

REVIEW

A systematic review of commercial serological antibody detection tests for the diagnosis of extrapulmonary tuberculosis

Karen R Steingart, Megan Henry, Suman Laal, Philip C Hopewell, Andrew Ramsay, Dick Menzies, Jane Cunningham, Karin Welding, Madhukar Pai

Postgrad Med J 2007;83:705–712. doi: 10.1136/thx.2006.075754

Conventional diagnostic tests for tuberculosis have several limitations and are often unhelpful in establishing the diagnosis of extrapulmonary tuberculosis. Although commercial serological antibody based tests are available, their usefulness in the diagnosis of extrapulmonary tuberculosis is unknown. A systematic review was conducted to assess the accuracy of commercial serological antibody detection tests for the diagnosis of extrapulmonary tuberculosis. In a comprehensive search, 21 studies that reported data on sensitivity and specificity for extrapulmonary tuberculosis were identified. These studies evaluated seven different commercial tests, with Anda-TB IgG accounting for 48% of the studies. The results showed that (1) all commercial tests provided highly variable estimates of sensitivity (range 0.00–1.00) and specificity (range 0.59–1.00) for all extrapulmonary sites combined; (2) the Anda-TB IgG kit showed highly variable sensitivity (range 0.26–1.00) and specificity (range 0.59–1.00) for all extrapulmonary sites combined; (3) for all tests combined, sensitivity estimates for both lymph node tuberculosis (range 0.23–1.00) and pleural tuberculosis (range 0.26–0.59) were poor and inconsistent; and (4) there were no data to determine the accuracy of the tests in children or in patients with HIV infection, the two groups for which the test would be most useful. At present, commercial antibody detection tests for extrapulmonary tuberculosis have no role in clinical care or case detection.

having a lower frequency of sputum smear positivity, even when the lungs are involved.^{7–9}

The diagnosis of extrapulmonary tuberculosis is often difficult to establish, especially for patients in resource limited areas. Signs and symptoms are non-specific and microscopic examination for acid-fast bacilli, the cornerstone of diagnosis for pulmonary tuberculosis in most parts of the world, lacks sensitivity for extrapulmonary disease.^{10–11} Mycobacterial culture and histological examination for caseating granulomas are more sensitive but not commonly available. Invasive procedures that are complex and costly may be required to obtain the necessary diagnostic specimens.^{11–12} In a retrospective study of patients in Tanzania with extrapulmonary tuberculosis, bacteriological or histological confirmation of diagnosis was found in only 18%.¹³ Because of these difficulties, misdiagnosis of extrapulmonary tuberculosis is common in all countries and may result in unnecessary treatment if falsely diagnosed, or greater morbidity and mortality if the diagnosis is missed, especially in persons with HIV infection.^{11–14–16}

Immune based tests would seem to offer the potential to improve the diagnosis of extrapulmonary tuberculosis as some of the test formats (eg, immunochromatographic test) are practical for resource limited areas. Blood or urine based assays avoid the problems of obtaining a specimen of the affected organ for microbiological or histological assay, are simpler to perform than smear microscopy and the results can be available within hours.^{8–17} Efforts to develop immune based tests for the detection of antibodies, antigens and immune complexes have been underway for decades and their performance described in several reviews and textbook chapters.^{18–27} The most common of these tests concentrate on the detection of the humoral (serological) antibody immune response to *Mycobacterium tuberculosis* (the subject of this review), as opposed to the T cell based cellular immune response (eg, interferon-gamma release assays) or direct detection of antigens in specimens other than serum (eg, lipoarabinoman-

See end of article for authors' affiliations

Correspondence to: Dr Madhukar Pai, Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montréal, Canada, H3A 1A2; madhukar.pai@mcgill.ca

Received 28 November 2006
Accepted 9 February 2007

Although tuberculosis most commonly affects the lungs, any organ or tissue may be involved. In the USA about 20% of incident cases in 2005 had only extrapulmonary sites of disease and an additional 9% had both pulmonary and extrapulmonary involvement.¹ Globally, the proportion of extrapulmonary cases reported by countries ranges from 15% to 25% with greater proportions occurring in countries with a high prevalence of HIV infection.² In addition to being proportionately greater in persons with HIV infection,^{3–5} extrapulmonary involvement occurs with greater relative frequency in children than in adults.⁶ In children and in persons with HIV infection, extrapulmonary tuberculosis compounds the diagnostic difficulty imposed by their

Abbreviations: AUC, area under the curve; FPR, false positive rate; TPR, true positive rate; ROC, receiver operating characteristic; SROC, summary receiver operating characteristic

This is a reprint of a paper that appeared in *Thorax*, October 2007, Volume 62, pages 911–8. Reprinted with kind permission of the authors and publisher.

nan detection in urine^{28, 29}). It is tempting to speculate that a combination of both humoral and T cell based diagnostic tests could provide the highest diagnostic efficacy, although this has not been evaluated to date.

A number of in-house antibody detection tests have been developed but are not marketed. These tests use different antigens and distinct protocols and techniques.

Currently, dozens of commercial serological antibody detection tests (hereafter referred to as commercial tests) are marketed in low income countries where diagnostic tests are rarely subjected to regulatory review or approval.^{30, 31} The extent of their use is unknown; however, companies report sales volumes between 3000 and 300 000 tests per year.³² These tests differ in several respects, including antigen composition and source (eg, native or recombinant), chemical composition (eg, protein, carbohydrate or lipid), extent and manner of antigen purification, class of immunoglobulin (eg, IgG, IgA or IgM), and test format (eg, enzyme-linked immunosorbent assay (ELISA) and immunochromatographic test). Most of the studies investigating the use of antibody detection tests have focused on pulmonary tuberculosis; only a subset has also included patients with extrapulmonary tuberculosis.

To our knowledge, the body of literature evaluating commercial tests for the diagnosis of extrapulmonary tuberculosis has not been synthesised. We therefore conducted a systematic review to summarise the evidence on accuracy (sensitivity and specificity) of commercial tests according to the guidelines and methods proposed for diagnostic systematic reviews and meta-analyses.³³ We specifically addressed the following questions:

- Overall, how accurate are commercial tests for the diagnosis of extrapulmonary tuberculosis?
- How accurate are commercial tests for the diagnosis of a specific form of extrapulmonary tuberculosis?

METHODS

Search strategy

The following electronic databases were searched for primary studies in the English language: PubMed (1990 to December 2006), BIOSIS (1990 to December 2005), Embase (1990 to December 2005) and Web of Science (1990 to December 2005). The search terms used included "tuberculosis", "*Mycobacterium tuberculosis*", "immunological tests", "serological tests", "antibody detection", "antigen detection", "ELISA", "western blot" and "sensitivity and specificity". Additional studies were identified by contacting experts in the field and by searching reference lists from primary studies, review articles and textbook chapters.

