Diagnostic accuracy of nucleic acid amplification tests for tuberculous meningitis: a systematic review and meta-analysis
Pai M, Flores L L, Pai N, Hubbard A, Riley L W, Colford J M

CRD summary
This well-conducted review aimed to determine the accuracy of nucleic acid amplification (NAA) tests for tuberculous meningitis. The authors concluded that, based on current evidence, commercial NAA tests have the potential to confirm a diagnosis of tuberculous meningitis. However, their low sensitivity means that they cannot be used to rule out tuberculous meningitis with certainty. These conclusions appear reliable.

Authors' objectives
To determine the accuracy of nucleic acid amplification (NAA) tests for tuberculous meningitis.

Searching
MEDLINE (1985 to 2002), EMBASE (1988 to 2002), Web of Science (1990 to 2002), BIOSIS Previews (1993 to 2002), the Cochrane Library (2002) and LILACS (1990 to 2002) were searched; the search terms were listed. Additional attempts to locate studies were made through contact with experts and the test manufacturers, and the screening of reference lists of retrieved studies. Studies of all languages were eligible for inclusion. Conference abstracts were excluded.

Study selection
Study designs of evaluations included in the review
No inclusion criteria relating to the study design were specified.

Specific interventions included in the review
Studies of NAA tests in cerebrospinal fluid (CSF) specimens were eligible for inclusion. The tests were classified as in-house or commercial. The commercial tests evaluated were Amplicor and Cobas Amplicor (Roche Molecular Systems), MTD (Gen-Probe Inc.) and LCx (Abbott Laboratories). The in-house tests were all based on polymerase chain reaction (PCR) or nested PCR.

Reference standard test against which the new test was compared
The studies had to include a reference standard for the detection of M.tuberculosis in CSF specimens to be eligible for inclusion, although no specific reference standard was specified. The reference standards reported by included studies were classed as: class I - microbiological diagnosis alone (culture alone or culture and/or microscopy); class II - combination of microbiological and clinical diagnosis; and class III - clinical diagnosis, response to therapy and other laboratory tests (included data such as clinical features, CSF analyses, imaging studies, history of contact with tuberculosis, presence of extrameningeal tuberculosis, and response to tuberculosis therapy).

Participants included in the review
Studies that included at least 10 CSF samples were eligible for inclusion.

Outcomes assessed in the review
The studies had to report sufficient data to calculate the sensitivity and specificity.

How were decisions on the relevance of primary studies made?
Two independent reviewers assessed studies for eligibility. Any disagreements were resolved by consensus.

Assessment of study quality
The studies were assessed for the following methodological features: study design (case-control versus cross-sectional...
or cohort); nature of the reference standard; blinding (single- or double-blinded versus unblinded interpretation of test and reference standard results); potential for verification bias (complete versus partial verification of test results by reference standard); patient sampling method (consecutive or random versus non-consecutive and non-random); and prospective data collection. Two independent reviewers, blinded to publication details, assessed the validity of the studies. Any disagreements were resolved by consensus.

Data extraction
Two independent reviewers, blinded to publication details, extracted the data using a piloted data extraction form. Any disagreements were resolved by consensus. The authors used a hierarchical approach to select one pair of sensitivity and specificity estimates from each study: if the study used a class II reference standard these data were preferentially included, followed by data from studies that used a class I reference standard, and lastly those that used a class III reference standard.

Methods of synthesis
How were the studies combined?
Random-effects models were used to pool the sensitivity, specificity, positive likelihood ratios (LRs), negative LRs and diagnostic odds ratios (DORs). Summary receiver operating characteristic curves were calculated to account for the correlation between sensitivity and specificity. Summary LRs, which are generally considered to be easier to interpret in practice, were selected as the primary outcome measure. Publication bias was investigated by producing funnel plots and applying the Egger test (see Other Publications of Related Interest).

How were differences between studies investigated?
Chi-squared tests were used to test for the presence of heterogeneity. Heterogeneity was investigated using stratified subgroup analyses. The following factors were specified a priori as potential sources of heterogeneity: commercial versus in-house tests and individual quality items. The effect of the reference standard on estimates of accuracy was investigated by pooling results for all studies and for only those that used a class I or II reference standard (i.e. excluding those studies that used a clinical reference standard).

Results of the review
Forty-five studies reporting 49 evaluations were included in the review. Fourteen assessed commercial tests and 35 assessed in-house tests. The median study size was 42 specimens or individuals (range: 15 to 392).

The studies of commercial tests were all cross-sectional and 86% collected data prospectively. All were published after 1996 and 86% involved small numbers (less than 10) of specimens from patients with confirmed tuberculous meningitis. Around half of the studies of in-house tests were cross-sectional and prospective. In addition, about half of the studies of in-house tests were published before 1996, and about 49% involved small numbers of specimens from tuberculous meningitis patients.

Commercial tests (n=14).

The pooled positive LR was 35.1 (95% confidence interval, CI: 19.0, 64.6). There was little evidence of heterogeneity between the studies. However, the pooled negative LR was poor (0.44, 95% CI: 0.33, 0.60) and there was some evidence of heterogeneity (P=0.07). The pooled values did not appear to be affected by the type of reference standard used.

In-house tests (n=35).

There was strong evidence of heterogeneity in both the positive and negative LRs (P<0.001). The pooled positive LR was 11.5 (95% CI: 6.8, 19.7), while the pooled negative LR was slightly better than that for the commercial tests (0.21, 95% CI: 0.11, 0.40). The authors stressed the problems associated with interpreting these summary values given the heterogeneity between studies. Heterogeneity was investigated using a ratified analysis that involved pooling the DOR separately for studies that did and did not meet specific quality criteria. Case-control studies produced DORs twice as large as those obtained from cross-sectional studies. Studies that did not use blinding produced DORs almost
twice as large as those that did.

There was strong evidence of publication bias, both for the commercial tests (P=0.01) and the in-house tests (P=0.02). This was supported by the funnel plots, which showed asymmetry.

**Authors' conclusions**
Based on current evidence, commercial NAA tests show a potential role in confirming the diagnosis of tuberculous meningitis, although their overall low sensitivity precludes their use to rule out tuberculous meningitis with certainty.

**CRD commentary**
This was an excellent review of the area. Appropriate methods for systematic reviews of test accuracy studies were followed, full details of the review process were reported, and an appropriate synthesis was undertaken. The authors' conclusions are supported by the results presented.

**Implications of the review for practice and research**
Practice: The authors stated that current evidence suggests a role for NAA tests in confirming a diagnosis of tuberculous meningitis. The results of their review suggested that a negative NAA test should not be used alone as justification to discontinue therapy. This finding underscores the need to use NAA tests in combination with other tests such as culture and smear. The clinical implications for institutions that use in-house PCR tests are much less clear. Since the accuracy of in-house tests are widely variable, clinicians may have to rely on research data from their own institutions to produce clinically useful estimates of test accuracy.

Research: The authors stated that there is a lack of data on the incremental gain of using NAA tests over and above the diagnostic performance achieved by the use of conventional methods alone.

**Bibliographic details**

**PubMedID**
14522262

**Other publications of related interest**

**Indexing Status**
Subject indexing assigned by NLM

**MeSH**
Humans; Mycobacterium tuberculosis /genetics; Nucleic Acid Amplification Techniques /standards; Odds Ratio; Predictive Value of Tests; Sensitivity and Specificity; Tuberculosis, Meningeal /diagnosis

**AccessionNumber**
12003002116

**Date bibliographic record published**
28/02/2006

**Date abstract record published**
28/02/2006

**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.