Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world-wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

_____ _____

Emilia Ilieva-Hughes Date

The Tuberculin Skin Test: Within-Subject Variability, Boosting, and Comparison with the QuantiFERON-TB Gold In-Tube Test

By

Emilia Ilieva-Hughes

Degree to be Awarded: Master of Public Health
Program: Executive MPH Program

_____
Andrew Voetsch, PhD
Thesis Committee Chair

_____
Gerald Mazurek, MD
Thesis Field Advisor

_____
William C. Whitworth, MPH
Thesis Field Advisor

_____
Leah Tompkins, MS, M.Ed., CHES
EMPH Program

The Tuberculin Skin Test: Within-Subject Variability, Boosting, and Comparison with the QuantiFERON-TB Gold In-Tube Test

By

Emilia Ilieva-Hughes

DMD
Medical University of Sofia
1982

Thesis Committee Chair: Andrew Voetsch, PhD
Thesis Field Advisors: Gerald Mazurek, MD and William C. Whitworth, MPH

An abstract of
a thesis submitted to the faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in the Executive Master of Public Health Program,
2017

Abstract

The Tuberculin Skin Test:
Within-Subject Variability, Boosting, and Comparison with the QuantiFERON-TB Gold In-Tube Test

By

Emilia Ilieva-Hughes

Introduction/Rationale:
Tuberculin skin test (TST) and interferon-gamma release assays, such as the QuantiFERON®-TB Gold In-Tube Test (QFT-GIT), are used to detect *Mycobacterium tuberculosis* infection. Information on the relative variability of TST and QFT-GIT, and effect of tuberculin injection on subsequent test results is limited.

Methods:
To assess 1) within-subject variability of TST when given simultaneously in the right vs. left arm, 2) agreement between simultaneously performed TST and QFT-GIT, 3) effect of initial TST on subsequent TST when performed a week apart, and 4) effect of initial TST on QFT-GIT when performed a week apart, we enrolled healthy adults with a prior positive TST but no TST in 3 previous years. All testing was performed blindly by healthcare workers with documented proficiency. Paired analyses compared categorical test interpretations using a 10mm TST cutoff and a 0.35 IU/mL QFT-GIT cutoff. Significance in differences of proportions was assessed using McNemar's test.

Results:
There were 158 total subjects available for the analysis. Of those with analyzable results, 75/154 (49%), 80/155 (52%), and 31/149 (21%) were positive by initial TST on the right arm, initial TST on the left arm, and initial QFT-GIT, respectively. When repeated 1 week later, 72/124 (58%) TSTs were positive and 71/153 (46%) QFT-GITs were positive.

1) TSTs performed simultaneously in the right and left arm were discordant in 14% of subjects, while previous analyses of simultaneously performed QFT-GITs in this population demonstrated 5% discordance ($p<0.01$).

2) As compared to initial TST on the left arm, initial QFT-GIT results were discordant for 66 (45%) subjects. As compared to initial TST on the right arm, initial QFT-GIT results were discordant for 57 (39%) subjects.

3) Repeat TST was discordant with initial TST on the left arm for 34 (28%) subjects, with the majority of discordance (22%) due to TST conversion (i.e., negative to positive). Of 71 subjects with negative initial left arm TST, 27 (38%) converted to positive when TST was administered a week later. Repeat TST was discordant with initial TST on the right arm for 36 (29%) subjects, again with the majority of discordance (25%) due to TST conversion. Of 77 subjects with negative initial right arm TST, 31 (40%) converted to positive when TST was administered a week later.

4) Repeat QFT-GIT was discordant with initial QFT-GIT for 40 (27%) subjects with the majority of discordance (26.7%, all but one subject) due to conversion. Of 115 subjects with negative initial QFT-GIT prior to TST, 39 (34%) converted to positive when QFT-GIT was performed a week after TST.

5) The proportions of subjects with conversion (38% to 40% for TST, and 34% for QFT-GIT) were greater than within-subject variability when the tests were performed simultaneously (14% for TST and 5% for QFT-GIT, $p < 0.0001$).

Conclusions:
In a population with prior positive TSTs, TST was more variable than QFT-GIT when pairs of each test were performed simultaneously. TSTs may trigger conversion of subsequent TST and QFT-GIT.

The Tuberculin Skin Test: Within-Subject Variability, Boosting, and Comparison with the QuantiFERON-TB Gold In-Tube Test

By

Emilia Ilieva-Hughes

DMD
Medical University of Sofia
1982

Thesis Committee Chair: Andrew Voetsch, PhD
Thesis Field Advisors: Gerald Mazurek, MD and William C. Whitworth, MPH

A thesis submitted to the faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health in the Executive Master of Public Health Program,
2017

Acknowledgments

I would like to express my sincerest gratitude to my committee chair, Dr. Andrew Voetsch and my field advisors, Dr. Gerald Mazurek and Mr. William Whitworth. I would also like to express my gratitude to the Centers for Disease Control and Prevention, Division of Tuberculosis Elimination's Clinical Research Branch, with whom I had the honor and privilege of working and learning. I would also like to thank the students and faculty at Emory University and the Rollins School of Public Health – EMPH program for helping me and providing me valuable knowledge and skills. I would finally like to thank my family for their constant understanding, patience, and support.

