

Published in final edited form as:

*Lancet Infect Dis.* 2012 January ; 12(1): 45–55. doi:10.1016/S1473-3099(11)70210-9.

## Predictive value of interferon- $\gamma$ release assays for incident active tuberculosis: a systematic review and meta-analysis

Molebogeng X Rangaka, Katalin A Wilkinson, Judith R Glynn, Daphne Ling, Dick Menzies, Judith Mwansa-Kambafwile, Katherine Fielding, Robert J Wilkinson, and Madhukar Pai (M X Rangaka MBChB), **Clinical Infectious Diseases Research Initiative, Institute of Infectious Diseases and Molecular Medicine** (M X Rangaka, K A Wilkinson PhD, J Mwansa-Kambafwile MBChB, R J Wilkinson FRCP), **Department of Medicine, School of Health Sciences** (R J Wilkinson), **University of Cape Town, South Africa; London School of Hygiene and Tropical Medicine, London, UK** (M X Rangaka, Prof J R Glynn PhD, K Fielding PhD); **Medical Research Council (MRC) National Institute for Medical Research, Mill Hill, London, UK** (K A Wilkinson, R J Wilkinson); **Division of Medicine, Imperial College London, UK** (R J Wilkinson); and **McGill University and Montreal Chest Institute, Montreal, Canada** (D Ling MPH, D Menzies MD, M Pai MD)

### Summary

**Background**—We aimed to assess whether interferon- $\gamma$  release assays (IGRAs) can predict the development of active tuberculosis and whether the predictive ability of these tests is better than that of the tuberculin skin test (TST).

**Methods**—Longitudinal studies of the predictive value for active tuberculosis of in-house or commercial IGRAs were identified through searches of PubMed, Embase, Biosis, and Web of Science and complementary manual searches up to June 30, 2011. Eligible studies included adults or children, with or without HIV, who were free of active tuberculosis at study baseline. We summarised incidence rates in forest plots and pooled data with random-effects models when appropriate. We calculated incidence rate ratios (IRR) for rates of disease progression in IGRA-positive versus IGRA-negative individuals.

**Findings**—15 studies had a combined sample size of 26 680 participants. Incidence of tuberculosis during a median follow-up of 4 years (IQR 2–6), even in IGRA-positive individuals, was 4–48 cases per 1000 person-years. Seven studies with no possibility of incorporation bias and reporting baseline stratification on the basis of IGRA results showed a moderate association between positive results and subsequent tuberculosis (pooled unadjusted IRR 2.10, 95% CI 1.42–3.08). Compared with test-negative results, IGRA-positive and TST-positive results were much the same with regard to the risk of tuberculosis (pooled IRR in the five studies that used both was 2.11 [95% CI 1.29–3.46] for IGRA vs 1.60 [0.94–2.72] for TST at the 10 mm cutoff). However,

---

Correspondence to: Dr Madhukar Pai, McGill University, Department of Epidemiology and Biostatistics, 1020 Pine Avenue West, Montreal, QC H3A 1A2, Canada madhukar.pai@mcgill.ca.

**Contributors** MXR, JG, and MP designed the study. MXR and MP did the data searches and study selection. MXR and DL extracted the data. MXR did the data analyses. All authors interpreted the data. MXR wrote the first draft and all authors helped to revise the paper.

**Conflicts of interest** MXR, DM, and MP participated in the WHO Expert Group Meeting on IGRAs, held in Geneva, in July 2010. Their role as systematic reviewers was to present the evidence for IGRAs in a series of systematic reviews. They had no role in making the final recommendations on IGRAs. They had no role in the Strategic and Technical Advisory Group for Tuberculosis endorsement of final WHO Expert Group Meeting recommendations on IGRAs. MP is as an external consultant for the Bill & Melinda Gates Foundation, and is as co-chair of the Stop TB Partnership's New Diagnostics Working Group—neither organisation is involved in this work. All other authors declare that they have no conflicts of interest.

the proportion of IGRA-positive individuals in seven of 11 studies that assessed both IGRAs and TST was generally lower than TST-positive individuals.

**Interpretation**—Neither IGRAs nor the TST have high accuracy for the prediction of active tuberculosis, although use of IGRAs in some populations might reduce the number of people considered for preventive treatment. Until more predictive biomarkers are identified, existing tests for latent tuberculosis infection should be chosen on the basis of relative specificity in different populations, logistics, cost, and patients' preferences rather than on predictive ability alone.

**Funding**—Special Programme for Research and Training in Tropical Diseases (WHO), Wellcome Trust, Canadian Institutes of Health Research, UK Medical Research Council, and the European and Developing Countries Clinical Trials Partnership.

---

## Introduction

A third of the world's population is estimated to be infected with *Mycobacterium tuberculosis*,<sup>1</sup> providing a very large reservoir for future active tuberculosis. The tuberculin skin test (TST) has traditionally been used to identify people with latent *M tuberculosis* infection who will benefit from isoniazid preventive treatment.<sup>2</sup> Despite its usefulness and simplicity, the TST has limitations—its specificity is affected by BCG vaccination and its predictive value for incident tuberculosis disease is low.

T-cell-based interferon- $\gamma$  release assays (IGRAs) can also be used for the diagnosis of *M tuberculosis* infection, and have been available for the past decade. Two licensed IGRAs are commercially available: QuantiFERON TB Gold in tube (Cellestis, Carnegie, Victoria, Australia) and T-SPOT.TB (Oxford Immunotec, Abingdon, UK). Studies of these assays have been reviewed for their ability to identify latent infection and to diagnose active disease in various populations.<sup>3–7</sup> IGRAs, like the TST, are a surrogate marker for *M tuberculosis* infection, indicating a cellular immune response to recent or remote sensitisation. However, neither assay can distinguish between latent and active tuberculosis.<sup>8</sup> Nevertheless, use of IGRAs in tuberculosis programmes is encouraged in many countries with low or intermediate incidence.<sup>9</sup>

The clinical benefit of IGRAs can be proven only if individuals identified as having latent tuberculosis infection by IGRA are at increased risk of active tuberculosis compared with test-negative individuals and if these individuals benefit substantially from preventive treatment. TST positivity is a surrogate marker for risk of subsequent tuberculosis (with those testing positive having a relative risk of about 2), and TST-positive individuals benefit from isoniazid preventive treatment.<sup>2,10–15</sup> To show equivalent or superior clinical value to TSTs, IGRAs should be assessed in various at-risk subgroups. Such longitudinal data are emerging, but have not been systematically reviewed.

We previously published systematic and narrative reviews on IGRA accuracy and performance in various subgroups.<sup>4–8</sup> We did a systematic review and subsequent meta-analysis to assess whether IGRAs can prospectively predict the development of active tuberculosis in individuals without active disease at baseline and whether that predictive ability is higher than that of the TST. Our secondary objectives were to compare rates of tuberculosis in IGRA-positive individuals with TST-negative or TST-positive results who received isoniazid preventive treatment, to assess the effect of immunological phenotypes of discordant-concordant TST and IGRA results at baseline on subsequent tuberculosis rates, to establish whether a gradient association exists between quantitative interferon- $\gamma$  response and rates of progression to tuberculosis disease, and to assess estimates of false-positive or false-negative IGRA results versus TST results.

## Methods

### Search strategy and study selection

We updated the database searches (with the same terms) that were done in previous systematic reviews<sup>4–8</sup> and searched PubMed, Embase, Biosis, and Web of Science for relevant IGRA studies (up to June 30, 2011) that reported data on IGRA predictive value in all settings. We reviewed citations of all original articles published in all languages. In addition to electronic database searches, we reviewed bibliographies of previous reviews and guidelines on IGRA and screened the citations of relevant original articles. We contacted experts to obtain relevant citations. No language restrictions were made and full-length papers, conference proceedings, and abstracts were included. When necessary, we contacted authors of primary studies to obtain additional information.

Studies were eligible for inclusion if they were longitudinal studies of adults or children, with or without HIV, who were free of active disease at study baseline that specifically stated assessment of the predictive ability of IGRA as a primary objective, had any longitudinal study design (eg, prospective or retrospective cohort) in any setting (low-income, middle-income, or high-income country), and described either active or passive follow-up of patients for any duration. Index tests assessed were any IGRA for *M tuberculosis* infection (whole-blood ELISA, enzyme-linked immunosorbent spot [ELISPOT], in-house laboratory-developed non-commercial assay, or the latest generation of commercially licensed assays) that included at least one region of difference 1 (RD1) antigen (eg, early secretory antigenic target 6 [ESAT6], culture filtrate protein 10 [CFP10], Rv2654c [antigen TB7.7]). The study endpoint assessed was active tuberculosis caused by *M tuberculosis* (we did not include studies that assessed non-tuberculous mycobacterium diseases), and the reference standard for the endpoint was any diagnosed incident active tuberculosis. Studies were included even if they did not stratify results into culture-confirmed tuberculosis and clinically diagnosed tuberculosis.

