

Reproducibility of Interferon Gamma (IFN- γ) Release Assays

A Systematic Review

Saloua Tagmouti¹, Madeline Slater², Andrea Benedetti^{3,4}, Sandra V. Kik^{3,4}, Niaz Banaei⁵, Adithya Cattamanchi⁶, John Metcalfe⁶, David Dowdy⁷, Richard van Zyl Smit⁸, Nandini Dendukuri⁹, Madhukar Pai^{3,4}, and Claudia Denkinge^{3,10}

¹Division of Respiriology, ³McGill International Tuberculosis Centre and Department of Epidemiology, Biostatistics, and Occupational Health, and ⁹Division of Clinical Epidemiology, McGill University, Montreal, Quebec, Canada; ²Division of Infectious Diseases and Geographic Medicine, Department of Medicine, and ⁵Department of Medicine and Pathology, Stanford University School of Medicine, Stanford, California; ⁴Respiratory Epidemiology and Clinical Research Unit, Montreal Chest Institute, Montreal, Quebec, Canada; ⁶Division of Pulmonary and Critical Care Medicine and Curry International Tuberculosis Center, San Francisco General Hospital, University of California San Francisco, San Francisco, California; ⁷Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland; ⁸Division of Pulmonology and University of Cape Town Lung Institute, University of Cape Town, South Africa; and ¹⁰Division of Infectious Disease, Beth Israel Deaconess Medical Center, Boston, Massachusetts

Abstract

Rationale: Interferon gamma (IFN- γ) release assays for latent tuberculosis infection result in a larger-than-expected number of conversions and reversions in occupational screening programs, and reproducibility of test results is a concern.

Objectives: Knowledge of the relative contribution and extent of the individual sources of variability (immunological, preanalytical, or analytical) could help optimize testing protocols.

Methods: We performed a systematic review of studies published by October 2013 on all potential sources of variability of commercial IFN- γ release assays (QuantiFERON-TB Gold In-Tube and T-SPOT.TB). The included studies assessed test variability under identical conditions and under different conditions (the latter both overall and stratified by individual sources of variability). Linear mixed effects models were used to estimate within-subject SD.

Measurements and Main Results: We identified a total of 26 articles, including 7 studies analyzing variability under the same

conditions, 10 studies analyzing variability with repeat testing over time under different conditions, and 19 studies reporting individual sources of variability. Most data were on QuantiFERON (only three studies on T-SPOT.TB). A considerable number of conversions and reversions were seen around the manufacturer-recommended cut-point. The estimated range of variability of IFN- γ response in QuantiFERON under identical conditions was ± 0.47 IU/ml (coefficient of variation, 13%) and ± 0.26 IU/ml (30%) for individuals with an initial IFN- γ response in the borderline range (0.25–0.80 IU/ml). The estimated range of variability in noncontrolled settings was substantially larger (± 1.4 IU/ml; 60%). Blood volume inoculated into QuantiFERON tubes and preanalytic delay were identified as key sources of variability.

Conclusions: This systematic review shows substantial variability with repeat IFN- γ release assays testing even under identical conditions, suggesting that reversions and conversions around the existing cut-point should be interpreted with caution.

Keywords: tuberculosis; IFN- γ release assays; diagnostics; reproducibility

(Received in original form May 3, 2014; accepted in final form July 2, 2014)

Supported by the Canadian Institutes of Health Research grant CIHR MOP 130330; the European and Developing Countries Clinical Trials Partnership EDCTP-TBNEAT grant (M.P.), the Fonds de recherche du Québec – Santé (FRQS) (M.P.), a Chercheur Boursier salary award from the FRSQ (N.D.), a Richard Tomlinson Fellowship at McGill University (C.M.D.), a fellowship of the Burroughs-Wellcome Fund from the American Society of Tropical Medicine and Hygiene (C.M.D.), and the U.S. National Institutes of Health career development awards K23 HL HL094141 (A.C.) and K23 AI 094251 (J.M.).

The sponsor of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The views expressed by the authors do not reflect the official position of the institutions to which the authors are affiliated.

Author Contributions: All authors contributed to the study design and the content of the manuscript and approved the manuscript version submitted for publication. S.T., M.S., and C.D. contributed to the extraction and analyses of data. M.P. provided funding and overall supervision.

Correspondence and requests for reprints should be addressed to Claudia Denkinge, M.D., Ph.D., FIND, Chemins des Mines 9, 1202 Geneva, Switzerland. E-mail: claudia.denkinge@finddiagnostics.org

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Ann Am Thorac Soc Vol 11, No 8, pp 1267–1276, Oct 2014

Copyright © 2014 by the American Thoracic Society

DOI: 10.1513/AnnalsATS.201405-188OC

Internet address: www.atsjournals.org

Screening and treatment of latent tuberculosis infection is a key component of tuberculosis control, particularly in low-burden settings. Traditionally, the tuberculin skin test has been used for latent tuberculosis testing, but IFN- γ release assays (IGRAs) such as the QuantiFERON-TB Gold in-tube (QuantiFERON; Cellestis Limited/Qiagen, Valencia, CA), and the T-SPOT.TB (T-SPOT; Oxford Immunotec, Abingdon, UK) are more recently developed alternatives with both strengths and limitations (1). IGRAs are considered to be more specific than the tuberculin skin test. However, there is uncertainty about the interpretation of results when IGRAs are performed serially in subgroups with very low annual risk of tuberculosis infection, including in health-care workers. Conversions (change from negative to positive test results) and reversions (change from positive to negative test results) have been reported to occur more commonly for IGRAs than skin test (3–8). Uncertainty in the interpretation of repeat results has also been reported for the skin test. Biologic variability in response as well as differences in administration/reading contribute to a variability of ± 6 mm in the skin test induration in 95% of subjects (9). This has resulted in the recommendation for a conversion to be diagnosed only if a greater than 6 mm (in Canada) and 10 mm (in the United States) increase in induration is observed, over the baseline value, along with a change from negative to positive.

Currently, variability in IGRA results is not incorporated into the interpretation of serial test results. In contrast to the skin test, where a substantial increase in induration over a baseline value is necessary for a conversion, the IGRA manufacturers (and guidelines from the CDC) recommend a single cut-point to define a positive test as well as define conversions (10, 11).

A previous systematic review from 2009 (12) concluded that although reproducibility data were scarce, there was significant within-person IGRA variability. Since 2009, many new studies have emerged on overall variability of tests under identical and routine conditions as well as varying individual factors that contribute to variability. Therefore, we systematically reviewed the existing evidence on IGRA variability.