The findings and conclusions in this document are solely those of the author and do not necessarily represent the official views of the Centers for Disease Control and Prevention, The U. S. Department of Defense, or the U. S. Air Force.

Table of Contents

Chapter I: Introduction

Infection and Disease Due to *Mycobacterium Tuberculosis*

Historical Perspective

Tuberculosis (TB) is an infectious disease that has ravaged humanity for ages and continues to kill millions of people each year. TB usually affects the lungs and is transmitted from human-to-human through inhalation. The bacterium that is primarily responsible for causing TB, *Mycobacterium tuberculosis* (MTB), was first identified and described in 1882 by Robert Koch. Infections with MTB (MTBI) can be classified as either an active disease (referred to as TB) with clinical symptoms and pathological signs, or as an asymptomatic non-contagious state, often called “latent” TB infection (LTBI). People with TB may present with constitutional symptoms, such as weight loss, fever, chills, night sweats, and weakness, but may also present with symptoms associated with disease in specific organs. The lung is most often affected as evidence by symptoms of cough, sputum production, hemoptysis, and chest pain, that are typically accompanied by an abnormal chest x-ray. Pulmonary TB accounts for approximately 80% of newly diagnosed TB in the U. S. while 20% to 30% involves extra-pulmonary sites. Traditional methods used to diagnose TB rely primarily on sputum smear microscopy, bacteriological culture, and clinical examination, although newer and rapid molecular tests are being used with increasing frequency. A combination of sputum culture and clinical exam are considered the gold standard for definitive TB diagnosis. TB is curable with a standard six month course of four antimicrobial agents (isoniazid, rifampicin,

pyrazinamide, and ethambutol), but, if left untreated, 70% of cases worldwide will not survive 10 years (World Health Organization, 2016).

Epidemiology

According to the latest estimates by the World Health Organization (WHO), 10.4 million people became ill with TB in 2015 (World Health Organization, 2016). Six countries account for 60% of the total global burden: India, Indonesia, China, Nigeria, Pakistan, and South Africa (World Health Organization, 2016). Of those with TB in 2015, 1.2 million (11%) were persons living with HIV (World Health Organization, 2016). Although not a direct cause and effect relationship, HIV is the strongest risk factor for TB, and TB-HIV coinfection is greatest in countries which have a high HIV burden. In 2015, the proportion of TB – HIV cases was highest in the WHO African Region (31%), and exceeded 50% in parts of southern Africa. Other significant TB risk factors are diabetes, alcohol abuse, poverty, crowded living conditions, inadequate nutrition, and poor indoor air quality. An estimated 27% of TB cases worldwide are attributable to inadequate nutrition and 22% are attributable to indoor air pollution (Lonnroth et al., 2010; Pai et al., 2016).

TB prevalence estimates from national surveys performed in high-burden countries (such as in Africa) are as high as 500 per 100,000 (World Health Organization, 2016). In contrast, TB prevalence in the U. S. is currently 3 cases per 100,000 (Centers for Disease Control and Prevention, 2017). TB incidence in high-burden countries is approximately 200 per 100,000 (World Health Organization, 2016). High-income countries, including most in western Europe,

Canada, the U. S., Australia, and New Zealand, have the lowest incidence of TB disease, typically less than 10 cases per 100,000 per year (Pai et al., 2016). TB incidence and prevalence also vary widely among different portions of the population within countries. For example, in contrast to the overall TB prevalence in the United States (3 cases per 100,000) the rate among Asians in the U. S. is 18 cases per 100,000 persons, approximately 30 times that of U. S. whites (0.6 cases per 100,000) (Centers for Disease Control and Prevention, 2017).

Accurate estimation of LTBI prevalence using currently available methods is difficult. It is estimated that 2 to 3 billion persons (approximately 33% worldwide) are latently infected with MTB (World Health Organization, 2016). The LTBI rate in the U. S. is much lower than the global rate, estimated to be about 4.5% in 2015, with most being attributable to foreign birth (Mancuso, Diffenderfer, Ghassemieh, Horne, & Kao, 2016).

Although TB is curable with antibiotics, it remains one of the world's most deadly diseases. Prior to 2014, deaths resulting from AIDS were greater than deaths resulting from TB. However, as of 2014, TB surpassed AIDS as a leading cause of cause of death worldwide (World Health Organization, 2016). The death toll from TB in 2015 was estimated at 1.4 million, whereas the death toll from HIV during the same period was estimated at 1.1 million (World Health Organization, 2016). Over 95% of deaths are in developing countries, and TB is a leading cause of death in HIV-positive persons (35% of HIV deaths in 2015) (World Health Organization, 2016). Perhaps the main modifiable behavioral risk factor

for TB mortality is smoking, to which more than 20% of deaths in TB cases worldwide is attributable (World Health Organization, 2016).

The financial burden of TB is also high. The estimated global resource requirement for a full response to the TB epidemic in low- and middle-income countries (LMIC, where TB is most prevalent) in 2016 was approximately $8.3 billion (U. S.), and the cost per patient treated is usually in the range of $100–$1000 (U. S.) for drug-susceptible TB and $2,000–$20,000 (U. S.) for multidrug-resistant (MDR) TB (World Health Organization, 2016).