Study selection

Extrapulmonary tuberculosis was defined as tuberculosis in which the major site of involvement was outside the lungs. Thereafter, nine types of extrapulmonary disease were classified: lymph node, pleural, meninges and/or central nervous system (CNS), bone and/or joint, disseminated/miliary, genitourinary, abdominal, skin and other sites. We included only those studies that based the diagnosis of extrapulmonary tuberculosis on (1) isolation of *M tuberculosis* on culture or, for studies conducted without culture in tuberculosis endemic countries, the presence of acid-fast bacilli detected by smear microscopy; or (2) the presence of caseating granulomas in histopathological specimens.

Studies that relied solely on clinical and/or radiological features or improvement while on antituberculosis treatment as the diagnostic criteria were excluded. We further excluded: (1) studies published before 1990 for the reason that many studies

used crude extracts or obsolete immunological methods; (2) studies with <50 subjects (at least 25 cases and 25 participants without tuberculosis for inclusion); (3) studies in which data were only provided for pulmonary and extrapulmonary cases combined; (4) studies of fluids other than serum (eg, cerebrospinal fluid); (5) studies of latent tuberculosis infection; (6) studies focused on non-tuberculous mycobacteria; (7) studies of antibody responses during or after tuberculosis treatment; (8) investigations using non-immunological methods for detection of antibodies; (9) basic science literature that focused on cloning of new antigens or their immunological properties (eg, epitope mapping) or other new methods of antibody detection; and (10) case reports and reviews.

Initially, two reviewers screened citations retrieved from all sources. To identify eligible studies, a second screen was done of full texts from citations found relevant in the first screen. A list of excluded studies and reasons for their exclusion is available from the authors.

Data extraction

One reviewer extracted data on the following qualities: study design, methodological quality, study population, reference standard, site of involvement, antigen and antibody characteristics, laboratory technique, and sensitivity and specificity. To verify the reproducibility of data extraction, a second reviewer independently extracted data from 24% of the included studies. The inter-rater agreement for sensitivity and specificity estimates was 100%. Data not clearly reported were coded as "not reported". When necessary, we attempted to contact authors for additional information.

Although some authors evaluated test performance using several different types of comparison groups, we preferentially selected only one type of comparison group for each study in the following order: (1) patients in whom extrapulmonary tuberculosis was initially suspected but who were later found to have a disease other than tuberculosis; (2) patients in whom pulmonary tuberculosis was initially suspected but who were later found to have non-tuberculous respiratory disease; (3) patients with a variety of diseases other than tuberculosis (mixed disease); (4) healthy persons from tuberculosis endemic countries; (5) contacts of patients with tuberculosis; (6) participants from categories 1–5 combined; and (7) healthy persons from non-endemic countries. We felt this hierarchy prioritised the groups in which the test would be used and provide more clinically relevant results.

Assessment of study quality

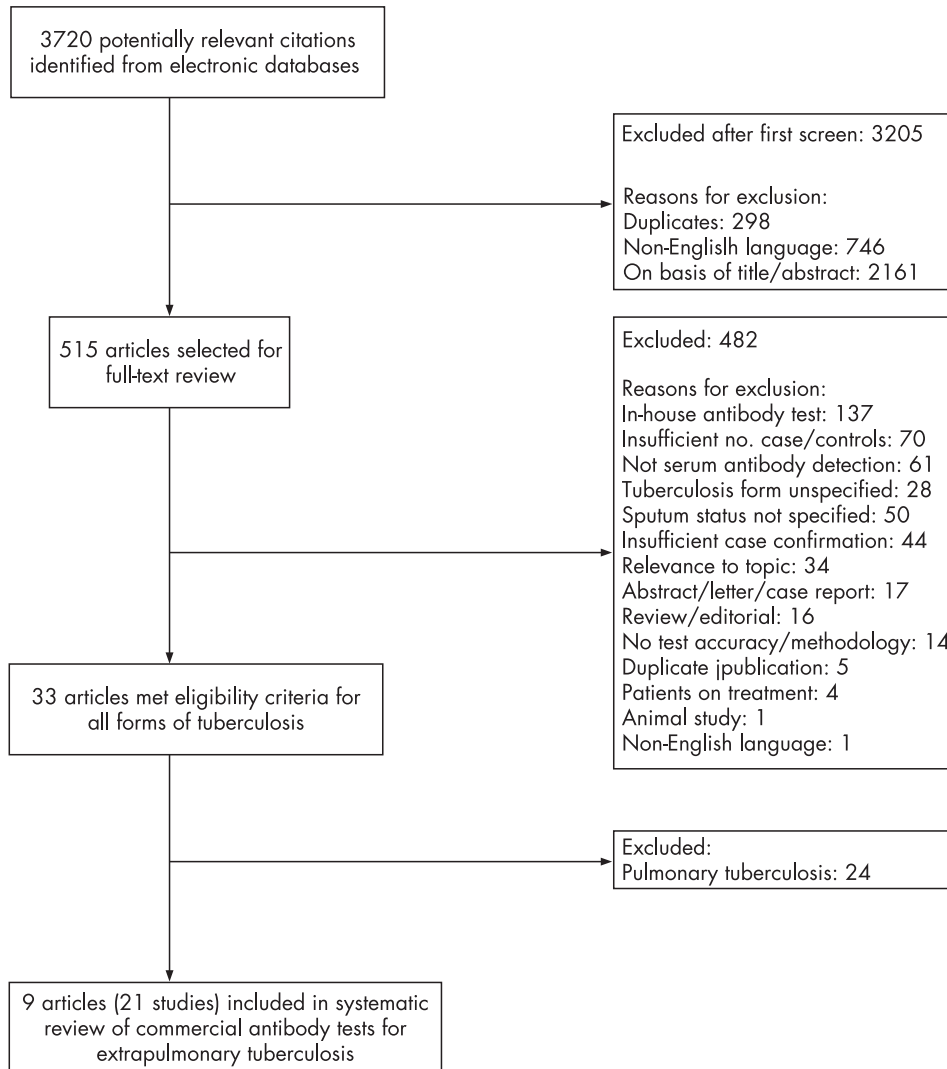
The quality of studies was assessed using the following criteria, suggested as important for diagnostic studies:^{33, 34}

- Was the commercial test result performed and recorded by technicians who were unaware (blinded) of the results of the reference standard?
- Did the whole sample or a randomly selected subset of the sample receive verification using the reference standard?
- Did the study prospectively recruit consecutive patients suspected of having extrapulmonary tuberculosis?

Data collation and meta-analysis

We used standard methods recommended for meta-analyses of diagnostic test evaluations.^{33, 35} As studies were heterogeneous, particularly with respect to the site of involvement, antigen composition of the tests, antibody class (IgG, IgM, or IgA) and control groups, we first grouped studies by type of commercial test and then further stratified by immunoglobulin class and location of disease. If insufficient data (ie, <25 cases) were provided for a specific disease site, we combined data from

Figure 1 Flow diagram for study selection.



several sites into a “multiple site” category with at least 25 cases. To calculate sensitivity and specificity of the commercial tests, we cross-tabulated each result against the reference standard. Sensitivity refers to the proportion of extrapulmonary tuberculosis cases with a positive result on a specific commercial test; specificity refers to the proportion of tuberculosis negative participants that had negative results on a specific commercial test. In calculations of sensitivity we included studies that used smear positivity or histological

characteristics as the reference standard together with studies that used culture.

