ORIGINAL ARTICLE

Optimization and Interpretation of Serial QuantiFERON Testing to Measure Acquisition of *Mycobacterium tuberculosis* Infection

Elisa Nemes[1,2]*, Virginie Rozot[1,2]*, Hennie Geldenhuys[1,2], Nicole Bilek[1,2], Simbarashe Mabwe[1,2], Deborah Abrahams[1,2], Lebohang Makhethe[1,2], Mzwandile Erasmus[1,2], Alana Keyser[1,2], Asma Toefy[1,2], Yolundi Cloete[1,2], Frances Ratangee[1,2], Thomas Blauenfeldt[3], Morten Ruhwald[3], Gerhard Walzl[4], Bronwyn Smith[4], Andre G. Loxton[4], Willem A. Hanekom[1,2], Jason R. Andrews[5], Maria D. Lempicki[6], Ruth Ellis[6], Ann M. Ginsberg[6], Mark Hatherill[1,2]*, and Thomas J. Scriba[1,2]*; the C-040-404 Study Team, and the Adolescent Cohort Study Team

[1]South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular Medicine, and [2]Division of Immunology, Department of Pathology, University of Cape Town, Cape Town, South Africa; [3]Statens Serum Institut, Copenhagen, Denmark; [4]South Africa Department of Science and Technology–National Research Foundation Centre of Excellence for Biomedical Tuberculosis Research, South African Medical Research Council Centre for Tuberculosis Research, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa; [5]Division of Infectious Diseases and Geographic Medicine, Stanford University School of Medicine, Stanford, California; and [6]AERAS, Rockville, Maryland

ORCID IDs: 0000-0003-1662-4961 (E.N.); 0000-0001-9757-1501 (V.R.); 0000-0002-0641-1359 (T.J.S.).

Abstract

Rationale: Conversion from a negative to positive QuantiFERON-TB test is indicative of *Mycobacterium tuberculosis* (*Mtb*) infection, which predisposes individuals to tuberculosis disease. Interpretation of serial tests is confounded by immunological and technical variability.

Objectives: To improve the consistency of serial QuantiFERON-TB testing algorithms and provide a data-driven definition of conversion.

Methods: Sources of QuantiFERON-TB variability were assessed, and optimal procedures were identified. Distributions of IFN-γ response levels were analyzed in healthy adolescents, *Mtb*-unexposed control subjects, and patients with pulmonary tuberculosis.

Measurements and Main Results: Individuals with no known *Mtb* exposure had IFN-γ values less than 0.2 IU/ml. Among individuals with IFN-γ values less than 0.2 IU/ml, 0.2–0.34 IU/ml, 0.35–0.7 IU/ml, and greater than 0.7 IU/ml, tuberculin skin test positivity results were 15%, 53%, 66%, and 91% ($P < 0.005$), respectively. Together, these findings suggest

that values less than 0.2 IU/ml were true negatives. In short-term serial testing, "uncertain" conversions, with at least one value within the uncertainty zone (0.2–0.7 IU/ml), were partly explained by technical assay variability. Individuals who had a change in QuantiFERON-TB IFN-γ values from less than 0.2 to greater than 0.7 IU/ml had 10-fold higher tuberculosis incidence rates than those who maintained values less than 0.2 IU/ml over 2 years ($P = 0.0003$). By contrast, "uncertain" converters were not at higher risk than nonconverters ($P = 0.229$). Eighty-seven percent of patients with active tuberculosis had IFN-γ values greater than 0.7 IU/ml, suggesting that these values are consistent with established *Mtb* infection.

Conclusions: Implementation of optimized procedures and a more rigorous QuantiFERON-TB conversion definition (an increase from IFN-γ <0.2 to >0.7 IU/ml) would allow more definitive detection of recent *Mtb* infection and potentially improve identification of those more likely to develop disease.

Keywords: IFN-γ release assay; QuantiFERON; conversion; tuberculosis; variability

At a Glance Commentary

Scientific Knowledge on the Subject: Interpretation of serial QuantiFERON-TB Gold In-Tube tests to assess recent *Mycobacterium tuberculosis* infection is complicated by immunological and technical variability.

What This Study Adds to the Field: We report optimized procedures stricter than the manufacturer's recommendations to enhance the consistency of longitudinal QuantiFERON-TB Gold In-Tube assays and data-driven recommendations to interpret test conversions. Distribution of IFN-γ values in populations with differential exposure to *M. tuberculosis*, as well as tuberculin skin test results, interassay variability, short-term conversion rates, and long-term risk of tuberculosis disease, support a more rigorous definition of QuantiFERON-TB conversion as change in IFN-γ values from less than 0.2 to greater than 0.7 IU/ml.

QuantiFERON-TB Gold In-Tube (QFT) (Qiagen, Hilden, Germany) measures immunological sensitization to *Mycobacterium tuberculosis* (*Mtb*) as a biomarker for *Mtb* infection. Compared with persistent QFT negativity, recent conversion from a negative to positive QFT test is associated with higher tuberculosis (TB) disease incidence (1, 2).

In settings with a low incidence of TB disease, individuals with recent infection, detected as QFT conversion from negative to positive, are prioritized for preventive therapy (3). However, QFT test results are confounded by immunological and analytical variability (4, 5), which can result in false-positive conversions. In the absence of risk factors, interpretation of QFT is problematic (6–9), and false-positive QFT conversion may trigger unnecessary preventive therapy.

In high-burden settings, QFT conversion is also likely to represent *Mtb* infection (1, 10), although preventive therapy is not universally recommended, owing to high risk of reinfection and low cost-effectiveness (3, 11). However, such

settings may be ideal for conducting clinical trials of prevention of *Mtb* infection (POI) by novel vaccines or TB preventive therapy, using QFT conversion as an efficacy endpoint. The POI strategy has been proposed as a cost-effective proof-of-concept trial design to test vaccine candidates prior to more resource-intensive efficacy trials based on TB disease endpoints (12, 13). A POI trial (registered with www.clinicaltrials.gov [NCT02075203]) is being conducted to test whether bacillus Calmette-Guérin revaccination or boost vaccination with the novel subunit vaccine H4:IC31 can protect South African adolescents against acquisition of *Mtb* infection. The primary efficacy endpoint is QFT conversion from negative to positive, defined by the positivity cutoff IFN-γ value of 0.35 IU/ml, as recommended by the manufacturer and the CDC (14). However, the immunological and analytical variability of QFT tests (4, 5) potentially confounds interpretation of QFT conversion as a clinical trial endpoint.

Several investigators have proposed caution in interpreting QFT results close to the assay cutoff, where assay variability is more likely associated with discordant results upon serial testing, and they have suggested the introduction of a zone of uncertainty (9, 15). Observations based on serial QFT testing in South Africa (16) and India (7, 17) suggest that this "uncertainty zone" lies between 0.2 and 0.7 IU/ml.

We aimed to identify and minimize sources of QFT assay variability to improve consistency of serial testing. We also report epidemiological and immunological evidence that supports a more rigorous definition of QFT conversion to account for the uncertainty of interpreting values between 0.2 and 0.7 IU/ml. We propose that this definition of QFT conversion is applicable as a more accurate estimate of subsequent risk of disease and as an endpoint for POI clinical trials.

Methods

Study Populations and Procedures

Five separate cohorts of HIV-negative participants with varying degrees of *Mtb* exposure were studied to address different questions. Inclusion and exclusion criteria for each cohort are described in the online supplement.

Cohort 1. Healthy adolescents were recruited at the South African Tuberculosis Vaccine Initiative (SATVI) in South Africa in an epidemiological and immunological cohort study described elsewhere (18–20). The QFT assay was conducted according to the manufacturer's instructions. Baseline IFN-γ values were compared with tuberculin skin test (TST) (RT23; Statens Serum Institut, Copenhagen, Denmark) results, and serial IFN-γ values were used to stratify participants to assess prospective risk of TB progression.

Cohort 2. Healthy adults were recruited at SATVI to assess the effects of variation in the QFT manufacturer's standard operating procedures (SOP), as described in Table 1.

Optimized QFT SOP stricter than the manufacturer's specifications (Table 1) were subsequently implemented and tested in cohorts 3, 4, and 5, which are detailed below.

Cohort 3. Healthy adults with no known exposure to *Mtb* were recruited at the Statens Serum Institut in Denmark.