Transmission

Transmission of TB from human-to-human (and subsequent infection in the new host) begins when “droplet nuclei” (about 1 – 5 µm in diameter) that are carrying MTB are aerosolized (usually by coughing) by a person with active pulmonary disease. Because of their small size, the droplets can penetrate deeply into the alveoli of uninfected individuals. It is estimated that fewer than 10 bacteria can cause infection (Nicas, Nazaroff, & Hubbard, 2005). In the alveoli, the bacteria are consumed by phagocytic immune cells, called alveolar macrophages (Ahmad, 2011). There are several factors that determine the likelihood of successful transmission. These include 1) immune status of the exposed individual, 2) infectiousness of the transmitting individual, 3) environmental factors such ventilation and humidity, and 4) exposure, such as proximity to the transmitting individual and frequency/duration of exposure (Centers for Disease Control and Prevention, 2013).

Mycobacterium Tuberculosis Infection

Exposure to *M. tuberculosis* leads to two general outcomes regarding the pathogen within the body: 1) elimination or 2) persistence (Pai et al., 2016). In many people, the bacteria are eliminated through non-specific innate mechanisms that are present before exposure. In others, adaptive immune responses (immune mechanisms that develop in response to a specific foreign substance) many eliminate or control the infection (Pai et al., 2016). For MTBI, the adaptive immune response begins when macrophages ingest the bacteria and present MTB components to naive lymphocytes. This results in a clonal proliferation of lymphocytes with affinity to specific MTB components. Development of this adaptive immune response typically controls the MTB infection and is evidenced by a positive tuberculin skin test (TST) or interferon-gamma (IFN-γ) release assay (IGRA). While the immune system controls MTB infection for most people, it does not eliminate the pathogen, resulting in the LTBI state. Even with initial control of the infection, conditions that disrupt the immune system (and for reasons not yet understood), latent infections may progress to active TB disease, a process referred to as reactivation. The lifetime risk of reactivation is estimated to be 5% to 10% (Horseburg 2004). The risk of reactivation is estimated to be 2.4% to 5% in the first 5 years after infection (Sloot, Schim van der Loeff MF, Kouw, & Borgdorff, 2014; Horseburg 2004), but the risk is higher among children, those co-infected with HIV, and some other groups, such as smokers and diabetics (American Thoracic Society & Centers for Disease Control and Prevention, 2000; Horsburgh, Jr., 2004; Centers for Disease

Control and Prevention, 2011a; World Health Organization, 2016; Shea, Kammerer, Winston, Navin, & Horsburgh, Jr., 2014). Estimates that 80% of TB cases arise from reactivation of LTBI have been confirmed by the use of genotyping (Pai & Behr, 2016).

Treatment and Control

TB is curable. A standard six-month course of four antimicrobial agents (isoniazid, rifampicin, pyrazinamide, and ethambutol) cures up to 95% of disease caused by susceptible MTB, with greater success among those receiving directly observed therapy (DOT) (Pai et al., 2016). However, if left untreated, 70% of cases will not survive 10 years (World Health Organization, 2016). Lower treatment success rates are observed among patients with MDR TB, which is defined as resistance to both isoniazid and rifampicin

LTBI can also be effectively treated. Effective treatment lowers the risk that people with LTBI will progress to TB. The following regimens are recommended by WHO for the treatment of LTBI (World Health Orgainzation, 2017):

- 6-month or 9-month isoniazid, daily
- 3-month rifapentine plus isoniazid, once weekly
- 3-month or 4-month isoniazid plus rifampicin, daily
- 3-month or 4-month rifampicin alone, daily

Of these regimens, the recently-approved 3-month rifapentine plus isoniazid given once weekly (“3-HP”) is especially attractive, as it reduces daily dosage to once per week, requires treatment for only three months (12 total doses), has a lower risk of hepatotoxicity, and higher treatment completion rates (Sterling et al., 2011; Sandul et al., 2017). This regimen has recently been recommended by the

Centers for Disease Control and Prevention (CDC) as an alternative to the standard nine months of daily isoniazid (Centers for Disease Control and Prevention, 2011b). The 3-HP regimen has also shown to be cost-effective vs. the standard regimen (Shepardson & MacKenzie, 2014; Shepardson et al., 2013).

The diagnosis of LTBI is important because of the large percentage of TB cases attributable to LTBI reactivation. Preventive treatment of persons diagnosed with LTBI can reduce the risk of subsequent tuberculosis by as much as 93% (Pape, Jean, Ho, Hafner, & Johnson, 1993; Huebner, Schein, Hall, & Barnes, 1994; Centers for Disease Control and Prevention, 2000; Comstock, Livesay, & Woolpert, 1974; Stead, 1995; Nardell, McInnis, Thomas, & Weidhaas, 1986; Centers for Disease Control and Prevention, 2011a). Diagnosis and treatment of LTBI is one of the interventions recommended by the WHO to end the worldwide TB epidemic, and is one of the elements of the post-2015 End TB Strategy (Pai & Behr, 2016). The primary goal of TB control in the U. S. has been to reduce the pool of infection through LTBI diagnosis and subsequent treatment (Pai & Behr, 2016).

