



# Utility of quantitative T-cell responses versus unstimulated interferon- $\gamma$ for the diagnosis of pleural tuberculosis

K. Dheda<sup>\*,#,\*†</sup>, R.N. van Zyl-Smit<sup>\*</sup>, L.A. Sechi<sup>†</sup>, M. Badri<sup>\*</sup>, R. Meldau<sup>\*</sup>, S. Meldau<sup>\*</sup>, G. Symons<sup>\*</sup>, P.L. Semple<sup>\*</sup>, A. Maredza<sup>\*</sup>, R. Dawson<sup>\*</sup>, H. Wainwright<sup>§</sup>, A. Whitelaw<sup>f</sup>, Y. Vallie<sup>\*</sup>, P. Raubenheimer<sup>\*</sup>, E.D. Bateman<sup>\*</sup> and A. Zumla<sup>†</sup>

**ABSTRACT:** The clinical utility of antigen-specific interferon (IFN)- $\gamma$  release assays (IGRAs) using pleural mononuclear cells, for the diagnosis of tuberculosis (TB), requires clarification.

We compared the diagnostic utility of unstimulated pleural IFN- $\gamma$  levels with several pleural antigen-specific T-cell IGRAs (early secretory antigenic target-6 and culture filtrate protein-10 (T-SPOT.<sup>®</sup>TB, QuantiFERON<sup>®</sup>-TB Gold In-tube), purified protein derivative (PPD) and heparin-binding haemagglutinin (HBHA)) in 78 South African TB suspects. Test results were compared against a clinical score and a reference standard.

Out of 74 evaluable subjects 48, seven and 19 had definite, probable and no TB, respectively. 11 (15%) out of 74 pleural samples (nine (19%) out of 48 of the definite TB cases) had total cell counts that were inadequate for T-cell processing. In the remaining 63 samples, the sensitivity, specificity, positive predictive value and negative predictive value of different diagnostic methods were as follows. Maximal bioclinical score: 54, 89, 92 and 43%, respectively; T-SPOT.<sup>®</sup>TB: 86, 60, 84 and 64%, respectively; QuantiFERON<sup>®</sup>-TB Gold In-tube: 57, 80, 87 and 44%, respectively; HBHA-specific IGRA: 59, 31, 64 and 27%, respectively; PPD-specific IGRA: 81, 40, 76 and 46%, respectively; and pleural fluid unstimulated IFN- $\gamma$ : 97, 100, 100 and 94%, respectively.

Unstimulated IFN- $\gamma$  was the most accurate test for distinguishing TB from non-TB effusions in a high-burden setting. The antigen-specific T-cell IGRAs were limited by suboptimal accuracy and the inability to isolate sufficient mononuclear cells to perform the assay.

**KEYWORDS:** Diagnosis, diagnostic accuracy, human, interferon- $\gamma$ , pleural effusion, tuberculosis

Pleural effusions are common in clinical practice in developing countries and have various aetiologies, including tuberculosis (TB). Annually, over half a million cases of TB-related pleural effusion (TB-PE) occur worldwide [1]. TB is out of control in Africa and accounts for a significant proportion of pleural effusions in this setting [2]; thus, making an accurate diagnosis becomes essential. Smear and culture of the pleural fluid is helpful in a minority (~30%) of cases [3]. Currently, testing for adenosine deaminase (ADA), a biomarker for TB, is used in several laboratories, but this facility is not widely available in high-burden developing countries and it is not specific for TB [3, 4]. Closed pleural biopsy has a good yield (~80%) but it is less sensitive, invasive, relatively expensive, and may be associated with complications [3]. Consequently, empirical anti-TB treatment is often administered. There is a need for alternative, rapid and less invasive methods to diagnose TB-PE, especially in developing countries.

Recently, candidate biomarkers interferon (IFN)- $\gamma$ -inducible protein-10 (a 10-kDa T-helper cell (Th) type 1 associated chemokine) [5] and unstimulated IFN- $\gamma$  levels [6] have shown promise. The latter has shown good discriminatory value in intermediate and low-burden settings [6]. However, data about the utility and performance outcome of pleural fluid unstimulated IFN- $\gamma$  levels in a high TB prevalence setting are limited.

More recently, peripheral blood quantitative region of difference (RD)-1 antigen-specific (early secretory antigenic target (ESAT)-6 and culture filtrate protein (CFP)-10) T-lymphocyte responses have become established as sensitive, specific and rapid immunodiagnostic tests for TB infection [7–9]. However, they cannot distinguish latent infection from active disease [8]. By contrast, at the site of active pleural [10] or pulmonary [11, 12] disease, the frequency of clonally expanded antigen-specific T-cells is over 10-fold higher than in peripheral

## AFFILIATIONS

\*Lung Infection and Immunity Unit and Centre for Tuberculosis Research and Innovation, University of Cape Town Lung Institute and Division of Pulmonology, Dept of Medicine, University of Cape Town,

#Institute of Infectious Disease and Molecular Medicine, University of Cape Town

§National Health Laboratory Service, Division of Anatomical Pathology, and

fNational Health Laboratory Service, Division of Medical Microbiology, Dept of Clinical Laboratory Sciences, University of Cape Town, Cape Town, South Africa.

†Centre for Infectious Disease and International Health, Dept of Infection, University College Medical School, London, UK.

\*Dept of Biomedical Sciences, University of Sassari, Sassari, Italy.

## CORRESPONDENCE

K. Dheda, Lung Infection and Immunity Unit, Division of Pulmonology, Dept of Medicine, University of Cape Town, J flr, Old Main Bldg, Grootte Schuur Hospital, Observatory, Cape Town, 7925 South Africa  
E-mail: keertan.dhedaa@uct.ac.za

Received:

Jan 13 2009

Accepted after revision:

April 07 2009

First published online:

April 22 2009

European Respiratory Journal

Print ISSN 0903-1936

Online ISSN 1399-3003

blood and should theoretically not be present in non-TB pleural effusions. We therefore hypothesised that TB antigen-specific responses could distinguish active pleural TB from alternative diagnoses and should have better discriminatory value than unstimulated IFN- $\gamma$ , a nonspecific biomarker of inflammation. A recent case-control study from Europe showed a high diagnostic accuracy with the T-SPOT.<sup>®</sup>TB assay (Oxford Immunotec, Abingdon, UK) [13] but overall data on the utility of two known IFN- $\gamma$  release assays (IGRAs), T-SPOT.<sup>®</sup>TB and QuantiFERON.<sup>®</sup>-TB Gold In-tube (QFT-TB-GIT; Cellestis, Carnegie, Australia), including a head-to-head comparison, for the diagnosis of TB-PE are limited [13–16]. There are several other unresolved questions, including the optimal cell number and pleural fluid volume required, optimal cut-off point and utility in high HIV-prevalence settings [17]. Other novel antigens, like heparin-binding haemagglutinin (HBHA), shown to be discriminatory in peripheral blood [18, 19], have hitherto not been evaluated using pleural fluid lymphocytes.

In this study, the main aim was to evaluate the diagnostic utility, in clinical practice, of several different IGRAs using pleural fluid T-cells in a high-burden TB and HIV setting. We prospectively studied the comparative performance outcomes of unstimulated pleural IFN- $\gamma$  levels against four antigen-specific IGRAs, including a head-to-head comparison between the T-SPOT.<sup>®</sup>TB and QFT-TB-GIT, using pleural fluid mononuclear cells. Comparative peripheral blood responses were also evaluated.