### Data extraction

Two reviewers (MXR and MP) independently assessed eligible articles for inclusion; disagreements were resolved by consensus. All articles included were assessed by a reviewer (MXR), who also extracted data, including study design, participants, country, period of recruitment, proportion of participant who had received BCG vaccination, IGRA methods (assay used, test version, cutoff-point used), TST methods (dose of purified protein derivative [PPD], cutoff-point used), and outcome data (eg, baseline TST and IGRA positivity rates, IGRA or TST concordance or discordance, and rates of progression to active tuberculosis). DL independently verified the extracted data on studies' general characteristics, test characteristics, and main results.

### Quality assessment

IGRA predictive value studies are not focused on diagnostic test accuracy. We therefore used a modified Newcastle-Ottawa quality assessment scale for non-randomised observational studies<sup>16</sup> rather than the quality assessment tool for diagnostic accuracy studies.<sup>17</sup> Studies were assessed for selection of study groups, comparability of study groups, and ascertainment of either the exposure or outcome of interest. Specific modifications were made to the selection and ascertainment of outcome items; high quality studies were those in which all cases of active tuberculosis were microbiologically confirmed, IGRA results were not incorporated into the reference standard (ie, no incorporation bias), and clinicians who assessed participants for active tuberculosis were masked to IGRA results. Incorporation of index tests and non-blinded assessments of possible tuberculosis cases could lead to relative risk estimates biased in favour of positive

IGRA results. IGRA-positive results could instigate tuberculosis investigations that might be more extensive than they would be for individuals with IGRA-negative results, resulting in further differential work-up bias (webappendix p 3). We did sensitivity analyses to explore the effect of key quality items on our main results.

### Statistical analysis

The main outcome of interest was person-years incidence rates of disease (incidence density), stratified by test results. We calculated incidence rate ratios (IRR) for rates of disease progression in IGRA-positive versus IGRA-negative individuals (we did the same for TST-positive vs TST-negative individuals). We also calculated risk ratios (cumulative incidence ratios) because several studies did not report rates or provide adequate information to allow computation of rates. We calculated DerSimonian and Laird random-effects pooled relative risks with 95% CIs;<sup>18</sup> 0.5 was added to correct for zero values in two by two tables.<sup>19</sup>

Because confounding is a concern in observational studies, we assessed whether adjusted IRR estimates were different from unadjusted estimates. We had planned to use the best-reported adjusted estimate of relative risk from each study if different from unadjusted estimates. However, most studies did not report adjusted estimates, and for those studies that did multivariable analyses, adjusted estimates were much the same as unadjusted estimates. Therefore, we used unadjusted estimates. Except when clearly indicated, all assays in the analyses were presented as IGRA irrespective of whether they were whole-blood ELISA or ELISPOT, in-house or commercial.

Heterogeneity was assessed with the  $I^2$  statistic and defined as low ( $I^2 < 25\%$ ), moderate ( $25\% < I^2 < 50\%$ ), or high ( $I^2 > 50\%$ ).<sup>20,21</sup> We did not report pooled estimate measures or interpreted them cautiously when the  $I^2$  value was greater than 25%. When heterogeneity was identified, reasons were explored by calculation of effect measures stratified by three prespecified subgroups: key study quality items, country-level stratifications, and study-level stratifications. Country-level stratifications were high-income versus low-income or middle-income,<sup>22</sup> estimated tuberculosis incidence per 100 000 individuals, and HIV prevalence. Study-level stratifications were study population or cohort followed, retrospective or prospective design, proportion of participants with BCG scar, age strata (adult or children), inclusion of individuals with HIV, provision of isoniazid preventive treatment to individuals in the study, whole-blood ELISA versus ELISPOT assay, in-house or commercial assay, assay incubation period, and TST status of participants.

Diagnostic accuracy estimates for progression to disease were chosen as surrogates for patient-relevant outcomes.<sup>23</sup> A false-positive test result could result in unnecessary treatment in an individual who would not have progressed to tuberculosis disease, whereas a false-negative result would mean progression to active tuberculosis disease that could have been prevented. Bivariate random-effects regression was done to obtain global summaries of sensitivity and specificity separately for studies that did IGRA (whole-blood ELISA or ELISPOT) and TST. We used Stata (version 10/MP) for all analyses.<sup>24</sup>

### Role of the funding source

The sponsors of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. MXR and MP had full access to all the data in the study and had final responsibility for the decision to submit for publication.

## Results

Figure 1 shows the study selection process, reported according to PRISMA guidelines.<sup>25</sup> 15 studies met our inclusion criteria (figure 1);<sup>26–40</sup> results of one study were described in two reports.<sup>29,41</sup> Of the 15 included studies, four were done in low-income countries, five in middle-income countries, and six in high-income countries; three had retrospective cohorts (table 1). The 15 included studies had a combined cohort size of 26 680 participants, with the study in South Africa<sup>38</sup> being the largest as of June 30, 2011 (11 988 person-years of follow-up).

All study cohorts were at high risk of tuberculosis (table 1). Two studies<sup>26,33</sup> followed up individuals with HIV (one of which, the study by Jonnalagadda and colleagues,<sup>33</sup> also studied infants who were followed up; main outcome results were stratified into mothers and infants when possible), one<sup>36</sup> followed up a group of patients with silicosis, one<sup>31</sup> followed up a cohort of asylum seekers, two<sup>34,40</sup> followed up cohorts of health-care workers, one study<sup>38</sup> followed up adolescents in a high-incidence area, and eight<sup>27–29,30,32,35,37,39</sup> followed up case contacts (table 1). All studies in high-income countries used data collected during routine care of patients. TST positivity was a prerequisite for IGRA testing in the studies by Kik and colleagues<sup>35</sup> and Harstad and colleagues.<sup>31</sup>

Median study duration was 3 years (IQR 2–5). In three studies,<sup>29,31,39</sup> more than 20% of the cohort was lost to follow-up. In one study,<sup>29</sup> 276 (72%) of 381 lost to follow-up were IGRA negative and had no information about development of tuberculosis. 0–76% of participants received isoniazid preventive treatment, which was not given to test-positive individuals in eight studies.<sup>26,28,30,32,33,35,38,40</sup>

Ten of 15 studies did a whole-blood ELISA, two of which<sup>28,30</sup> used in-house assays and eight<sup>26,29,31,34,35,38–40</sup> used the third-generation QuantiFERON-TB Gold in tube technology (table 2). Six studies assessed the ELISPOT assay—three used in-house assays<sup>27,32,37</sup> and three used T-SPOT.TB.<sup>33,35,36</sup> All studies used incubation periods of less than 24 h, except for Doherty and colleagues<sup>30</sup> (5 days) and del Corral and colleagues (7 days).<sup>28</sup> 11 of 15 studies included TST,<sup>27,29,31,32,34,35–38,40</sup> but only eight<sup>27,29,32,34,36–38,40</sup> reported results of the predictive value of TST (webappendix p 8).

Table 3 shows key quality characteristics of included studies. Individuals selected seemed to be representative of specific high-risk groups of interest within the population (eg, case contacts or health-care workers); both IGRA-positive and IGRA-negative individuals came from the same high-risk for progression groups. However, studies varied in quality, especially with respect to ascertainment of incident tuberculosis and the potential for incorporation and differential work-up biases. Only in eight studies<sup>29,34–40</sup> were at least half the active tuberculosis cases microbiologically confirmed (webappendix p 5). In seven studies,<sup>26,29,31,35,38–40</sup> positive IGRA results were incorporated in the case definitions of active tuberculosis. Thus, IGRA-positive individuals were more likely than were IGRA-negative individuals to be investigated for tuberculosis or diagnosed with tuberculosis.

Nine studies<sup>27,28,32–38,45</sup> reported incidence rates per person-time of follow-up (all of which except Kik and colleagues<sup>35</sup> reported rates stratified by IGRA status), the remaining six reported cumulative incidence risk (webappendix p 5). Tuberculosis rates in IGRA-positive individuals were 4–48 cases per 1000 person-years, whereas rates in IGRA-negative individuals were 2–24 cases per 1000 person-years during a median period of observation of 4 years (IQR 2–6; figure 2).<sup>33,34</sup> None of the studies reported rates of tuberculosis in IGRA-positive individuals, with both TST-negative and TST-positive results, who were given preventive treatment.

We compared unadjusted cumulative incidence risk ratios (RR) estimates because adjusted estimates were not available for most studies. The pooled RR for IGRA-positive results was 3.54 (95% CI 2.23–5.60), but heterogeneity was high (figure 3).