Data were analysed using SPSS (Version 14.0.1.366)³⁶ and Meta-DiSc software (Version 1.4).³⁷ Sensitivity and specificity estimates were calculated for the commercial tests along with their 95% confidence intervals. In addition, true positive rates (TPR = sensitivity) and false positive rates (FPR = 1 – specificity) were summarised using a summary receiver operating characteristic (SROC) curve. Each data point in the

Table 1 Commercial serological antibody detection tests for the diagnosis of extrapulmonary tuberculosis

Name of test (number of studies in review)	Antigen(s)	Source of antigen	Ig class	Laboratory technique	Manufacturer	Address/URL
Anda-TB* (17)	Antigen 60	Native	IgG, IgA, IgM	ELISA	Anda Biologicals S.A.	Strasbourg, France www.andabiologicals.com
AMRAD ICT TB (1)	38 kDa and four proprietary antigens	Recombinant	IgG	Immunochromatographic test card	ICT Diagnostics	Balgowlah, NSW, Australia
Pathozyme Myco (1)	Lipoarabinomannan (LAM) and 38 kDa	Native LAM and recombinant 38 kDa	IgG	ELISA	Omega Diagnostics http://omegadiagnostics.com	Alloa, UK
Pathozyme TB Complex Plus (1)	38 kDa and 16 kDa	Recombinant	IgG	ELISA	Omega Diagnostics http://omegadiagnostics.com	Alloa, UK
SEVA TB (1)	31 kDa	Native glycoprotein antigen from culture filtrate of <i>M tuberculosis</i> H37Rv	IgG	ELISA	Jamnalal Bajaj Tropical Disease Research Centre, Mahatma Gandhi Institute of Medical Sciences	Sevagram-442 102 (Wardha) M S (India)

Ig, immunoglobulin; ELISA, enzyme-linked immunosorbent assay; kDa, kilodalton.
*Anda-TB tests include: IgG (n=10), IgM (n=5), IgA (n=1), IgG and IgM (n=1).

Table 2 Selected characteristics of studies investigating Anda-TB (Anda Biologicals, Strasbourg, France) for the diagnosis of extrapulmonary tuberculosis

Reference	Data collection	Verification	Reference standard	Disease site	Country	Comparison group	Ig class	No. of participants*	Sensitivity (95% CI)	Specificity (95% CI)
Alifano (1998, a) ⁴¹	Retrospective	Differential	Culture and/or histology	Multiple	Italy	Mixed disease	IgG	42/44	0.74 (0.58 to 0.86)	0.93 (0.81 to 0.99)
Alifano (1998, b) ⁴¹	Retrospective	Differential	Culture and/or histology	Multiple	Italy	Mixed disease	IgA	42/44	0.69 (0.53 to 0.82)	0.89 (0.75 to 0.96)
Banerjee (2003, a) ⁴²	Retrospective	NR	Histology	Lymph node	India	Healthy	IgG	30/32	0.43 (0.25 to 0.63)	0.59 (0.41 to 0.76)
Caminero (1993) ⁴³	Prospective	Complete	Culture and/or histology	Pleura	Spain	Pleural TB suspects	IgG	30/48	0.53 (0.34 to 0.72)	1.00 (0.93 to 1.00)
Caminero (1994) ⁴⁴	Prospective	Differential	Culture and/or histology	Multiple	Canary Islands	Non-TB respiratory	IgG	56/31	0.32 (0.20 to 0.46)	0.94 (0.79 to 0.99)
Gevaudan (1992, a) ⁴⁵	Retrospective	Differential	Culture and/or histology	Lymph node	France	Mixed disease	IgG	26/194	1.00 (0.87 to 1.00)	0.76 (0.69 to 0.82)
Gevaudan (1992, b) ⁴⁵	Retrospective	Differential	Culture and/or histology	Lymph node	France	Mixed disease	IgM	26/194	0.23 (0.09 to 0.44)	0.95 (0.91 to 0.98)
Gevaudan (1992, c) ⁴⁵	Retrospective	Differential	Culture	Disseminated/miliary (primary)	France	Mixed disease	IgG	56/194	0.95 (0.85 to 0.99)	0.76 (0.69 to 0.82)
Gevaudan (1992, d) ⁴⁵	Retrospective	Differential	Culture	Disseminated/miliary (primary)	France	Mixed disease	IgM	56/194	0.32 (0.20 to 0.46)	0.95 (0.91 to 0.98)
Gevaudan (1992, e) ⁴⁵	Retrospective	Differential	Culture	Disseminated/miliary (post-primary)	France	Mixed disease	IgG	25/194	1.00 (0.86 to 1.00)	0.76 (0.69 to 0.82)
Gevaudan (1992, f) ⁴⁵	Retrospective	Differential	Culture	Disseminated/miliary (post-primary)	France	Mixed disease	IgM	25/194	0.00 (0.00 to 0.14)	0.95 (0.91 to 0.98)
Gevaudan (1992, g) ⁴⁵	Retrospective	Differential	Culture	Genitourinary	France	Mixed disease	IgG	34/194	1.00 (0.90 to 1.00)	0.76 (0.69 to 0.82)
Gevaudan (1992, h) ⁴⁵	Retrospective	Differential	Culture	Genitourinary	France	Mixed disease	IgM	34/194	0.00 (0.00 to 0.10)	0.95 (0.91 to 0.98)
Kunter (2003, a) ⁴⁶	Prospective	Differential	Culture	Pleura	Turkey	Non-TB respiratory	IgG	88/37	0.26 (0.17 to 0.37)	0.86 (0.71 to 0.96)
Kunter (2003, b) ⁴⁶	Prospective	Differential	Culture	Pleura	Turkey	Non-TB respiratory	IgM	88/37	0.48 (0.37 to 0.59)	0.92 (0.78 to 0.98)
Kunter (2003, c) ⁴⁶	Prospective	Differential	Culture	Pleura	Turkey	Non-TB respiratory	IgG and IgM	88/37	0.59 (0.48 to 0.70)	0.81 (0.65 to 0.92)
Luh (1996) ⁴⁷	Prospective	NR	Culture	Multiple	Taiwan	Mixed disease	IgG	35/224	0.63 (0.45 to 0.79)	0.92 (0.88 to 0.96)

TB, tuberculosis; Ig, immunoglobulin; NR, not reported.

The treatment status of cases was not reported for reference 45. For all other studies, patients were either not on treatment or on treatment for <14 days at the time serum was obtained.

*Number of participants with tuberculosis/without tuberculosis.