Targeted Testing

Although diagnosis of LTBI can be an important TB control strategy, because only a small fraction of those with LTBI will ever develop active disease, and also because the positive-predictive value (PPV) of tests for LTBI decreases with decreasing prevalence, it is neither practical nor cost-effective to screen

everybody for LTBI. For this reason, it is recommended by the CDC and American Thoracic Society that only persons that are at high risk of developing TB disease should be tested, a process referred to as “targeted testing” (American Thoracic Society & Centers for Disease Control and Prevention, 2000; Centers for Disease Control and Prevention, 2011a). Testing and treating those at highest risk for TB will have the greatest impact on TB elimination.

The CDC has issued guidelines that identify high risk groups recommended for targeted LTBI testing (American Thoracic Society & Centers for Disease Control and Prevention, 2000; Centers for Disease Control and Prevention, 2011a). Persons at high risk for developing TB disease fall into two broad categories: 1) those who have an increased likelihood of exposure to persons with TB disease and 2) those with clinical conditions or other factors associated with an increased risk of progression from LTBI to TB disease (American Thoracic Society & Centers for Disease Control and Prevention, 2000; Centers for Disease Control and Prevention, 2011a). Those who have an increased likelihood of exposure to persons with TB disease include:

- Known close contacts of a person with infectious TB disease
- Persons who have immigrated from TB-endemic regions of the world
- Persons who work or reside in facilities or institutions with people who are at high risk for TB, such as hospitals that care for TB patients, homeless shelters, correctional facilities, nursing homes, or residential facilities for patients with HIV infection/AIDS

Those with clinical conditions or other factors associated with an increased risk of progression from LTBI to TB disease include those with:

- HIV infection
- Injection drug use
- Radiographic evidence of prior healed TB
- Low body weight (10% below ideal)
- Other medical conditions such as silicosis, diabetes mellitus, chronic renal failure, receiving hemodialysis, and other conditions

Immunologic Tests for *Mycobacterium Tuberculosis* Infection

LTBI is diagnosed by responses to *in vivo* or *in vitro* MTB antigen stimulation using either the TST or IGRA (Getahun, Matteelli, Chaisson, & Raviglione, 2015). Both are based on the quantification of the immune system's memory T-cell response reaction to TB antigens (small pieces of protein from the MTB bacillus). If an immune reaction is present, this then indicates infection.

Tuberculin Skin Test (TST)

The TST has been used since the early 1900s and, until recently, was the only practical immunological test for infection by MTB (American Thoracic Society & Centers for Disease Control, 2000; Lee & Holzman, 2002). The most common version of the test (Mantoux) is performed by injecting 0.1 ml of tuberculin purified protein derivative (PPD) intradermally (between the layers of the dermis), usually on the forearm. PPD is an extract of the media used to grow MTB and contains a mixture of MTB protein antigens. In infected individuals, PPD elicits a delayed-type hypersensitivity reaction that is evidence by induration at the injection site. Measurement of induration diameter 48 to 72 hours after injection is used to interpret the TST. A test is considered positive if the induration diameter is greater than or equal to a predetermined cutoff point. A

positive TST implies MTB infection and an increased risk of currently having, or subsequently developing TB (Edwards & Edwards, 1960; Antonucci, Girardi, Raviglione, & Ippolito, 1995; Selwyn et al., 1992). However, false-positive TSTs may occur, and causes may include exposure or infection with non-tuberculous mycobacteria (NTM), vaccination with Bacillus Calmette–Guérin (BCG), or errors in placing or interpreting the TST (Edwards & Edwards, 1960; Judson & Feldman, 1974; Snider, Jr., 1985). Such reactions result in lower TST specificity and lower PPV. TST screening of low-risk persons is discouraged because the PPV decreases as the prevalence of infection decreases (Centers for Disease Control and Prevention, 2000; Jensen, Lambert, Iademarco, Ridzon, & Centers for Disease Control and Prevention, 2005). Although TST testing programs should be conducted only among high-risk groups, certain persons may require a TST for situations such as employment or school attendance. Diagnosis and treatment of LTBI should always be tied to risk assessment (American Thoracic Society & Centers for Disease Control and Prevention, 2000).

A risk-stratified interpretation of TST is used to increase the predictive value of the test. A person's medical risk factors determine at which increment (5 mm, 10 mm, or 15 mm) of induration diameter above which the result is considered positive. This is referred to as risk-stratified interpretation (Huebner, Schein, & Bass, 1993; American Thoracic Society & Centers for Disease Control and Prevention, 2000; Snider, Jr., 1982). Using a high TST cutoff such as 15 for those at low risk of infection and progression, increases specificity and the likelihood that a positive result is a true positive. Using a low TST cutoff such as

5 mm for those at greatest risk of infection or progression to TB if infected, increases test sensitivity. Risk-stratified interpretation of TST as applied in the U. S. is summarized as follows:

High Risk: ≥ 5 mm induration is interpreted as positive in:

- HIV-infected persons
- Close contacts to an infectious TB case
- Persons with chest radiographs consistent with prior untreated TB
- Organ transplant recipients
- Other immunosuppressed patients (e.g. , those taking the equivalent of > 15 mg/day of prednisone for 1 month or those taking TNF-α antagonists)