Stratification of studies by whether incorporation bias or differential work-up bias were or were not possible yielded a subgroup RR in the possibly biased studies higher than that in the subgroup of those in which such bias was not possible (figure 3). Studies with possible incorporation or differential work-up bias were therefore omitted from the rest of the review's main outcomes.

Although we had planned to explore heterogeneity by study characteristics, after exclusion of studies with possible incorporation or differential work-up biases, which included all those from high-income countries and with low national tuberculosis incidence, we were left with eight studies,<sup>27,28,30,32–34,36,37</sup> all from low-income or middle-income countries. For these studies, we detected no statistical difference across strata for country-level subgroups (World Bank income classification, tuberculosis case-detection rate, and HIV prevalence) or study-level subgroups (tuberculosis case contacts *vs* other cohorts, prospective *vs* retrospective study, BCG scar, children *vs* adults, any individuals with HIV in cohort, any isoniazid preventive treatment given, whole-blood ELISA *vs* ELISPOT assay, in-house *vs* commercial assay, and assay incubation period; data not shown).

In seven studies<sup>27,28,32–34,36,37</sup> reporting tuberculosis incidence rates stratified by IGRA status, individuals with positive IGRA results at baseline had a higher incidence of active tuberculosis than did those with negative results (n=9530; figure 4). For three studies,<sup>34,36,37</sup> the pooled IRR in microbiologically confirmed tuberculosis cases was higher (3.5, 95% CI 1.3–10.0) than for all cases of tuberculosis diagnosed (2.2, 1.1–4.5), but the difference was not significant. In five studies that reported results of multivariable analyses,<sup>27,32,33,36,37</sup> the adjusted IRR were much the same as the crude IRR (webappendix p 6).

Five studies<sup>27,32,36–38</sup> without biases in incorporation or differential work-up stratified tuberculosis incidence by IGRA and TST results status at baseline. All studies used 0.1 mL of 2TU PPD RT23, but different cutoffs for positivity were selected in these studies, according to national guidelines. Relative risks were stratified by TST cutoff, when possible. The IRR of incident tuberculosis in test-positive individuals compared with test-negative individuals was slightly higher for IGRA than for the TST, but was not significant because of overlapping CIs (figure 5).

Assessment of discordant results can help to find out whether IGRAs are better than TST in the prediction of tuberculosis disease.<sup>50</sup> Four studies<sup>27,32,34,37</sup> further explored rates of tuberculosis in paired concordant and discordant TST and IGRA results (table 4). Rates of tuberculosis were slightly higher in discordant pairs when IGRA was positive than in pairs when TST was the positive test; however, the association with incident active tuberculosis in the individual studies was weak (table 4).

With a 7 day whole-blood ELISA, one study<sup>28</sup> assessed whether an exposure-gradient relation existed between baseline quantitative IGRA strata and subsequent rates of tuberculosis (webappendix p 7). Rates across the three highest strata were much the same and the CIs overlapped (incidence of 7 cases per 1000 person-years [95% CI 2.6–15.2] for 22–99 pg/mL, 6.7 [3.4–11.7] for 100–999 pg/mL, and 7.7 [5.1–11.3] for 1000 pg/mL); the incidence per 1000 person-years in the lowest stratum was 4.1 (1.7–8.5).

Three studies<sup>27,34,37</sup> assessed whether baseline median IGRA responses in individuals who subsequently developed tuberculosis were higher than responses in those who did not develop tuberculosis but had positive tests. In one study,<sup>37</sup> median responses were 250 spot-

forming cells (SFC) per million peripheral blood mononuclear cells (PBMC) in individuals who subsequently developed tuberculosis and 50 SFC per million PBMC in those who did not develop tuberculosis ( $p=0.02$ ). By contrast, two studies<sup>27,34</sup> noted no differences in median responses in individuals who developed tuberculosis compared with those who did not.

Findings from seven<sup>27,29,31,35,37,38,40</sup> of 11 studies that assessed IGRA and TST suggest that the proportion of IGRA-positive participants who scored positive at the initial baseline visit might be lower than in those who were TST-positive (at investigator-selected cutoffs; webappendix p 9).

For studies that used the ELISPOT assay, the sensitivity for developing active tuberculosis was 72% (95% CI 58–82) and specificity was 50% (41–58).<sup>27,32,36,37</sup> Estimates for TST in those same studies were 72% (58–83) and 41% (30–54).<sup>27,32,36,37</sup> However, estimates for both tests were imprecise. On the basis of these estimates, the false-positive rate for the ELISPOT and TST were much the same—50% (95% CI 42–59) for ELISPOT and 59% (46–70) for TST (false-positive rate=100–specificity). False-negative rates for the two tests were also similar. An estimate for studies that used the whole-blood ELISA was available from only two studies that were possibly not subject to incorporation or work-up bias.<sup>28,34</sup> In one study,<sup>28</sup> which used a 7 day whole-blood ELISA assay, sensitivity was 79% (95% CI 61–91) and specificity was 34% (32–36), but in the other study,<sup>34</sup> which used an assay with a shorter incubation, reported sensitivity was 43% (18–72) and specificity was 59% (55–62).

## Discussion

The strength of the association between positive IGRA results and development of active tuberculosis in the studies identified was weak to moderate, with relative risks of about 2–3. The incidence of tuberculosis, even in IGRA-positive individuals, was low, suggesting that most IGRA-positive individuals did not progress to tuberculosis disease during follow-up. This finding is similar to that for TST in this meta-analysis and in historical studies.<sup>14</sup> Thus, the most important finding in this review is that no available tests for latent *M tuberculosis* infection have high prognostic value.

However, in some populations the proportion of IGRA-positive individuals might generally be lower than the proportion of TST-positive individuals (as shown in seven of 11 studies assessed). This occurrence could be because of either higher IGRA specificity for *M tuberculosis* infection or lower sensitivity than with TST. Higher specificity would suggest that even though use of both TST and IGRA is imperfect for informing a decision about who should receive preventive treatment, the number of individuals identified for preventive treatment could be less with IGRA than with TST because fewer people are IGRA-positive than are TST-positive. This characteristic of IGRAs might be useful in settings where TST specificity is compromised by cross-reactivity with environmental mycobacteria, BCG vaccination after infancy, or multiple BCG vaccinations.<sup>51</sup> In such settings, TST-positive and IGRA-negative discordance is likely to be common. As noted by Zwerling and colleagues,<sup>5</sup> high reversion rates of IGRA are common in settings with both low and high tuberculosis incidence, which might contribute to the low proportion of IGRA positivity (compared with TST) in many studies.

We had planned to explore heterogeneity by study characteristics. However, stratification by studies with possible incorporation or differential work-up biases, left us with eight studies with a statistically homogeneous pooled estimate. Further assessment of these remaining studies did not show any statistically significant differences across subgroup strata. Studies that use routine care data, from mostly high-income countries where commercial IGRA are

already included in tuberculosis guidelines, are likely to report results that are in favour of IGRA. Although assessment of the tests in routine practice should be encouraged, the potential issues of incorporation bias or differential work-up should be anticipated early and be mitigated either in study design or at analysis. In the excluded studies, incorporation bias would have been mitigated if a definite diagnosis of incident tuberculosis was made for all cases on the basis of microbiological methods instead of diagnoses that relied on subjective clinical interpretation (unblinded to IGRA results). Complete and identical diagnostic work-up of all participants who entered follow-up, index-test positive and negative, would mitigate differential work-up bias.<sup>52</sup>

Overall, most of the studies did not fully answer the question of whether IGRA, as a surrogate marker for future risk, can predict subsequent active tuberculosis. They merely showed a slight positive association between initial positive IGRA results and subsequent tuberculosis. Most studies were likely to yield exaggerated results in favour of a positive association largely because of ascertainment bias of active tuberculosis and by not fully accounting for other risk factors for tuberculosis. Furthermore, no studies assessed the ability of IGRA to adequately discriminate and therefore predict individuals at risk of developing disease from those not at risk. Discriminatory ability, rather than association, is of primary interest in predictive studies of disease. One measure of association does not imply that the new test can accurately discriminate individuals at risk of disease and those who will not develop disease.<sup>53-56</sup> This fact is widely acknowledged in biomarker studies related to cancer and cardiovascular epidemiology.<sup>56-58</sup> Perfect discrimination will occur only when the distribution curves of the marker in individuals with disease versus individuals without disease do not overlap.<sup>56</sup> One marker or test has to be strongly associated with disease to be useful for disease prediction.<sup>53,56-58</sup> A possible reason could be that even strongly associated risk factors or markers (eg, TST or IGRA) can perform poorly as tests of disease prediction if little variation exists in exposure (eg, high tuberculosis infection) within the population being studied.<sup>53,55</sup>

A biomarker such as interferon- $\gamma$  could indicate *M tuberculosis* sensitisation (rather than disease) but might not, on its own, be adequate to predict active tuberculosis disease, especially in countries with a high burden and rates of reinfection. Interferon- $\gamma$  might be necessary but not sufficient on its own for prediction of disease. This finding is probably true even when interferon- $\gamma$  is regarded as a correlate of protection in vaccine studies.<sup>59</sup> Interferon- $\gamma$  alone might not be sufficient as a biomarker because antigen-specific interferon- $\gamma$  response is elicited in almost all stages of the tuberculosis spectrum.<sup>60,61</sup> Therefore, the identification of more predictive biomarkers is important, as is measurement of an array of biomarkers or incorporation of biomarkers with other known risk factors into a composite scoring system. For example, age, recent infection in young children, recent contact with smear-positive active case, or HIV infection, in combination with IGRA results, might have much higher predictive ability than just interferon- $\gamma$  response alone. Future studies need to assess multivariable risk prediction,<sup>54,62</sup> as attempted in algorithms such as the Online TST/IGRA interpreter,<sup>63</sup> and the discriminatory incremental yield of new tests to existing clinical algorithms, before predictive ability is declared.