Methods

We developed a systematic review protocol, building on the previous review (12), before commencing the current review, following standard guidelines (13–15).

Search Strategy

We searched MEDLINE, EMBASE, Biosis, and Web of Knowledge for citations of all original articles and abstracts published from January 2005 until October 2013. We reviewed reference lists of included articles and identified review articles. In addition, we contacted experts to identify additional studies and abstracts. We restricted our search to studies with at least an abstract in English or French. The search terms included combinations of: tuberculosis, reproducibility, variability, repeatability, IFN- γ release assays (*see* online supplement for further details).

Inclusion Criteria

We included studies that repeated the IGRA and reported on any of the possible sources of variability (e.g., immunological, preanalytical, and analytical sources). Studies were included when they (1) assessed 10 or more subjects, (2) included immunocompetent subjects only, (3) were done in a low tuberculosis-prevalent setting (<25 cases/100,000 population) and included subjects without a known tuberculosis exposure or were done in a high tuberculosis-prevalence setting and repeat tests were done within 1-month interval to avoid possible confounding of the test result through new tuberculosis exposure, and (4) used one of the two currently available commercial assays (QuantiFERON and T-SPOT). We excluded studies that conducted the IGRA as part of a contact investigation or outbreak investigation or repeated IGRAs during antituberculosis treatment (for latent or active tuberculosis).

Study Selection and Data Extraction

Two review authors independently assessed articles for inclusion, extracted data on study methodology and characteristics, and test results (qualitative and quantitative). We resolved any discrepancies by consulting a third review author. We contacted authors to obtain quantitative IGRA data for subjects relevant to the review and for information not reported in the published studies.

Assessment of Methodological Quality

We assessed study quality using a modified quality appraisal tool for studies of diagnostic reliability (QAREL), a tool that has been developed for the assessment of reproducibility studies (16) and found to be reliable (17). Key components of the quality assessment were blinding, the explanation of missing data, and reporting of erroneous results. The quality assessment questions are available in the online supplement.

Categorization of Studies

We categorized studies into those that assessed:

- Variability under identical conditions: repeat testing with the same IGRA assay on the same sample under identical conditions (e.g., same preincubation delay, incubation time, and operator) to assess the isolated impact of analytical sources on variability (i.e., variability inherent to the test).
- Variability due to changing one or more components at a time: on the same sample from the same individual varying one or more components at a time to assess their impact on variability. We assume the impact of uncontrolled sources of variability equally affects the different study groups and therefore does not impact the overall findings.
- Variability over time (under different conditions): from the same individual a second sample is taken within 4 weeks. For repeat testing, variations in testing within the recommendations specified in the package insert are expected to occur (e.g., blood volume, preincubation delay, different ELISA operators). Some studies only repeated testing on positive baseline results.
- Boosting due to skin test: on the same individual one IGRA is done before and a second within a short time frame (100 d) after a skin test is administered. Only patients with negative skin test were included.

Outcomes and Analysis

We listed all identified sources of IGRA variability for which supporting data were available from at least one published study or abstract. We excluded uninterpretable test results (i.e., indeterminate or erroneous) from the analyses. Results were classified according to the manufacturer-recommended

cut-point on repeat testing in the different study categories, and kappa statistic were calculated for agreement between repeat tests (18). We assessed the number of conversions and reversions within predefined strata of the quantitative baseline value. If more than one repeat measurement was done within one subject, the repeat test result was always compared with the immediately preceding result except in the assessment of skin test boosting, where we compared all repeat test results to the results before skin test.

Assessment of quantitative values was only pursued for QuantiFERON, given the limited number of studies that assessed T-SPOT. We calculated three summary measures of test variability for both overall IFN- γ response and for IFN- γ response between 0.25 and 0.80 IU/ml: (1) the coefficient of variation (CV), defined as the SD of the differences in within-person measurements divided by the grand mean, multiplied by 100 to give a percentage. The lower the coefficient of variation (closer to zero), the more consistent are the measurements; (2) the intraclass correlation coefficient (ICC), defined as the ratio of variance between measurements to the total variance observed in a repeated measurements (within-subject variance/total variance); the higher the ICC (closer to 1), the more likely that a test gives consistent repeat measurements within the same individual (low within-subject variance), relative to differences in measurements observed across subjects (high between-subject variance); and (3) the “normal expected range of variability” (i.e., ± 1.96 SD; this is the range around an individual’s mean value that would include 95% of repeat measurements for that individual) (19). We constructed a Bland-Altman plot to analyze agreement between repeat measurements (20).

The ICC and the normal expected range were assessed using a linear mixed-effects model (incorporating subject effects as random effects and, where appropriate, study effects as a fixed effect given the small number of studies) fit to the numerical IFN- γ response values. This model assumes that whereas IFN- γ response values vary by subject and by study, the within-subject variability is the same for all subjects and all studies. This allowed the global estimation of within-subject variability by calculating the SD weighted for the correlation structure of the repeated measurements (21). Separate linear mixed-effects models

were fit within strata defined by baseline IFN- γ values. As per the manufacturer’s suggestion, IFN- γ values greater than 10 were truncated at 10 IU/ml, and negative values were rescaled to zero. Analyses were performed using STATA 12 (STATA Corporation, College Station, TX).

The Canadian Institutes of Health Research, funder of the study, had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Characteristics and Quality of Included Studies

Our literature search identified a total of 2,562 citations (Figure 1). Twenty-six studies fulfilled our inclusion criteria. All studies used QuantiFERON, and three studies evaluated T-SPOT as well. Quantitative values of individual test results were available from authors for 19 of 26 studies that evaluated QuantiFERON and were partially available from 1 study. Characteristics of included studies are shown in Tables 1 and 2. Overall, the quality of the selected studies was very good (Figure 2).

QuantiFERON Variability under Identical Conditions

A total of seven studies were included (6, 22–27). Individual test data were available for five studies with a total of 739 subjects, with each subject providing two test results obtained in identical settings. In total, 6.2% (46 of 739; 95% confidence interval [CI], 4.6–8.2%) of results were discrepant when repeated. The kappa statistic for agreement between repeat measurements was 0.88 (95% CI, 0.80–0.95). There was a higher variability around the cut-point; 26% (14 of 55) of samples converted if the baseline IFN- γ results ranged between 0.25 and 0.34 IU/ml, and 18% (17 of 94) of samples reverted if baseline IFN- γ results were between 0.35 and 0.8 IU/ml (Tables 3 and 4).

