Accuracy of Helicobacter pylori diagnostic tests in patients with bleeding peptic ulcer: a systematic review and meta-analysis

Gisbert J P, Abraira V

CRD summary
This review assessed the accuracy of diagnostic tests for Helicobacter pylori in patients with upper gastrointestinal bleeding. The review concluded that biopsy methods have high specificity, while the urea breath test has a very high accuracy and the stool antigen test is less accurate. Given the variation between the studies, the review’s findings should be interpreted with some degree of caution.

Authors’ objectives
To assess the accuracy of diagnostic tests for Helicobacter pylori (H. pylori) in patients with upper gastrointestinal bleeding.

Searching
EMBASE, MEDLINE and CINAHL were searched up to May 2005; the search terms were reported. The reference lists of included studies and review articles were checked for additional studies.

Study selection
Study designs of evaluations included in the review
Inclusion criteria were not specified in terms of the study design. Study design was not reported in the review.

Specific interventions included in the review
Studies assessing one or more diagnostic tests for H. pylori were eligible for inclusion. The tests included in the review were the rapid urease test, histology, culture and other biopsy-based methods, the urea breath test, the stool antigen test and serology. For tests that involved biopsy, biopsy sites were either antrum or corpus.

Reference standard test against which the new test was compared
Eligible reference standards had to be based on at least one independent diagnostic method. Most of the included reference standards were based on at least two methods.

Participants included in the review
Eligible participants had a bleeding peptic ulcer as proven by endoscopic examination. The included participants had either duodenal or gastric ulcers; most studies included a combination of both types of ulcer. The majority of studies included some participants (34 to 74%) who were undergoing treatment with non-steroidal anti-inflammatory drugs.

Outcomes assessed in the review
Eligible studies had to report sufficient data to construct a 2x2 table. The outcomes reported in the review were the sensitivity, specificity, and positive and negative likelihood ratios (LRs).

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
The authors did not report that they assessed validity.

Data extraction
Two reviewers extracted data from the studies; any discrepancies were resolved through consensus. The data were
extracted as 2x2 tables of test performance. The sensitivity, specificity, and positive and negative LRs were calculated for each study, along with 95% confidence intervals (CIs).

**Methods of synthesis**

**How were the studies combined?**
The studies were combined using a random-effects model. Pooled sensitivity, specificity, and positive and negative LRs with 95% CIs were calculated.

**How were differences between studies investigated?**
The studies were grouped according to diagnostic test. Statistical homogeneity was assessed visually using a forest plot and using the chi-squared and I-squared statistics. The following subgroup analyses were planned a priori: number of tests used in the reference standard (one or at least two); biopsy site (corpus or antrum); and the presence or absence of blood in the stomach.

**Results of the review**

Twenty-three studies were included in the review.

All of the meta-analyses showed evidence of statistical heterogeneity.

**Rapid urease test** (16 studies, n=1,417).

The sensitivity ranged from 0.41 to 0.94 and the specificity from 0.78 to 1.00. The pooled sensitivity was 0.67 (95% CI: 0.64, 0.70; I-squared 89%) and the pooled specificity 0.93 (95% CI: 0.90, 0.96; I-squared 70%). The pooled positive and negative LRs were 9.6 (95% CI: 5.1, 18.1; p=0.05) and 0.31 (95% CI: 0.22, 0.44; p<0.001), respectively.

**Histology** (10 studies, n=827).

The sensitivity ranged from 0.33 to 0.98 and the specificity from 0.45 to 1.00; mean prevalence was 80% (range: 59 to 100%). The pooled sensitivity was 0.70 (95% CI: 0.66, 0.74; I-squared 93%) and the pooled specificity 0.90 (95% CI: 0.85, 0.94; I-squared 81%). The pooled positive and negative LRs were 6.7 (95% CI: 2.5, 18.4; p=0.05) and 0.23 (95% CI: 0.12, 0.46; p<0.001), respectively. The subgroup analysis showed that heterogeneity was reduced when only samples obtained from both the antrum and corpus were considered.

**Culture and other biopsy-based methods** (3 studies, n=314).

The sensitivity ranged from 0.34 to 0.88 and the specificity from 0.98 to 1.00 (2 studies); mean prevalence was 79% (range: 71 to 100%). The pooled sensitivity was 0.45 (95% CI: 0.39, 0.51; I-squared 95%) and the pooled specificity 0.98 (95% CI: 0.92, 1.00; p=0.52). The pooled positive and negative LRs were 19.6 (95% CI: 4.96; p=0.89) and 0.31 (95% CI: 0.05, 1.9; p<0.001), respectively.

