Clinical utility of existing and second-generation interferon-γ release assays for diagnostic evaluation of tuberculosis: an observational cohort study

Hilary S Whitworth, Amarjit Badhan, Aime A Boakye*, Yemisi Takwoingi*, Melanie Rees-Roberts*, Christopher Partlett, Heather Lambie, John Innes, Graham Cooke, Marc Lipman, Christopher Conlon, Derek Macallan, Felix Chua, Frank A Post, Martin Wiselka, Gerrit Woltmann, Jonathan J Deeks, Onn Min Kon†, Ajit Lalvani†, on behalf of the Interferon-γ Release Assays for Diagnostic Evaluation of Active Tuberculosis study group‡*

Summary

Background **The clinical utility of interferon-γ release assays (IGRAs) for diagnosis of active tuberculosis is unclear, although they are commonly used in countries with a low incidence of tuberculosis. We aimed to resolve this clinical uncertainty by determining the accuracy and utility of commercially available and second-generation IGRAs in the diagnostic assessment of suspected tuberculosis in a low-incidence setting.**

Methods **We did a prospective cohort study of adults with suspected tuberculosis in routine secondary care in England. Patients were tested for *Mycobacterium tuberculosis* infection at baseline with commercially available (T-SPOT.*TB* and QuantiFERON-TB Gold In-Tube [QFT-GIT]) and second-generation (incorporating novel *M tuberculosis* antigens) IGRAs and followed up for 6–12 months to establish definitive diagnoses. Sensitivity, specificity, positive and negative likelihood ratios, and predictive values of the tests were determined.**

Findings **Of the 1060 adults enrolled in the study, 845 were included in the analyses and 363 were diagnosed with tuberculosis. Sensitivity of T-SPOT.*TB* for all tuberculosis diagnosis, including culture-confirmed and highly probable cases, was 81·4% (95% CI 76·6–85·3), which was higher than QFT-GIT (67·3% [62·0–72·1]). Second-generation IGRAs had a sensitivity of 94·0% (90·0–96·4) for culture-confirmed tuberculosis and 89·2% (85·2–92·2) when including highly probable tuberculosis, giving a negative likelihood ratio for all tuberculosis cases 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% (75·6–83·8) for second-generation IGRAs.**

Interpretation **Commercially available IGRAs do not have sufficient accuracy for diagnostic evaluation of suspected tuberculosis. Second-generation tests, however, might have sufficiently high sensitivity, low negative likelihood ratio, and correspondingly high negative predictive value in low-incidence settings to facilitate prompt rule-out of tuberculosis.**

Funding **National Institute for Health Research.**

Introduction

Prompt diagnosis and treatment of active tuberculosis are essential for optimal patient outcomes and preventing onward transmission in the community and in health-care facilities.[1] However, diagnostic assessment of suspected tuberculosis can be lengthy, costly, and burdensome for patients and health-care systems,[2] often resulting in substantial delays in diagnosis and treatment of other diseases in cases when suspected tuberculosis is eventually ruled out. Therefore, improving and accelerating diagnostic evaluation remains a clinical and public health priority in high-income, low-incidence countries, as well as in high-burden regions. Over the past decade, advances in molecular diagnostics, such as the GeneXpert assay (Cepheid, Sunnyvale, CA, USA), have improved the speed and accuracy of microbiological diagnosis and enabled prediction of antibiotic susceptibility.[3] However, although these tests have high specificity (which is important to rule in tuberculosis), they have insufficient sensitivity to rule out tuberculosis and require clinical specimens from anatomical disease sites, which are often obtained with resource-intensive invasive procedures.[4] A blood test of high diagnostic sensitivity could help to promptly (eg, in 24 h) triage patients at clinical presentation (appendix), which would address a major unmet clinical need and has been prioritised by WHO.[5] Given the paucibacillary nature of most cases of culture-negative tuberculosis, such a test would probably be based on measurement of immune responses to *Mycobacterium tuberculosis* rather than direct detection of the bacteria or nucleic acids.

Interferon-γ release assays (IGRAs) are regulatory-approved, immune-based blood tests for detecting *M tuberculosis* infection. By measuring T-cell responses to two strongly immunogenic but highly specific

Lancet Infect Dis 2019; 19: 193–202

Published Online January 14, 2019 http://dx.doi.org/10.1016/S1473-3099(18)30613-3

This online publication has been corrected; the first correction appeared at thelancet.com/infection on January 30, 2019. The second correction appeared on January 31, 2019

See Comment page 121

*Authors contributed equally

†Authors contributed equally

‡Members of study group listed at the end of the Article

Tuberculosis Research Centre, National Heart and Lung Institute (H S Whitworth PhD, A Badhan MSc, A A Boakye MSc, M Rees-Roberts PhD, H Lambie MSc, Prof O M Kon MD, Prof A Lalvani DM) **and National Institute for Health Research Health Protection Research Unit in Respiratory Infections** (A Badhan, A A Boakye, M Rees-Roberts, Prof O M Kon, Prof A Lalvani)**, Imperial College London, London, UK; Department of Clinical Research, London School of Hygiene and Tropical Medicine, London, UK** (H S Whitworth)**; Institute of Applied Health Research, University of Birmingham, Birmingham, UK** (Y Takwoingi PhD, C Partlett PhD, Prof J J Deeks PhD)**; Centre for Health Services Studies, University of Kent, Canterbury, UK** (M Rees-Roberts)**; Heart of England National Health Service (NHS) Foundation Trust, Birmingham, UK** (J Innes FRCP)**; Department of Infectious Diseases, St Mary's Hospital, Imperial College Healthcare Trust, London, UK** (Prof G Cooke PhD)**; Department of Respiratory Medicine, Royal Free London NHS Foundation**

