Diagnostic Accuracy of Clinical, Laboratory, and Radiological Findings in Incident versus Recurrent Acute Pancreatitis

Abdullah A Almamari*, Ahmer Akbar Memon², John A Windsor¹ and Maxim S Petrov¹

¹Department of Surgery, University of Auckland, Auckland, New Zealand
²Armed Forces Hospital, Muscat, Oman

Abstract

Background: The overall incidence of acute pancreatitis (AP) is deemed to be increasing in the past three decades. However, while the incidence of incident AP appears to be increasing, the incidence of recurrent AP appears to be decreasing. The aim of this study was to evaluate and compare the diagnostic accuracy of the clinical, laboratory, and radiological findings in diagnosing of incident versus recurrent AP.

Methods: All consecutive adult patients with a nonelective hospital admission and elevated total blood amylase and/or pancreatic amylase were prospectively enrolled over a 2-year study period. The diagnostic accuracy metrics were sensitivity (Se), specificity (Sp), positive likelihood ratio (PLR), negative likelihood ratio (NLR), positive predictive value (PPV), and negative predictive value (NPV).

Results: Of the total 1392 patients included in this study, 223 (16%) patients met the recommended diagnostic criteria for AP. When determining the relative accuracy of clinical domain, specificity, PLR, and NPV were significantly lower (p < 0.05) in patients with incident AP as compared with recurrent AP whereas PPV was significantly higher in patients with incident AP. Similar differences in terms of specificity, PLR, and PPV were observed when determining the relative accuracy of the radiological and laboratory domains in incident versus recurrent AP. Sensitivity and NLR did not differ significantly between the groups.

Conclusion: There is a significant variance in accuracy of clinical, laboratory, and radiological criteria when used in diagnosing of incident versus recurrent AP. The consistently lower specificity and PLR of all the three recommended domains in patients with incident AP suggests that the reported incidence of AP might be inflated.

Introduction

Acute pancreatitis (AP) is the most common hospital discharge diagnosis of all gastrointestinal diseases [1] and is a considerable financial burden on health care systems. A study from the United States estimated that the total cost of AP admissions in 2003 was $2.2 billion, mean cost per hospitalisation was $9870, and mean cost per hospital day was $1670 [2]. Several studies published in the last decade suggested that the overall incidence of AP is increasing [3-10]. This has been attributed to the increasing prevalence of risk factors (cholelithiasis, obesity, diabetes), aging population, and improved diagnostic workup [5,7,8,11-14].

Another possible reason for the increasing incidence of AP is over-diagnosing of this disease, though this issue has been poorly investigated. Conventionally, the diagnosis of AP is based on two of three criteria: abdominal pain consistent with AP, elevated serum pancreatic enzymes, and/or radiological features of AP. And these criteria have been re-affirmed in the 2013 AGA guidelines and the 2013 IAP/APA guidelines [15,16]. Although widely accepted, these diagnostic criteria are based on the assumption that the accuracy of the clinical, laboratory, and radiological domains is similar, and that they have equivalent accuracy when applied to patients with incident and recurrent AP. This is an important issue to resolve because large-scale cohort studies report opposing trends. The global incidence of incident AP appears to be increasing independent of geographical location [5] compared with the incidence of recurrent AP which appears to be decreasing [17,18]. We hypothesised that the difference in accuracy for the three diagnostic criteria in patients with incident, as compared with recurrent attack, contributes to over-diagnosing of AP.

Citation: Abdullah A Almamari, *et al*.”Diagnostic Accuracy of Clinical, Laboratory, and Radiological Findings in Incident versus Recurrent Acute Pancreatitis”. EC Gastroenterology and Digestive System 5.8 (2018): 674-681.
Thus, the overall aim of this study was to compare the diagnostic accuracy of the clinical, laboratory, and radiological criteria in patients with incident and recurrent attacks of AP.

