Predicting the Severity of Acute Pancreatitis by Rapid Measurement of Trypsinogen-2 in Urine

Marko Lempinen,1 Marja-Leena Kylänpää-Bäck,1 Ulf-Håkan Stenman,2 Pauli Puolakkainen,1 Reijo Haapiainen,1 Patrik Finne,2 Armi Korvuo,3 and Esko Kemppainen1*

Background: Early identification of patients at risk of developing a severe attack of acute pancreatitis (AP) is of great importance because rapid therapeutic interventions improve outcome. At a cutoff of 50 μg/L, trypsino- gen-2 measured by a rapid urinary dipstick is a sensitive and specific diagnostic test in AP. The trypsino- gen-2 concentration correlates with the severity of the disease, and a test with a higher cutoff might therefore be useful for prediction of disease severity.

Methods: We increased the detection limit of the urinary trypsino-gen-2 test strip (Actim Pancreatitis) from 50 μg/L to 2000 μg/L and evaluated the prognostic value of this test. The results were compared with those obtained with serum C-reactive protein and the acute physiology and chronic health evaluation II (APACHE II) score. The study population consisted of 150 consecutive patients with AP (42 with severe disease).

Results: The sensitivity of the rapid urinary test strip (detection limit, 2000 μg/L) for prediction of severe AP, both on admission and at 24 h, was 62%; specificities were 87% and 85%, respectively, positive predictive values were 65% and 62%, and negative predictive values were 85% and 85%. C-Reactive protein had a sensitivity of only 38% on admission, but at 24 h, it was 83%; specificities were 90% and 70%, respectively, whereas positive predictive values were 59% and 52%, and NPVs were 79% and 91%, respectively. On admission the positive-likelihood ratio for the urinary trypsino-gen-2 test strip was 4.8, and at 24 h it was 4.2; for C-reactive protein, the values were 3.7 and 2.7, respectively.

Conclusions: The urinary trypsino-gen-2 dipstick is a simple and rapid method for prediction of severe acute pancreatitis.

In most cases, acute pancreatitis (AP) is a mild self-limiting disease without complications, but acute necrotizing pancreatitis, which occurs in ~20% of the attacks, is a life-threatening disease with high morbidity and mortality. Improved outcome in the severe form of AP is based on early identification of disease severity and subsequent management of high-risk patients in intensive care units with early endoscopic retrograde cholangiopancreatography in gallstone-induced disease, early enteral feeding, and prophylactic antibiotics (1–4).

Early prediction of the severity of AP is difficult (5). Several biochemical tests have been evaluated in the severity assessment of AP, including C-reactive protein (CRP), trypsino-gen-activation peptide (TAP), interleukin-6, tumor necrosis factor-α, and platelet-activating factor (6–8). CRP has been most useful, but it is not specific for pancreatitis, and 48 h may elapse before it increases. Urinary TAP concentrations have also been shown to correlate with the severity of AP at admission (8–10), but its measurement by a manual enzyme immunoassay limits its use as an emergency test. Clinico-biochemical scores like the Ranson prognostic signs (11) require 48 h of follow-up, which is considered too late for therapeutic decision-making. The acute physiology and chronic health evaluation II score (APACHE II) can be measured sequentially, but it is cumbersome, which limits its use in clinical practice (12). Computed tomography (CT) is currently the most accurate noninvasive single method to

1 Second Department of Surgery and 2 Department of Clinical Chemistry, University Central Hospital Helsinki, Haartmaninkatu 4, FIN-00290 Helsinki, Finland.

3 Medix Biochemica, FIN-02700 Kauniainen, Finland.

*Author for correspondence. Fax 358-9-47174675; e-mail esko.kemppainen@hus.fi.

Received June 26, 2001; accepted August 20, 2001.

4 Nonstandard abbreviations: AP, acute pancreatitis; CRP, C-reactive protein; TAP, trypsino-gen-activation peptide; APACHE II, acute physiology and chronic health evaluation II score; CT, computed tomography; PPV, positive predictive value; NPV, negative predictive value; and PLR, positive-likelihood ratio.
evaluate the severity of the disease (13), but limited availability, high costs, and potential side effects limit its utility (14). Thus, there is a need for a marker with good prognostic accuracy, already in the early phase of the disease, that can be measured rapidly and inexpensively.