Trust, London, UK (M Lipman MD)**; University College London Respiratory, Division of Medicine, University College London, London, UK** (M Lipman)**; Nuffield Department of Medicine, Oxford University Hospitals NHS Trust, Oxford, UK** (Prof C Conlon MD)**; Infection Care Group** (Prof D Macallan PhD) **and Department of Respiratory Medicine** (F Chua PhD)**, St George's University Hospitals NHS Foundation Trust, London, UK; Institute of Infection and Immunity, St George's, University of London, London, UK** (Prof D Macallan)**; Department of Sexual Health and HIV, King's College Hospital NHS Foundation Trust, London, UK** (Prof F A Post PhD)**; Department of Infection and Tropical Medicine, University Hospitals of Leicester NHS Trust, Leicester, UK** (Prof M Wiselka PhD)**; and Department of Infection, Immunity and Inflammation, Respiratory Biomedical Research Centre, Institute for Lung Health, University of Leicester, Leicester, UK** (G Woltmann MD)

Correspondence to: Prof Ajit Lalvani, Tuberculosis Research Centre, National Heart and Lung Institute, Imperial College London, London W2 1PG, UK
a.lalvani@imperial.ac.uk

See Online for appendix

Research in context

Evidence before this study

Although the utility of interferon-γ release assays (IGRAs) for the diagnosis of active tuberculosis is unclear, their use in clinical practice is common. A comprehensive systematic review and meta-analysis published in 2011 described data from studies published between January, 2001, and November, 2009, that evaluated the diagnostic accuracy of IGRAs for active tuberculosis. Therefore, we searched PubMed for original research studies published in any language between Dec 1, 2009, and June 30, 2018, using the search terms (Tuberculosis OR TB) AND (active OR disease) AND (interferon-γ release assay OR T-SPOT.*TB* OR QuantiFERON) AND (diagnosis OR evaluation OR rule-in OR rule-out). The evidence base suggests that commercially available IGRAs have insufficient specificity to rule in tuberculosis and insufficient sensitivity to rule out tuberculosis. However, this finding is derived primarily from studies that are either small, low quality, or not representative of patient populations in routine clinical practice. Only one large prospective cohort study embedded in routine practice was identified, but in a high-incidence setting of tuberculosis. Thus, 15 years after the introduction of IGRAs, the ability of policy makers in low-incidence settings to generate recommendations and guidelines for the use of IGRAs to diagnose active tuberculosis is still hampered by a paucity of reliable and informative evidence.

Added value of this study

This study is the largest prospective study to investigate the utility of IGRAs for the diagnosis of active tuberculosis in a low tuberculosis-incidence setting. Because the study was done at multiple centres in England, and patients representing the full natural clinical spectrum of tuberculosis were recruited, the results are generalisable to other high-income, low-incidence settings. By showing that existing IGRAs have no useful role in the diagnosis of active tuberculosis, this study aims to resolve a major clinical uncertainty and represents a substantial new high-quality component of the evidence base. Simultaneous evaluation of second-generation IGRAs (incorporating novel *Mycobacterium tuberculosis* antigens) identifies a potentially useful high-sensitivity triage test that meets, we believe, a major unmet clinical need.

Implications of all the available evidence

Existing IGRAs do not have sufficient sensitivity or negative predictive value to rule out a diagnosis of tuberculosis. They also have low specificity and, as a result, are unable to rule in a diagnosis of tuberculosis. Thus, existing IGRAs do not have a clinically useful role in the diagnostic investigation of tuberculosis. The finding that second-generation IGRAs might have sufficiently high sensitivity, low negative likelihood ratio, and high negative predictive value to be used as a triage test to help rule out a diagnosis of tuberculosis within 24 h (the time needed to obtain the test result) 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 postlicensure.

M tuberculosis antigens (ESAT-6 and CFP-10), they are not confounded by previous BCG vaccination and provide higher diagnostic specificity than the tuberculin skin test.[6] Given that *M tuberculosis* infection is a prerequisite for tuberculosis disease, a negative IGRA result could potentially rule out a diagnosis of tuberculosis (ie, exclude tuberculosis from the differential diagnosis), although previous evidence suggests the sensitivity of available IGRAs might be insufficient to fulfil this triage function.[1,7–9]

Although established as the new standard of care for diagnosing latent tuberculosis infection, IGRAs are not recommended for the diagnosis of active tuberculosis other than in specific scenarios (eg, in children), with caveats around interpretation of the results and the expertise required.[10,11] However, development of definitive recommendations has been hindered by a paucity of robust and informative evidence. Most studies of the diagnostic accuracy of IGRAs in active tuberculosis, to our knowledge, are retrospective reviews of hospital records and tuberculosis registry data or small case-control studies, typically not representative of the heterogeneous patient population seen in clinical practice. Although one large (n=746) prospective cohort study embedded in routine practice and including a head-to-head comparison of T-SPOT.*TB* and QuantiFERON-TB Gold In-Tube (QFT-GIT) tests was published in 2018, the study was done in a setting with a high incidence of tuberculosis.[12] Prospective cohort studies done in low-incidence settings have been substantially smaller.[1,8]

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

Discovery of other highly specific *M tuberculosis* antigens that are as strongly immunogenic as ESAT-6 and CFP-10 presents an opportunity to develop second-generation IGRAs of higher sensitivity.[13,14] Furthermore, they might allow development of an ESAT-6-free IGRA for application in populations vaccinated with new ESAT-6-based tuberculosis vaccines, as previously described.[15] Evidence from previous studies suggests that the 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 done in routine practice in a low tuberculosis incidence setting.

Therefore, we sought to evaluate the clinical utility of existing IGRAs—T-SPOT.*TB* (Oxford Immunotec, Abingdon, UK) and QFT-GIT (Qiagen, Hilden, Germany)—and second-generation IGRAs incorporating novel *M tuberculosis* antigens in patients presenting with suspected tuberculosis in English hospitals.

Methods

Study design

We did a prospective, multicentre, cohort study in routine clinical practice in England. A within-patient design was used to compare test accuracy: blood samples from each study participant were tested with all IGRAs, with the presence or absence of active tuberculosis verified with 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 (reference 11/H0722/8). All participants provided written, informed consent. The study protocol is available online and a standards for reporting studies of diagnostic accuracy checklist is provided in the appendix.