Methods

Study setting and data collection

The study was conducted at Auckland City Hospital (Auckland, New Zealand), a publicly funded teaching hospital with a capacity of 710 adult beds, from January 1, 2010 to January 1, 2012. All consecutive adult patients (18 years or older) with elevated total blood amylase (normal reference range: 25 - 135 U/L) and/or pancreatic amylase (normal reference range: 8 - 53 U/L) were prospectively enrolled into this study. Patients admitted electively were excluded. Medical records of all patients were comprehensively reviewed during their hospital admission and the following parameters were collected using a pre-specified data collection form: age, gender, initial level of pancreatic enzymes (serum total amylase and/or pancreatic amylase), presenting signs and symptoms, and radiology reports for all patients who underwent abdominal ultrasound scan and/or abdominal CT scan.

Definitions

Acute pancreatitis cases were defined as those patients who met the diagnostic criteria for AP recommended by the 2013 AGA guidelines and the 2013 IAP/APA guidelines [15,16]. Namely, 2 of the 3 domains (clinical, laboratory, radiological) had to be positive to establish the diagnosis of AP:

- The clinical domain was defined as signs and symptoms suggestive of AP such as acute upper abdominal pain, radiation of pain to the back, nausea, and vomiting.
- The laboratory domain was defined as elevation of serum pancreatic enzymes (total amylase and/or pancreatic amylase) of more than 3 times the upper reference of normal limit (URL).
- The radiological domain was defined as radiological findings indicative of AP on either CT scan and/or US scan. CT scan findings indicative of AP included (but were not limited to) pancreatic enlargement with diffuse edema, peripancreatic stranding and fluid collection and heterogeneity of pancreatic parenchyma. US scan findings indicative of AP included (but were not limited to) diffuse glandular pancreatic enlargement and hypo-echoic pancreatic texture.

An incident attack of AP was defined as first diagnosis of AP, without any history of a prior diagnosis of AP. A recurrent attack of AP was defined as a repeat diagnosis of AP, with a prior history of AP. All patients who did not meet the diagnostic criteria for AP were regarded as "Non-AP" cases.

Statistical Analysis

Descriptive data were presented as median and range for continuous data (age) and as proportions for categorical data (gender). Sensitivity (Se), specificity (Sp), positive likelihood ratio (PLR), negative likelihood ratio (NLR), positive predictive value (PPV), and negative predictive value (NPV) were calculated using the MedCalc software version 12.3.0 (MedCalc Software, Broekstraat 52, 9030 Mariakerke, Belgium). Receiver operator characteristics (ROC) curve analyses were undertaken using SPSS 20.0 software (SPSS Inc., Chicago, III, USA). ROC curves were constructed and area under the curve (AUC) was used to determine and compare the overall accuracy of serum total amylase and pancreatic amylase measurements in discriminating patients with incident and recurrent AP. Results were considered statistically significant if p-value was below 0.05.

Results

Patient characteristics

There were 1392 adult patients admitted to Auckland City Hospital during the study period who had an elevated total serum amylase and/or pancreatic serum amylase. A total of 223 (16%) patients met the diagnostic criteria for AP. Of these, 199 (89%) had incident AP and 24 (11%) had recurrent AP. The median age of patients with incident AP was 57 years (range, 19 - 96) and the median age of patients with recurrent AP was 50 years (range, 27 - 88) (p = 0.68). Ninety-seven patients were males (48.7%) among patients with incident AP and 16 (66.7%) patients were males among patients with recurrent AP (p = 0.35).

Citation: Abdullah A Almamari., et al. "Diagnostic Accuracy of Clinical, Laboratory, and Radiological Findings in Incident versus Recurrent Acute Pancreatitis". EC Gastroenterology and Digestive System 5.8 (2018): 674-681.
Diagnostic accuracy of the clinical domain

When determining the relative accuracy of the clinical domain in diagnosing incident AP, there were a total of 264 out of 1368 (19.3%) patients presented with clinical signs and symptoms suggestive of AP. Forty-eight (18%) patients with clinical symptoms of AP were identified correctly as having AP whereas 216 (82%) patients were identified incorrectly as having AP. One thousand and eighty-six (98.4%) patients without clinical symptoms of AP were identified correctly as not having AP and 18 (1.6%) patients without clinical symptoms of AP were identified incorrectly as not having AP.

