Staging accuracy of esophageal cancer by endoscopic ultrasound: a meta-analysis and systematic review


CRD summary
This review concluded that EUS had excellent sensitivity and specificity in staging esophageal cancer, performing better in advanced (T4) than early (T1) disease; FNA improved accuracy of N stage assessment. The reliability of the conclusions was limited by poor reporting of the review process and characteristics of the included studies, and the lack of an appropriate quality assessment.

Authors' objectives
To evaluate the accuracy of endoscopic ultrasound (EUS) in the staging of esophageal cancer.

Searching
MEDLINE, CINAHL, ACP journals, HealthStar, DARE, International Pharmaceutical Abstracts and Cochrane Central Register of Controlled Trials were searched; search terms were reported, but search dates were not.

Study selection
Diagnostic accuracy studies of EUS used for T (T1 to T4) or N staging (definitions supplied) of esophageal cancer compared to surgery or appropriate follow-up that reported sufficient data to construct a 2x2 table were eligible for inclusion. All included studies used dedicated EUS machines and surgery as the reference standard; no further population details were provided. The authors did not state how many reviewers performed the study selection.

Assessment of study quality
Completeness of data and inclusion criteria were used as quality criteria on which studies were selected. The authors did not state who applied the validity criteria.

Data extraction
True and false positive and negative results were extracted and 2x2 tables constructed for each study. Two reviewers independently extracted data; differences were resolved by discussion.

Methods of synthesis
Pooled estimates of sensitivity, specificity, likelihood ratios and diagnostic odds ratios with 95% confidence intervals (CI) were calculated using both fixed-effect and random-effects models; those from the fixed-effect model were presented. Sensitivity analyses were conducted where studies were grouped by date of publication as a surrogate for technological advancement. Pooled estimates were calculated separately for EUS alone and EUS with fine needle aspiration (FNA). Heterogeneity was assessed using either the likelihood ratio test (sensitivity and specificity) or Cochran Q (likelihood ratios and diagnostic odds ratios). Summary receiver operating curves (sROC) were plotted and the area under the curve (AUC) calculated. Publication bias was assessed using funnel plots and the Egger and Begg-Mazumdar tests.

Results of the review
Forty-nine studies met the inclusion criteria (n=2,558; range not reported); enrolment was consecutive in 33 studies and prospective in 16.

Pooled estimates of sensitivity and specificity for EUS were: 82 per cent (95% CI: 78%, 85%; 36 studies) and 0.99 (95% CI: 0.99, 1.00; 36 studies) for T1 stage; 81 per cent (95% CI: 78%, 85%; 38 studies) and 96 per cent (95% CI: 95%, 97%; 38 studies) for T2 stage; 91 per cent (95% CI: 89%, 93%; 40 studies) and 94 per cent (95% CI: 93%, 95%; 40 studies) for T3 stage; 92 per cent (95% CI: 89%, 95%; 38 studies) and 97 per cent (95% CI: 97%, 98%; 38 studies) for T4 stage.

For N stage oesophageal cancer, pooled estimates of sensitivity and specificity were: 85 per cent (95% CI: 83%, 86%;
44 studies) and 85% (95% CI: 83%, 86%; 44 studies) for EUS alone; and 97 per cent (95% CI: 92%, 99%; four studies) and 96 per cent (95% CI: 91%, 98%; four studies) for EUS with FNA. Pooled results for diagnostic odds ratios, likelihood ratios and AUC for each stage and results of sensitivity analyses were also presented.

There was no statistically significant heterogeneity between studies for any analysis. The authors stated that there was no evidence for publication bias.

**Authors' conclusions**

EUS had excellent sensitivity and specificity in accurately diagnosing the TN stage of oesophageal cancer. It performed better in advanced (T4) than early (T1) disease. FNA improved accuracy of N stage assessment.

**CRD commentary**

The authors addressed a clear research question. An appropriate reference standard was required for a study to be included. Several appropriate databases were searched, but EMBASE was not included. Neither search dates nor language restrictions were reported. There was no attempt to locate unpublished data. Therefore, publication and language bias may have been introduced. Publication bias was investigated, and the authors state that there was no evidence of such bias, but some of the statistical tests showed significant results for some outcomes at some stages of disease. A relatively large number of studies were included, but the overall number of patients included was relatively small and there was no way of ascertaining the number of patients in any individual study and, therefore, assessing the reliability of the individual study results.

Data extraction was conducted in duplicate, but it was unclear whether similar methods were used to reduce error and bias during study selection. The authors stated that criteria for RCTs were not appropriate for assessing the study designs included, specified only two criteria on which study quality was assessed, and indicated how patients were enrolled in the studies. QUADAS criteria, a validated quality assessment tool for diagnostic accuracy studies, were not applied. The authors stated that there was no statistical heterogeneity between studies, but with the distinct lack of study or population details it was not possible to assess clinical heterogeneity and so the appropriateness of pooling the studies could not be assessed.

Pooled sensitivities were similar for disease stages T3 and T4, and higher than those for T1 and T2 (indicating a lower number of false negative results in later stage disease). Pooled specificities were similar for all disease stages. The pooled positive likelihood ratio was greater in T1 (indicating fewer false positive results with early stage disease). However, confidence intervals for each outcome overlapped across all stages. Therefore, the implication that EUS was better at diagnosing certain stages was unreliable. Further, the implication that EUS was better at detecting the smallest/least invasive (early stage) and largest/most invasive (late stage) cancers may have been more an effect of the reliability of the data than the accuracy of EUS in clinical practice. The conclusion that FNA improved accuracy was based on the results of only four studies of unknown size and quality.

The reliability of the conclusions of this review was limited by poor reporting of the review process and characteristics of the included studies, and the lack of an appropriate quality assessment.

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

The authors did not state implications for practice or research.

**Funding**

Not stated.

**Bibliographic details**


**PubMedID**
Indexing Status
Subject indexing assigned by NLM

MeSH
Endosonography; Esophageal Neoplasms /pathology /ultrasonography; Humans; Neoplasm Staging /methods;
Publication Bias; Sensitivity and Specificity

AccessionNumber
12008105750

Date bibliographic record published
23/12/2008

Date abstract record published
15/04/2009

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.