Study participants

Adult inpatients and outpatients presenting with suspected active tuberculosis (based on signs and symptoms assessed by the attending hospital clinician) were consecutively enrolled from ten National Health Service hospitals in five English 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 pretest probabilities for active tuberculosis were included. Exclusion criteria were restricted to age younger than 16 years and inability or unwillingness to provide informed consent. Centres were selected to ensure the population was representative of various ethnic groups and comorbidities.

Participant enrolment and follow-up

Participants were first observed by research nurses at enrolment. Following consent, a baseline blood sample was drawn and data were collected in a case report form on the demographics and medical history of the participant and on investigations performed during routine diagnostic work-up. Participants were followed up for 2 months and 6 months thereafter, with data collected on any subsequent investigations, test results, diagnoses, and response to tuberculosis treatment if initiated. Patients with a definitive non-tuberculosis diagnosis who were discharged from routine care were not required to attend follow-up visits, but when necessary, data were collected from hospital records up to 12 months after enrolment to identify final diagnoses made by hospital clinicians.

For the **study protocol** see https://njl-admin.nihr.ac.uk/document/download/2006627

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 with a composite reference standard[1] by a panel of four or more respiratory medicine and infectious disease clinicians specialising in tuberculosis. The panel assessed anonymised clinical

	Criteria	Number of patients (n/N)
1: Culture-confirmed tuberculosis*	Microbiological culture of *Mycobacterium tuberculosis*, and clinical and radiological findings suggestive of tuberculosis	261/845 (31%)
2: Highly probable tuberculosis*	Clinical and radiological features highly suggestive of tuberculosis and unlikely to be caused by other diseases, a decision to treat made by a clinician, appropriate response to therapy, and histological evidence if available	102/845 (12%)
3: Clinically indeterminate diagnosis	Final diagnosis of tuberculosis neither highly probable nor reliably excluded	43/845 (5%)
4: Active tuberculosis excluded		
4A: Inactive tuberculosis	Stable chest x-ray changes, TST positive† (if available), bacteriologically negative (if available), and no clinical evidence of active disease	7/845 (1%)
4B: One or more risk factors for tuberculosis exposure‡, TST positive†	TST positive†, bacteriologically negative (if available), and no clinical evidence of active disease	48/845 (10%)
4C: One or more risk factors for tuberculosis exposure‡, TST negative	History of exposure to tuberculosis and TST negative (if available)	267/845 (32%)
4D: No risk factors for tuberculosis exposure‡, TST negative	No history of exposure to tuberculosis and TST negative (if available)	117/845 (14%)

TST=tuberculosis skin test. **M tuberculosis* culture is the gold standard for diagnosis of active tuberculosis; however, given that culture does not detect all tuberculosis cases, our previously validated standard reference includes a second category for culture-negative but highly probable active tuberculosis diagnoses, made on the basis of other available evidence.[1] †TST by use of Mantoux test, with a threshold ≥15 mm considered positive. ‡Risk factors for exposure to tuberculosis: recent exposure to a patient with active tuberculosis, born in country of high prevalence, or belonging to an ethnic group with a very high prevalence of tuberculosis (incidence >100 per 100 000 people).

Table 1: **Predefined criteria for case definitions and diagnostic categories[1]**

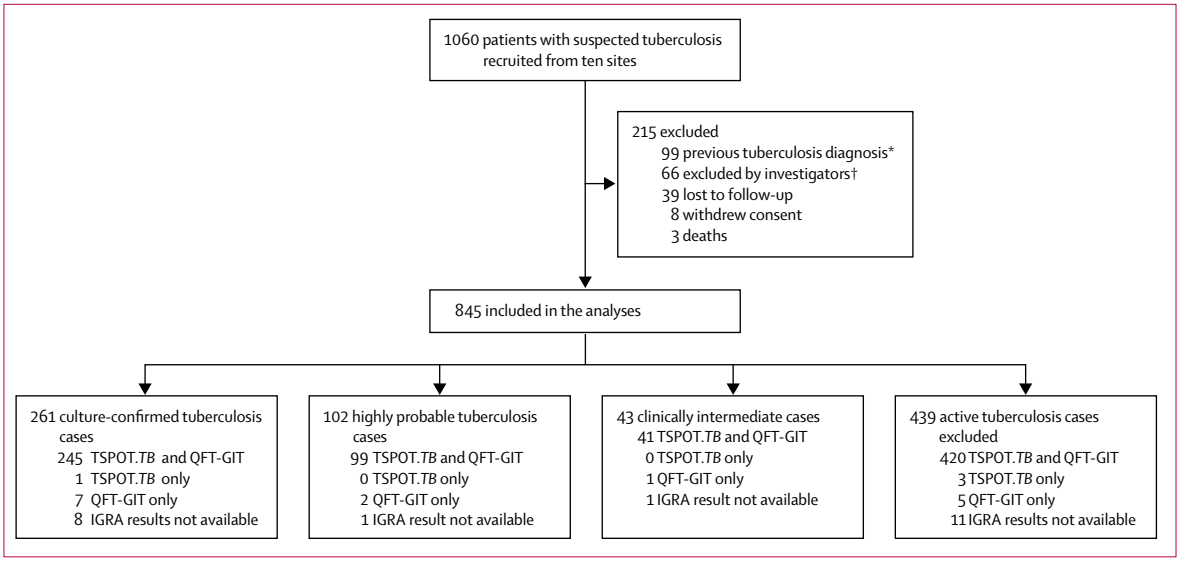

Figure: **Study selection for patients with suspected active tuberculosis**
QFT-GIT=QuantiFERON-TB Gold In-Tube. IGRAs=interferon-γ release assays. *Patients with previously diagnosed tuberculosis were excluded from analyses because results from IGRAs cannot be reliably interpreted in these patients; this decision was made by the expert diagnostic panel and study management group in consultation with the independent study steering committee before unblinding of IGRA and second-generation IGRA results. †Upon advice from the study steering committee, and following consultation between the study management group and data management groups, a total of 66 patients were excluded from analyses; patients recruited from one study site (n=40) were excluded because they all had diagnoses of confirmed or highly probable tuberculosis (categories 1 and 2) due to an error of implementation of recruitment criteria at this site (ie, patients with suspected tuberculosis were not being recruited). A further subset of patients (n=26) were, upon review, considered by the expert diagnostic panel to be ineligible (before unblinding IGRA results) on the basis that they were being investigated for tuberculosis (incidental atypical chest x-ray, known contact with active tuberculosis, or screening for anti-tumour necrosis factor treatment), but did not present with symptoms or signs suggestive of tuberculosis.

data (patient demographics, medical history, tuberculosis symptoms, previous tuberculosis information, history of exposure to tuberculosis, current medication, HIV status, relevant clinical correspondence, test results during diagnosis and follow-up, and any other relevant clinical information) while being masked to all IGRA results (including IGRAs done as part of routine practice at the recruitment sites).

