TITLE: An observational cohort study to evaluate the clinical utility of current and second-generation interferon-gamma release-assays in diagnostic evaluation of tuberculosis

AUTHORS

Hilary S Whitworth (PhD),[1,2] Amarjit Badhan (MSc),*[1,3] Aime A Boakye (MSc),*[1,3] Yemisi Takwoingi (PhD),*[4] Melanie Rees-Roberts (PhD),*[1,3,5] Christopher Partlett (PhD),[4] Heather Lambie (MSc Hons.),[1] Professor John Innes (FRCP),[6] Professor Graham Cooke (DPhil),[7] Marc Lipman (MD),[8,9] Professor Christopher Conlon (MD),[10] Professor Derek Macallan (PhD),[11] Felix Chua (PhD),[11] Professor Frank Post (PhD),[12] Professor Martin Wiselka (PhD),[13] Gerrit Woltmann (MD),[14] Professor Jonathan J Deeks (PhD),[4] Professor Onn Min Kon (MD)[#1,3,7] and Professor Ajit Lalvani (DM)[#1,3] on behalf of the IGRAs for Diagnostic Evaluation of Active TB (IDEA) Study Group.

[1]Tuberculosis Research Centre, National Heart and Lung Institute, Imperial College London, St Mary's Campus, Norfolk Place, London W2 1PG.

[2]Department of Clinical Research, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT.

[3]NIHR Health Protection Research Unit in Respiratory Infections, Imperial College London, St Mary's Campus, Norfolk Place, London W2 1PG.

[4]Institute of Applied Health Research, Public Health Building, University of Birmingham, Birmingham B15 2TT.

[5]Centre for Health Services Studies, University of Kent, Canterbury, CT2 7NZ.

[6]Heart of England NHS Foundation Trust. Bordesley Green East, Birmingham B9 5SS.

[7]St Mary’s Hospital, Imperial College Healthcare Trust, Praed Street, London W2 1NY.

[8]Royal Free London NHS Foundation Trust, Pond Street, London NW3 2QG.

[9]University College London (UCL) Respiratory, Division of Medicine, UCL, London.

[10]Oxford University Hospitals NHS Trust, Oxford, OX3 9DU.

[11]St George's Healthcare NHS Trust, Cranmer Terrace, London SW17 0RE.

[12]King's College Healthcare NHS Trust, Cutcombe Road, London SE9 5RJ.

[13] Department of Infection and Tropical Medicine. University Hospitals of Leicester NHS Trust. Leicester LE1 5WW.

[14]Respiratory Biomedical Research Centre, Institute for Lung Health, Department of Infection, Immunity and Inflammation, University of Leicester, Leicester LE3 9QP, UK.

*These four authors contributed equally to this work.

[#]These two authors contributed equally to this work.

Corresponding author: Professor Ajit Lalvani

Address - Director, Tuberculosis Research Centre, National Heart and Lung Institute, Imperial College London, St Mary's Campus, Norfolk Place, London W2 1PG.

Email - a.lalvani@imperial.ac.uk

Telephone - 0207 594 0883

RESEARCH IN CONTEXT

Evidence before study

Although the role of IGRAs in diagnosis of active TB is unclear, their use in clinical practice is common. A comprehensive systematic review and meta-analysis published in 2011 describes data from studies evaluating diagnostic accuracy of IGRAs in active TB up to November 2009. We therefore searched PubMed for original research studies published in any language between December 2009 and June 2018, using search terms for tuberculosis AND interferon gamma release assays, T-SPOT.TB or Quantiferon AND diagnosis, evaluation, rule-in or rule-out. The evidence-base to-date suggests that current IGRAs have insufficient specificity to rule in TB and insufficient sensitivity to rule out TB. However, this is derived primarily from studies that are either small, low quality, or not representative of patient populations seen in real-life clinical practice. Only one large prospective cohort study embedded in routine practice was identified, but in a high TB-incidence setting. Thus, fifteen years after introduction of IGRAs, the ability of policy-makers in low TB-incidence settings to generate recommendations and guidelines for the role of IGRAs in active TB is still hampered by a paucity of reliable and informative evidence.

Added value of this study

This is the largest prospective study specifically to define the role of IGRAs in diagnosis of active TB in a low TB incidence setting. Because the study was multicentre and embedded in routine clinical practice in England, and recruited patients representing the full natural clinical spectrum of TB, the results are generalisable to other high income, low incidence settings. By demonstrating that existing IGRAs have no useful role in diagnosis of active TB, it resolves a major clinical uncertainty and represents a significant new high-quality component of the evidence-base. Simultaneous evaluation of second-generation IGRA identifies this as a potentially useful high-sensitivity triage test that meets a major unmet clinical need.

Implications of all the available evidence

Results from this and previous studies can now be used to generate evidence-based national guidelines and recommendations for TB diagnosis. Specifically, neither T-SPOT.TB nor QFT-GIT have sufficient sensitivity or negative predictive value (NPV) to rule out a diagnosis of TB. Taken together with their low specificity and consequent inability to rule in a diagnosis of TB, existing IGRAs do not have a clinically useful role in the diagnostic work-up of TB. The finding that the second-generation IGRA may have sufficiently high sensitivity, low negative likelihood ratio and high NPV to serve as a triage test to help rule-out a diagnosis of TB within 24 hours indicates a clinically useful role for this novel test and provides the basis for evidence-based guidelines on its use in low incidence settings once it is widely available post-licensure.

ABSTRACT

Background

The role of interferon-gamma release assays (IGRAs) in diagnosis of active tuberculosis (TB) is unclear, yet they are commonly used in low-TB-incidence countries. This study sought to resolve this clinical uncertainty by determining the diagnostic accuracy and role of current and second-generation IGRAs in the diagnostic assessment of suspected TB in a low-incidence setting.

Methods

This was a prospective cohort study of 1,060 adults with suspected TB, conducted in routine secondary care in England. Patients were tested for *M. tuberculosis* (Mtb) infection at baseline using current and second-generation IGRAs, the latter incorporating novel Mtb antigens, and followed up for 6-12m to establish definitive diagnoses. Sensitivity, specificity and positive and negative likelihood ratios (LRs) and predictive values (PVs) of the tests for TB were determined.

Findings

TB was diagnosed in 363 (43%) of 845 patients included in analyses. Sensitivity of T-SPOT.TB was 81.4% (95%CI 76.6-85.3%), higher than Quantiferon-Gold In-Tube at 67.3% (95%CI 62.0-72.1%). Second-generation IGRA had higher sensitivity than current tests, at 94.0% (95%CI 90.0–96.4%) for culture-confirmed TB and 89.2% (95%CI 85.2–92.2%) when including highly-probable TB, giving a negative LR for all TB of 0.13 (95%CI 0.10-0.19). Specificity ranged from 86.2% (95%CI 82.3-89.4%) for T-SPOT.TB to 80.0% (95%CI 75.6-83.8%) for second-generation IGRA.