When determining the relative accuracy of the clinical domain in diagnosing recurrent AP, there were a total of 111 out of 1193 (9.3%) patients presented with clinical signs and symptoms indicative of AP. Seven (6.3%) patients with clinical symptoms of AP were identified correctly as having AP whereas 104 (93.7%) patients with clinical symptoms of AP were identified incorrectly as having AP. One thousand and eighty-two (100%) patients without clinical symptoms of AP were identified correctly as not having AP and none of the patients without clinical symptoms of AP were identified incorrectly as having AP.

Table 1 presents the diagnostic accuracy metrics for the clinical domain in patients with incident AP as compared with recurrent AP. Specificity, PLR, and NPV of the clinical domain were significantly higher (p < 0.05) in patients with recurrent AP whereas PPV was significantly higher (p < 0.05) in patients with incident AP. The remaining accuracy metrics did not differ significantly between patients with incident and recurrent AP.

<table>
<thead>
<tr>
<th>Domains</th>
<th>Se (95% CI)</th>
<th>Sp (95% CI)</th>
<th>PLR (95% CI)</th>
<th>NLR (95% CI)</th>
<th>PPV (95% CI)</th>
<th>NPV (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incident AP</td>
<td>72.7 (60.4 - 82.9)</td>
<td>83.8 (81.3 - 85.4)</td>
<td>4.4 (3.6 - 5.3)</td>
<td>0.3 (0.2 - 0.5)</td>
<td>18.2 (13.7 - 23.4)</td>
<td>98.4 (97.4 - 99.0)</td>
</tr>
<tr>
<td>Recurrent AP</td>
<td>100 (58.9 - 100)</td>
<td>91.2 (89.5 - 92.8)</td>
<td>11.4 (9.5 - 13.7)</td>
<td>N/E</td>
<td>6.3 (2.6 - 12.6)</td>
<td>100 (99.7 - 100)</td>
</tr>
<tr>
<td>Amylase &gt; 3URL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incident AP</td>
<td>82.8 (70.6 - 91.4)</td>
<td>84.9 (82.9 - 86.8)</td>
<td>5.5 (4.6 - 6.5)</td>
<td>0.2 (0.1 - 0.4)</td>
<td>19.6 (14.8 - 25.1)</td>
<td>99.1 (98.4 - 99.6)</td>
</tr>
<tr>
<td>Recurrent AP</td>
<td>70.0 (34.8 - 92.9)</td>
<td>91.1 (89.3 - 92.7)</td>
<td>7.9 (5.1 - 12.3)</td>
<td>0.3 (0.1 - 0.9)</td>
<td>6.3 (2.6 - 12.5)</td>
<td>99.7 (99.2 - 99.9)</td>
</tr>
<tr>
<td>p. Amylase &gt; 3URL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incident AP</td>
<td>87.7 (76.3 - 94.9)</td>
<td>73.1 (70.6 - 75.5)</td>
<td>3.3 (2.9 - 3.7)</td>
<td>0.2 (0.1 - 0.3)</td>
<td>12.5 (9.4 - 16.2)</td>
<td>99.3 (98.5 - 99.7)</td>
</tr>
<tr>
<td>Recurrent AP</td>
<td>100 (68.9 - 100)</td>
<td>80.8 (78.4 - 82.9)</td>
<td>5.2 (4.6 - 5.8)</td>
<td>N/E</td>
<td>4.2 (2.1 - 7.7)</td>
<td>100 (99.6 - 100)</td>
</tr>
<tr>
<td>Radiological</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incident AP</td>
<td>36.6 (28.4 - 45.5)</td>
<td>96.8 (95.6 - 97.7)</td>
<td>11.3 (7.8 - 16.6)</td>
<td>0.7 (0.6 - 0.8)</td>
<td>54.6 (43.6 - 65.2)</td>
<td>93.5 (92.0 - 94.8)</td>
</tr>
<tr>
<td>Recurrent AP</td>
<td>43.8 (19.8 - 70.1)</td>
<td>99.1 (98.3 - 99.5)</td>
<td>46.8 (20.8 - 105.1)</td>
<td>0.6 (0.4 - 0.9)</td>
<td>38.9 (17.4 - 64.2)</td>
<td>99.2 (98.6 - 99.7)</td>
</tr>
</tbody>
</table>

Table 1: Diagnostic accuracy metrics of the three domains in diagnosing of incident versus recurrent acute pancreatitis.

