Commercial serological antibody detection tests for the diagnosis of pulmonary tuberculosis: a systematic review


CRD summary
This well-conducted review concluded that the accuracy of commercial tests for the detection of pulmonary tuberculosis varies greatly and none perform well enough to replace sputum smear microscopy. These conclusions are likely to be reliable.

Authors’ objectives
To determine the accuracy of commercial antibody detection tests for the diagnosis of pulmonary tuberculosis (TB).

Searching
PubMed, EMBASE, BIOSIS Previews and Web of Science were searched from 1990 to October 2005 (May 2006 for PubMed). The search terms, which included a diagnostic filter, were listed. The reference lists of primary studies, review articles and textbook chapters were screened and experts in the area were contacted for additional studies. Only studies published in English from 1990 onwards were included.

Study selection
Study designs of evaluations included in the review
Diagnostic accuracy studies of any design (diagnostic cohort or case-control) that included at least 50 participants (25 patients and 25 controls), with prospective or retrospective data collection, were eligible for inclusion.

Specific interventions included in the review
Studies that assessed commercial tests for the serological diagnosis of pulmonary TB were eligible for inclusion. Studies that assessed latent TB, nontuberculous mycobacteria, antibody response during or after TB treatment, and non-immunologic methods for detection of antibodies were excluded, as were basic science studies that focused on cloning new antigens or their immunologic properties, or other new methods of antibody detection. Where reported, most studies carried out antibody detection on frozen sera; one study used fresh sera. The included studies evaluated nine different commercial tests, either alone or in combination: Anda-TB, Detect-TB, ICT TB Tests, Kaolin agglutination tests, MycoDot, Pathozyme TB Complex Plus, Pathozyme Myco, TB enzyme immunoassay, TB glycolipid ISA. Full details of the tests were reported.

Reference standard test against which the new test was compared
Studies in which the reference standard was either isolation of Mycobacterium tuberculosis on culture or the presence of acid-fast bacilli detected by sputum smear microscopy, or a combination of these, were eligible for inclusion. Some studies used alternative reference standards, such as chest radiographs, to assess control patients.

Participants included in the review
Inclusion criteria were not defined in terms of the participants. The participants in the included studies had suspected TB, confirmed TB, nontuberculous respiratory disease or mixed disease, or were contacts of patients with TB and healthy people. All included participants were aged 15 years or over. None of the studies included patients infected with the human immunodeficiency virus. The studies were conducted in Saudi Arabia, Italy, India, Taiwan, Japan, Ghana, Brazil, Spain, Pakistan, Argentina, Egypt, Turkey, Guinea-Bissau and Tanzania.

Outcomes assessed in the review
Studies had to report data separately for smear-positive and smear-negative patients. The outcomes reported in the review were the sensitivity, specificity, area under the receiver operating characteristic curve (AUC) and Q*.

How were decisions on the relevance of primary studies made?
Two reviewers screened studies for inclusion in the review; it is unclear whether this was carried out independently. [A: Two reviewers independently completed screening and study selection.]
Assessment of study quality
Studies were assessed for methodological quality according to the following criteria: comparison of the index test with an appropriate independent reference standard; index test interpreted blind to the results of the reference standard; whole sample or randomly selected sub-sample received the reference standard; and prospective enrolment of consecutive patients with suspected TB. The quality assessment was carried out as part of the data extraction, which was carried out by one reviewer with a second reviewer independently assessing 15% of the included studies.

Data extraction
One reviewer extracted the data onto a piloted data extraction form, while a second reviewer independently extracted data from 15% of the included studies. If necessary, authors were contacted for additional information. If data were reported for different control groups without TB, data were extracted for one control group based on the following order of preference: patients in whom pulmonary TB was initially suspected but subsequently ruled out; patients diagnosed with diseases other than TB; healthy people from endemic countries; contacts of patients with TB; mixed groups from any of the earlier categories; healthy people from non-endemic countries. Where possible, data were extracted as 2x2 tables of test performance and used to calculate the sensitivity and specificity with 95% confidence intervals (CIs).

