Diagnostic performance of urea breath test, rapid urea test, and histology for Helicobacter pylori infection in patients with partial gastrectomy: a meta-analysis

Tian XY, Zhu H, Zhao J, She Q, Zhang GX

CRD summary
This review concluded that histological examination performed best, followed by rapid urea tests, for diagnosis of H. pylori infection after partial gastrectomy. Given the number of limitations of this review and the included studies, and that the studies that evaluated histology seemed to suffer from incorporation bias that would over-estimate accuracy, the conclusion cannot be considered reliable.

Authors' objectives
To determine the performance of different commonly used tests for the diagnosis of Helicobacter pylori infection in patients after partial gastrectomy.

Searching
MEDLINE, EMBASE and Science Direct were searched for studies published in English; search terms were reported. Bibliographies of included articles and contents of major relevant journals (not specified) were searched.

Study selection
Cross-sectional and case-control studies that evaluated the accuracy of at least one test for diagnosing H. pylori in patients who had undergone partial gastrectomy were eligible for inclusion. Studies had to use at least one independent test as the reference standard and report sufficient data to construct 2x2 tables of test performance.

The included studies evaluated the urea breath test, rapid urease test, the stool antigen test and/or histology. All of the included studies used histology as the reference standard, with or without other tests. Most of the included studies used both Bilroth I and Bilroth II techniques during the gastrectomy; the operation was most commonly used to treat gastric cancer and peptic ulcer. Prevalence of H. pylori ranged from 31.3% to 80.3%. Most of the studies were conducted in Asia.

The authors did not state how many reviewers selected studies for the review.

Assessment of study quality
Study quality was assessed by two independent reviewers using the 14-point QUADAS tool.

Data extraction
Data were extracted by two independent reviewers in order to construct 2x2 tables of test performance. Sensitivity, specificity, positive and negative likelihood ratios (LR+/-) and the diagnostic odds ratio (DOR) were calculated along with 95% confidence intervals (CI).

Methods of synthesis
Summary estimates of sensitivity, specificity and likelihood ratios were calculated along with 95% CIs using a frequentist meta-analytical technique. Summary receiver operating characteristic (SROC) curves were produced from these the area under the curve (AUC) and Q-value were calculated. The models used for these analyses were not reported. Only the urea breath test, rapid urease test and histology were included in the analysis. Heterogeneity was assessed using the Q and I² statistics. Spearman’s rank correlation was used to investigate any threshold effect. Meta-regression was used to investigate the impact of study quality on the results.

Results of the review
Nineteen datasets (11 studies) were included in the review.

Urea breath test (nine datasets; 565 patients): The pooled estimate of sensitivity was 77% (95% CI 72% to 82%; I²=84.3%), specificity was 89% (95% CI 85% to 93%; I²=69.8%), LR+ was 6.32 (95% CI 3.22 to 12.42; I²=66.5%),
LR- was 0.27 (95% CI 0.17 to 0.43; \(I^2=79.1\%\)) and DOR was 27.86 (95% CI 13.27 to 58.49; \(I^2=41.1\%\)). AUC was 0.91. Q was 0.84.

**Rapid urea test (seven datasets; 432 patients):** The pooled estimate of sensitivity was 79% (95% CI 72% to 84%; \(I^2=62.3\%\)), specificity was 94% (95% CI 90% to 97%; \(I^2=44.4\%\)), LR+ was 10.21 (95% CI 5.94 to 17.54; \(I^2=0\%\)), LR- was 0.28 (95% CI 0.22 to 0.36; \(I^2=38.4\%\)) and DOR was 49.02 (95% CI 24.24 to 99.14; \(I^2=0\%\)). AUC was 0.93. Q was 0.87.

**Histology (three datasets; 337 patients):** The pooled estimate of sensitivity was 93% (95% CI 88% to 97%; \(I^2=58.0\%\)), specificity was 85% (95% CI 73% to 93%; \(I^2=51.8\%\)), LR+ was 5.88 (95% CI 3.26 to 10.60; \(I^2\) not reported), LR- was 0.09 (95% CI 0.05 to 0.15; \(I^2\) not reported) and DOR was 97.28 (95% CI 34.30 to 275.95; \(I^2\) not reported). The AUC was 0.96 and Q was 0.91.

None of the quality criteria showed statistically significant correlations with the diagnostic odds ratio for any of the index tests. Results of further (apparently post hoc)sensitivity analyses were reported.

**Authors’ conclusions**

Among the three commonly used tests, histological examination performed best followed by rapid urea tests for diagnosis of *H. pylori* infection after partial gastrectomy.

**CRD commentary**

The review addressed a clear research question supported by broad but appropriate inclusion criteria. Several relevant sources were searched. The search was restricted to articles published in English so language and publication biases may have been present. Data extraction and quality assessment were conducted in duplicate but it was unclear whether such steps to avoid selection bias were taken during study selection. Appropriate criteria were used to assess study quality and the results were published in full and taken into consideration in the analysis. Five datasets appeared to suffer from incorporation bias. The quality assessment identified three studies suffering from incorporation bias, two of which did not have the index test listed as part of the reference standard. Therefore, it was unclear which studies suffered from this bias. The conduct of several reference standard tests could indicate the presence of differential verification bias; all the studies passed this QUADAS criterion. Two studies that evaluated the stool antigen test seemed to be excluded from the review; this appeared to be a post hoc decision to exclude tests with fewer than three studies.

The summary estimates of sensitivity and specificity seemed to be derived from frequentist meta-analytical methods. Heterogeneity was moderate to high for most of these analyses. The model used to produce the SROC curves was not reported; it appeared to be the Moses-Littenberg model. More robust SROC models are available from which summary estimates of sensitivity and specificity could have been derived. Results of the threshold tests were not reported, and the relationship between study quality and the diagnostic odds ratio were reported as non-significant correlations. Although some studies assessed more than one index test in the same patients, these direct comparisons were not investigated by the authors. There were some discrepancies in numbers between the text and figures, but these were small and did not alter the clinical interpretation.

There were limitations of this review and the included evidence. These issues together with the three studies that evaluated histology and seemed to suffer from incorporation bias that would over-estimate accuracy mean that the conclusion and recommendations for practice cannot be considered reliable.

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

**Practice:** The authors proposed a diagnostic algorithm using the rapid urease test followed by histology in those with a negative result, followed by a urea breath test if both former tests were negative.

**Research:** The authors stated that studies with larger sample sizes were needed to evaluate the accuracy of histology further.

**Funding**

Natural Science Funds of China.

**Bibliographic details**

PubMedID
22392025

DOI
10.1097/MCG.0b013e318249c4cd

Original Paper URL

Indexing Status
Subject indexing assigned by NLM

MeSH
Breath Tests /methods; Gastrectomy; Helicobacter Infections /diagnosis; Helicobacter pylori /isolation & purification; Humans; Sensitivity and Specificity; Stomach Neoplasms /microbiology /surgery; Urea /analysis; Urease /analysis

AccessionNumber
12012017957

Date bibliographic record published
01/06/2012

Date abstract record published
23/10/2012

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.