Interferon-γ release assays for the diagnosis of latent 
*M. tuberculosis* infection: A systematic review and meta-analysis

Diel R¹, Goletti D², Ferrara G³, Bothamley G⁴, Cirillo D⁵, Kampmann B⁶, Lange C⁷, Losi M⁸, Markova R⁹, Migliori GB¹⁰, Nienhaus A¹¹, Ruhwald M¹², Wagner D¹³, Zellweger JP¹⁴, Huitric E¹⁵, Sandgren A¹⁵, Manissero D¹⁵

¹Department of Pneumology, Medical School Hannover, Hannover, Germany; ²Department of Epidemiology and Preclinical Research, National Institute for Infectious Diseases L. Spallanzani, Rome, Italy; ³Section of Respiratory Diseases, Department of Internal Medicine, S Maria Hospital, University of Perugia, Terni, Italy; ⁴Homerton University Hospital, NHS Foundation Trust, London, UK; ⁵Emerging Bacterial Pathogens Unit, San Raffaele Scientific Institute, Milan, Italy; ⁶Academic Department of Paediatrics, Imperial College, London, United Kingdom; ⁷Division of Clinical Infectious Diseases, Research Center Borstel, Borstel, Germany; ⁸Department of Oncology, Hematology and Respiratory diseases, University of Modena and Reggio Emilia, Modena, Italy; ⁹Immunology and Allergology, National Center of Infectious and Parasitic Diseases, Sofia, Bulgaria; ¹⁰WHO Collaborating Centre for Tuberculosis and Lung Diseases, Tradate, Italy; ¹¹Institute for Health Service Research in Dermatology and Nursing; University Medical Center Hamburg-Eppendorf, Germany; ¹²Clinical Research Centre, Hvidovre Hospital, Copenhagen University, Copenhagen, Denmark; ¹³Division of Infectious Diseases and Centre of Chronic Immunodeficiency, University Medical Center, Freiburg, Germany; ¹⁴Department of ambulatory care and community medicine, University of Lausanne, Lausanne, Switzerland. ¹⁵European Centre for Disease Prevention and Control (ECDC), Stockholm, Sweden

**Corresponding author contact information:** Christoph Lange, MD, PhD; Division of Clinical Infectious Diseases, Medical Clinic, Research Center Borstel, Parkallee 35, 23845 Borstel, Germany, phone +49 4537 188-332, fax +49 4537 188-313, email: clange@fz-borstel.de
Abstract (200 words)
We conducted a systematic review and meta-analysis to compare the accuracy of the QuantiFERON-TB Gold In-Tube (QFT-G-IT) and the T-SPOT.TB assays with the tuberculin skin test (TST) for the diagnosis of latent Mycobacterium tuberculosis infection (LTBI). Medline, EMBASE and Cochrane databases were explored for relevant articles through November 2009. Specificities, negative (NPV) and positive (PPV) predictive values of interferon-γ release assays (IGRAs) and the TST and the exposure gradient influences on test results among bacillus-Calmette-Guérin (BCG) vaccinees were evaluated. Specificity of IGRAs varied from 98%-100%. In non-immunocompromized adults, NPV for progression to tuberculosis within 2 years were 97.8% for T-SPOT.TB and 99.8% for QFT-G-IT, respectively. When test performance of an immunodiagnostic test was not restricted to prior positivity of another test, progression rates to tuberculosis among IGRA-positive individuals followed for 19-24 months varied between 8-15%, exceeding those reported for the TST (2-3%). In multivariate analyses, the odd ratios for TST positivity following BCG vaccination varied from 3-25, whereas IGRA results remained uninfluenced and IGRA positivity was clearly associated with exposure to contagious tuberculosis cases. IGRAs may have a relative advantage over the TST in detecting LTBI and allow the exclusion of M. tuberculosis infection with higher reliability.

Keywords:
Interferon-γ release assay, meta-analysis, latent Mycobacterium tuberculosis infection, systematic review, tuberculin skin test, tuberculosis

Introduction
Improvement of diagnostic methods for latent infection with Mycobacterium tuberculosis (LTBI) is an important step in the progress towards reaching the goal of tuberculosis elimination, brought forward by the WHO Stop TB strategy [1] and the Framework Action Plan to Fight TB in the European Union by The European Centre for Disease Prevention and Control (ECDC) [2]. As part of reaching this goal, individuals infected with M. tuberculosis need to be identified and offered preventive therapy to stop the progression to active tuberculosis and prevent further M. tuberculosis transmission [3]. Thus, there is a need to develop more accurate methods for the detection of LTBI, and to provide evidence-based guidance on the use for such methods before they can be adopted by national tuberculosis screening programmes [4, 5].

The identification of LTBI relies in most areas in Europe upon the tuberculin skin test (TST). This diagnostic test has been assessed comprehensively in terms of potentials and limitations for its use in preventive strategies for tuberculosis elimination [6]. However, the TST does not discriminate between potential infection with M. tuberculosis and prior vaccination with the bacillus of Calmette-Guérin (BCG), or possible infection with non-tuberculous mycobacteria (NTM).