Analysis of the quantitative test results showed a normal expected range of variability around an individual mean of ± 0.47 IU/ml (95% CI, 0.45–0.50; CV, 13%) for all values (irrespective of the initial IFN- γ value) and ± 0.26 IU/ml (95% CI, 0.23–0.29; CV, 30%) for individuals with an initial IFN- γ response in the borderline range (0.25–0.80 IU/ml [Table 5]; see

Figure E1 in the online supplement, Bland-Altman plot). The variability was largest for high (>1.5 IU/ml) baseline IFN- γ values (normal expected range of variability ± 0.79 IU/ml). Overall, the ICC was high (0.99) despite the high rate of conversions and reversions seen. In the two studies for which the quantitative data were not available from authors, an overall discordant rate of 4.8% on 145 subjects was reported in one study (27) and no discordance in the other (35 subjects) (22).

QuantiFERON Variability under Different Conditions

A total of 10 studies were included (4–6, 22, 24, 25, 28–31). Quantitative data from the authors were available for eight studies, with data on a total of 932 subjects. Variability between measurements was substantially higher when the QuantiFERON was repeated within 4 weeks compared with when it was done on an aliquot from the same sample (as above). Overall 3.3% (8 of 243; 95% CI, 1.4–6.4%) of samples showed conversions and 44.4% (363 of 818; 95% CI, 40.9–47.9%) showed reversions. The kappa value was 0.27 (95% CI, 0.22–0.31). Again, variability was especially pronounced around the cut-off. Fourteen percent (1 of 7; 95% CI, 0–58%) of subjects converted on repeat testing if the baseline IFN- γ results ranged between 0.25 and 0.34 IU/ml and 57% (239 of 420; 95% CI, 52.0–61.7%) of subjects had a reversion if the baseline IFN- γ results were between 0.35 and 0.8 IU/ml (Table 3). However, reversions occurred even with higher baseline IFN- γ results: 33.3% (7 of 27; 95% CI, 11.1–46.3%) and 14% (9 of 64; 95% CI, 6.6–25.0%) when baseline results ranged between 5 to 6.9 and 7 to 10 IU/ml, respectively (Table 3; Figure E2, Bland-Altman plot).

Based on an analysis of the quantitative test results, we estimated that the normal expected range on repeat testing for all subjects is ± 1.4 IU/ml (95% CI, 1.30–1.42; CV, 60%) and ± 0.70 IU/ml (95% CI, 0.66–0.75; CV, 84%) for subjects with initial QuantiFERON responses in the borderline range (0.25–0.80 IU/ml) (Table 5). Overall, repeat results were more likely to be lower than initial results (overall average mean difference: -0.18 ; 95% CI, -0.23 to -0.13), which partially explained the large number of reversions with retesting. To assess whether this was due to a regression-to-the-mean phenomenon, an analysis was

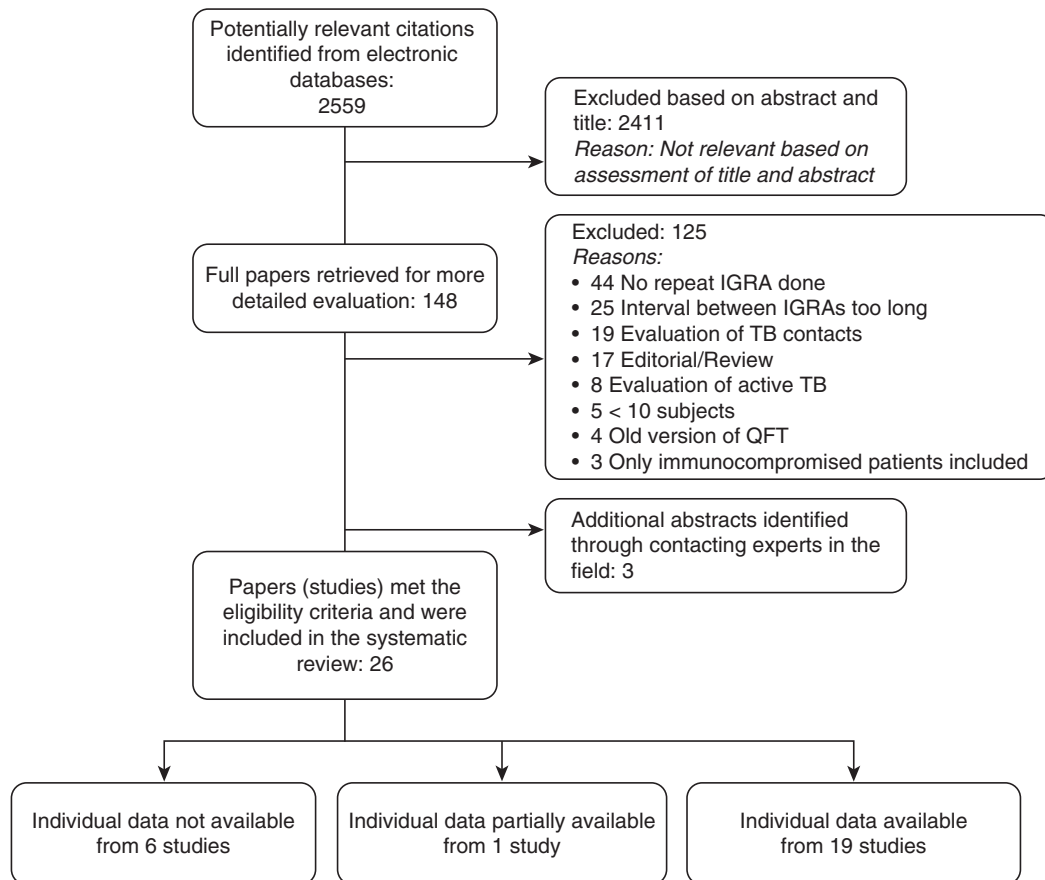


Figure 1. Flow chart of the study selection process. The literature search identified a total of 2,559 citations. After contacting experts in the field, three additional abstracts were added. Twenty-six studies fulfilled our inclusion criteria. Quantitative results for individual test results were available from authors for 19 of 26 studies that evaluated QuantiFERON and were partially available from 1 study. IGRA = IFN- γ release assay; TB = tuberculosis.

done excluding studies that only repeated tests when the initial result was positive (4, 5, 31). In this subanalysis, the total number of reversions decreased to 5.5% (19 of 344) and the kappa statistic for agreement between the baseline and first repeat measurement increased to 0.75 (95% CI, 0.61–0.88), suggesting that a regression-to-the-mean phenomenon was present. Overall, the ICC was slightly reduced (0.89 if all studies were included and 0.90 if studies retesting only positive results were excluded) compared with the ICC in testing under identical conditions (0.99).

