Accuracy of diagnostic tests read with and without clinical information: a systematic review

Loy C T, Irwig L

CRD summary
This review assessed whether the provision of clinical information increases the accuracy of reading diagnostic tests. The authors concluded that the current practice of reading diagnostic tests in conjunction with clinical information appears warranted, but that further research is required. These conclusions seem reasonable.

Authors' objectives
To assess whether the provision of clinical information increases the accuracy of diagnostic tests.

Searching
MEDLINE was searched from 1966 to December 2003; the search terms were stated. The reference lists of all included studies were checked. Articles citing the identified studies were examined on ISI Web of Science.

Study selection
Study designs of evaluations included in the review
The included studies were diagnostic accuracy studies.

Specific interventions included in the review
Studies were eligible if they compared diagnostic tests read by the same readers twice, once with and once without clinical information. The studies had to use identical conditions apart from the intervention. The included studies assessed the following diagnostic tests: cytology of bronchial washings; radiographs of the chest, abdomen and bones; mammography; and computed tomography of the head. The included studies used actual or constructed clinical information. In the majority of studies the time between readings of tests was several months.

Reference standard test against which the new test was compared
Studies were included if they used any reference standard. The reference standards employed included, for example, clinical, laboratory, surgery and autopsy follow-up, biopsy and final clinical diagnosis (details of the reference standards used in the included studies were tabulated).

Participants included in the review
Inclusion criteria were not specified in terms of the participants. The included studies used multiple readers to read tests from patients with a variety of medical conditions.

Outcomes assessed in the review
The studies had to assess test accuracy in comparison with a reference standard, and report either the sensitivity and specificity or a receiver operating characteristic (ROC) curve. The review used the area under the ROC curve as an outcome.

How were decisions on the relevance of primary studies made?
One reviewer selected studies for inclusion.

Assessment of study quality
The studies were assessed using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool (see Other Publications of Related Interest), along with additional criteria: actual or constructed clinical information provided; use of balanced study design; time interval between test readings; and whether alternative methods of providing clinical information were considered. The authors did not state how the papers were assessed for quality, or how many reviewers performed the quality assessment.
Data extraction
One reviewer extracted the data and the second reviewer examined articles that were difficult to interpret. Any disagreements were resolved by discussion. The data extracted included study characteristics and test accuracy. Where possible, the area under the ROC curve and the corresponding confidence intervals (CIs) were calculated for each intervention using the standard errors and variances reported in the studies. Data on the sensitivity and specificity were only extracted for studies that did not report data on the area under the ROC curve. Where studies reported subsets of diseases or readers, these were treated as separate entries.

Methods of synthesis
How were the studies combined?
Studies reporting areas under ROC curves and studies reporting only sensitivity and specificity were considered separately. The number of studies reporting significant differences in the area under the ROC curve between tests read with and without clinical information was presented. The sensitivity and specificity values for individual studies were plotted for readings with and without clinical information.

How were differences between studies investigated?
The proportion of studies reporting a significant improvement on area under the ROC curve with clinical information was reported for each type of information (actual versus constructed). Any differences between studies in the area under the ROC curve were discussed with respect to type of radiograph selected (routine or subtle), details in clinical information (comprehensive or short suggestive cues) and specialty of the test reader.

Results of the review
Sixteen studies were included (1,060 tests for ROC curve and 400 tests used for sensitivity and specificity).

Nine of the 16 included studies were conducted by three groups of researchers.

Eleven of the 16 studies used well-described reference standard tests on the entire sample within a reasonable time period. The tests were performed blinded to the results of the reference test, but only one study described methods to ensure the independence of the reference standard from clinical information and test. Fifteen studies included uninterpretable or intermediate results. Ten studies used actual clinical information. Ten studies used a balanced design. Two studies considered alternative methods of providing clinical information.

Area under ROC curve (11 studies with 16 entries): 6 studies reported only the area or the modelled ROC curves; the other 5 studies plotted individual points of sensitivity and specificity on the ROC curves. Nine of the 16 entries found that the addition of clinical information significantly improved the areas under the ROC curve compared with no provision of such clinical information. None of the entries suggested a reduction in the area under the ROC curve when using clinical information.

The results varied among studies. Three of 9 studies using actual clinical information reported significant improvements in comparison with 6 of 7 studies using constructed information.

Sensitivity and specificity (5 studies): all of the studies used actual clinical information. Four of the 5 studies reported that the provision of clinical information improved sensitivity without a loss of specificity.

Authors' conclusions
The current practice of reading diagnostic tests in conjunction with clinical information appears warranted. Further research is required.

CRD commentary
The review question was clear in terms of the intervention and outcomes. The inclusion criteria were broadly defined
in terms of the study design. Restricting the search to one electronic database, references and an internet site might have resulted in the omission of relevant studies. It was not stated whether any attempts were made to minimise language bias. Only one reviewer selected studies and extracted all the data, and this raises the possibility of bias and errors. Quality was assessed using some of the relevant criteria, but there was no discussion of the inter-reader agreement or the validity of the reference standard test.

Given the differences among studies, the methods used to combine the studies appeared appropriate. However, the results were not discussed with study quality taken into consideration. Some potential sources of differences among areas under the ROC curves were discussed, but differences in the sensitivity and specificity were not. The authors' conclusions about the value of adding clinical information to readings of tests appears reasonable.

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

Practice: The authors stated that the current practice of reading diagnostic tests in conjunction with clinical information appears justified.

Research: The authors stated that future studies should examine the most effective method of providing clinical information. Studies should quantify the accuracy of the clinical information used, present ROC curves with identified data points, and examine a wide range of diagnostic tests.

**Funding**

National Medical Research Council, grant number 211205.

**Bibliographic details**


**PubMedID**

15467063

**DOI**

10.1001/jama.292.13.1602

**Original Paper URL**

http://jama.ama-assn.org/

**Other publications of related interest**


**Indexing Status**

Subject indexing assigned by NLM

**MeSH**

Diagnostic Techniques and Procedures; Humans; Medical History Taking; ROC Curve; Reproducibility of Results; Sensitivity and Specificity

**AccessionNumber**

12004008679

**Date bibliographic record published**

30/09/2005
Date abstract record published
30/09/2005

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.