Interferon-γ release assays (IGRAs) are in vitro immune assays that have been introduced in recent years as an alternative to the TST for the diagnosis of LTBI. IGRAs are based on the detection of a T-cell immune response towards M. tuberculosis complex specific antigens (early secretory antigenic target [ESAT]-6 and culture filtrate protein [CFP]-10 ± TB7.7). To date there are two commercially available platforms that measure IFN-γ production following ex-vivo antigen stimulation: In the QuantiFERON-TB Gold In-Tube test (QFT-G-IT) and its predecessor the QuantiFERON-TB Gold (QFT-G) test (Cellestis Ltd., Carnegie, Australia) [7] the amount of IFN-γ released into the supernatant is quantified using an enzyme-linked
immunosorbent assay (ELISA) whereas the proportion of blood cells releasing IFN-γ is
determined with the T-SPOT.TB assay (Oxford Immunotec Ltd., Abingdon, UK) [8] using an
enzyme-linked immunospot (ELISPOT) technique [9]. As the antigens used in IGRAs are
almost exclusively expressed by the M. tuberculosis complex – with the exception of M.
kansasii, M. szulgai, M. marinum and M. riyadhense – IGRAs are less likely to be
confounded by prior BCG vaccination and/or exposure to NTMs [10].
For any new test to replace the TST, evidence of the test’s higher diagnostic accuracy is
needed. Importantly, such a test should show a higher specificity (ability to exclude M. tuberculosiss infection) especially in subjects with confounding factors such as BCG vaccination and at least a similar sensitivity. However, the lack of a gold standard for
diagnosing LTBI is a problem when trying to determine the diagnostic accuracy of both, the
IGRAs and the TST [11]. An alternative way to evaluate sensitivity, which cannot be assessed
directly in this context, is through comparison of the risk to develop active tuberculosis in
subjects with an increased risk of progression from LTBI to active disease.
The adoption and implementation of IGRAs into national tuberculosis screening programmes
should be evidence-based. Several countries have to date adopted IGRAs for the diagnosis of
LTBI within tuberculosis screening programmes and guidance documents on their use are
available [e.g. 4, 12, 13]. Clinical guidance for the diagnosis of LTBI is especially important
for the care of immunocompromized individuals, e.g. patients undergoing tumor-necrosis-
factor (TNF)-antagonists therapies [14], who carry an increased risk for progression to active
tuberculosis. Several studies and consequent systematic reviews and meta-analyses have been
conducted evaluating the IGRAs accuracy in diagnosing LTBI [15-18]. Furthermore,
evidence on the actual predictive value of the IGRAs (i.e. the relationship between having a
positive or negative IGRA result and developing/not developing tuberculosis) is emerging
[19-22]. To further support public health programmes, a continuous assessment of the use of
IGRA for the diagnosis of LTBI is therefore needed.
This manuscript provides important up-to-date information on the use of IGRAs to support
the development of evidence-based guidance for tuberculosis screening programmes. These
data may also be helpful for conducting more reliable cost effectiveness and cost benefit-
analyses with respect to the use of IGRA or TST solely or in combination.
Particularly, the diagnostic specificity of IGRAs relative to the TST; the IGRAs’ negative
(NPV) and positive (PPV) predictive values for the progression to active tuberculosis and the
association between exposure to patients with tuberculosis and IGRA and TST test results in
contacts are being evaluated.
We present a systematic review and meta-analysis of the literature according to evidence-
based highest standard criteria on the accuracy of commercially available IGRAs for the
diagnosis of LTBI.
Methods
We conducted this systematic review according to the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [23] and the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) checklist [24].

Search strategy and selection criteria
We searched MEDLINE, EMBASE and the Cochrane Central Register of Controlled Trials (CENTRAL; Cochrane Library 2009, issue 3) through November 15, 2009 for studies published in English. The following search terms: “Tuberculosis” OR “TB infection” OR “TB disease” AND “QuantiFERON” OR “Elispot” OR “T-SPOT” OR “interferon-gamma assay” OR “interferon-γ release assay” OR “T cell assay” AND “ESAT-6” OR “CFP-10”, were separately combined with the terms “NPV” OR “negative predictive value” OR “negative PV”, “PPV” OR “positive predictive value” OR “predictive” as well as “TST” AND “BCG”. No restrictions were made with respect to basic study design or data collection (prospective or retrospective). Secondary references were made with respect to basic study design or data collection (prospective or retrospective). Secondary references were made with respect to basic study design or data collection (prospective or retrospective). Secondary references were made with respect to basic study design or data collection (prospective or retrospective). Secondary references were made with respect to basic study design or data collection (prospective or retrospective).

From studies identified as potentially relevant, original articles or letters to the editor were chosen that met the following criteria: Original data were presented, the most recent commercially available IGRAs (QFT-G-IT and T-SPOT.TB) were used with European cut-offs for positive test results; active disease in participants was confirmed by culture and/or NAAT and/or histopathological examination; if comparing results of two or more tests, these had to be performed in the same individuals.

The following studies were excluded: case reports, editorials, immunological (laboratory) and animal studies; studies not following the manufacturers’ instructions, e.g. in vitro incubation >24 h or freezing of cells; studies involving patients whose diagnosis of active tuberculosis relied solely on clinical evaluation or radiological features, improvement of symptoms while on anti-tuberculosis therapy, and/or on smear positivity; mixed studies, in which the number of unconfirmed and confirmed patients was not presented separately, thus possibly resulting in a selection bias. Three reviewers performed the review of titles and/or abstracts and 53 publications were retrieved for full text review (see figure 1).

Data extraction and quality assessment
Both study selection and data extraction were conducted independently, in duplicate, by members of the authors, in order to reduce the risk of errors. A data extraction form was created including a subset of predefined items covering inclusion and exclusion criteria as described above. These criteria included study setting (country), year(s) of performance, type of patients and patient method used to recruit participants (including number of eligible and subjects), study duration (from/to), and loss of participants during the observation period follow-up.

For analysis of specificity, NPV, NPV for progression and PPV for progression (see section below on “Data synthesis and analysis” for definitions), the following study contents were documented: IGRA(s) used, mean age or age composition, numbers of participants finally included, numbers of subjects for whom valid test results were available (after subtraction of indeterminate, i.e. uninterpretable results), duration of follow up, number of test results

1In US FDA-approved criteria a person is T-SPOT.TB negative if they have ≤4 spots above the Nil control, positive if ≥8, and borderline (or equivocal) for 5, 6, or 7 spots.
considered true negative and false positive in persons at very low risk of *M. tuberculosis* infection (specificity), number of confirmed tuberculosis cases with negative IGRA scores (i.e. number of false negatives in active tuberculosis patients to measure NPV), number of confirmed non-tuberculosis cases with negative IGRA scores (i.e. number true negatives in suspects of active tuberculosis to measure NPV), number of score-negative test results among persons at risk and screened for LTBI who later developed active disease (i.e. false-negatives for NPV for progression), number of score-negative persons screened for LTBI that did not develop active tuberculosis (true negative for NPV for progression calculation), untreated score-positive persons who developed active tuberculosis during follow-up (PPV for progression), and lastly number of score-positive persons screened for LTBI with a confirmed relationship to their index case as demonstrated by highly discriminating molecular genotyping.