SROC space represents an individual study. The SROC curve is obtained by fitting a regression curve to pairs of TPR and FPR.³⁵

The SROC curve and the area under the curve (AUC) present an overall summary of test performance and display the trade off between sensitivity and specificity. An AUC of 1.00 indicates perfect discriminatory ability of the diagnostic test. In addition, the Q* index is another useful global summary of the SROC curve and test performance. The Q* index, defined by the point where sensitivity equals specificity on the SROC curve, is the point that is intersected by the antidiagonal, the top left corner of the SROC region. A Q* value of 1.00 indicates 100% accuracy.^{35 38 39}

In meta-analyses of studies of diagnostic tests, heterogeneity refers to a high level of variability in study results.⁴⁰ Such

heterogeneity could be a result of variability in thresholds, laboratory technique, disease spectrum, study design and/or quality between studies.⁴⁰ In the presence of significant heterogeneity, pooled or summary estimates from meta-analyses are difficult to interpret. Given the anticipated variability in accuracy, we decided a priori to avoid the pooling of sensitivity and specificity. Also, as described previously, we addressed heterogeneity by using subgroup (stratified) analyses.

RESULTS

Description of included studies

Of the 3720 citations identified in the literature search, 9 publications describing the results of 21 independent studies

Table 3 Selected characteristics of studies investigating commercial tests for the diagnosis of extrapulmonary tuberculosis

Reference	Data collection	Verification	Reference standard	Disease site	Country	Comparison group	Name of test*	No. of participants‡	Sensitivity (95% CI)	Specificity (95% CI)
Banerjee (2003, b) ⁴²	Retrospective	NR	Histology	Lymph node	India	Healthy	SEVA TB	30/32	0.77 (0.58 to 0.90)	0.88 (0.71 to 0.97)
McConkey (2002) ⁴⁸ †	Prospective	Complete	Culture	Meninges	Egypt	Mixed disease	AMRAD ICT	56/74	0.48 (0.35 to 0.62)	0.82 (0.72 to 0.90)
Nsanze (1997, a) ⁴⁹	Prospective	Differential	Culture and/or smear	NR	UAE	Mixed disease and healthy	Pathozyme Myco	35/35	0.51 (0.34 to 0.69)	1.00 (0.90 to 1.00)
Nsanze (1997, b) ⁴⁹	Prospective	Differential	Culture and/or smear	NR	UAE	Mixed disease and healthy	Pathozyme TB Complex Plus	35/35	0.11 (0.03 to 0.27)	1.00 (0.90 to 1.00)

NR, not reported.

*All commercial tests detected immunoglobulin G antibody.

†Blinded study.

‡Number of participants with tuberculosis/without tuberculosis.

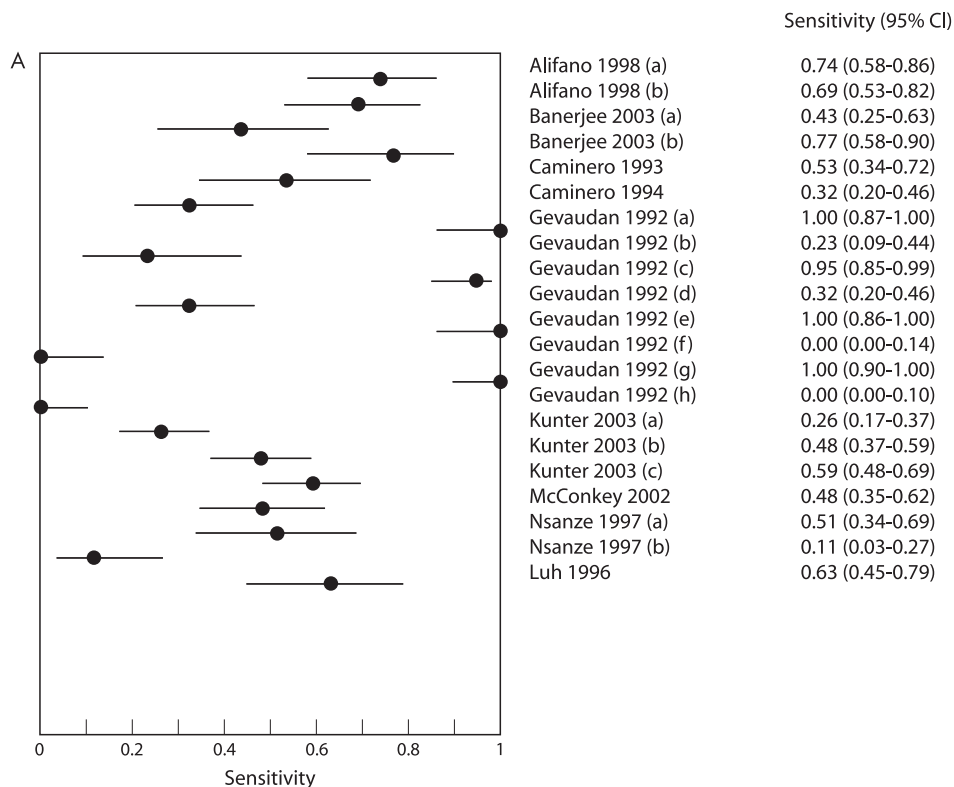
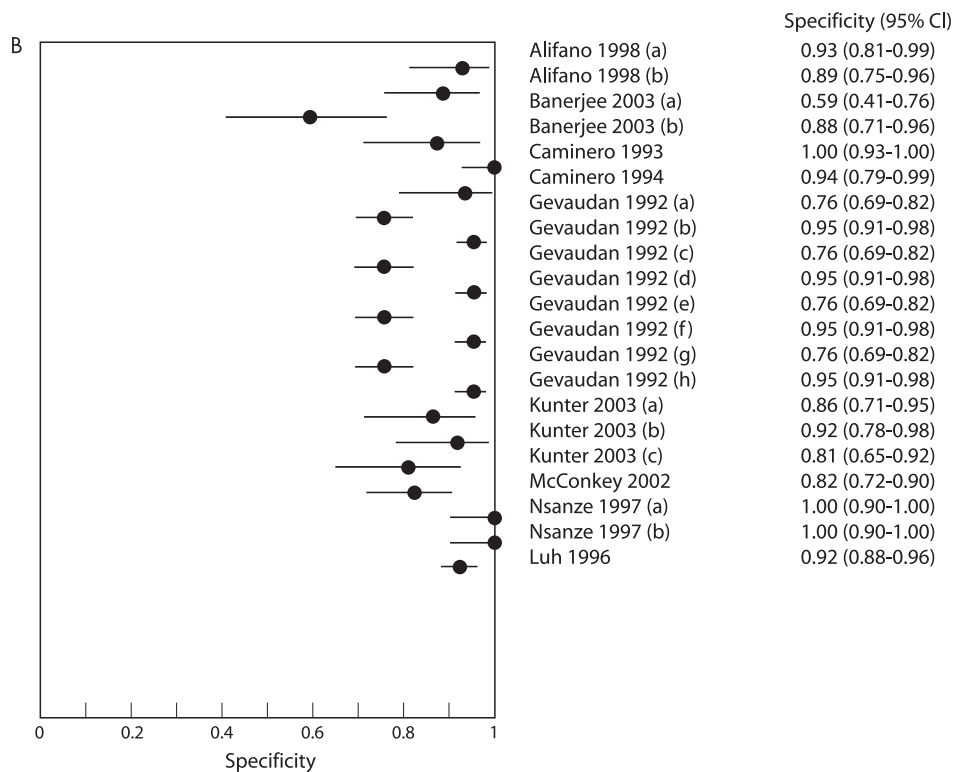