Footnotes: Se: Sensitivity; Sp: Specificity; PLR: Positive Likelihood Ratio; NLR: Negative Likelihood Ratio; PPV: Positive Predictive Value; NPV: Negative Predictive Value and 95% CI: 95% Confidence Interval; N/E: Not Estimable.

Statistically significant differences between incident and recurrent AP are in bold.

Diagnostic accuracy of the laboratory domain

When determining the relative accuracy of the laboratory domain in diagnosing incident AP, there were a total of 1355 of 1368 (99%) who had both total amylase and pancreatic amylase data available, 3 patients had pancreatic amylase measured only and 10 patients had total amylase measured only. A total of 1334 of 1355 (98.5%) patients presented with an elevated total amylase and 893 of 1355 (66%) patients presented with an elevated pancreatic amylase.

When the threshold of 3URL for total amylase was applied, 48 (19.6%) patients with total amylase above 3URL were identified correctly as having AP and 197 (80.4%) patients with total amylase above 3URL were identified incorrectly as having AP. One thousand one hundred and ten (99%) patients with total amylase below the threshold of 3URL were identified correctly as not having AP and 10 (1%) patients with total amylase below 3URL were identified incorrectly as not having AP.

When the threshold of 3URL for pancreatic amylase was applied, 50 (12.5%) patients with pancreatic amylase above 3URL were identified correctly as having AP and 350 (87.5%) patients with pancreatic amylase above 3URL were identified incorrectly as having AP. Nine hundred and fifty-one (99.3%) patients with pancreatic amylase below the threshold of 3URL were identified correctly as not having AP and 7 (0.7%) patients with pancreatic amylase below 3URL were identified incorrectly as not having AP.

When determining the relative accuracy of the laboratory domain in diagnosing recurrent AP, there were a total of 1182 of 1193 (99%) patients presented with an elevated total amylase and 720 of 1182 (60.9%) patients presented with an elevated pancreatic amylase.

When the threshold of 3URL for total amylase was applied, seven (6.3%) patients with total amylase above 3URL were identified correctly as having AP and 105 (93.7%) patients with total amylase above the threshold of 3URL were identified incorrectly as having AP. One thousand and seventy-six (99.7%) patients with total amylase below 3URL were identified correctly as not having AP and three (0.3%) patients with total amylase below 3URL were identified incorrectly as not having AP. When the threshold of 3URL for pancreatic amylase was applied, 10 (4.2%) patients with pancreatic amylase above the threshold of 3URL were identified correctly as having AP and 226 (95.8%) patients with pancreatic amylase above 3URL were identified incorrectly as having AP. Nine hundred and forty-eight (100%) patients with pancreatic amylase below 3URL were identified correctly as not having AP and none of the patients with pancreatic amylase below 3URL were identified incorrectly as not having AP.

Table 1 presents the diagnostic accuracy metrics for the laboratory domain in patients with incident AP as compared with recurrent AP. When the threshold of 3URL for total amylase was used, specificity was significantly higher in patients with recurrent AP (p < 0.05) whereas PPV was significantly higher (p < 0.05) in patients with incident AP. The remaining accuracy metrics did not differ significantly between patients with incident versus recurrent AP. When the threshold of 3URL for pancreatic amylase was used, specificity and PLR were significantly higher (p < 0.05) in patients with recurrent AP whereas PPV was significantly higher (p < 0.05) in patients with incident AP. The remaining accuracy metrics did not differ significantly between patients with incident versus recurrent AP.

