

RESEARCH ARTICLE

Interferon-gamma release assays for the diagnosis of extrapulmonary tuberculosis: a systematic review and meta-analysis

Lin Fan¹, Zhou Chen², Xiao-Hui Hao¹, Zhong-Yi Hu¹ & He-Ping Xiao¹

¹Tuberculosis Center for Diagnosis and Treatment, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai, China; and

²Department of Pharmacology, Fujian Medical University, Fuzhou, Fujian, China

Correspondence: He-ping Xiao, Tuberculosis Center for Diagnosis and Treatment, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai, 200433, China. Tel.: +86 216 511 5006; fax: +86 216 511 1298; e-mail: fanautomatic@gmail.com

Received 5 January 2012; revised 2 April 2012; accepted 2 April 2012.
Final version published online 18 June 2012.

DOI: 10.1111/j.1574-695X.2012.00972.x

Editor: Patrick Brennan

Keywords

IGRAs; IFN- γ ; extrapulmonary tuberculosis; meta-analysis.

Abstract

Interferon -gamma release assays (IGRAs) provide a new diagnostic method for *Mycobacterium tuberculosis* (TB) infection. However, the diagnostic value of IGRAs for extrapulmonary TB (EPTB) has not been clarified. We searched several databases and selected papers with strict inclusion criteria, evaluated the evidence of commercially available IGRAs (QuantiFERON[®]-TB Gold QFT-G or QFT-GIT and T-SPOT[®].TB) on blood and the tuberculin skin test (TST) using random effects models. Twenty studies with 1711 patients were included. After excluding indeterminate results, pooled sensitivity for the diagnosis of EPTB was 72% [95% confidence interval (CI) 65–79%] for QFT-G or GIT and 90% (95% CI, 86–93%) for T-SPOT; in high-income countries the sensitivity of QFT-G or GIT (79%, 95% CI 72–86%) was much higher than that (29%, 95% CI 14–48%) in low/middle-income countries. Pooled specificity for EPTB was 82% (95% CI 78–87%) for QFT-G or GIT and 68% (95% CI 64–73%) for T-SPOT. Pooled sensitivity of TST from four studies in high-income countries was lower than that of IGRAs. T-SPOT was more sensitive in detecting EPTB than QFT-G or GIT and TST. However, both IGRAs and TST have similar specificity for EPTB. IGRAs have limited value as diagnostic tools to screen and rule out EPTB, especially in low/middle-income countries. The immune status of patients does not affect the diagnostic accuracy of IGRAs for EPTB.

Introduction

Extrapulmonary tuberculosis (EPTB) remains a serious problem, with an estimated 9.4 million TB cases worldwide in 2009 (Lawn & Zumia, 2011). Moreover, the HIV epidemic altered the proportion of EPTB among TB patients, increasing new EPTB infection to over 15% of total TB cases (Small *et al.*, 1991). EPTB mainly includes tuberculous lymphadenitis, pleural TB, skeletal TB, central nervous system TB, abdominal TB, genitourinary TB, tuberculous pericarditis, etc. (Golden & Vikram, 2005). Early diagnosis of EPTB is essential for further treatment and control of this infectious disease. However, low bacterium detection rate and difficult access to pathological evidence make the diagnosis of EPTB more elusive than pulmonary TB.

Interferon-gamma release assays (IGRAs) are new immunologic diagnostic tools for TB based on detection

of a T-cell immune response to the TB antigens, which include early secretory antigen target (ESAT-6), culture filtrate protein (CFP-10) and a third antigen, TB7.7 (included in the newer version of the QTF Gold test *in vitro*). ESAT-6, CFP-10 and TB7.7 are encoded by genes in the region of difference-1 (RD1) which is present in *Mycobacterium tuberculosis* but absent in *Myobacterium bovis* *Bacilli Calmette-Guérin* (BCG) and most nontuberculous mycobacteria (NTM), and these antigens had been shown to be more specific in the diagnosis of TB than others (Pai *et al.*, 2004; Goletti *et al.*, 2006). At the time of writing, two commercial IGRAs are available: the enzyme-linked immunosorbent assay (ELISA)-based QuantiFERON-TB[®] Gold in-tube assay (QFT-GIT) and its predecessor the QuantiFERON[®]-TB Gold (QFT-G) test (Cellestis Ltd, Carnegie, Vic., Australia) in which interferon-gamma (IFN- γ) is measured quantitatively by

ELISA; and the ELISpot-based T-SPOT.TB™ test (Oxford Immunotec Inc., Abingdon, UK) in which the number of IFN- γ -producing T cells is measured by ELISPOT. In recent years, IGRAs have been widely used in the diagnosis of latent TB infection and active TB in some countries. Meta-analysis data have shown that the overall specificity of IGRAs was 98–100% in latent TB infection (Diel *et al.*, 2011). Pooled sensitivity for active TB in blood was 80% [95% confidence interval (CI) 75–84%] for QFT-G-IT and 81% (95% CI 78–84%) for T-SPOT TB; pooled specificity for active TB in blood was 79% (95% CI 75–82%) for QFT-GIT and 59% (95% CI 56–62%) for T-SPOT TB (Sester *et al.*, 2011). IGRAs may have a slight advantage over the tuberculin skin test (TST) in detecting latent TB infection (Diel *et al.*, 2011). However, IGRAs cannot efficiently distinguish latent TB infection from active TB, and a recent meta-analysis concluded that IGRAs had limited accuracy in diagnosing active TB in HIV-infected patients (Chen *et al.*, 2011). Whether IGRAs have similar limitation in diagnosing EPTB to the overall diagnostic accuracy of IGRAs for EPTB on blood has not yet been clearly defined. This is a problem for physicians trying to interpret these results in patients with suspected EPTB.

We performed a systematic review and meta-analysis of the accuracy of commercially available IGRAs on blood for the diagnosis of EPTB according to evidence-based criteria.

Materials and methods

This meta-analysis was conducted according to the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Moher *et al.*, 2009); we evaluated the quality of studies with two methods of standard study quality assessment: the Standards for Reporting of Diagnostic Accuracy (STARD) and the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) checklist (Bossuyt *et al.*, 2003; Whiting *et al.*, 2003).

Search strategy

We identified studies that evaluated the diagnostic value of IGRAs in patients with EPTB. We searched all studies published in English in electronic databases as follows: PubMed (1980–2011), Embase (1980–2011), Web of Science (1990–2011) and Cochrane-controlled central register of controlled trials (1990–2011). All searches were up to date as of June 2011. The following research terms were included: ‘Tuberculosis/diagnosis’ [MeSH] and ‘IGRA’ (text term), ‘EPTB, interferon gamma’ (text term), ‘interferon-gamma, diagnosis tuberculosis’ (text term), ‘pleural tuberculosis, IGRA’ (text term), ‘tuberculosis,

pleural/diagnosis’ [MeSH], ‘interferon gamma’ (text term), ‘T-cell interferon-gamma release assay, tuberculosis’ (text term), ‘T-SPOT, tuberculosis’ (text term), ‘T-cell response, tuberculosis, diagnosis’ (text term). We also identified further studies from the references in these papers.