Figure 2 (A) Sensitivity and (B) specificity estimates of commercial tests for the diagnosis of extrapulmonary tuberculosis (21 studies). Point estimates of sensitivity and specificity from each study are shown as solid circles. The solid lines represent the 95% confidence intervals (CI).



met our eligibility criteria (fig 1).⁴¹⁻⁴⁹ None of the studies reported the method (eg, consecutive or random) of subject selection. Only one study⁴⁸ reported blinded interpretation. No studies involved children younger than 15 years or patients with documented HIV infection. Six studies (29%)^{41 46 48} were performed in HIV negative patients and 15 (71%) in patients in whom HIV status was unknown or not reported.^{42-45 47 49} The median sample size was 35 cases (interquartile range (IQR) 30-56) and 48 participants without tuberculosis (IQR 37-194).

ELISA was used in 20 studies and immunochromatography in one study.⁴⁸ All investigators adhered to standard laboratory methods (eg, mean \pm 2SD measured in a healthy population and receiver operating characteristic (ROC) curves) for determining the cut off values, as recommended by the manufacturers. Table 1 shows the commercial tests and their respective antigens. Anda-TB IgG was the most frequently studied test (n = 10 studies (48%)). Tables 2 and 3 show design, quality, performance characteristics and disease site.

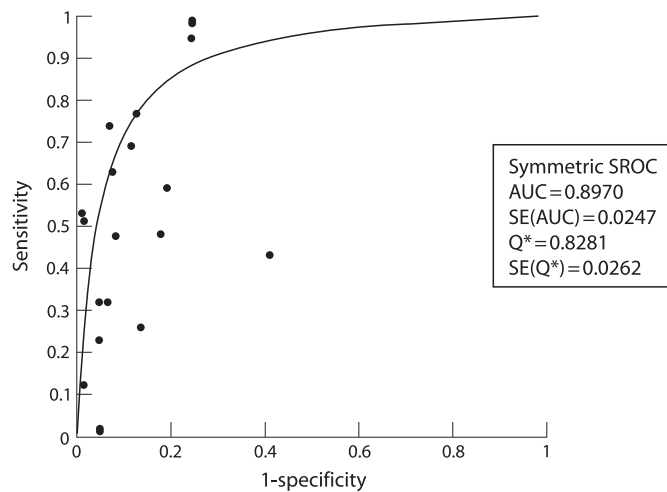


Figure 3 Summary receiver operating characteristic (SROC) curve for commercial tests (21 studies). Each solid circle represents each study in the meta-analysis. The curve is the regression line that summarises the overall diagnostic accuracy. AUC, area under the curve; SE (AUC), standard error of AUC; Q^* , an index defined by the point on the SROC curve where the sensitivity and specificity are equal, which is the point closest to the top left corner of the ROC space; SE(Q^*), standard error of Q^* index.

Overall, how accurate are commercial tests for the diagnosis of extrapulmonary tuberculosis?

For all 21 studies the sensitivity estimates ranged from 0.00 to 1.00 and specificity estimates from 0.59 to 1.00 (fig 2). Both sensitivity and specificity varied widely among studies using the same commercial test and among studies using different commercial tests. Confidence intervals for the sensitivity and specificity values of individual studies, depicted graphically by horizontal lines in the forest plots, show poor overlap, suggesting the presence of significant heterogeneity. As shown in fig 3, the accuracy of commercial tests was modest, the symmetrical SROC curve showing a trade off between sensitivity and specificity, with much greater variability in sensitivity estimates.

We identified 10 studies that assessed the accuracy of Anda-TB IgG.^{41–47} As seen in table 2 (and in online supplementary fig S1A available at <http://thorax.bmj.com/supplemental>), sensitivity estimates ranged from 0.26 to 1.00 and specificity estimates from 0.59 to 1.00. The specificity forest plot (online supplementary fig S1B available at <http://thorax.bmj.com/supplemental>) includes 7 unique studies as 4 of the total 10 studies were conducted with the same comparison population.⁴⁵ The results for both sensitivity and specificity among studies by different investigators were highly variable. Individual study results for the other commercial tests in the review are shown in table 3.

How accurate are commercial tests for the diagnosis of a specific form of extrapulmonary tuberculosis?

Four studies determined the accuracy of commercial tests for the diagnosis of lymph node tuberculosis.^{42–45} As seen in online supplementary figs S2A and B (available at <http://thorax.bmj.com/supplemental>), sensitivity estimates ranged from 0.23 to 1.00 and specificity estimates from 0.59 to 0.95. Four studies determined the accuracy of commercial tests for the diagnosis of pleural tuberculosis.^{43–46} Sensitivity estimates ranged from 0.26 to 0.59 and specificity estimates, from 0.81 to 1.00 (see online supplementary figs S3A and B available at <http://thorax.bmj.com/supplemental>). For both lymph node and pleural tuberculosis, commercial tests showed inconsistent results.

Only one study assessed the accuracy of a commercial test in patients with meningitis.⁴⁸ In this prospective blinded study (56 culture confirmed patients and 74 controls), the sensitivity of the AMRAD ICT TB test was 0.48 (95% CI 0.35 to 0.62) and the specificity was 0.82 (95% CI 0.72 to 0.90).

DISCUSSION

Principal findings

Our systematic review of 21 studies examining the performance of commercial tests for the diagnosis of extrapulmonary tuberculosis showed that (1) all commercial tests provided highly variable estimates of sensitivity (range 0.00–1.00) and specificity (range 0.59–1.00) for all extrapulmonary sites combined; (2) Anda-TB IgG showed highly variable sensitivity (range 0.26–1.00) and specificity (range 0.59–1.00) for all extrapulmonary sites combined; (3) for all commercial tests combined, sensitivity estimates for both lymph node tuberculosis (range 0.23–1.00) and pleural tuberculosis (range 0.26–0.59) were poor and inconsistent; and (4) there were no data to determine the accuracy of commercial tests for the diagnosis of extrapulmonary tuberculosis in children or patients with HIV infection. Although commercial serological tests—by virtue of being rapid, simple to use and non-invasive—are appealing, this review did not find sufficient evidence to justify their use for the diagnosis of extrapulmonary tuberculosis.

Our systematic review had several strengths. First, the comprehensive search strategy with various overlapping approaches enabled us to retrieve relevant studies published since 1990. Moreover, two reviewers independently completed screening and study selection. To verify reproducibility of data extraction, a second reviewer independently extracted data on five (24%) of the included studies. Whenever possible we selected a control population with disease, in lieu of healthy subjects, to evaluate how well commercial tests performed in patients suspected of having tuberculosis. Authors were contacted for missing data. Finally, we analysed data within specific subgroups to lessen the effect of heterogeneity.