Interpretation

Currently-available IGRAs lack sufficient accuracy for diagnostic evaluation of suspected TB. Second-generation tests, however, may have sufficiently high sensitivity, low negative LR and correspondingly high negative PV in low-incidence settings to facilitate prompt rule-out of TB.

Funding

This study was funded by the National Institute for Health Research.

INTRODUCTION

Prompt diagnosis and treatment of active tuberculosis (TB) are essential for optimal patient outcomes and preventing onward transmission in the community and healthcare facilities.[1] However, diagnostic assessment of suspected TB can be lengthy, costly and burdensome for patients and healthcare systems,[2] often resulting in significant delays in diagnosis and treatment of other diseases in cases where suspected TB is eventually ruled out. Improving and accelerating diagnostic evaluation thus remains a clinical and public health priority in high-income, low-incidence countries, as well as high-burden regions. Recently, great advances in molecular diagnostics, such as GeneXpert (Cepheid Inc, Sunnyvale, CA, USA), have improved the speed and accuracy of microbiologic diagnosis and enabled prediction of antibiotic susceptibility.[3] However, whilst such tests have high specificity (which is important for 'rule-in'), they have insufficient sensitivity to rule out TB and require clinical specimens from anatomical disease sites, often requiring resource-intensive invasive procedures.[4] A blood test of high diagnostic sensitivity could help to promptly (e.g. in 24h) triage patients at clinical presentation (Appendix: Supplementary Panel, page 1); this would address a major unmet clinical need and has been prioritised by the World Health Organisation (WHO).[5] Given the paucibacillary nature of most cases of culture-negative TB, such a test would likely be based on measurement of immune responses to *M. tuberculosis* (Mtb) rather than direct detection of the bacteria or nucleic acids.

Interferon-gamma release-assays (IGRAs) are regulatory-approved immune-based blood tests for detecting Mtb infection. By measuring T-cell responses to two strongly immunogenic but highly specific Mtb antigens (ESAT-6 and CFP-10), they are not confounded by prior *Bacillus Calmette-Guérin* (BCG) vaccination and provide higher diagnostic specificity than the tuberculin skin test (TST).[6] Since Mtb infection is a pre-requisite for TB disease, a negative IGRA result could potentially rule-out a diagnosis of TB disease (i.e. exclude TB from the

differential diagnosis), though prior evidence suggests the sensitivity of current IGRAs may be insufficient to fulfil this triage function.[1,7–9]

Although established as the new standard-of-care for diagnosing latent TB infection (LTBI), IGRAs are currently not recommended in diagnosis of active TB other than in specific scenarios, such as paediatric TB, with caveats around interpretation and level of expertise required.[10,11] However, development of definitive recommendations has been hindered by a lack of robust and informative evidence. Most studies of diagnostic accuracy of IGRAs in active TB to date are retrospective reviews of hospital records and TB registry data or small-scale case-control studies, typically not representative of the heterogeneous patient population seen in real-life clinical practice. Although one large prospective cohort study embedded in routine practice and including head-to-head comparison of T-SPOT.TB and QFT-GIT was recently published, this was in a high TB-incidence setting.[12] Prospective cohort studies conducted in low-incidence settings have been substantially smaller.[1,8]

Given the shortfalls associated with available TB diagnostics, IGRAs continue to be used widely in clinical practice in the UK, albeit resulting in complexities and challenges in interpretation of results.[11] A large-scale prospective head-to-head comparison of diagnostic performance of IGRAs in routine practice is therefore required to conclusively define what, if any, clinical role they have in diagnosis of active TB, allowing development of evidence-based and authoritative recommendations in this setting.

Discovery of other highly-specific Mtb antigens as strongly immunogenic as ESAT-6 and CFP-10 presents an opportunity to develop second-generation IGRAs of higher sensitivity.[13,14] Furthermore, they may allow development of an ‘ESAT-6-free’ IGRA for application in populations vaccinated with new ESAT-6-based TB vaccines, as previously described.[15] Studies suggest adaptation of existing IGRAs with these novel antigens is possible,[1,14,16] but

no large-scale prospective clinical evaluation of this novel approach has been conducted in routine practice in a low TB-incidence setting.

We therefore sought to evaluate the clinical utility of existing IGRAs, T-SPOT.TB (Oxford Immunotec plc, Abingdon, UK) and Quantiferon-Gold In-tube (QFT-GIT; Qiagen NV), and second-generation IGRAs in patients presenting with suspected TB in UK clinical practice.

METHODS

We conducted a prospective, multicentre, cohort study in routine clinical practice to determine the diagnostic accuracy of commercially-available and second-generation IGRAs in active TB. A within-patient design was used to compare test accuracy by performing all IGRAs on blood samples from each study participant, with the presence or absence of active TB verified using a composite reference standard (Table 1).[1] This design minimises between-patient variability. The study was approved by Camden and Islington National Research Ethics Committee (11/H0722/8). The study protocol is available at https://njl-admin.nihr.ac.uk/document/download/2006627, and a STARD checklist is provided in the Appendix (Supplementary Checklist, pages 2-3).

Study participants

Adult inpatients and outpatients presenting with suspected active TB (based on signs and symptoms assessed by the attending hospital clinician) were consecutively enrolled from ten National Health Service (NHS) hospitals in five UK cities (London, Slough, Oxford, Leicester and Birmingham). Patients were enrolled at presentation to infectious disease and respiratory medicine secondary care services, before a final diagnosis was made, and a wide spectrum of pre-test probabilities for active TB were included. Exclusion criteria were limited to age <16y and inability/unwillingness to provide informed consent. Centres were selected to ensure the population was representative of ethnic mix and range of co-morbidities.

Participant enrolment and follow-up

Participants were first seen by research nurses at enrolment. Following consent, a baseline blood sample was drawn and data collected in a case report form on the demographics and medical history of the participant, and investigations performed in their routine diagnostic work-up. Participants were followed up two and six months thereafter with data collected on any subsequent investigations, test results and clinical diagnoses, and response to TB treatment if initiated. Patients with a definitive non-TB diagnosis who were discharged from routine care were not required to attend follow-up visits but, where necessary, data were collected from hospital records up to 12 months after enrolment to identify final diagnoses made by hospital clinicians.

Diagnosis and diagnostic categorisation

Participants were investigated in routine practice under the direction of the infectious disease or respiratory medicine attending physician. After completion of follow-up in this routine hospital setting, participants' final diagnoses were verified using a composite reference standard[1] by a panel of ≥4 respiratory medicine and infectious disease clinicians specialising in TB. The panel assessed anonymised clinical data (patient demographics, medical history, TB symptoms, previous TB information, TB exposure history, current medication, human immunodeficiency virus (HIV) status, relevant clinical correspondence, test results during diagnosis and follow-up, and any other relevant clinical information) whilst blinded to all IGRA results (including IGRAs carried out as part of routine practice at recruiting sites). Diagnoses of all participants were categorised into the following groups, as previously defined[1] (Table 1): definite TB (category 1); highly-probable TB (category 2); clinically indeterminate (category 3); and non-TB (category 4). Category 4 participants were sub-divided based on risk factors for LTBI (Table 1). Final diagnoses and diagnostic categories were determined by consensus across the panel.