## METHODS

### Patient recruitment, characterisation and routine laboratory tests

Study approval was obtained from the University of Cape Town Health Sciences Faculty (Cape Town, South Africa) research ethics committee. After obtaining informed consent and giving counselling, 78 consecutive patients with clinically suspected TB-PE were prospectively recruited at the Groote Schuur, Somerset and Victoria hospitals in Cape Town over a 12-month period (ending April 30, 2008). All patients had a history taken and physical examination performed, with routine investigations including testing for HIV infection, chest radiography, microbiological sputum examination whenever possible, and aspiration of ~20 mL of pleural fluid (or closest obtainable volume) for biochemical, cytological and detailed microbiological evaluation. Four patients were excluded from further analysis (see figure 1 for details including summary of recruitment); thus, 74 patients contributed evaluable samples. For accurate characterisation of disease, multiple closed pleural biopsies were undertaken. Although offered, in 16 patients biopsies were not performed because of patient refusal, a contraindication, a positive culture of fluid or histology from another site prior to attempted pleural biopsy.

The reference standard used for diagnosis of TB was culture positivity for *Mycobacterium tuberculosis* (using pleural fluid or tissue) and/or histology suggestive of TB (caseous necrosis with acid-fast bacilli, with or without granuloma formation). Patients were thus characterised as 1) definite TB (meeting the reference standard), 2) non-TB (no microbiological or histological evidence for TB, alternative diagnosis made, not treated for TB and did not develop TB over 6-month follow-up), and 3) probable TB (empirical anti-TB treatment but not

meeting the criteria for definite TB). The definite and non-TB groups were used for sensitivity and specificity calculations. All assays were performed by an experienced laboratory technician who was blinded to patient and clinical details.

### Antigen-specific IGRAs

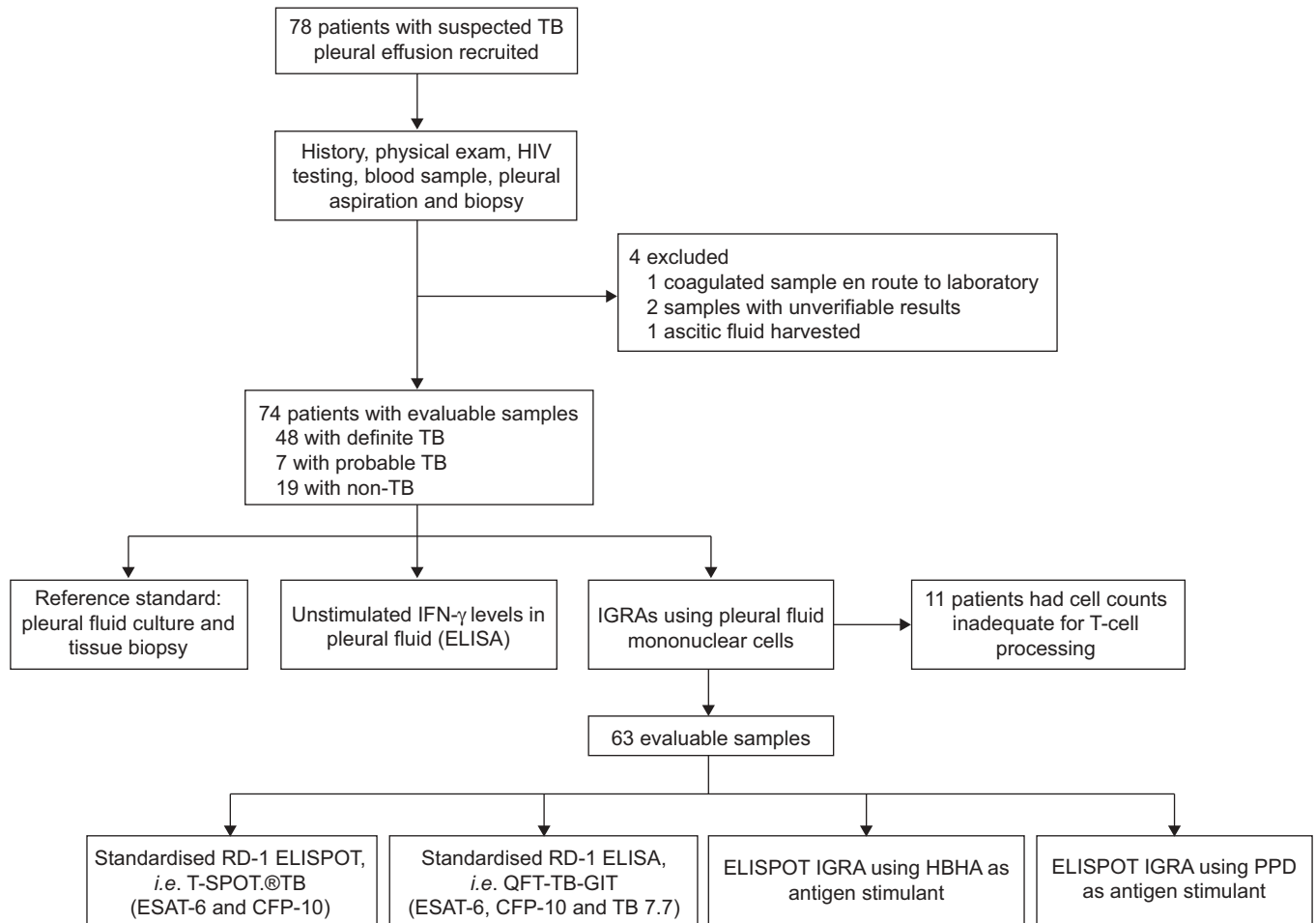
Four different IGRAs were evaluated. 1) IFN- $\gamma$  ELISPOT responses to ESAT-6/CFP-10 peptide pools were performed according to manufacturer's instructions (T-SPOT.<sup>®</sup>TB), using peripheral blood (250,000 cells·well<sup>-1</sup>) and pleural mononuclear cells (200,000 cells·well<sup>-1</sup>), and enumerated using an ELISPOT reader. Data from nine patients showed that  $2 \times 10^5$  cells·well<sup>-1</sup> produced optimal IFN- $\gamma$  responses, compared with  $1 \times 10^5$  cells·well<sup>-1</sup> (149 versus 39 spot-forming cells (SFCs) for ESAT-6;  $p=0.002$ ; and 108 versus 36 SFCs for CFP-10;  $p=0.01$ ). Using  $4 \times 10^5$  cells·well<sup>-1</sup> did not significantly increase the spot counts. 2) IFN- $\gamma$  responses using the QFT-TB-GIT assay were performed using 1 mL of blood and  $1 \times 10^6$  pleural mononuclear cells were re-suspended in 1 mL of serum-free medium in each tube. Data from six patients showed that antigen-specific responses were significantly higher when using at least  $1 \times 10^6$  pleural mononuclear cells per tube compared with  $1 \times 10^5$  cells (2.53 versus 0.026 IU·mL<sup>-1</sup>;  $p=0.008$ ); suspension of pleural cells in smaller volumes (0.25 and 0.5 mL) had no effect on measured IFN- $\gamma$  concentrations. 3) IFN- $\gamma$  ELISPOT responses to HBHA were performed using IFN- $\gamma$  pre-coated ELISPOT plates (Mabtech, Nacka Strand, Sweden). Dose-response experiments revealed that optimum IFN- $\gamma$  secretion occurred at a HBHA dose of  $100 \mu\text{g}\cdot\text{mL}^{-1}$  using peripheral mononuclear cells. Methylated HBHA antigen was produced in cultures of *Mycobacterium smegmatis* 3.38 as previously described [20]. 4) IFN- $\gamma$  ELISPOT responses to purified protein derivative (PPD;  $100 \mu\text{g}\cdot\text{mL}^{-1}$ ; Statens Serum Institut, Copenhagen, Denmark) were performed using IFN- $\gamma$  pre-coated ELISPOT plates. All assays were incubated for 16–20 h.