Moderate Risk: ≥ 10 mm induration is interpreted as positive in:

- Recent immigrants
- Injection drug users
- Residents or employees of congregate settings
- Mycobacteriology laboratory personnel
- Persons with clinical conditions that place them at high risk
- Children < 4 years; infants, children, and adolescents exposed to adults at high-risk

Low Risk: ≥ 15 mm induration is interpreted as positive in:

- Persons with no known risk factors for TB

Interferon-Gamma Release Assay (IGRA)

IGRAs have become clinically-acceptable alternatives to the TST over the past two decades. IGRAs quantify IFN-γ response when blood or peripheral blood mononuclear cells (PBMCs) from blood are stimulated *in vitro* with MTB antigens (Belknap & Daley, 2014). In these assays, IFN-γ responses are expressed as either 1) plasma concentration in International Units (IU) per milliliter (mL) of plasma, or 2) counts of cells producing IFN-γ per million PBMCs. Responses above a certain cutoff value are considered to be indicating infection.

Two IGRAs are currently approved by the U. S. Food and Drug Administration (FDA) as aids for diagnosing LTBI. The T-SPOT® TB Test (T-Spot; Oxford Immunotec) and the QuantiFERON®-TB Gold In-Tube Test (QFT-GIT, Quiagen). Both tests assess response to manufactured peptides with overlapping sequences representing the specific MTB antigens early secreted antigenic target 6 (ESAT-6) and culture-filtrate protein 10 (CFP-10), but QFT-GIT also assesses response to the TB7.7 antigen. A newer version of the whole blood IGRA called QuantiFERON-TB Gold Plus has recently been approved by the FDA. It assesses response to the same manufactured peptides with overlapping sequences representing ESAT-6 and CFP-10 used in QFT-GIT, but does not asses response to TB7.7. It includes a tube with shorted peptides representing ESAT-6 and CFP-10, which are included to detect IFN-γ from CD8 T-cell lymphocytes (QUIAGEN, 2016a). The main difference in T-Spot and QFT-GIT is in how the IFN-γ response is measured. QFT-GIT uses an enzyme-linked immunosorbent assay (ELISA) to measure differences in plasma IFN-γ concentrations whereas the T-Spot measures differences in the number of PBMCs expressing IFN-γ using an enzyme-linked spot assay. QFT-GIT is preferred by many health departments over T-Spot, as it uses a spectrophotometer to record ELISA (using optical density) rather than a subjective visual counting method. It is also regarded to be easier to perform (there is no white blood cell separation step), and does not require shipment to a central facility, which can affect the viability of the white blood cells (QUIAGEN, 2016b).

IGRAs have several advantages over the TST. IGRAs require a single patient visit, while the TST requires two visits. IGRA results can be available in 24 hours, quicker than for the TST, which requires 48 to 72 hours for a result. IGRAs may be more specific than the TST because they assess response to antigens that are not in BCG or most non-tuberculosis mycobacteria (NTM) (DeKeyser, DeKeyser, & DeBaets, 2014). IGRAs also do not require a subjective visual measurement, which may be biased due to digit preference. However, IGRAs are more expensive than the TST (Nienhaus, Schablon, Costa, & Diel, 2011; Dewan et al., 2006), require sophisticated equipment and software, and are not recommended for children < 5 years of age.

A number of studies have compared the sensitivity and specificity of the TST and IGRAs. These studies have assessed specificity among subjects at low-risk for LTBI, and assessed sensitivity among those with culture-confirmed TB. Pooled sensitivity estimates from these studies are 89% for TST, 83% for QFT-GIT, and 90% for T-Spot; while pooled specificity estimates are 85% for TST, 99% for QFT-GIT, and 88% for T-Spot (Mazurek et al., 2010).

IGRAs are being used with greater frequency and are replacing the TST in many health departments in the U. S. Guidelines and recommendations for the use of IGRAs in the U. S. have been published by the CDC (Mazurek et al., 2010). Other high-income countries have also incorporated these tests in their national guidelines (Denkinger, Dheda, & Pai, 2011). The WHO has recently published recommendations against their use in LMIC (World Health Organization, 2011). This is because most IGRA studies have been done in high-

income countries and the application of their results to LMIC settings with high background TB infection rates may not be appropriate (World Health Organization, 2011). Systematic reviews have also suggested that IGRA performance differs in high- versus low-TB, and high- versus low-HIV incidence settings, with generally lower sensitivity in high-burden settings (World Health Organization, 2011). Another important aspect against their use in LMIC is cost. Given similar performance but higher costs, the use of IGRAs as a replacement for the TST in areas with a limited budget is not recommended (Trajman, Steffen, & Menzies, 2013).

