard of imaging in acute pancreatitis? JOP. J Pancreas (Online) 2005; 6:464-6. [PMID 16186670]

9. Balthazar EJ. Acute pancreatitis: assessment of severity with clinical and CT evaluation. Radiology 2002; 223:603-13. [PMID 12034923]

10. Zhou XH. Correcting for verification bias in studies of a diagnostic test's accuracy. Stat Methods Med Res. 1998; 7:337-53. [PMID 9871951]

11. Obuchowski NA. Estimating and comparing diagnostic tests' accuracy when the gold standard is not binary. Acad Radiol 2005; 12:1198-204. [PMID 16099683]

12. Zhou XH, Castelluccio P, Zhou C. Nonparametric estimation of ROC curves in the absence of a gold standard. Biometrics 2005; 61:600-9. [PMID 16011710]

## REPLY

Dear Sir:

We would like to thank Brestas *et al*. for their interest in our article published in JOP. J Pancreas (Online) [1]. In this study, The results of an imaging method (computed tomography), a clinical scoring system (APACHE II) and a biochemical measurement (serum CRP concentration) were evaluated in a group of patients with acute pancreatitis. The accuracy of these variables for predicting the clinical course of the disease was calculated. This had initially been

classified using the Atlanta criteria. Brestas *et al*. presented their opinions and suggestions mainly on two subjects:

- the cut-off values of the computed tomography severity index (CTSI) and APACHE II;
- evaluation of the results using the receiver operating characteristics (ROC) curve.

We chose these cut-off values because of our personal experience [2] on acute pancreatitis cases (length of hospital stay, morbidity and mortality, etc.) and the results of previous studies. Larwin and McMahon [3] found that patients with an APACHE II score greater than 7 were likely to have a severe clinical course while Connor *et al*. [4] reported increasing mortality in patients with APACHE II scores greater than 7. Balthazar *et al*. [5] reported that acute pancreatitis cases with a CTSI of 0 to 2 points had 2% morbidity and no mortality while morbidity and mortality rates for patients with a CTSI ranging from 0 to 3 were 8% and 3%, those with a CTSI of 4-6 were 35% and 6% and those with a CTSI of 7-10 were 92% and 17%, respectively. Therefore, aiming to have a lower mortality rate for our patients, we preferred to use a CTSI score of 3 and an APACHE II value of 7 as our cut-off points.

The variables of our study were compared by valid statistical methods. The sensitivity, specificity, positive and negative predictive values and accuracy were also calculated for each variable. We agree with Brestas *et al*. that the ROC curve could have been used, especially in assessment of the imaging technique. Thus, we applied the test to our results.

The area under the ROC curve (AUC±SE evaluated by SPSS 10.0) for CTSI was 0.969±0.024, clearly supportive of the high accuracy of this index in predicting the severity of acute pancreatitis (Figure 1). A value of the CTSI equal to 3 was identified as the best cut-off using the procedure proposed by Pezzilli *et al*. [6] (the maximum likelihood ratio, LR, was 10.3). This value corresponds to previously published values of sensitivity and specificity [1]. The APACHE II gave an