In the two studies where the quantitative data were not available, one study showed a discordance rate of 6% (158 subjects) (29); the second study did not report change of QuantiFERON over time (28).

QuantiFERON Variability by Varying Individual Sources of Variability

Different preanalytical or analytical components of the test or immunological

host factors identified as potential sources that possibly contributed to overall test variability were evaluated. Variability in the delay before incubation and the blood volume included in the test tubes had the most impact on the QuantiFERON results.

Two studies with quantitative data examined preincubation delay: 6-hour (148 subjects) and 12-hour delay (168 subjects) compared with immediate incubation (all within the specifications of the manufacturers) (32, 33). A large number of reversions were observed with delayed incubation: 50% (3 of 6; 95% CI, 12–88%) and 67% (4 of 6; 95% CI, 22–96%) for 6-hour and 12-hour delay, respectively, when immediate incubation results were in the range of 0.35 to 0.8 IU/ml (Table 3). Analysis of the quantitative test results with a linear mixed model demonstrated that every 6-hour delay in incubation compared with immediate incubation decreased the IFN- γ response by 0.24 IU/ml (95% CI, 0.15–0.33). Furthermore, a delay in

incubation resulted in an overall increased variability (normal expected range, ± 1.62 IU/ml; 95% CI, 1.50–1.76) compared with what was observed in studies that repeated testing under identical conditions (± 0.47 ; 95% CI, 0.45–0.50) (Table 6). One additional study, for which individual test data were not available, also found significantly lower IFN- γ responses with longer incubation delays (11–12 h vs. <1 h) (27).

The impact of varying blood volumes in test tubes (0.8 ml, 1 ml, and 1.2 ml; all three within the range allowed by the manufacturer) was assessed in one study on 50 subjects for whom individual test data were available (34). An increased blood volume in the QuantiFERON tube resulted in a substantial decrease of the IFN- γ response measured (0.11 for every 0.2-ml increase in blood volume; 95% CI, 0.04–0.18). Variability in blood volume also contributed to an increased overall variability beyond what was observed in

Table 1. Characteristics of studies for which additional data were obtained from the authors

Author/yr (Reference)	Country/Prevalence	Participants	Variable Assessed	Type of IGRA	No. of Subjects	Time Points (d)
Slater/2013 (31)	USA/low	Health-care workers	Repeat over time; only positive results repeated*	QFT	537	0, 1–30
Joshi/2012 (5)	USA/low	Health-care workers	Repeat over time; only positive results repeated*	QFT	45	0, 2–30
Ringshausen/2011 (22) [†]	Germany/low	Health-care workers	Repeat over time	QFT	35	0, 7, 14, 21, 28
				T-SPOT	35	0, 7, 14, 21, 28
Gandra/2010 (4)	USA/low	Health-care workers	Repeat over time; only positive results repeated*	QFT	135	0, within 28 d
Ritz/2011 (41)	Australia/low	Students	Boosting by skin test	QFT	15	0, 42, 70
Baker/2009 (40)	USA/low	Immigrants	Boosting by skin test	QFT	60	0, 14–185
Cummings/2009 (44)	USA/low	Health-care workers	Boosting by skin test	QFT	64	0, 9
Leyten/2007 (42)	Netherlands/low	Volunteers	Boosting by skin test	QFT	14	0, 3
Doberne/2011 (32)	USA/low	Volunteers	Preatalytical delay (0, 6, 12 h)	QFT	128 (6 h), 127 (8 h)	0
				QFT	20 (6 h), 41 (12 h)	0
Herrera/2010 (33)	USA/low	Volunteers/health-care workers	Preatalytical delay (0, 6, 12 h)	QFT	363	0
Metcalfe/2013 (23)	USA/low	Petrochemical workers	Same conditions	QFT	19	0
Scott/2010 (26) [‡]	South Africa/high	Students/health-care workers	Same conditions	QFT	14	0
Shanaube/2010 (37)	Zambia/high	Volunteers	Effect of storage temperature 37°C vs. < 25°C	QFT	14	0
Hang/2009 (35)	Viet Nam/high	Volunteers	Between-operators variability	QFT	23	0
Min/2013 (36)	South Korea/medium	Volunteers	Incubation time (24, 48, 72 h)	QFT	33	0
Dorman/2014 (6)	USA/low	Health-care workers	Repeat over time	QFT	137	0, 5–21
				T-SPOT	122	0, 5–21
				QFT	139	0
				T-SPOT	125	0
				QFT	66	0, 7–21
				T-SPOT	63	0, 7–21
Gaur/2013 (34)	USA/low	Volunteers	ELISA rerun	QFT	245	0, 8 (0–10)
				QFT	50	0
				QFT	50	0
van Zyl-Smit/2009 (30)	South Africa/high	Health-care workers	Incubation time (16, 20, 24 h)	QFT	40	0
				QFT	40	0
				QFT	40	0
Detjen/2009 (24)	South Africa/high	Health-care workers	Shaking (gentle vs. vigorous)	QFT	19	0
				QFT	19	0
Veerapathran/2008 (25)	India/high	Health-care workers	Repeated reading (0, 2.5, 5, 30, 60, 90, 120 min)	QFT	14	0, 7, 14
				QFT	14	0, 7, 14
Detjen/2009 (24)	South Africa/high	Health-care workers	Repeat over time	QFT	15	0, 3
				QFT	27	0
Veerapathran/2008 (25)	India/high	Health-care workers	Same conditions (6 duplicates)	QFT	21	0
				QFT	19	0
Veerapathran/2008 (25)	India/high	Health-care workers	Plasma storage Fresh vs. 4°C vs. –80°C	QFT	19	0
				QFT	19	0
Veerapathran/2008 (25)	India/high	Health-care workers	Between-operator variability	QFT	14	0, 3, 9, 12
				QFT	14	0

Definition of abbreviations: IGRA = IFN-γ release assay; QFT = QuantiFERON-TB Gold In-Tube, T-SPOT = T-SPOT.TB.

*Subjects with indeterminate results removed.

[†]Quantitative data only available for a part of this study.

[‡]Study published as abstract or conference presentation.