Trypsinogen occurs as two major isoenzymes, trypsinogen-1 (cationic) and trypsinogen-2 (anionic) (15). The inflammatory process in AP leads to leakage of pancreatic enzymes into circulation, and trypsinogen-2 is increased in both serum and urine (16, 17). Premature trypsinogen activation within acinar cells is thought to play a crucial role in the pathogenesis of AP by activating other pancreatic enzymes and initiating autodigestion (18). It has previously been shown that increased trypsinogen-2 concentrations in urine are useful in diagnosing AP and that they also reflect disease severity (15, 19, 20). The median urinary trypsinogen-2 concentration is 5600–10 000 µg/L in severe and 130–890 µg/L in mild AP (17, 20, 21).

A rapid dipstick test for trypsinogen-2 in urine with a fixed detection limit of 50 µg/L is a sensitive and specific diagnostic marker for AP (19, 21, 22), but at that concentration, it does not differentiate between mild and severe disease. To achieve the differentiation we increased the detection limit of the test strip up to 2000 µg/L by a dilution method. The objective of the study was to compare the prognostic value of the rapid urinary test with CRP and the APACHE II score in patients with AP.

**Patients and Methods**

The study was approved by the Committee on Research Ethics at the Helsinki University Central Hospital. The prospective study population consisted of 150 consecutive patients with AP and <72 h of acute symptoms, admitted to the emergency unit at Helsinki University Central Hospital between August 1997 and May 2000. There were 46 women and 104 men with a median age of 45 years (range, 20–94).

**STUDY DESIGN**

Urine samples were obtained on admission and 24 h thereafter in the emergency department. The samples were stored at −20 °C until tested with the strip and measured quantitatively. The tests were performed without knowledge of the clinical presentation of the patients.

The diagnosis of AP was made on the basis of a typical clinical picture and serum amylase at least more than threefold the upper reference limit and/or typical findings on CT. In patients with marginally increased serum amylase concentrations (300–900 U/L) and abdominal pain, pancreatitis was ruled out or confirmed on the basis of repeated serum amylase measurements and clinical follow-up, as well as CT and ultrasonographic studies. APACHE II score values were calculated on admission and at 24 h. The severity of AP was categorized retrospectively by the clinically based classification of the 1992 Atlanta symposium (23).

**RESULTS**

Of the 150 patients, 42 (28%) had a severe disease and 108 (72%) a mild disease according to the Atlanta classification (23). The median duration of pain before hospitalization was 24 h. Thirty-one patients with severe AP had alcohol-induced disease, 5 patients had a gallstone-related disease, and 6 had AP of unknown etiology. In the mild group, 66 patients had alcohol-induced disease, 26 patients had gallstone-related pancreatitis, and 16 patients had AP of other etiologies. In the severe-disease group, 27 (64%) of 42 patients had >30% pancreatic necrosis or other local complications on contrast-enhanced CT, 23 (55%) had organ failure, and 8 (19%) died. The length of the hospital stay of patients with severe AP was 30 days (range, 2–130) and with mild AP, 6 days (range, 1–17).

The median urinary trypsinogen-2 concentration on admission in severe AP was 26-fold higher than that in mild AP and at 24 h the difference was 31-fold (both P <0.0001; Table 2). On admission the sensitivity and specificity of the urinary trypsinogen-2 dipstick in predicting severe AP were 62% and 87%, respectively, whereas they were 62% and 77% for the quantitative assay. At 24 h the sensitivity and specificity were 62% and 85%, respectively, for the rapid dipstick and 71% and 78%, respectively, for the quantitative assay. The corresponding positive predictive values (PPVs) and negative predictive values (NPVs) are shown in Table 1, as are predictive values for combined testing of serum CRP and the rapid

**TRYPsinogen-2 MEASUREMENTS**

The principle of the immunochromatographic urinary trypsinogen-2 dipstick tests has been described (21). When testing the samples with the Actim Pancreatitis test strip (Medix Biochemica), the urine samples were diluted 40-fold with phosphate-buffered saline (8.0 mmol/L Na₂HPO₄, 1.5 mmol/L KH₂PO₄, 137 mmol/L NaCl, 2.7 mmol/L KCl, pH 7.4) containing 1 g/L bovine serum albumin, protease inhibitors, and preservative to increase the detection limit from 50 µg/L to 2000 µg/L. The test was considered positive when the blue line was detected within 5 min. A control line indicated proper functioning of the strip. The concentration of trypsinogen-2 in the samples was also measured by a quantitative immunoenzymometric assay (in-house assay; Medix Biochemica).