Diagnoses of all participants were categorised into the following groups (table 1): definite tuberculosis (category 1), highly probable tuberculosis (category 2), clinically indeterminate diagnosis (category 3), and non-tuberculosis (category 4). Category 4 participants were subdivided on the basis of risk factors for latent tuberculosis infection. Final diagnoses and diagnostic categories were determined by consensus across the panel.

Laboratory procedures

Blood samples (35 mL) were collected from all participants at enrolment in heparinised and QFT-GIT blood collection tubes. QFT-GIT and T-SPOT.*TB* were done and interpreted in real time at the Tuberculosis Research Centre (Imperial College London, London, UK) according to the manufacturer's instructions and as described by Whitworth and colleagues.[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. The laboratory scientists were masked to all clinical information, diagnoses, and personal identifiers.

Statistical analysis

This 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 of 75% for QFT-GIT.[1,7–9] Accounting for the paired nature of the data and assuming independence of errors,[17] 855 patients (after loss to follow-up, withdrawal, or exclusion because of missing or invalid index or 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 tuberculosis in the study population. We calculated sensitivity, specificity, positive and negative predictive values, and positive and negative likelihood ratios for each test. 95% CIs were calculated with the Wilson method for proportions,[18,19] and likelihood ratios with the method by Simel and colleagues.[20] All patients in diagnostic categories 1, 2, and 4 were included in analyses (table 1); data for patients in category 3 were reported but not included in the analyses.

Patients with indeterminate IGRA or borderline T-SPOT.*TB* results were excluded from primary analyses but included as test positives in sensitivity analyses. Sensitivity analyses were also done to investigate the effects of excluding category 2 patients on IGRA sensitivity and of excluding category 4A–C patients on IGRA specificity. To compare the accuracy of two IGRAs, we fit

	Culture-confirmed tuberculosis (n=261)	Highly probable tuberculosis (n=102)	Clinically indeterminate (n=43)	Active tuberculosis excluded (n=439)	Total (n=845)
Clinical setting					
Outpatient	171 (66%)	72 (71%)	32 (74%)	269 (61%)	544 (64%)
Inpatient	90 (34%)	30 (29%)	11 (2·6%)	170 (3·9%)	301 (3·6%)
Age, years	32 (26–42)	36 (29–45)	38 (27–56)	44 (33–56)	38 (30–51)
Male	177 (68%)	53 (52%)	21 (4·9%)	250 (57%)	501 (59%)
Female	84 (32%)	49 (48%)	22 (51%)	189 (43%)	344 (41%)
Ethnic origin					
Indian subcontinent	167 (64%)	61 (60%)	16 (37%)	168 (38%)	412 (49%)
Black	50 (19%)	22 (22%)	10 (23%)	102 (23%)	184 (22%)
White	22 (8%)	9 (9%)	12 (2·8%)	126 (29%)	169 (20%)
Asian	16 (6%)	6 (6%)	5 (12%)	14 (3%)	41 (5%)
Middle Eastern	4 (2%)	0	0	12 (3%)	16 (2%)
Mixed	1 (<1%)	4 (3·9%)	0	8 (2%)	13 (2%)
Hispanic	1 (<1%)	0	0	7 (2%)	8 (1%)
Unknown	0	0	0	2 (<1%)	2 (<1%)
Years in the UK	4·8 (2·3–10·7)	6·1 (3·1–16·3)	10·6 (5·8–34·5)	13·1 (6·4–32·1)	8·3 (3·3–20·4)
Profession*					
Paid employment	130 (50%)	52 (51%)	21 (4·9%)	214 (4·9%)	417 (49%)
Unemployed	62 (24%)	24 (24%)	16 (37%)	164 (37%)	266 (31%)
Student	50 (19%)	13 (13%)	3 (7%)	26 (6%)	92 (11%)
Health-care or laboratory worker	16 (6%)	9 (9%)	2 (5%)	24 (5%)	51 (6%)
Social or prison worker	1 (<1%)	1 (1%)	0	2 (<1%)	4 (<1%)
Sex worker	0	1 (1%)	0	2 (<1%)	3 (<1%)
Unknown	2 (1%)	2 (2%)	1 (2%)	7 (2%)	12 (1%)
Height, m	1·7 (1·6–1·8)	1·7 (1·6–1·8)	1·6 (1·6–1·7)	1·7 (1·6–1·8)	1·7 (1·6–1·8)
Weight, kg	63 (55–72)	64 (54–75)	71 (57–85)	68 (58–79)	65 (57–77)
Body-mass index, kg/m²	22 (20–25)	22 (21–26)	24 (20–32)	24 (21–28)	23 (20–27)
BCG vaccinated	194 (74%)	79 (77%)	36 (8·4%)	340 (77%)	649 (77%)
BCG scar visible					
Yes	172 (6·6%)	72 (7·1%)	29 (67%)	283 (64%)	556 (66%)
No	12 (5%)	3 (3%)	3 (7%)	19 (4%)	37 (4%)
Unknown	16 (6%)	8 (8%)	6 (14%)	44 (10%)	74 (9%)
Data not available	61 (23%)	19 (1·9%)	5 (1·2%)	93 (21%)	178 (21%)
Known contact with tuberculosis	70 (27%)	25 (25%)	12 (28%)	83 (19%)	190 (22%)
Other pre-existing conditions or comorbidities†					
None	169 (65%)	61 (60%)	19 (44%)	169 (38%)	418 (49%)
HIV infected	13 (5%)	12 (12%)	2 (4%)	108 (25%)	135 (16%)
Diabetes	22 (8%)	5 (5%)	8 (19%)	53 (12%)	88 (10%)
Asthma	12 (5%)	5 (5%)	4 (9%)	50 (11%)	71 (8%)
Cancer	1 (<1%)	1 (1%)	0	12 (3%)	14 (2%)
Chronic or end-stage kidney disease	5 (2%)	1 (1%)	2 (5%)	4 (1%)	12 (1%)
Hepatitis C	1 (<1%)	1 (1%)	0	10 (2%)	12 (1%)
Hepatitis B	5 (2%)	1 (1%)	0	5 (1%)	11 (1%)
Organ transplant	0	0	0	2 (<1%)	2 (<1%)
Sarcoidosis	1 (<1%)	0	0	0	1 (<1%)
Other	74 (28%)	37 (36%)	20 (47%)	228 (52%)	359 (42%)