ROC curves were constructed for both total amylase and pancreatic amylase to determine the relative accuracy of both tests in diagnosing incident and recurrent AP. The area under ROC curve for total amylase and pancreatic amylase in diagnosing incident AP was 0.88 (95% CI, 0.83-0.93) and 0.90 (95% CI, 0.86 - 0.93), respectively (Figure 1). The difference between total amylase and pancreatic amylase was not statistically significant (p > 0.05). The area under ROC curve for total amylase and pancreatic amylase in diagnosing recurrent AP was 0.92 (95% CI, 0.87-0.97) and 0.95 (95% CI, 0.92 - 0.98), respectively (Figure 2). The difference between total amylase and pancreatic amylase was not statistically significant (p > 0.05).

Diagnostic accuracy of the radiological domain

When determining the relative accuracy of the radiological domain in diagnosing incident AP, there were a total of 588 (43%) patients who had either US or CT scans done during their hospital admission and the remaining 780 (57%) had either another form of radiology or no radiological imaging. Forty-eight (54.5%) patients with radiological findings suggestive of AP were identified correctly as having AP whereas 40 (45.5%) patients with radiological findings suggestive of AP were identified incorrectly as having AP. A total of 1197 (93.5%) patients without radiological findings suggestive of AP were identified correctly as not having AP and 83 (6.5%) patients without radiological findings suggestive of AP were identified incorrectly as not having AP.

When determining the relative accuracy of the radiological domain in diagnosing recurrent AP, there were a total of 426 (36%) who had either US or CT scans done during their hospital admission and the remaining 767 (64%) had either another form of radiology or no radiological imaging. Eleven (61%) patients with radiological findings suggestive of AP were identified correctly as having AP and seven (39%) patients with radiological findings suggestive of AP were identified incorrectly as not having AP. One thousand one hundred and sixty-six (99.2%) patients without radiological findings suggestive of AP were identified correctly as not having AP and nine (0.8%) patients without radiological findings suggestive of AP were identified incorrectly as not having AP.

Figure 1: Receiver operator characteristic (ROC) curves for serum total amylase and serum pancreatic amylase in diagnosing of incident acute pancreatitis.

Figure 2: Receiver operator characteristic (ROC) curves for serum total amylase and serum pancreatic amylase for diagnosing of recurrent acute pancreatitis.

Citation: Abdullah A Almamari, et al. "Diagnostic Accuracy of Clinical, Laboratory, and Radiological Findings in Incident versus Recurrent Acute Pancreatitis." EC Gastroenterology and Digestive System 5.8 (2018): 674-681.
Table 1 presents the diagnostic accuracy metrics for the radiological domain in patients with incident AP as compared with recurrent AP. Specificity and PLR of the radiological domain were significantly higher (p < 0.05) in patients with recurrent AP whereas PPV was significantly higher in patients with incident AP (p < 0.05). The remaining accuracy metrics did not differ significantly between patients with incident versus recurrent AP.

### Discussion

This large prospective clinical study has compared, for the first time, the accuracy of the clinical, laboratory, and radiological diagnostic criteria in patients with incident versus recurrent AP. It shows that 4 out of 6 diagnostic accuracy metrics (specificity, PLR, PPV, NPV) differ significantly between incident and recurrent cases. Specificity and PLR are affected the most, with all 3 domains (clinical, laboratory, and radiological) being significantly lower in patients with incident AP. Further, PPV is significantly different in regards to clinical and laboratory domains whereas NPV is significantly different in regards to clinical and radiological domains. Sensitivity and NPV of all the 3 domains do not differ significantly in patients with incident versus recurrent AP. These findings have important clinical implications.

When comparing the relative diagnostic accuracy of the clinical domain, the specificity was significantly (p < 0.05) lower for incident AP (83;95% CI, 81 - 85) than recurrent AP (91; 95% CI, 89 - 93). This difference in specificity indicates that 8% more patients without AP will be incorrectly diagnosed with AP based on clinical signs and symptoms. Similarly, when comparing the relative diagnostic accuracy of the laboratory and radiological domains the difference in specificity varies between 3% and 8% and is significantly lower in those with incident AP. These differences may lead to the over-diagnosis of AP in patients with incident attack, as up to 8% will not have AP despite fulfilling the diagnostic criteria. Similar to the findings regarding specificity, the data relating to PLR highlights the problem of over-diagnosis. Considering PLR of the radiological criteria, there was a significant difference (p < 0.05) between patients with incident AP (11; 95% CI, 8 - 16) and those with recurrent AP (47; 95% CI, 21 - 105). In other words, there is a four times lower likelihood of making a correct diagnosis of AP in patients with an incident attack, as compared with a recurrent attack, using radiological criteria. The difference is also significant for the clinical criteria (2 times lower likelihood) and pancreatic enzyme criteria (1.5 times lower likelihood) in patients with an incident attack of AP.

