Diagnostic value of adenosine deaminase in cerebrospinal fluid for tuberculous meningitis: a meta-analysis

H-B. Xu,* R-H. Jiang,† L. Li,* W. Sha,† H-P. Xiao†
*Department of Clinical Pharmacy, Shanghai Tenth People’s Hospital, Tongji University School of Medicine, Shanghai, †Department of Tuberculosis, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai, China

OBJECTIVE: To determine the accuracy of adenosine deaminase (ADA) measurements in the diagnosis of tuberculous meningitis (TBM).

DESIGN: After a systematic review of English language studies, the sensitivity, specificity and accuracy of ADA concentrations in the diagnosis of cerebrospinal fluid (CSF) were evaluated using random effects models. Summary receiver operating characteristic curves were used to summarise overall test performance.

RESULTS: Ten studies met our inclusion criteria. The sensitivity of ADA in the diagnosis of TBM was 0.79 (95%CI 0.75–0.83), specificity 0.91 (95%CI 0.89–0.93), positive likelihood ratio 6.85 (95%CI 4.11–11.41), negative likelihood ratio 0.29 (95%CI 0.19–0.44) and diagnostic odds ratio 26.93 (95%CI 12.73–56.97).

CONCLUSION: Our data suggest that ADA in the CSF can be a sensitive and specific target and a critical criteria for the diagnosis of TBM.

KEY WORDS: adenosine deaminase; cerebrospinal fluid; tuberculosis

TUBERCULOSIS (TB) is a major cause of morbidity and mortality, the global incidence of which is increasing by 0.4% per annum.1 Tuberculous meningitis (TBM) is one of the most harmful infectious diseases, and accounts for 1% of all forms of TB. About 30% of TBM patients die despite anti-tuberculosis chemotherapy, and early institution of treatment is essential for satisfactory improvement.2,3 The diagnosis of TBM is difficult due to the low sensitivity of acid-fast bacilli (AFB) microscopy using Ziehl-Neelsen staining of the cerebrospinal fluid (CSF), and defining the growth of Mycobacterium tuberculosis bacilli is time consuming.4,5 A quick, non-invasive test to confirm TBM would therefore be helpful.

Adenosine deaminase (ADA) is an enzyme widely distributed in tissues and body fluid used in the diagnosis of TB in pleural, meningeal and pericardial fluids.6 The measurement of ADA was initiated by Giusti in 1981 and applied extensively in clinical practice.7 Although its use has become popular in many countries, there is no consensus regarding its current usefulness in clinical practice.

We performed a systematic review of the literature and a meta-analysis to determine the usefulness of ADA measurement in the diagnosis of TBM.

MATERIALS AND METHODS

Search strategy and study selection
As the present study was a meta-analysis based on published articles, we did not require patient consent or the approval of institutional review boards. We searched the following electronic databases: Medline (1966–2009), Embase (1980–2009), Web of Science (1990–2009), BIOSIS (1994–2009) and LILACS (1980–2009). All searches were up to date as of May 2009. The search terms used included ‘tuberculosis’, ‘Mycobacterium tuberculosis’, ‘meningitic/meningitis’, ‘cerebrospinal effusion/cerebrospinal fluid’, ‘adenosine deaminase/ADA’, ‘sensitivity and specificity’ and ‘accuracy’. We contacted experts in the speciality and searched reference lists from primary and review articles. Articles published in English were reviewed and analysed, while conference abstracts and letters to journal editors were excluded due to their limited data content.

Studies were included in the present meta-analysis if they provided both the sensitivity (true-positive rate) and the specificity (false-positive rate) of ADA in the diagnosis of TBM, or if they provided ADA values in a dot-plot form, allowing test results to be extracted for individual study subjects. As very small studies may be vulnerable to selection bias, only studies including at least 10 TBM specimens were selected for inclusion in the study. Two reviewers (HBX and RHJ) independently judged study eligibility while screening the citations. Disagreements were resolved by consensus.