For analysis of the correlation between TST- and IGRA-results and infection-exposure among BCG-vaccinated or unvaccinated subjects, the following items were also included (or recalculated if possible from the available data): time of BCG vaccination, assessment of BCG vaccination among the study participants, TST units, type and chosen cut-off (induration diameter), number of participants finally tested by TST and at least one IGRA, assay used, number of subjects for whom test results were available, TST and IGRA results (if available, also in BCG-vaccinated or unvaccinated subgroups), number of indeterminate results (if any), results on IGRA/IGRA/TST-concordance and discordance, statistical agreement (κ statistics) in total and subgroups, and lastly, odd ratios (ORs) of independent predictors for IGRA- and TST-positivity. In the case of TST-performance, results using the cut-off chosen by the authors, as part of the respective study design, were taken for our review. Additional clarifications on studies were requested by personal correspondence with the authors. Any discrepancies were resolved by consensus with the help of the team coordinator and thus obtaining an inter-reviewers agreement of 100%.

**Data synthesis and analysis**

The following definitions were used for analysis:

The *specificity* is defined as the number of true negatives divided by the sum of true negatives and false positives. As there is no gold standard to ascertain LTBI in a study subject, specificity studies are typically conducted in low prevalence countries and settings in which the population is at no risk, suspicion or evidence of LTBI or prior tuberculosis (any individual with a positive test score being considered false-positive). Specificity is therefore the measure of a test’s ability to score LTBI-free, healthy subjects as test-negative. This definition holds for both IGRA and the TST, irrespective of BCG vaccination status.

**Negative predictive value (NPV):** Generally, the NPV of a screening test is defined as the number of true negative test results, divided by the sum of true and false negative results, all in individuals suspected to have a disease. With regards to diagnosing active tuberculosis, this refers to the degree to which a test does not score a person with active tuberculosis, as test negative i.e. the certainty that a score-negative person does not have active tuberculosis. Translating this definition to persons suspected of LTBI, the NPV measures the degree to which a test does not score a person infected with *M. tuberculosis*, as score-negative; i.e. the certainty that a score-negative person does not have LTBI. Again, due to the lack of a gold standard for LTBI identification, studies commonly determine the NPV among patients with confirmed active tuberculosis, using the proportion of patients with false-negative test-scores as a surrogate for the proportion of false-negative scores in LTBI-suspects. The number of true negative cases is consequently the number of individuals, suspected of active tuberculosis that are subsequently confirmed not to have active disease.

**NPV for progression:** NPV for progression is defined as the proportion of IGRA score-negative subjects that do not progress to active tuberculosis in a longitudinal, follow-up study.
of individuals tested for LTBI. Subsequently, this value reflects a test’s ability to correctly predict that an LTBI score-negative individual will not develop active tuberculosis later on in life, provided no further exposure to infection. Seeing that only subjects infected with \(M.\) \(tuberculosis\) can develop disease, the NPV can be measured by following subjects that scored negative for LTBI, over time, and quantifying the number that have remained free from active disease.

**Positive Predictive Value (PPV) for progression:** The PPV for progression is defined as the proportion of IGRA score-positive subjects that develop active tuberculosis in a longitudinal, follow-up study of individuals tested for LTBI. Subsequently, this value reflects a test’s ability to correctly predict that an LTBI score-positive individual is at risk of developing active tuberculosis later on in life. As for NPV progression studies, this value is measured by following subjects that scored positive for LTBI, over time, and quantifying the number that did develop active disease.

Estimates of specificity, NPV and PPV and their respective 95% confidence intervals (95% CI) were calculated for each of the included studies and used, if applicable, to calculate pooled estimates. Pooled estimates are weighted averages built by using the study sample sizes as implicit weights. Weighted pooled estimates were compared by the Pearson’s \(\chi^2\)-squared test (or Fisher’s exact test, when the expected cell sizes were smaller than five). All \(p\) values reported are based on two-tailed comparisons with the statistical significance set at \(p<0.05\).

Statistical heterogeneity between studies was assessed using the \(I^2\) statistic described by Higgins et al. [18]. The test describes the percentage of total variability, within a set of effect-sizes that is caused by true heterogeneity rather than by sampling error (chance). \(I^2=100\% \times \frac{(Q-df)}{Q}\), where \(Q\) is the chi-squared statistic and \(df\) enumerates the degrees of freedom. If applicable, “forest plots” were built to investigate whether outcomes were stable over a series of studies. Statistical analyses were performed by MetaDisc software, version 1.4 [25], and by applying SPSS version 18.0 for Windows (SPSS Inc, Chicago, Illinois, USA).
Results

The study selection process is shown in figure 1. Four hundred and thirty-two citations were identified, from which 60 articles were eligible for final inclusion for one or more of the following topics. Overall the quality, evaluated by the QUADAS tool, was very high and 11/12 (91.7%) study quality indicators were met by 100% of the included studies, thereby increasing the strength of scientific evidence of the review.

Specificity

Only four published studies evaluated the QFT-G-IT and/or T-SPOT.TB specificity in a total of 346 subjects with valid IGRA results [26-29], fulfilling the inclusion criteria. The specificity of the IGRAAs ranged from 98% [95% CI: 86.8-99.9%] [26] for the T-SPOT.TB to 100% [95% CI: 97.6–100%] [29] for the QFT-G-IT, for which specificity estimates were available in all four studies of 99.4% [95% CI: 97.9-99.9%] (Table 1). We excluded studies from analysis that either used criteria from the USA for scoring T-SPOT.TB results [30], or that were performed in intermediate tuberculosis burden countries. In the latter, even individuals belonging to low risk groups may have been exposed to unknown M. tuberculosis infection with a potentially higher risk of infection than that expected in low burden countries [31-33]. Regarding the comparison of IGRAAs with TST, one report [28] was excluded because the TST was not performed in all the enrolled subjects. From the data that could be analysed, TST specificity ranged from 55% [95%CI, 38.5–70.7] [26] to 95% [CI, 87.7–97.2%] [29], with a pooled specificity of 88.7% [95% CI, 84.6- 92.0]. When comparing the IGRAAs and TST results, only a few subjects unexposed to M. tuberculosis scored positive by IGRAAs compared to TST, in particular individuals with known prior BCG vaccination or confirmed NTM infection (Table 1).