Cohort 4. Adult patients with microbiologically confirmed TB disease were recruited by the Immunology Research Group at Stellenbosch University and at SATVI in Cape Town, South Africa.

Banked samples from cohorts 3 and 4 were analyzed at SATVI in a blinded fashion by the same operator and using the same ELISA batch to describe the distribution of IFN-γ values in populations with known *Mtb* infection status.

Cohort 5. Healthy adolescents were recruited at SATVI in the POI vaccine trial (registered with www.clinicaltrials.gov [NCT02075203]), as described in the introductory section of the main text above. QFT was performed at screening and repeated 3 and 12 months later (*see* online supplement for study procedures). Interassay variability, distribution of baseline IFN-γ values, and short-term conversion rates were determined in this cohort.

QFT

All quantitative QFT results were expressed as IFN-γ concentration (IU/ml; calculated as TB antigen minus Nil), and qualitative results were defined according to the manufacturer's algorithm. To minimize the effects of known sources of variability (4) on QFT results and improve the consistency of serial testing, restrictions and

Table 1. QuantiFERON-TB Gold In-Tube Standard Operating Procedures

Procedure	Standard Operating Procedures as per QFT Package Insert	Implemented in Optimized Standard Operating Procedures			Reference
		Based on Literature	Upon Further Testing	QC*	
Blood collection tube order	Not specified	Mitogen → TB Ag → Nil	Discard tube → Nil → TB Ag → Mitogen	Right tubes collected	(26)
Blood volume	0.8–1.2 ml		1 ml for Nil and TB Ag, 0.8–1.2 ml for mitogen	Laboratory staff visually check volumes	(5, 27)
Tube mixing	Shake 10 times	Invert 10 times at clinic + 5 min on tube rotator before incubation		Tick on processing log	(27)
Maximum delay between blood collection and incubation start	16 h	2 h		Time recorded, delay check	(5, 28–30)
Duration of incubation at 37°C	16–24 h		16–20 h	Time recorded, delay check	(27, 30)
Delay between end of incubation and centrifugation	Up to 3 d at 4–27°C	Centrifuge and harvest immediately		Time recorded, delay check	
Plasma storage and delay before testing	At 2–8°C for 28 d or −20°C indefinitely	At 2–8°C for 5 d (for operational reasons)	−80°C	Check QFT tube → aliquot	(31)
ELISA standard curve†	Eight standards Four standards		Four standards	Standard curve pass criteria as per manufacturer's instructions	
IFN-γ ELISA	Four standards (rest of the world)	Centrifuge thawed plasma before loading, then follow instructions	Each sample in single well (vs. triplicate)	Check aliquot → plate, Include IQC‡ Check values software → report§, Test new ELISA batch‖	(5, 32)

Definition of abbreviations: Ag = antigen; IQC = internal quality control samples (low, medium, high); QC = quality control; QFT = QuantiFERON-TB Gold In-Tube; TB = tuberculosis.
*Each QC step is routinely documented, and results are not released if specifications are not met.
†QFT package insert commercialized in the United States indicates the use of eight IFN-γ standards (at 8, 4, 2, 1, 0.5, 0.25, 0.125, and 0 IU/ml), whereas in the rest of the world, four standards (at 4, 1, 0.25, and 0 IU/ml) are recommended.
‡Mean and SD calculated over first 11 runs with each new ELISA batch. In addition to manufacturer's instructions to interpret a valid run: (1) if one or more IQC falls between ±2 SD and ±3 SD → repeat samples with TB Ag − Nil = 0.2–0.7 IU/ml (uncertainty zone); and (2) if two or more IQC are outside ±3 SD → repeat plate.
§In addition to manufacturer's instructions to interpret results, rerun samples if: Mitogen < Nil; Nil >> TB Ag; and Nil > TB Ag and inverting values result in different qualitative result (suspect aliquot mixup). If same result, rebleed participant.
‖New IFN-γ ELISA batch must be validated prior to use in parallel with current batch: (1) Run multiple IQC sets (five or more for each low, medium, and high): The majority should be ⩽2 SD of current batch. (2) Run at least 10 sets of unstimulated and stimulated plasma samples (with IFN-γ ⩽4 IU/ml), calculate IFN-γ values in stimulated − unstimulated. No significant (Wilcoxon paired test) difference should be observed when sample sets are analyzed with new versus current batch.
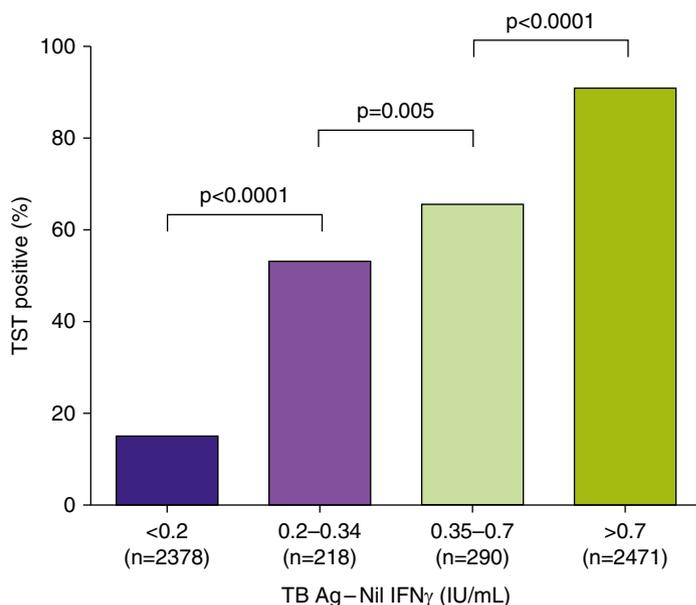
Figure 1. Tuberculin skin test (TST) positivity rates, stratified by QuantiFERON-TB Gold In-Tube (QFT) IFN-γ values. TST and QFT were performed in parallel in cohort 1 (n = 5,357 healthy adolescents). Proportions of adolescents with positive TST results (≥5 mm) were stratified by QFT IFN-γ values at baseline. *Purple shading* denotes negative test results, and *green shading* denotes positive test results, according to the assay cutoff at 0.35 IU/ml. IFN-γ values were further stratified according to an uncertainty zone between 0.2 and 0.7 IU/ml. *P* values were calculated using chi-square tests. Ag = antigen; TB = tuberculosis.

quality control checks were implemented within the boundaries of the manufacturer's instructions (Table 1). These procedures were implemented without further testing when evidence derived from published studies was considered to be robust (Table 1); otherwise, they were tested in cohort 2.

In addition to well-defined training and validation criteria for new operators, as well as participation in an external quality assurance program (run by the College of American Pathologists), an internal quality assurance (IQA) program was implemented. Large volumes of internal quality control

(IQC) samples with low, medium, and high IFN-γ concentrations were generated by appropriate dilution of plasma from stimulated whole blood in plasma from unstimulated whole blood. Hundreds individual plasma aliquots were cryopreserved and included in each ELISA run.

Statistical Analyses

Prism software (version 7; GraphPad Software, La Jolla, CA) was used to perform statistical analyses as described in each figure legend. In cohort 1, incidence of TB disease and rate ratios were calculated for

different QFT conversion classes and compared with nonconverters (*see* online supplement for details).

Results

Immunological Evidence for Uncertainty Zone

Definition of a zone of uncertainty for QFT results has been based primarily on considerations about assay reproducibility (15, 16). To determine if such an uncertainty zone was also supported by an independent immunological test (TST), we analyzed QFT and TST agreement in cohort 1 (*see* Figure E1 in the online supplement) (18).

Among 5,357 adolescents who had QFT and TST performed at study baseline, 15% of participants with QFT IFN-γ values less than 0.2 IU/ml displayed a positive TST (TST+; induration, ≥5 mm; cutoff defined elsewhere [18]). By contrast, 53% of those with a negative QFT, but whose IFN-γ value fell between 0.2 and 0.34 IU/ml, were TST+. In QFT+ adolescents, the proportions of TST+ individuals were 66% among those with IFN-γ values between 0.35 and 0.7 IU/ml and 91% among those with IFN-γ values greater than 0.7 IU/ml (Figure 1). Overall, 43% of individuals with QFT IFN-γ values between 0.2 and 0.7 IU/ml had discordant QFT and TST results, whereas concordance between QFT and TST was observed in greater than 85% of participants with QFT IFN-γ values less than 0.2 or greater than 0.7 IU/ml.