As discussed elsewhere,<sup>64</sup> another possible explanation for the poor predictive ability of existing tests for latent *M tuberculosis* infection is that a single or cross-sectional TST or IGRA result cannot resolve the underlying phenotypes because they do not capture information about when infection occurred and how the infection was fully, partly, or not eliminated by the host. All but two studies<sup>27,28</sup> included in this review reported results of only one IGRA or TST test at baseline. Serial IGRA testing might show interesting underlying phenotypes that have different histories and trajectories.<sup>65</sup> Without serial testing,



the underlying phenotypes are not distinguishable, undermining the predictive value of a single test result.

Our systematic review had limitations. Data included in our review did not allow for formal assessment of publication bias with methods such as funnel plots or regression asymmetry tests. We, therefore, assume some degree of publication bias is likely because the number of studies of IGRA is rapidly increasing and new studies will soon become available—we are aware of at least three ongoing studies (in South Africa [NCT00463086],<sup>66</sup> Zambia [ZAMSTAR; ISRCTN36729271 ],<sup>67</sup> and the UK<sup>44</sup>) that could not be included. Anecdotal reports exist of unpublished negative studies of IGRA—many (about 50%) IGRA studies have some industry involvement or support,<sup>5</sup> meaning that studies with negative findings might not be published (or publication might be delayed). Data used to obtain our main summary measures were restricted to low-income to middle-income countries, which largely limits interpretation to those settings or similar individuals in high-income countries. IGRAs might have superior predictive ability in high-income settings with low tuberculosis incidence, but we were unable to identify this effect because of likely incorporation or work-up biases. A meta-analysis of individual patient-data would have allowed for a multivariable assessment of discriminatory value and better adjustment of confounding, and thus provide a better interpretation of available data, but this was not possible. Although inclusion of non-commercial IGRAs might be a limitation of our analysis, the inclusion of all studies of RD1-based assays enabled the most comprehensive synthesis of IGRA predictive ability to date.

Further research is needed to identify more predictive biomarkers to improve existing tests for latent *M tuberculosis* infection. Indeed, the revised Global Plan to Stop TB (2011–15) has set 2015 as the goal for such predictive tests.<sup>68</sup> Until then, the following strategies might be useful to improve the predictive value of existing tests: testing only those individuals at high risk of tuberculosis; serial testing to identify new infections (ie, conversions); incorporation of biomarkers with known risk factors (age, recent exposure, HIV infection, etc) into risk prediction models; and use of a higher cutoff for prediction of disease (as compared with diagnosis).

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The findings of this review were presented at a WHO Expert Group Meeting on IGRA, held in July 2010, organised by the Stop TB Department of the WHO. We thank WHO and Expert Group members for their constructive feedback on the review. We also thank the authors of the studies included in this review for providing additional information upon request. This work was supported in part by the Special Programme for Research and Training in Tropical Diseases, WHO, as part of the WHO Expert Group Meeting process to develop policy recommendations on IGRAs for low-income and middle-income countries, and in part by the Canadian Institutes of Health Research (grant number MOP-81362). MXR and RJW were supported by the Wellcome Trust (grants 084323, 084670, and 088316) and the European and Developing Countries Clinical Trials Partnership (EDCTP). RJW was also supported by the European Union (SANTE/2005/105-061-102). KAW was supported by the UK Medical Research Council. MP was a recipient of a Canadian Institutes of Health Research New Investigator Award and recipient of support from EDCTP (TB-NEAT grant). DM is a recipient of a Fonds de la recherche en santé du Québec (FRSQ) career award.

## References

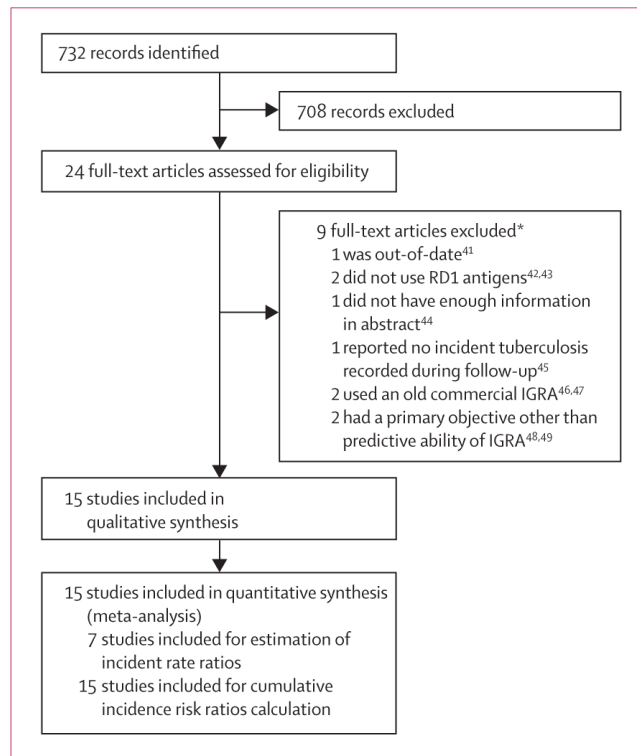
1. Dye C, Scheele S, Dolin P, Pathania V, Raviglione MC. Consensus statement: global burden of tuberculosis—estimated incidence, prevalence, and mortality by country (WHO global surveillance and monitoring project). *JAMA*. 1999; 282:677–86. [PubMed: 10517722]

2. Woldehanna S, Volmink J. Treatment of latent tuberculosis infection in HIV infected persons. *Cochrane Database Syst Rev.* 2004; 1:CD000171. [PubMed: 14973947]
3. Diel R, Loddenkemper R, Nienhaus A. Evidence-based comparison of commercial interferon-gamma release assays for detecting active tuberculosis: a metaanalysis. *Chest.* 2009; 137:952–68. [PubMed: 20022968]
4. Pai M, Riley LW, Colford JM Jr. Interferon-gamma assays in the immunodiagnosis of tuberculosis: a systematic review. *Lancet Infect Dis.* 2004; 4:761–76. [PubMed: 15567126]
5. 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; 149:177–84. [PubMed: 18593687]
6. Zwerling A, van den Hof S, Scholten J, Cobelens F, Menzies D, Pai M. Interferon-gamma release assays for tuberculosis screening of healthcare workers: a systematic review. *Thorax.* 2011 published online Jan 12. DOI:10.1136/thx.2010.143180.
7. Cattamanchi A, Smith R, Steingart KR, et al. Interferon-gamma release assays for the diagnosis of latent tuberculosis infection in HIV-infected individuals: a systematic review and meta-analysis. *J Acquir Immune Defic Syndr.* 2011; 56:230–38. [PubMed: 21239993]
8. Metcalfe JZ, Everett CK, Steingart KR, et al. Interferon-gamma release assays for active pulmonary TB diagnosis in low-and middle-income countries: systematic review and meta-analysis. *J Infect Dis.* (in press).
9. Denkinger CM, Dheda K, Pai M. Guidelines on interferon-gamma release assays for tuberculosis infection: concordance, discordance or confusion? *Clin Microbiol Infect.* 2011; 17:806–14. [PubMed: 21682801]
10. Comstock GW, Livesay VT, Woolpert SF. The prognosis of a positive tuberculin reaction in childhood and adolescence. *Am J Epidemiol.* 1974; 99:131–38. [PubMed: 4810628]
11. Jeyakumar D. Tuberculin reactivity and subsequent development of tuberculosis in a cohort of student nurses. *Med J Malaysia.* 1999; 54:492–95. [PubMed: 11072468]
12. Leung CC, Yew WW, Chang KC, et al. Risk of active tuberculosis among schoolchildren in Hong Kong. *Arch Pediatr Adolesc Med.* 2006; 160:247–51. [PubMed: 16520443]
13. Watkins RE, Brennan R, Plant AJ. Tuberculin reactivity and the risk of tuberculosis: a review. *Int J Tuberc Lung Dis.* 2000; 4:895–903. [PubMed: 11055755]
14. Ferebee SH. Controlled chemoprophylaxis trials in tuberculosis: a general review. *Bibl Tuberc.* 1970; 26:28–106. [PubMed: 4903501]
15. Samandari TAT, Nyirenda S, Tedla Z, et al. 6-month versus 36-month isoniazid preventive treatment for tuberculosis in adults with HIV infection in Botswana: a randomised, double-blind, placebo-controlled trial. *Lancet.* 2011; 377:1588–98. [PubMed: 21492926]
16. Wells, GAOCD.; Pertersen, J.; Welch, V.; Losos, M.; Tugwell, P. [accessed Feb 20, 2010] The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies and meta-analyses. [http://www.ohri.ca/programs/clinical\\_epidemiology/oxford.htm](http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm)
17. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. 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. [PubMed: 14606960]
18. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986; 7:177–88. [PubMed: 3802833]
19. Sheehe PR. Combination of log relative risk in retrospective studies of disease. *Am J Public Health Nations Health.* 1966; 56:1745–50. [PubMed: 5951494]
20. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med.* 2002; 21:1539–58. [PubMed: 12111919]
21. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003; 327:557–60. [PubMed: 12958120]
22. [accessed Feb 15, 2010] World Bank. Country Classification. <http://data.worldbank.org/about/country-classifications>
23. Schunemann HJ, Oxman AD, Brozek J, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ.* 2008; 17:336, 1106–10.