Laboratory procedures

Blood samples (35ml) were collected into heparinised and QFT-GIT blood collection tubes from all participants at enrolment, before any diagnosis was made. QFT-GIT and T-SPOT.TB were carried out and interpreted in real-time at the TB Research Centre (Imperial College London) according to the manufacturer's instructions, and as described in Whitworth et al.[6] The second-generation IGRA used the T-SPOT.TB platform and incorporated ESAT-6, CFP-10 and Rv3615c; the 'ESAT-6-free' IGRA incorporated CFP-10, Rv3615c and Rv3879c. Further details on assay methods and interpretation of results are provided in the Appendix (Supplementary Methods, pages 4-5). Laboratory scientists performing study IGRAs were blinded to all clinical information, diagnoses and personal identifiers.

Statistical Analyses

The study was powered to detect a 10% difference in sensitivity between T-SPOT.TB and QFT-GIT, assuming a sensitivity of 85% for T-SPOT.TB and 75% for QFT-GIT.[1,7–9] Accounting for the paired nature of the data and assuming independence of errors,[16] 855 patients (after loss-to-follow-up (LTFU)/withdrawal and missing/excluded index/reference test results) were required to detect this difference at the 5% significance level (two-tailed) with 90% power, based on a predicted 40% prevalence of active TB in the study population. Sensitivity, specificity, positive and negative predictive values (PPV; NPV), and positive and negative likelihood ratios (PLR; NLR) for each test were calculated. Ninety-five percent confidence intervals (CIs) were calculated using the Wilson method for proportions[18,19] and the method by Simel *et al* for LRs.[20] All patients in diagnostic categories 1, 2 and 4 were included in analyses (Table 1); category 3 patients were reported but not included in analyses.

Patients with indeterminate IGRA or borderline TSPOT-TB results were excluded from primary analyses, but included as test-positives in sensitivity analyses. Sensitivity analyses were also conducted to investigate the impact of (1) excluding category 2 patients on IGRA

sensitivity and (2) excluding category 4A-C patients on IGRA specificity. To compare the accuracy of two IGRAs, we fitted separate generalized estimating equation (GEE) models for patients with and without active TB to estimate differences in sensitivity and specificity, respectively. This approach exploits the paired nature of the data whilst allowing use of all available data if test results were missing for either IGRA. We computed ratios of sensitivities (relative-sensitivity) and specificities (relative-specificity) from the GEE models using a post-estimation procedure with CIs computed using the delta method. Analyses were performed using Stata, version 13.0 (Stata, College Station, Texas).

Role of the funding source

The study funder, the National Institute for Health Research (NIHR), played no role in study design, data collection, analysis or interpretation, or writing of the report. The corresponding author had full access to all data in the study and had final responsibility for the decision to submit for publication.

RESULTS

Participant flow

Participant flow in the study is shown in Figure 1. Patients (n=1,060) with suspected active TB were consented and enrolled between 25 November 2011 and 31 August 2013. Those with a history of prior TB diagnosis (n=99) were excluded from analyses, as in previous studies.[12] Additionally, 116 patients were excluded for reasons provided in Figure 1, giving a final study population of 845 patients.

Demographic & clinical characteristics

Demographic and clinical characteristics for the final study population are shown in Table 2. The median age of the cohort was 38y (range 16-86y); 501/845 (59%) were male, and 412/845 (48%) were of Indian Subcontinent origin. One or more co-morbidities were reported in

427/845 (51%) participants (Table 2). Medications at presentation are shown in the Appendix (Supplementary Table 1, page 6). The most common symptoms reported at presentation were cough, weight-loss and lethargy (Appendix: Supplementary Table 2, page 7).

Diagnostic classification of patients

Among the study cohort, 363/845 patients (43%) had a final diagnosis of active TB (Table 1); 261/845 (31%) culture-confirmed (category 1), and 102/845 (12%) highly-probable (Category 2). Of all active TB cases (categories 1 and 2), 129/363 (36%) were pulmonary, 189/363 (52%) were extra-pulmonary and 45/363 (12%) were both (Table 3); most 154/363 (42%) had lymph node involvement. Of Mtb isolates undergoing drug-susceptibility testing, 21/261 (6%) were drug-resistant and one was multi-drug-resistant. TB was excluded (category 4) in 439/845 (52%) patients. These were sub-classified according to risk factors for LTBI or inactive TB into categories 4A-D in decreasing likelihood of having Mtb infection (Table 1).[1] Most common non-TB diagnoses are listed in Table 3. Only 43/845 patients (5.1%) were classified as clinically indeterminate (category 3).

Diagnostic accuracy of T-SPOT.TB and QFT-GIT

T-SPOT.TB and QFT-GIT results were available for 809/845 (96%) and 820/845 (97%) study participants, respectively; 805/845 (95%) patients had data for both IGRAs. Diagnostic sensitivity, specificity, PPV, NPV, PLR and NLR are shown in Table 4, with a cross-tabulation of T-SPOT.TB and QFT-GIT results in patients with active TB and non-TB diagnoses provided in the Appendix (Supplementary Table 3, page 8). Sensitivity of T-SPOT.TB was 84.9% (95%CI 79.5-89.0%) for culture-confirmed TB and 81.4% (95%CI 76.6-85.3%) for all TB, giving an NPV of 84.7% (95%CI 80.6-87.9%) and NLR of 0.22 (95%CI 0.17-0.27) for all TB. Specificity was 86.2% (95%CI 82.3-89.4%) for all non-TB patients and 93.5% (95%CI 86.6-97.0%) for cases with no risk factors for LTBI (category 4D). Sensitivity of QFT-GIT was 70.6% (95%CI 64.4-76.1%) for culture-confirmed TB and 67.3% (95%CI 62.0-72.1%) for all

TB, giving an NPV of 74.0% (95%CI 69.5-78.0%) and NLR of 0.41 (95%CI 0.35-0.48) for all TB. Specificity was 80.4% (95%CI 76.1-84.1%) for all non-TB patients and 93.4% (95%CI 86.4-96.9%) for cases with no risk factors for LTBI. Sensitivity and specificity of T-SPOT.TB were superior to QFT-GIT; relative sensitivity was 1.20 (95%CI 1.12-1.29) with p<0.0001, and relative specificity was 1.07 (95%CI 1.02-1.12) with p=0.004.