Table 2. Characteristics of studies for which available data used as reported in the publication

Author/yr (Reference)	Country/Prevalence	Subjects	Variable Assessed	Type of IGRA	No. of Subjects	Time Points (d)	Study Results Summary
Whitworth/2014 (39)	USA/low	Air Force and CDC staff	Automated vs. manual ELISA	QFT	146	0	7 (4.8%) Subjects had discordant automated interpretations and 10 (6.9%) in manual interpretations ($P = 0.17$).
Whitworth/2012 (38)	USA/low	Air Force and CDC staff	Variation between 3 laboratories	QFT	91	0	7 (7.7%) Subjects were discordant; 6 of them had QFT responses within 0.25 IU/ml of the cut-off.
Sauzullo/2011 (43)	Italy/low	Health-care workers	Boosting by skin test	QFT	69	0, 7, 14, 28, 42	Conversion in 1 subject in Day 14 with subsequent reversion.
Powell/2011 (28)	USA/low	Volunteers	Repeat over time	QFT	864	0, 14– 84	Study did report change of QFT results only for unusual IFN- γ measurements and did not assess change of IGRA in repeated testing.
Whitworth/2012 (27)*	USA/low	Air Force and CDC staff	Incubation delay (<1 h vs.11–12 h)	QFT	149	0	Discordant rates of 6.8%. Significantly lower QFT responses with longer incubation delays ($P = 0.001$)
			Incubation time (23–24 vs.16–17 h)	QFT	152	0	Discordant rates of 5.3%. Significantly lower QFT responses were seen with shorter incubation times ($P = 0.002$).
			Incubation temperature (35°C vs. 37°C)	QFT	102	0	Discordant rates of 7.8%.
Mazurek/2012 (29)*	USA/low	Unclear	Same conditions	QFT	145	0	Discordant rates of 4.8%.
			Blood collection time: A.M. vs. P.M.	QFT	146	0	Discordance in 8.9% of subjects.
Ringshausen/2011 (22)†	Germany/low	Health-care workers	Repeat over time	QFT	149	0, 6	Discordance in 6% of subjects.
			Same conditions	QFT	35	0	There were no conversions or reversions.
			Repeated reading (0, 5, 15, 30, 60 min)	QFT	35	0	There were no conversions or reversions.

Definition of abbreviations: IGRA = IFN- γ release assay; QFT = QuantiFERON-TB Gold In-Tube.

*Study published as abstract or conference presentation.

†Quantitative data only available for a part of this study.

studies that repeated testing under identical conditions (normal expected range, ± 1.38 IU/ml; 95% CI, 1.20–1.56) (Table 6).

All other variables identified in this systematic review appeared to have no significant effect on qualitative

QuantiFERON results, but some components, when insufficiently standardized, caused a larger variability on repeat testing. Whether

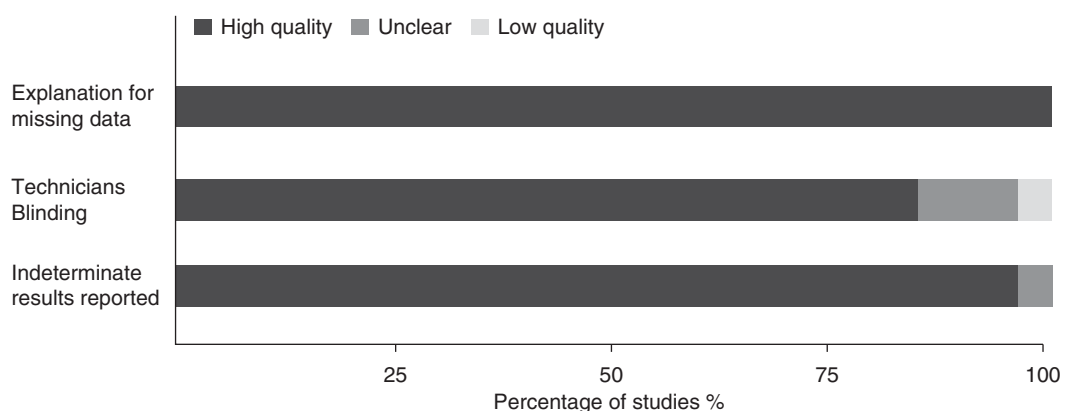


Figure 2. Quality of included studies. The overall quality of the selected studies as assessed by the modified quality appraisal tool for studies of diagnostic reliability (QAREL) tool (16).

Table 3. Variability in semiquantitative results around manufacturer-defined cut-point: results for QuantiFERON-TB Gold In-Tube assay

Variable Assessed	No. of Studies (Total No. of Samples)	Baseline Tuberculosis-Specific IFN- γ (IU/ml)								
		0–0.1	0.11–0.24	0.25–0.34	0.35–0.8	0.81–1.49	1.5–2.9	3–4.99	5–6.99	7–10
		% Conversions (Total N)				% Reversions (Total N)				
Retesting under identical conditions	5 (739)	2 (255)	6 (73)	26 (55)	18 (94)	2 (62)	7 (63)	0 (27)	0 (16)	0 (94)
Repeat testing within 4 wk	8 (1,061)	2 (205)	10 (31)	14 (7)	57 (420)	40 (158)	34 (106)	23.3 (43)	33.3 (27)	14 (64)
Repeat ELISA within 2 h	1 (114)	0 (72)	0 (12)	—	0 (18)	0 (6)	0 (6)	—	—	—
Incubation delay 12 h vs. immediate	2 (168)	2 (122)	0 (8)	33(3)	67 (6)	38 (8)	33 (6)	20 (5)	0 (1)	0 (9)
6 h vs. immediate	2 (148)	1(106)	0 (8)	33(3)	50 (6)	38 (8)	25 (4)	20 (5)	0 (1)	0 (7)
Interoperator variability	2 (42)	0 (21)	0 (4)	—	0 (2)	25 (4)	0 (2)	0 (1)	0 (1)	0 (7)
Plasma storage 4°C vs. 0°C	1 (288)	0 (155)	0 (13)	0 (2)	36 (72)	27 (22)	15 (13)	—	0 (5)	0 (6)
–80°C vs. 0°C	1 (21)	0 (6)	0 (3)	—	0 (2)	25 (4)	0 (1)	—	0 (1)	0 (4)
Boosting by negative skin test	5 (219)	4(193)	10 (10)	33 (3)	43 (7)	20 (5)	0 (1)	—	—	—

tubes were shaken gently or vigorously appeared to affect variability (normal expected range, ± 1.90 IU/ml; 95% CI, 1.53–2.36) compared with what was observed in studies that retested under identical conditions (Table 6) (34). Having two different operators also increased the variability (normal expected range, 1.39 IU/ml; 95% CI, 1.12–1.72) based on 42 comparisons in two studies (24, 35). In addition, including a storage step for plasma (at -80°C or 4°C) contributed to an increase in variability: normal expected range, ± 0.66 IU/ml (95% CI, 0.61–0.72) if a sample was frozen at -80°C (done on 21 subjects) and ± 0.83 IU/ml (95% CI, 0.55–0.73) if a sample was kept at 4°C (done on 267 subjects) compared with fresh plasma processing (Table 6).

