Systematic review of endoscopic ultrasound in gastro-oesophageal cancer

Authors' objectives
To review the literature on the use of endoscopic ultrasound (EUS) for the pre-operative staging of gastro-oesophageal cancer, especially regarding staging performance and staging impact. In addition, evidence was sought on the health economics, therapeutic impact and effect on patient outcome of EUS in any clinical application.

Searching
MEDLINE and ISI (via BIDS)(from 1981 until October 1997), Inside Information Plus (from 1993), the Cochrane Library, EMBASE, SIGLE and FirstSearch were searched for relevant trials; the search terms were provided. Bibliographic listings of all retrieved articles were handsearched. In addition, authors of abstracts, academic centres, leading centres of EUS, manufacturers and an EUS e-mail discussion group were contacted for unpublished information. Only English language trials and full reports were included.

Study selection
Study designs of evaluations included in the review
Diagnostic studies with more than 10 patients, which reported sufficient raw data, study information and the most recent report of a patient group used, were eligible for inclusion.

Specific interventions included in the review
EUS using conventional probes (radial scanning echoendoscopes or linear/curved array technology), EUS-guided fine-needle aspiration (FNA) or miniprobes were eligible for inclusion. The ultrasound frequencies ranged from 7.5 to 20 MHz.

Reference standard test against which the new test was compared
Studies using pathology or histology as the reference standard of diagnosis were eligible for inclusion.

Participants included in the review
Studies of patients with squamous cell carcinoma, or adenocarcinoma located in the oesophagus, stomach or cardia, were eligible for inclusion. Only 5 of the 36 studies provided information on the grouped tumour, node and metastases (TNM) stage.

Outcomes assessed in the review
The included studies were required to report sufficient data for the construction of 2x2 contingency tables. The outcomes reported in the review were staging performance, staging impact, therapeutic impact, patient outcomes and health economics. Staging performance comprised sensitivity, specificity, positive and negative predictive values (PPV and NPV, respectively), accuracy and the odds ratio (OR).

How were decisions on the relevance of primary studies made?
The abstracts were read and a decision was made about the relevance of each study, in terms of the applicability to EUS and the fulfilment of the inclusion and exclusion criteria. The authors did not state how many reviewers performed the selection.

Assessment of study quality
A list of 20 potential biases was drawn up. This included biases due to patient selection (referral, patient filtering or patient cohort bias); biases associated with the application of the 'gold' standard (verification, work-up or incorporation bias); biases due to the measurement of results (disease progression, withdrawal or observer variability bias); and independence of interpretation biases (diagnostic review, tests review, comparator review or clinical review bias). Validity was assessed using a questionnaire developed and tested for this review. The authors did not state how...
many reviewers performed the validity assessment.

**Data extraction**
The main reviewer extracted the data. Data were extracted on article details, study cohort, sample size, clinical description of the participants, homogeneity of the diagnostic application, technical quality, and procedural quality. The data were extracted from the selected studies using data extraction forms. Numerical values of staging performance for the completion of 2x2 contingency table were also extracted. Descriptive summaries were prepared for the other types of study where quantitative analysis was not feasible.

**Methods of synthesis**
How were the studies combined?
The results of each primary study were expressed using the following summary statistics: sensitivity, specificity, PPV, NPV, accuracy and OR. Receiver operator characteristic (ROC) curves for the differentiation of tumour stages T1 and T2 from T3 and T4 were plotted using the methodology developed by Moses et al. and Irwig et al. (see Other Publications of Related Interest nos.1-3). The data synthesis was conducted in three stages: in stage 1, an ROC curve scatterplot was constructed; in stage 2, a summary ROC curve was fitted; and in stage 3, a conventional summary ROC curve with true-positive rate and false-positive rate axes was plotted to summarise the combined results. A summary statistic ($Q^*$), which balanced sensitivity and specificity, was read from the curve. A similar analysis was performed for the discrimination of lymph node stage N0 from stages N1 and above. A quantitative synthesis was not applicable for the studies of staging impact, therapeutic impact, patient outcome, or health economics.

How were differences between studies investigated?
The robustness of the results was investigated using regression techniques to incorporate bias risk and other factors (e.g. use of protocol) into the quantitative analysis.

**Results of the review**
Twenty-seven studies (2,508 participants) evaluated the staging performance of EUS. Seven studied EUS FNA (365 participants) and three studied miniprobes (one also looked at EUS and is also included in the 27 studies; 120 participants).