Data extraction and quality assessment
The final set of English language articles was assessed independently by two reviewers (HBX and RHJ). The reviewers were blinded to publication details, and
disagreements were resolved by consensus. Data retrieved from the reports included participant characteristics, test methods, sensitivity and specificity data, cut-off values, publication year and methodological quality. The numbers of true-positive, false-positive, false-negative and true-negative results are shown in Table 1.

We assessed the methodological quality of the studies using guidelines published by the standards for reporting diagnostic accuracy (STARD) initiative (maximum score 25) and the quality assessment for studies of diagnostic accuracy (QUADAS) tool (maximum score 14). In addition, for each study the following study design characteristics were also retrieved: 1) cross-sectional design (vs. case-control design), 2) consecutive or random sampling of patients, 3) blinded (single or double) interpretation of determination and reference standard results, and 4) prospective data collection. If no data on the above criteria were reported in the primary studies, we requested the information from the authors. If the authors did not respond to our letters, the ‘unknown’ items were treated as ‘no’.

**Statistical analysis**

The kappa (κ) coefficient was calculated as a measure of agreement between two reviewers. We used standard methods recommended for meta-analyses of diagnostic test evaluations. The following measures of test accuracy for each study, e.g., sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR) and diagnostic odds ratio (DOR) were computed using several statistical software programmes. The analysis was based on a summary receiver operating characteristic (SROC) curve. The sensitivity and specificity for the single test threshold identified for each study were used to plot an SROC curve. A random effects model was used to calculate the average sensitivity, specificity and other measures across studies. The term ‘heterogeneity’ when used in relation to meta-analyses refers to the degree of variability. Fisher’s exact tests were used to detect significant heterogeneity. To assess the effects of

---

### Table 1 Summary of studies included

<table>
<thead>
<tr>
<th>Study, year, reference</th>
<th>Patients n</th>
<th>Assay method</th>
<th>Cut-off IU/l</th>
<th>Test results</th>
<th>Quality score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chotmongkol, 2006</td>
<td>177</td>
<td>Giusti method</td>
<td>15.5</td>
<td>TP: 12, FP: 11, FN: 4, TN: 42</td>
<td>9, 7</td>
</tr>
<tr>
<td>Kashyap, 2007</td>
<td>153</td>
<td>Giusti method</td>
<td>5</td>
<td>TP: 55, FP: 12, FN: 11, TN: 75</td>
<td>14, 9</td>
</tr>
<tr>
<td>Lopez-Cortes, 1995</td>
<td>180</td>
<td>Kinetic determination</td>
<td>10</td>
<td>TP: 10, FP: 11, FN: 10, TN: 149</td>
<td>12, 7</td>
</tr>
<tr>
<td>Corral, 2004</td>
<td>62</td>
<td>Spectrophotometric method</td>
<td>8.5</td>
<td>TP: 8, FP: 6, FN: 6, TN: 42</td>
<td>15, 10</td>
</tr>
<tr>
<td>Mishra, 1995</td>
<td>35</td>
<td>Giusti method</td>
<td>5</td>
<td>TP: 24, FP: 3, FN: 3, TN: 5</td>
<td>11, 9</td>
</tr>
<tr>
<td>Choi, 2002</td>
<td>182</td>
<td>Giusti method</td>
<td>10.4</td>
<td>TP: 21, FP: 6, FN: 15, TN: 140</td>
<td>15, 11</td>
</tr>
<tr>
<td>Rohani, 1995</td>
<td>119</td>
<td>Giusti method</td>
<td>9</td>
<td>TP: 14, FP: 13, FN: 0, TN: 92</td>
<td>13, 8</td>
</tr>
<tr>
<td>Ribera, 1987</td>
<td>205</td>
<td>Giusti method</td>
<td>9</td>
<td>TP: 32, FP: 2, FN: 0, TN: 171</td>
<td>9, 8</td>
</tr>
<tr>
<td>Coovadia, 1986</td>
<td>78</td>
<td>Giusti method</td>
<td>10</td>
<td>TP: 11, FP: 18, FN: 4, TN: 45</td>
<td>13, 11</td>
</tr>
</tbody>
</table>

IU = international unit; TP = true-positive; FP = false-positive; FN = false-negative; TN = true-negative; STARD = standards for reporting diagnostic accuracy; QUADAS = quality assessment for studies of diagnostic accuracy.