Diagnostic accuracy of second-generation and ESAT-6-free IGRA

Second-generation and ESAT-6-free IGRA results were available for 809/845 (96%) patients (Table 4). Sensitivity of second-generation IGRA was 94.0% (95%CI 90.0-96.4%) for culture-confirmed TB and 89.2% (95%CI 85.2-92.2%) for all TB, giving an NPV of 90.0% (95%CI 86.2-92.8%) and NLR of 0.13 (95%CI 0.10-0.19) for all TB. Specificity was 80.0% (95%CI 75.6-83.8%) for all non-TB patients and 91.3% (95%CI 83.8-95.5%) for cases with no risk factors for LTBI. Sensitivity of ESAT-free IGRA was 93.4% (95%CI 89.2-96.0%) for culture-confirmed TB and 88.0% (95%CI 83.8-91.2%) for all TB, giving an NPV of 89.2% (95%CI 85.4-92.1) and NLR of 0.15 (95%CI 0.11-0.21) for all TB. Specificity was 79.6% (95%CI 75.2-83.4%) for all non-TB patients and 90.3% (95%CI 82.6-94.8%) for cases with no risk factors for LTBI. Comparing second-generation IGRA with T-SPOT.TB, relative sensitivity was 1.08 (95%CI 1.04–1.11) with p<0.0001, and relative specificity was 0.94 (95%CI 0.91–0.96) with p<0.0001. For ESAT-6-free IGRA versus T-SPOT.TB, relative sensitivity was 1.07 (95%CI 1.03–1.10) with p=0.0002, and relative specificity was 0.93 (95%CI 0.90–0.96) with p<0.0001. A cross-tabulation of second-generation IGRA against T-SPOT.TB results and table of response magnitudes for each individual antigen are provided in the Appendix (Supplementary Tables 4 (page 9) and 5 (page 10) respectively).

Test performance in key patient subgroups

Of culture-confirmed TB cases with available smear microscopy results, 165/232 (71%) were smear-negative (57/165 with pulmonary TB, 80/165 with extra-pulmonary TB and 28/165 with

both). Sensitivities of T-SPOT.TB, QFT-GIT, second-generation IGRA and ESAT-6-free IGRA in this population were 85.9% (95%CI 79.2%-90.7%), 68.6% (95%CI 60.9%-75.4%), 93.8% (95%CI 88.5%-96.7%) and 92.9% (95%CI 87.4%-96.1%), respectively.

Among HIV-infected study participants, 25/135 (19%) had a final diagnosis of active TB and 108/135 (80%) had TB excluded; 27/88 (31%) diabetic participants had a final diagnosis of TB (Table 2). Sensitivity and specificity of all IGRAs for active TB in patients with HIV-infection and diabetes are shown in the Appendix (Supplementary Tables 6 (page 11) and 7 (page 12), respectively).

Indeterminate and borderline results

There was a trend towards a higher indeterminate rate for QFT-GIT (79/820; 9.6%) than T-SPOT.TB (57/809; 7.0%; p=0.07), and rates for QFT-GIT were higher than second-generation IGRA (55/809; 6.8%; p=0.04) and ESAT-6-free IGRA (55/809; 6.8%; p=0.04). Most indeterminate results occurred in non-TB patients (Appendix: Supplementary Tables 3 (page 8) and 4 (page 9)). T-SPOT.TB results were borderline in 17/345 (4.9%) patients with active TB and 16/423 (3.8%) with non-TB diagnoses. Lowering the cut-off of T-SPOT.TB from eight to five SFCs (thereby scoring all borderline results as positive) did not improve diagnostic performance of T-SPOT.TB or either of the second-generation IGRAs, giving only a marginal increase in sensitivity at the cost of a decrease in specificity (Supplementary Table 8; page 13). Scoring both indeterminate and borderline results as positives also did not affect test performance in sensitivity analyses (Table 4, footnote f).

DISCUSSION

This is the largest prospective cohort study embedded in real-life clinical practice to assess and compare the role of IGRAs in the evaluation of suspected pulmonary and extrapulmonary TB in a low TB-incidence setting. Although T-SPOT.TB had significantly higher sensitivity than

QFT-GIT, neither assay had sufficient sensitivity or NPV to rule out a diagnosis of active TB. In contrast, the second-generation IGRA, incorporating Rv3615c alongside ESAT-6 and CFP-10, had significantly higher diagnostic sensitivity than T-SPOT.TB and QFT-GIT. Interestingly, and reflecting common practice despite the absence of good evidence or guidelines supporting use of IGRAs in this setting, 35% of study patients, distributed across the recruiting sites, had IGRAs performed as part of their routine diagnostic work-up for active TB (data not shown).

The NLR of 0.13 for second-generation IGRA means a negative test result would reduce the odds of TB post-test by a clinically-meaningful factor of 7.7-fold compared to pre-test. The NPV for all TB, including highly-probable cases, was 90% despite the 43% prevalence in this population presenting to urban infectious diseases and respiratory medicine services with suspected TB. Since our study was performed in routine clinical practice and encompassed the full, natural clinical spectrum of TB and non-TB diagnoses, the results are likely generalizable across clinical practice in high-income, low-incidence countries. Accordingly, in clinical settings with a low-to-moderate pre-test probability of TB, such as general medical inpatient and outpatient services or primary care, second-generation IGRA has sufficiently low NLR to almost rule out TB. For example, a negative test result would convert pre-test probabilities of 20% and 10% to post-test probabilities of 3.1% and 1.4%, respectively. This would provide a useful prompt triage of patients on initial presentation, similar to the role played by other diagnostic tests of high sensitivity and limited specificity, such as serum D-dimer to triage patients with low-to-moderate suspicion of venous thromboembolism.[21] To our knowledge, other currently-available tests for TB lack required diagnostic sensitivity to fulfil this role. Although Xpert MTB/RIF Ultra has shown diagnostic sensitivity of 88%, its sensitivity in smear-negative, culture-positive TB is only 63%[3] (and sensitivity of Xpert even lower[4]), compared to 93.8% (CI 88.6%-96.7%) for second-generation IGRA in this diagnostically

challenging subgroup who frequently have paucibacillary disease. However, the very high specificity of molecular tests such as Xpert provides high PPV, enabling rule-in of active TB. Second-generation IGRA may thus play a complementary role to rapid molecular tests in the diagnostic work-up of suspected TB.

Given that IGRAs are the standard-of-care for detecting LTBI,[10,11] they will inevitably identify LTBI in cases where active TB has been excluded. Because most people with possible TB in low-burden countries are from ethnic groups with a high prevalence of LTBI,[22] as in our study, the diagnostic specificity for active TB is low for all IGRAs, and would be lower still in high-burden countries. The enhanced diagnostic sensitivity of the second-generation IGRA was accompanied by only a modest reduction in specificity to 80%, similar to QFT-GIT. Our study confirms that the low specificity and PLR of current and second-generation IGRAs mean that a positive result cannot rule in a diagnosis of TB. Interestingly, the specificity of all IGRAs increased to 90-93% in patients with active TB excluded and no risk factors for LTBI (Category 4D). Thus, a positive IGRA result may help to keep a diagnosis of active TB in the differential diagnosis in populations with a very low prevalence of LTBI, which however is not usually the case in patient populations being assessed for possible TB.