**Urea breath test** (8 studies, n=520).

The sensitivity ranged from 0.86 to 1.00 and the specificity from 0.78 to 1.00; mean prevalence was 70% (range: 32 to 100%). The pooled sensitivity was 0.93 (95% CI: 0.90, 0.95; I-squared 44%) and the pooled specificity 0.92 (95% CI: 0.87, 0.96; I-squared 68%). The pooled positive and negative LRs were 9.5 (95% CI: 3.9, 23.3; p=0.05) and 0.11 (95% CI: 0.07, 0.16; p=0.46), respectively.

**Stool antigen test** (6 studies, n=377).

The sensitivity ranged from 0.74 to 1.00 and the specificity from 0.33 to 0.90; mean prevalence was 63% (range: 42 to 100%). The pooled sensitivity was 0.87 (95% CI: 0.82, 0.91; I-squared 73%) and the pooled specificity 0.70 (95% CI: 0.62, 0.78; I-squared 85%). The pooled positive and negative LRs were 2.3 (95% CI: 1.4, 4; p<0.001) and 0.2 (95% CI: 0.13, 0.3; p=0.63), respectively.
Serology (9 studies, n=803).

The sensitivity ranged from 0.68 to 1.00 and the specificity from 0.11 to 0.84; mean prevalence was 72% (range: 46 to 100%). The pooled sensitivity was 0.88 (95% CI: 0.85, 0.90; I-squared 79%) and the pooled specificity 0.69 (95% CI: 0.62, 0.75; I-squared 77%). The pooled positive and negative LRs were 2.5 (95% CI: 1.6, 4.1; p<0.001) and 0.25 (95% CI: 0.19, 0.33; p=0.32), respectively.

Authors' conclusions
In individuals with bleeding upper gastrointestinal ulcers, biopsy methods such as the rapid urease test, histology and culture have high specificity but low sensitivity. The urea breath test has a very high accuracy, but the stool antigen test is less accurate. Serology seems unaffected by upper gastrointestinal bleeding but is not recommended as a first diagnostic test.

CRD commentary
This review addressed a clear review question. However, the reviewers only searched published data, so the findings might have been affected by publication bias. It was also unclear as to what extent the review methods were open to bias and error in terms of the study selection process, although two reviewers checked the data extraction. The authors also failed to assess the quality of the studies and to report the study designs used, thus it was difficult to assess the reliability of the data.

Given the apparent high degree of statistical heterogeneity between the studies and the clinical heterogeneity in some instances, the reliability of the pooled effect sizes is questionable. Overall, taking into account these limitations, the review findings should be interpreted with some degree of caution.

Implications of the review for practice and research
Practice: The authors stated that treatment with proton-pump inhibitors (PPIs) can affect the accuracy of tests, with the exception of serology; however, H2-receptor antagonists appear to have little effect on test accuracy. The authors recommended the use of noninvasive tests and stated that the urease test should be the method of choice (with ideally a biopsy performed before or immediately after treatment with PPIs to rule out false-negative results). However, other additional noninvasive tests (urea breath test or serology) are recommended before ruling out infection, or in cases where biopsies are not feasible. If a breath test taken while the patient is on PPIs is found to be negative, it should be repeated 2 weeks after stopping the drugs to definitely rule out infection. The stool antigen test is not recommended, nor is serology as a first-line test.

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

Funding
Funded in part by the Instituto de Salud Carlos III, grant numbers C03/02 and G03/90.

Bibliographic details

PubMedID
16494583

DOI
10.1111/j.1572-0241.2006.00528.x

Indexing Status

Subject indexing assigned by NLM

MeSH
Antigens, Bacterial /analysis; Breath Tests; Feces /microbiology; Helicobacter Infections /diagnosis; Helicobacter pylori /isolation & purification; Humans; Peptic Ulcer Hemorrhage /microbiology; Sensitivity and Specificity; Serologic Tests; Urea /analysis; Urease /analysis

AccessionNumber
12006001700

Date bibliographic record published
31/10/2007

Date abstract record published
31/10/2007

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.