24. Sterne, JA. Meta-analysis in Stata: an updated collection from the Stata journal. Stata Press; Atlanta: 2009.
25. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 2009; 21:6, e1000097.
26. Aichelburg MC, Rieger A, Breitenecker F, et al. Detection and prediction of active tuberculosis disease by a whole-blood interferon-gamma release assay in HIV-1-infected individuals. *Clin Infect Dis.* 2009; 48:954–62. [PubMed: 19245343]
27. Bakir M, Millington KA, Soysal A, et al. Prognostic value of a T-cell-based, interferon-gamma biomarker in children with tuberculosis contact. *Ann Intern Med.* 2008; 149:777–87. [PubMed: 18936496]
28. del Corral H, Paris SC, Marin ND, et al. IFN-gamma response to *Mycobacterium tuberculosis*, risk of infection and disease in household contacts of tuberculosis patients in Colombia. *PLoS One.* 2009; 4:e8257. [PubMed: 20011589]
29. Diel R, Loddenkemper R, Niemann S, Meywald-Walter K, Nienhaus A. Negative and positive predictive value of a whole-blood interferon- $\gamma$  release assay for developing active tuberculosis: an update. *Am J Respir Crit Care Med.* 2011; 183:88–95. [PubMed: 20802162]
30. Doherty TM, Demissie A, Olobo J, et al. Immune responses to the *Mycobacterium tuberculosis*-specific antigen ESAT-6 signal subclinical infection among contacts of tuberculosis patients. *J Clin Microbiol.* 2002; 40:704–06. [PubMed: 11826002]
31. Harstad I, Winje BA, Heldal E, Oftung F, Jacobsen GW. Predictive values of QuantiFERON-TB Gold testing in screening for tuberculosis disease in asylum seekers. *Int J Tuberc Lung Dis.* 2010; 14:1209–11. [PubMed: 20819271]
32. Hill PC, Jackson-Sillah DJ, Fox A, et al. Incidence of tuberculosis and the predictive value of ELISPOT and Mantoux tests in Gambian case contacts. *PLoS One.* 2008; 3:e1379. [PubMed: 18167540]
33. Jonnalagadda S, Lohman Payne B, Brown E, et al. Latent tuberculosis detection by interferon gamma release assay during pregnancy predicts active tuberculosis and mortality in human immunodeficiency virus type 1-infected women and their children. *J Infect Dis.* 2010; 202:1826–35. [PubMed: 21067370]
34. Joshi R, Narang U, Zwerling A, et al. Predictive value of latent TB tests in Indian health-care workers: a cohort study. *Eur Respir J.* (in press).
35. Kik SV, Franken WP, Mensen M, et al. Predictive value for progression to tuberculosis by IGRAs and TST in immigrant contacts. *Eur Respir J.* 2010; 35:1346–53. [PubMed: 19840963]
36. Leung CC, Yam WC, Yew WW, et al. T-Spot TB outperforms tuberculin skin test in predicting tuberculosis disease. *AJRCCM.* 2010; 182:834–40.
37. Lienhardt C, Fielding K, Hane AA, et al. Evaluation of the prognostic value of IFN-gamma release assay and tuberculin skin test in household contacts of infectious tuberculosis cases in Senegal. *PLoS One.* 2010; 5:e10508. [PubMed: 20463900]
38. Mahomed H, Hawkrigde T, Verver S, et al. The tuberculin skin test versus QuantiFERON TB Gold in predicting tuberculosis disease in an adolescent cohort study. *PLoS One.* 2011; 6:e17984. [PubMed: 21479236]
39. Yoshiyama T, Harada N, Higuchi K, Sekiya Y, Uchimura K. Use of the QuantiFERON TB Gold Test for screening tuberculosis contacts and predicting active disease. *Int J Tuberc Lung Dis.* 2010; 14:819–27. [PubMed: 20550763]
40. Costa JTSR, Ringshausen F, Nienhaus A. Screening for tuberculosis and prediction of disease in Portuguese healthcare workers. *J Occup Med Toxicol.* 2011; 6:19. [PubMed: 21658231]
41. Diel R, Loddenkemper R, Meywald-Walter K, Niemann S, Nienhaus A. Predictive value of a whole blood IFN-gamma assay for the development of active tuberculosis disease after recent infection with *Mycobacterium tuberculosis*. *Am J Respir Crit Care Med.* 2008; 177:1164–70. [PubMed: 18276940]
42. Elliott AM, Hodsdon WS, Kyosiimire J, et al. Cytokine responses and progression to active tuberculosis in HIV-1-infected Ugandans: a prospective study. *Trans R Soc Trop Med Hyg.* 2004; 98:660–70. [PubMed: 15363646]