Negative predictive value in patients with active tuberculosis

Eighteen articles satisfied our inclusion criteria for evaluating the NPV; 13 assessed the NPV among confirmed tuberculosis cases [22, 26, 34-44] and six assessed the prospective outcome in IGRA-negative individuals after an average of two years [19-22, 45, 46]. Among the studies in which the NPV was evaluated in patients with confirmed tuberculosis (387 tuberculosis cases with a valid T-SPOT.TB result and 304 with valid QFT-G-IT result), the NPV varied greatly, irrespective of the IGRA used. Among the 13 articles, the NPV ranged between 74.4% [36] and 100% [22, 34], with a pooled value of 94% [95% CI, 92.1–95.6%]) for the T-SPOT.TB and 88% [95% CI, 84.6–91.5%] for the QFT-G-IT (figures 2 and 3).

In the six longitudinal studies assessing the NPV among healthy persons, a total of 1442 individuals scored negative by the QFT-G-IT and 182 scored negative by the T-SPOT.TB (figures 4 and 5) [19-22] [45, 46].

Negative predictive value for progression

Although subjects at increased risk for developing tuberculosis were included in all studies, the pooled NPV for progression was high in the five studies performed in low-incidence countries. For these five studies, the pooled NPV for the QFT-G-IT was 99.8% [95% CI, 99.4–100%] (3 individuals contracted tuberculosis among the 1442 scored negatives) and 97.8% for the T-SPOT.TB [95% CI, 94.5–99.4%] (4 individuals contracted active tuberculosis among the 182 scored negatives). Conversely, the NPV of the included study [46] performed in Thailand, an intermediate burden country, was 88% [95% CI, 63.6–98.5%]. In a majority of the evaluated studies, it was not possible to compare the NPV estimates for IGRAAs to that of the TST results [19, 21, 22] [45, 46]. In particular, in two studies performed
on HIV-infected persons, TST-negative individuals were excluded from the follow-up [19, 22]. Only one study performed a comprehensive follow-up that included TST- and IGRA-scored negative subjects [20]. In this study, only 1 out of 354 TST-score negative close contacts developed tuberculosis, resulting in an NPV for the TST of 99.7% [95% CI, 98.5–100%] compared to 100% for the QFT-G-IT (95% CI, 99.4–100%).

**Positive predictive value for progression**

Four studies, investigating the probability of developing active disease in subjects with a positive IGRA result were included for analysis (Table 2). Two studies were screening-studies among HIV-positive subjects, while the remaining two consisted of large contact investigations among native and immigrant subjects in Germany and the Netherlands. Only one study [20] included children aged below 16 years. The rate of progression to active disease among subjects having tested positive for LTBI, and having refused preventive treatment, ranged between 2.3-3.3% for TST, 2.8-14.3% for QFT-G-IT and 3.3-10% for T-SPOT.TB (Table 2).

In a study by Clark et al. [22], 201 HIV-infected patients diagnosed with LTBI were followed for a median period of 12 months. Two of the 20 HIV-infected patients with positive T-SPOT.TB results, and who did not receive preventive treatment, developed active tuberculosis, resulting in a PPV for progression of 10% (3 and 10 months post-testing, no detail or comparative results with the TST were reported).

In an extensive contact investigation study, Dièl et al. [20] followed the outcome of 41 QFT-G-IT-positive close contacts, all of which did not receive preventive therapy, for a period of up to 2 years. Four of the individuals were TST-negative and 3 were children under the age of 16 years. Within this period, 6 of the 41 untreated, QFT-G-IT-positive subjects developed active tuberculosis (PPV for progression of 14.6%), including one QFT-G-IT-positive, TST-negative contact. In comparison, of the 219 untreated TST-positive contacts followed, only 5 developed tuberculosis (PPV for progression of 2.3%). The progression rate increased to 5.6% (5/90), when a 10 mm TST cut-off was applied.

Aichelburg et al. [19] enrolled 830 HIV-infected individuals in an extensive LTBI screening study. Among the study group, 36 were QFT-G-IT-positive, subsequently diagnosed with LTBI and followed for a mean period of 19 months without receiving preventive treatment. An additional 7 individuals also tested positive at the time of screening, however were then diagnosed with active tuberculosis and subsequently excluded from the follow-up. Of the 36 QFT-G-IT-positive individuals, three had developed active tuberculosis within the follow-up period, resulting in a PPV for progression of 8.3%. Comparison between TST and QFT-G-IT results with respect to progression of disease was not possible due to the fact that TST was not administered in all individuals.

In their study, Kik et al. [21] included 433 close immigrant, adult contacts in the Netherlands, none of which were immunosuppressed. QFT-G-IT or a T-SPOT.TB was performed only in TST-positive patients or in persons with a known prior TST result ≥10 mm (n=17). All TST-positive contacts were followed for up to 24 months. In total, 339 (78.3%) of the 433 contacts were TST-positive. IGRA was performed on 327 of the TST-positive contacts, of which 179 (54.4%) tested positive using the QFT-G-IT and 181 (60.5%) tested T-SPOT.TB positive. Nine (2.6%) of the 339 contacts developed active tuberculosis; IGRA had been performed in eight of these contacts of which six had scored positive. The PPV value for developing active tuberculosis was 3.1% (9/288; 95% CI, 1.4–5.8%) for TST (cut-off ≥10 mm), 2.8% (5/178; 95% CI, 0.9–6.4) for QFT-G-IT and 3.3% (6/181; 95% CI, 1.2–7.0) for T-SPOT.TB.