Data are n (%) or median (IQR). *Some patients had more than one profession. †Some patients had multiple comorbidities.

Table 2: **Demographics and clinical characteristics of participants**

separate generalised estimating equation models for patients with active tuberculosis to estimate differences in sensitivity and without active tuberculosis to investigate differences in specificity. This approach exploits the paired nature of the data while 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 generalised estimating

equation models using a postestimation procedure, with 95% CIs calculated with the delta method. Analyses were done with Stata (version 13.0).

Role of the funding source

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

Results

Between Nov 25, 2011, and Aug 31, 2013, 1060 patients with suspected active tuberculosis provided consent and were enrolled. Patients with a history of tuberculosis diagnosis (n=99) were excluded from analyses, as in previous studies.[12] Additionally, 116 patients were also excluded (figure), giving a final study population of 845 patients.

Demographic and clinical characteristics for the final study population are shown in table 2. The median age of the cohort was 38 years (IQR 30–51); over half of the patients were men, and almost half were from the Indian subcontinent. One or more comorbidities were reported in 427 (51%) of 845 participants (table 2). Medications at presentation are shown in the appendix. The most common symptoms reported at presentation were cough, weight loss, and lethargy (appendix).

363 (43%) of 845 patients had a final diagnosis of active tuberculosis (categories 1 and 2; table 1). Of these, 129 (36%) had pulmonary tuberculosis, 189 (52%) had extrapulmonary tuberculosis, and 45 (12%) had both (table 3); most patients had lymph node involvement. Of 261 patients who had drug-susceptibility testing done on *M tuberculosis* isolates, 21 (6%) had drug-resistant tuberculosis and one (<1%) had multidrug-resistant tuberculosis. Tuberculosis was excluded (category 4) in 439 (52%) of 845 patients (table 1). These patients were subclassified according to risk factors for latent tuberculosis infection or inactive tuberculosis into categories 4 A–D in decreasing likelihood of having *M tuberculosis* infection (table 1).[1] The most common non-tuberculosis diagnoses are listed in table 3. Only 43 (5%) of 845 patients were classified as clinically indeterminate (category 3).

TSPOT.*TB* results were available for 809 (96%) of 845 study participants and QFT-GIT results for 820 (97%); data from both tests were available for 805 (95%) patients (figure). Diagnostic sensitivity, specificity, positive and negative predictive values, and positive and negative likelihood ratios are shown in table 4, with a cross-tabulation of T-SPOT.*TB* and QFT-GIT results in patients with active tuberculosis and non-tuberculosis diagnoses provided in the appendix. Sensitivity of T-SPOT.*TB* was 84·9% (95% CI 79·5–89·0) for culture-confirmed tuberculosis and 81·4% (76·6–85·3) for all tuberculosis diagnoses, giving a negative predictive value of 84·6% (80·6–87·9) and negative likelihood value of 0·22 (0·17–0·27) for all tuberculosis cases. Specificity was 86·2% (82·3–89·4) for all patients without tuberculosis and 93·5% (86·6–97·0) for cases with no risk factors for latent tuberculosis infection (category 4D, table 4). Sensitivity of QFT-GIT was 70·6% (64·4–76·1) for culture-confirmed tuberculosis and 67·3% (62·0–72·1) for all tuberculosis cases, giving a negative predictive value of 74·0% (69·5–78·0) and negative likelihood ratio of 0·41 (0·35–0·48) for all tuberculosis cases (table 4). Specificity was 80·4% (76·1–84·1) for all patients without tuberculosis and 93·4% (86·4–96·9) for cases with no risk factors for latent tuberculosis infection. Sensitivity and specificity of T-SPOT.*TB* were superior to QFT-GIT; the relative sensitivity was 1·20 (95% CI 1·12–1·29, p<0·0001) and the relative specificity was 1·07 (1·02–1·12, p=0·0039).

Second-generation and ESAT-6-free IGRA results were available for 809 (96%) of 845 patients. Sensitivity of second-generation IGRA was 94·0% (95% CI 90·0–96·4) for culture-confirmed tuberculosis and 89·2%

	Confirmed or highly probable tuberculosis (n=363)	Active tuberculosis excluded (n=439)*
All cases	363 (100%)	..
Pulmonary	129 (36%)	..
Extrapulmonary	189 (52%)	..
Pulmonary and extrapulmonary	45 (12%)	..
Site of disease†		
Lungs	174 (48%)	..
Lymph node	154 (42%)	..
Pleura	26 (7%)	..
Spine	16 (4%)	..
Miliary tuberculosis (disseminated)	11 (3%)	..
Abdomen	9 (2%)	..
Pericardium	6 (2%)	..
Brain	6 (2%)	..
Musculoskeletal system	5 (1%)	..
Chest wall	2 (1%)	..
Other	31 (9%)	..
Pneumonia	..	104 (24%)
Sarcoidosis	..	38 (9%)
Cancer	..	36 (8%)
Lower respiratory tract infection	..	23 (5%)
Reactive lymphadenopathy	..	18 (4%)
Chest infection	..	16 (4%)
Exacerbation of asthma	..	14 (3%)
Upper respiratory tract infection	..	13 (3%)
Non-tuberculosis mycobacterium infection	..	12 (3%)
Exacerbation of bronchiectasis	..	11 (3%)
Exacerbation of chronic obstructive pulmonary disease	..	8 (2%)
Other‡	..	158 (36%)

*Some patients had multiple diagnoses. †Some patients had tuberculosis at multiple anatomical sites. ‡Less than five cases per diagnosis.