Methods of synthesis
How were the studies combined?
A summary receiver operating characteristic (SROC) analysis, based on the Moses-Littenberg model, was used to pool the data. SROC curves were constructed and plotted in ROC space together with individual study estimates of sensitivity and specificity. The AUC was calculated as an overall indicator of accuracy. Q*, the point where sensitivity equals specificity, was also calculated as an overall indicator of accuracy.

How were differences between studies investigated?
Forest plots of sensitivity and specificity were used to visually assess heterogeneity. Heterogeneity was investigated using subgroup analysis in which studies were grouped by type of test and then further according to immunoglobulin class and smear status. Pooling of sensitivity and specificity was avoided because of the anticipated heterogeneity of the results.

Results of the review
Twenty-seven publications reporting 68 studies were included in the review. The total number of participants was unclear because the overlap between patient groups was unclear. The median number of TB patients per study was 41 (interquartile range: 38 to 75) and the median number of controls was 45 (interquartile range: 40 to 107).

The data were collected prospectively in 32 studies. Twenty-four studies enrolled patients randomly or consecutively. Thirty-one studies reported that the commercial test results were interpreted blind to the results of the reference standard. The reference standard was independent of the commercial test in all studies. Thirty-nine studies reported that all patients received the same reference standard. In 18 studies the reference standard differed for the cases and controls. Seventeen studies met all four quality criteria. Mycobacterial culture was the reference standard in 51 studies and sputum smear microscopy in 17 studies.

The overall sensitivity ranged from 10 to 90% and specificity from 47 to 100%. The AUC for all tests combined was 0.89 (95% CI: 0.86, 0.92), suggesting moderate accuracy. The variation in accuracy was observed both within and between tests. Although the AUC was higher for patients with smear-positive pulmonary TB (0.90, 95% CI: 0.86, 0.94) than for patients with smear-negative pulmonary TB (0.84, 95% CI: 0.77, 0.91), this difference was not statistically significant. Studies that used healthy controls reported higher specificity (range: 86 to 100%) than those using patients with nontuberculous respiratory disease (range: 47 to 100%).

Authors' conclusions
The accuracy of commercial tests for the detection of pulmonary TB varies greatly and none perform well enough to replace sputum smear microscopy.
CRD commentary
This was generally a well-conducted and clearly reported review. The review question was focused and inclusion criteria were clearly defined. A number of relevant databases were searched, but the search was limited to studies published in English and included a diagnostic filter. It is therefore likely that relevant studies have been missed and the review may be subject to language and publication bias. Methodological quality was assessed using appropriate criteria and the results reported. Some steps were taken to minimise bias and errors in the review process, but the involvement of two independent reviewers at all stages and for all papers would have been preferable. Study details were tabulated, but further details about the participants in the included studies would have helped determine the generalisability of the results. The methods used to synthesise the results were appropriate and the inclusion of several graphical displays clearly portrayed the results of the included studies. Although the review suffered from some minor limitations, particularly in relation to the literature search, it is unlikely that improvements in these areas would have changed the results of this review. The authors’ conclusions are therefore likely to be reliable.

Implications of the review for practice and research
Practice: The authors stated that commercial tests have little role in the diagnosis of pulmonary TB.

Research: The authors stated that research into the discovery of new antigens with immunodiagnostic potential should be intensified.

Funding

Bibliographic details

PubMedID
17564490

DOI
10.1371/journal.pmed.0040202

Original Paper URL
http://medicine.plosjournals.org/perlserv/?request=get-document&amp;doi=10.1371%2Fjournal.pmed.0040202

Indexing Status
Subject indexing assigned by NLM

MeSH
Adult; Agglutination Tests; Antibodies, Bacterial /blood; Blood Preservation; Blotting, Western; Child; Comorbidity; Developing Countries; Enzyme-Linked Immunosorbent Assay /methods /standards; HIV Infections /epidemiology; Humans; Immunoglobulin G /blood; Kaolin; Mycobacterium tuberculosis /growth & development /immunology /isolation & purification; Predictive Value of Tests; Reagent Kits, Diagnostic; Reproducibility of Results; Research Design; Sensitivity and Specificity; Sputum /microbiology; Tuberculosis, Pulmonary /blood /diagnosis /epidemiology /immunology

AccessionNumber
12007008209

Date bibliographic record published
08/11/2007

Date abstract record published
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.