**Association of IGRA and the TST with MTB exposure and BCG vaccination**

Thirty-four studies [27, 39, 45, 47-77] from a total of 18 different countries fulfilled the inclusion criteria for this section of the review. Four studies were from intermediate-TB
burden countries (Korea [55], Lithuania [59], Iran [63], Turkey [72]) and three studies from high-burden countries (India [56], South Africa [61] and Vietnam [65]). Seven studies were based solely on pediatric cohorts [39, 51, 54, 59, 73, 75] and three other studies [47, 57, 61] included children in the study population. Seventeen of the 34 articles calculated the odd ratios (OR) assessing the influence of different factors on test positivity in multivariate analyses. In nine of these, TST positivity was significantly associated with BCG vaccination as an independent predictor in multivariate analysis [47, 50, 55, 62, 67, 71, 72, 74, 77] with odd ratios ranging between 3.8 [95% CI, 1.0–13.9] [50] and 24.7 [95% CI, 11.7–52.5] [48] regardless of the epidemiological context. No correlation was found between IGRA positivity and BCG vaccination. In nine of the ten studies comparing the odds for test positivity with M. tuberculosis exposure gradients, or using chest radiography lesions as a surrogate of prior exposure, the IGRA-associated better with exposure than the TST, irrespective of the setting’s disease-burden. Furthermore, there was generally a poor agreement between IGRA and TST results (for κ-statistics, see Table 3).
Discussion
In the last few years a large number of studies have been published on the performance of IGRAs for the diagnosis of LTBI in different cohorts at risk for progression to active tuberculosis. We performed a systematic review and meta-analysis to assess the accuracy of commercially available IGRAs in terms of specificity, NPV and PPV as the main estimates. Despite the volume of publications on the subject, only a limited number of studies fulfilled the inclusion criteria for our systematic review and meta-analysis. This is due to the use of non-commercial IGRAs or to methodological limits in the design of individual studies. Based on the results obtained from our analysis we identified that: 1) the IGRAs show a higher specificity as compared to the TST; 2) the NPV, when using patients with active tuberculosis as a surrogate of LTBI, is high (although with a high variability in values); 3) the ability of the IGRAs to predict that a test-negative individual will not develop disease is even higher; and lastly, 4) the predictive value for progression to active disease when testing positive is higher for the IGRAs as compared to the TST. However, NPV and PPV remain to be established in the paediatric population and in immunocompromized individuals, e.g. patients with advanced HIV-infection, transplant recipients or candidates for TNF antagonists therapies.

The pooled specificity (99.4% [95% CI, 97.8–99.9%]) when measured in low-risk population groups using the QFT-G-IT (the only IGRA for which results could be analysed in all included studies) was clearly higher than the pooled specificity for the TST (88.7% [95% CI, 84.6–92.0]). These results suggest that the IGRAs are more certain to correctly identify individuals not infected with *M. tuberculosis* as compared to the TST. However, the estimates have to be interpreted with caution due to the low number of included studies. An optimal LTBI diagnostic tool should be predictive of an individual’s risks for developing or not developing active disease when tested. When analyzing the studies using cohorts of active tuberculosis cases as surrogates of LTBI, a high overall NPV value was measured (pooled NPV 94% for T-SPOT.TB and 88% for QFT-G-IT). This would suggest that IGRAs, especially the T-SPOT.TB, are effective at ruling out *M. tuberculosis* infection. However, caution must be taken when interpreting these results as there was a high variability between the studies (ranging from 70-98% for the QFT-G-IT and 92-100% for the T-SPOT.TB) likely due to the corresponding prevalence of active tuberculosis cases, thus limiting the utility of pooling data across studies. This would therefore refute the conclusion that the IGRAs are an effective “rule-out-test” for LTBI.

Conversely, longitudinal NPV studies on mostly non-immunocompromized adults in low prevalence countries, who are at risk for LTBI, suggest that one can be reasonably confident that, when scoring negative with an IGRA, the likelihood of false negative results is low. High NPVs were found for both the QFT-G-IT and the T-SPOT.TB (pooled NPV of 97.8% for T-SPOT.TB and 99.8% for QFT-G-IT). This indicates that an individual testing negative will most likely not develop tuberculosis in the future. However, a limitation to this analysis was the low number of individuals included in the studies (total 1442), the short durations of follow up, and the fact that similar studies have not yet been performed in high-burden settings.

The QFT-G-IT and the T-SPOT.TB both showed a similar PPV for progression, with a slightly wider range in values as compared to the TST (2.8%-14.3% for QFT-G-IT; 3.3-10% for T-SPOT.TB and 2.3-3.3% for TST), suggesting that the IGRAs have a higher predictive value for progression to active disease as compared to the TST. If IGRA performance was not conditionally restricted to prior TST positivity (as was the case in one study), the progression rates were higher (8%, 10% and 15%) than those reported for TST, with two years follow up [19, 20, 22]. In the study by Kik et al. [21], a lower PPV value for progression was measured
(3%) however, the study was limited to only assessing the IGRA among TST-positive individuals, thereby presenting a source of bias.

Studies on predictive value of the IGRAs with regard to developing active tuberculosis upon testing positive are still very few, with only four studies having fulfilled our inclusion criteria, and they often vary substantially in design and rely on empirical observations of subjects refusing LTBI treatment. Additional larger-sized studies including new markers for LTBI are needed to evaluate the predictive values of IGRAs and other biomarkers in groups with the highest risk of progression to tuberculosis, especially in children and in immunocompromised hosts.

The review of studies investigating the influence of BCG vaccination on TST positivity and the influence of exposures for LTBI clearly confirmed – despite the large heterogeneity in the design of the analysed studies – that, unlike the TST, the newest commercial IGRAs are not affected by prior BCG vaccination. They are subsequently more likely associated with exposure to tuberculosis cases compared to the TST. These results were not influenced by the epidemiological context where the studies were performed (high vs. low incidence countries).

