Urinary trypsinogen-2 for diagnosing acute pancreatitis: a meta-analysis
Jin T, Huang W, Jiang K, Xiong JJ, Xue P, Javed MA, Yang XN, Xia Q

CRD summary
This review concluded that urinary trypsinogen-2 could potentially be used as a rapid test for the diagnosis of post-ERCP pancreatitis and to an extent for the diagnosis of acute pancreatitis. Lack of a clear reference standard and uncertainty about the reliability of the included studies make the reliability of the conclusions uncertain.

Authors' objectives
To assess the value of urinary trypsinogen-2 in differentiating acute pancreatitis from other acute abdominal disease and in predicting post-endoscopic retrograde cholangiopancreatography (ERCP) pancreatitis.

Searching
PubMed, EMBASE and Web of Science were searched from January 1990 to April 2012. Major gastrointestinal journals were searched. Search terms were reported. Only full-text reports published in English in peer-reviewed journals were eligible for inclusion.

Study selection
Studies that assessed the value of urinary trypsinogen-2 in differentiating acute pancreatitis from other acute abdominal disease or in the prediction of post-ERCP pancreatitis were included. Studies had to report enough data for a 2x2 table to be extracted. Studies had to use patients with extrapancreatic acute abdominal disease as controls. Case reports, reviews and other studies without primary data were excluded. Studies that assessed the prognostic value of urinary trypsinogen-2 in predicting postoperative pancreatitis.

Studies were conducted in European countries (Finland, Italy, Sweden, Turkey and Spain), China and USA. All studies used an immunofluorometric technique to test urinary trypsinogen-2 levels and all except one study used a diagnostic cut-off of 50μg/L. Comparators were serum amylase or serum lipase. The reference standard was not reported for included studies. Patient selection criteria and prevalence of acute pancreatitis were not reported. Studies included patients with mixed aetiologies; biliary and alcoholic origins were the most common causes. Most patients were male and in most cases had mild rather than severe disease. Mean ages ranged from 42 to 70 years.

Three authors independently assessed the studies for inclusion. Disagreements were resolved through consensus.

Assessment of study quality
Two reviewers independently assessed the reporting quality of the studies using the guidelines of the Standards for Reporting of Diagnostic Accuracy (STARD) initiative.

Data extraction
2x2 data on true and false positivity and negativity for diagnosis of either acute pancreatitis or post-ERCP pancreatitis were extracted. These parameters were calculated for serum amylase and lipase where reported. Three independent reviewers performed the data extraction using standardised forms.

Methods of synthesis
Pooled sensitivity, specificity and diagnostic odds ratios (DOR) were calculated, each with 95% confidence intervals. Statistical heterogeneity was evaluated using Cochran's Q test and the I² statistic. A fixed-effect model analysis was used where I² was less than 50% and otherwise a random-effects model was used. Receiver operating characteristics (ROC) were derived and the area under the curve (AUC) was calculated.

Subgroup analyses were conducted for studies with a STARD quality score of at least 16, a sample size of at least 50 patients, admission at up to 72 hours after onset of symptoms, or that assessed prediction of severe acute pancreatitis. Sensitivity analyses were used to explore the impact of excluding each study in turn.

Results of the review
Eighteen studies were included. Fifteen studies assessed the diagnosis of acute pancreatitis. Three studies assessed diagnosis of post-ERCP pancreatitis. Sample sizes ranged from 17 to 156. STARD scores ranged from 14 to 20 for the acute pancreatitis studies and from 13 to 17 for the post-ERCP studies.

**Diagnosis of acute pancreatitis:** There were 14 studies (2,659 patients and 852 cases).

The pooled sensitivity for differentiating acute pancreatitis from other acute abdominal disease was 80% (95% CI 0.77 to 0.82; $I^2 = 87.7\%$) and specificity was 92% (95% CI 0.91 to 0.94; $I^2 = 62.7\%$). The area under the curve was 0.96. The diagnostic odds ratio was 65.63 (95% CI 31.65 to 139.09).

Serum lipase showed better pooled performance on all parameters (nine studies). The pooled specificity, sensitivity and area under the curve of serum amylase were comparable (10 studies) but the diagnostic odds ratio for urinary trypsinogen-2 was reported to be better.

**Diagnosis of post-ERCP pancreatitis:** There were three studies (285 patients).

The pooled sensitivity of urinary trypsinogen-2 for predicting post-ERCP pancreatitis was 86% (95% CI 0.67 to 0.96). Specificity was 94% (95% CI 0.91 to 0.97). Area under the curve was 0.92 and the diagnostic odds ratio was 77.68 (95% CI 24.99 to 241.48). There was no evidence of significant statistical heterogeneity ($I^2=0\%$).

Results of subgroup and sensitivity analyses were also reported.

**Authors’ conclusions**
Urinary trypsinogen-2 could potentially be used as a rapid test for diagnosis of post-ERCP pancreatitis and to an extent for diagnosis of acute pancreatitis.

**CRD commentary**
The review had a clear question supported by appropriate inclusion criteria. Three relevant databases were searched. The limitation to published studies in English may have meant that some relevant studies were not included. The authors reported that they used methods designed to reduce reviewer error and bias at all stages of the review process. They used a checklist that primarily examines reporting quality rather than study bias or generalisability to assess the validity of the included studies. As only summary scores were reported for this tool it was not clear whether they truly reflected the validity of the studies. The reference standard used was not reported for the included studies and neither were patient selection criteria or prevalence of disease. The synthesis appeared reasonable. However, the pooled estimates of sensitivity and specificity may not be very meaningful because the individual study estimates varied widely (particularly for sensitivity).

The authors’ conclusions reflected the evidence presented. However, the unclear quality of the included studies, the lack of key study details and wide variations between study estimates are reasons for uncertainty about the reliability of the review conclusions.

**Implications of the review for practice and research**
**Practice:** The authors stated that urinary trypsinogen-2 may be a useful tool in the early diagnosis of acute pancreatitis in countries where laboratory facilities are sparse.

**Research:** The authors stated that further studies with stringent inclusion criteria and enrolling around 20% of patients with severe disease were required to confirm the usefulness of urinary trypsinogen-2 in the diagnosis of both acute and post-ERCP pancreatitis. The need for larger studies was particularly important in the case of post-ERCP pancreatitis.

**Funding**
National Natural Science Foundation of China; National Institute of Health Research, UK.

**Bibliographic details**
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.