Twenty-seven primary studies addressing the performance of EUS for the pre-operative staging of gastro-oesophageal cancer satisfied the inclusion criteria.

The performance of EUS in T-staging gastro-oesophageal cancer was $Q^*$ 0.91. $Q^*$ was 0.93 for gastric T-staging and 0.89 for oesophageal T-staging. The value for $Q^*$ was significantly lower for studies performed in the 1990s than for those conducted in the 1980s ($P<0.05$). The presence of stenosis resulting in non-traversability was found to reduce the staging performance of EUS slightly, but significantly ($P<0.05$). Radial probes performed better than linear probes in staging gastric cancer, although, in staging oesophageal cancer there was no significant difference in probe performance.

The performance of EUS in N (lymph node)-staging associated with gastro-oesophageal cancer was $Q^*$ 0.79. $Q^*$ was 0.76 for N-staging associated with gastric cancer and 0.82 for N-staging associated with oesophageal cancer. Studies that reported attempts to perform some form of blinding achieved a significantly better performance than those that did not ($P<0.05$).

There was insufficient information on M staging (staging of metastases) and grouped TNM staging for data synthesis. There was also insufficient information on the use of miniprobes (for subanalysing T1 tumours). There was little information about the use of FNA specifically applicable to gastro-oesophageal cancer.

Eight studies compared the staging performance of EUS with that of incremental computed tomography (CT), but the CT aspects of these were poorly performed and no measure of the staging impact of EUS could be determined. There was very little evidence on therapeutic impact, patient outcomes and health economics.
Cost information
Yes. Two studies contained economic information. One retrospective study reported that EUS is cost-effective, but the analysis presented used the incorrect formulae and did not present marginal cost-effectiveness. The other study also did not perform a formal cost-effectiveness analysis, but reported that the 'EUS first' strategy is more costly; however, clinical data indicated it might be more effective.

Authors' conclusions
EUS is highly effective for the discrimination of stages T1 and T2 from T3 and T4 in both the oesophagus and the stomach. Initial indications are that the performance of T staging at the cardia is less good.

Non-traversable stenosis reduces the staging performance of EUS, but evidence on whether this reduction justifies the risk of dilatation was not available. The studies available on the use of miniprobes reported a high performance for discrimination between mucosal and submucosal cancer. No evidence regarding the subsequent impact of these findings was available. Lymph node staging with EUS has a lower performance than that of tumour staging.

Staging for metastases using EUS alone is not satisfactory.

CRD commentary
This was a good-quality review with clear objectives and selection criteria. The search strategy was extensive and included methods designed to identify grey literature. However, the restriction of the included studies to those reported in English might have resulted in the loss of relevant data. The validity of the included studies was assessed using a comprehensive checklist, specifically developed for the assessment of biases affecting studies of diagnostic tests. The inter-observer variation of this checklist was evaluated prior to its use in the review. Only one reviewer extracted the data, which could have resulted in errors.

The raw data from the included studies were reported in full, thus allowing analyses to be reproduced. The methods used to pool the studies and to assess the potential impact of bias in the primary studies were appropriate, clearly reported and rigorously applied. The subject of the review is of clinical relevance and the conclusions appear to follow on from the results presented.

Implications of the review for practice and research
Practice: EUS is highly effective for the discrimination of stages T1 and T2 from T3 and T4 in both the oesophagus and the stomach.

Research: Well-designed studies that use optimal protocols for both EUS and CT are needed to compare staging performance; such studies must also investigate the complementary use of the modalities. There is also a need for further investigation of the use of FNA in gastro-oesophageal cancer in a study concentrating on lymph nodes. New studies specifically designed to measure staging impact, therapeutic impact and patient outcome are also required, because evidence in these areas is not currently available.

Funding
NHS R&D Health Technology Assessment (HTA) Programme, project number 94/44/03.

Bibliographic details

PubMedID
9728296
Other publications of related interest

Indexing Status
Subject indexing assigned by NLM

MeSH
Age Factors; Child; Child, Preschool; Female; Humans; Infant; Language Development Disorders /diagnosis /epidemiology; Male; Prevalence; Speech Disorders /diagnosis /epidemiology

AccessionNumber
11999008454

Date bibliographic record published
30/06/2005

Date abstract record published
30/06/2005

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.