***Table 3*: Final diagnoses of patients with and without active tuberculosis**

	T-SPOT.*TB*		QFT-GIT		ESAT, CFP-10, and Rv3615c		CFP-10, Rv3615c, and Rv3879c	
	n/N	Estimate (95% CI)	n/N	Estimate (95% CI)	n/N	Estimate (95% CI)	n/N	Estimate (95% CI)
Sensitivity for active tuberculosis								
All	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 tuberculosis	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 tuberculosis*	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 tuberculosis†	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 tuberculosis†‡	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 tuberculosis	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)
Extrapulmonary tuberculosis	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 tuberculosis								
Active tuberculosis 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 tuberculosis 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 tuberculosis								
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 tuberculosis								
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)

25 of 845 QuantiFERON-TB Gold In-Tube (QFT-GIT) tests, 36 of 845 T-SPOT.*TB* tests, and 36 second-generation IGRAs were missing because of blood draw difficulties, samples being unsuitable for testing, or samples being destroyed for laboratory reasons; missing results were spread across all diagnostic categories. Indeterminate and borderline IGRA results were excluded from the analysis and are not presented in this table; numbers of indeterminate and borderline results for T-SPOT.*TB* or QFT-GIT and second-generation IGRA are presented in the appendix. When indeterminate and borderline results were included as test positives in sensitivity analyses (ie, the result could not exclude a tuberculosis diagnosis), sensitivity results for all tuberculosis cases were: 83·2% (95% CI 78·9–86·8) for T-SPOT.*TB*, 69·7% (64·7–74·2) for QFT-GIT, 90·4% (86·9–93·1) for second-generation IGRA (ESAT-6, CFP-10, and Rv3615c), and 89·6% (85·9–92·4) for ESAT-6-free IGRA (CFP-10, Rv3615c, and Rv3879c). IGRAs=interferon-γ release assays. LTBI=latent tuberculosis infection. TST=tuberculin skin test. *Highly probable tuberculosis includes culture-negative tuberculosis cases plus ten patients with a final diagnosis of tuberculosis who were not tested for *Mycobacterium tuberculosis*; sensitivity (95% CIs) results for culture-negative tuberculosis alone were: 69·9% (95% CI 59·3–78·7) for T-SPOT.*TB*, 57·1% (46·5–67·2) for QFT-GIT, 75·0% (64·5–83·2) for second-generation IGRA (ESAT-6, CFP-10, and Rv3615c), and 73·1% (62·3–81·7) for ESAT-6-free IGRA (CFP-10, Rv3615c, and Rv3879c). †Smear microscopy not done for 56 of 845. ‡Among 165 patients who were smear-negative but culture-positive, 122 of 142 participants were T-SPOT.*TB*-positive; 105 of 153 were QFT-GIT-positive; 135 of 144 were positive in second-generation IGRA, and 131 of 141 were positive in ESAT-6-free IGRA.

Table 4: **Diagnostic accuracy of commercially available and second-generation IGRAs for diagnosis of active tuberculosis**

(85·2–92·2) for all tuberculosis cases, giving a negative predictive value of 90·0% (86·2–92·8) and negative likelihood ratio of 0·13 (0·10–0·19) for all tuberculosis cases. Specificity was 80·0% (75·6–83·8) for all patients without tuberculosis and 91·3% (83·8–95·5) for cases with no risk factors for latent tuberculosis infection. Sensitivity of ESAT-free IGRA was 93·4% (89·2–96·0) for culture-confirmed tuberculosis and 88·0% (83·8–91·2) for all tuberculosis cases, giving a negative predictive value of 89·2% (85·4–92·1) and negative likelihood ratio of 0·15 (0·11–0·21) for all tuberculosis cases. Specificity was 79·6% (75·2–83·4) for patients without tuberculosis and 90·3% (82·6–94·8) for cases with no risk factors for latent tuberculosis infection.

Comparing second-generation IGRA with T-SPOT.*TB*, relative sensitivity was 1·08 (95% CI 1·04–1·11, p<0·0001) and relative specificity was 0·94 (0·91–0·96, p<0·0001). For ESAT-6-free IGRA versus T-SPOT.*TB*, relative sensitivity was 1·07 (1·03–1·10, p=0·0002) and relative specificity was 0·93 (0·90–0·96, 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. Responses to Rv3615c were similar to or higher than those to ESAT-6 and CFP-10.

Of the 232 patients with culture-confirmed tuberculosis and with available smear microscopy results, 165 (71%) had smear-negative results (57 [35%] of 165 patients had pulmonary tuberculosis, 80 [48%] patients had extrapulmonary tuberculosis, and 28 [17%] patients had both). Sensitivities of the tests in this population were: 85·9% (95% CI 79·2–90·7) for T-SPOT.*TB*, 68·6% (60·9–75·4) for QFT-GIT, 93·8% (88·5–96·7) for second-generation IGRA, and 92·9% (87·4–96·1) for ESAT-6-free IGRA.

Of the 135 study participants with HIV, 25 (19%) had a final diagnosis of active tuberculosis and 108 (80%) had a diagnosis that excluded tuberculosis. 27 (31%) of 88 participants with diabetes had a final diagnosis of tuberculosis (table 2). Sensitivity and specificity of all IGRAs for active tuberculosis in patients with HIV and diabetes are shown in the appendix.