Study selection

We read all the abstracts to select the appropriate studies on blood samples with full text available online to diagnosis EPTB. The following types of studies or letters to the editors were included that satisfied the following conditions: (1) clinical diagnosis studies using either two commercially available IGRA platforms (QFT-GIT and its predecessor QuantiFERON-TB® Gold, QFT-G test or T-SPOT®.TB) with data demonstrating diagnostic sensitivity and specificity; (2) active EPTB in participants was confirmed by culture, nucleic acid amplification and histopathological findings; (3) clinical research on EPTB (rather than latent TB, active TB without any indication of infection site and pulmonary tuberculosis); (4) studies on humans; (5) original research with data presented.

The following types of studies were excluded: (1) studies where active TB was not confirmed by *M. tuberculosis* culture or *Mycobacterium* TB (MTB) nucleic acid amplification, or characteristic histopathological findings (in mixed studies, i.e. those without these strict criteria, data were analysed for the confirmed cases separately). However, if the number of unconfirmed and confirmed patients was not presented separately, those studies were excluded to avoid selection bias. (2) Studies performed with noncommercial assays rather than commercially available QFT-G/GIT or T-SPOT.TB; (3) studies presenting data on body fluids rather than blood; (4) studies not performed according to the manufacturers’ instructions; and (5) studies not published in English, reviews, case reports, letters and studies with limited data, studies on animals, not exclusively for diagnostic research and studies with fewer than 10 subjects, papers published before 1980, and studies investigating immune response to new antigens.

Citations were independently screened by two investigators (L.F. and Z.C.) by examining titles and abstracts to identify potentially relevant studies. Disagreements were resolved by consensus. Three reviewers performed the review of titles and/or abstracts and 42 publications were retrieved for full-text review (Fig. 1).

Data extraction and quality assessment

An electronic form for data extraction was designed. Two investigators (L.F. and Z.C.) extracted the data from

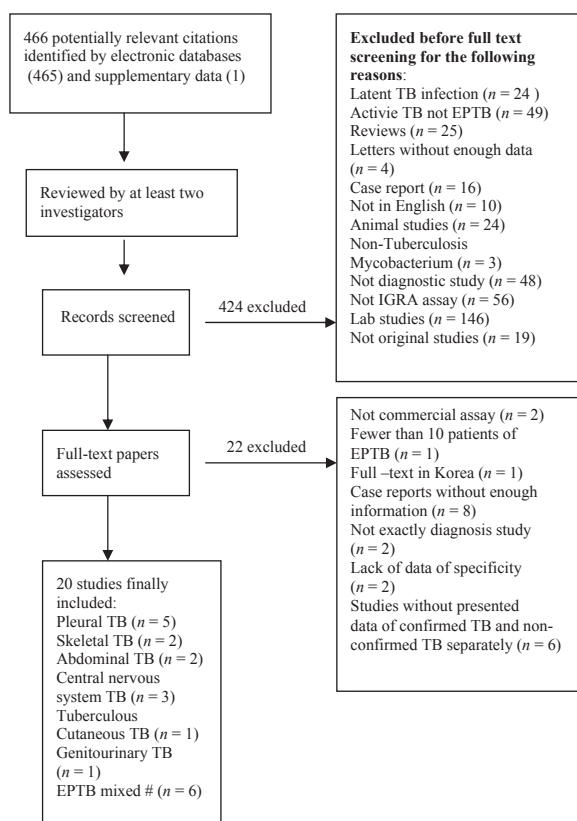


Fig. 1. Flow diagram for study selection. #EPTB mixed refers to subjects with mixed different extrapulmonary sites of active TB.

selected papers and assessed the methodological quality of studies independently. Differences were resolved by consensus.

The following data were extracted: first author, publication year, country where the study was conducted, study period, age and sex of the enrolled participants, proportion of HIV-positive participants, proportion of participants with immunosuppressive status, number of cases, BCG history, IGRA method (T-SPOT TB or QFT-GIT, QFT-G) and data (true-positive, false-positive, false-negative, true-negative and indeterminate results) from IGRAs and TST on blood, and study design (cross-sectional study or not, consecutive or not, blinded or not blinded, prospective or retrospective study). The reference standard for the diagnosis of EPTB in all included studies was positive culture or MTB PCR confirmation. The classification of countries was made with reference to The World Bank (<http://data.worldbank.org/country>), and the income of each country is inversely proportional to TB incidence (per 100 000 people), which can represent the TB burden of each country (available at <http://data.worldbank.org/about/country-classifications/country-and-lending-groups>) in which TB incidence can be searched.

Immunosuppressive status was based on participants having diabetes, liver cirrhosis, chronic renal failure, solid tumours, haematologic malignancy, HIV infection and receiving immunosuppressive treatment within 3 months (Song *et al.*, 2009).

The methodological quality of the studies was assessed using the STARD (maximum score 25) and the QUADAS (maximum score 14) tools (Bossuyt *et al.*, 2003; Whiting *et al.*, 2003, 2006). We contacted the study authors for further details about their studies when we were unable to find them in papers. If we did not get a response or did not get a satisfactory response, the data were qualified as 'unknown'.

Data synthesis and statistical analysis

We followed standard methods of the PRISMA checklist and flowchart recommended for meta-analyses and didactic guidelines for diagnostic tests (Moher *et al.*, 2009). For each study, we recalculated the sensitivity and specificity using data from the original papers; sensitivity is as used here is the proportion of individuals with a positive IGRA who had culture-confirmed EPTB, and specificity is the proportion of individuals with a negative IGRA who had TB based on another diagnosis or if there was clinical improvement without anti-TB therapy. For indeterminate tests, these results were excluded before calculation of sensitivity and specificity.

Combined estimates of sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR) and diagnostic odds ratio (DOR), together with their 95% CIs, were computed collecting all the available data (number of true positives, true negatives, false positives, and false negatives). We analysed and compared the estimates of diagnostic value from two subgroups, namely the ELISA-based QFT-GIT or its predecessor QFT-G, and ELISPOT-based T-SPOT.TB.