Prognostic Significance of "Uncertain" QFT Conversion

We previously reported that recent QFT conversion, defined as a negative test (IFN-γ

Table 2. QuantiFERON-TB Gold In-Tube Conversion and Prospective Risk of Tuberculosis

QFT Class	TB Cases	n	Observation Years	Incidence (Cases/100 Person-Years)	95% Confidence Interval	P Value	IRR	95% Confidence Interval
Stringent nonconverters*	2	648	1,289.79	0.16	(0.02–0.56)	Reference	Reference	Reference
Stringent persistent positives†	19	989	1,953.07	0.97	(0.59–1.52)	0.005	6.27	(1.51–55.55)
Stringent converters‡	14	485	874.3	1.60	(0.88–2.69)	0.0003	10.33	(2.37–93.62)
"Uncertain" converters§	3	310	453.3	0.66	(0.14–1.95)	0.229	4.27	(0.49–51.10)

Definition of abbreviations: IRR = incidence rate ratio; QFT = QuantiFERON-TB Gold In-Tube; TB = tuberculosis.
*IFN-γ (TB Ag − Nil) less than 0.2 IU/ml at Days 0, 360, and 720.
†IFN-γ (TB Ag − Nil) greater than 0.7 IU/ml at Days 0, 360, and 720.
‡IFN-γ (TB Ag − Nil) less than 0.2 IU/ml at Day 0 and IFN-γ greater than 0.7 IU/ml at Day 360.
§IFN-γ (TB Ag − Nil) less than 0.35 IU/ml at Day 0 and IFN-γ greater than or equal to 0.35 IU/ml at Day 360, with at least one result within 0.2–0.7 IU/ml.

<0.35 IU/ml) followed by a positive test (IFN-γ ≥0.35 IU/ml), was associated with significantly increased risk of TB disease (1). In the present study, we aimed to determine how application of a more rigorous definition of QFT conversion, which accounts for uncertainty of IFN-γ values between 0.2 and 0.7 IU/ml, influenced prospective TB risk in cohort 1. Compared with stringent QFT nonconverters (who maintained IFN-γ <0.2 IU/ml at Days 0, 360, and 720), stringent QFT converters (IFN-γ <0.2 IU/ml at Day 0 and >0.7 IU/ml at Day 360) had a significantly higher risk of developing TB disease (Table 2). By contrast, the risk of TB disease in "uncertain" QFT converters (IFN-γ <0.35

IU/ml at Day 0 and ≥0.35 IU/ml at Day 360, with at least one result within the uncertainty zone) was not different from that of stringent QFT nonconverters. Similar results (i.e., significantly higher TB risk in stringent QFT converters but not in "uncertain" QFT converters) were obtained when applying the conventional cutoff of 0.35 IU/ml to define QFT nonconverters (data not shown) or when defining "uncertain" QFT converters as IFN-γ less than 0.2 IU/ml at Day 0 and IFN-γ 0.2–0.7 IU/ml at Day 360 (Table E1). These results suggest that "uncertain" converters include a substantial number of "false-positive" converters, potentially resulting from immunological and assay variability.

QFT Preanalytical Variability
An important source of variability in IFN-γ values derives from laboratory procedures. To reduce analytical variability of the QFT assay and improve consistency of longitudinal QFT testing, we investigated stricter SOP parameters (Table 1). First, we assessed the combined effects of blood volume and incubation time by incubating sets of QFT tubes (Nil, TB antigen [Ag], mitogen) from eight healthy donors in cohort 2 (four QFT− and four QFT+) containing 0.8 ml, 1 ml, or 1.2 ml of blood for 16, 20, or 24 hours each (Figure 2A). IFN-γ values (TB Ag minus Nil) appeared reproducible across the conditions tested. When comparing each condition against all others, the highest combined variability was
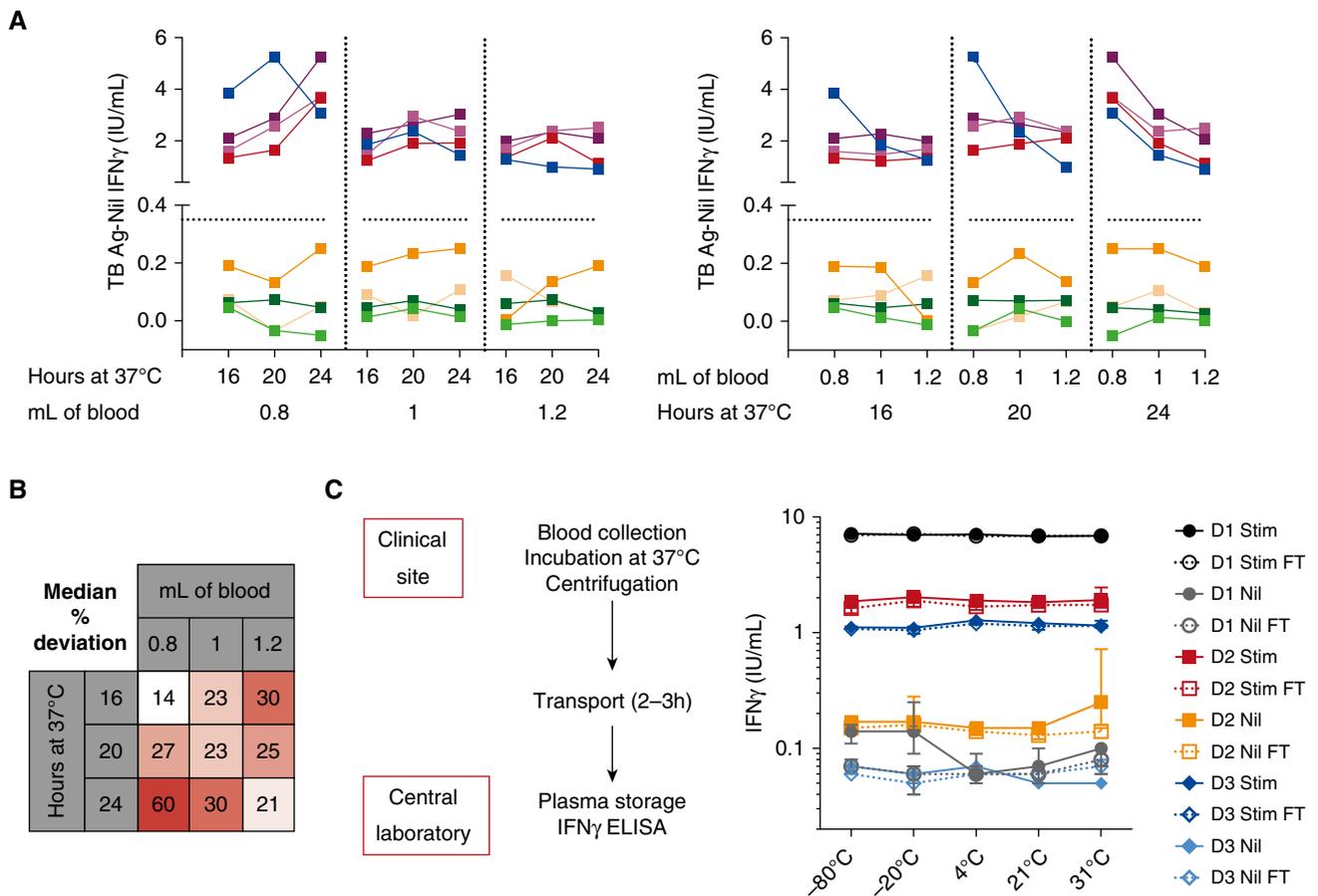

Figure 2. Preanalytical variability in QuantiFERON-TB Gold In-Tube (QFT) results relating to blood volume, incubation time, and plasma storage conditions. (A) Variability introduced by incubating different blood volumes (0.8, 1, and 1.2 ml) from eight healthy donors (color coded; from cohort 2) in QFT tubes at 37°C for 16, 20, or 24 hours. Dotted lines denote assay cutoff at IFN-γ values of 0.35 IU/ml. (B) Variability was calculated as median absolute percentage deviation for each participant, considering each condition (i.e., combination of blood volume and incubation time) versus all other conditions. Median absolute percent deviation across all participants is shown, with darker red identifying higher variability (color scale from low [white] to high [red]). (C) Comparison of IFN-γ values in plasma stored at a range of temperatures. Plasma from unstimulated (Nil) or stimulated (Stim) whole blood from healthy donors (n = 3; from cohort 2) was immediately cryopreserved at −80°C or −20°C, kept at 4°C, or incubated at 21°C or 31°C for 3 hours prior to storage at 4°C. IFN-γ values were measured the following day. To assess the impact of freezing/thawing (FT), all samples were then stored at −20°C for 2 weeks, thawed, and analyzed. Symbols denote median and range across triplicate ELISA wells. Ag = antigen; D = donor; TB = tuberculosis.

observed at lower blood volumes (0.8 ml) and longer incubation times (24 h) (Figure 2B). We therefore restricted the blood volume to be collected into Nil and TB Ag tubes to 1 ml and constrained the incubation time to 16–20 hours for all subsequent tests.