The review only assessed TST results in the context of IGRA studies and thus the results on prior studies on accuracy performing the TST solely could not be taken into consideration. Furthermore, we did not evaluate the ability of IGRAs to discriminate recent from remote LTBI and the phenomenon of conversion/reversion over time or after therapy. Although this would be a very important point in a public health programme, the data available today [78-82] are limited and mostly obtained with experimental techniques and in particular settings, making it difficult to interpret their potential clinical value. In conclusion, the current systematic review and meta-analysis on the accuracy of the IGRAs for LTBI diagnosis confirm the concept that the IGRAs are a valid alternative to the TST. The superior specificity and the good NPV make them the first choice, especially in BCG-vaccinated subjects. The current evidence is however still limited to determine whether the IGRAs have a stronger predictive value for developing active disease later in life, as compared to the TST. In the development of guidance on the use of IGRAs in tuberculosis screening programmes, it is vital that such evidence is considered when describing diagnostic algorithms for the diagnosis of LTBI. In this way we can provide support to assure that only assays with proven accuracy are introduced into national screening programmes for further advancement towards the elimination of tuberculosis.

**Acknowledgment:** The authors thank Prof. Martina Sester, Homburg (Germany) and Claudia Giehl (Saarbrücken, Germany) for managing and editorial support and Prof. Giovanni Sotgiu Sassari (Italy) for statistical advice. TBNET is a Clinical Research Collaboration (CRC) of the European Respiratory Society (ERS).
References


76. Vassilopoulos D, Stamoulis N, Hadziyannis E, Archimandritis AJ. Usefulness of enzyme-linked immunospot assay (Elisport) compared to tuberculin skin testing for latent tuberculosis screening in rheumatic patients scheduled for anti-tumor necrosis factor treatment. *J Rheumatol* 2008: 35(7): 1271-1276.


Table 1: Specificity for LTBI diagnosis in individuals with very low risk for LTBI tested by QFT-G-IT and/or T-SPOT.TB

<table>
<thead>
<tr>
<th>Study, Year</th>
<th>Country</th>
<th>Type of participants</th>
<th>IGRA(s) used</th>
<th>Mean age (yrs ±SD)</th>
<th>No. subjects assigned as low risk for MTB infection</th>
<th>No. subjects with negative tests/all tested subjects*</th>
<th>No. subjects considered false positive</th>
<th>Specificity % (rounded), [CI 95%] of IGRA(s)/TST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detjen et al. 2007 [26]</td>
<td>Germany (Low burden)</td>
<td>Children with confirmed non-TB lymphadenitis (44 median) or respiratory infections</td>
<td>T-SPOT.TB/ QFT-G-IT</td>
<td>44 months (median); Others: 52.5 months</td>
<td>45</td>
<td>T-SPOT.TB: 39/40; QFT-G-IT: 40/40; TST: 22/40</td>
<td>T-SPOT.TB: 1; QFT-G-IT: 0; TST: 18 (all 18 with NTM lymphadenitis)</td>
<td>98% [86.8–99.9%]; QFT-G-IT: 100% [92.8–100%]; TST: 55% [38.5–70.7%]; QFT-G-IT: 99% [95.8–99.9%]; TST: 94% [88.6–97.1%]</td>
</tr>
<tr>
<td>Franken et al. 2007 [27]</td>
<td>Netherlands (Low burden)</td>
<td>Dutch Armed Forces personnel</td>
<td>QFT-G-IT</td>
<td>19.6 ±2.8</td>
<td>171</td>
<td>QFT-G-IT: 166/168; TST: 136/145</td>
<td>QFT-G-IT: 2; TST: 9</td>
<td>99% [95.8–99.9%]; TST: 94% [88.6–97.1%]</td>
</tr>
<tr>
<td>Palazzo et al. 2008 [28]</td>
<td>Italy (Low burden)</td>
<td>Healthy blood donors as controls for TB suspects</td>
<td>QFT-G-IT</td>
<td>37.2 ±2 (BCG unvacc.); 35.2± 2 (BCG vacc.)</td>
<td>24</td>
<td>QFT-G-IT: 14/14; TST: N.A.</td>
<td>QFT-G-IT: 0</td>
<td>100% [80.7–100%]; TST: N.A.</td>
</tr>
<tr>
<td>Ruhwald et al. 2008 [29]</td>
<td>Denmark (Low burden)</td>
<td>86 high school students and 38 high school staff</td>
<td>QFT-G-IT</td>
<td>17.6±1.3; Staff: 54.5±8.5</td>
<td>124</td>
<td>QFT-G-IT:124/124; TST: 116/124</td>
<td>QFT-G-IT: 0; TST: 8</td>
<td>100% [97.6–100%]; TST: 95% [87.7–97.2%]</td>
</tr>
</tbody>
</table>

Pooled specificity QFT-G-IT: 99.4% [95% CI 97.9–99.9%]
Chi-Square=2.90; df=3 (p=0.4069)
Inconsistency (I-Square)=0.0%

Pooled specificity TST: 88.7% [95% CI 84.6–92.0%]
Chi-Square=2.90; df=2 (p=0.000)
Inconsistency (I-Square)=94.5%

*after subtraction of invalid results/indeterminates

*Among recruits, 5/171 results were QFT-G-IT positive. Of these recruits, two had previously been treated for tuberculosis and one was foreign born without information of the tuberculosis prevalence in his home country, thus reducing the number of false positive subjects and the number of tested persons eligible for analysis accordingly.