### Measurement of unstimulated IFN- $\gamma$

IFN- $\gamma$  concentrations were measured in the unstimulated pleural fluid supernatant using the QuantiFERON.<sup>®</sup> ELISA kit.

### Statistical analysis

Categorical variables and continuous variables were compared using the appropriate parametric and nonparametric tests. The McNemar Chi-squared test was used to compare sensitivities and specificities of the different diagnostic tests [21]. To ascertain the relative value of newer tests in a high-burden setting, a regression model was used to develop a bioclinical score [22] appropriate for a resource-poor setting. Thus, to ascertain the relative value of newer tests in a high-burden setting, regression models were fitted to identify variables independently associated with risk of TB, taking into account findings from the history, physical examination and pleural fluid biochemical data (the variables considered in the analysis included age, race, smoking, HIV status, history of previous TB, cell counts and pleural fluid protein level). Multicollinearity was assessed, and variables contributing to the best fit of the final model, or most cost-effective and widely available in our setting, were retained in the final model. The final bioclinical scoring rule was developed by assigning a relative score or points to each of the variables included in the



**FIGURE 1.** Summary and flow chart of recruited patients. In addition, blood samples were taken for the region of difference (RD)-1 ELISPOT and RD-1 ELISA interferon (IFN)- $\gamma$  release assays (IGRAs). TB: tuberculosis; ESAT: early secretory antigenic target; CFP: culture filtrate protein; QFT-TB-GIT: QuantiFERON $\text{®}$ -TB Gold In-tube; HBHA: heparin-binding haemagglutinin; PPD: purified protein derivative. The T-SPOT. $\text{®}$ TB assay and QFT-TB-GIT were manufactured by Oxford Immunotec, Abingdon, UK and Cellestis, Carnegie, Australia, respectively.

final multivariate model. Reporting of the study was done using the Standards for Reporting of Diagnostic Accuracy (STARD) template [23].

**RESULTS**

**Clinical, demographic data and biochemical data**

There were 48, 19 and seven patients with definite TB, non-TB and probable TB, respectively. Smear, pleural fluid culture and biopsy (tissue culture and histology) were positive in one, 27 and 41 of the 48 definite TB cases, respectively, and, by definition, in none of the non-TB cases. None of the probable TB cases were culture or biopsy positive but all were treated empirically for TB based on clinical suspicion. Effusions in the non-TB group were due to several causes (two lymphoma, two myeloproliferative disorders, nine adeno- or small cell carcinoma, three parapneumonic and three due to other causes). 21% (16 out of 74) of patients did not have a pleural biopsy (refused by four patients, contraindicated in two patients, one in whom a liver biopsy was done, and nine in whom the culture result was positive prior to a biopsy being done (six sputum culture positive and three pleural fluid culture positive)). Of all subjects in the study cohort tested for HIV,

46.4% (26 out of 56) were HIV positive. Clinical, demographic and biochemical data are summarised in table 1 for the 67 patients in the definite TB or non-TB groups from which sensitivities and specificities were calculated.

In the final multivariate logistic regression model, only age (<42 yrs; odds ratio (OR) 3.89, 95% confidence interval (CI) 1.01–14.90;  $p=0.04$ ) and pleural fluid protein levels ( $>53 \text{ g}\cdot\text{L}^{-1}$ ; OR 3.59, 95% CI 1.02–12.56;  $p=0.04$ ) were independently associated with the risk of TB. These variables, when incorporated into a bioclinical score, had a maximal sensitivity and specificity of 54 and 89%, respectively (table 2).

The median (25th and 75th percentiles) pleural fluid cell count was  $1.75 \times 10^6 \text{ cells}\cdot\text{mL}^{-1}$  ( $1.03 \times 10^6$  and  $5.45 \times 10^6 \text{ cells}\cdot\text{mL}^{-1}$ ), and the median volume of fluid obtained was 20 mL (10 and 25 mL). The pleural fluid ADA had a sensitivity and specificity of 96 and 69%, respectively, for the diagnosis of TB, and further details are published elsewhere [24].

**Performance outcomes of pleural fluid unstimulated IFN- $\gamma$**

The median values for unstimulated IFN- $\gamma$  were significantly higher in TB *versus* non-TB effusions ( $10.95 \text{ versus}$

**TABLE 1** Baseline characteristics of the definite tuberculosis (TB) and non-TB patients

Characteristic	Definite TB	Non-TB	p-value
<b>Patients n</b>	48	19	
<b>Age yrs</b>	37 ± 14.4	55.3 ± 15.7	<0.0001 <sup>#</sup>
<b>Sex</b>			0.99 <sup>†</sup>
Male	23 (47.9)	9 (47.4)	
Female	25 (52.1)	10 (52.6)	
<b>Race</b>			0.03 <sup>+</sup>
Black African	25 (52.1)	4 (21.1)	
Mixed	23 (47.9)	15 (78.9)	
<b>HIV status</b>			0.07 <sup>+</sup>
HIV positive	22 (45.8)	3 (15.8)	
HIV negative	16 (33.3)	10 (52.6)	
Refused testing	10 (20.8)	6 (31.6)	
<b>History of previous TB</b>			0.01 <sup>+</sup>
Yes	3 (6.2)	6 (31.6)	
No	45 (93.8)	13 (68.3)	
<b>BCG status</b>			0.44 <sup>+</sup>
Positive	17 (35.4)	8 (42.1)	
Negative	19 (37.6)	9 (47.4)	
Unknown	12 (25.0)	2 (10.5)	
<b>Current smoker</b>			0.02 <sup>†</sup>
Yes	11 (22.9)	10 (52.6)	
No	37 (77.1)	9 (47.4)	
<b>Employment status</b>			0.84 <sup>†</sup>
Employed	29 (60.4)	12 (63.2)	
Unemployed	19 (39.6)	7 (36.8)	
<b>Pleural fluid protein g·L<sup>-1</sup></b>	58.9 ± 15.7	43.4 ± 18.5	0.003 <sup>#</sup>

Data are presented as mean ± SD or n (%), unless otherwise stated. BCG: bacille Calmette–Guerin. <sup>#</sup>: unpaired t-test; <sup>†</sup>: Chi-squared test; <sup>+</sup>: Fisher exact test.