Other possible sources of variability appeared to have no impact on QuantiFERON results: variation of incubation duration (comparing 20 h and 24 h to a baseline of 16 h [34], and 48 h and 72 h to 24 h [36]), or storage temperature of tubes before testing (37°C for 3 months compared with $<25^{\circ}\text{C}$) (37). Simply repeating optical density measurements of the ELISA test result after a time interval (up to 2 h) did not result in any variability (22, 30).

Three other studies, for which quantitative data were not available, assessed additional sources of QuantiFERON variability: diurnal variability, interlaboratory variation, and variability comparing the use of an

automated versus manual ELISA (29, 38, 39). Results are summarized in Table 2.

QuantiFERON Boosting by Tuberculin Skin Test

QuantiFERON boosting after a negative skin test was assessed in six studies (6, 40–44). Individual test data were available from five studies on 219 subjects (6, 40–42, 44). Most baseline QuantiFERON results before skin test were negative (206), and on repeat testing after skin test only nine tests converted (4.4%). Out of 13 tests that were positive at baseline, 4 reverted on retesting (30.8%). The quantitative data analysis showed that there was no significant boosting (0.01; 95% CI, -0.004 to 0.02) and only minimal variability

Table 4. Variability in semiquantitative results around manufacturer-defined cut-point: results for T-SPOT.TB

Variable Assessed	No. of Studies (Total No. of Samples)	Baseline Spot-Forming Units								
		0–2	2.1–4	4.1–7.9*	8–10	10.1–14.9	15–29.9	30–49.9	50–69.9	70–100
		% Conversions (Total N)			% Reversions (Total N)					
Repeat testing within 4 wk	3 (270)	2 (173)	4 (26)	- (17)	15 (13)	33 (6)	8 (13)	0 (2)	14(7)	7.6(13)
Retesting under identical conditions	1 (125)	1 (84)	7 (14)	- (8)	0 (4)	40 (5)	0 (10)	—	—	—
Boosting by negative skin test	1 (63)	4 (56)	0 (5)	- (2)	—	—	—	—	—	—

*Borderline range as defined by manufacturer; results that fall into borderline range were not classified as reversions or conversions.

Table 5. Variability in quantitative results under the same and different conditions

	Analysis	No. of Valid Comparisons	Mean of Difference of Repeat Tests in IU/ml (95% CI)	Within-Subject SD of Variation in IU/ml (95% CI)	Coefficient of Variation (%)
Identical conditions, same sample	All values	739	0.01 (−0.01 to 0.04)	±0.24 (0.23–0.25)	13
	Group with baseline < 0.1	250	0.02 (0.01 to 0.03)	±0.06 (0.06–0.07)	263
	Group with baseline 0.1–0.25	77	0.02 (−0.03 to 0.08)	±0.17 (0.15–0.20)	99
	Group with baseline 0.25–0.8	148	−0.002 (−0.03 to 0.03)	±0.13 (0.12–0.15)	30
Repeat over times	All values	932	−0.18 (−0.23 to −0.13)	±0.69 (0.66–0.72)	60
	Group with baseline < 0.1	134	0.04 (0.01 to 0.07)	±0.31 (0.29–0.33)	588
	Group with baseline 0.1–0.25	21	0.01 (−0.09 to 0.1)	±0.28 (0.21–0.37)	151
	Group with baseline 0.25–0.8	418	−0.11 (−0.15 to −0.08)	±0.36 (0.34–0.38)	84

Data for all analyses included only the first and second valid comparisons.

(normal expected range, ±0.14; 95% CI, −0.13 to 0.16).

Variability of T-SPOT

A total of three studies used T-SPOT (6, 22, 30) and assessed variability under the same conditions (125 subjects) on repeat testing over time (270 subjects) and boosting by negative skin test (63 subjects). Percent of conversions and reversions are summarized in Table 4; variation of results on repeat testing were not reported if the baseline test was in the borderline range (4.1–7.9 spot-forming units).

Discussion

Many guidelines now recommend the use of IGRAs interchangeably with the skin test or

suggest that IGRAs can replace the skin test, including in serial testing (45). However, the studies reviewed here demonstrate that on repeat testing, IGRA conversions occur more frequently than what is epidemiologically expected, especially in settings with low tuberculosis incidence (6, 31) (<1%). Sample processing factors, such as delays in incubation and variations in blood volume, may be responsible for an important part of this variability (4–6, 22–26, 28–37, 40–44, 46).

Among people whose baseline QuantiFERON IFN-γ result was between 0.25 and 0.8 IU/ml, we estimated that on testing of an aliquot of the same sample, at the same time, a result would fall within ±0.26 IU/ml (95% CI, 0.23–0.29) of the

first value for 95% of samples. Thus, a substantial number of reversions and conversions could be expected in these individuals even under the most ideal conditions. As further variability is introduced through routine variations in phlebotomy, sample preparation, and test procedures, on repeat testing at a later time (within 4 wk) even greater variability could be expected for samples. For 95% of these samples the repeat result would fall within ±0.70 IU/ml (95% CI, 0.66–0.75%) of the baseline value (between 0.25–0.8 IU/ml). Also, reversions are even seen with very high initial values. This suggests that even the introduction of an uncertainty range, as has been suggested by various authors, may not be sufficient (47, 48).

Table 6. Variability in quantitative results within categories defined by individual sources of variability

Analysis	No. of Valid Comparisons	Mean of Difference of Repeat Tests in IU/ml (95% CI)	Within-Subject SD of Variation (95% CI)	Coefficient of Variation (%)
Pretest storage	14	−0.05 (−0.09 to 0.18)	±0.18 (0.13 to 0.27)	4
Preincubation delay (0, 6, 12 h)	168	−0.24 (−0.33 to −0.15)	±0.83 (0.77 to 0.90)	153
Incubation time (16, 20, 24 h)	100	−0.001 (−0.01 to 0.008)	±0.18 (0.15 to 0.20)	49
Incubation time (24, 48, 72 h)	66	0.0001 (−0.02 to 0.03)	±0.24 (0.20 to 0.29)	21
Interoperator variability	42	−0.05 (−0.35 to 0.26)	±0.71 (0.57 to 0.88)	34
Blood volume	100	−0.11 (−0.18 to −0.04)	±0.70 (0.61 to 0.81)	84
Plasma storage (fresh vs. −80°C)	21	−0.03 (−0.08 to 0.14)	±0.37 (0.28 to 0.47)	68
Plasma storage (fresh vs. 4°C)	288	0.04 (−0.01 to 0.10)	±0.34 (0.31 to 0.37)	57
Shaking	40	0.37 (−0.06 to 0.79)	±0.97 (0.78 to 1.21)	83
ELISA repeat reading (immediate vs. 2 h)	19	−0.004 (−0.01 to 0.01)	±0.02 (0.01 to 0.02)	6
Skin test boosting	219	0.01 (−0.004 to 0.02)	±0.07 (0.07 to 0.08)	190

A better understanding of both the frequency and impact of key factors that increase IGRA variability will lead to better interpretation rules for serial testing and knowledge of how best to standardize IGRA testing where such standardization is possible.