Forest plots were constructed to graphically assess both the variability of the estimates of the diagnostic parameters and the weight of sample sizes in the calculation of the pooled estimates (weighted means). A random-effects meta-analysis was performed to account for the expected between-study variability, along with a pooled estimate using the statistical analysis software METADISC version 1.4 (Zamora *et al.*, 2006). Inconsistency (statistical heterogeneity) among studies was assessed by the conventional chi-squared test for heterogeneity and by calculating the I^2 to highlight the effect of true variability rather than sampling error on the overall variation in diagnostic estimates. To assess the effects of the immunosuppressive status of participants and different study designs on the diagnostic values of IGRA for EPTB, we calculated the covariates in covariance meta-regression analysis. This

was performed using METADISC software in which the relative DOR (RDOR) was calculated to determine the change in diagnostic precision in the study per unit increase in the covariate (Suzuki *et al.*, 2004).

Results

Literature search

A total of 465 studies were screened for patients with EPTB diagnosed by IGRAs (Fig. 1). After independent review, 20 publications investigating the diagnosis of patients with EPTB using commercially available IGRAs (QFT-G or GIT and T-SPOT.TB) met the inclusion criteria ($n = 1711$ patients; Table 1). Among them there were four studies presenting data of TST results. After reading the papers carefully, we excluded eight studies that were done using IGRAs for the diagnosis of EPTB because one paper did not use a commercial IGRA assay and seven papers failed to present the data separately from culture-confirmed EPTB cases (Chegou *et al.*, 2008; Thomas

et al., 2008; Ang *et al.*, 2009; Dheda *et al.*, 2009; Ball *et al.*, 2010; Kim *et al.*, 2010, 2011; Ates *et al.*, 2011). In this meta-analysis, extrapulmonary sites of active TB included pleura, bone, abdomen, central nervous system, skin and genitourinary system. There were 13 studies (65%) using the ELISPOT-based T-SPOT.TBTM test, and seven studies (35%) using ELISA-based QFT-G or QFT-GIT assay. In one study (Kim *et al.*, 2009b), the subjects were classified into two groups: immunocompromised and immunocompetent, so the data were presented separately for each group; we extracted two groups of data from this study.

Characteristics of the selected studies

Table 1 shows data for 20 selected studies carried in seven different countries. One was a multicentre study from several countries, predominately India. South Korea was the most frequent country (seven of 20, 35%), followed by Taiwan (five of 20, 25%), Japan (three of 20 studies, 15%), South Africa (two of 20, 10%), Germany (one of

Table 1. Characteristics of the included studies

References	Country	Diagnosis	Methods of IGRAs used	Age [years; mean \pm SD or mean (range)]
Nishimura <i>et al.</i> (2008)	Japan	EPTB in mixed sites	QFT-G	56 (18–56)
Kobashi <i>et al.</i> (2009)*	Japan	EPTB in mixed sites	QFT-G	62 \pm 13.3
Liao <i>et al.</i> (2009)	Taiwan	EPTB in mixed sites	T-SPOT.TB	53.2 (13–93)
Lai <i>et al.</i> (2011b)	Taiwan	Skeletal TB	T-SPOT.TB	62 \pm 15.4
Song <i>et al.</i> (2009)	Korea	EPTB in mixed sites	QFT-G	47 (16–96)
Kim <i>et al.</i> (2009b)*	Korea	EPTB in mixed sites	T-SPOT.TB	48.3 \pm 17.8 (TB) 46.2 \pm 17.8 (Not TB)
Lai <i>et al.</i> (2010)	Taiwan	Genitourinary TB	T-SPOT.TB	62.9 \pm 12.4
Kim <i>et al.</i> (2007)*	Korea	EPTB in mixed sites	T-SPOT.TB	51.2 \pm 15.2 (TB) 47.7 \pm 17.3 (Not TB)
Kim <i>et al.</i> (2009a)	Korea	Abdominal TB	T-SPOT.TB	48.5 \pm 16.6 (TB) 53.7 \pm 14.9 (Not TB)
Kim <i>et al.</i> (2008)	Korea	Central nervous system TB	T-SPOT.TB	45.5 \pm 16.5 (TB) 39.3 \pm 16.5 (Not TB)
Cho <i>et al.</i> (2010)*	Korea	Osteoarticular TB	T-SPOT.TB	52.3 \pm 15.5 (TB) 58.8 \pm 16.6 (Not TB)
Ariga <i>et al.</i> (2007)	Japan	Pleural TB	QFT-G	67.9 \pm 17.1
Losi <i>et al.</i> (2007)	Germany	Pleural TB	T-SPOT.TB	69 (21–72)
Baba <i>et al.</i> (2008)	South Africa	Pleural TB	QFT-GIT	39 (20–70)
Lee <i>et al.</i> (2009)	Taiwan	Pleural TB	T-SPOT.TB	60.5 (21–102)
Cho <i>et al.</i> (2011)	Korea	Peritonitis TB	T-SPOT.TB	48.6 \pm 19.6 (TB) 49.9 \pm 16.0 (Not TB)
Patel <i>et al.</i> (2010)	South Africa	TB meningitis	T-SPOT TB	33.5 \pm 9.5 (TB) 32.9 \pm 9.7 (Not TB)
Vidhate <i>et al.</i> (2011)	India	TB meningitis	QFT-G-IT	28.9 \pm 11.8 (TB) 34.5 \pm 17.1 (Not TB)
Losi <i>et al.</i> (2011)	Italy	Pleural TB	QFT-G-IT	32 \pm 16 (TB) 71.5 \pm 15 (Not-TB)
Lai <i>et al.</i> (2011a)	Taiwan	Cutaneous TB	T-SPOT.TB	55.2 \pm 16.4

*Studies with TST data.

20), India (one of 20) and Italy (one of 20). Most of the included studies (17/20, 85%) were conducted in high-income countries (Japan, Republic of Korea, Taiwan, Germany, Italy), with the remainder in low/middle-income countries (South Africa, India). Most of the studies recruited adult patients with the exception of one study which recruited a few of children with tuberculous meningitis (Vidhate *et al.*, 2011). The median (range) sample size was 85.6 ± 52.1 . Males are more represented than females among the 20 studies (total ratio of males and females of 1.14). The mean age of participants was 50.7 ± 12.8 years in the QFT-G or QFT-GIT studies, and 53.1 ± 9.7 years in T-SPOT studies. Participants with BCT vaccination history, which was presented in only three studies, comprised 43.35% of the total recruited cases. All included studies were prospective (Table 2).

Study quality

Overall quality, evaluated via the STARD tool, was high and study quality indicators were met by 62.5% (15/24 items in the STARD tool) in all included studies. When

evaluated by the QUADAS tool, overall quality was also high, study quality indicators were met by 85.7% (12/14 items in the QUADAS tool) in 13 (65%) studies and 64.3% (9/14 items) in another seven studies, thereby increasing the scientific strength of the review.