0.105 IU·mL<sup>-1</sup>;  $p < 0.001$ ). At a receiver operating curve (ROC)-derived cut-off point of 0.31 IU·mL<sup>-1</sup>, the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (95% CI) of unstimulated IFN- $\gamma$  was 97% (85–99), 100% (81–100), 100% (90–100) and 94% (73–99), respectively. Accuracy was 98% (90–100) and area under the ROC was 0.99. A scatter plot and area under the ROC of unstimulated IFN- $\gamma$  levels in pleural fluid are shown in figure 2. When probable and definite TB cases were grouped together the sensitivity of the assay was 98%.

#### Performance outcome of pleural versus peripheral blood T-SPOT.<sup>®</sup>TB and QFT-TB-GIT

Pleural T-cell assays could only be performed in 63 (85%) out of the 74 patients because inadequate cell numbers were isolated from 11 patients during the pleural fluid processing stage (median cell count in this group was  $1.3 \times 10^4$  cells·mL<sup>-1</sup> and the median volume of pleural fluid received was 14 mL). Of these 11 patients, nine had definite TB; thus, nine (19%) out of the 48 definite TB cases were unsuitable for T-cell processing. This left 63 evaluable patients where a T-cell assay had been performed on the pleural fluid.

The performance outcomes of both commercial IGRAs (T-SPOT.<sup>®</sup>TB and QFT-TB-GIT) for pleural fluid (using

**TABLE 2** Performance outcomes of a bioclinical score for use in a high-burden setting

Score	Sensitivity	Specificity	PPV	NPV	Accuracy
<b>0</b>	19 (10–31)	58 (36–76)	53 (31–73)	22 (12–35)	29 (20–42)
<b>4</b>	81 (68–89)	58 (32–76)	82 (69–91)	55 (34–78)	74 (63–83)
<b>8</b>	54 (40–67)	89 (68–97)	92 (77–98)	43 (29–59)	64 (52–74)

Data are presented as percentage (95% confidence interval). Age <42 yrs or pleural fluid protein level >53 g·L<sup>-1</sup> scored four points each; thus, a score of 0, 4 or 8 was possible. PPV: positive predictive value; NPV: negative predictive value.

mononuclear cells for both assays) and whole blood are shown in table 3. Of the seven probable TB cases (data not shown), two and three had positive pleural T-SPOT.<sup>®</sup>TB and QFT-TB-GIT results, respectively. The scatter plots for the pleural T-SPOT.<sup>®</sup>TB and pleural mononuclear cell-derived QFT-TB-GIT assays, including area under the ROC, are shown in figure 3. In summary, the pleural T-SPOT.<sup>®</sup>TB assay, compared with the pleural mononuclear cell-derived QFT-TB-GIT assay, had a better sensitivity (80 *versus* 51%;  $p = 0.002$ , McNemar Chi-squared test) but poorer specificity (65 *versus* 94%) and PPV (83 *versus* 95%).

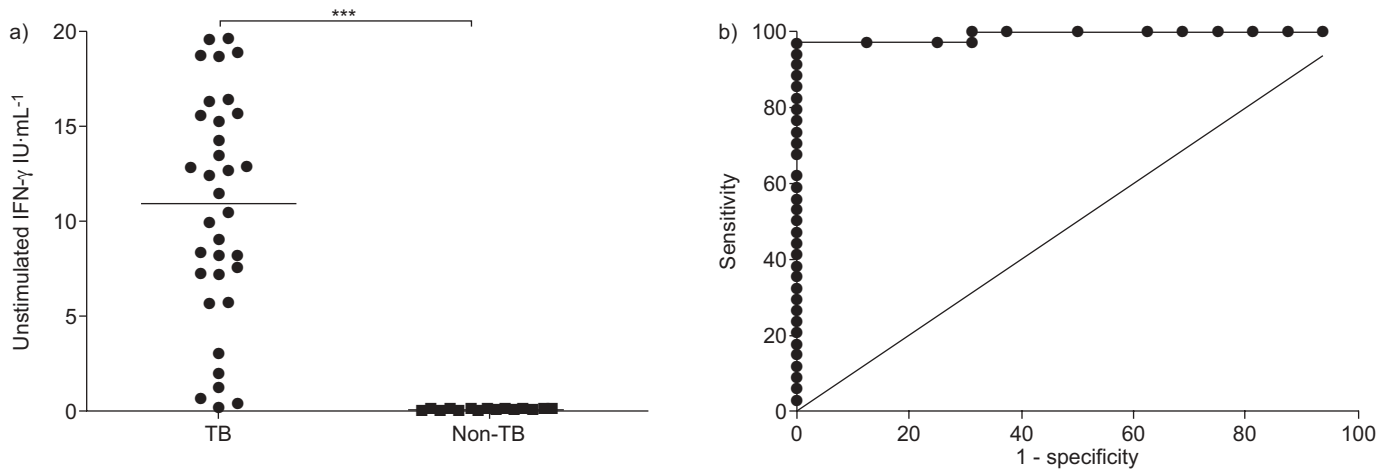
In an attempt to improve the specificity of the assay, we corrected pleural antigen-specific T-cell responses for peripheral antigen-specific T-cell counts (pleural RD-1 spot count/peripheral spot count ratio). This ratio, although higher in TB compared with non-TB patients (6.56 *versus* 2.37), was not significant ( $p = 0.11$ ). Based on the ROC (area under the ROC of 0.68), the optimal ratio cut-off point was 3.9 for ESAT-6-specific responses, which produced a sensitivity of 64% and specificity of 75% and was thus not discriminatory.

In total, 63 participants had a least one pleural T-cell assay, and their outcome categorisation by disease status is shown in table 4. There was no difference in the number of indeterminate results using the QFT-TB-GIT pleural mononuclear cell-derived assay compared with the pleural T-SPOT.<sup>®</sup>TB assay (four out of 60 *versus* six out of 61;  $p = 0.40$ , Fisher exact test). A head-to-head comparison of 53 paired pleural T-SPOT.<sup>®</sup>TB and QFT-TB-GIT assays, in TB and non-TB subjects, is shown in table 5.

There was modest agreement between the pleural RD-1 assays (56 and 53% in TB *versus* non-TB patients, respectively). Inter-assay concordance was poor (table 5;  $\kappa = 0.21$ ). When probable and definite TB cases were grouped together, the sensitivity of assays remained unchanged (85 and 81% *versus* 53 and 50%, respectively, for the two T-SPOT.<sup>®</sup>TB and QFT-TB-GIT cut-off points specified in table 3).

#### Pleural HBHA and PPD ELISPOT responses

The performance outcomes of HBHA and PPD-driven ELISPOT responses, using pleural mononuclear cells, are shown in table 3, while a scatter plot of the data together with area under the ROC is shown in figure 3f–h. In summary, both antigens had poor discriminatory value when using pleural mononuclear cells.



**FIGURE 2.** a) Scatter plot of unstimulated interferon (IFN)-γ using pleural fluid from patients with tuberculous (TB) and non-TB pleural effusions. b) The receiver operating curve (ROC) for TB versus non-TB pleural effusion. The area under the ROC was 0.99. \*\*\*: p<0.001.

