CRD summary
This review assessed the performance of clinical guidelines in selecting patients for upper endoscopy where endoscopic findings were the reference standard and found that appropriate indication had low specificity for relevant findings/cancer. Cancer detection rate was low in inappropriate indications. The conclusions reflect the data, but should be viewed cautiously due to methodological limitations in the review.

Authors' objectives
To assess the performance of American Society for Gastrointestinal Endoscopy (ASGE) and European Panel on the Appropriateness of Gastrointestinal Endoscopy (EPAGE) guidelines in selecting patients for upper endoscopy (OGD) based on the detection of relevant endoscopic findings (particularly gastro-oesophageal cancer).

Searching
MEDLINE was searched from 1992 (date of publication of ASGE guidelines) to September 2008. Search terms were reported. Bibliographies of included studies were screened for additional articles.

Study selection
Studies of use of ASGE and/or EPAGE guidelines to assess the appropriateness of OGD indication prior to endoscopy were eligible for inclusion. Included studies were required to define relevant endoscopic findings and report prevalence of relevant endoscopic findings and cancer by guideline indication of appropriateness to enable construction of 2x2 tables of appropriateness status compared with endoscopic findings. Relevant endoscopic findings were defined similarly across studies and included gastro-duodenal ulcers, erosive gastro-duodenitis, strictures, cancer, and signs of portal hypertension.

Most included studies (six out of eight) were conducted in Europe. Mean age of study participants, where reported, ranged from 38 to 60 years. The proportion of male participants ranged from 42% to 65%.

Two reviewers independently assessed studies for inclusion. Disagreements were resolved by discussion with a third author.

Assessment of study quality
The authors stated that as all the studies shared the same design (consecutive recruitment of patients with non-selective use of OGD as gold standard), no further assessment of the quality of studies was performed.

Data extraction
Data were extracted on: numbers of participants with relevant endoscopic findings whose OGDs were judged appropriate by the guidelines; numbers of participants with relevant endoscopic findings whose OGDs were judged inappropriate by the guidelines; numbers of participants with gastro-oesophageal cancer whose OGDs were judged appropriate by the guidelines; and numbers of participants with gastro-oesophageal cancer whose OGDs were judged inappropriate by the guidelines. Data were also extracted on the rate of inappropriate OGDs.

Two investigators jointly performed data extraction using predefined forms; any disagreements were resolved through consultation with a third reviewer.

Methods of synthesis
Pooled estimates of the sensitivity and specificity, with 95% confidence intervals (CIs), of guidelines for identifying appropriateness of OGD (which patients will have relevant findings/cancer on OGD) were calculated using a random-effects model. Pooled estimates of positive and negative likelihood ratios (LRs) and a summary receiver operating characteristic (SROC) curve were presented.

Between-study heterogeneity was assessed using $X^2$ and $I^2$ statistics. The following variables were considered as
potential sources of heterogeneity: type of guidelines adopted (ASGE/EPAGE); number of patients included (more/less than 1,000); number of centres included (monocentric/multicentric); and geographical localization of the study (Europe or not).

Egger’s test for funnel plot asymmetry and a Glabraith plot were used to investigate whether publication bias or other small study effects may have affected the results.

**Results of the review**

Eight studies (n=13,856 participants, range 80 to 6,270) were included in the review. In all studies the appropriateness of the indication for OGD was pre-endoscopically assessed by endoscopists; endoscopic findings served as the reference standard. Six studies used ASGE guidelines to assess appropriateness and two studies used EPAGE guidelines.

According the guidelines, 10,643 OGD indications were appropriate and 3,010 (22%) were inappropriate; 203 participants were excluded from the analysis because their indications were not included in the current guidelines.

**Accuracy of appropriateness guidelines for relevant endoscopic findings**: Pooled estimate of sensitivity was 85% (95% CI 84% to 86%). Pooled estimate of specificity was 28% (95% CI 27% to 29%). Pooled positive likelihood ratio was 1.18 (95% CI 1.1 to 1.3). Pooled negative likelihood ratio was 0.6 (95% CI 0.5 to 0.7). There was evidence of significant between-study heterogeneity, which was not explained by any of the variables considered. Egger’s test for publication bias was not significant.

**Accuracy of appropriateness guidelines for cancer**: All studies reported data on cancer detection; 267 carcinomas were detected in the included studies (prevalence=2%). Pooled estimate of sensitivity was 97% (95% CI 94% to 98%). Pooled estimate of specificity was 22% (95% CI 22% to 23%). Pooled positive likelihood ratio was 1.2 (95% CI 1.1 to 1.4). Pooled negative likelihood ratio was 0.2 (95% CI 0.05–0.9). There was evidence of significant between-study heterogeneity, which was not explained by any of the variables considered. Egger’s test for publication bias was significant.

**Authors’ conclusions**

Endoscopic referral for an inappropriate indication seemed to be an inefficient strategy in terms of cancer detection. The relatively high rate of false positive OGDs related with appropriate indications could be reduced by adoption of a more specific prioritisation list based on alarm features.

**CRD commentary**

The review question was defined by appropriate inclusion criteria. The literature search was limited to a single bibliographic database, no attempts to identify unpublished studies were reported and it was unclear whether any language restrictions were applied, which meant that relevant studies may have been omitted from this review. Measures were taken throughout the review process to minimise error and/or bias. Although the design of included studies met a number of key aspects of methodological quality, no formal quality assessment was undertaken and issues such as blinding of endoscopists who performed the reference standard investigation to the appropriateness assessment were not addressed (the authors stated that appropriateness assessment was performed by endoscopists). The presence of significant between-study heterogeneity meant that the estimation of overall estimates of sensitivity, specificity and so on were of questionable value.

Overall, the authors’ conclusions reflect the data presented, but should be viewed cautiously given the limitations of the review.

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

**Practice**: For inappropriate OGD, the very low likelihood of cancer argued against endoscopic referral and the low specificity substantially reduced the predictive value of an appropriate indication for both cancer and relevant endoscopic findings.

**Research**: The authors did not state any recommendations for future research.

**Funding**
None stated.

**Bibliographic details**

**PubMedID**
19497799

**DOI**
10.1016/j.dld.2009.04.019

**Original Paper URL**
http://dx.doi.org/10.1016/j.dld.2009.04.019

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Adult; Age Factors; Endoscopy, Gastrointestinal; Esophageal Neoplasms /diagnosis; False Positive Reactions; Humans; Middle Aged; Patient Selection; Practice Guidelines as Topic; ROC Curve; Referral and Consultation; Stomach Neoplasms /diagnosis

**AccessionNumber**
12010001265

**Date bibliographic record published**
23/06/2010

**Date abstract record published**
03/11/2010

**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.