Sensitivity of IGRAs

The median (SD) proportion of indeterminate results in QFT-G or GIT was 6.23% (8.54), slightly higher than 2.48% (3.28) in T-SPOT.TB. After excluding the indeterminate results, the pooled sensitivity of QFT-G or QFT-GIT was 72% (95% CI 65–79%, $I^2 = 80.5\%$, $P = 0.0000$; Fig. 2a); that of QFT-G or QFT-GIT in high-income countries was higher (79%, 95% CI 72–86%, $I^2 = 0\%$, $P = 0.59$) than that (29%, 95% CI 14–48%, $I^2 = 72.5\%$, $P = 0.06$) in low/middle-income countries. The pooled sensitivity of T-SPOT.TB was 90% (95% CI 86–93%, $I^2 = 38.4\%$, $P = 0.07$, Fig. 2b); T-SPOT.TB in high-income countries showed a sensitivity of 89% (95% CI 85–93%, $I^2 = 29.7\%$, $P = 0.15$), which was similar to that in the one study (89%, 95% CI 75–97%) conducted in

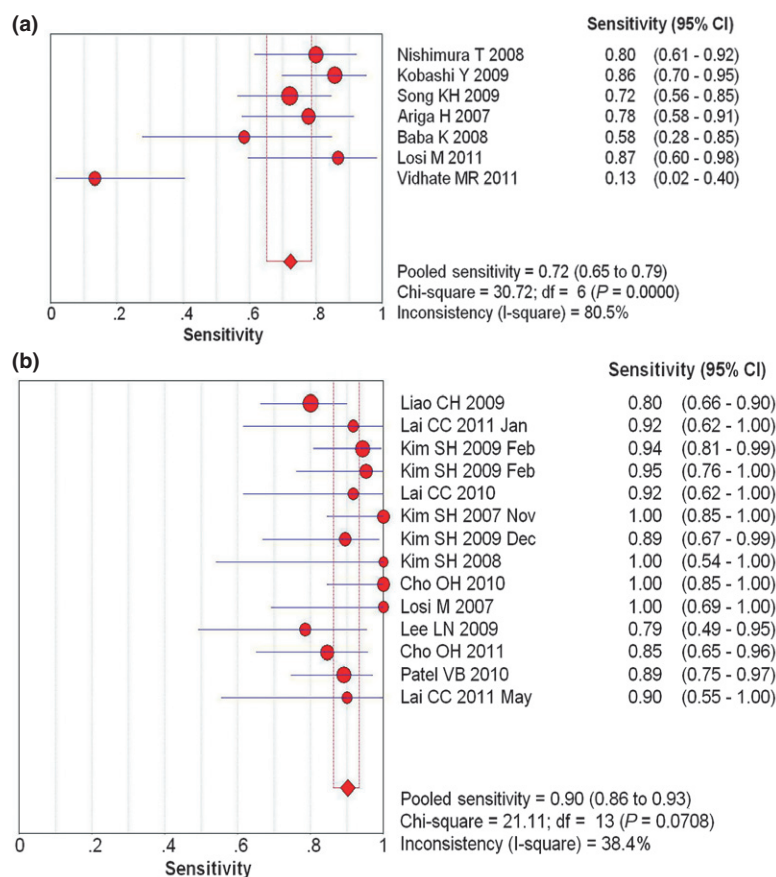


Fig. 2. Forest plot of estimate of sensitivity using QFT-G or GIT (a) and T-SPOT.TB (b) for the diagnosis of EPTB.

low/middle-income countries. If indeterminate results were deemed negative, both results decreased: sensitivity of QFT-G or QFT-GIT to 69% (95% CI 62–76%, $I^2 = 80.8\%$, $P = 0.000$), and sensitivity of T-SPOT.TB to 89% (95% CI 85–92%, $I^2 = 29.4$, $P = 0.14$). If we analyse the subgroups according to immune status, the pooled sensitivities in studies without immunosuppressive hosts were 69% (95% CI 58–78%) for QFT and 96% (95% CI 85–99%) for T-SPOT, similar to values in studies with immunocompromised patients (76%, 95% CI 65–84% for QFT; 89%, 95% CI 85–93% for T-SPOT).

Specificity of IGRAs

The pooled specificity of QFT-G or GIT was 82% (95% CI 78–87%, $I^2 = 76.8\%$, $P < 0.001$, Fig. 3a); QFT-G or GIT conducted in high-income countries showed a specificity of 83% (95% CI 78–87%, $I^2 = 80.2\%$, $P < 0.001$), slightly higher than (71%, 95% CI 48–89%, $I^2 = 74.7\%$, $P = 0.05$) in low/middle-income countries. The pooled specificity of T-SPOT.TB was 68% (95% CI 64–73%,

$I^2 = 74.2\%$, $P < 0.001$; Fig. 3b), T-SPOT.TB conducted in high-income countries showed a specificity of 73% (95% CI 68–77%, $I^2 = 46.4\%$, $P = 0.03$), much higher than that of 34% (95% CI 21–49%) conducted in only one low/middle-income country. If we analyse the specificity according to immune status subgroups, the pooled specificity in studies without immune-compromised hosts was 80% (95% CI 74–85%) for QFT and 65% (95% CI 54–75%) for T-SPOT, similar to values in studies with immune-compromised patients (87%, 95% CI 79–93% for QFT; 69%, 95% CI 64–74% for T-SPOT).

Characteristics of TST described in the studies

Among the studies included, four used a TST test, all of which were conducted in high-income countries (Republic of Korea, Japan). The pooled sensitivity of TST from these four studies was 59% (95% CI 51–68%, $I^2 = 65.9\%$, $P = 0.02$), pooled specificity was 71% (95% CI 64–77%, $I^2 = 78.5\%$, $P = 0.001$). Excluding the data from immunocompromised patients in one study which