---

**RESULTS**

After independent review, 16 publications dealing with ADA concentrations for the diagnosis of TBM were considered to be eligible for inclusion in the analysis. Studies were excluded if ADA concentration was determined only in TBM patients, they recruited <10 patients with confirmed TBM, or they did not calculate the sensitivity or specificity. Some of the data might reappear in the publication. A final 10 studies involving 375 patients with TBM and 989 non-TBM patients were available for analysis; the clinical characteristics of these studies, along with their QUADAS scores, are outlined in Table 1.

**Quality of reporting and study characteristics**

The average internal agreement between the two reviewers for items in the quality checklist was 0.9. The average sample size of the included studies was 136 (range 35–281). In all 10 studies, the diagnoses of all patients with TBM were made based on positive smears or cultures of *M. tuberculosis* in the CSF and/or on clinical symptoms. There were also increased proteins, decreased glucose and good response to anti-tuberculosis drugs in the CSF. Our initial data were affected by the poor quality of reporting in the primary studies. To overcome this problem, we contacted all authors of the 10 studies by airmail as well as e-mail where e-mail addresses were available. Six
authors responded who could provide additional data. As shown in Table 2, 5 of the 10 studies (50%) were cross-sectional in design. In eight studies (80%), the samples were collected from consecutive patients. Eight studies (80%) were prospective.

Diagnostic accuracy

Figure 1 shows the forest plot of sensitivity and specificity of the 10 ADA assays in the diagnosis of TBM. Sensitivity and specificity ranged from 0.50 to 1.00 (mean 0.79, 95% confidence interval [CI] 0.75–0.83) and from 0.63 to 0.99 (mean 0.91, 95% CI 0.89–0.93), respectively. PLR was 6.85 (95% CI 4.11–11.41), NLR 0.29 (95% CI 0.19–0.44) and DOR 26.92 (95% CI 12.72–56.97). χ² values of sensitivity, specificity, PLR, NLR and DOR were respectively 44.80 (P < 0.001), 63.60 (P < 0.001), 51.66 (P < 0.001), 34.82 (P < 0.001) and 29.75 (P < 0.001), indicating significant heterogeneity for sensitivity, specificity, PLR, NLR and DOR between the studies.

Unlike a traditional receiver operating characteristic (ROC) plot that explores the effect of varying thresholds on sensitivity and specificity in a single study, each data point in the summary ROC (SROC) plot represents a separate study. The SROC curve presents a global summary of test performance and shows the trade-off between sensitivity and specificity. A graph of the SROC curve for the ADA determination showing true-positive vs. false-positive rates from individual studies is shown in Figure 2. As a global measure of test efficacy we used the Q-value, the intersection point of the SROC curve with a diagonal line from the left upper corner to the right lower