Next, we evaluated the effects of transport and storage temperatures on IFN-γ values in plasma samples following centrifugation (Figure 2C). Neither (1) plasma storage at 21°C, 31°C (ambient temperature), or 4°C for 3 hours (average transport time) prior to storage at 4°C nor (2) immediate cryopreservation of plasma at −20°C or −80°C had marked effects on IFN-γ values. Importantly, no effects of freeze thawing were observed (Figure 2C).

IFN-γ ELISA Standard Curve

QFT kits sold in the United States (and approved by the Food and Drug Administration) contain eight IFN-γ standards (ranging from 0 to 8 IU/ml), whereas the assay available in other countries contains only four IFN-γ standards (ranging from 0 to 4 IU/ml). To evaluate the effects of applying these different standard curves, we compared IFN-γ values extrapolated from four or eight IFN-γ standards for QFT assays, run in parallel on the same plate, on 20 healthy donors in cohort 2. Samples analyzed with an 8-point standard curve yielded significantly higher IFN-γ values than those calculated with a 4-point standard curve (Figure 3A). However, IFN-γ values were highly correlated and deviated from equivalence only at concentrations greater than 4 IU/ml (Figures 3B and E2).

QFT Analytical Variability

We also considered whether analyzing samples in single ELISA plate wells, as recommended by the manufacturer, rather than in triplicate was acceptable, because significant intraassay variability under identical conditions has been reported for QFT assays (5). In our hands, triplicate assays showed similar or lower variability than that reported in a published meta-analysis (Figures 3C and 3D and [5]). Triplicate measurements less than 0.8 IU/ml yielded a coefficient of variation below 30% (which dropped below 10% for higher values) and an SD below 0.15 IU/ml, with a few exceptions. On the basis of these results and the cost of the QFT ELISA, our protocol conformed to the manufacturer's
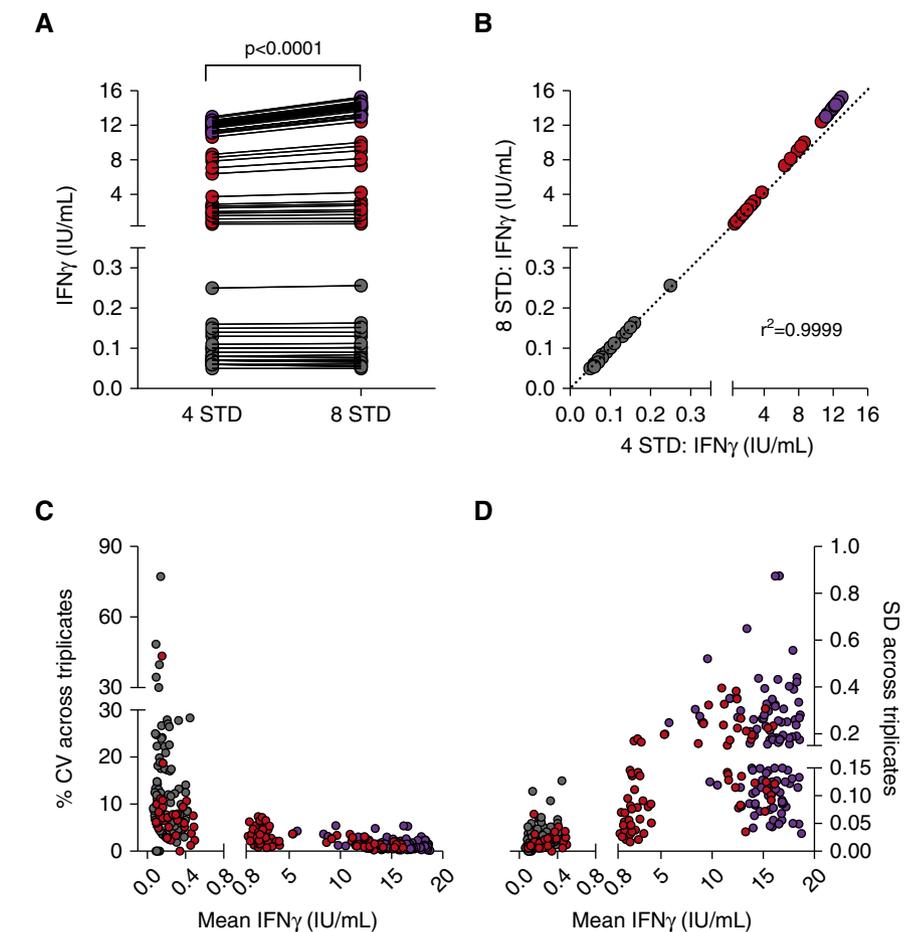

Figure 3. Analytical variability of QuantiFERON-TB Gold In-Tube (QFT) assay. (*A* and *B*) Comparison of IFN-γ values of 20 healthy donors in cohort 2, derived using two different sets of IFN-γ standards. IFN-γ standards were reconstituted and diluted according to different QFT package inserts to generate a 4-point standard curve (4 STD; 4, 1, 0.25, and 0 IU/ml) and an 8-point standard curve (8 STD; 8, 4, 2, 1, 0.5, 0.25, 0.125, and 0 IU/ml), which were used in parallel to calculate IFN-γ values for Nil (*gray dots*), tuberculosis antigen (*red dots*), and mitogen (*purple dots*) conditions from each sample (*A*). The *P* value was calculated using the Wilcoxon signed-rank test. (*B*) The *dotted line* is set at 45 degrees to allow visualization of deviation from equivalence. The r^2 value was calculated by linear regression. (*C* and *D*) Intraassay variability, depicted as (*C*) coefficient of variation (%CV) and (*D*) SD, calculated from triplicate ELISA wells, as a factor of mean IFN-γ concentration in 324 individual QFT samples from cohort 2.

recommendation to analyze samples in single wells.

QFT Interassay Variability

IFN-γ ELISA variability over time was monitored by a newly established IQA program. Analyses of a first set of IQC samples included in each ELISA run for 15 consecutive months revealed that the ELISA batch was a source of significant variability (Figure E3A). We then implemented a preassessment procedure in which new ELISA kit batches were tested before those QFT kits were procured, with consistent performance relative to previous