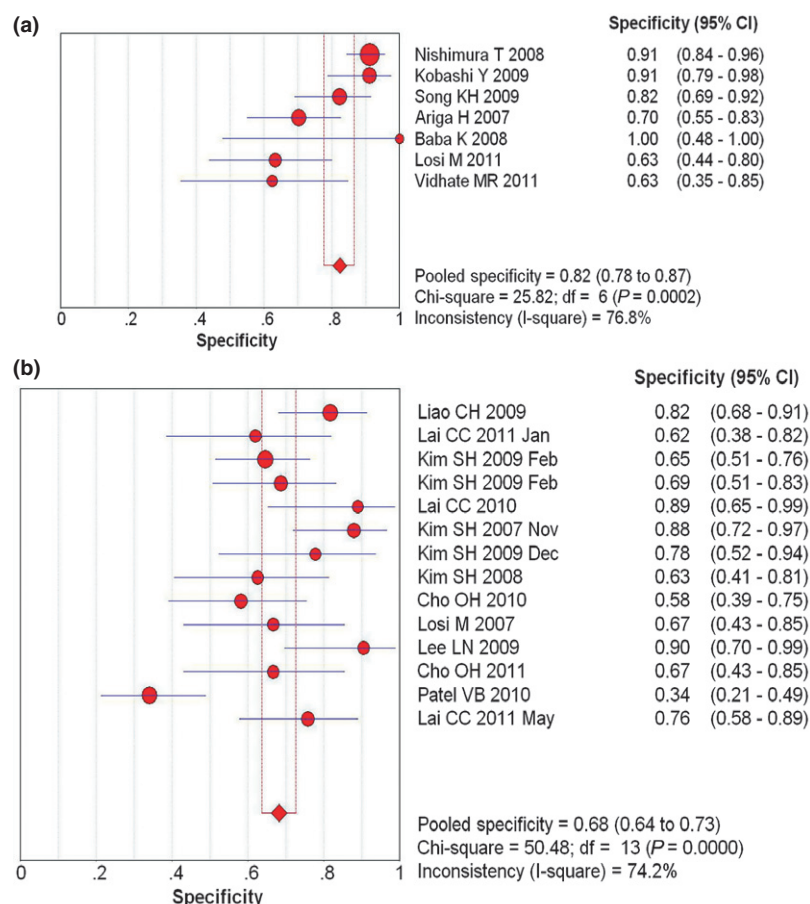


Fig. 3. Forest plot of estimate of specificity of QFT-G or GIT (a) and T-SPOT.TB (b).

have two groups of participants with different immune status (Kim *et al.*, 2009b), the pooled sensitivity increased to 63% (95% CI 54–72%, $I^2 = 57.8\%$, $P = 0.07$; Fig. 4a), heterogeneity of sensitivity decreased and pooled specificity decreased to 68% (95% CI 60–75%, $I^2 = 73.6\%$, $P = 0.01$, Fig. 4b), indicating the immunocompromised status of participants may affect the diagnostic sensitivity of TST.

Multiple regression analysis

Table 3 assesses the potential effect of immunosuppressive status and different study designs on diagnostic accuracy of IGRAs using metaregression analysis. Table 3 demonstrates that factors such as immunosuppressive status, blinded design, cross-sectional design and consecutive did not affect the diagnostic accuracy of IGRAs for EPTB (all P values non-significant).

Discussion

EPTB is difficult to diagnose because of its atypical clinical presentation and low rate of culture positivity. Pathological diagnosis usually requires invasive surgical procedures, which are expensive and traumatic for patients. New immunodiagnostic IGRAs are required to efficiently diagnose EPTB. However, in this meta-analysis, the data showed that neither ELISA-based IGRAs nor ELISPOT-based IGRAs can provide enough sensitivity and specificity for the diagnosis of EPTB, especially for patients in low/middle-income countries which have high TB burdens; this is in agreement with other researchers

(Metcalf *et al.*, 2010; Rangaka *et al.*, 2012). As it is well known that IGRAs cannot separate TB infection from active TB disease, it may explain the poor specificity of IGRAs in distinguishing EPTB from latent TB infection and other diseases.

Four important messages may be derived from this meta-analysis (1) the sensitivity of QFT-G or GIT in high-income countries was much higher than that in low/middle-income countries, but it is still not high enough to screen for EPTB; (2) the sensitivity of QFT-G/GIT or specificities of QFT-G/GIT and T-SPOT.TB were very low in low/middle-income countries; (3) in high-income countries both T-SPOT.TB and QFT-G/GIT had advantages over screening for EPTB than TST; and (4) immunosuppressive status of patients does not affect the diagnostic accuracy of IGRAs for EPTB.

Regarding diagnosis of culture-confirmed EPTB, the overall sensitivity of T-SPOT.TB (90%) was higher than that of QFT-G or GIT (72%); the sensitivity of T-SPOT.TB was similar in high- and low/middle-income countries while the sensitivity of QFT-G or GIT in high-income countries was obviously higher (79%) than that (29%) in low/middle-income countries. This demonstrates that T-SPOT was superior to QFT-G or GIT in diagnostic sensitivity for EPTB, in agreement with another meta-analysis of IGRAs for latent TB infection (Pai *et al.*, 2008). However, a meta-analysis of IGRAs in the diagnosis of active TB in HIV-positive patients (Chen *et al.*, 2011) showed different results, with similar sensitivity of QFT-GIT and T-SPOT (76.7% and 77.4%, respectively). We conjectured that the different analysis results may result from different groups of participants in different

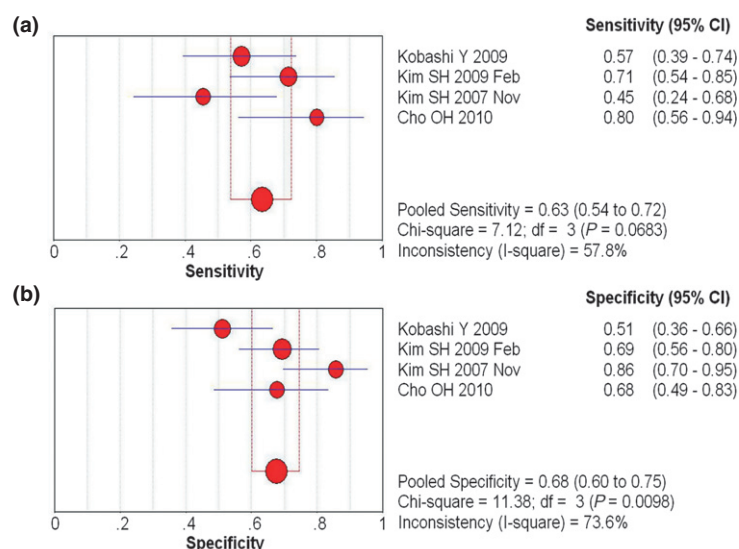


Fig. 4. Forest plot of estimate of sensitivity (a) and specificity (b) of TST from four studies in high-income countries.