The proportion of indeterminate test results did not differ between QFT-GIT (79 [10%] of 820 patients) and TSPOT.*TB* (57 [7%] of 809 patients, p=0·061), but was higher for QFT-GIT than for second-generation IGRA (55 [7%] of 809 patients, p=0·038) and ESAT-6-free IGRA (55 [7%] of 809 patients, p=0·038). Most indeterminate results occurred in patients without tuberculosis (appendix). T-SPOT.*TB* results were borderline in

17 (5%) of 345 patients with active tuberculosis and 16 (4%) of 423 of patients with a non-tuberculosis diagnosis. Lowering the cutoff of T-SPOT.*TB* from eight to five spot-forming cells (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. Scoring both indeterminate and borderline results as positives also did not affect test performance in sensitivity analyses (table 4).

Discussion

To our knowledge, this study is the largest prospective cohort study embedded in routine clinical practice to assess and compare the role of IGRAs in the evaluation of suspected pulmonary and extrapulmonary tuberculosis in a low-incidence setting. Although T-SPOT.*TB* had significantly higher sensitivity than QFT-GIT, neither assay had sufficient sensitivity or negative predictive value to rule out a diagnosis of active tuberculosis. By contrast, the second-generation IGRA, incorporating Rv3615c with ESAT-6 and CFP-10, had significantly higher diagnostic sensitivity than T-SPOT.*TB* and QFT-GIT. Also reflecting common clinical practice (despite the absence of good evidence or guidelines supporting use of IGRAs in this setting), IGRAs were used as part of routine diagnostic investigation for active tuberculosis in 35% of study patients; these patients were distributed across the recruiting sites (data not shown).

The negative likelihood ratio of 0·13 for second-generation IGRA means a negative test result would reduce the odds of tuberculosis post-test by a clinically meaningful factor of 7·7 times compared with pretest (appendix). The negative predictive value for all tuberculosis cases, including highly probable cases, was 90%, despite the 43% prevalence in this population presenting to urban infectious diseases and respiratory medicine services with suspected tuberculosis. Since our study was done in routine clinical practice and encompassed the full, natural clinical spectrum of tuberculosis and non-tuberculosis diagnoses, the results are probably generalisable across clinical practice in high-income, low-incidence countries. Accordingly, in clinical settings with a low-to-moderate pretest probability of tuberculosis, such as general medical inpatient and outpatient services or primary care, second-generation IGRA has sufficiently low negative likelihood ratio to almost rule out tuberculosis. For example, a negative test result would convert pretest probability of 20% to a post-test probability of 3·1%, and a 10% pretest probability to 1·4% post-test probability (appendix). This test would therefore provide a useful prompt triage of patients on initial presentation, similar to the role other diagnostic tests of high sensitivity and restricted specificity have, such as serum D-dimer in patients with low-to-moderate suspicion of venous thromboembolism.[21] To our knowledge, other available tests for tuberculosis do not have the 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 tuberculosis is only 63%[3] (and sensitivity of Xpert MTB/RIF was even lower),[4] compared with 93·8% (95% CI 88·5–96·7) 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 positive predictive value, enabling rule in of active tuberculosis. Second-generation IGRA might, thus, have a complementary role to rapid molecular tests in the diagnostic investigation of suspected tuberculosis.

Given that IGRAs are the standard of care for detecting latent tuberculosis infection,[10,11] they will inevitably identify latent tuberculosis infection when active tuberculosis has been excluded. Because most people with possible tuberculosis in low-burden countries are from ethnic groups with a high prevalence of latent tuberculosis infection,[22] similar to our study, the diagnostic specificity for active tuberculosis 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 positive likelihood ratios of available and second-generation IGRAs mean that a positive result cannot rule in a diagnosis of tuberculosis. However, the specificity of all IGRAs increased to 90–93% in patients with active tuberculosis excluded and no risk factors for latent tuberculosis infection (category 4D). Thus, a positive IGRA result might help to keep a diagnosis of active tuberculosis in the differential diagnosis in populations with a very low prevalence of latent tuberculosis infection, which, however, is not usually the case in patient populations being assessed for possible tuberculosis.

Two of the leading new tuberculosis vaccine candidates, Hybrid 1-IC31[23] and H56:IC31,[24] contain ESAT-6 and might 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 people who are not *M tuberculosis* infected, analogous to false-positive tuberculosis skin test results in *M tuberculosis*-uninfected people with previous BCG vaccination. Diagnostic accuracy of ESAT-6-free IGRA was very similar to second-generation IGRA and, therefore, has the potential to replace other IGRAs in populations immunised against tuberculosis with ESAT-6-based vaccines.

Two of the most important global risk factors for tuberculosis are HIV co-infection[25] and diabetes,[26] both of which have been reported to adversely affect the performance of IGRAs.[27,28] The performance of commercially available IGRAs in patients with HIV and diabetes in this study was insufficient to be of value in

the diagnosis of active tuberculosis. Performance seemed to be lower in HIV-infected and diabetic subgroups, but the small numbers of patients with tuberculosis in these subgroups precluded statistical comparisons, which was also the case for other types of immunosuppression associated with tuberculosis, such as chronic kidney disease and immunosuppressive medication.

Strengths of our study include the rigorous case definitions, including 6-month follow-up to confirm that a diagnosis of tuberculosis was excluded when a non-tuberculosis diagnosis was not made at presentation. For highly probable tuberculosis, we used a composite reference standard[1] that was applied by a panel of experts and experienced clinicians, masked to IGRA results. Despite this stringent case definition, a proportion of patients without tuberculosis were likely to have been incorrectly categorised as having highly probable tuberculosis, which would explain why all IGRAs had lower sensitivity for highly probable tuberculosis than for all tuberculosis cases, including culture-confirmed cases. Therefore, our estimates of diagnostic sensitivity for all tuberculosis cases, including highly probable tuberculosis, are likely to be conservative, which highlights the significance of increased IGRA sensitivity in culture-confirmed tuberculosis (and the importance of including this subgroup in study analyses) since this is the only population in whom tuberculosis diagnoses are definitive.

Our study also has some limitations. First, it does not include children, in whom the unmet clinical need for improved diagnosis of tuberculosis is high. Second, the numbers of patients with risk factors associated with immunosuppression that affect (eg, HIV) or might affect (eg, diabetes) test performance were modest, precluding clear conclusions about test performance in these subpopulations. Finally, although blood collection and assays were done strictly in accordance with manufacturers' instructions, IGRAs were not done in a routine diagnostic service laboratory, and retesting of new samples was not done for cases when 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 done, its diagnostic accuracy does not appear to be significantly better than that of QFT-GIT and no evidence is available to suggest 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 tuberculosis is unaffected by availability of QFT-GIT-Plus.