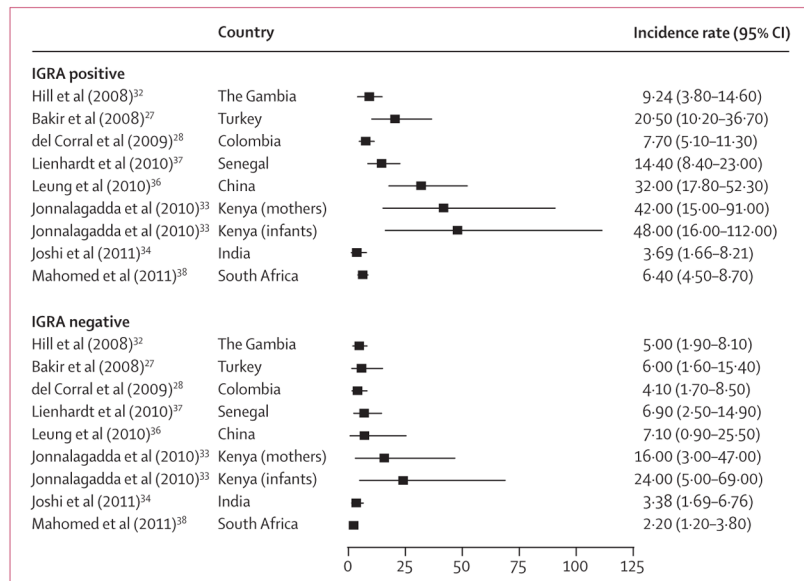
43. Ordway DJ, Costa L, Martins M, et al. Increased interleukin-4 production by CD8 and gammadelta T cells in health-care workers is associated with the subsequent development of active tuberculosis. *J Infect Dis.* 2004; 190:756–66. [PubMed: 15272404]
44. Haldar P. Contact screening with single-step TIGRA testing and risk of active TB infection: the Leicester cohort. *Thorax.* 2009; 64(suppl 4):A5–A74.
45. Santin M, Casas S, Saumoy M, et al. Detection of latent tuberculosis by the tuberculin skin test and a whole-blood interferon-gamma release assay, and the development of active tuberculosis in HIV-seropositive persons. *Diagn Microbiol Infect Dis.* 2011; 69:59–65. [PubMed: 21146715]
46. Higuchi K, Harada N, Mori T, Sekiya Y. Use of QuantiFERON-TB Gold to investigate tuberculosis contacts in a high school. *Respirology.* 2007; 12:88–92. [PubMed: 17207031]
47. Higuchi K, Kondo S, Wada M, et al. Contact investigation in a primary school using a whole blood interferon-gamma assay. *J Infect.* 2009; 58:352–57. [PubMed: 19342102]
48. Clark SA, Martin SL, Pozniak A, et al. Tuberculosis antigen-specific immune responses can be detected using enzyme-linked immunospot technology in human immunodeficiency virus (HIV)-1 patients with advanced disease. *Clin Exp Immunol.* 2007; 150:238–44. [PubMed: 17672869]
49. Elliott JH, Vohith K, Saramony S, et al. Immunopathogenesis and diagnosis of tuberculosis and tuberculosis-associated immune reconstitution inflammatory syndrome during early antiretroviral therapy. *J Infect Dis.* 2009; 200:1736–45. [PubMed: 19874177]
50. Pai M, O'Brien R. Serial testing for tuberculosis: can we make sense of T cell assay conversions and reversions? *PLoS Med.* 2007; 4:e208. [PubMed: 17564491]
51. Zwerling A, Behr MA, Verma A, Brewer TF, Menzies D, Pai M. The BCG world atlas: a database of global BCG vaccination policies and practices. *PLoS Med.* 2011; 8:e1001012. [PubMed: 21445325]
52. Oostenbrink R, Moons KG, Bleeker SE, Moll HA, Grobbee DE. Diagnostic research on routine care data: prospects and problems. *J Clin Epidemiol.* 2003; 56:501–06. [PubMed: 12873643]
53. Greenland P. When is a new prediction marker useful? a consideration of lipoprotein-associated phospholipase A<sub>x</sub> and C-reactive protein for stroke risk. *Arch Intern Med.* 2005; 165:2454–56. [PubMed: 16314539]
54. Kattan MW. Judging new markers by their ability to improve predictive accuracy. *J Natl Cancer Inst.* 2003; 95:634–35. [PubMed: 12734304]
55. Wald NJ, Hackshaw AK, Frost CD. When can a risk factor be used as a worthwhile screening test? *BMJ.* 1999; 319:1562–65. [PubMed: 10591726]
56. Ware JH. The limitations of risk factors as prognostic tools. *N Engl J Med.* 2006; 355:2615–17. [PubMed: 17182986]
57. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. *Clin Chem.* 2008; 54:17–23. [PubMed: 18024533]
58. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol.* 2004; 159:882–90. [PubMed: 15105181]
59. Dockrell HM. Gamma-interferon key, but not sufficient for protection against TB? *Microbiology Today.* 2007:172–75.
60. Barry CE 3rd, Boshoff HI, Dartois V, et al. The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat Rev Microbiol.* 2009; 7:845–55. [PubMed: 19855401]
61. Mack U, Migliori GB, Sester M, et al. LTBI: latent tuberculosis infection or lasting immune responses to *M tuberculosis*? A TBNET consensus statement. *Eur Respir J.* 2009; 33:956–73. [PubMed: 19407047]
62. Moons KG, Biesheuvel CJ, Grobbee DE. Test research versus diagnostic research. *Clin Chem.* 2004; 50:473–36. [PubMed: 14981027]
63. Menzies D, Gardiner G, Farhat M, Greenaway C, Pai M. Thinking in three dimensions: a web-based algorithm to aid the interpretation of tuberculin skin test results. *Int J Tuberc Lung Dis.* 2008; 12:498–505. [PubMed: 18419884]
64. Pai M. Spectrum of latent tuberculosis—existing tests cannot resolve the underlying phenotypes. *Nat Rev Microbiol.* 2010; 8:242. [PubMed: 20084059]

65. Andersen P, Doherty TM, Pai M, Weldingh K. The prognosis of latent tuberculosis: can disease be predicted? *Trends Mol Med.* 2007; 13:175–82. [PubMed: 17418641]
66. Rangaka, MX.; Boulle, A.; van Cutsem, G., et al. [accessed Aug 10, 2011] Isoniazid preventive therapy plus combined antiretroviral therapy to prevent tuberculosis in HIV-infected persons (ART-IPT Study): a pragmatic randomized trial. <http://www.thelancet.com/protocol-reviews/09PRT-2885>
67. STAG-TB. Report of the tenth meeting: diagnostic policies. World Health Organization; Geneva: 2010.
68. Stop TB partnership. The global plan to stop TB (2011–2015). World Health Organization; Geneva: 2011.



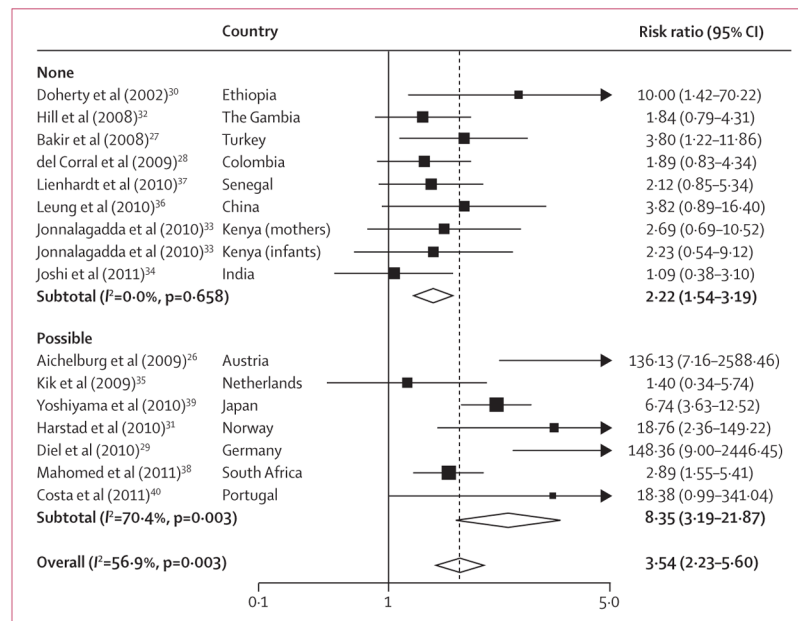
**Figure 1. Study selection**

\*See webappendix p 4 for further exclusion details.



**Figure 2. Unadjusted incidence rates for all tuberculosis diagnoses stratified by interferon- $\gamma$  release assay (IGRA) status**

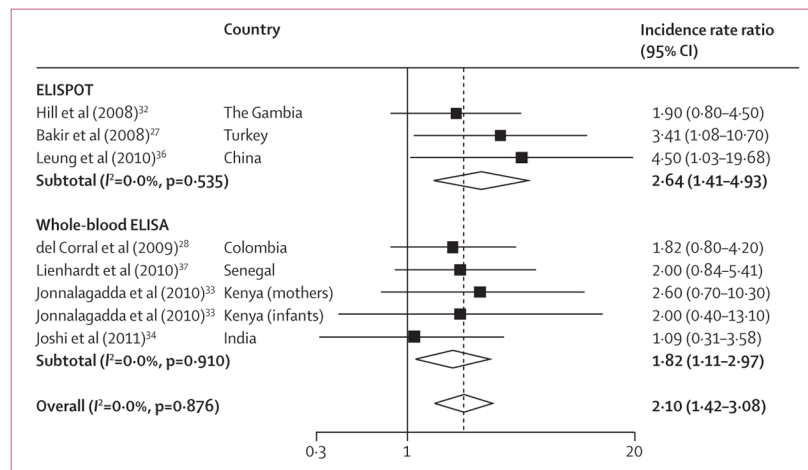
Incidence rate estimates are per 1000 person-years from individual studies that provided person-time data stratified by IGRA status at baseline.