Pleural fluid IFN-γ responses did not differ significantly in HIV-infected and HIV-uninfected patients (unstimulated IFN-γ median (range) 10.87 (0.7–20.1) versus 12.4 (0.1–20.1) IU·mL<sup>-1</sup>; QFT-TB-GIT 1.64 (0.0–13.2) versus 0.74 (0.0–14.0) IU·mL<sup>-1</sup>; ESAT-6 292 (0–1,250) versus 420 (0–1,250) SFCs per 10<sup>6</sup> cells; CFP-10 182 (0–1,250) versus 397 (0–1,250) SFCs per 10<sup>6</sup> cells).

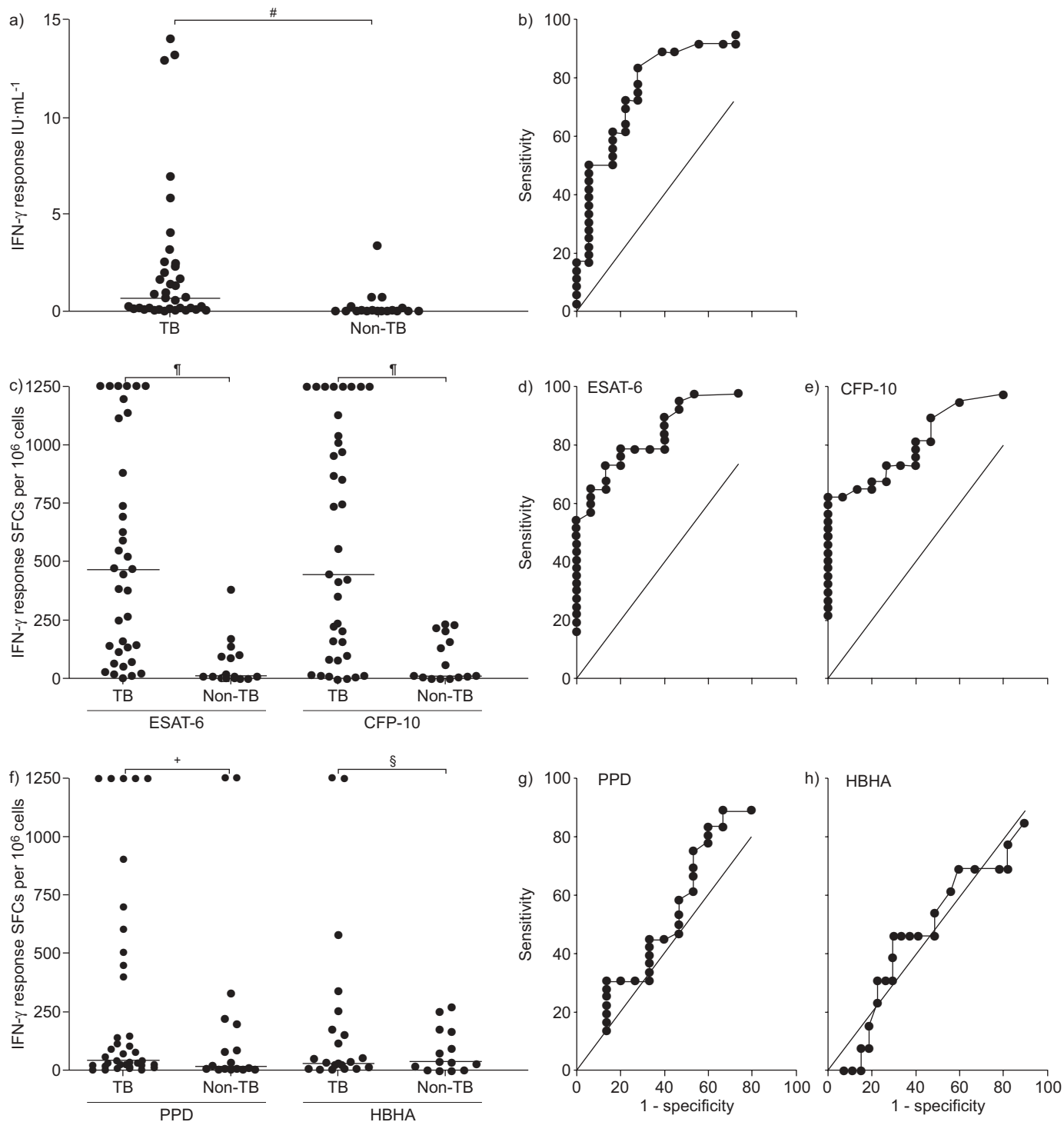
**DISCUSSION**

In this study we prospectively evaluated the clinical diagnostic utility of four different IGRAs (RD-1 ELISPOT, RD-1 ELISA,

HBHA and PPD ELISPOT) using pleural fluid mononuclear cells, which were compared against pleural fluid unstimulated IFN-γ levels, in South African patients suspected to have TB. A comparative study of the different RD-1 IGRAs, and evaluation of HBHA responses, has not hitherto been undertaken using pleural fluid lymphocytes. The IGRAs, which are technically more demanding and expensive, performed suboptimally and, at best, missed 15% of TB cases and incorrectly diagnosed a further 20%. By contrast, unstimulated IFN-γ levels in pleural fluid distinguished, with a high level of accuracy, between TB and non-TB effusions. The excellent outcomes with IFN-γ are

<b>TABLE 3</b> Performance outcomes of interferon (IFN)-γ release assays at different cut-off points in pleural tuberculosis suspects							
	Cut-off point	Sensitivity	Specificity	PPV	NPV	Accuracy	AUC
<b>Pleural fluid</b>							
T-SPOT. <sub>®</sub> TB							
Combined	≥24 SFCs <sup>#</sup>	86 (71–94)	60 (36–80)	84 (69–92)	64 (39–84)	78 (65–88)	NA
	>100 SFCs <sup>‡</sup>	80 (65–90)	65 (41–83)	83 (67–92)	61 (39–80)	76 (62–85)	
ESAT-6							0.88
	≥24 SFCs <sup>#</sup>	86 (71–94)	67 (42–84)	86 (71–94)	67 (42–85)	80 (68–89)	
	>137 SFCs <sup>‡</sup>	72 (56–84)	93 (70–98)	96 (82–100)	58 (38–76)	78 (65–87)	
CFP-10							0.84
	≥24 SFCs <sup>#</sup>	78 (62–88)	60 (36–80)	82 (66–92)	53 (31–73)	73 (59–83)	
	>225 SFCs <sup>‡</sup>	56 (40–71)	93 (70–99)	95 (77–99)	47 (30–64)	67 (53–78)	
QFT-TB-GIT							0.80
	≥0.35 IU·mL <sup>-1#</sup>	57 (41–72)	80 (55–93)	87 (68–96)	44 (28–63)	64 (50–76)	
	>0.73 IU·mL <sup>-1‡</sup>	51 (36–67)	94 (72–94)	95 (75–99)	47 (31–64)	65 (51–76)	
HBHA stimulation							0.51
	>17 SFCs <sup>‡</sup>	59 (41–75)	31 (13–58)	64 (45–80)	27 (11–52)	59 (40–75)	
	>158 SFCs <sup>‡</sup>	22 (11–41)	69 (31–83)	60 (31–83)	30 (17–48)	38 (24–53)	
PPD stimulation							0.61
	>12 SFCs <sup>‡</sup>	81 (65–90)	40 (20–64)	76 (61–87)	46 (23–71)	69 (55–80)	
	>362 SFCs <sup>‡</sup>	24 (14–38)	87 (61–96)	85 (58–96)	27 (16–41)	39 (28–52)	
<b>Peripheral whole blood</b>							
T-SPOT. <sub>®</sub> TB							
	≥24 SFCs <sup>#</sup>	83 (68–92)	56 (33–77)	81 (66–91)	60 (36–80)	75 (62–85)	NA
QFT-TB-GIT							0.76
	≥0.35 IU·mL <sup>-1#</sup>	87 (70–95)	69 (42–87)	87 (70–95)	69 (42–87)	81 (67–90)	