Two of the leading new TB vaccine candidates, Hybrid 1-IC31[23] and H56:IC31,[24] contain ESAT-6 and may induce conversion of IGRA results in vaccinated individuals. If these vaccines show protective efficacy in ongoing clinical trials and achieve licensure, ESAT-6-containing IGRAs will give false-positive results in vaccinated persons who are not Mtb-infected, analogous to false-positive TST results in Mtb-uninfected persons with prior BCG vaccination. Diagnostic accuracy of ESAT-6-free IGRA was very similar to second-generation IGRA and thus has potential to replace other IGRAs in populations immunised against TB with ESAT-6-based vaccines.

Two of the most important global risk factors for TB are HIV co-infection[25] and diabetes,[26] both of which have been reported to adversely affect IGRA performance.[27,28] Performance of current IGRAs in patients with HIV-infection and diabetes in this study was insufficient to be of value in the diagnosis of active TB. Performance appeared to be lower in HIV-infected and diabetic subgroups, but the small numbers of patients with TB in these subgroups precluded statistical comparisons. This was also the case for other types of immunosuppression associated with TB, such as chronic kidney disease and immunosuppressive medication.

Strengths of our study include the rigorous case definitions, including six-months follow-up to confirm that a diagnosis of TB was excluded where a non-TB diagnosis was not made at presentation. For highly-probable TB, we used a composite reference standard[1] that was applied by a panel of expert and experienced clinicians, blinded to IGRA results. Despite this stringent case definition, it is likely that a proportion of patients without TB were incorrectly categorised as having highly-probable TB, which would explain why all IGRAs had lower sensitivity for highly-probable TB than for all TB, which includes culture-confirmed cases. Thus, our estimates of diagnostic sensitivity for all TB, which includes highly-probable TB, are likely conservative. This highlights the significance of increased IGRA sensitivity in culture-confirmed TB (and the importance of including this sub-group in study analyses) as this is the only population in whom TB diagnoses are definitive.

Our study has some limitations. First, it does not include children, in whom the unmet clinical need for improved diagnosis of TB is high. Second, the numbers of patients with risk-factors associated with immunosuppression that do (e.g. HIV-infection) or might (e.g. diabetes) affect test performance were modest, precluding clear conclusions about test performance in these subpopulations. Third, whilst blood collection and assays were performed strictly in accordance with manufacturers' instructions, IGRAs were not performed in a routine

diagnostic service laboratory, and re-testing of new samples was not performed in cases where initial results were indeterminate or borderline (as recommended by manufacturers).

Although the QFT-GIT has been replaced by the QFT-GIT-Plus since our study was conducted, its diagnostic accuracy does not appear to be significantly better than QFT-GIT and there is no evidence it is as sensitive as T-SPOT.TB.[29,30] Therefore, our conclusion that neither existing IGRA has a clinically useful role in the evaluation of suspected active TB is unaffected by availability of QFT-GIT-Plus.

In conclusion, our study provides conclusive and generalizable evidence that existing IGRAs do not have a useful role as rule-in or rule-out tests in routine clinical practice. However, second-generation IGRAs have higher sensitivity and NPV which may help to rule out a diagnosis of TB in clinical settings with a low-to-moderate prevalence of TB.

CONTRIBUTORS

HSW was responsible for day-to-day management of the IDEA study, including oversight of clinical and laboratory data collection and management. AB managed participant recruitment, follow-up activities and clinical data collection, and contributed to data management. AAB contributed to laboratory data collection, data management and quality assurance. YT led statistical analyses and producing data tables and figures, and contributed to data interpretation. MRR led the study set-up and initial management, and built the study databases. CP contributed to statistical analyses and producing data tables and figures. HL contributed to laboratory data collection and managed the laboratory database. JI led the expert clinical panel. GC, ML, CC, DM, FC, FP, MW and GW contributed to patient recruitment, data collection and the study expert diagnostic clinical panel. JD contributed to study design, data analyses and interpretation of results. OMK and AL co-led study conceptualisation, design, oversight and interpretation of

results. Writing of the manuscript was co-led by HSW and AL, and all authors contributed to its drafting and revision.

STUDY GROUP MEMBERS

David Abdoyeku, Northwest London Hospitals NHS Trust

Robert Davidson, Northwest London Hospitals NHS Trust

Martin Dedicoat, University Hospitals Birmingham Foundation Trust

Heinke Kunst, Queen Mary University Blizard Institute

Michael R Loebingher, Chelsea and Westminster Hospital NHS Foundation Trust

William Lynn, Ealing NHS Trust

Nazim Nathani, Sandwell and West Birmingham Hospitals NHS Trust

Rebecca O' Connell, Bart's Healthcare NHS Trust

Anton Pozniak, Chelsea and Westminster Hospital NHS Foundation Trust

Sarah Menzies, Wexham Park Hospital, Frimley Health NHS Foundation Trust

DECLARATION OF INTERESTS

HSW, AB, AAB, MRR and HL were employed by Imperial College London using grant funding from NIHR to conduct the work described in this paper. JD, YT and CP, whilst working for the University of Birmingham, received grant funding from NIHR to conduct the work described in this paper. The funding contributed to their salary costs. JI, GC, ML, CC, DM, FC, FP, MW and GW have no conflicts of interest to declare. JD reports grants from NIHR during the conduct of the study outside the submitted work. OMK is employed by Imperial Healthcare Trust and was partially paid by the NIHR grant from Imperial College London. OMK received other grants from NIHR during the conduct of the study and has received speaker fees from Oxford Immunotec. He chairs a non-remunerated independent

committee that organizes an annual educational symposium on tuberculosis, sponsored by Qiagen. AL is named inventor on patents pertaining to T cell-based diagnosis, including current and second-generation IGRA technologies. Some of these patents were assigned by the University of Oxford to Oxford Immunotec plc, resulting in royalty entitlements for the University of Oxford and AL.

FUNDING AND ACKNOWLEDGEMENTS

The above study was funded by the NIHR HTA programme (grant number: 08/106/02). AL, OMK and GC are supported in part by the NIHR Imperial BRC HPRU. We would like to thank the Research Nurses at each of the study sites for their work in recruitment of participants, and the Research Assistants at the TB Research Centre, in particular Ms Bianca Bartlett, Dr Ann-Kathrin Reuschl and Dr Luigi Marongiu, who contributed to processing of participant samples and performing the study assays. We additionally would like to thank the members of the Study Steering Committee, Professor Khalid Khan, Professor Stephen Gordon, Dr James Gray, Dr Johannes Reitsma and Ms Nisha Karnani (lay member) for their support and advice throughout the study.