This review also had limitations. There were an insufficient number of studies for most of the specific commercial tests or the specific disease sites to provide meaningful summary measures of performance. Our use of stringent criteria for eligibility is probably the main reason that we identified only one study on tuberculous meningitis.⁴⁸ Fifty-six studies of tuberculous meningitis identified by our search strategy were considered ineligible because they involved fewer than 25 cases; investigated studies of fluids (eg, cerebrospinal fluid) other than serum; involved antigen detection; or relied on clinical features and/or treatment response for case confirmation. Our choice of a bacteriological and/or histopathological reference standard may have limited the inclusion of studies involving children. Paediatric tuberculosis is difficult to diagnose on a bacteriological basis because of the paucibacillary nature of the disease.⁶ In addition, the number of specific antigens included in the commercial tests in this review was limited (table 1) compared with the number of potential antigens for serodiagnosis.^{22 50 51}

Another set of problems involved shortcomings in study design and quality. No studies reported the method for recruitment of subjects, so it was not possible to ascertain if studies used the sound probabilistic sampling framework found in consecutive or random sampling designs. Only one study⁴⁸ reported a blinded interpretation of the results of the commercial test and the reference standard. Lack of blinding may have resulted in an overestimation of the accuracy of the commercial test.⁵² Variability in study design and study quality might account for some of the observed heterogeneity evident in the results. Although statistical tests and graphical methods

are available to detect potential publication bias in meta-analyses of randomised control trials, such techniques have not been adequately evaluated for diagnostic data.⁵³ It is therefore difficult to rule out publication bias in our review. In addition, our search strategy may have missed some relevant studies by excluding non-English language publications.

Developing an immunological diagnostic test for tuberculosis presents a formidable challenge in part because both the stage of tuberculosis infection and the tissues involved may alter the profile of genes expressed by the organism and, thus, the antibody responses to these gene products may differ. Studies during the last decade have provided ample evidence that *M tuberculosis* adapts to its environment by altering the profile of genes that it expresses, that these profiles are modulated as infection progresses and the in vivo environment changes,⁵⁴⁻⁵⁷ and that some genes of *M tuberculosis* are differentially expressed in different host tissues.⁵⁸ Consequently, antibodies developed in response to pulmonary tuberculosis may not be the appropriate targets for diagnosing extrapulmonary involvement. The choice of reagents for all current assays was most likely determined in patients with pulmonary tuberculosis. Although a systematic investigation of the humoral immune responses of pulmonary tuberculosis has been performed and several antigenic proteins that are recognised by antibodies at different stages of pulmonary disease have been identified,⁵⁵⁻⁵⁹ no similar analysis of antigens expressed during extrapulmonary replication of *M tuberculosis* has been attempted so far. Thus, identification and study of *M tuberculosis* genes expressed in the different environments that characterise different sites of involvement may be able to provide the optimal reagents for devising a diagnostic test for extrapulmonary forms of the disease.

It will also be important to examine proteins expressed by *M tuberculosis* in HIV infected patients as smear negative pulmonary and extrapulmonary disease are disproportionately higher in HIV positive than in HIV negative individuals.^{3-5 60 61} Given that memory B cells are relatively independent of T cell help, antibody detection based diagnostic tests would be significant assets for identification of forms of paucibacillary disease. Indeed, despite the dysfunctional cellular immune responses and the presence of HIV induced hypergammaglobinaemia,⁶²⁻⁶⁴ the presence of antibodies to *M tuberculosis* antigens has been reported by several investigators.^{51 55 65-67} It is also possible that antibodies to *M tuberculosis* antigens are elicited in patients who get infected before their immune system is significantly compromised and the CD4 numbers are still high, but we are unaware of studies that have addressed this possibility.

CONCLUSIONS AND POLICY IMPLICATIONS

The evidence presented in this systematic review shows that, at present, commercial tests produce highly variable sensitivity and specificity results and therefore cannot be recommended as a sole test for the diagnosis of extrapulmonary tuberculosis. It is particularly disappointing that there are no studies of commercial tests that are of sufficient quality to enable their evaluation in patients with HIV infection or in children, as it is in these groups that the tests could be most useful. Our findings should be interpreted in the context of the variability in design and the quality of the studies in this review. Recent articles have attested to the mediocre quality of diagnostic studies for tuberculosis.^{30 68} The use of guidelines such as the Standards for Reporting of Diagnostic Accuracy (STARD)⁶⁹ and the tool for Quality Assessment of Diagnostic Accuracy Studies (QUADAS)³⁴ may lead to improvements in the quality of future studies. Guidelines specifically for the evaluation of diagnostic tests for infectious diseases have recently been published.⁷⁰

ACKNOWLEDGEMENTS

The authors thank Nandini Dendukuri, Maya Bhat-Gregerson and Anna Meddaugh for technical assistance, Madelyn Hall for help in acquiring the publications in the review, and Izabela Suder-Dayao and Melissa Anthony of WHO/TDR for administrative assistance.



Further data are given in the online figures available at <http://thorax.bmj.com/supplemental>.

Authors' affiliations

Karen R Steingart, Philip C Hopewell, Division of Pulmonary and Critical Care Medicine, San Francisco General Hospital, University of California, and Francis J Curry National Tuberculosis Center, San Francisco, California, USA

Megan Henry, County of Sacramento Department of Health and Human Services, Sacramento, California, USA

Suman Laal, Veterans Affairs Medical Center and Departments of Pathology and Microbiology, New York University School of Medicine, New York, USA

Andrew Ramsay, Jane Cunningham, UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR), World Health Organization, Geneva, Switzerland

Dick Menzies, Madhukar Pai, Respiratory Epidemiology and Clinical Research Unit, Montréal Chest Institute and Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montréal, Canada

Karin Welding, Statens Serum Institut, Department of Infectious Disease Immunology, Copenhagen S, Denmark

This work was supported by the UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR). AR and JC, both with TDR, contributed to the conception and design of the systematic review, and critical revision and decision to publish the manuscript. AR also participated in data interpretation.

Competing interests: KW is co-inventor on a number of patents relating to mycobacterial antigens which may be used for serological assays. All rights have been assigned to Statens Serum Institut.