Data are presented as percentage (95% confidence interval), unless otherwise stated. 61 T-SPOT.<sub>®</sub>TB (Oxford Immunotec, Abingdon, UK) and 60 QuantiFERON<sub>®</sub>-TB Gold In-tube (QFT-TB-GIT; Cellestis, Carnegie, Australia) assays were performed. PPV: positive predictive value; NPV: negative predictive value; AUC: area under receiver operating curve; ESAT: early secretory antigenic target; CFP: culture filtrate protein; HBHA: heparin-binding haemagglutinin; PPD: purified protein derivative; SFCs: spot-forming cells; NA: not applicable (as this is a dichotomised variable the AUC result cannot be calculated). <sup>#</sup>: manufacturer-derived cut-off point for whole blood; <sup>‡</sup>: AUC-derived cut-off point. Comparative values for unstimulated IFN-γ levels are given in figure 2 and in the main text.



**FIGURE 3.** Scatter plots (a, c and f) of the frequency (IU·mL<sup>-1</sup> or spot-forming cells (SFCs) per 10<sup>6</sup> pleural mononuclear cells) of the antigen-specific interferon (IFN- $\gamma$ ) responses and respective area under the receiver operating curve analyses (b, d, e, g and h) for tuberculosis (TB) versus non-TB pleural effusions. a and b) Using the QuantiFERON<sup>®</sup>-TB Gold In-tube assay (Cellestis, Carnegie, Australia). c-e) Using the T-SPOT<sup>®</sup>-TB assay (Oxford Immunotec, Abingdon, UK); d) early secretory antigenic target (ESAT)-6 and e) culture filtrate protein (CFP)-10. f and g) Purified protein derivative (PPD)-specific and f and h) heparin-binding haemagglutinin (HBHA)-specific IFN- $\gamma$  responses. For data values, see table 3. #: p=0.0003; †: p<0.0001; +: p=0.21; §: p=0.94.

borne out by several other studies from geographical areas outside Africa (Asia, Europe and South America), which have been summarised in a recent meta-analysis [6], where the pooled sensitivity and specificity of IFN- $\gamma$  in unstimulated

pleural fluid was 89% and 97%, respectively (individual studies and their outcomes are outlined in the meta-analysis [6]). We confirm this finding in African HIV-infected subjects. Collectively, these data suggest that a point-of-care test

**TABLE 4** Performance outcomes, including indeterminate results, of T-SPOT.<sup>®</sup>TB<sup>#</sup> and QFT-TB-GIT<sup>†</sup>, by disease status, using pleural fluid mononuclear cells

	TB	Non-TB	Probable TB	Total
<b>T-SPOT.<sup>®</sup>TB positive</b>	31	6	3	40
<b>QFT-TB-GIT positive</b>	20	3	2	25
<b>T-SPOT.<sup>®</sup>TB negative</b>	5	9	1	15
<b>QFT-TB-GIT negative</b>	15	12	4	31
<b>T-SPOT.<sup>®</sup>TB indeterminate</b>	3	2	1	6
<b>QFT-TB-GIT indeterminate</b>	1	3	0	4

Data are presented as n. TB: tuberculosis; QFT-TB-GIT: QuantiFERON<sup>®</sup>-TB Gold In-tube assay. The T-SPOT.<sup>®</sup>TB assay and QFT-TB-GIT were manufactured by Oxford Immunotec, Abingdon, UK and Cellestis, Carnegie, Australia, respectively. <sup>#</sup>: n=61; <sup>†</sup>: n=60.

incorporating IFN- $\gamma$  testing holds promise for the diagnosis of TB-PE in a high-burden setting. This approach, however, does not allow determination of drug-susceptibility status and, hence, microbiological investigation in appropriate patients may still be required. Thus, IFN- $\gamma$  levels cannot necessarily replace culture of *M. tuberculosis*, but could be used as an adjunctive test to make rapid diagnosis and possibly reduce the need for further investigations to exclude other causes of the pleural effusion.

Two recent preliminary studies from low-burden countries (Europe and Japan) suggested that RD-1-based commercial IGRAs, using pleural mononuclear cells, appear promising as diagnostic tools for TB-PE [13, 14]. In the last several months, two case-control studies from high-burden countries (South Africa) and using QFT-TB-GIT have been published [15, 16]. Both showed that the QFT-TB-GIT assay had suboptimal sensitivity (27 and 57%, respectively). However, the T-SPOT.<sup>®</sup>TB test, thought to be more sensitive [7], was not

evaluated in these studies. Furthermore, details about optimisation and rationale of laboratory protocols are unclear, the total cell number in the QFT-TB-GIT assay was not controlled for, and pleural biopsy was not undertaken or was limited at the discretion of the operator, thus limiting the accuracy of patient categorisation. In the current study, preliminary experiments were performed to determine optimal cell number and antigen concentration, pleural biopsy was performed unless refused or contraindicated, and both RD-1 commercial platforms were evaluated.

Overall, the T-SPOT.<sup>®</sup>TB assay was more sensitive than the QFT-TB-GIT assay, in keeping with similar observations in peripheral blood [7], presumably because ELISPOT is generally regarded as a more sensitive technique [9]. Although the QFT-TB-GIT had a high PPV, it missed almost half of the TB cases. By contrast, using pleural fluid unstimulated IFN- $\gamma$  levels, a nonspecific Th1 marker of inflammation, achieved a higher PPV. Thus, the specificity of IGRAs was suboptimal despite the use of TB-specific antigens. How do we explain these results? One possibility is the translocation of blood RD-1-specific T-cells, through an inflamed or “leaky” pleura, in subjects with non-TB pleuritis (no active TB) who had a concomitant high frequency of peripheral antigen-specific T-cells and hence latent TB infection (LTBI). Six out of seven non-TB patients who had a higher than expected frequency of antigen-specific T-cells in the pleural space also had a high frequency of peripheral antigen-specific T-cells, and hence LTBI. We hypothesise that the ELISPOT was more prone to this effect because it is a more sensitive technique. Alternative explanations for the poor specificity include transient exposure to *M. tuberculosis*, organisms in a state of nonreplicating persistence in the pleural space, or dual pathology (active TB and malignancy), although on follow-up (6–12 months), we detected no clinical evidence of TB in those with alternative diagnoses. Compensating for the frequency of peripheral antigen-specific T-cells or the total pleural mononuclear cells (pleural RD-1 spot count/peripheral spot count ratio) did not significantly improve the specificity of the assay. In this