Problems Addressed by this Research

Within-Subject TST Variability

As with any diagnostic test, the within-subject variability of TST, especially around the cutoff, is an area of concern. Erroneous reclassification of a negative test as positive (or vice versa) can be due to test variability, and not due a change infection status. Sources of variation can be due to differences in test reagents (such as the brand and lot of PPD used) and random biologic variation. However, the most common sources of this variation are in the administration (placing) and reading of the test (Menzies, 1999). Standard deviations in TST induration of 1.3 mm to 1.9 mm have been reported when read twice by the same reader (Bearman, 1964; Menzies, 1999), and discordance in test result interpretation as low as 1% to 2% have been reported (Furcolow, Watson, Charron, & Lowe, 1967). According to Menzies, et.al, biologic variation is small

compared to the variability resulting from placing and reading (Menzies, 1999). Variation due to all sources should result in standard deviations of less than 3 mm induration diameter (Menzies, 1999). In other words, when TSTs are repeated, chance variation should result in differences of less than 6 mm (which represents 2 standard deviations) in 95% of subjects (Menzies, 1999). Assessments of within-subject TST variability using results from two tests given in the same person, however, have rarely been performed (only two previous studies).

TST – QFT Agreement

Because IGRAs are being used instead of TST, there is a need to assess the head-to-head agreement. To date, the assessment of TST and QFT-GIT agreement has been performed in a limited number of populations, including healthcare workers (HCW), TB contacts, and persons with HIV or other diseases. It is currently not known how these tests would agree in persons who have had a prior positive TST.

Boosting of Immunologic Responses as Measured by TST

With time following MTB infection, the immune response to MTB can weaken, as evidence by an increase in the frequency of negative TST results with time (Menzies, 1999). However, the injection of MTB antigens (such as when PPD is injected for a TST) may stimulate, boost, or reawaken the immune response. Consequently, a TST given shortly after the initial TST may be larger, and in some situations, convert from negative to positive MTB (Figure 1).

Boosting is an increase in induration due to stimulation of an amnestic immunologic response. Detection of boosting requires measurement of an increase in induration size that significantly exceeds nonspecific variability, and requires an assumption of the absence of new infection occurring in the test interval between the initial and subsequent test. TST boosting is maximal if test interval is between one and five weeks (Menzies, 1999), although TST boosting has been detected as long as one year after the initial TST (Bass & Serio, 1981; Thompson, Glassroth, Snider, Jr., & Farer, 1979). Additional assessments of boosting include 1) determining the frequency that TST converts from negative to positive, and 2) determining the frequency that TST induration diameter increases ≥ 6 mm (Menzies, 1999). This “recall of waned cell-mediated immunity” is analogous to the secondary anamnestic serologic antibody response (Menzies, 1999). TST boosting has been described in persons with previous MTB infection, BCG vaccination, or NTM exposure (Menzies, Vissandjee, Rocher, & St Germain, 1994).

Because of the possibility of boosting, the CDC recommends a “two-step” TST procedure the with initial testing for individuals who may be tested serially (e.g., healthcare workers), as shown in Figure 2 (Centers for Disease Control and Prevention, 2014). If the initial TST is positive, the individual is considered positive for MTB infection. If the initial TST is negative, then the test is repeated in one to three weeks. If the second test is negative, then the individual is considered uninfected with MTB. If the second test is positive, then the individual is considered previously infected (i.e., a “boosted” reaction from a past infection).

Initial two-step testing is required to detect boosting, decrease false-negative TSTs, and avoid misinterpretation of subsequent positive TST reactions as new infections (Centers for Disease Control and Prevention, 2000).

Boosting of Immunologic Responses as Measured by QFT-GIT

Boosting of the second test is not limited to the second test being a TST, but can also occur when the second test is an IGRA (van Zyl-Smit, Zwerling, Dheda, & Pai, 2009). Injection of PPD is expected to boost IGRA responses to ESAT-6 and CFP-10 for some people who were previously infected with MTB. Such boosting would not be expected following BCG vaccination or infection with an NTM because these organisms do not stimulate a primary immune response to ESAT-6 or CFP-10. QFT-GIT boosting may have clinical significance in that boosting of negative QFT-GIT results to positive may prompt unnecessary treatment. The conclusions from prior studies assessing TST boosting of IGRA responses have been conflicting. Brock and colleagues did not observe boosting when people were retested with an IGRA three days after TST (Brock, et. al., 2001). However, Zyl-Smit and colleagues observed boosting of IGRA responses when people were tested 7 days after TST. It would therefore, be beneficial to assess the frequency of boosting, not only when TSTs are performed a week apart, but also when QFT-GITs are performed a week apart and the second QFT-GIT is a week after the initial TST.

Analysis Goals

This analysis is performed on data from a 2010 study jointly conducted by the CDC's Division of Tuberculosis Elimination and the United States Air Force (USAF). In this study, a series of TSTs and QFT-GITs were administered to healthy subjects (CDC or USAF employees) recruited at the CDC in Atlanta, GA or at Lackland Air Force Base in San Antonio, TX. The main goal of the study was to assess factors affecting QFT-GIT and TST variability. Manuscripts from the initial analysis of QFT-GIT variability from this study have been published (Whitworth et al., 2012; Whitworth et al., 2014). However, analyses of the TST component of the study has not yet been performed.

The four objectives of this study were to assess:

I. Within-subject TST variability, by comparing results of TSTs performed at the same time but in different arms.

II. TST - QFT-GIT agreement, by comparing results of TST and QFT-GIT that were initiated on the same day.

III. TST boosting of a subsequent TST, by comparing results of two TSTs performed one week apart.

IV. TST boosting of a subsequent QFT-GIT, by comparing results of two QFT-GITs performed one week apart, with the second QFT-GIT being one week after the initial TST.