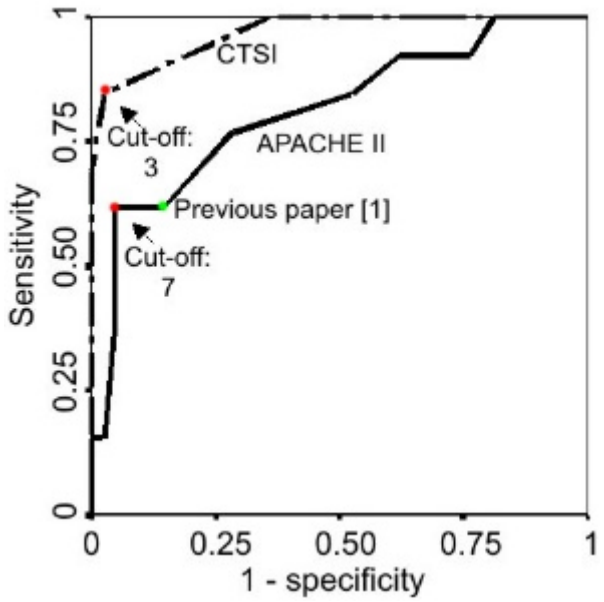


**Figure 1.** Receiver operating characteristics (ROC) curves of the computed tomography severity index (CTSI) and APACHE II in distinguishing between mild and severe acute pancreatitis. Data from Gurleyik *et al*. [1]. Red bullets show the best cut-off values. Green bullets shows the point of the ROC curve corresponding to the values of previously published APACHE II score [1].

AUC value of 0.812±0.074 and a best cut-off value of 7 (the maximum LR was 3.6). Sensitivity, specificity, and frequency of cases correctly identified by applying this cut-off are 61.5% (8 out of 13 severe acute pancreatitis patients), 95.2% (40 out of 42 mild pancreatitis patients), and 87.3% (48 out of 55 overall acute pancreatitis patients), respectively.

Finally, it should be noted that the AUC of the CTSI was significantly higher when compared with the APACHE II score (P=0.044, z-test).

In conclusion, the use of ROC curve analysis confirmed our previous results [1] showing that values of the CTSI greater than 3 are highly indicative of severe acute pancreatitis. On the other hand, the ROC curve applied to our data indicates values of an APACHE II score greater than 7 are necessary in order to identify severe acute pancreatitis patients instead of values equal to or greater than 7 as we used in our paper [1]. Moreover, both the CTSI and the APACHE II scores are highly accurate in predicting the natural outcome of

specific patients with acute pancreatitis, with the indicated cut-off values clearly specific for the purpose. Finally, the CTSI is significantly more accurate than the APACHE II score.

Gunay Gurleyik[1]
Gamze Kilicoglu[2]

*[1]Department of Surgery and [2]Department of Radiology, Haydarpasa Numune Education and Research Hospital. Istanbul, Turkey*

Received January 24th, 2006

**Keywords** APACHE; Pancreatitis, Acute Necrotizing; ROC Curve; Tomography, X-Ray Computed

**Correspondence**
Günay Gürleyik
Eski Bagdat cad. 29/9
Altintepe 34840 Istanbul
Turkey
Phone: +90-216.489.8325
Fax: +90-216.373.1096
E-mail: ggurleyik@yahoo.com

**References**

1. Gurleyik G, Emir S, Kilicoglu G, Arman A, Saglam A. Computed tomography severity index, APACHE II score, and serum CRP concentration for predicting the severity of acute pancreatitis. JOP. J Pancreas (Online) 2005; 6:562-7. [PMID 16286706]

2. Gurleyik G, Cirpici OZ, Aktekin A, Saglam A. The value of Ranson and APACHE II scoring systems, and serum levels of interleukin-6 and C-reactive protein in the early diagnosis of the severity of acute pancreatitis. Ulus Travma Derg 2004; 10:83-8. [PMID 15103565]

3. Larvin M, Mc Mahon MJ. APACHE -II score for assessment and monitoring of acute pancreatitis. Lancet 1989; 2:201-5. [PMID 2568529]

4. Connor S, Ghaneh P, Raraty M, Rosso E, Hartley MN, Garvey C, et al. Increasing age and APACHE II scores are the main determinants of outcome from pancreatic necrosectomy. Br J Surg 2003; 90:1542-7. [PMID 14648734]

5. Balthazar EJ, Robinson DL, Megibow AJ, Ranson JH. Acute pancreatitis: value of CT in establishing prognosis. Radiology 1990; 174:331-6. [PMID 2296641]

6. Pezzilli R, Billi P, Miniero R, Fiocchi M, Cappelletti O, Morselli-Labate AM, et al. Serum interleukin-6, interleukin-8, and beta 2-microglobulin in early assessment of severity of acute pancreatitis. Comparison with serum C-reactive protein. Dig Dis Sci 1995; 40:2341-8. [PMID 7587812]