ELISAs (*see* footnote below Table 1 for acceptance criteria). A second set of IQC samples was then used to monitor 150 ELISA tests for over 18 months (Figure 4A). Overall, less than 15% variability was observed between assays and across operators and ELISA batches for IQC samples with medium and high IFN-γ values (0.52 IU/ml and 1.62 IU/ml, respectively). An SD of 0.05 IU/ml was observed for the low IQC sample (0.07 IU/ml). Using the same ELISA batch, interoperator variability was less than 15% for medium and high IQC samples and less than 30% for the low IQC (*see* Figure E3B).
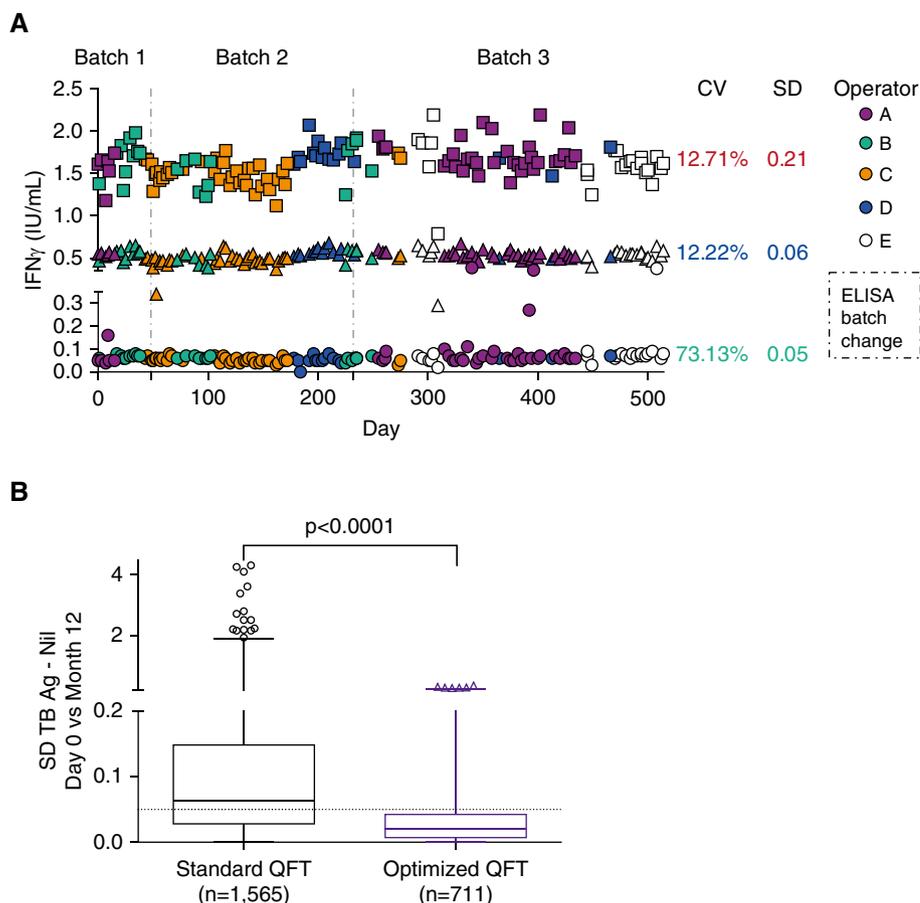
A

B

Figure 4. QuantiFERON-TB Gold In-Tube (QFT) interassay variability. (*A*) IFN-γ ELISA variability assessed by longitudinal measurements of three internal quality control (IQC) samples in 150 ELISA assays performed over 18 months by five different operators (indicated by the colors) using three different IFN-γ ELISA batches. Coefficient of variation (CV) and SD across all measurements were calculated for identical IQC plasma samples with low (*circles*, *green text*; mean, 0.07 IU/ml), medium (*triangles*, *blue text*; mean, 0.52 IU/ml), and high (*squares*, *red text*; mean, 1.62 IU/ml) IFN-γ values. (*B*) Variability between two sequential QFT assays, calculated as SD between IFN-γ values at Day 0 and Month 12 for each individual. QFT was performed according to standard manufacturer's recommendations in an infant cohort (described in Reference 2) or according to optimized procedures described in Table 1 for cohort 5 (adolescents enrolled in the prevention of *Mycobacterium tuberculosis* infection trial). All included individuals had negative QFT test results (IFN-γ <0.35 IU/ml) at both time points, did not undergo tuberculin skin testing, and were not diagnosed with incident tuberculosis (TB) disease. *Box-and-whisker plots* denote median, interquartile range, and 1st–99th percentiles, respectively. *P* value was calculated by Mann-Whitney *U* test. The *dotted line* denotes the SD observed in the lower IQC sample shown in (*A*), which represents the expected interassay variability solely due to the IFN-γ ELISA used. Ag = antigen.

Additional pass/fail criteria, adapted from the Westgard rules (21), were also implemented to allow real-time assessment of each ELISA run prior to the release of each assay's results (*see* footnote below Table 1). Implementation of optimized QFT procedures (shown in Table 1) yielded significantly lower interassay variability than that indicated by historical data (2) obtained using standard QFT procedures (Figure 4B).

Interpretation of QFT Values within the Uncertainty Zone

No gold standard for detection of asymptomatic (latent) *Mtb* infection exists. To determine if the QFT assay cutoff at 0.35 IU/ml is optimal in a TB-endemic population, we employed the optimized SOP described in Table 1 to measure IFN-γ values in healthy, *Mtb*-unexposed control subjects from Denmark (cohort 3; n = 50) and microbiologically confirmed TB cases

from South Africa (cohort 4; n = 68) (Figure 5A). The standard QFT cutoff of 0.35 IU/ml yielded 100% specificity and 91% sensitivity. Because all *Mtb*-unexposed control subjects had IFN-γ values less than 0.2 IU/ml, we explored whether a lower QFT cutoff would increase sensitivity. Applying a cutoff at 0.2 IU/ml, which coincides with the lower limit of the uncertainty zone, yielded 96% sensitivity, whereas the specificity remained at 100%. Interestingly, 9% of IFN-γ values from TB cases fell between 0.2 and 0.7 IU/ml, whereas 87% of patients with TB had IFN-γ greater than 0.7 IU/ml.

Finally, we implemented the optimized QFT SOP to screen healthy adolescents for the POI trial (cohort 5; n = 2,432). About 50% were QFT positive (Figure 5B). IFN-γ values were evenly distributed around the assay cutoff of 0.35 IU/ml, and approximately 7% of IFN-γ values fell into the uncertainty zone between 0.2 and 0.7 IU/ml. The IFN-γ values around this zone were continuously distributed with no natural inflection points, even when the ratio (instead of subtraction) of TB Ag and Nil was considered (Figure 5C).

Interpretation of Serial QFT Results

Remarkably, 5.3% of QFT-negative adolescents enrolled in cohort 5 converted to a positive QFT result when retested 3 months later (Figure 6A), whereas no incident TB occurred. More than half of these converters (56.3%) had one of two consecutive IFN-γ values within the uncertainty zone (0.2–0.7 IU/ml). When variability across these consecutive paired measurements was evaluated, a considerable overlap between some QFT nonconverters and "uncertain" QFT converters was noted (Figure 6B), suggesting that a substantial proportion of "uncertain" conversions results from immunological and technical variability around the assay cutoff. Of note, 49% percent of adolescents with an initial IFN-γ value within 0.2–0.34 IU/ml converted to a positive test, whereas only 3% with an initial IFN-γ value less than 0.2 IU/ml converted (Figure 6C). Among converters, adolescents with an initial IFN-γ value less than 0.2 IU/ml were more likely to convert to an IFN-γ value greater than 0.7 IU/ml (81%, Figure 6D) than those with an initial IFN-γ value within 0.2–0.34 IU/ml (18%; *P* < 0.0001).
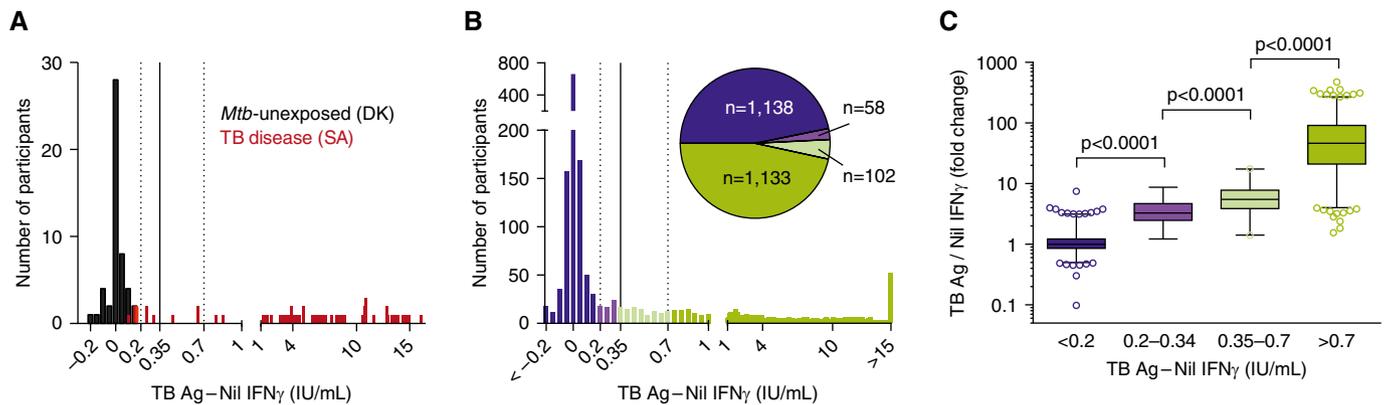
Figure 5. Distribution of IFN-γ values in populations with differential exposure to *Mycobacterium tuberculosis* (*Mtb*). (*A*) Distribution of IFN-γ values among *Mtb*-unexposed healthy Danish (DK) adults (cohort 3; *black*; n = 50) and South African (SA) patients with microbiologically confirmed tuberculosis (TB) disease (cohort 4; *red*; n = 68). (*B*) Distribution of IFN-γ values in adolescents screened for possible inclusion in the prevention of *Mtb* infection trial (cohort 5; n = 2,432) based on assay threshold (negative if TB Ag − Nil <0.35 IU/ml, positive if TB Ag − Nil ≥0.35 IU/ml) and uncertainty zone around the threshold (0.2 ≥ TB Ag − Nil ≤ 0.7 IU/ml). The pie chart represents the relative proportions (and numbers) of IFN-γ values falling into the four respective categories indicated by color codes in the histogram. In *A* and *B*, the *solid vertical lines* represent the assay positivity cutoff, and the *dotted vertical lines* represent the uncertainty zone around the threshold at 0.2 and 0.7 IU/ml. (*C*) Fold change in IFN-γ values between stimulated (TB Ag) and unstimulated (Nil) blood, stratified by the four categories of IFN-γ values in *B*. *Box-and-whisker plots* denote median, interquartile range, and 1st–99th percentiles, respectively. *P* values were calculated by Mann-Whitney *U* test. Ag = antigen.