Table 2. Quality and design of studies included

References	No. of participants	Cross-sectional	Consecutive or random	Blinded design	Prospective study	Immunosuppressive status (%)*
Nishimura <i>et al.</i> (2008)	144	Yes	Yes	Yes	Yes	0
Kobashi <i>et al.</i> (2009)	110	Yes	Yes	Unknown	Yes	41.8
Liao <i>et al.</i> (2009)	138	Yes	Yes	No	Yes	50.7
Lai <i>et al.</i> (2011b)	36	Yes	Yes	Unknown	Yes	44.7
Song <i>et al.</i> (2009)	100	Yes	Yes	No	Yes	15
Kim <i>et al.</i> (2009b)	179	Yes	Yes	No	Yes	0 and 100†
Lai <i>et al.</i> (2010)	36	Yes	Yes	Unknown	Yes	66.7
Kim <i>et al.</i> (2007)	179	Yes	Yes	No	Yes	49.3
Kim <i>et al.</i> (2009a)	48	Yes	Yes	Yes	Yes	27.1
Kim <i>et al.</i> (2008)	37	Yes	Yes	No	Yes	24.3
Cho <i>et al.</i> (2010)	65	Yes	Yes	Yes	Yes	10.9
Ariga <i>et al.</i> (2007)	155	Yes	Yes	Unknown	Yes	0
Losi <i>et al.</i> (2007)	41	Yes	Unknown	Unknown	Yes	0
Baba <i>et al.</i> (2008)	34	No	Yes	No	Yes	82
Lee <i>et al.</i> (2009)	40	Yes	Yes	Unknown	Yes	62.5
Cho <i>et al.</i> (2011)	64	Yes	Yes	Yes	Yes	38.2
Patel <i>et al.</i> (2010)	140	Yes	Yes	Yes	Yes	94.19
Vidhate <i>et al.</i> (2011)	52	No	No	Unknown	Yes	0
Losi <i>et al.</i> (2011)	68	Yes	Yes	Unknown	Yes	0
Lai <i>et al.</i> (2011a)	45	Yes	Yes	Unknown	Yes	57.8

*Ratio of participants with immunocompromised status to total cases.

†The participants in this study included two groups: one group had 100% with immunocompromised status, one group with immunocompetent status.

Table 3. Metaregression (inverse variance weights)

Var	Coeff.	SE	P	RDOR	95% CI
(a) Metaregression on seven studies with QFT-G or QFT-GIT					
Cte	-2.281	1.2930	0.1759	—	—
S	-0.495	0.4188	0.3220	—	—
Immunosuppressive*	1.363	1.0921	0.2801	3.91	0.19–81.03
Cross sectional	4.143	1.7214	0.0738	63.00	0.53–7497.74
Blinded	1.917	1.7512	0.3352	6.80	0.05–878.81
Consecutive	4.143	1.7214	0.0738	63.00	0.53–7497.74
(b) Metaregression on 13 studies with T-SPOT.TB					
Cte	3.526	1.6131	0.0566	—	—
S	0.074	0.2314	0.7553	—	—
Immunosuppressive*	-0.491	0.3426	0.1853	0.61	0.28–1.33
Consecutive	0.200	1.6326	0.9053	1.22	0.03–49.06
Blinded	-1.151	0.5693	0.0739	0.32	0.09–1.15

*Immunosuppressive status was evaluated based on the percentage of subjects with immunosuppressive status enrolled studies: 0, 0–40% of subjects with immunosuppressive status; 1, 41–80% of subjects with immunosuppressive status; 2, 81–100% of subjects with immunosuppressive status. In (a), all studies were performed prospectively, whereas in (b) all studies were performed prospectively and were cross-sectional.

countries and populations, or that HIV infection decreased the ability of T-SPOT to screen for active TB. The sensitivity of QFT was very low in low/middle-income countries in the present meta-analysis; this might be due to depressed immune response to TB antigens in severe or disseminated active TB as there were more severe TB cases in low/middle-income countries, and a decreased IFN- γ response might result in more false

negative QFT results in the diagnosis of EPTB (Vekemans *et al.*, 2001; Vidhate *et al.*, 2011). The other possibility might be cut-off values or spots of IGRAs: all included studies performed IGRAs according to standard manufacturer's instructions, and cut-off values among the different countries were slightly different, but this did not influence heterogeneity of meta-analysis as we calculated the threshold effects (data not shown). Although IGRAs

might not be influenced by BCG and most NTM, the optimal cut-off values in countries with a high TB burden need to be explored further.

The overall specificity of QFT-G/GIT (82%) was slightly higher than that of T-SPOT (68%), and values for QFT-G/GIT and T-SPOT in high-income countries (83% and 73%, respectively) were higher than those in low/middle-income countries (71% and 34%). The low specificity of T-SPOT in one study (Patel *et al.*, 2010) from a low/middle-income country indicates that IGRAs completely failed to efficiently distinguish EPTB from latent TB infection, indicting populations from counties with high TB burden might influence the diagnostic specificity of IGRAs for EPTB, a conclusion also reached in another meta-analysis of IGRAs for pulmonary TB (Metcalf *et al.*, 2011).

Compared with data for TST for the diagnosis of EPTB from four studies, all of which were conducted in high-income countries, the data indicated that IGRAs are superior for detecting culture-confirmed EPTB, especially in high-income countries: the sensitivity of QFT-G/GIT (79%) or T-SPOT (90%) was obviously higher than that of TST (59%). This conclusion was also reached by Diel *et al.* (2010) and Sester *et al.* (2011). However, the specificity of the two commercial IGRAs (83% and 73%) had no obvious advantage over TST (71%), and the low and similar specificity of IGRAs for EPTB in low/middle-income countries demonstrated that IGRAs cannot replace TST in the diagnosis of active TB (Lee *et al.*, 2011; Machingaidze *et al.*, 2011; Mahomed *et al.*, 2011; Mandalakas *et al.*, 2011). However, we considered that IGRAs had one advantage over TST because IGRAs results remained uninfluenced by BCG (Sun *et al.*, 2011).

Stratified analysis, in which we divided participants into three subgroups according to the proportion of participants with immunosuppressive status in total cases and subgroup analysis, showed that immunosuppressive status did not influence the diagnosis accuracy of IGRAs for EPTB. One study (Kim *et al.*, 2009a, b) also found that immunosuppressive condition did not affect the diagnostic sensitivity of T-SPOT for EPTB, but further research is needed to clarify this.

There were some potential limitations in our meta-analysis. We selected studies published in English only, limiting our ability to gather more valuable papers in non-English. Most of the studies included were conducted in high-income countries with a low TB burden; the number of studies from low/middle-income countries that met our included criteria was limited. In addition, we focused only on IGRAs on blood in this analysis rather than body fluid due to the heterogeneous protocols for the latter. In addition, we subdivided the studies into those from high-income and low/middle-income coun-

tries: heterogeneity decreased greatly after subgroup analysis but still exists in some results, which may need to be explored further.

In conclusion, this systematic review and meta-analysis shows that IGRAs cannot provide sufficient sensitivity and specificity for the diagnosis of EPTB, especially in low/middle-income countries. Immunosuppressive condition does not affect the diagnostic accuracy of IGRAs for EPTB. Further modifications are necessary to improve the diagnostic efficiency of IGRAs.