<table>
<thead>
<tr>
<th>Study, year, reference</th>
<th>TB patients/ non-TB subjects</th>
<th>Reference standard</th>
<th>Cross-sectional design</th>
<th>Consecutive or random</th>
<th>Blinded design</th>
<th>Prospective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chotmongkol, 2006⁸</td>
<td>15/53</td>
<td>Bac/clin</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Kashyap, 2007⁹</td>
<td>66/87</td>
<td>Bac/clin</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Lopez-Cortes, 1995¹⁰</td>
<td>20/160</td>
<td>Bac/clin</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Corral, 2004¹¹</td>
<td>14/48</td>
<td>Bac</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Kashyap, 2006¹²</td>
<td>117/164</td>
<td>Bac/clin</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Mishra, 1995¹³</td>
<td>27/84</td>
<td>Bac/clin</td>
<td>No</td>
<td>Unknown</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Choi, 2002¹⁴</td>
<td>36/146</td>
<td>Bac/clin</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Rohani, 1995¹⁵</td>
<td>14/105</td>
<td>Bac/clin</td>
<td>Unknown</td>
<td>Yes</td>
<td>Unknown</td>
<td>Yes</td>
</tr>
<tr>
<td>Ribera, 1982¹⁶</td>
<td>32/173</td>
<td>Bac</td>
<td>Unknown</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Coovadia, 1986¹⁷</td>
<td>15/63</td>
<td>Bac/clin</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Yes</td>
</tr>
</tbody>
</table>

TB = tuberculosis; bac = bacteriology; clin = clinical course.
corner of the ROC space, which corresponds to the highest common value of sensitivity and specificity for the test. This point does not indicate the only or even the best combination of sensitivity and specificity for a particular clinical setting, but represents an overall measure of the discriminatory power of a test. Our data showed that the SROC curve was positioned near the desirable upper left corner of the SROC curve, and the $Q$ value was 0.85, while the area under the curve (AUC) was 0.92, indicating a high level of overall accuracy.

Multiple regression analysis and publication bias
Using STARD guidelines, a quality score for every study was compiled on the basis of the title, introduction, methods, results and discussion (Table 1).\(^1\) Quality was scored using QUADAS, where 1, 0 or −1 were given if all criteria were fulfilled, unclear or not achieved, respectively (Table 1).\(^2\) These scores were used in the meta-regression analysis to assess the effect of study quality on the RDOR of ADA in the diagnosis of TBM. As shown in Table 3, there was no statistical significance in RDOR values between studies of higher (STARD score $\geq 13$, QUADAS score $\geq 10$) and lower quality. There was no significant difference between studies that were or were not blinded, cross-sectional, consecutive/random and prospective. The evaluation of publication bias showed that the Egger test was not significant ($P = 0.171$). The funnel plots for publication bias (Figure 3) also did not show any asymmetry. These results indicate a lack of potential for publication bias.

DISCUSSION
Making a differential diagnosis between TBM and non-TBM is a critical clinical problem. Conventional methods, such as the direct examination of CSF, are positive in only 5~20% of cases. The rate of positivity is about 40% in culture, which takes about 6 weeks.\(^4\),\(^5\) Cerebrospinal levels of some biomarkers have been proposed to be helpful in the diagnosis of TBM, including ADA. The present meta-analysis shows that the mean values of the sensitivity and specificity of the ADA assays were respectively 0.79 and 0.91, and that the maximum joint sensitivity and specificity ($Q$ value) was 0.85, while the AUC was 0.92, indicating a high level of overall accuracy. We also noted that three studies\(^16\),\(^29\),\(^32\) showed relatively low sensitivity (<0.70) and six studies\(^15\),\(^28\),\(^29\),\(^31\) demonstrated low specificity (<0.90) for the detection of ADA in diagnosing TBM.

DOR is a single indicator of test accuracy that combines the data from sensitivity and specificity into a single number and the ratio of the odds of positive test results in the patient with disease or without disease.\(^31\) DOR values range from 0 to infinity, with higher values indicating better discriminatory test performance (i.e., higher accuracy). In the present meta-analysis, we found that the mean DOR was 26.92, also indicating a high level of overall accuracy, and also presented both PLR and NLR as diagnostic accuracy, as the SROC curve and the DOR are not easy to interpret and use in clinical practice, and ratios are considered to be more clinically meaningful.\(^32\),\(^33\) A PLR of 6.85 suggests that TBM patients have an approximately 7-fold higher chance of being ADA assay-positive as compared to patients without TBM. However, if the ADA assay result is negative, the

---

**Table 3** Weighted meta-regression of the effects of methodological quality and study design on diagnostic precision of cerebrospinal adenosine deaminase in 10 assays