Serum CRP concentrations were determined by an immunoturbidimetric method. The detection limit was 2 mg/L and the reference limit 10 mg/L. The CV was <8.5% (24).

**STATISTICAL ANALYSIS**

Continuous data were compared by the Mann–Whitney U-test. P values <0.05 were considered significant. Agreement between quantitative trypsinogen-2 concentrations and the test strip was evaluated with the κ statistic (κ <0.20 indicates poor agreement, κ = 0.61–0.80 good agreement, and κ >0.81 very good agreement) (25). We compared the accuracy of the different markers by the McNemar test and by logistic regression analysis.
The positive likelihood ratio (PLR) of the urinary trypsinogen-2 dipstick test was 4.8 on admission and 4.2 at 24 h. Comparisons of the results for the dipstick test with the quantitative assay indicate fairly good agreement (Fig. 1). The $H_9260$ value was 0.68 for the trypsinogen-2 test strip (detection limit, 2000 $H_9262$ g/L).

The median serum CRP concentration on admission in severe AP was approximately sevenfold that in mild AP ($P < 0.012$; Table 2). With the predetermined cutoff concentration of 150 mg/L (9, 26), the specificity was 90%, but the sensitivity only 38% (Table 1). At 24 h, the median serum CRP concentration in severe AP was approximately threefold that in mild AP ($P < 0.001$; Table 2), and the sensitivity and specificity were 83% and 70%, respectively. The corresponding PPVs and NPVs are shown in Table 1. On admission, the PLR for CRP was 3.7 and at 24 h it was 2.7, respectively. The accuracy of CRP was also analyzed at cutoffs of 100 mg/L and 125 mg/L, but the best PLR was achieved with the cutoff 150 mg/L (data not shown).

On admission the sensitivity and specificity of the APACHE II score were 52% and 87%, respectively, in predicting severe AP; at 24 h the values were 45% and 86%, respectively. The APACHE II score was 8 at admission.

<table>
<thead>
<tr>
<th>Scoring system</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>PPV, %</th>
<th>NPV, %</th>
<th>PLR</th>
<th>95% CI $^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>At admission</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trypsinogen-2 dipstick, $b$ 2000 $\mu$g/L</td>
<td>62</td>
<td>87</td>
<td>65</td>
<td>85</td>
<td>4.8</td>
<td>2.8–8.2</td>
</tr>
<tr>
<td>Quantitative trypsinogen-2, 2000 $\mu$g/L</td>
<td>62</td>
<td>77</td>
<td>51</td>
<td>84</td>
<td>2.7</td>
<td>1.8–4.1</td>
</tr>
<tr>
<td>Serum CRP $&gt;150$ mg/L</td>
<td>38</td>
<td>90</td>
<td>59</td>
<td>79</td>
<td>3.7</td>
<td>1.9–7.2</td>
</tr>
<tr>
<td>APACHE II $&gt;8$</td>
<td>52</td>
<td>87</td>
<td>61</td>
<td>82</td>
<td>3.6</td>
<td>2.1–6.3</td>
</tr>
<tr>
<td>24 h after admission</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trypsinogen-2 dipstick, $b$ 2000 $\mu$g/L</td>
<td>62</td>
<td>85</td>
<td>62</td>
<td>85</td>
<td>4.2</td>
<td>2.5–7.0</td>
</tr>
<tr>
<td>Quantitative trypsinogen-2, 2000 $\mu$g/L</td>
<td>71</td>
<td>78</td>
<td>57</td>
<td>89</td>
<td>3.4</td>
<td>2.2–5.1</td>
</tr>
<tr>
<td>Serum CRP $&gt;150$ mg/L</td>
<td>83</td>
<td>70</td>
<td>52</td>
<td>91</td>
<td>2.7</td>
<td>2.0–3.7</td>
</tr>
<tr>
<td>APACHE II $&gt;8$</td>
<td>45</td>
<td>86</td>
<td>56</td>
<td>80</td>
<td>3.7</td>
<td>2.1–6.6</td>
</tr>
<tr>
<td>Serum CRP $&gt;150$ mg/L or trypsinogen-2 dipstick$^b$ positive</td>
<td>93</td>
<td>65</td>
<td>51</td>
<td>96</td>
<td>2.6</td>
<td>2.0–3.5</td>
</tr>
<tr>
<td>Serum CRP $&gt;150$ mg/L and trypsinogen-2 dipstick$^b$ positive</td>
<td>51</td>
<td>90</td>
<td>68</td>
<td>83</td>
<td>5.4</td>
<td>2.8–10.4</td>
</tr>
</tbody>
</table>

$^a$ CI, confidence interval.