Acknowledgements

We thank Lauren Mcvov at New York University Langone Medical Center for corrections to the English text.

References

- Ang M, Htoon HM & Chee SP (2009) Diagnosis of tuberculous uveitis: clinical application of an interferon-gamma release assay. *Ophthalmology* **116**: 1391–1396.
- Ariga H, Kawabe Y, Nagai H *et al.* (2007) Diagnosis of active tuberculous serositis by antigen-specific interferon-gamma response of cavity fluid cells. *Clin Infect Dis* **45**: 1559–1567.
- Ates G, Yildiz T, Ortakoylu MG, Ozekinci T, Erturk B, Akyildiz L & Caglar E (2011) Adapted T cell interferon-gamma release assay for the diagnosis of pleural tuberculosis. *Respiration* **82**: 351–357.
- Baba K, Sørnes S, Hoosen AA, Lekabe JM, Mpe MJ, Langeland N & Dyrhol-Riise AM (2008) Evaluation of immune responses in HIV infected patients with pleural tuberculosis by the QuantiFERON TB-Gold interferon-gamma assay. *BMC Infect Dis* **8**: 35.
- Ball PM, Pernollet M, Bouillet L, Maurin M, Pavese P, Quesada JL, Romanet JP & Chiquet C (2010) Usefulness of an *in-vitro* tuberculosis interferon- γ release assay (T-SPOT. TB) in the first-line check-up of uveitis patients. *Ann Med* **42**: 546–554.
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie Dde Vet HC (2003) Standards for reporting of diagnostic accuracy: towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ* **326**: 41–44.
- Chegou NN, Walzl G, Bolliger CT, Diacon AH & van den Heuvel MM (2008) Evaluation of adapted whole-blood interferon-gamma release assays for the diagnosis of pleural tuberculosis. *Respiration* **76**: 131–138.
- Chen J, Zhang R, Wang J, Liu L, Zheng Y, Shen Y, Qi T & Lu H (2011) Interferon-gamma release assays for the diagnosis of active tuberculosis in HIV-infected patients: a systematic review and meta-analysis. *PLoS ONE* **6**: e26827.
- Cho OH, Park SJ, Park KH *et al.* (2010) Diagnostic usefulness of a T-cell-based assay for osteoarticular tuberculosis. *J Infect* **61**: 228–234.

- Cho OH, Park KH, Park SJ *et al.* (2011) Rapid diagnosis of tuberculous peritonitis by T cell-based assays on peripheral blood and peritoneal fluid mononuclear cells. *J Infect* **62**: 462–471.
- Dheda K, van Zyl-Smit RN, Sechi LA *et al.* (2009) Utility of quantitative T-cell responses versus unstimulated interferon- γ for the diagnosis of pleural tuberculosis. *Eur Respir J* **34**: 1118–1126.
- Diel R, Loddenkemper R & Nienhaus A (2010) Evidence-based comparison of commercial interferon-gamma release assays for detecting active TB: a meta analysis. *Chest* **137**: 952–968.
- Diel R, Goletti D, Ferrara G *et al.* (2011) Interferon- γ release assays for the diagnosis of latent *Mycobacterium tuberculosis* infection: a systematic review and meta-analysis. *Eur Respir J* **37**: 88–99.
- Golden MP & Vikram HR (2005) Extrapulmonary tuberculosis: an overview. *Am Fam Physician* **72**: 1761–1768.
- Goletti D, Carrara S, Vincenti D, Saltini C, Rizzi EB, Schinina V, Ippolito G, Amicosante M & Girardi E (2006) Accuracy of an immune diagnostic assay based on RD1 selected epitopes for active tuberculosis in a clinical setting: a pilot study. *Clin Microbiol Infect* **12**: 544–550.
- Kim SH, Choi SJ, Kim HB, Kim NJ, Oh MD & Choe KW (2007) Diagnostic usefulness of a T-cell based assay for extrapulmonary tuberculosis. *Arch Intern Med* **167**: 2255–2259.
- Kim SH, Chu K, Choi SJ, Song KH, Kim HB, Kim NJ, Park SH, Yoon BW, Oh MD & Choe KW (2008) Diagnosis of central nervous system tuberculosis by T-cell-based assays on peripheral blood and cerebrospinal fluid mononuclear cells. *Clin Vaccine Immunol* **15**: 1356–1362.
- Kim SH, Cho OH, Park SJ, Ye BD, Sung H, Kim MN, Lee SO, Choi SH, Woo JH & Kim YS (2009a) Diagnosis of abdominal tuberculosis by T-cell-based assays on peripheral blood and peritoneal fluid mononuclear cells. *J Infect* **59**: 409–415.
- Kim SH, Song KH, Choi SJ, Kim HB, Kim NJ, Oh MD & Choe KW (2009b) Diagnostic usefulness of a T-cell-based assay for extrapulmonary tuberculosis in immunocompromised patients. *Am J Med* **122**: 189–195.
- Kim SH, Cho OH, Park SJ *et al.* (2010) Rapid diagnosis of tuberculous meningitis by T cell-based assays on peripheral blood and cerebrospinal fluid mononuclear cells. *Clin Infect Dis* **50**: 1349–1358.
- Kim YK, Uh Y, Lee NS, Cho MY, Eom M & Kim HY (2011) Whole-blood interferon-gamma release assay for diagnosis of tuberculous lymphadenitis. *Tohoku J Exp Med* **224**: 189–193.
- Kobashi Y, Mouri K, Yagi S, Obase Y, Miyashita N & Oka M (2009) Clinical utility of a T cell-based assay in the diagnosis of extrapulmonary tuberculosis. *Respirology* **14**: 276–281.
- Lai CC, Tan CK, Lin SH, Liao CH, Huang YT, Wang CY, Wang JY, Lin HI & Hsueh PR (2010) Diagnostic value of an enzyme-linked immunospot assay for interferon- γ in genitourinary tuberculosis. *Diagn Microbiol Infect Dis* **68**: 247–250.
- Lai CC, Tan CK, Lin SH, Liu WL, Liao CH, Huang YT & Hsueh PR (2011a) Diagnostic value of an enzyme-linked immunospot assay for interferon- γ in cutaneous tuberculosis. *Diagn Microbiol Infect Dis* **70**: 60–64.
- Lai CC, Tan CK, Liu WL, Lin SH, Huang YT, Liao CH & Hsueh PR (2011b) Diagnostic performance of an enzyme-linked immunospot assay for interferon- γ in skeletal tuberculosis. *Eur J Clin Microbiol Infect Dis* **30**: 767–771.
- Lawn SD & Zumia AI (2011) Tuberculosis. *Lancet* **378**: 57–72.
- Lee LN, Chou CH, Wang JY, Hsu HL, Tsai TH, Jan IS, Hsueh PR & Yang PC (2009) Enzyme-linked immunospot assay for interferon-gamma in the diagnosis of tuberculous pleurisy. *Clin Microbiol Infect* **15**: 173–179.
- Lee JE, Kim HJ & Lee SW (2011) The clinical utility of tuberculin skin test and interferon- γ release assay in the diagnosis of active tuberculosis among young adults: a prospective observational study. *BMC Infect Dis* **11**: 96.
- Liao CH, Chou CH, Lai CC, Huang YT, Tan CK, Hsu HL & Hsueh PR (2009) Diagnostic performance of an enzyme-linked immunospot assay for interferon-gamma in extrapulmonary tuberculosis varies between different sites of disease. *J Infect* **59**: 402–408.
- Losi M, Bossink A, Codecasa L *et al.* (2007) European Tuberculosis Network TBNET. Use of a T-cell interferon-gamma release assay for the diagnosis of tuberculous pleurisy. *Eur Respir J* **30**: 1173–1179.
- Losi M, Bocchino M, Matarese A *et al.* (2011) Role of the quantiferon-TB test in ruling out pleural tuberculosis: a multi-centre study. *Int J Immunopathol Pharmacol* **24**: 159–165.
- Machingaidze S, Wiysonge CS, Gonzalez-Angulo Y, Hatherill M, Moyo S, Hanekom W & Mahomed H (2011) The utility of an interferon gamma release assay for diagnosis of latent tuberculosis infection and disease in children: a systematic review and meta-analysis. *Pediatr Infect Dis J* **30**: 694–700.
- Mahomed H, Hawkridge T, Verver S, Abrahams D, Geiter L, Hatherill M, Ehrlich R, Hanekom WA & Hussey GD (2011) The tuberculin skin test versus QuantiFERON TB Gold® in predicting tuberculosis disease in an adolescent cohort study in South Africa. *PLoS ONE* **6**: e17984.
- Mandalakas AM, Detjen AK, Hesseling AC, Benedetti A & Menzies D (2011) Interferon-gamma release assays and childhood tuberculosis: systematic review and meta-analysis. *Int J Tuberc Lung Dis* **15**: 1018–1032.
- Metcalfe JZ, Cattamanchi A, Vittinghoff E, Ho C, Grinsdale J, Hopewell PC, Kawamura LM & Nahid P (2010) Evaluation of quantitative IFN-gamma response for risk stratification of active tuberculosis suspects. *Am J Respir Crit Care Med* **181**: 87–93.
- Metcalfe JZ, Everett CK, Steingart KR, Cattamanchi A, Huang L, Hopewell PC & Pai M (2011) Interferon- γ release assays for active pulmonary tuberculosis diagnosis in adults in low- and middle-income countries: systematic review and meta-analysis. *J Infect Dis* **204**(Suppl 4): S1120–S1129.