<table>
<thead>
<tr>
<th>Covariates</th>
<th>Studies</th>
<th>$n$</th>
<th>Coefficient</th>
<th>RDOR</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>STARD $\geq 13$</td>
<td>6</td>
<td>−0.168</td>
<td>0.84 (0.44–1.63)</td>
<td>0.5628</td>
<td></td>
</tr>
<tr>
<td>QUADAS $\geq 10$</td>
<td>4</td>
<td>−0.137</td>
<td>0.87 (0.39–1.97)</td>
<td>0.7041</td>
<td></td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>5</td>
<td>0.427</td>
<td>1.53 (0.33–7.10)</td>
<td>0.5316</td>
<td></td>
</tr>
<tr>
<td>Consecutive or random</td>
<td>8</td>
<td>2.243</td>
<td>9.42 (0.61–144.42)</td>
<td>0.0952</td>
<td></td>
</tr>
<tr>
<td>Blinded</td>
<td>2</td>
<td>−0.411</td>
<td>0.66 (0.13–3.39)</td>
<td>0.5702</td>
<td></td>
</tr>
<tr>
<td>Prospective</td>
<td>8</td>
<td>0.231</td>
<td>1.26 (0.23–6.87)</td>
<td>0.7565</td>
<td></td>
</tr>
</tbody>
</table>

RDOR = relative diagnostic odds ratio; STARD = standards for reporting diagnostic accuracy; QUADAS = quality assessment for studies of diagnostic accuracy.
probability that the patient has TBM is approximately 29%, which is not low enough to rule out TBM. These data suggest that a negative ADA assay result should not be used alone as a justification to exclude or discontinue anti-tuberculosis treatment. The choice of therapeutic strategy should be based on the results of microscopic examination of a smear or culture for M. tuberculosis, as well as other clinical data, such as response to anti-tuberculosis treatment.

An exploration of the reasons for heterogeneity rather than the computation of a single summary measure is an important goal of meta-analysis.34 In our meta-analysis, both STARD and QUADAS scores were used to assess the effect of study quality on RDOR. We found significant heterogeneity for sensitivity, specificity, PLR, NLR, and DOR among the studies analysed, although the exact mechanism responsible for the significance was inexplicable. There were no differences between studies with or without blinded, cross-sectional, consecutive/random and prospective design, which may contribute to the quality of the test performance.

Publication bias may also have been introduced by the inflation of diagnostic accuracy estimates, as studies with positive results were more likely to be accepted for publication. TBM is not always diagnosed by microbiological examination. The heterogeneity detected in the present analysis may have limitations. TBM was diagnosed in some patients with TBM only on clinical course. This can cause non-random misclassification, leading to biased results.

In conclusion, current evidence suggests a potential role of ADA assays in confirming a diagnosis of TBM. There is a great need to develop CSF biomarkers specific for TBM. The results of ADA assays should be interpreted with clinical findings and other examinations.

Reference

Objectives: To determine the diagnostic value of adenosine deaminase (ADA) in tuberculous meningitis (TBM).

Methods: A systematic review of English language studies was performed to evaluate the sensitivity, specificity, and diagnostic accuracy of ADA concentrations in cerebrospinal fluid (CSF) for the diagnosis of TBM. The summary receiver operating characteristic curves were used to assess the global efficacy of the test.

Results: Ten studies met the inclusion criteria. The sensitivity of ADA for the diagnosis of TBM was 0.79 (95% CI 0.75–0.83), specificity 0.91 (95% CI 0.89–0.93), positive likelihood ratio 6.85 (95% CI 4.11–11.41), negative likelihood ratio 0.29 (95% CI 0.19–0.44), and diagnostic odds ratio 26.93 (95% CI 12.73–56.97).

Conclusion: Our data suggest that CSF ADA concentration can be a sensitive and specific marker and a critical criterion for the diagnosis of TBM.