b13 BCG-unvaccinated and 1 BCG-vaccinated control persons as shown in Table 2 of the publication. Evaluation of TST results was not performed because TST was done in only 7 of the 14 persons included in the study.
Table 2: Positive predictive value of commercial IGRA(s) for progression in those with latent tuberculosis infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design</th>
<th>Number and type of participants</th>
<th>Country and burden of TB disease</th>
<th>IGRA(s) used</th>
<th>Mean age (yrs ±SD)</th>
<th>Performance of IGRA(s)/TST testing</th>
<th>Follow up period (months)</th>
<th>PPV: No. later TB cases among untreated IGRA/TST positives (%; [CI 95%])*</th>
<th>No. later TB cases among IGRA/TST positives with confirmed relationship to index cases by RFLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clark et al. 2007 [22]</td>
<td>Case-control</td>
<td>201 immune-deficient HIV-1 positive adults</td>
<td>UK (Low burden)</td>
<td>T-SPOT.TB</td>
<td>median 40 (range: 33–46)</td>
<td>N.A.</td>
<td>24</td>
<td>2/20 (10% [0.012–0.32])/ N.A. for TST positives</td>
<td>N.A.</td>
</tr>
<tr>
<td>Diel et al. 2008 [20]</td>
<td>Cohort</td>
<td>601 close contacts of AFB positive TB cases (66 -10.9% - children &lt;16yrs)</td>
<td>Germany (Low burden)</td>
<td>QFT-G-IT</td>
<td>27.7 (±12.0) (range: 1–56)</td>
<td>Simultaneously with TST</td>
<td>24 (mean)</td>
<td>6/41 QFT-G-IT positives (14.6%, [0.06–0.29])/ 5/219 TST positives &gt;5mm (2.3% [0.007–0.052])</td>
<td>QFT-G-IT: 2/6 (33.3%)/ TST: 1/5 (20%)</td>
</tr>
<tr>
<td>Aichelburg et al. 2009 [19]</td>
<td>Cohort</td>
<td>830 HIV-1 positive adults</td>
<td>Austria (Low burden)</td>
<td>QFT-G-IT</td>
<td>39 (inter-quartile range: 32–47)</td>
<td>IGRA first, TST only if QFT positive</td>
<td>19 (mean)</td>
<td>3/36 (8.3% [0.018–0.22])/ N.A. for TST positives</td>
<td>N.A.</td>
</tr>
<tr>
<td>Kik et al. 2010 [21]</td>
<td>Cohort</td>
<td>433 adult close immigrant and BCG vaccinated Dutch-born contacts</td>
<td>The Netherlands (Low burden)</td>
<td>QFT-G-IT, T-SPOT.TB</td>
<td>N.A.</td>
<td>TST first, IGRA only if TST positive (at a cut-off ≥5 mm)*</td>
<td>22 (median)*</td>
<td>QFT-G-IT: 5/178 (2.8% [0.009–0.064]); T-SPOT.TB: 6/181 (3.3% [0.012–0.07])/ TST≥10mm: 9/288 (3.1% [0.014–0.058])*</td>
<td>IGRA: 3/3 (100%)/ TST: 6/6 (100%)*</td>
</tr>
</tbody>
</table>

N.A.: not assessed
*Among 44 QFT-G-IT-positive patients, seven were diagnosed to have active tuberculosis during the first investigation and one was lost within the first three weeks of follow-up.
†Conventional waiting period of 8 week was neither respected for IGRA(s) nor for the TST. TST was repeated 8-12 weeks later if negative initially, but not the IGRA if the IGRA result was negative initially.
‡Only 53.4% (95/178) of the QFT-G-IT positive contacts and only 55.8% (101/181) of the T-SPOT.TB positive contacts were followed for 21 months.
§As calculated by the authors although a TST was considered positive if induration diameter was at least 5 mm.
Six out of 9 patients developing tuberculosis were culture-confirmed, 2 of them were negative by both IGRA and in one tuberculosis patient, culture confirmed, IGRA testing was not performed.
<table>
<thead>
<tr>
<th>Study</th>
<th>Country and IGRA used</th>
<th>No. and type of participants</th>
<th>No. BCG vaccinees</th>
<th>No. IGRA positives/total tested</th>
<th>IGRA (OR, 95% CI)</th>
<th>No. TST positives/total tested</th>
<th>TST (OR, 95% CI)</th>
<th>Statistical agreement (κ, 95% CI) between the results of IGRA(s) and those of the TST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bartalesi et al. 2009 [50]</td>
<td>Italy; QFT-G-IT</td>
<td>LTB1 screening in 398 IMID subjects</td>
<td>16/393 (4.1%)</td>
<td>52/398</td>
<td>5.3 (1–28)</td>
<td>75/323 (23.2%)</td>
<td>6.5 (1.1–36.9)</td>
<td>κ (total)**: 0.55 (0.44–0.66)</td>
</tr>
<tr>
<td>Casas et al. 2009 [53]</td>
<td>Spain; T-SPOT.TB/ QFT-G-IT</td>
<td>147 HCWs</td>
<td>23/147 (15.6%)</td>
<td>T-SPOT.TB: 57/157; QFT-G-IT: 43/147</td>
<td>T-SPOT: 2.1 (1.5–4.1); QFT: 1.82 (0.88–3.8)</td>
<td>103/147 (70.1%)</td>
<td>OR n.s*.: 0.94 (0.46–1.93)</td>
<td>T-SPOT.TB: κ (total): 0.32 (SE 0.061); κ (BCG): 0.17 (SE 0.09); QFT-G-IT: κ (total): 0.29 (SE 0.052); κ (BCG): 0.085 (SE 0.05)</td>
</tr>
<tr>
<td>Hesseling et al. 2009 [61]</td>
<td>South Africa; T-SPOT.TB/ QFT-G-IT</td>
<td>82 close contacts (29 children and 53 adults)</td>
<td>62/82 (75.6%)</td>
<td>29/74</td>
<td>T-SPOT.TB: 38.40 (7.59–616.11); QFT-G-IT: 14.94 (4.02–55.58)</td>
<td>54/78 (69.2%)</td>
<td>3.83 (1.05–14.03)</td>
<td>T-SPOT.TB: κ (total): 0.12 (0.11–0.36); QFT-G-IT: κ (total): 0.45 (0.28–0.62)</td>
</tr>
<tr>
<td>Laffitte et al. 2009 [64]</td>
<td>Switzerland; T-SPOT.TB</td>
<td>LTB1 screening in 50 psoriasis patients before anti-TNF-alpha therapy</td>
<td>46/50 (92.8%)</td>
<td>10/50</td>
<td>Contact history: 5.67 (1.25–25.7); Lesions in CXR: 25.3 (2.41–267)</td>
<td>20/50 (40.0%)</td>
<td>OR n.s.: Cutoff &gt;5mm: 2.14 (0.55–8.3); &gt;10mm: 1.67 (0.43–6.5)</td>
<td>κ (total): 0.33 (± SD 0.13)</td>
</tr>
<tr>
<td>Lien et al. 2009 [65]</td>
<td>Vietnam; QFT-G-IT</td>
<td>300 HCWs</td>
<td>&gt; one third (data not shown)</td>
<td>142/300</td>
<td>1.94 (1.04–3.64)</td>
<td>176/288 (61.1%)</td>
<td>not calculated</td>
<td>κ (total): 0.44 (SE 0.06); κ (BCG): 0.29 (SE 0.09)</td>
</tr>
<tr>
<td>Matulis et al. 2008 [67]</td>
<td>Switzerland; QFT-G-IT</td>
<td>LTB1 screening in 142 rheumatic patients</td>
<td>118/142 (83.1%)</td>
<td>17/142</td>
<td>Contact history: 17.8 (2.06–154); Lesions in CXR: 66.8 (10.1–441);</td>
<td>46/115 (40.0%)</td>
<td>6.23 (1.18–33.01)</td>
<td>κ (total): 0.17 (0.02–0.32)</td>
</tr>
</tbody>
</table>
### Study Details