Our review identified blood volume and delay before incubation to be significant factors contributing to increased QuantiFERON variability. Although control of the delay before incubation is feasible with portable incubators, further standardization of blood volume may be more difficult. Differences in QuantiFERON tube shaking also contributed to variability, although the number of samples evaluated was small. Standardization and automation of key steps in the operating procedure could possibly reduce QuantiFERON variability and the rates of conversions and reversions and improve the overall performance of the test, especially if used in serial testing. However, the feasibility and cost of this standardization and automation need to be weighed against the added value of IGRAs over the skin test for serial testing, especially as the key advantage of an improved specificity of IGRAs over the skin test is probably minimal in the setting of serial testing when an exposure to bacillus Calmette-Guérin between repeat tests is unlikely. Although our review primarily assessed sources of variability for the QuantiFERON, due to the limited

availability of T-SPOT data, similar conclusions may apply to the ELISPOT technique.

Our results also put the results of a one-off IGRA test into question. As with the tuberculin skin test, it is important to interpret test results in the context of the patient's risk for tuberculosis, and, given the limitations of the test, only high-risk groups should be tested with tuberculin skin test or IGRA. With the incidence of TB reaching historic low levels in North America, most health-care workers might not be at high risk of TB exposure. Therefore, regardless of the test used, current policies on serial testing of all health-care workers should be reconsidered.

Strengths of our review include a standard protocol and strict inclusion criteria (to exclude possible tuberculosis exposures confounding the results), availability of quantitative data from most studies, and the use of a correlated data analysis. However, our review also has several limitations. We acknowledge that we may have missed studies despite the comprehensive search. We also did not include an assessment of variability due to sources in the manufacturing process, although there is evidence that QuantiFERON lot-to-lot variability may occur (49). In addition, low numbers of subjects tested for some of the variables limits the precision of our estimates, and

a component of between-subject variability might confound the estimates of within-subject variability. Furthermore, the persistent negative trend on repeat testing observed with high values could be a result of the truncation of high values at 10 IU/ml, a regression toward the mean because studies oversample people with positive QuantiFERON results or due to factors that are biasing the results that have not been accounted for.

In summary, this review demonstrated that a substantial number of conversions/reversions can be observed with repeat measurement of IGRAs due to factors inherent to the test that cannot be controlled. Furthermore, sample-processing factors, such as delay in incubation and variations in blood volume, add substantial additional variability. Our results may help with the derivation of alternative cut-offs for IGRA conversions and reversions, which account for the observed limitations in reproducibility. Further standardization of key contributors to variability may improve IGRA performance and optimize interpretation of test results. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Acknowledgment: The authors thank all the corresponding authors of primary studies who provided clarifications or shared additional data.

References

- 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.
- McMullen SE, Pegues DA, Shofer FS, Sheller AC, Wiener EB. Performance of QuantiFERON-TB Gold and tuberculin skin test relative to subjects' risk of exposure to tuberculosis. *Clin Infect Dis* 2014;58:1260–1266.
- Fong KS, Tomford JW, Teixeira L, Fraser TG, van Duin D, Yen-Lieberman B, Gordon SM, Miranda C. Challenges of interferon-gamma release assay conversions in serial testing of health-care workers in a TB control program. *Chest* 2012;142:55–62.
- Gandra S, Scott WS, Somaraju V, Wang H, Wilton S, Feigenbaum M. Questionable effectiveness of the QuantiFERON-TB Gold Test (Cellestis) as a screening tool in healthcare workers. *Infect Control Hosp Epidemiol* 2010;31:1279–1285.
- Joshi M, Monson TP, Woods GL. Use of interferon-gamma release assays in a health care worker screening program: experience from a tertiary care centre in the United States. *Can Respir J* 2012;19: 84–88.
- Dorman SE, Belknap R, Graviss EA, Reves R, Schluger N, Weinfurter P, Wang Y, Cronin W, Hirsch-Moverman Y, Teeter LD, *et al*. Interferon-gamma release assays and tuberculin skin testing for diagnosis of latent tuberculosis infection in healthcare workers in the united states. *Am J Respir Crit Care Med* 2014;189:77–87.
- Pai M, Kik SV, Banaei N. Occupational screening for tuberculosis: a testing time for interferon-gamma release assays. *Ann Am Thorac Soc* 2014;11:399–401.
- Joshi M, Monson TP, Joshi A, Woods GL. IFN-gamma release assay conversions and reversions: challenges with serial testing in US health care workers. *Ann Am Thorac Soc* 2014;11:296–302.
- Menzies D. Interpretation of repeated tuberculin tests: boosting, conversion, and reversion. *Am J Respir Crit Care Med* 1999;159: 15–21.
- Cellestis. Quantiferon-TB Gold (in-Tube Method) package insert. 2011, Nov 7 [accessed 2014 Sep 24]. Available from: <http://www.cellestis.com/irm/content/PI/QFT/2PK/AU.pdf>.
- 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.
- van Zyl-Smit RN, Zwerling A, Dheda K, Pai M. Within-subject variability of interferon-g assay results for tuberculosis and boosting effect of tuberculin skin testing: a systematic review. *PLoS ONE* 2009;4: e8517.
- Gatsonis C, Paliwal P. Meta-analysis of diagnostic and screening test accuracy evaluations: methodologic primer. *AJR Am J Roentgenol* 2006;187:271–281.
- Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM. Systematic reviews of diagnostic test accuracy. *Ann Intern Med* 2008;149:889–897.
- Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Analysing and presenting results. In: Deeks JJ, Bossuyt PM,