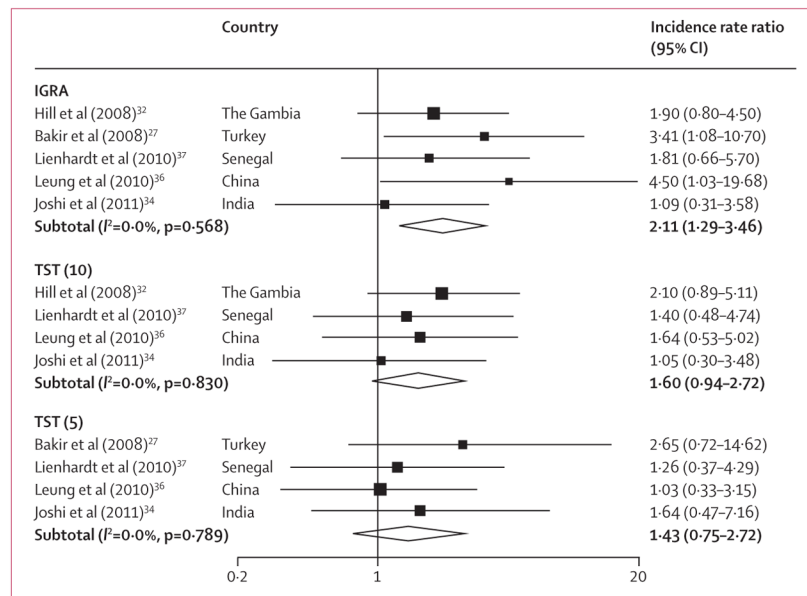
**Figure 3. Unadjusted cumulative incidence risk ratios for positive versus negative interferon- $\gamma$  release assay (IGRA) results, by possibility of incorporation bias**

One study<sup>35</sup> did not report whether tuberculosis diagnoses methods included IGRA and was therefore included along with the six studies<sup>26,29,31,38-40</sup> in which IGRA formed part of tuberculosis diagnoses methods and, therefore, incorporation bias could not be ruled out. Data from Kenya has been stratified into HIV-exposed infants and their mothers who had HIV.<sup>33</sup> Pooled risk ratio (RR) estimate with Netherlands T-Spot.TB results (rather than QuantiFERON Gold in tube): RR=3.61 (95% CI 2.29-5.69),  $I^2=67.4\%$ ,  $p=0.005$ .





**Figure 4.** Unadjusted incidence rate ratios for positive versus negative interferon- $\gamma$  release assay results, by type of assay



**Figure 5. Unadjusted incidence rate ratios for positive versus negative test result, by test type** Tuberculin skin test (TST) is stratified by cutoff for studies that provided values in the original paper or on request. TST (10)=TST more than 10 mm. TST (5)=TST more than 5 mm.

**Table 1**  
**Study characteristics of subpopulations included**

	Country (income status)	Age group (years)	Individuals with HIV in cohort (%)	Population	Individuals assessed (n)	Individuals followed up and included in analysis (n)	IPT given (%) <sup>*</sup>	Tuberculosis diagnoses included
Doherty et al (2002) <sup>30</sup>	Ethiopia (LIC)	Adults (15–65)	No; exclusion criterion	Tuberculosis case-contacts	38	24	No	Smear and culture
Hill et al (2008) <sup>32</sup>	The Gambia (LIC)	Adults and children (0.5–100)	Yes (2%)	Tuberculosis case-contacts	2381	2348	No	TST, smear, and culture
Bakir et al (2008) <sup>27</sup>	Turkey (MIC)	Children (0–16)	Not stated	Tuberculosis case-contacts	1024	908	Yes (76% of 908)	Smear and culture
Aichelburg et al (2009) <sup>26</sup>	Austria (HIC)	Adults (IQR 31–46)	Yes (100%)	Outpatients with HIV	834	822	No	IGRA and culture
Kik et al (2009) <sup>35</sup>	Netherlands (HIC)	Adults (16–45+)	No; exclusion criterion	Tuberculosis case-contacts	433	339	No; exclusion	Smear and culture
Del Corral et al (2009) <sup>28</sup>	Colombia (MIC)	Adults and children (IQR 10–42)	Unknown <sup>†</sup>	Tuberculosis case-contacts	2060	2060	No	Smear and culture
Lienhardt et al (2010) <sup>37</sup>	Senegal (LIC)	Adults and children (18–71)	Unknown <sup>†</sup>	Tuberculosis case-contacts	2762	2679	Yes (% NS)	Smear and culture
Yoshiyama et al (2010) <sup>39</sup>	Japan (HIC)	Adults and children (0–60+)	Unknown <sup>†</sup>	Tuberculosis case-contacts (retrospective)	NS	5676	Yes (20% of 3102)	IGRA <sup>‡</sup>
Leung et al (2010) <sup>36</sup>	China (MIC)	Adults (mean 60)	Unknown <sup>†</sup>	Outpatients with silicosis	331	308	Yes (33% of 203)	Smear and culture
Harstad et al (2010) <sup>31</sup>	Norway (HIC)	Adults (18–50+)	Unknown <sup>†</sup>	Asylum seekers	NS	823	Yes (3%)	IGRA <sup>‡</sup>
Diel et al (2010) <sup>29</sup>	Germany (HIC)	Adults and children (1–62)	No; exclusion criterion	Tuberculosis case-contacts	1417	1335	Yes (% NS)	TST, IGRA, and culture
Jonnalagadda et al (2010) <sup>33</sup>	Kenya (LIC)	Adults (24–26)	Yes (100%)	HIV cohort with no prior tuberculosis (retrospective)	333	258	No	Self-reported
Jonnalagadda et al (2010) <sup>33,§</sup>	Kenya (LIC)	Infants (<1)	Unknown	HIV-exposed infants (retrospective)	327	250	No	Report by the mother
Joshi et al (2011) <sup>34</sup>	India (MIC)	Adults (18–40)	Unknown	Health-care workers with no prior tuberculosis (retrospective)	726	719	Yes (17% of 360)	Self-reports (confirmed)
Mahomed et al (2011) <sup>38</sup>	South Africa (MIC)	Adolescents (12–18)	Unknown	Individuals with no prior tuberculosis	6363	5244	No	TST, IGRA, smear, and culture
Costa et al (2011) <sup>40</sup>	Portugal (HIC)	Adults (<25–50+)	No	Health-care workers	2889	2876	Yes (2% of 2876)	TST, IGRA, and culture

HIC=high-income country. IGRA=interferon- $\gamma$  release assay. IPT=isoniazid preventive treatment. LIC=low-income country. MIC=middle-income country. NS=not stated. TST=tuberculin skin test. The proportion of participants who completed follow-up was more than 80% for all studies except for those done in Norway<sup>31</sup> (cannot estimate), Japan<sup>39</sup> (55%) and Germany<sup>29</sup> (70%)—follow-up rates were not reported in the studies done in Netherlands<sup>35</sup> and Norway.<sup>31</sup>

\* Bakir et al<sup>27</sup> IPT given on the basis of age and positive skin tests; Aichelburg et al<sup>26</sup> all eligible participants refused; Yoshiyama et al<sup>39</sup> criteria were non-random and non-standard; Leung et al<sup>36</sup> given on the basis of TST positivity; Harstad et al<sup>31</sup> and Diel et al<sup>29</sup> given if IGRA positive; Joshi et al<sup>34</sup> given if TST or IGRA positive; and Costa et al<sup>40</sup> given to those who had conversion in past 2 years from negative to positive on TST, IGRA, or both, as confirmed by study investigators.

<sup>†</sup> HIV prevalence unknown but likely to be low, except for the Norwegian cohort where prevalence is probably high.

<sup>‡</sup> Confirmed in personal communication with investigators.

<sup>§</sup> Additional information requested from the authors.