Discussion

QFT conversion should identify individuals with recent *Mtb* infection who are at high risk of disease progression, allowing targeted intervention. However, interpretation of serial QFT tests is confounded by immunological and technical variability. In this paper, we highlight technical, immunological, and epidemiological factors that influence the interpretation and clinical significance of QFT assay conversions, supporting the establishment of an uncertainty zone between 0.2 and 0.7 IU/ml and a more rigorous definition of QFT conversion.

Stringent conversion from IFN-γ less than 0.2 to greater than 0.7 IU/ml was associated with a 10-fold higher incidence rate ratio of TB disease than stringent nonconversion with serial IFN-γ values less than 0.2 IU/ml over 2 years. By contrast, "uncertain" converters, with at least one of two serial test results within 0.2–0.7 IU/ml, were not at higher risk than nonconverters, although this analysis was powered to detect only substantially increased risks (greater than fivefold rate ratios). We previously showed that "uncertain" QFT conversions to IFN-γ values less than 0.7 IU/ml were associated with 50% probability of reversion (19). Whether QFT reverters experience lower incidence of TB disease than persistent QFT+ and sustained converters is still unclear (19), but it might

contribute to the lower risk of TB disease observed in "uncertain" converters. The data in our present study suggest that a substantial proportion of "uncertain" QFT conversions may result from immunological and technical variability around the assay cutoff rather than from *Mtb* infection.

To minimize assay variability for longitudinal testing, we assessed preanalytical and analytical sources of variation, including previously unreported effects of transport temperature, standard curve, and ELISA batch, and we adapted our procedures accordingly. We showed that a 4-point standard curve was acceptable to calculate IFN-γ values well above the assay cutoff of 0.35 IU/ml. We previously showed by serial plasma dilutions that, even when extrapolated from a 4-point standard curve, IFN-γ values are quantitative up to 12 IU/ml (2). At the time of implementation of our QFT IQA, we were not aware of any published guidelines about such programs. The IQA program proved invaluable for monitoring consistency of results and identifying sources of analytical variability, including continuous evaluation of trained operators and preassessment of IFN-γ ELISA batches. Taken together, our results support the implementation of IQA programs to monitor intralaboratory performance of QFT, and QFT control panel reagents are now commercially available. Implementation of optimized

QFT procedures, which are more stringent than the standard procedures recommended by the manufacturer, resulted in significant reduction of serial testing variability compared with historical data obtained by standard procedures.

In cohort 5 (POI trial), we observed that 5% of adolescents who were QFT− at screening converted to a positive test within 3 months, which extrapolates to an annual infection rate of 20%. This is considerably higher than the previous estimate of 14% (19). Overlap in short-term interassay variability between some nonconverters and "uncertain" converters was observed, suggesting that at least some conversions may have resulted from assay variability, despite implementation of optimized laboratory procedures. Indeed, approximately half of adolescents with baseline IFN-γ values between 0.2 and 0.34 IU/ml converted to greater than or equal to 0.35 IU/ml within 3 months, but 82% of these conversions were still within the uncertainty zone (IFN-γ <0.7 IU/ml). These results, together with the finding that 43% of individuals with IFN-γ values between 0.2 and 0.7 IU/ml had a concomitant discordant TST result in cohort 1, suggest that QFT values in this range may also reflect heterogeneous immunological sensitization due to *Mtb* infection. This interpretation is also supported by the higher rates of QFT conversion after TST administration
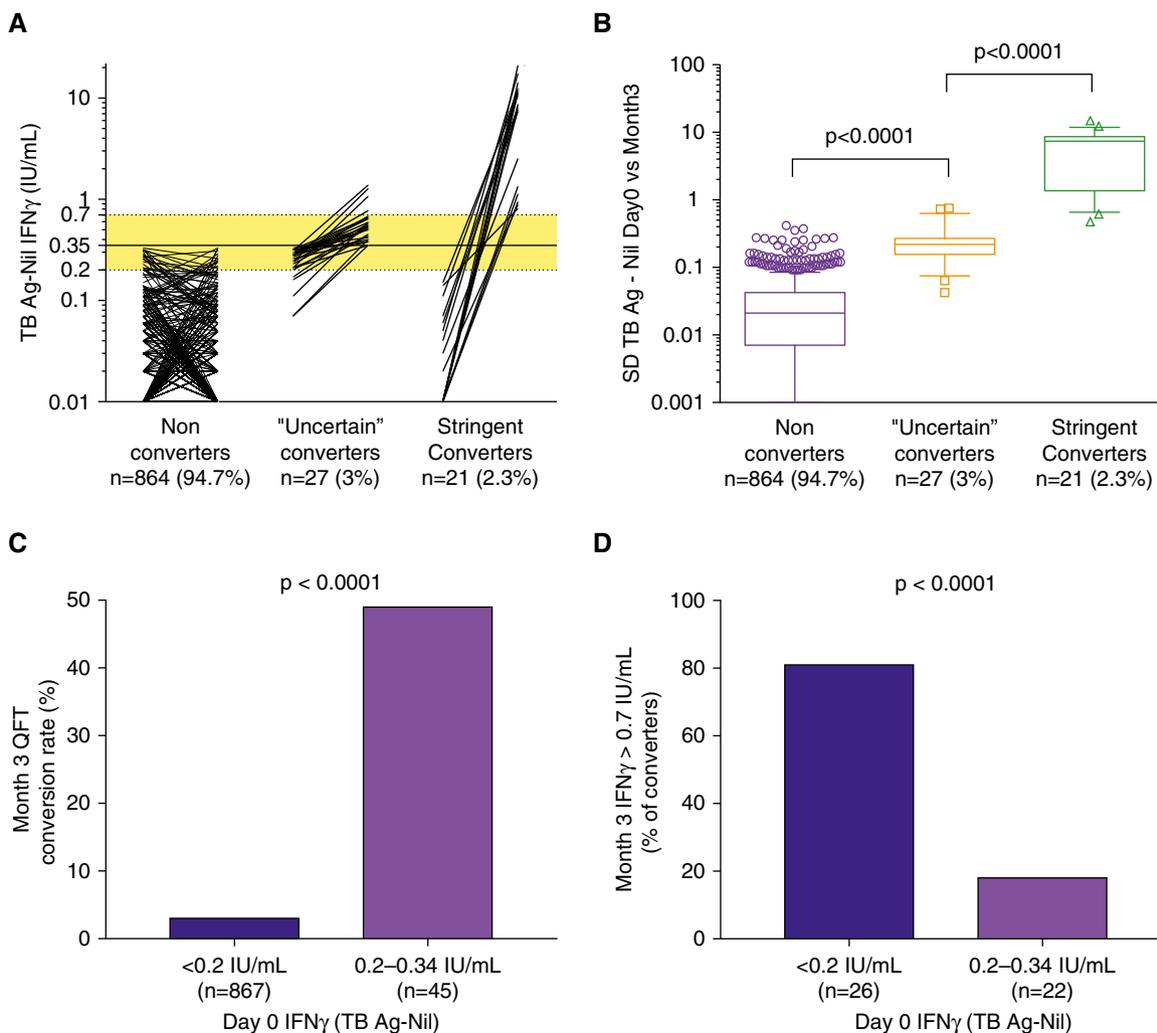
Figure 6. Interpretation of serial QuantiFERON-TB Gold In-Tube (QFT) assays. IFN-γ values were measured in adolescents who were QFT negative at screening for the prevention of *Mycobacterium tuberculosis* infection trial (cohort 5) and retested 3 months thereafter (n = 912) to assess conversion rates in an endemic population. (*A*) Paired IFN-γ values, from *left* to *right*, for adolescents who did not convert (IFN-γ <0.35 IU/ml at both time points), those who converted with one of the two measurements falling in the uncertainty zone ("uncertain" converters), and those who converted from below to above the uncertainty zone (stringent converters). *Dotted horizontal lines* and *yellow shading* denote the uncertainty zone (0.2–0.7 IU/ml), and *solid line* denotes the assay cutoff (0.35 IU/ml). Zeros and negative values have been set to 0.01 to allow visualization on a logarithmic scale. (*B*) Variability across serial QFT assays, calculated as SD between IFN-γ values at Day 0 and Month 3 for each individual. *Box-and-whisker plots* denote median, interquartile range, and 10th–90th percentiles, respectively. *P* values were calculated by Mann-Whitney *U* test. (*C*) Month 3 conversion rates, stratified by IFN-γ values at Day 0. (*D*) Rates of converters with Month 3 IFN-γ values greater than 0.7 IU/ml, stratified by IFN-γ values at Day 0. In *C* and *D*, *P* values were calculated by Fisher's exact test. Ag = antigen; TB = tuberculosis.

