Nucleic acid amplification tests for the diagnosis of tuberculous lymphadenitis: a systematic review

Daley P, Thomas S, Pai M

CRD summary
This review found that studies assessing nucleic acid amplification tests for tuberculous lymphadenitis showed inconsistent results. Although these conclusions are supported by the data, they should be interpreted with caution given the likelihood that relevant studies were missed by the searches. A more sophisticated analysis might have produced more meaningful estimates of accuracy.

Authors’ objectives
To determine the accuracy of nucleic acid amplification tests (NAAT) for the diagnosis of tuberculous (TB) lymphadenitis.

Searching
A database produced for previous reviews of NAAT for TB diagnosis was used as an initial source of studies. This database was produced by searching MEDLINE (1985 to August 2002), EMBASE (1988 to August 2002), Web of Science (1990 to August 2002), BIOSIS Previews (1993 to August 2002), the Cochrane Library (August 2002) and LILACS (1990 to August 2002). The search terms, which were reported, included a diagnostic filter. The MEDLINE search was updated to July 2007 but was restricted by setting the clinical queries filter to ‘diagnosis’ and ‘broad/sensitive’. The Indian Journal of Tuberculosis was handsearched from 1990 to 2006. Only studies reported in English were included.

Study selection
Studies that used an NAAT (commercial or in-house) on lymph node samples from at least 10 patients with suspected TB lymphadenitis, and that included a reference standard, were eligible for inclusion. Studies that used clinical information alone as the reference standard were excluded. Two additional studies were excluded based on the reference standard: in one the reference standard was response to therapy but not all patients were treated, while in the other the reference standard was biopsy but data were only reported for four biopsy specimens.

The included studies were prospective and retrospective and used both diagnostic cohort and case-control designs. They used commercial (Roche Amplicor, Abbott LCx, Gen probe MTDT, BD Probe Tec) and in-house polymerase chain reaction NAAT tests, and fine-needle aspiration, biopsy or paraffin-embedded tissue samples or a combination of these specimens. The outcomes reported were the sensitivity, specificity and diagnostic odds ratios (DOR).

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 studies were assessed for methodological quality using 11 of the 14 Quality Assessment of Diagnostic Accuracy Studies (QUADAS) criteria. It appears that the studies were assigned a score based on the number of QUADAS items fulfilled.

The authors did not state how many reviewers performed the quality assessment.

Data extraction
Each set of 2x2 data was considered as a separate study, e.g. studies reporting more than one NAAT comparison or more than one specimen type. Inhibitory results were considered as failed rather than negative. A variety of reference standards were used and were assigned to the following hierarchy: culture > smear > histology > response to therapy. If studies applied more than one reference standard then data for the most valid were selected. If data were reported at both the patient and sample level, sample data were included. Data were extracted as the sensitivity and specificity.
DORs were also calculated. Where cells contained a 0 value, 0.5 was added to these cells to prevent division by 0.

Two reviewers independently extracted the data; any disagreements were resolved through consensus.

**Methods of synthesis**

Studies reporting data on sensitivity and specificity were plotted on forest plots together with 95% confidence intervals (CIs) around these estimates. Pooled DORs were calculated using DerSimonian and Laird random-effects models. A summary receiver operating characteristic (SROC) curve was estimated; the methods used to generate such curves were not reported.

Heterogeneity in estimates of accuracy was assessed using the $I^2$ statistic. Values greater than 50% were considered to indicate substantial heterogeneity. Subgroup analysis, where DORs were pooled separately according to subgroup, was carried out to investigate the effects of the following: type of reference standard (incorporated culture or not), study design (cross-sectional versus case-control), blinded interpretation of the NAAT and reference standard, type of NAAT used (commercial versus in-house), specimen examined (fine-needle aspiration or biopsy), volume of template used in NAAT ($\geq$20 μL or <20 μL), use of discrepant analysis and use of IS6110 locus as target for amplification.

**Results of the review**

Thirty-six studies reporting 49 sets of 2x2 data were included (1,829 samples).

The QUADAS scores ranged from 3 to 10 out of 11. The quality of reporting was generally poor.

The sensitivity ranged from 2 to 100% and the specificity from 28 to 100%. There was strong evidence of heterogeneity in both sensitivity and specificity ($I^2$$>86\%$).

Studies using commercial NAATs were more accurate than studies using in-house NAATs (DOR 45.5 versus 7.2). Studies using a volume of template $>\geq$20 μL were more accurate than those using a lesser volume (DOR 63.0 versus 9.0). Studies using discrepant analysis showed higher accuracy than those using conventional analysis (DOR 212 versus 9.9). None of the other subgroups investigated showed significant differences between the groups. Heterogeneity remained for all subgroup analyses conducted ($I^2=35\%$ to $I^2=87\%$; $p<0.001$ to $p=0.1$).

**Authors' conclusions**

The results from studies assessing NAAT for TB lymphadenitis are highly variable and inconsistent. It was not possible to produce clinically meaningful estimates of accuracy.

**CRD commentary**

The review addressed a focused question and inclusion criteria were reported. However, these were somewhat confusing, especially in relation to the population and study design: the inclusion criteria suggested that only studies in patients with suspected disease were eligible for inclusion (i.e. diagnostic cohort studies), but diagnostic case-control studies were also included. Some studies were excluded at a later stage of the review for reasons relating to the reference standard, which do not appear to have been pre-specified. The literature search was based on an extensive search conducted for previous reviews on a similar topic, but it included a diagnostic filter. The updated search for the 5 years since the previous reviews were conducted was limited to MEDLINE and was restricted by the use of the clinical queries 'diagnosis' filter. Relevant studies are therefore likely to have been missed, especially those published more recently. No attempts were made to locate unpublished studies and the review was restricted to studies in English, so there is the possibility of language and publication bias. Although appropriate criteria were used to assess study quality, it is not clear exactly which criteria were applied as three QUADAS items were omitted. The number of QUADAS items fulfilled was used to generate a summary quality score which has been shown to be inappropriate for this tool; the results for individual items were not shown, thus the quality of the included studies remains unclear. Appropriate steps were taken to minimise bias and errors in the extraction of data, but it is unclear whether such steps were also applied at the study selection and quality assessment stages. The methods used to synthesise the results were acceptable but the use of more sophisticated methods, such as the bivariate or hierarchical SROC models, might have
been more informative. Although the authors’ conclusions are supported by the data presented, a more sophisticated analysis may have allowed more clinically meaningful estimates of accuracy to have been produced. The conclusions should be interpreted with caution given the likelihood that relevant studies were missed by the searches.

Implications of the review for practice and research
Practice: The authors stated that NAATs need to be applied in conjunction with conventional methods and interpreted in the context of clinical suspicion.

Research: The authors stated that future studies should assess the incremental yield of NAAT, over and above that obtained by conventional tests. Future studies should also address whether NAAT yield can be increased by changing specimen type and template volume, and whether fine-needle aspiration can be used instead of biopsy. Studies should define reference standards and standardise laboratory protocols, use appropriate controls, assess 'real time' NAAT and isothermal amplification techniques, and evaluate newer and simpler NAAT methods.

Funding
Not stated.

Bibliographic details

PubMedID
17958977

Indexing Status
Subject indexing assigned by NLM

MeSH
Humans; Nucleic Acid Amplification Techniques; Sensitivity and Specificity; Tuberculosis, Lymph Node /diagnosis

AccessionNumber
12007004141

Date bibliographic record published
09/08/2008

Date abstract record published
03/11/2008

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.