$^b$ The urine samples were diluted 40-fold, increasing the detection limit to 2000 $\mu$g/L.

Table 1. Sensitivity, specificity, PPVs, NPVs, and PLRs for differentiation between severe (42 patients) and mild AP (108 patients) for trypsinogen-2 dipsticks, the APACHE II score, and serum CRP.

---

The median serum CRP concentration on admission in severe AP was approximately sevenfold that in mild AP ($P = 0.012$; Table 2). With the predetermined cutoff concentration of 150 mg/L (9, 26), the specificity was 90%, but the sensitivity only 38% (Table 1). At 24 h, the median serum CRP concentration in severe AP was approximately threefold that in mild AP ($P < 0.001$; Table 2), and the sensitivity and specificity were 83% and 70%, respectively. The corresponding PPVs and NPVs are shown in Table 1. On admission, the PLR for CRP was 3.7 and at 24 h it was 2.7, respectively. The accuracy of CRP was also analyzed at cutoffs of 100 mg/L and 125 mg/L, but the best PLR was achieved with the cutoff 150 mg/L (data not shown).

On admission the sensitivity and specificity of the APACHE II score were 52% and 87%, respectively, in predicting severe AP; at 24 h the values were 45% and 86%, respectively. The APACHE II score was 8 at admission.

Fig. 1. Correlation between the urinary trypsinogen-2 dipstick tests and the quantitative assay of trypsinogen-2 to assess severity of AP on admission and at 24 h after hospitalization.

Panels A and B describe an assay with a detection limit of 2000 $\mu$g/L. $\cdots \cdots$ indicates the nominal detection limits.
86%, respectively. The corresponding PPVs and NPVs are shown in Table 1. The APACHE II score had a PLR of 3.6 on admission and a PLR of 3.7 at 24 h.

The prognostic accuracy of CRP was compared with that of the dipstick at a cutoff, giving the same sensitivity (62%) for both tests. The dipstick test showed better specificity than CRP on admission (P = 0.00036 with the McNemar test), but at 24 h the difference was not significant (sensitivity of each test, 61%; P = 0.291).

To evaluate whether the combination of CRP and the dipstick test provided more information than either test alone, we analyzed the data by logistic regression. This showed that the combination of the tests provided no advantage over the dipstick tests alone on admission (P = 1.00) or at 24 h after admission (P = 0.424). When either CRP was >150 mg/L or the trypsinogen-2 dipstick was positive at 24 h, 93% of the patients developing severe diseases were detected. On the other hand, when at 24 h CRP was >150 mg/L and the dipstick was positive, the probability that the patients developed severe AP was 90% (Table 1).

**Discussion**

Improved outcome in severe AP is based on early identification of disease severity. In the study of Brivet et al. (1), patients admitted to the intensive care unit after a delay of >24 h had a fourfold risk of dying. Accurate predictors of severity in the early phase of the disease are needed to direct appropriate measures for immediate treatment and subsequent management of these high-risk patients in clinics specialized in the treatment of patients with severe AP. On the contrary, patients accurately identified to have mild disease may be fairly safely treated on low-cost wards, which is important because of limited healthcare resources.