REFERENCES

1. Dosanjh DP, Hinks TS, Innes JA, Deeks JJ, Pasvol G, Hackforth S, *et al.* Improved diagnostic evaluation of suspected tuberculosis. *Ann Intern Med* 2008;**148**:325-36.
2. Whitworth HS, Aranday-Cortes, Lalvani A. Biomarkers of tuberculosis: A research roadmap. *Biomark Med* 2013;**7**(3):349-62.
3. Dorman SE, Schumacher SG, Alland D, Nabeta P, Armstrong DT, King B *et al.* Xpert MTB/RIF Ultra for detection of Mycobacterium tuberculosis and rifampicin resistance:

a prospective multicentre diagnostic accuracy study. *Lancet Infect Dis* 2018;**18**(1):76-84.

4. Sohn H, Aero AD, Menzies D, Behr M, Schwartzman K, Alvarez GG *et al*. Xpert MTB/RIF testing in a low tuberculosis incidence, high-resource setting: limitations in accuracy and clinical impact. *Clin Infect Dis*. 2014;**58**(7):970-6.
5. García-Basterio AL and Cobelens F. Triage tests: a new priority for tuberculosis diagnostics. *Lancet* 2015;**3**(3):177-178.
6. Whitworth HS, Scott M, Connell DW, Donges B, Lalvani A. IGRAs–the gateway to T cell based TB diagnosis. *Methods* 2013;**61**:52-62.
7. 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.
8. Goletti D, Carrara S, Butera O, Amicosante M, Ernst M, Sauzullo I, *et al.* Accuracy of immunodiagnostic tests for active tuberculosis using single and combined results: a multicenter TBNET-Study. *PLoS One* 2008;**3**:e3417.
9. Sester M, Sotgiu G, Lange C, Giehl C, Girardi E, Migliori GB et al. Interferon-γ release assays for the diagnosis of active tuberculosis: a systematic review and meta-analysis. Eur Respir J. 2011;37(1):100-11.
10. UK ICGT. Tuberculosis: Prevention, diagnosis, management and service organisation. *National Institute for Health and Care Excellence (UK) NICE guidelines [NG33]* 2016.
11. European Centre for Disease Prevention and Control. Use of interferon-gamma release assays in support of TB diagnosis. Stockholm: ECDC; 2011.
12. Du F, Xie L, Zhang Y, Gao F, Zhang H, Chen W *et al*. Prospective comparison of QFT-GIT and T-SPOT.TB assays for diagnosis of active Tuberculosis. *Sci Rep* 2018;**8**(1):5882.

13. Liu XQ, Dosanjh D, Varia H, Ewer K, Cockle P, Pasvol G *et al*. Evaluation of T cell responses to novel RD1 and RD2 encoded *Mycobacterium tuberculosis* gene products for specific detection of human tuberculosis infection. *Infect Immun* 2004;**72**:2574-81.
14. Millington KA, Fortune SM, Low J, Garces A, Hingley-Wilson SM, Wickremasinghe M *et al*. Rv3615c is a highly immunodominant RD1 (Region of Difference 1)-dependent secreted antigen specific for Mycobacterium tuberculosis infection. *Proc Natl Acad Sci USA* 2011;**108**(14):5730-5.
15. Ruhwald M, de Thurah L, Kuchaka D, Zaher MR, Salman AM, Abdel-Ghaffer AR *et al*. Introducing the ESAT-6 free IGRA, a companion diagnostic for TB vaccines based on ESAT-6. *Sci Rep* 2017;**7**:45969.
16. Li G, Li F, Zhao HM, Wen HL, Li HC, Li CL *et al*. Evaluation of a new IFN-γ release assay for rapid diagnosis of active tuberculosis in a high incidence setting. *Front Cell Infect Microbiol* 2017;**7**:117.
17. Alonzo TA, Pepe MS, Moskowitz CS. Sample size calculations for comparative studies of medical tests for detecting presence of disease. *Stat Med* 2002;**21**:835-52.
18. Wilson EB. Probable inference, the law of succession, and statistical inference. *J Amer Stat Assoc* 1927;**22**:209-12.
19. Brown LD, Cai TT, DasGupta A. Interval Estimation for a Binomial Proportion. *Stat Sci* 2001;**16**:101-17.
20. Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. *J Clin Epidemiol* 1991;**44**:763-70.
21. Van der Hulle T, Cheung WY, Kooij S, Beenen LFM, van Bemmel T, van Es J *et al*. Simplified diagnostic management of suspected pulmonary embolism (the YEARS study): a prospective, multicentre cohort study. *Lancet* 2017;390(10091):289-297.

22. Pareek M, Greenaway C, Noori T, Munoz J, Zenner D. The impact of migration on tuberculosis epidemiology and control in high-income countries: a review. *BMC Med* 2016;**14**;48.

23. Mearns H, Geldenhuys HD, Kagina BM, Musvosvi M, Little F, Ratangee F *et al*. H1:IC31 vaccination is safe and induces long-lived TNF-α^{+}IL-2^{+}CD4 T cell responses in M. tuberculosis infected and uninfected adolescents: A randomised trial. *Vaccine* 2017;**35**(1):132-141.

24. Luabeya AK, Kagina BM, Tameris MD, Geldenhuys H, Hoff ST, Shi Z *et al.* First-in-human trial of the post-exposure tuberculosis vaccine H56:IC31 in Mycobacterium tuberculosis infected and non-infected healthy adults. *Vaccine* 2015;**33**(33):4130-40.

25. Esmail H, Riou C, Bruyn ED, Lai RP, Harley YXR, Meintjes G *et al*. The immune response to Mycobacterium tuberculosis in HIV-1-coinfected persons. *Annu Rev Immunol* 2018;**36:**603-638.

26. Kuo MC, Lin SH, Lin CH, Mao IC, Chang SJ, Hsieh MC. Type 2 diabetes: an independent risk factor for tuberculosis: a nationwide population-based study. *PLoS One* 2013;**8**(11):e78924.

27. Cattamanchi A, Smith R, Steingart KR, Metcalfe JZ, Date A, Coleman C 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(3):230-8.

28. Faurholt-Jepson D, Aabye MG, Jenson AV, Rangge N, PrayGod G, Jeremiah K *et al*. Diabetes is associated with lower tuberculosis antigen-specific interferon gamma release in Tanzanian tuberculosis patients and non-tuberculosis controls. *Scand J Infect Dis*. 2014;**46**:384-391.

29. Takasaki J, Manabe T, Morino E, Muto Y, Hashimoto M, Iikura M et al. Sensitivity and specificity of QuantiFERON-TB Gold Plus compared to QuantiFERON-TB Gold In-Tube and T-SPOT.TB on active tuberculosis in Japan. J Infect Chemother. 2018;24(3):188-192.
30. Barcellini L, Borroni E, Brown J, Brunetti E, Codecasa L, Cugnata F, et al. First independent evaluation of QuantiFERON-TB Plus performance. Eur Respir J. 2016;47:1587-90.

