Diagnostic accuracy of adenosine deaminase in tuberculous pleurisy: a meta-analysis
Liang Q L, Shi H Z, Wang K, Qin S M, Qin X J

CRD summary
The review produced summary measures of test accuracy, using heterogeneous data sets, from studies with a range of diagnostic thresholds. The authors concluded that adenosine deaminase assays are likely to be useful in tuberculosis pleurisy diagnosis. The wide range of sensitivity and specificity values reported by included studies, and weaknesses in the meta-analytical methods, limit the reliability of these conclusions.

Authors' objectives
To determine the accuracy of adenosine deaminase in pleural fluid for the early diagnosis of tuberculosis pleurisy.

Searching
MEDLINE, EMBASE, Web of Science and the Cochrane databases were searched from inception to March 2007. Search terms were reported and included methodological terms (relating to test accuracy studies). Additional studies were sought by searching the bibliographies of included studies and through contacts with experts in the field. Only studies published in English were included.

Study selection
Included studies were required to include a minimum of ten specimens and to report the sensitivity and specificity of adenosine deaminase for the diagnosis of tuberculosis pleurisy, or sufficient data to calculate these parameters. Conference abstracts and letters were excluded.

Adenosine deaminase assay methods were reported as Guisti or non-Guisti. Diagnostic thresholds ranged from 20 to 71 IU/L. The reference standards used, where reported, were bacterial and/or histological studies, or clinical course (including pleural fluid analysis, radiology, presentation and response to treatment).

Two reviewers independently assessed studies for inclusion and disagreements were resolved by consensus.

Assessment of study quality
The methodological quality of included studies was assessed using STARD (Standards for Reporting of Diagnostic Accuracy) and QUADAS (Quality Assessment of Diagnostic Accuracy Studies Assessment). Overall quality scores were presented for both, with maximum values of 25 and 14, respectively. Studies scoring at least 10 on QUADAS, or at least 11 on STARD, were considered to be of high quality.

Two blinded reviewers independently assessed study quality and disagreements were resolved by consensus.

Data extraction
The assay method (Guisti or non-Guisti), diagnostic threshold and true positives, false positives, false negatives and true negatives were extracted for each study. Sensitivity, specificity, positive and negative likelihood ratios (LRs) and diagnostic odds ratio (DOR) were calculated for each study.

Data were extracted independently by two blinded reviewers and disagreements were resolved by consensus.

Methods of synthesis
A summary receiver operating characteristic (sROC) curve was constructed using the method of Moses et al. and a random effects model was used to generate pooled estimates of sensitivity, specificity and other parameters, with 95% confidence intervals (CIs).

Heterogeneity was assessed using the $\chi^2$ and Fisher's exact tests.
Univariate regression analyses (weighted by inverse variance) were used to assess the impact of quality scores and other study characteristics on test performance, as indicated by relative diagnostic odds ratio (RDOR).

Publication bias was assessed using funnel plots and the Egger test.

**Results of the review**

Sixty three studies, with a total of 8,036 participants, were included in the review. The mean sample size was 138 (range 28 to 600). STARD quality scores ranged from 5 to 11, with three studies scoring 11. QUADAS quality scores ranged from 4 to 10, with one study scoring 10.

The pooled estimate of sensitivity was 0.92 (95% CI: 0.90, 0.93) and 0.90 (95% CI: 0.89, 0.91) for specificity. The ranges reported by included studies were 0.47 to 1.00 for sensitivity and 0.41 to 1.00 for specificity. The pooled estimate of positive likelihood ratios was 9.03 (95% CI: 7.19, 11.35) and 0.10 (95% CI: 0.07, 0.14) for negative likelihood ratios. The pooled estimate of diagnostic odds ratio was 110.08 (95% CI: 69.96, 173.20), which indicates a high level of overall accuracy. There was evidence of significant heterogeneity (p<0.001) for all parameters.

From the sROC curve, the maximal joint sensitivity and specificity was estimated to be 0.91, indicating relatively high overall accuracy.

Regression analyses indicated that quality score, assay method, diagnostic threshold (>40 IU/L versus 40 IU/L or less), adequacy of the reference standard and characteristics of the non tuberculosis pleurisy group did not affect diagnostic accuracy.

Tests indicated a potential for publication bias.

**Authors’ conclusions**

Adenosine deaminase is a relatively sensitive and specific test for the diagnosis of tuberculosis pleurisy. Current evidence suggests a potential role for adenosine deaminase assays in confirming a diagnosis of tuberculosis pleurisy, when used alongside clinical findings and the results of microbiological examinations and biopsy.

**CRD commentary**

The review addressed a clearly stated research question, though inclusion criteria were limited to outcome measures and sample size. A wide range of sources were searched to identify studies for inclusion. However, the limitation to English language studies and the use of methodological search terms for test accuracy studies may have resulted in the omission of relevant data and left the review susceptible to language bias. The review process was clearly reported and included measures to minimise error and bias. The methodological quality of included studies was assessed using validated tools, but results were then inappropriately used to generate overall quality scores (using QUADAS to generate summary quality scores has been shown to be invalid). The results of included studies were clearly reported but the generation of pooled estimates of accuracy measures is of doubtful value, given the significant heterogeneity present and the apparent variation in diagnostic threshold across studies. This is reflected in the wide confidence intervals. The generation of an sROC curve was a more useful approach, although the model used is not currently recommended as the optimum method. Regression analyses were used to investigate the impact of potentially relevant variables upon accuracy but better methods of conducting these analyses (which are able to assess impact on sensitivity and specificity individually rather than using RDOR) are available. The use of overall quality scores in regression analyses can mask the effects of individual components of study quality. The range of sensitivity and specificity values reported by the included studies, along with limitations in the methods of analysis, mean that the authors’ conclusions regarding the likely usefulness of adenosine deaminase assays should be treated with caution.

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

Practice: The measurement of adenosine deaminase in pleural effusion is likely to be useful in the diagnosis of tuberculosis pleurisy. The results of adenosine deaminase assays should be interpreted in parallel with clinical findings and the results of other tests.

Research: The authors make no recommendations for future research.
Funding
National Natural Science Foundation of China, grant number 30660064; Natural Science Foundation of Guangxi Zhuang Autonomous Zone, grant number 0639044; New Century Excellent Talents in Chinese Universities, programme NCET-04-0835.

Bibliographic details

PubMedID
18222681

DOI
10.1016/j.rmed.2007.12.007

Other publications of related interest


Indexing Status
Subject indexing assigned by NLM

MeSH
Adenosine Deaminase /analysis; Biomarkers /analysis; Clinical Enzyme Tests; Humans; Mycobacterium tuberculosis; Odds Ratio; Pleural Effusion /enzymology; ROC Curve; Sensitivity and Specificity; Tuberculosis, Pleural /diagnosis

AccessionNumber
12008106440

Date bibliographic record published
03/02/2009

Date abstract record published
10/06/2009

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.