Use of the rapid urinary trypsinogen-2 test on diluted urine samples is a novel way to predict severity in AP. If the test was positive on admission, the PLR of having severe AP was 4.8. A positive trypsinogen-2 result increases the likelihood of severe AP from 28% (pretest probability) to 65% (posttest probability). On admission, PLR for a CRP value >150 mg/L is also fairly good (3.7), but its sensitivity is only 38%. Various prognostic markers to assess the severity of AP have been evaluated. Most of these may be considered experimental, and they cannot be routinely used in emergency settings. In a recent European multicenter study on the prognostic value of TAP, sensitivity 24 h after onset of symptoms was 58%, specificity was 73%, PPV was 39%, and NPV was 86%, respectively, and the PLR was 2.1 (9). At 24 h after admission, the sensitivity was 68%, specificity was 74%, PPV was 44%, NPV was 89%, and PLR was 2.6, respectively. These values suggest that TAP is inferior to the trypsinogen-2 test. Both markers are useful in the early phase of the disease, and their time course profiles are similar (9,27). However, the result of the trypsinogen-2 dipstick test is available within 5 min, whereas TAP requires a laborious ELISA method, which takes several hours and requires skilled laboratory personnel. The rapid urinary trypsinogen-2 test requires no laboratory equipment and can be performed by nursing personnel. In these respects, the rapid trypsinogen-2 measurement shows a clear advantage over urinary TAP.

Although the trypsinogen-2 dipstick test was clearly better than CRP at presentation, approximately one-third of the cases developing severe AP were missed with the cutoff used. However, earlier studies have shown that, when used with undiluted urine and a cutoff of 50 μg/L, the dipstick test identifies all cases of severe AP. Thus, by lowering the cutoff, it will be possible to increase sensitivity, which naturally is associated with loss in specificity. The agreement between the quantitative assay and the test strip result was good (κ = 0.68) (25), but better agreement (κ = 0.92) was obtained in an earlier study with undiluted urine with a cutoff of 50 μg/L for diagnosis of AP (22). The reading of the stick result was considered easy, but we cannot rule out mistakes during the manual dilution of urine.

CRP is currently the most widely used laboratory method for differentiation between severe and mild AP. However, it peaks only ~48–72 h after onset of symptoms (9,28), and in the present study, the sensitivity at admission was only 38%. Larvin (6) suggested that a panel of different tests, such as pancreatic enzymes and inflammatory markers, could be better than individual tests. We estimated the combined value of the dipstick test and CRP by logistic regression. This showed that the accuracy for predicting the outcome was not significantly improved.

**Table 2. Concentrations of urinary trypsinogen-2 and serum CRP in 150 patients with AP.**

<table>
<thead>
<tr>
<th></th>
<th>Mild AP</th>
<th></th>
<th>Severe AP</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>IQR*</td>
<td>Median</td>
<td>IQR</td>
<td>p*</td>
</tr>
<tr>
<td>On admission</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urinary trypsinogen-2, μg/L</td>
<td>220</td>
<td>35–1570</td>
<td>5780</td>
<td>340–19 000</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Serum CRP, mg/L</td>
<td>14</td>
<td>6–60</td>
<td>96</td>
<td>6–230</td>
<td>0.012</td>
</tr>
<tr>
<td>24 h after admission</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urinary trypsinogen-2, μg/L</td>
<td>180</td>
<td>15–1230</td>
<td>5650</td>
<td>1670–17 000</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Serum CRP, mg/L</td>
<td>71</td>
<td>33–170</td>
<td>216</td>
<td>152–332</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

* IQR, interquartile range.

* Mann–Whitney U test.
either at admission or at 24 h. However, a positive result with either the dipstick or a CRP > 150 mg/L within 24 h after admission detected 93% of the cases with severe AP, giving a NPV of 96%. It remains to be determined whether improvement can be achieved by altering the cutoff of the dipstick test by the use of various dilutions of the sample. In clinical practice, high sensitivity (and NPV) is needed to justify treatment of low-risk patients in low-cost wards. On the contrary, a high specificity (and PPV) is desirable when deciding on treatments in intensive care units or specific interventions like endoscopic retrograde cholangiopancreatography in biliary AP.

Multifactorial scoring systems such as the Ranson prognostic signs (11) and APACHE II scoring system are often used in clinical trials to assess the severity of AP. Although APACHE II has the advantage over other clinico-biochemical scoring systems of immediate use on admission and serial measurement, its complexity limits its use in clinical practice (12). In the present study the APACHE II score, with a cutoff of 8, did not reach the accuracy of the dipstick tests in predicting severe AP within 24 h after admission.

In conclusion, the rapid urinary trypsinogen-2 test strip is a promising new way of rapidly identifying patients with severe AP in the early phase of the disease.

Supported by grants from Helsinki University Central Hospital, Sigrid Juselius Foundation, and Stockman’s Foundation. Ulf-Håkan Stenman holds a patent for the use of trypsinogen-2 assays in urine for diagnosis of acute pancreatitis.

References