REFERENCES

- 1 **Centers for Disease Control and Prevention**. *Reported tuberculosis in the United States, 2005*. Atlanta: US Department of Health and Human Services, CDC, 2006:64.
- 2 **World Health Organization**. *TB epidemiology and surveillance workshop, 2006*, http://www.who.int/tb/surveillanceworkshop/trend_analysis/increasing_decreasing_over_under_diagnosis_of_extrapulmonary_tb.htm (accessed 6 October 2006).
- 3 **Iseman MD**. Tuberculosis in relation to HIV and AIDS. In: Iseman MD, ed. *A clinician's guide to tuberculosis*. Philadelphia: Lippincott Williams and Wilkins, 2000:207-9.
- 4 **Hopewell PC**. Tuberculosis and other mycobacterial diseases. In: Mason RJ, Broaddus VC, Murray JF, Nadel JA, eds. *Murray and Nadel's textbook of respiratory medicine*, 4th edn. Philadelphia: Saunders, 2005:984-1026.
- 5 **Talavera W**, Lessnau K-D KL, Handwerker S. Extrapulmonary tuberculosis. In: Friedman LN, eds. *Tuberculosis*, 1st edn. Boca Raton: CRC Press, 1994:113-51.
- 6 **Shingadia D**, Novelli V. Diagnosis and treatment of tuberculosis in children. *Lancet Infect Dis* 2003;**3**:624-32.
- 7 **Hopewell PC**, Pai M, Maher D, et al. International standards for tuberculosis care. *Lancet Infect Dis* 2006;**6**:710-25.
- 8 **Perkins MD**, Roscigno G, Zumla A. Progress towards improved tuberculosis diagnostics for developing countries. *Lancet* 2006;**367**:942-3.
- 9 **Hargreaves NJ**, Kadzakanja O, Whitty CJ, et al. 'Smear-negative' pulmonary tuberculosis in a DOTS programme: poor outcomes in an area of high HIV seroprevalence. *Int J Tuberc Lung Dis* 2001;**5**:847-54.
- 10 **Iseman MD**. Extrapulmonary tuberculosis in adults. In: Iseman MD, ed. *A clinician's guide to tuberculosis*. Philadelphia: Lippincott Williams and Wilkins, 2000:145-97.
- 11 **Shafer RW**, Kim DS, Weiss JP, et al. Extrapulmonary tuberculosis in patients with human immunodeficiency virus infection. *Medicine (Baltimore)* 1991;**70**:384-97.
- 12 **Golden MP**, Vikram HR. Extrapulmonary tuberculosis: an overview. *Am Fam Physician* 2005;**72**:1761-8.
- 13 **Richter C**, Ndosi B, Mwammy AS, et al. Extrapulmonary tuberculosis—a simple diagnosis? A retrospective study at Dar es Salaam, Tanzania. *Trop Geogr Med* 1991;**43**:375-8.