**TABLE 5** Head-to-head comparison of paired pleural fluid T-SPOT.<sup>®</sup>TB and QFT-TB-GIT results

	T-SPOT. <sup>®</sup> TB positive	T-SPOT. <sup>®</sup> TB negative	T-SPOT. <sup>®</sup> TB indeterminate	Total
<b>Definite TB<sup>#</sup></b>				
QFT-TB-GIT positive	17	0	3	20
QFT-TB-GIT negative	12	3	0	15
QFT-TB-GIT indeterminate	0	1	0	1
Total	29	4	3	36
<b>Non-TB<sup>†</sup></b>				
QFT-TB-GIT positive	2	1	0	3
QFT-TB-GIT negative	4	6	1	11
QFT-TB-GIT indeterminate	0	2	1	3
Total	6	9	2	17

Data are presented as n. TB: tuberculosis; QFT-TB-GIT: QuantiFERON<sup>®</sup>-TB Gold In-tube assay. The T-SPOT.<sup>®</sup>TB assay and QFT-TB-GIT were manufactured by Oxford Immunotec, Abingdon, UK and Cellestis, Carnegie, Australia, respectively. <sup>#</sup>: n=36;  $\kappa=0.21$ , 95% confidence interval (CI) -0.02–0.41, p=0.03; agreement 56%. <sup>†</sup>: n=17;  $\kappa=0.20$ , 95% CI -0.27–0.67, p=0.20; agreement 53%.

compensation, we reasoned that in cases of active pleural TB there would be a greater proportion of antigen-specific cells and hence a high ratio, and in cases of LTBI there would be fewer such cells and hence the ratio would be low. Adjustment of the cut-off point improved the specificity but compromised on the sensitivity.

By contrast, the unstimulated pleural fluid IFN- $\gamma$  was highly sensitive and specific for pleural TB. This potent, *M. tuberculosis*-driven, compartment-specific, Th1 response is not characteristic of non-mycobacterial or tumour antigens. Thus, we believe, based on the data presented here and from other studies [6], that unstimulated IFN- $\gamma$  should be used as a discriminatory diagnostic tool in clinical practice. Further research is now required to develop user-friendly formats to measure unstimulated IFN- $\gamma$  levels in pleural fluid.

The optimal cell number and pleural fluid volume required to perform pleural T-cell assays is unclear [17]. Hence, an additional aim of our study was to clarify this aspect. We show that obtaining adequate cell numbers to perform the IGRA was an important practical limitation of the assay. In most cases, <10 mL of pleural fluid was more than adequate for the required experiments. However, in some cases, where effusions were fibrinous or loculated, a limited volume of pleural fluid was obtainable (<20 mL), or the fluid was pauci-cellular, thus limiting the number of cells that were isolated. Additional factors, probably more frequent in high-burden settings, that may have modulated the number of viable cells obtained include the following: unequal distribution of cells in the pleural fluid, cell loss during processing, lymphocyte-poor and neutrophil-predominant effusions, excessive debris and degenerating cells in complicated effusions, and pauci-cellular pleural inflammation in HIV-positive subjects (six out of the eight patients with inadequate cell numbers, who were tested, had advanced HIV disease). Thus, for successful T-cell assays to be performed, it is crucial that adequate volumes of pleural fluid are obtained. In contrast to previous reports [15], to enable precise quantification of the number of cells being used in the assay, we washed pleural mononuclear cells and re-incubated them in serum-free media. This had the additional advantage of minimising indeterminate results due to high readouts in the negative control well/tube. There was only modest inter-assay agreement/concordance and this may reflect differences in the technique (ELISA *versus* ELISPOT), cell numbers and antigen cocktails used (QFT-TB-GIT has an additional peptide antigen, TB 7.7). For this study all samples were processed in a category 3 containment laboratory. However, this may not always be possible in a resource-poor setting where safety precautions may be in line with local policy and regulations.

To meaningfully evaluate the relative clinical value of newer and established tests, we compared their utility to a simple bioclinical score, generated through regression analysis, and relevant to a resource-poor setting [25, 26]. We also investigated the effect of HIV status on test performance outcomes. All tests performed equally well in HIV-positive and -negative patients.

In the ELISPOT assay we also tested a novel antigen, HBHA, a virulence factor that mediates binding of *M. tuberculosis* to alveolar epithelial cells and whose antigenicity is dependent on methylation of the C-terminal domain, which is absent from

recombinant forms of HBHA [19]. HBHA responses were recently found to be a sensitive marker of TB infection [18, 27]. We therefore used *M. smegmatis* electroporated with a plasmid vector to produce methylated *M. tuberculosis* HBHA [20, 28, 29]. Human tuberculous pleural fluid was recently shown to contain high titres of anti-HBHA antibody [30], but the utility of antigen-specific HBHA responses for the diagnosis of TB-PE is unknown. We found generally poor responses to methylated HBHA compared with the RD-1 antigens. The reasons for this are unclear but may be related to the altered pleural fluid regulatory T-cell profile in TB-PE [31], which is known to attenuate HBHA responses [32] and may underlie the proposed ability of this antigen to distinguish latent from active TB [18]. The differential effect of *M. smegmatis* *versus* *M. tuberculosis* methylated HBHA on IFN- $\gamma$  responses, if any, is unclear.

Investigator bias in our study was minimised through several steps that ensured our study's validity, including consecutive recruitment with universally applied and pre-specified inclusion criteria, an experienced operator blinded to clinical details, invasive procedures to ensure accurate classification of patient and control subgroups, and use of a pre-specified reference standard. However, our results are probably generalisable only to high TB/HIV burden settings. Thus, further and larger studies are required in order to evaluate whether outcomes are different in low-burden settings using the cut-off points identified here, and the effect of HIV infection on IFN- $\gamma$ -related pleural assays. The high background rate of LTBI and possible transient infection are likely to have impacted heavily on the T-cell assay results in the non-TB group.

In conclusion, unstimulated IFN- $\gamma$  was the most accurate test to distinguish TB from non-TB effusions in a high-burden setting. In this well categorised cohort of unselected patients from a high-burden setting, the antigen-specific T-cell IGRAs were limited by suboptimal accuracy and the inability to isolate sufficient mononuclear cells to perform the assay.

#### SUPPORT STATEMENT

K. Dheda is supported by the South African National Research Foundation (SARChI), the South African Medical Research Council, the EC Framework 7 programme and the European and Developing Countries Clinical Trials Partnership (EDCTP), and by the University College London–University of Cape Town Collaboration Initiative. A. Zumla is in receipt of funding from the EC Framework 7 programme and EDCTP, from the Comprehensive Biomedical Research Centre and National Institute for Health Research (London, UK), and from EuropeAID (Brussels, Belgium).

#### STATEMENT OF INTEREST

None declared.

#### ACKNOWLEDGEMENTS

The authors thank P. Govender, H. Khalfey, A. Aboo, I. Carr (all Dept of Medicine, Groote Schuur Hospital, Cape Town, South Africa), R. van Zyl-Smit (Dept of Medicine, Somerset Hospital, Cape Town), and registered nurses C. Cooper and M. Isaacs (both Dept of Medicine, Groote Schuur Hospital) for facilitating patient recruitment. We are also grateful to S. Zanetti (Dept of Biomedical Sciences, University of Sassari, Sassari, Italy) and G. Delogu (Institute of Microbiology, Catholic University of the Sacred Heart, Rome, Italy) for the gift of the pMV 3-38.