reported in individuals with HIV infection who had IFN-γ values between 0.15 and 0.34 IU/ml, compared with those with lower IFN-γ values (22). Taken together, these data suggest that a substantial proportion of *Mtb*-sensitized individuals display a weak immune response falling in the lower range of the QFT uncertainty zone.

Because the QFT assay measures *Mtb*-specific immune responses that are typically not bimodally distributed (23) and are heterogeneous within and across populations, interpreting quantitative measurements close to the cutoff as evidence for *Mtb* infection is fraught with uncertainty (24, 25). In this paper, we provide evidence that IFN-γ values within the zone of uncertainty are consistent with both weak immune responses to *Mtb* and technical variability. These possible causes cannot be further teased apart, and failure to take this uncertainty zone into account when defining QFT conversion will

inevitably lead to misclassifications. CDC guidelines for interpreting T-SPOT.*TB* (Oxford Immunotec, Marlborough, MA) test results already include criteria for interpreting borderline results, indicative of uncertain *Mtb* infection (14). Our study supports the introduction of an uncertainty zone for QFT.

The epidemiological data that provide the basis for our interpretation of serial QFT results are subject to some limitations, which may restrict the generalizability of our

findings. First, cohort 1 was one of the first large longitudinal studies assessing the performance of serial QFT in settings with a high TB burden. At that time, there was little information about sources of QFT variability, and therefore the assay was performed according to manufacturer's instructions and not with the optimized SOP used to generate the other results presented in this paper. Further, despite performing serial QFT in about 3,000 adolescents over 2 years, this sample size was still limited to accurately calculate the prospective risk of TB disease once participants were stratified by different QFT conversion definitions. Second, *Mtb*-unexposed Danish negative control subjects (cohort 3) were recruited in a setting with low TB burden and may have lower baseline IFN-γ levels than South African volunteers, owing to other genetic and/or environmental factors. Third, although patients with microbiologically confirmed TB are certainly infected with *Mtb*, they do not represent an optimal gold standard for asymptomatic *Mtb* infection. Finally, participants in the POI trial (cohort 5) were vaccinated with saline placebo, bacillus Calmette-Guérin, or H4:IC31 (1:1:1) between baseline and following QFT tests. Because we are currently blinded to treatment allocation, we cannot exclude the possibility that vaccination may have influenced either QFT conversion rates or IFN-γ values at conversion, although neither of the investigational products contains antigens included in the QFT assay.

This work was undertaken to implement serial QFT testing to measure efficacy of novel TB vaccines to prevent *Mtb* infection. However, the results are relevant to other clinical and research settings where longitudinal assay consistency is essential. Because the risk of incident TB disease was

highest for individuals with QFT conversion values from below to above the uncertainty zone (0.2–0.7 IU/ml), and because results within 0.2–0.7 IU/ml reflect weak immune responses and/or technical variability, we propose that stringent QFT conversions should be considered those with an initial IFN-γ value less than 0.2 IU/ml followed by an IFN-γ value greater than 0.7 IU/ml. This more rigorous QFT conversion definition should be applied in clinical trials as an exploratory endpoint to define acquisition of *Mtb* infection. Applicability of this more rigorous QFT conversion definition requires evaluation in independent studies in low– and high–TB-endemic settings to identify individuals at high risk of TB disease for targeted preventive treatment. Because the QFT-Plus assay, newly approved by the Food and Drug Administration, is in most respects identical to the QFT (Gold-In-Tube), we believe that our findings will likely also be generally applicable to the QFT-Plus assay. However, this requires experimental verification, and the boundaries for the uncertainty zone will need to be established.

Individuals who convert according to the manufacturer's specifications but do not meet this conversion definition should be closely followed and retested at a later date to confirm sustained QFT conversion. CDC guidelines for borderline IFN-γ release assay results could be applied to one-time "uncertain" QFT results (i.e., repeating IFN-γ release assay or performing a TST [14]). ∎

Members of the C-040-404 Study Team: Susan Rossouw, Carolyn Jones, Elisma Schoeman, Yolande Gregg, Elizabeth Beyers, Sandra Kruger, Helen Veltdsman, Sophie Keffers, Sandra Goliath, Mariana Mullins, Michele Tameris, Angelique Luabeya, Ashley Veldsman, Humphrey Mulenga, Angelique Hendricks, Fajwa Opperman, Elma Van Rooyen, Julia Noble, Samentra Braaf, Rose Ockhuis, Emerencia Vermeulen, Alessandro Companie, Xoliswa Kelepu, Maigan Ratangee, Abraham Pretorius, Henry Issel, Phumzile Langata, Ilse Davids, Roxanne Herling, Hadn Africa, Marcia Steyn, Lungisa Nkantso, Noncedo Xoyana, Bongani Diamond, Margareth Erasmus, Jane Hughes, Denise van der Westhuizen, Lydia Makunzi, Natasja Botes, Julia Amsterdam, Clive Maqubela, Portia Dlakavu, Pamela Mangala, Charmaine Abrahams, Petrus Tyambetyu, Diann Gempies, Cindy Elbring, Elizabeth Hamilton, Fadia Alexander, Sindile Wiseman Matiwane, Cashwin September, Christel Petersen, Yulande Herselman, Johanna Hector, Terence Esterhuizen, Lauren Mactavie, Elize van der Riet, Debbie Pretorius, Carolyn Jones, Justin Shenje, Anne Swarts, Eunice Sinandile, Janelle Botes, Terence Esterhuizen, Constance Schreuder, Jateel Kassiem, Onke Xasa, Boitumelo Mosito, Rodney Raphela, Denis Arendsen, Palesa Dolo, and Elizabeth Filander (all affiliated with the South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular Medicine and Division of Immunology, Department of Pathology, University of Cape Town, Cape Town, South Africa).

Members of the Adolescent Cohort Study Team: Hassan Mahomed, Fazlin Kafaar, Leslie Workman, Humphrey Mulenga, Rodney Ehrlich, Sizulu Moyo, Sebastian Gelderbloem, Miche Tameris, and Gregory Hussey (all affiliated with the South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular Medicine, and Division of Immunology, Department of Pathology, University of Cape Town, Cape Town, South Africa).