Chapter II: Literature Review

Measures of Variability and Boosting Reported in Prior Studies

Studies assessing within-subject TST variability and TST - QFT-GIT interpretation (positive or negative) agreement have commonly reported, as indices, percent agreement (concordance), Cohen's kappa coefficient (*k*, or kappa), and percent discordance (100 minus percent agreement). *K*, also referred to as agreement beyond chance, is a statistic with values between 0 and 1.0 that adjusts for the possibility of agreement by chance (McHugh, 2012; Viera & Garrett, 2005). Several subjective scales have been developed to serve as guides for interpreting *k*. One such scale for interpreting *k* is as follows: 0 – 0.20 = no agreement, 0.21 - 0.39 = minimal, 0.40 – 0.59 = weak, 0.60 – 0.79 = moderate, 0.80 – 0.90 = strong, and > 0.90 = almost perfect (McHugh, 2012). Most studies examining TST boosting have focused on changes in positive or negative interpretations (qualitative changes) rather than changes in numeric values of induration or IFN-g concentration (quantitative changes). These studies typically assessed the percentage of subjects with a negative initial test result that was followed by a positive result (i.e., "percent of initial negatives that boosted to positive") which is the number of subjects who converted from negative to positive divided by the total number of subjects who were negative at initial testing. Studies assessing boosting using quantitative changes have usually defined a boosted reaction as being an increase in numeric values of induration or IFN-g concentration beyond some pre-determined numeric threshold.

Within-Subject TST Variability (Within-Subject Comparison of TST Results)

There have been three main categories of studies assessing TST variability. Firstly, there are a few studies that assess the difference between different test parameters in the same subject, such as reading methodologies (Geldenhuys et al., 2010; Bouros, Maltezakis, Tzanakis, Tzortzaki, & Siafakas, 1992; Longfield et al., 1984) and types of reagents used (Erdtmann, Dixon, & Llewellyn, 1974). These studies thus determine the variability that is caused by a change in a test parameter. Although informative, these studies do not provide comparative findings by which to assess the findings from the present study (i.e., the variability of two identical tests performed in the same individual at the same time using identical reading methodologies and test reagents).

Secondly, there are a number of TST inter-reader/rater (or inter-observer or interrater reliability) studies (Villarino et al., 1999; Villarino et al., 2000; Bearman, 1964; Mancuso et al., 2012; Perez-Stable & Slutkin, 1985; Pouchot et al., 1997; Kahwati et al., 2016; Longfield et al., 1984). These studies assess the agreement of a single TST given in a subject among two or more readers, which is fairly common in the literature. Of these TST interrater reliability studies, kappa statistics indicate moderate to substantial agreement between two observers (0.52 to 0.95). These studies assess TST within-subject variability to a degree, as it is only the variability attributed to different readers reading the same test. However, these studies do not take into account the possible biologic variability

that may occur when two tests are administered in the same person, which is what was done in the present study.

Thirdly are the studies in which two or more TSTs are administered at the same time on a subject, and the agreement among/between these tests determined. There have been only two of these studies to date (Furcolow et al., 1967; Chaparas et al., 1985). These studies assess the random variability of identically-performed tests administered in the same subject. This type of study assesses test precision (i.e., how close the measurements from the identically performed test are to each other). The findings from these two studies are most applicable and comparable to those in this present study.

In the 1967 study by Furcolow, et al. (Furcolow et al., 1967), each subject received four TSTs by three different methods (tine test, Mono-Vacc, and standard Mantoux method). The fourth test was a duplicate of one of the three. Two readers interpreted each test. The comparison of duplicate TSTs by the Mantoux method is of particular interest because this is the standard TST method and the method used in the present study. Among 212 hospital patients and employees who had duplicate TSTs by the Mantoux method, results indicated discordance/kappa of 1.9%/0.96 for the 1st reading and 1.0%/0.98 for the 2nd reading. One reason for this high amount of agreement (low discordance) may be that both tests in each subject were read by the same person.

In the 1985 study by Chaparas, et al. (Chaparas et al., 1985), two TSTs using the Mantoux method were administered in healthy subjects, read 48 hours (n = 1,036) and 72 hours (n = 892) later by four readers who interpreted the tests

using cutoffs of both 5 mm and 10 mm. For each subject, two readers read one reaction and two readers read the other reaction, and the average of the two readings per reaction was used for the analysis. Results from this study are shown in the following table:

Delay to Reading	Interpretation Cutoff	Agreement	Kappa	Discordance
48 hour	5 mm	79.2%	0.78	22.1%
48 hour	10 mm	83.0%	0.84	17.0%
72 hour	5 mm	78.9%	0.81	21.1%
72 hour	10 mm	85.5%	0.87	14.5%

Within-Subject Comparison of TST and QFT-GIT Results

Assessment of TST and QFT-GIT agreement has been included in a number of studies involving a variety of populations. Some studies included only patients with culture-confirmed TB, which facilitated comparison of test sensitivity without focusing on test agreement. Other studies involved cohorts of patients at minimal risk of infection, which facilitated comparison of test specificity. Some studies have compared these tests among people with varying levels of exposure to TB, and compared association of test results with level of exposure. However, for most studies comparing TST and QFT-GIT, evidence supporting the accuracy of one test over the other is lacking, leaving investigators to report indices of agreement.