- 14 **d'Arminio Monforte A**, Vago L, Gori A, et al. Clinical diagnosis of mycobacterial diseases versus autopsy findings in 350 patients with AIDS. *Eur J Clin Microbiol Infect Dis* 1996;**15**:453-8.
- 15 **Rana FS**, Hawken MP, Mwachari C, et al. Autopsy study of HIV-1-positive and HIV-1-negative adult medical patients in Nairobi, Kenya. *J Acquir Immune Defic Syndr* 2000;**24**:23-9.
- 16 **Weir MR**, Thornton GF. Extrapulmonary tuberculosis. Experience of a community hospital and review of the literature. *Am J Med* 1985;**79**:467-78.
- 17 **Menzies D**. What is the current and potential role of diagnostic tests other than sputum microscopy and culture? In: Frieden T, eds. *Toman's tuberculosis: case detection, treatment, and monitoring - questions and answers*, 2nd ed. Geneva: World Health Organization, 2004:88.
- 18 **Bothamley GH**. Serological diagnosis of tuberculosis. *Eur Respir J Suppl* 1995;**20**:676-88s.
- 19 **Chan ED**, Heifets L, Iseman MD. Immunologic diagnosis of tuberculosis: a review. *Tuberc Lung Dis* 2000;**80**:131-40.
- 20 **Daniel TM**. Rapid diagnosis of tuberculosis: laboratory techniques applicable in developing countries. *Rev Infect Dis* 1989;**11**(Suppl 2):S471-8.
- 21 **Daniel TM**. The rapid diagnosis of tuberculosis: a selective review. *J Lab Clin Med* 1990;**116**:277-82.
- 22 **Gennaro ML**. Immunologic diagnosis of tuberculosis. *Clin Infect Dis* 2000;**30**(Suppl 3):S243-6.
- 23 **Iseman MD**. Immunity and pathogenesis. In: Iseman MD, eds. *A clinician's guide to tuberculosis*. Philadelphia: Lippincott Williams and Wilkins, 2000:63-96.
- 24 **Palomino JC**. Nonconventional and new methods in the diagnosis of tuberculosis: feasibility and applicability in the field. *Eur Respir J* 2005;**26**:339-50.
- 25 **Pai M**, Kalantri S, Dheda K. New tools and emerging technologies for the diagnosis of tuberculosis: part II. Active tuberculosis and drug resistance. *Expert Rev Mol Diagn* 2006;**6**:423-32.
- 26 **Laal S**. Immunodiagnosis. In: Rom WN, Garay SM, eds. *Tuberculosis*. Philadelphia: Lippincott, Williams & Wilkins, 2004:185-91.
- 27 **Laal S**, Skeiky YA. Immune-based methods. In: Cole ST, eds. *Tuberculosis and the tubercle bacillus*. Washington, DC: ASM Press, 2005:71-83.
- 28 **Boehme C**, Molokova E, Minja F, et al. Detection of mycobacterial liparabinomannan with an antigen-capture ELISA in unprocessed urine of Tanzanian patients with suspected tuberculosis. *Trans R Soc Trop Med Hyg* 2005;**99**:893-900.
- 29 **Tessema TA**, Hamasur B, Bjun G, et al. Diagnostic evaluation of urinary liparabinomannan at an Ethiopian tuberculosis centre. *Scand J Infect Dis* 2001;**33**:279-84.
- 30 **Small PM**, Perkins MD. More rigour needed in trials of new diagnostic agents for tuberculosis. *Lancet* 2000;**356**:1048-9.
- 31 **Walsh A**, McNerney R. Guidelines for establishing trials of new tests to diagnose tuberculosis in endemic countries. *Int J Tuberc Lung Dis* 2004;**8**:609-13.
- 32 **World Health Organization Special Programme for Research and Training in Tropical Diseases, Foundation for Innovative New Diagnostics**. *Diagnostics for tuberculosis: global demand and market potential*. Geneva: World Health Organization, 2006:61.
- 33 **Pai M**, McCulloch M, Enanoria W, et al. Systematic reviews of diagnostic test evaluations: what's behind the scenes? *ACP J Club* 2004;**141**:A11-3.
- 34 **Whiting P**, Rutjes AW, Reitsma JB, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;**3**:25.
- 35 **Littenberg B**, Moses LE. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. *Med Decis Making* 1993;**13**:313-21.
- 36 **SPSS for Windows**. 14.0.1.366 version. Chicago: SPSS Inc, 2006.
- 37 **Zamora J**, Abairra V, Muriel A, et al. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol* 2006;**6**:31.
- 38 **Jones CM**, Athanasiou T. Summary receiver operating characteristic curve analysis techniques in the evaluation of diagnostic tests. *Ann Thorac Surg* 2005;**79**:16-20.
- 39 **Walter SD**. Properties of the summary receiver operating characteristic (SROC) curve for diagnostic test data. *Stat Med* 2002;**21**:1237-56.
- 40 **Lijmer JG**, Bossuyt PM, Heisterkamp SH. Exploring sources of heterogeneity in systematic reviews of diagnostic tests. *Stat Med* 2002;**21**:1525-37.
- 41 **Alifano M**, De Pascalis R, Sofia M, et al. Detection of IgG and IgA against the mycobacterial antigen A60 in patients with extrapulmonary tuberculosis. *Thorax* 1998;**53**:377-80.
- 42 **Banerjee S**, Gupta S, Shende N, et al. Serodiagnosis of tuberculosis using two ELISA systems. *Indian J Clin Biochem* 2003;**18**:48-53.
- 43 **Camirero JA**, Rodriguez de Castro F, Carrillo T, et al. Diagnosis of pleural tuberculosis by detection of specific IgG anti-antigen 60 in serum and pleural fluid. *Respiration* 1993;**60**:58-62.
- 44 **Camirero JA**, Rodriguez de Castro F, Carrillo T, et al. Value of ELISA using A60 antigen in the serodiagnosis of tuberculosis. *Respiration* 1994;**61**:283-6.
- 45 **Gevaudan MJ**, Bollet C, Charpin D, et al. Serological response of tuberculosis patients to antigen 60 of BCG. *Eur J Epidemiol* 1992;**8**:666-76.
- 46 **Kunter E**, Cerrahoglu K, Ilvan A, et al. The value of pleural fluid anti-A60 IgM in BCG-vaccinated tuberculous pleurisy patients. *Clin Microbiol Infect* 2003;**9**:212-20.
- 47 **Luh KT**, Yu CJ, Yang PC, et al. Tuberculosis antigen A60 serodiagnosis in tuberculous infection: application in extrapulmonary and smear-negative pulmonary tuberculosis. *Respirology* 1996;**1**:145-51.
- 48 **McConkey SJ**, Youssef FG, Azem E, et al. Evaluation of a rapid-format antibody test and the tuberculin skin test for diagnosis of tuberculosis in two contrasting endemic settings. *Int J Tuberc Lung Dis* 2002;**6**:246-52.
- 49 **Nsanze H**, Ameen AS, Fares E, et al. Serodiagnosis of tuberculosis and leprosy by enzyme immunoassay. *Clin Microbiol Infect* 1997;**3**:236-9.
- 50 **Houghton RL**, Lodes MJ, Dillon DC, et al. Use of multi-epitope polyproteins in serodiagnosis of active tuberculosis. *Clin Diagn Lab Immunol* 2002;**9**:883-91.
- 51 **Weldingh K**, Rosenkrands I, Okkels LM, et al. Assessing the serodiagnostic potential of 35 Mycobacterium tuberculosis proteins and identification of four novel serological antigens. *J Clin Microbiol* 2005;**43**:57-65.
- 52 **Lijmer JG**, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;**282**:1061-6.
- 53 **Tatsioni A**, Zarin DA, Aronson N, et al. Challenges in systematic reviews of diagnostic technologies. *Ann Intern Med* 2005;**142**:1048-55.
- 54 **McMurray DN**. Disease model: pulmonary tuberculosis. *Trends Mol Med* 2001;**7**:135-7.
- 55 **Samanich K**, Belisle JT, Laal S. Homogeneity of antibody responses in tuberculosis patients. *Infect Immun* 2001;**69**:4600-9.
- 56 **Singh KK**, Zhang X, Patibandla AS, et al. Antigens of Mycobacterium tuberculosis expressed during preclinical tuberculosis: serological immunodominance of proteins with repetitive amino acid sequences. *Infect Immun* 2001;**69**:4185-91.
- 57 **Talaat AM**, Lyons R, Howard ST, et al. The temporal expression profile of Mycobacterium tuberculosis infection in mice. *Proc Natl Acad Sci USA* 2004;**101**:4602-7.
- 58 **Delogu G**, Sanguinetti M, Pusceddu C, et al. PE_PGRS proteins are differentially expressed by Mycobacterium tuberculosis in host tissues. *Microbes Infect* 2006;**8**:2061-7.
- 59 **Sartain MJ**, Slayden RA, Singh KK, et al. Disease state differentiation and identification of tuberculosis biomarkers via native antigen array profiling. *Mol Cell Proteomics* 2006;**5**:2102-13.
- 60 **Elliott AM**, Namaambo K, Allen BW, et al. Negative sputum smear results in HIV-positive patients with pulmonary tuberculosis in Lusaka, Zambia. *Tuberc Lung Dis* 1993;**74**:191-4.
- 61 **Johnson JL**, Vjecha MJ, Okwera A, et al. Impact of human immunodeficiency virus type-1 infection on the initial bacteriologic and radiographic manifestations of pulmonary tuberculosis in Uganda. Makerere University-Case Western Reserve University Research Collaboration. *Int J Tuberc Lung Dis* 1998;**2**:397-404.
- 62 **De Milito A**, Nilsson A, Titanji K, et al. Mechanisms of hypergammaglobulinemia and impaired antigen-specific humoral immunity in HIV-1 infection. *Blood* 2004;**103**:2180-6.
- 63 **Martinez-Maza O**, Crabb E, Mitsuyasu RT, et al. Infection with the human immunodeficiency virus (HIV) is associated with an in vivo increase in B lymphocyte activation and immaturity. *J Immunol* 1987;**138**:3720-4.
- 64 **Nagase H**, Agematsu K, Kitano K, et al. Mechanism of hypergammaglobulinemia by HIV infection: circulating memory B-cell reduction with plasmacytosis. *Clin Immunol* 2001;**100**:250-9.
- 65 **Laal S**, Samanich KM, Sonnenberg MG, et al. Surrogate marker of preclinical tuberculosis in human immunodeficiency virus infection: antibodies to an 88-kDa secreted antigen of Mycobacterium tuberculosis. *J Infect Dis* 1997;**176**:133-43.
- 66 **Hendrickson RC**, Douglass JF, Reynolds LD, et al. Mass spectrometric identification of Mtb81, a novel serological marker for tuberculosis. *J Clin Microbiol* 2000;**38**:2354-61.
- 67 **Singh KK**, Dong Y, Belisle JT, et al. Antigens of Mycobacterium tuberculosis recognized by antibodies during incipient, subclinical tuberculosis. *Clin Diagn Lab Immunol* 2005;**12**:354-8.
- 68 **Pai M**, O'Brien R. Tuberculosis diagnostics trials: do they lack methodological rigor? *Expert Rev Mol Diagn* 2006;**6**:509-14.
- 69 **Bossuyt PM**, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Ann Intern Med* 2003;**138**:40-4.
- 70 **Banoos S**, Bell D, Bossuyt P, et al. Evaluation of diagnostic tests for infectious diseases: general principles. *Nat Rev Microbiol* 2006;**4**:S21-31.