<table>
<thead>
<tr>
<th>Study</th>
<th>Country and IGRA used</th>
<th>No. and type of participants</th>
<th>No. BCG vaccinees</th>
<th>No. IGRA positives/total tested</th>
<th>IGRA (OR, 95% CI)</th>
<th>No. TST positives/total tested</th>
<th>TST (OR, 95% CI)</th>
<th>Statistical agreement (κ, 95% CI) between the results of IGRA(s) and those of the TST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seyhan et al. 2010 [72]</td>
<td>Turkey; QFT-G-IT</td>
<td>LTBI screening in 100 hemodialysis patients</td>
<td>72/100 (72.0%)</td>
<td>43/100</td>
<td>Contact history: 5.08 (1.2–21.2); Lesions in CXR: 3.06 (2.1–11.9)</td>
<td>34/100 (34.0%)</td>
<td>not calculated</td>
<td>κ (total): 0.26; κ (BCG): 0.15</td>
</tr>
<tr>
<td>Triverio et al. 2009 [74]</td>
<td>Switzerland: T-SPOT.TB; QFT-G-IT</td>
<td>LTBI screening in 62 ERDS patients</td>
<td>14/62 (22.6%)</td>
<td>T-SPOT.TB: 18/62; QFT-G-IT: 13/62</td>
<td>T-SPOT.TB: 1.2 (0.3–4.8); QFT-G-IT: 4.6 (1.2–18.1)</td>
<td>12/62 (19.4%)</td>
<td>OR n.s.: 1.4 (0.3–7.0)</td>
<td>QFT-G-IT: κ (total): 0.16; T-SPOT: κ (total): 0.32</td>
</tr>
<tr>
<td>Vinton et al. 2009 [77]</td>
<td>Australia; QFT-G-IT</td>
<td>LTBI screening in 481 hospital staff members from 5 hospitals</td>
<td>375/481 (78.0%)</td>
<td>32/381</td>
<td>5.6 (1.42–22.0)</td>
<td>120/364 (33.0%)</td>
<td>OR n.s.: 1.96 (0.68–5.63)</td>
<td>κ (total): 0.16</td>
</tr>
<tr>
<td>Zellweger et al. 2005 [48]</td>
<td>Switzerland; T-SPOT.TB</td>
<td>91 contact persons (residents/staff members) in an institution for alcoholic patients</td>
<td>78/91 (85.7%)</td>
<td>15/91</td>
<td>5.0 (1.05–23.86)</td>
<td>40/91 (44.0%)</td>
<td>OR n.s.: 1.85 (0.78–4.36)</td>
<td>κ (total): 0.232</td>
</tr>
</tbody>
</table>

*n.s.: not significant

** total: no division into subgroups
Figure legends

Figure 1: Flow diagram for study selection

432 potentially relevant citations identified by electronic databases and supplementary sources:
- 123 Specificity
- 34 Negative predictive value
- 144 Positive predictive value
- 131 TST/IGRA and exposure

334 full text studies screened for further eligibility criteria

60* studies finally included into review:
- 4 Specificity
- 18 Negative predictive value
- 4 Positive predictive value
- 34 TST/IGRA and TB exposure
*7 studies also belonging to another category

Excluded before full text screening
- Specificity: Reviews (9), Animal study (1), Cost analysis (7), Guideline (1)
- PPV: No TB investigated (42), Reviews (9), Animal TB infection (4), Cost analysis (1)
- TST/IGRA and exposure: Editorial (1), Cost analyses (2), Animal TB infection (6), Reviews (14), Guideline (1)

Excluded before in depth examination
- Specificity: Prior TB exposure/no low burden country (64), No commercial IGRA (29), Laboratory studies (4), Other reasons (4)
- NPV: No commercial IGRA (9), No differentiation between latent and active TB (3), Other reasons (4)
- PPV: Laboratory studies (7), No data on progression (71), Letters without original data (3), Other Reasons (3)
- TST/IGRA and TB exposure: Laboratory studies (7), Experimental study (1), No commercial IGRA (54), Only active TB cases (2), No differentiation between latent and active TB (3), Other reasons (6)

Figure 2: NPV of T-SPOT.TB in tuberculosis suspects
Figure 3: NPV of QFT-G-IT in tuberculosis suspects

Figure 4: NPV for progression in QFT-G-IT negative subjects
Figure 5: NPV for progression in T-SPOT.TB negative subjects

Clark et al. 2007: 1.00 (0.92 - 1.00) 47/47
Kik et al. 2010: 0.98 (0.94 - 1.00) 116/118
Lee et al. 2009: 0.88 (0.64 - 0.99) 15/17

Pooled NPV (T-Spot.TB) = 0.98 (0.94 to 0.99)
Chi-square = 5.88; df = 2 (p = 0.0533)
Inconsistency (I-square) = 65.9%