- Gatsonis C, editors. Cochrane handbook for systematic reviews of diagnostic test accuracy version: the Cochrane Collaboration; 2010.
- 16 Lucas NP, Macaskill P, Irwig L, Bogduk N. The development of a quality appraisal tool for studies of diagnostic reliability (QAREL). *J Clin Epidemiol* 2010;63:854–861.
 - 17 Lucas N, Macaskill P, Irwig L, Moran R, Rickards L, Turner R, Bogduk N. The reliability of a quality appraisal tool for studies of diagnostic reliability (QAREL). *BMC Med Res Methodol* 2013;13:111.
 - 18 Fleiss JL. Statistical methods for rates and proportions. New York: Wiley; 1981.
 - 19 American Society for Testing and Materials. E456 standard terminology relating to quality and statistics. 2012 [accessed 2014 Feb 13]. Available from: <http://www.Astm.Org/Standards/E456.htm>.
 - 20 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307–310.
 - 21 Diggle PJ. An approach to the analysis of repeated measurements. *Biometrics* 1988;44:959–971.
 - 22 Ringshausen FC, Nienhaus A, Torres Costa J, Knoop H, Schlosser S, Schultze-Werninghaus G, Rohde G. Within-subject variability of *Mycobacterium tuberculosis*-specific gamma interferon responses in German health care workers. *Clin Vaccine Immunol* 2011;18:1176–1182.
 - 23 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.
 - 24 Detjen AK, Loebenberg L, Grewal HM, 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.
 - 25 Veerapathran A, Joshi R, Goswami K, Dogra S, Moodie EE, Reddy MV, Kalantri S, Schwartzman K, Behr MA, Menzies D, et al. T-cell assays for tuberculosis infection: deriving cut-offs for conversions using reproducibility data. *PLoS One* 2008;3:e1850.
 - 26 Scott L. QuantiFERON-TB gold in-tube OD readings around the cut-off value: what do they mean? *Int J Tuberc Lung Dis* 2010;14:S31–S2.
 - 27 Whitworth WC, Mazurek G, Goodwin DJ. Assay parameters affecting variability Of QuantiFERON-TB Gold In-Tube assay results [abstract]. *Am J Respir Crit Care Med* 2012; 185:A4728.
 - 28 Powell RD III, Whitworth WC, Bernardo J, Moonan PK, Mazurek GH. Unusual interferon gamma measurements with QuantiFERON-TB Gold and QuantiFERON-TB Gold In-Tube tests. *PLoS ONE* 2011;6:e20061.
 - 29 Mazurek G, Whitworth WC, Goodwin DJ. Affect of blood collection time on QuantiFERON-TB Gold In-Tube test variability [abstract]. *Am J Respir Crit Care Med* 2012;185:A4735.
 - 30 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.
 - 31 Slater ML, Welland G, Pai M, Parsonnet J, Banaei N. Challenges with QuantiFERON-TB Gold assay for large-scale, routine screening of US healthcare workers. *Am J Respir Crit Care Med* 2013;188:1005–1010.
 - 32 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.
 - 33 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.
 - 34 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.
 - 35 Hang NT, Ishizuka N, Keicho N, Hong LT, Tam DB, Thu VT, Matsushita I, Harada N, Higuchi K, Sakurada S, et al. Quality assessment of an interferon-gamma release assay for tuberculosis infection in a resource-limited setting. *BMC Infect Dis* 2009;9:66.
 - 36 Min JW, Lee HY, Lee JS, Lee J, Chung JH, Han SK, Yim JJ. Effect of prolonged incubation time on results of the QuantiFERON TB gold in-tube assay for diagnosis of latent tuberculosis infection. *Clin Vaccine Immunol* 2013;20:1377–1380.
 - 37 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(R)-TB Gold In-Tube test in Zambia. *Int J Tuberc Lung Dis* 2010;14:828–833.
 - 38 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.
 - 39 Whitworth WC, Goodwin DJ, Racster L, West KB, Chuke SO, Daniels LJ, Campbell BH, Bohanon J, Jaffar AT, Drane W, et al. Variability of the QuantiFERON(R)-TB gold in-tube test using automated and manual methods. *PLoS ONE* 2014;9:e86721.
 - 40 Baker CA, Thomas W, Stauffer WM, Peterson PK, Tsukayama DT. Serial testing of refugees for latent tuberculosis using the QuantiFERON-gold in-tube: effects of an antecedent tuberculin skin test. *Am J Trop Med Hyg* 2009;80:628–633.
 - 41 Ritz N, Yau C, Connell TG, Tebruegge M, Leslie D, Curtis N. Absence of interferon-gamma release assay conversion following tuberculin skin testing. *Int J Tuberc Lung Dis* 2011;15:767–769.
 - 42 Leyten EM, Prins C, Bossink AW, Thijsen S, Ottenhoff TH, van Dissel JT, Arend SM. Effect of tuberculin skin testing on a *Mycobacterium tuberculosis*-specific interferon-gamma assay. *Eur Respir J* 2007;29:1212–1216.
 - 43 Sauzullo I, Massetti AP, Mengoni F, Rossi R, Lichtner M, Ajassa C, Vullo V, Mastroianni CM. Influence of previous tuberculin skin test on serial IFN-gamma release assays. *Tuberculosis (Edinb)* 2011;91:322–326.
 - 44 Cummings KJ, Smith TS, Shogren ES, Khakoo R, Nanda S, Bunner L, Smithmyer A, Soccorsi D, Kashon ML, Mazurek GH, et al. Prospective comparison of tuberculin skin test and QuantiFERON-TB Gold In-Tube assay for the detection of latent tuberculosis infection among healthcare workers in a low-incidence setting. *Infect Control Hosp Epidemiol* 2009;30:1123–1126.
 - 45 Denkinger CM, Dheda K, Pai M. Guidelines on interferon-gamma release assays for tuberculosis infection: concordance, discordance or confusion? *Clin Microbiol Infect* 2011;17:806–814.
 - 46 Whitworth WC, Mazurek GH, Goodwin DJ, Racster L, West K, Jaffar A, Daniels LJ, Campbell BH, Chuke SO. Within-subject inter-assay variability of QuantiFERON-TB Gold In-Tube assay results using automated and manual methods [abstract]. *Am J Respir Crit Care Med* 2011;183:A1192.
 - 47 Thanassi W, Noda A, Hernandez B, Newell J, Terpeluk P, Marder D, Yesavage JA. Delineating a retesting zone using receiver operating characteristic analysis on serial QuantiFERON tuberculosis test results in US healthcare workers. *Pulm Med* 2012;2012:291294.
 - 48 Pai M. Serial testing with TB interferon-gamma release assays: toward a nuanced understanding. *Chest* 2012;142:1366–1367.
 - 49 Slater M, Parsonnet J, Banaei N. Investigation of false-positive results given by the QuantiFERON-TB Gold In-Tube assay. *J Clin Microbiol* 2012;50:3105–3107.