The consequences of over-diagnosing of AP are important. Firstly, it inflates the incidence of AP and this may, at least in part, explain the reports of increasing incidence. Secondly, the inflated incidence of AP may deflate the case-fatality rate, giving a false impression that patients die less frequently because of AP. In fact, while most recent studies reporting on trends indicated a steady decrease in the case-fatality rate over time, several studies reported that the population mortality rate for AP has remained unchanged over time [5,8,19]. The likely explanation for this is that, given that the case-fatality rate is a proportion of deaths within a designated population of people with diagnosis of AP and the population mortality is a rate per 100,000 population, over-diagnosing of AP within a population will result in a decrease in case-fatality but not the population mortality rate. Thirdly, the over-diagnosis of AP will result in over-representation of patients in the mild severity category relative to the others categories [20-21]. Lastly, the over-diagnosis of AP may have cost implications increasing the cost of diagnostic workup and treatment as AP is one of the most costly gastrointestinal diseases [1,22].

A strength of this study is that the diagnosis of AP was based on currently recommended diagnostic criteria [15,16] rather than derived from an administrative database using discharge diagnoses, which have been shown to be incorrect in 20 - 25% of cases [23-25]. In particular, it has been shown by Saligram and colleagues [23] that at least one-fifth of patients who received a discharge diagnosis of AP did not fulfill the recommended diagnostic criteria. Further, a recently published study from Sweden [24] showed that only 17% of patients with discharge diagnosis of AP did not meet the recommended diagnostic criteria. Similarly, Spanier and colleagues [25] showed that the PPV of the discharge diagnosis of AP was only 78%.

This study has some limitations, which need to be acknowledged. The diagnostic accuracy of serum lipase was not determined because it was not routinely measured in our institution until very recently. This should be included in future studies on diagnostic accuracy. The other potential limitation is that only patients with elevated pancreatic enzymes were included in this study. We used this criterion because elevation of pancreatic enzymes is present in the overwhelming majority of AP cases. However, clinical studies in the 1980s showed that a small proportion of patients with AP may present with a normal level of pancreatic enzymes, especially in patients with a delayed presentation, and alcohol-induced or hypertriglyceridemia-induced AP [26-28].
Conclusion

In conclusion, there is a significant difference in the diagnostic accuracy of the clinical, laboratory, and radiological criteria in patients with incident AP as compared with recurrent AP. The consistently lower specificity and PLR across all the three diagnostic domains for incident cases indicates the probability of over-diagnosis. Further quality studies are required to elucidate its implications, both clinical and economic.

Study Highlights

What is current knowledge

- The diagnostic criteria for acute pancreatitis are well established.
- The accuracy of the clinical, laboratory, and radiological findings in patients with incident versus recurrent attack of acute pancreatitis is largely unknown.

What is new here

- Specificity, positive likelihood ratio, positive and negative predictive values differ significantly between incident and recurrent cases.
- Up to 8% of incident cases will not have acute pancreatitis despite fulfilling the diagnostic criteria.
- There is a up to 4 times lower likelihood of making a correct diagnosis of acute pancreatitis in patients with an incident attack, as compared with a recurrent attack.

Bibliography

Diagnostic Accuracy of Clinical, Laboratory, and Radiological Findings in Incident versus Recurrent Acute Pancreatitis


Volume 5 Issue 8 August 2018
©All rights reserved by Abdullah A Almamari., et al.

Citation: Abdullah A Almamari, et al. "Diagnostic Accuracy of Clinical, Laboratory, and Radiological Findings in Incident versus Recurrent Acute Pancreatitis". EC Gastroenterology and Digestive System 5.8 (2018): 674-681.