Endoscopic ultrasound: a meta-analysis of test performance in suspected biliary obstruction

CRD summary
The authors concluded that endoscopic ultrasound is accurate, safe and cost-effective, particularly for diagnosing choledocholithiasis. However, safety and costs were not assessed, which makes it impossible to comment on the authors' conclusion concerning these outcomes. Conclusions about the accuracy of endoscopic ultrasound appear to be supported by the review, but possible differences between the studies mean it is difficult to assess their reliability.

Authors' objectives
To assess the diagnostic performance of endoscopic ultrasound (EUS) in diagnosing biliary obstruction and detecting choledocholithiasis and malignant causes of obstruction.

Searching
MEDLINE (January 1987 to May 2006) and the bibliographies of relevant papers were searched. In addition, manual searches of published abstracts and relevant national meetings were undertaken and experts in the field contacted for unpublished papers. The search terms were reported.

Study selection
Study designs of evaluations included in the review
Studies other than case series were eligible for inclusion. Studies were excluded if they used reference standard tests only after positive EUS.

Specific interventions included in the review
Studies were eligible for inclusion if they compared EUS with single or multiple 'gold' standard tests.

Reference standard test against which the new test was compared
Studies were eligible for inclusion if they used one of the following gold standard tests: cholangiography (endoscopic retrograde cholangiopancreatography (ERCP), intra-operative cholangiography, percutaneous transhepatic cholangiography); surgical exploration; positive cytology and histopathology at resection; or a composite gold standard test such as computed tomography plus clinical follow-up of at least 6 months. Most (89%) of the included studies used ERCP; the other studies used surgery, intra-operative cholangiography and clinical follow-up.

Participants included in the review
Studies involving participants with suspected choledocholithiasis, biliary stricture, or a combination of both, were eligible for inclusion. The included studies identified participants with biliary obstruction, choledocholithiasis and malignancy.

Outcomes assessed in the review
Studies were eligible for inclusion if they provided sufficient data to construct 2x2 tables for detection rates of any obstructive biliary pathology, choledocholithiasis and malignancy.

How were decisions on the relevance of primary studies made?
The authors did not state how the papers were selected for the review, or how many reviewers performed the selection.

Assessment of study quality
Two reviewers independently assessed validity according to published criteria described by Irwig et al. These included items on blinding, consecutive participant recruitment, single versus composite gold standard, and whether all patients received gold standard tests. Any disagreements were resolved by a third reviewer.

Data extraction
The authors did not state how the data were extracted for the review, or how many reviewers performed the data
For each study, dichotomous data for each primary outcome were extracted and used to calculate the sensitivity, specificity, and positive and negative predictive values. To prevent problems caused by zero cells, 0.5 was added to each cell in the tables.

**Methods of synthesis**

How were the studies combined?

The sensitivity, specificity, and positive and negative predictive values were pooled and 95% confidence intervals (CIs) calculated using a bootstrap approach. Adjusted values of diagnostic accuracy values were presented. Summary receiver operating characteristic curves were constructed to account for potential variations in diagnostic thresholds. Studies were weighted using the inverse of the log diagnostic odds ratio to account for sample size and level of confidence. Q-points (the point of equal sensitivity and specificity) and area under the curve were also calculated for each outcome. Publication bias was assessed using funnel plots.

How were differences between studies investigated?

Patient characteristics were summarised and presented in tables. Studies assessing choledocholithiasis and malignancy were analysed separately. A linear regression model was used to assess the influence of potential confounding factors on Q-points, including sample size, quality disease prevalence and spectrum, pancreatitis, echoendoscope type and EUS era.

**Results of the review**

Thirty-six studies (n=3,532) were included in the review. Thirty-one studies assessed choledocholithiasis (n=3,075) and nine assessed malignancy (n=555).

Over 60% of the included studies met at least 3 of the 4 validity criteria (Irwig et al.). Twenty-five studies were blinded, 22 studies reported consecutive enrolment, in 33 studies the gold standard reference test was performed in all patients, and in 23 studies a single gold standard was used in all patients. There was no evidence of publication bias from funnel plots.

The overall performance of EUS in detecting biliary obstruction was reported as 88% (95% CI: 85, 91) for adjusted sensitivity and 90% (95% CI: 87, 93) for adjusted specificity. EUS performance was better for detecting choledocholithiasis than for malignancy (without fine-needle aspiration): the sensitivity was 89% (95% CI: 87, 91) versus 78% (95% CI: 69, 85) and the specificity was 94% (95% CI: 91, 96) versus 84% (95% CI: 78, 90).

Findings for Q-point analysis, area under the curve, and positive and negative likelihood ratios were reported. Significant differences in diagnostic test performance between studies were reported for sample size and patient spectrum, but not other evaluated potential confounders; the findings were presented in the review.

**Authors' conclusions**

EUS is accurate, safe and cost-effective for investigating suspected biliary obstruction, particularly the detection of choledocholithiasis. EUS without fine-needle aspiration was less effective for diagnosing malignancy.

**CRD commentary**

The review question was clear and was supported by appropriate inclusion criteria relating to the participants, interventions, comparators, outcomes and study designs. A literature search was undertaken using one database and other appropriate sources, and attempts were made to identify unpublished material. It was unclear whether publications were restricted by language, which means that language bias might have been introduced. Validity was assessed according to published criteria and the results reported. The review process was not always explicitly described and reviewer bias and error cannot be ruled out. The data were pooled and potential sources of heterogeneity investigated. However, information about study details and patient characteristics (reported as being available on the journal’s website) could not be accessed, and it is therefore unclear whether studies were similar enough for pooling to be
appropriate. Adverse effects and costs were not assessed in the review, so it is not possible to comment on the authors’ conclusion regarding the safety or cost-effectiveness of EUS. Conclusions about the accuracy of EUS appear to be supported by the review, but incomplete reporting of review methods and possible differences between the studies mean it is difficult to assess their reliability.

**Implications of the review for practice and research**

**Practice:** The authors stated that EUS without fine-needle aspiration is less accurate in detecting malignancy.

**Research:** The authors did not state any implications for further research.

**Funding**
Not stated.

**Bibliographic details**

**PubMedID**
17478348

**DOI**
10.1016/j.cgh.2007.02.027

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Cholestasis /etiology /ultrasonography; Endosonography; Humans; Neoplasms /complications; Predictive Value of Tests; ROC Curve

**AccessionNumber**
12007001687

**Date bibliographic record published**
10/03/2008

**Date abstract record published**
30/09/2008

**Record Status**
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.