## REFERENCES

- 1 Gopi A, Madhavan SM, Sharma SK, *et al.* Diagnosis and treatment of tuberculous pleural effusion in 2006. *Chest* 2007; 131: 880–889.
- 2 Sinzobahamvya N, Bhakta HP. Pleural exudate in a tropical hospital. *Eur Respir J* 1989; 2: 145–148.
- 3 Trajman A, Pai M, Dheda K, *et al.* Novel tests for diagnosing tuberculous pleural effusion: what works and what does not? *Eur Respir J* 2008; 31: 1098–1106.
- 4 Greco S, Girardi E, Masciangelo R, *et al.* Adenosine deaminase and interferon- $\gamma$  measurements for the diagnosis of tuberculous pleurisy: a meta-analysis. *Int J Tuberc Lung Dis* 2003; 7: 777–786.
- 5 Pokkali S, Das SD, Logamurthy R. Expression of CXC and CC type of chemokines and its receptors in tuberculous and non-tuberculous effusions. *Cytokine* 2008; 41: 307–314.
- 6 Jiang J, Shi HZ, Liang QL, *et al.* Diagnostic value of interferon- $\gamma$  in tuberculous pleurisy: a metaanalysis. *Chest* 2007; 131: 1133–1141.
- 7 Menzies D, Pai M, Comstock G. Meta-analysis: new tests for the diagnosis of latent tuberculosis infection: areas of uncertainty and recommendations for research. *Ann Intern Med* 2007; 146: 340–354.
- 8 Pai M, Dheda K, Cunningham J, *et al.* T-cell assays for the diagnosis of latent tuberculosis infection: moving the research agenda forward. *Lancet Infect Dis* 2007; 7: 428–438.
- 9 Pai M, Zwerling A, Menzies D. Systematic review: T-cell-based assays for the diagnosis of latent tuberculosis infection: an update. *Ann Intern Med* 2008; 149: 177–184.
- 10 Wilkinson KA, Wilkinson RJ, Pathan A, *et al.* *Ex vivo* characterization of early secretory antigenic target 6-specific T cells at sites of active disease in pleural tuberculosis. *Clin Infect Dis* 2005; 40: 184–187.
- 11 Jafari C, Ernst M, Strassburg A, *et al.* Local immunodiagnosis of pulmonary tuberculosis by enzyme-linked immunospot. *Eur Respir J* 2008; 31: 261–265.
- 12 Jafari C, Ernst M, Kalsdorf B, *et al.* Rapid diagnosis of smear-negative tuberculosis by bronchoalveolar lavage enzyme-linked immunospot. *Am J Respir Crit Care Med* 2006; 174: 1048–1054.
- 13 Losi M, Bossink A, Codecasa L, *et al.* Use of a T-cell interferon- $\gamma$  release assay for the diagnosis of tuberculous pleurisy. *Eur Respir J* 2007; 30: 1173–1179.
- 14 Ariga H, Kawabe Y, Nagai H, *et al.* Diagnosis of active tuberculous serositis by antigen-specific interferon- $\gamma$  response of cavity fluid cells. *Clin Infect Dis* 2007; 45: 1559–1567.
- 15 Baba K, Sornes S, Hoosen AA, *et al.* Evaluation of immune responses in HIV infected patients with pleural tuberculosis by the QuantiFERON TB-Gold interferon- $\gamma$  assay. *BMC Infect Dis* 2008; 8: 35.
- 16 Chegou NN, Walzl G, Bolliger CT, *et al.* Evaluation of adapted whole-blood interferon- $\gamma$  release assays for the diagnosis of pleural tuberculosis. *Respiration* 2008; 76: 131–138.
- 17 Joshi R, Pai M. Can pleural tuberculosis be diagnosed using interferon- $\gamma$  release assays? *Respiration* 2008; 76: 128–130.
- 18 Hougardy JM, Schepers K, Place S, *et al.* Heparin-binding-hemagglutinin-induced IFN- $\gamma$  release as a diagnostic tool for latent tuberculosis. *PLoS One* 2007; 2: e926.
- 19 Locht C, Hougardy JM, Rouanet C, *et al.* Heparin-binding hemagglutinin, from an extrapulmonary dissemination factor to a powerful diagnostic and protective antigen against tuberculosis. *Tuberculosis (Edinb)* 2006; 86: 303–309.
- 20 Delogu G, Bua A, Pusceddu C, *et al.* Expression and purification of recombinant methylated HBHA in *Mycobacterium smegmatis*. *FEMS Microbiol Lett* 2004; 239: 33–39.
- 21 Trajman A, Luiz RR. McNemar Chi-squared test revisited: comparing sensitivity and specificity of diagnostic examinations. *Scand J Clin Lab Invest* 2008; 68: 77–80.
- 22 Moons KG, Harrell FE, Steyerberg EW. Should scoring rules be based on odds ratios or regression coefficients? *J Clin Epidemiol* 2002; 55: 1054–1055.
- 23 Bossuyt PM, Reitsma JB, Bruns DE, *et al.* Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Standards for Reporting of Diagnostic Accuracy. *Clin Chem* 2003; 49: 1–6.
- 24 Dheda K, Van-Zyl Smit RN, Sechi LA, *et al.* Clinical diagnostic utility of IP-10 and LAM antigen levels for the diagnosis of tuberculous pleural effusions in a high burden setting. *PLoS One* 2009; 4: e4689.
- 25 Neves DD, Dias RM, Cunha AJ. Predictive model for the diagnosis of tuberculous pleural effusion. *Braz J Infect Dis* 2007; 11: 83–88.
- 26 Porcel JM, Vives M. Differentiating tuberculous from malignant pleural effusions: a scoring model. *Med Sci Monit* 2003; 9: CR175–CR180.
- 27 Temmerman S, Pethe K, Parra M, *et al.* Methylation-dependent T cell immunity to *Mycobacterium tuberculosis* heparin-binding hemagglutinin. *Nat Med* 2004; 10: 935–941.
- 28 Delogu G, Sanguinetti M, Posteraro B, *et al.* The hbhA gene of *Mycobacterium tuberculosis* is specifically upregulated in the lungs but not in the spleens of aerogenically infected mice. *Infect Immun* 2006; 74: 3006–3011.
- 29 Zanetti S, Bua A, Delogu G, *et al.* Patients with pulmonary tuberculosis develop a strong humoral response against methylated heparin-binding hemagglutinin. *Clin Diagn Lab Immunol* 2005; 12: 1135–1138.
- 30 Yang CS, Lee JS, Lee HM, *et al.* Differential cytokine levels and immunoreactivities against *Mycobacterium tuberculosis* antigens between tuberculous and malignant effusions. *Respir Med* 2008; 102: 280–286.
- 31 Chen X, Zhou B, Li M, *et al.* CD4+/CD25+/FoxP3+ regulatory T cells suppress *Mycobacterium tuberculosis* immunity in patients with active disease. *Clin Immunol* 2007; 123: 50–59.
- 32 Hougardy JM, Place S, Hildebrand M, *et al.* Regulatory T cells depress immune responses to protective antigens in active tuberculosis. *Am J Respir Crit Care Med* 2007; 176: 409–416.