Table 1: Pre-defined criteria for case definitions and diagnostic categories.[1]

Diagnostic category	Criteria	Number of Patients
1: Culture-confirmed TB[a]	Microbiological culture of M. tuberculosis, AND suggestive clinical and radiological findings.	261
2: Highly-probable TB[a]	Clinical and radiological features highly suggestive of TB unlikely to be caused by other disease, AND a decision to treat made by a clinician, AND appropriate response to therapy, AND histology supportive if available.	102
3: Clinically indeterminate	Final diagnosis of TB neither highly-probable, nor reliably excluded.	43
4: Active TB excluded		
Sub-classification		
4A: Inactive TB	Stable CXR changes, AND TST positive[b] (if done), AND bacteriologically negative (if done), AND no clinical evidence of active disease.	7
4B: One or more risk factors for TB exposure[c], TST positive[b]	TST positive[b], AND bacteriologically negative (if done) AND no clinical evidence of active disease.	48
4C: One or more risk factors for TB exposure[c], TST negative	History of TB exposure, AND TST negative (if done).	267
4D: No risk factors for TB exposure[c], TST negative	No history of TB exposure, AND TST negative (if done)	117
Total		**845**

CXR, chest radiograph; TB, tuberculosis; TST, tuberculin skin test.

[a]Mtb culture is the gold standard test for diagnosis of active TB. However, given that even culture does not detect all TB cases, our previously-validated reference standard includes a second category for culture-negative but highly-probable active TB diagnoses, made based on other available evidence.[1]

[b]TST using Mantoux test with threshold ≥15mm considered positive

[c]Risk factors for TB exposure: recent exposure to active TB patient; born in country of high prevalence; or belonging to an ethnic group with a very high prevalence of TB (incidence >100/100,000).

638 **Table 2: Demographics and clinical characteristics.** Column percentages for each characteristic are shown.

Characteristic	Diagnosis as per Reference Standard[1]				Total
	Culture-confirmed TB N=261	Highly-probable TB N=102	Clinically indeterminate N=43	Active TB excluded N=439	N=845
Clinical setting, n (%)					
Outpatient	171 (65.5)	72 (70.6)	32 (74.4)	269 (61.3)	544 (64.4)
Inpatient	90 (34.5)	30 (29.4)	11 (25.6)	170 (38.7)	301 (35.6)
Median age (range), years	32 (16–81)	36 (18–76)	38 (16–79)	44 (17–86)	38 (16–86)
Male, n (%)	177 (67.8)	53 (52.0)	21 (48.8)	250 (56.9)	501 (59.3)
Ethnic origin, n (%)					
Indian Subcontinent	167 (64.0)	61 (59.8)	16 (37.2)	168 (38.3)	412 (48.8)
Black	50 (19.2)	22 (21.6)	10 (23.3)	102 (23.2)	184 (21.8)
White	22 (8.4)	9 (8.8)	12 (27.9)	126 (28.7)	169 (20.0)
Asian	16 (6.1)	6 (5.9)	5 (11.6)	14 (3.2)	41 (4.9)
Middle Eastern	4 (1.5)	0	0	12 (2.7)	16 (1.9)
Mixed	1 (0.4)	4 (3.9)	0	8 (1.8)	13 (1.5)
Hispanic	1 (0.4)	0	0	7 (1.6)	8 (0.9)
Unknown	0	0	0	2 (0.5)	2 (0.2)
Median years in UK (range)	4.9 (0.1–52.9)	6.1 (0.3–59.7)	10.5 (0.4–56.9)	13.2 (0.0–60.3)	8.3 (0.0–60.3)
Profession, n (%)[a]					
Paid employment	130 (49.8)	52 (51.0)	21 (48.8)	214 (48.7)	417 (49.4)
Unemployed	62 (23.8)	24 (23.5)	16 (37.2)	164 (37.4)	266 (31.5)
Student	50 (19.2)	13 (12.8)	3 (7.0)	26 (5.9)	92 (10.9)
Healthcare/laboratory worker	16 (6.1)	9 (8.8)	2 (4.7)	24 (5.5)	51 (6.0)
Social/prison worker	1 (0.4)	1 (1.0)	0	2 (0.5)	4 (0.5)
Sex worker	0	1 (1.0)	0	2 (0.5)	3 (0.4)
Unknown	2 (0.8)	2 (2.0)	1 (2.3)	7 (1.6)	12 (1.4)
Median height (range), m	1.7 (1.4–2.0)	1.7 (1.5–1.9)	1.6 (1.5–1.8)	1.7 (1.3–2.0)	1.7 (1.3–2.0)
Median weight (range), kg	63 (35–127)	64 (40–116)	71 (37–110)	68 (38–157)	65 (35–157)
Median BMI (range)	22 (14–48)	22 (16–42)	24 (13–45)	24 (15–47)	23 (13–48)
BCG vaccinated, n (%)	194 (74.3)	79 (77.5)	36 (83.7)	340 (77.4)	649 (76.8)
BCG scar visible, n (%)					
Yes	172 (65.9)	72 (70.6)	29 (67.4)	283 (64.5)	556 (65.8)
No	12 (4.6)	3 (2.9)	3 (7.0)	19 (4.3)	37 (4.4)
Unknown	16 (6.1)	8 (7.8)	6 (14.0)	44 (10.0)	74 (8.8)

Missing	61 (23.4)	19 (18.6)	5 (11.6)	93 (21.2)	178 (21.1)
Recent known TB contact, n (%)	70 (26.8)	25 (24.5)	12 (27.9)	83 (18.9)	190 (22.5)
Other pre-existing conditions/co-morbidities, n (%)[b]					
None	169 (64.8)	61 (59.8)	19 (44.2)	169 (38.5)	418 (49.5)
HIV-infected	13 (5.0)	12 (11.8)	2 (4.7)	108 (24.6)	135 (16.0)
Diabetes	22 (8.4)	5 (4.9)	8 (18.6)	53 (12.1)	88 (10.4)
Asthma	12 (4.6)	5 (4.9)	4 (9.3)	50 (11.4)	71 (8.4)
Cancer	1 (0.4)	1 (1.0)	0	12 (2.7)	14 (1.7)
Chronic/end stage kidney disease	5 (1.9)	1 (1.0)	2 (4.7)	4 (0.9)	12 (1.4)
Hepatitis C	1 (0.4)	1 (1.0)	0	10 (2.3)	12 (1.4)
Hepatitis B	5 (1.9)	1 (1.0)	0	5 (1.1)	11 (1.3)
Organ transplantation	0	0	0	2 (0.5)	2 (0.2)
Sarcoidosis	1 (0.4)	0	0	0	1 (0.1)
Other	74 (28.4)	37 (36.3)	20 (46.5)	228 (51.9)	359 (42.5)

BMI, body mass index

[a]Some patients had more than one profession.