- Moher D, Liberati A, Tetzlaff J, Altman DG & PRISMA Group (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* **6**: e1000097.
- Nishimura T, Hasegawa N, Mori M, Takebayashi T, Harada N, Higuchi K, Tasaka S & Ishizaka A (2008) Accuracy of an interferon-gamma release assay to detect active pulmonary and extra-pulmonary tuberculosis. *Int J Tuberc Lung Dis* **12**: 269–274.
- Pai M, Riley LW & Colford JM Jr (2004) Interferon-gamma assays in the immunodiagnosis of tuberculosis: a systematic review. *Lancet Infect Dis* **4**: 761–776.
- Pai M, Zwerling A & Menzies D (2008) Systematic review: T-cell-based assays for the diagnosis of latent tuberculosis infection: an update. *Ann Intern Med* **149**: 177–184.
- Patel VB, Singh R, Connolly C, Coovadia Y, Peer AK, Parag P, Kasproicz V, Zumla A, Ndung'u T & Dheda K (2010) Cerebrospinal T-cell responses aid in the diagnosis of tuberculous meningitis in a human immunodeficiency virus- and tuberculosis-endemic population. *Am J Respir Crit Care Med* **182**: 569–577.
- Rangaka MX, Gideon HP, Wilkinson KA *et al.* (2012) No discriminatory value of interferon release added to smear negative HIV-tuberculosis algorithms. *Eur Respir J* **39**: 163–171.
- Sester M, Sotgiu G, Lange C *et al.* (2011) Interferon- γ release assays for the diagnosis of active tuberculosis: a systematic review and meta-analysis. *Eur Respir J* **37**: 100–111.
- Small PM, Schechter GF, Goodman PC, Sande MA, Chaisson RE & Hopewell PC (1991) Treatment of tuberculosis in patients with advanced human immunodeficiency virus with infection. *N Engl J Med* **324**: 289–294.
- Song KH, Jeon JH, Park WB, Kim SH, Park KU, Kim NJ, Oh MD, Kim HB & Choe KW (2009) Usefulness of the whole-blood interferon-gamma release assay for diagnosis of extrapulmonary tuberculosis. *Diagn Microbiol Infect Dis* **63**: 182–187.
- Sun L, Xiao J, Miao Q *et al.* (2011) Interferon gamma release assay in diagnosis of pediatric tuberculosis: a meta-analysis. *FEMS Immunol Med Microbiol* **63**: 165–173.
- Suzuki S, Moro-oka T & Choudhry NK (2004) The conditional relative odds ratio provided less biased results for comparing diagnostic test accuracy in meta-analyses. *J Clin Epidemiol* **57**: 461–469.
- Thomas MM, Hinks TS, Raghuraman S *et al.* (2008) Rapid diagnosis of *Mycobacterium tuberculosis* meningitis by enumeration of cerebrospinal fluid antigen-specific T-cells. *Int J Tuberc Lung Dis* **12**: 651–657.
- Vekemans J, Lienhardt C, Sillah JS, Wheeler JG, Lahai GP, Doherty MT, Corrah T, Andersen P, McAdam KP & Marchant A (2001) Tuberculosis contacts but not patients have higher gamma interferon responses to ESAT-6 than do community controls in The Gambia. *Infect Immun* **69**: 6554–6557.
- Vidhate MR, Singh MK, Garg RK, Verma R, Shukla R, Goel MM, Makker A & Jain A (2011) Diagnostic and prognostic value of *Mycobacterium tuberculosis* complex specific interferon gamma release assay in patients with tuberculous meningitis. *J Infect* **62**: 400–403.
- Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM & Kleijnen J (2003) The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* **3**: 25.
- Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PN & Kleijnen J (2006) Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Med Res Methodol* **6**: 9.
- Zamora J, Abaira V, Muriel A, Khan K & Coomarasamy A (2006) Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC Med Res Methodol* **6**: 31.