References

1. Machingaidze S, Verver S, Mulenga H, Abrahams DA, Hatherill M, Hanekom W, Hussey GD, Mahomed H. Predictive value of recent QuantiFERON conversion for tuberculosis disease in adolescents. *Am J Respir Crit Care Med* 2012;186:1051–1056.
2. Andrews JR, Nemes E, Tameris M, Landry BS, Mahomed H, McClain JB, Fletcher HA, Hanekom WA, Wood R, McShane H, *et al*. Serial QuantiFERON testing and tuberculosis disease risk among young children: an observational cohort study. *Lancet Respir Med* 2017;5: 282–290.
3. World Health Organization (WHO). Guidelines on the management of latent tuberculosis infection. Geneva, Switzerland: WHO; 2015.
4. Pai M, Denkinger CM, Kik SV, Rangaka MX, Zwerling A, Oxlade O, Metcalfe JZ, Cattamanchi A, Dowdy DW, Dheda K, *et al*. Gamma interferon release assays for detection of *Mycobacterium tuberculosis* infection. *Clin Microbiol Rev* 2014;27:3–20.
5. Tagmouti S, Slater M, Benedetti A, Kik SV, Banaei N, Cattamanchi A, Metcalfe J, Dowdy D, van Zyl Smit R, Dendukuri N, *et al*. Reproducibility of interferon gamma (IFN-γ) release assays: a systematic review. *Ann Am Thorac Soc* 2014;11:1267–1276.
6. Joshi M, Monson TP, Joshi A, Woods GL. IFN-γ release assay conversions and reversions: challenges with serial testing in U.S. health care workers. *Ann Am Thorac Soc* 2014;11:296–302.
7. Zwerling A, Joshi R, Kalantri SP, Dakshinamoorthy G, Reddy MV, Benedetti A, Schwartzman K, Menzies D, Pai M. Trajectories of tuberculosis-specific interferon-gamma release assay responses

among medical and nursing students in rural India. *J Epidemiol Glob Health* 2013;3:105–117.

8. Moses MW, Zwerling A, Cattamanchi A, Denkinger CM, Banaei N, Kik SV, Metcalfe J, Pai M, Dowdy D. Serial testing for latent tuberculosis using QuantiFERON-TB Gold In-Tube: a Markov model. *Sci Rep* 2016; 6:30781.

9. Pai M. Serial testing with TB interferon-γ release assays: toward a nuanced understanding. *Chest* 2012;142:1366–1368.

10. Zellweger JP, Sotgiu G, Block M, Dore S, Altet N, Blunschi R, Bogyi M, Bothamley G, Bothe C, Codecasa L, et al.; TBNET. Risk assessment of tuberculosis in contacts by IFN-γ release assays: a Tuberculosis Network European Trials Group study. *Am J Respir Crit Care Med* 2015;191:1176–1184.

11. Department of Health, Republic of South Africa. National tuberculosis management guidelines 2014. Pretoria, South Africa: Department of Health, Republic of South Africa; 2014.

12. Ellis RD, Hatherill M, Tait D, Snowden M, Churchyard G, Hanekom W, Evans T, Ginsberg AM. Innovative clinical trial designs to rationalize TB vaccine development. *Tuberculosis (Edinb)* 2015;95:352–357.

13. Hawn TR, Day TA, Scriba TJ, Hatherill M, Hanekom WA, Evans TG, Churchyard GJ, Kublin JG, Bekker LG, Self SG. Tuberculosis vaccines and prevention of infection. *Microbiol Mol Biol Rev* 2014; 78:650–671.

14. Mazurek GH, Jereb J, Vernon A, LoBue P, Goldberg S, Castro K. Updated guidelines for using interferon gamma release assays to detect *Mycobacterium tuberculosis* infection – United States, 2010. *MMWR Recomm Rep* 2010;59:1–25.

15. Metcalfe JZ, Cattamanchi A, McCulloch CE, Lew JD, Ha NP, Graviss EA. Test variability of the QuantiFERON-TB Gold In-Tube assay in clinical practice. *Am J Respir Crit Care Med* 2013;187:206–211.

16. van Zyl-Smit RN, Pai M, Peprah K, Meldau R, Kieck J, Juritz J, Badri M, Zumla A, Sechi LA, Bateman ED, et al. Within-subject variability and boosting of T-cell interferon-gamma responses after tuberculin skin testing. *Am J Respir Crit Care Med* 2009;180:49–58.

17. Pai M, Joshi R, Dogra S, Mendiratta DK, Narang P, Kalantri S, Reingold AL, Colford JM Jr, Riley LW, Menzies D. Serial testing of health care workers for tuberculosis using interferon-γ assay. *Am J Respir Crit Care Med* 2006;174:349–355.

18. Mahomed H, Hawkridge T, Verver S, Geiter L, Hatherill M, Abrahams DA, Ehrlich R, Hanekom WA, Hussey GD; SATVI Adolescent Study Team. Predictive factors for latent tuberculosis infection among adolescents in a high-burden area in South Africa. *Int J Tuberc Lung Dis* 2011;15:331–336.

19. Andrews JR, Hatherill M, Mahomed H, Hanekom WA, Campo M, Hawn TR, Wood R, Scriba TJ. The dynamics of QuantiFERON-TB Gold In-Tube conversion and reversion in a cohort of South African adolescents. *Am J Respir Crit Care Med* 2015;191:584–591.

20. Mahomed H, Ehrlich R, Hawkridge T, Hatherill M, Geiter L, Kafaar F, Abrahams DA, Mulenga H, Tameris M, Geldenhuys H, et al. Screening for TB in high school adolescents in a high burden setting in South Africa. *Tuberculosis (Edinb)* 2013;93: 357–362.

21. Mugan K, Carlson IH, Westgard JO. Planning QC procedures for immunoassays. *J Clin Immunoassay* 1994;17:216–222.

22. Esmail H, Thienemann F, Oni T, Goliath R, Wilkinson KA, Wilkinson RJ. QuantiFERON conversion following tuberculin administration is common in HIV infection and relates to baseline response. *BMC Infect Dis* 2016;16:545.

23. Penn-Nicholson A, Nemes E, Hanekom WA, Hatherill M, Scriba TJ. *Mycobacterium tuberculosis*-specific CD4 T cells are the principal source of IFN-γ in QuantiFERON assays in healthy persons. *Tuberculosis (Edinb)* 2015;95:350–351.

24. Woo KS, Choi JL, Kim BR, Han JY, Kim JM, Kim KH. Repeatability of QuantiFERON-TB Gold In-Tube assay results near cut-off points. *Ann Lab Med* 2016;36:76–78.

25. Banaei N, Gaur RL, Pai M. Interferon gamma release assays for latent tuberculosis: what are the sources of variability? J Clin Microbiol 2016;54:845–850.

26. Gaur RL, Banaei N. Inoculation of QuantiFERON-TB tubes with skin microbiota causes false-positive results. *Am J Respir Crit Care Med* 2014;190:834–837.

27. Gaur RL, Pai M, Banaei N. Impact of blood volume, tube shaking, and incubation time on reproducibility of QuantiFERON-TB Gold In-Tube assay. *J Clin Microbiol* 2013;51:3521–3526.

28. Doberne D, Gaur RL, Banaei N. Preanalytical delay reduces sensitivity of QuantiFERON-TB Gold In-Tube assay for detection of latent tuberculosis infection. *J Clin Microbiol* 2011;49: 3061–3064.

29. Herrera V, Yeh E, Murphy K, Parsonnet J, Banaei N. Immediate incubation reduces indeterminate results for QuantiFERON-TB Gold In-Tube assay. *J Clin Microbiol* 2010;48: 2672–2676.

30. Shanaube K, De Haas P, Schaap A, Moyo M, Kosloff B, Devendra A, Raby E, Godfrey-Faussett P, Ayles H. Intra-assay reliability and robustness of QuantiFERON®-TB Gold In-Tube test in Zambia. *Int J Tuberc Lung Dis* 2010;14:828–833.

31. Detjen AK, Loebenberg L, Grewal HMS, Stanley K, Gutschmidt A, Kruger C, Du Plessis N, Kidd M, Beyers N, Walzl G, et al. Short-term reproducibility of a commercial interferon gamma release assay. *Clin Vaccine Immunol* 2009;16:1170–1175.

32. Whitworth WC, Hamilton LR, Goodwin DJ, Barrera C, West KB, Racster L, Daniels LJ, Chuke SO, Campbell BH, Bohanon J, et al. Within-subject interlaboratory variability of QuantiFERON-TB Gold In-Tube tests. *PLoS One* 2012;7:e43790.