In summary, our study provides conclusive and generalisable evidence that existing IGRAs do not have a useful role as rule-in or rule-out tests for tuberculosis in routine clinical practice. However, second-generation IGRAs have a higher sensitivity and negative predictive value, which might help to rule out a diagnosis of tuberculosis in clinical settings with a low-to-moderate prevalence of tuberculosis.

Contributors

HSW was responsible for the daily management of the Interferon-γ Release Assays (IGRA) for Diagnostic Evaluation of Active Tuberculosis (IDEA) study group, 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 did the statistical analyses, produced tables and figures, and contributed to data interpretation. MR-R led the study set-up and initial management, and built the study databases. CP contributed to statistical analyses and producing 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, FAP, MW, and GW contributed to patient recruitment, data collection, and the study expert diagnostic clinical panel. JJD 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.

IDEA study group

David Adeboyeku, Robert Davidson (Northwest London Hospitals National Health Service [NHS] Trust, London); Martin Dedicoat (University Hospitals Birmingham Foundation Trust, Birmingham); Heinke Kunst (Queen Mary University Blizard Institute, London); Michael R Loebingher, Anton Pozniak (Chelsea and Westminster Hospital NHS Foundation Trust, London); William Lynn (Ealing NHS Trust, Southall); Nazim Nathani (Sandwell and West Birmingham Hospitals NHS Trust, Birmingham); Rebecca O' Connell (Bart's Healthcare NHS Trust, London); Sarah Menzies (Frimley Health NHS Foundation Trust, Camberley).

Declaration of interests

HSW, AB, AAB, MR-R, and HL were employed by Imperial College London, London, UK, with grant funding from the National Institute for Health Research (NIHR) to do the work described in this Article. JJD, YT and CP, while working for the University of Birmingham, Birmingham, UK, received grant funding from NIHR to do the work described in this Article. JJD reports grants from NIHR during the study, outside the submitted work. OMK is employed by Imperial Healthcare Trust and was partly 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 organises an annual educational symposium on tuberculosis, sponsored by Qiagen. AL is named inventor on patents (EP057292575, EP1735623[B1], US8,105,797[B2], EP2069792, EP2069792[B1], EP2005182, EP2005182[B1], US8,765,336[B2], EP10716313.1, EP2417456[B1], US9,377,460[B2], US9360480[B2], EP0941478[B2], EP1152012[B1], EP1735623[B1], US8105797[B2], EP1144447[B1], and US9005902[B2]) pertaining to T-cell-based diagnosis, including current and second-generation IGRA technologies. Some of these patents were assigned by the University of Oxford, Oxford, UK, to Oxford Immunotec plc, resulting in royalty entitlements for the University of Oxford and AL. All other authors declare no competing interests.

Acknowledgments

We thank the research nurses at each of the study sites for their work in recruitment of participants; the research assistants at the Tuberculosis Research Centre, in particular Bianca Bartlett, Ann-Kathrin Reuschl, and Luigi Marongiu, who contributed to processing of participant samples and did the study assays; and Alice Halliday who contributed to the data analysis. We also thank the members of the study steering committee, Khalid Khan, Stephen Gordon, James Gray, Johannes B Reitsma, and Nisha Karnani (lay member), for their support and advice throughout the study. We acknowledge the support of the National Institute for Health Research (NIHR) Imperial Biomedical Research Centre and the NIHR Health Protection Research Unit in Respiratory Infections.

References

1 Dosanjh DP, Hinks TS, Innes JA, 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:** 349–62.

3 Dorman SE, Schumacher SG, Alland D, 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:** 76–84.

4 Sohn H, Aero AD, Menzies D, 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:** 970–76.

5 Garcìa-Basterio AL, Cobelens F. Triage tests: a new priority for tuberculosis diagnostics. *Lancet* 2015; **3:** 177–78.

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, 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, et al. Interferon-γ release assays for the diagnosis of active tuberculosis: a systematic review and meta-analysis. *Eur Respir J* 2011; **37:** 100–11.

10 UK International Clinical Guidelines Team. Tuberculosis: prevention, diagnosis, management and service organisation. London: National Institute for Health and Care Excellence, 2016.

11 ECDC. Use of interferon-gamma release assays in support of TB diagnosis. Stockholm: European Centre for Disease Prevention and Control; 2011.

12 Du F, Xie L, Zhang Y, et al. Prospective comparison of QFT-GIT and T-SPOT.TB assays for diagnosis of active tuberculosis. *Sci Rep* 2018; **8:** 5882.

13 Liu XQ, Dosanjh D, Varia H, 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, 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:** 5730–35.

15 Ruhwald M, de Thurah L, Kuchaka D, 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, 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, et al. Simplified diagnostic management of suspected pulmonary embolism (the YEARS study): a prospective, multicentre cohort study. *Lancet* 2017; **390:** 289–97.

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, 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:** 132–41.

24 Luabeya AK, Kagina BM, Tameris MD, 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:** 4130–40.

25 Esmail H, Riou C, Bruyn ED, et al. The immune response to *Mycobacterium tuberculosis* in HIV-1-coinfected persons. *Annu Rev Immunol* 2018; **36:** 603–38.

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:** e78924.

27 Cattamanchi A, Smith R, Steingart KR, et al. Interferon-gamma release assays for the diagnosis of latent tuberculosis infection in HIV-infected individuals: a systematic review and meta-analysis. *J Acquir Immune Defic Syndr* 2011; **56:** 230–38.

28 Faurholt-Jepson D, Aabye MG, Jenson AV, 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:** 384–91.

29 Takasaki J, Manabe T, Morino E, 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:** 188–92.

30 Barcellini L, Borroni E, Brown J, et al. First independent evaluation of QuantiFERON-TB Plus performance. *Eur Respir J* 2016; **47:** 1587–90.