[b]Some patients had multiple co-morbidities.

Table 3: Final diagnoses of patients with and without active TB

Confirmed or highly probably TB	n (%)	**Active tuberculosis excluded**[b]	n (%)
N = 363		**N = 439**	
All TB	363 (100)	Pneumonia	104 (23.7)
Pulmonary	129 (35.5)	Sarcoidosis	38 (8.7)
Extrapulmonary	189 (52.1)	Cancer	36 (8.2)
Pulmonary and extrapulmonary	45 (12.4)	Lower respiratory tract infection	23 (5.2)
Site of disease[a]		Reactive lymphadenopathy	18 (4.1)
Lungs	174 (47.9)	Chest Infection	16 (3.6)
Lymph node	154 (42.4)	Exacerbation of asthma	14 (3.2)
Pleura	26 (7.2)	Upper respiratory tract infection	13 (3.0)
Spine	16 (4.4)	Non-tuberculosis mycobacterium infection	12 (2.7)
Miliary TB (disseminated)	11 (3.0)	Exacerbation of bronchiectasis	11 (2.5)
Abdomen	9 (2.5)	Exacerbation of COPD	8 (1.8)
Pericardium	6 (1.7)	Other[c]	158 (36.0)
Brain	6 (1.7)		
Musculoskeletal	5 (1.4)		
Chest wall	2 (0.6)		
Other	31 (8.5)		

COPD, Chronic obstructive pulmonary disease

[a]Some patients had TB at multiple anatomical sites.

[b]Some patients had multiple diagnoses.

[c]Less than five cases per diagnosis.

Table 4: Diagnostic accuracy of current and second-generation IGRAs for diagnosis of active TB. Sensitivity, specificity and predictive values are presented as percentages.

Test performance	T-SPOT.TB[a,e,f]		QFT-GIT[a,e,f]		ESAT+ CFP10 + Rv3615c[a,e,f]		CFP10 + Rv3615c + Rv3879c[a,e,f]	
	n/N	Estimate (95%CI)	n/N	Estimate (95%CI)	n/N	Estimate (95%CI)	n/N	Estimate (95%CI)
Sensitivity for active TB								
All TB	253/311	81.4 (76.6–85.3)	220/327	67.3 (62.0–72.1)	273/306	89.2 (85.2–92.2)	263/299	88.0 (83.8–91.2)
Culture-confirmed TB	185/218	84.9 (79.5–89.0)	163/231	70.6 (64.4–76.1)	203/216	94.0 (90.0–96.4)	197/211	93.4 (89.2–96.0)
Highly-probable TB[b]	68/93	73.1 (63.3–81.1)	57/96	59.4 (49.4–68.7)	70/90	77.8 (68.2–85.1)	66/88	75.0 (65.0–82.9)
Smear-positive TB[c]	45/55	81.8 (69.7–89.8)	42/56	75.0 (62.3–84.5)	48/51	94.1 (84.1–98.0)	47/50	94.0 (83.8–97.9)
Smear-negative TB[c,d]	169/206	82.0 (76.2–86.7)	148/222	66.7 (60.2–72.5)	183/207	88.4 (83.3–92.1)	176/202	87.1 (81.8–91.1)
Pulmonary TB	79/105	75.2 (66.2–82.5)	79/115	68.7 (59.7–76.5)	88/100	88.0 (80.2–93.0)	85/97	87.6 (79.6–92.8)
Extra-pulmonary TB	141/169	83.4 (77.1–88.3)	113/171	66.1 (58.7–72.8)	148/167	88.6 (82.9–92.6)	142/164	86.6 (80.5–91.0)
Specificity for active TB								
Active TB excluded	319/370	86.2 (82.3–89.4)	304/378	80.4 (76.1–84.1)	296/370	80.0 (75.6–83.8)	296/372	79.6 (75.2–83.4)
Active TB excluded, TST-negative, no risk factors for LTBI	87/93	93.5 (86.6–97.0)	85/91	93.4 (86.4–96.9)	84/92	91.3 (83.8–95.5)	84/93	90.3 (82.6–94.8)
Predictive values for all TB								
Positive predictive value	253/304	83.2 (78.6–87.0)	220/294	74.8 (69.6–79.5)	273/347	78.7(74.1–82.7)	263/339	77.6 (72.8–81.7)
Negative predictive value	319/377	84.6 (80.6–87.9)	304/411	74.0 (69.5–78.0)	296/329	90.0 (86.2–92.8)	296/332	89.2 (85.4–92.1)
Likelihood ratios for all TB								
Positive likelihood ratio		5.90 (4.55–7.66)		3.44 (2.76–4.27)		4.46 (3.62–5.49)		4.31 (3.51–5.28)
Negative likelihood ratio		0.22 (0.17–0.27)		0.41 (0.35–0.48)		0.13 (0.10–0.19)		0.15 (0.11–0.21)

LTBI, latent tuberculosis infection; TST, tuberculin skin test.

[a]25/845 QFT-GIT and 36/845 T-SPOT.TB and second-generation IGRA results were missing due to blood draw difficulties, samples being unsuitable for testing, or samples being destroyed for laboratory reasons. Missing results were spread across all diagnostic categories.

[b]'Highly-probable' TB includes culture-negative TB cases plus 10 patients with a final diagnosis of TB who did not have Mtb culture performed. Sensitivity (95%CI) results for culture-negative TB alone were as follows: T-SPOT.TB – 69.9% (59.3–78.7); QFT-GIT – 57.1% (46.5–67.2); second-generation IGRA (ESAT-6, CFP-10, Rv3615c) – 75.0% (64.5–83.2); ESAT-6-free IGRA (CFP-10, Rv3615c, Rv3879c) – 73.1% (62.3–81.7).

[c]56/845 participants did not undergo smear microscopy.

[d]Among 165 patients who were smear-negative but culture-positive, 122/142 were T-SPOT.TB-positive; 105/153 were QFT-GIT-positive; 135/144 were positive in second-generation IGRA and 131/141 were positive in ESAT-6-free IGRA.

[e]Indeterminate and borderline IGRA results were excluded from the analysis and thus also from data presented in this table. Numbers of indeterminate and borderline results for T-SPOT.TB/QFT-GIT and second-generation IGRA are presented in the Appendix (Supplementary Tables 3 (page5) and 4 (page 6), respectively).

[f]When indeterminate and borderline results were included as test positives in sensitivity analyses (positive on the basis that such a result could not exclude a TB diagnosis), sensitivity (95%CI) results for all TB were as follows: T-SPOT.TB – 83.2% (78.9-86.8); QFT-GIT – 69.7% (64.7–74.2); second-generation IGRA (ESAT-6, CFP-10, Rv3615c) – 90.4% (86.9–93.1); ESAT-6-free IGRA (CFP-10, Rv3615c, Rv3879c) – 89.6% (85.9–92.4).