**Table 2**  
**Test characteristics**

	Type of IGRA (in-house or commercial); <i>M tuberculosis</i> antigens or peptides*	IGRA cutoff †	TST cutoff(s)	Blinding to IGRA results
Doherty et al (2002) <sup>30</sup>	WBA, ELISA (in house); PPD, ESAT 6	100 pg/mL	..	Yes
Hill et al (2008) <sup>32</sup>	ELISPOT (in house); ESAT6, CFP-10	~32 SFC	10 mm	‡NS
Bakir et al (2008) <sup>27</sup>	ELISPOT (in house); ESAT6, CFP-10	~20 SFC	5 mm	Yes
Aichelburg et al (2009) <sup>26</sup>	WBA, ELISA (QFT-Gold in tube); ESAT6; CFP-10, Rv2654c (p38–55)	0.35 IU/mL	..	No
Kik et al (2009) <sup>35</sup>	ELISPOT (T-SPOT.TB); ESAT6, CFP-10	8 SFU after negative well subtraction (~20 SFC)	(inclusion criterion)	NS
Kik et al (2009) <sup>35</sup>	WBA, ELISA (QFT-Gold in tube); ESAT6, CFP-10, Rv2654c (p38–55)	0.35 IU/ml (~14 pg/mL)		
Del Corral et al (2009) <sup>28</sup>	WBA, ELISA (in house); CFP-10, CFP, Ag85A, Rv2031c	22 pg/mL	10 mm and 5 mm	‡NS
Lienhardt et al (2010) <sup>37</sup>	ELISPOT (in house); ESAT6, CFP-10	20 SFC after negative well subtraction and 32 SFC	10 mm	Yes
Yoshiyama et al (2010) <sup>39</sup>	WBA, ELISA (QFT-Gold in tube); ESAT6, CFP-10, Rv2654c (p38–55)	0.35 IU/mL (~14 pg/mL)	..	No
Leung et al (2010) <sup>36</sup>	ELISPOT (T SPOT-TB); ESAT6, CFP-10	6 SFU (~20 SFC)	5 mm, 10 mm, and 15 mm	Yes
Harstad et al (2010) <sup>31</sup>	WBA, ELISA (QFT-Gold in tube); ESAT6, CFP-10, Rv2654c (p38–55)	0.35 IU/mL (~14 pg/mL)	(inclusion criterion)	No
Diel et al (2010) <sup>29</sup>	WBA, ELISA (QFT-Gold in tube); ESAT6, CFP-10, Rv2654c (p38–55)	0.35 IU/mL (~14 pg/mL)	5 mm, 10 mm	No
Jonnalagadda et al (2010) <sup>33</sup>	ELISPOT (T.SPOT.TB); ESAT6, CFP-10	6 SFU (~20 SFC)	..	Yes
Joshi et al (2011) <sup>34</sup>	WBA, ELISA (QFT-Gold in tube); ESAT6, CFP-10, Rv2654c (p38–55)	0.35 IU/mL (~14 pg/mL)	10 mm	§Yes
Mahomed et al (2011) <sup>38</sup>	WBA, ELISA (QFT-Gold in tube); ESAT6, CFP-10, Rv2654c (p38–55)	0.35 IU/mL (~14 pg/mL)	10 mm	No
Costa et al (2011) <sup>40</sup>	WBA, ELISA (QFT-Gold in tube); ESAT6, CFP-10, Rv2654c (p38–55)	0.35 IU/mL (~14 pg/mL)	0 mm	¶No

CFP=culture filtrate protein. ESAT=early secreted antigenic target. HIC=high-income country. LIC=low-income country, MIC=middle-income country, NS=not stated. TST=tuberculin skin test. PPD=purified protein derivative. WBA=whole-blood assay.

\* All measured interferon-gamma. Studies done in Turkey,<sup>27</sup> Colombia,<sup>28</sup> and South Africa<sup>38</sup> described serial testing: interferon- $\gamma$  release assay (IGRA) repeated at 6 months of follow-up, at 2 months of follow-up, and repeated for individuals with suspected tuberculosis, respectively.

† Unit conversions given are equivalent to cutoffs defined in original papers: Hill et al<sup>32</sup> (positive well with eight spot-forming units [SFU]> negative well; one pool of overlapping peptides must be positive); Bakir et al<sup>27</sup> (> 5 mean spot-forming cells [SFC] in duplicate wells than negative wells and if number of SFCs was twice the mean of negative control wells). Incubation time was more than 24 h in the studies done in Ethiopia (120 h) and Colombia (168 h).

‡ Although not reported, study investigators confirmed that clinicians were blinded to IGRA results.

§ Cases were not diagnosed by researchers, but health-care workers were aware of IGRA results.

¶ Clinicians were not blinded to IGRA-positive results.

**Table 3**  
**Summary of study quality (modified Newcastle-Ottawa scale items)**

	Low or intermediate income (n=9)	High income (n=6)
<b>Selection</b>		
Representative sample	9	6
IGRA positive and negative from same source population	9	6
Assay described in detail	9	6
Active tuberculosis excluded at baseline*	8	4
Methods include smear and culture <sup>†</sup>	..	..
Whole or random sample screened for tuberculosis	0	0
IGRA incorporated into reference standard (or not reported) <sup>‡</sup>	2	6
<b>Comparability</b>		
Adjustment of identified confounders <sup>‡</sup>	3	1
<b>Outcome</b>		
Blind assessment and active follow-up by regular visits to the clinic or home to check for tuberculosis <sup>§</sup>	5	0
IGRA incorporated into reference standard (or not reported) <sup>¶</sup>	1	6
>50% incident cases culture-confirmed <sup>  </sup>	4	4
Study follow-up at least 1 year	9	6
80% of cohort followed up <sup>**</sup>	9	4
Outcome reported as incidence rate and rate ratio (person-time incidence) <sup>††</sup>	8	1

\* Studies in Germany,<sup>29</sup> Japan,<sup>39</sup> Kenya,<sup>33</sup> and Norway<sup>31</sup> did not clearly report whether active tuberculosis was adequately excluded at baseline.

<sup>†</sup> Not assessed—studies reported that participants were screened for tuberculosis symptoms and signs, and that only those with a positive screen were investigated further for tuberculosis with methods that may or may not have included smear, culture, or both.

<sup>‡</sup> Incorporation bias (Austria,<sup>26</sup> Japan,<sup>39</sup> Norway,<sup>31</sup> Germany,<sup>29</sup> India,<sup>34</sup> South Africa,<sup>38</sup> Portugal<sup>40</sup>) and not reported (Netherlands<sup>35</sup>).

<sup>‡</sup> No or poor adjustment of confounders (Ethiopia,<sup>30</sup> The Gambia,<sup>32</sup> Turkey [isoniazid preventive treatment only],<sup>27</sup> Austria,<sup>26</sup> Colombia,<sup>28</sup> Norway,<sup>31</sup> Kenya,<sup>33</sup> India,<sup>34</sup> and South Africa<sup>38</sup>).

<sup>§</sup> For ascertainment of any health complaint that could have been active tuberculosis: Turkey,<sup>37</sup> China,<sup>36</sup> Colombia (home visits),<sup>28</sup> Senegal (home visits),<sup>37</sup> and The Gambia<sup>22</sup> (home visits). Studies without active follow-up include the studies done in Kenya<sup>33</sup> and India,<sup>34</sup> which relied on self-reports, tuberculosis subsequently confirmed for nine of 14 in the Indian study.

<sup>¶</sup> Incorporation bias (Austria, Japan, Norway, Germany, South Africa, Portugal) and not reported (Netherlands).

<sup>||</sup> More than 50% cases culture-confirmed (Senegal,<sup>37</sup> China, Netherlands, Japan, Germany, South Africa, India, and Portugal).

<sup>\*\*</sup> Inadequate follow-up of cohort or poor description for three high-income countries—Japan (55%), Norway (cannot estimate), Germany (70%, about 70% of whom were IGRA negative).

<sup>††</sup> Cumulative incidence and risk ratios reported (Ethiopia, and all six studies done in high-income countries).

**Table 4**  
**Concordance in tuberculin skin test and interferon- $\gamma$  release assay results and incidence of tuberculosis**

	n/N	Person-years	Incidence per 1000 person-years (95% CI)	Incidence rate ratio (95% CI)
<b>The Gambia</b> <sup>32</sup> *				
IGRA+/TST-	4/177	322.6	12.4 (0.3–24.5)	1.29 (0.24–6.93)
IGRA-/TST+	4/230	416.7	9.6 (0.2–19)	1
IGRA+/TST+	7/428	790.1	8.86 (2.4–15.4)	2.2 (0.63–7.98)
IGRA-/TST-	6/183	1500	4.0 (0.8–7.2)	1
<b>Turkey</b> <sup>27</sup> †				
IGRA+/TST-	1/44	86	11.7 (0.3–65.1)	1.58 (0.03–30.38)
IGRA-/TST+	2/213	272	7.4 (0.9–26.6)	1
IGRA+/TST+	10/337	451	22.2 (10.6–40.8)	5 (1.07–46.93)
IGRA-/TST-	2/314	451	5.1 (0.6–18.4)	1
<b>Senegal</b> <sup>37</sup> ‡				
IGRA+/TST-	1/77	170	5.9 (0.83–41.87)	1.19 (0.02–22.94)
IGRA-/TST+	2/193	406	4.93 (1.23–19.7)	1
IGRA+/TST+	14/436	950	14.74 (8.73–24.89)	1.50 (0.47–6.2)
IGRA-/TST-	4/187	406	9.85 (3.79–26.25)	1
<b>India</b> <sup>34</sup> *				
IGRA+/TST-	1/58	345.6	2.89 (0.4–20.5)	1.1 (0.01–84.97)
IGRA-/TST+	1/63	374.1	2.67 (0.4–19.0)	1
IGRA+/TST+	5/217	1280.8	3.9 (1.6–9.4)	1.11 (0.28–4.10)
IGRA-/TST-	7/336	1991.7	3.5 (1.7–7.4)	1

Studies with possible incorporation bias excluded. IGRA=interferon- $\gamma$  release assay. SFC=spot-forming cells. TST=tuberculin skin test.

\* 10 mm TST cutoff, 14 pg/mL IGRA cutoff.

† 5 mm, 20 SFC  $\times 10^6$ .

‡ 10mm, 32 SFC  $\times 10^6$ .