Several TST – QFT-GIT agreement studies have been conducted in TB contacts (Song et al., 2014; Verhagen et al., 2014; Ayubi, Doosti-Irani, & Mostafavi, 2015). A 2015 systematic review and meta-analysis by Ayubi, et.al, examined 24 such studies (Ayubi et al., 2015). Results from this meta-analysis,

which included several studies conducted in children, indicated an overall kappa of 0.40 but did not indicate overall discordance. This analysis also showed that increasing values of the cutoff for the TST induration diameter (5 mm > 10 mm > 15 mm) resulted in improved agreement between the two tests. Of two studies of 2,982 adolescent TB contacts not included in this analysis Song, et al. found kappa of 0.38 using a 10mm TST cutoff and 0.56 using a 15 mm cutoff, confirming the observation made in the Ayubi study of greater kappa with increasing induration diameter cutoff (Song et al., 2014). The other, performed in 163 Venezuelan Amerindian pediatric contacts, found much higher agreement than in the other assessments of agreement in contacts (kappa = 0.76) (Verhagen et al., 2014).

Because of expected similarities with our study cohort, studies involving healthy HCWs are of particular interest (Hefzy, Wegdan, Elhefny, & Nasser, 2016; Mostafavi et al., 2016; Doosti-Irani, Ayubi, & Mostafavi, 2016; Bozkanat et al., 2016; Lamberti et al., 2015). In a 2015 systematic review and meta-analysis of TST- QFT-GIT agreement in HCW, Lamberti et al., examined 29 studies (Lamberti et al., 2015). Out of the 10,314 subjects in these studies, TST and QFT-GIT agreed for 6,893 and did not agree for 3,421 (33% discordance). TST+/QFT-GIT- discordance occurred approximately four times more often than TST-/QFT-GIT+ discordance [2,711 (26.3%) vs. 710 (6.9%)]. Kappas for the different studies ranged from 0.10 to 0.61, with an overall value of 0.28. In a 2016 meta-analysis of 30 TST - QFT-GIT agreement studies, kappa ranged from 0 to 0.93 with an overall discordance/kappa of 29%/0.27 (Doosti-Irani et al., 2016).

Studies conducted since (and not included in these HCW meta-analyses) have found varying discordance/kappa results for TST - QFT-GIT agreement: [6.5%/0.71, n=31 (Hefzy et al., 2016)], [22.5%/0.19, n=244 (Mostafavi et al., 2016)], and [55.9%/0.0, n=34 (Bozkanat et al., 2016)].

TST – QFT-GIT agreement has been assessed in a number of different diseased populations, including rheumatoid arthritis patients (Mehta, Zapantis, Petryna, & Efthimiou, 2015; Lee et al., 2014), kidney transplant recipients (Jambaldorj et al., 2017; Ayubi et al., 2017; Edathodu et al., 2017), and people living with HIV infection (Leung et al., 2016; Mathad et al., 2016; Kussen, Dalla-Costa, Rossoni, & Raboni, 2016; Khazraiyan et al., 2016; Mamishi, Pourakbari, Marjani, & Mahmoudi, 2014; Chkhartishvili et al., 2013). A 2015 review by Mamishi, et.al, identified 13 studies performed among HIV-infected patients, both adults and children (Mamishi et al., 2014). Discordance ranged from 11.0% to 79.0% and kappa ranged from 0.30 to 0.60. Other studies performed in people living with HIV infection since the above review have shown varying rates of discordance/kappa: [?%/0.29, n=240 (Chkhartishvili et al., 2013)], [26.1%/0.08, n=130 (Khazraiyan et al., 2016)], [14.2%/0.20, n=110 (Leung et al., 2016)], [25%/0.25, n=252 (Mathad et al., 2016)], and [11.6%/0.20, n=140 (Kussen et al., 2016)].

TST – QFT-GIT agreement in children has also been assessed (Howley et al., 2015; Rose et al., 2015; Masoumi Asl, Alborzi, Pourabbas, & Kalani, 2015). Discordance/kappa from these studies was determined as follows: [23.1%/0.20, n=2,520 immigrant children from Mexico, the Philippines, and Vietnam (Howley

et al., 2015)], [20.0%/0.58, n=103 children from a pediatric clinic (Rose et al., 2015)], and [5.8%/0.01, n=967 children randomly sampled from schools (Masoumi Asl et al., 2015)].

Studies that are the most comparable to this present study are those performed in healthy (non-HCW or non-contact) subjects. Ayubi's 2015 systematic review and meta-analysis included data from 22 studies comparing TST and QFT-GIT in healthy populations (Ayubi et al., 2015). The reported overall kappa of 0.35 was slightly lower than the overall kappa of 0.40 found in their assessment of agreement in TB contacts. In an assessment of 60 healthy subjects in Thailand, a discordance/kappa of 25.0%/0.16 was found (Reechaipichitkul, Pimrin, Bourpoern, Prompinij, & Faksri, 2015). In 107 healthy, male, Ethiopian medical and paramedical students, discordance/kappa was 8.4%/0.83 (Dagnew et al., 2012). In one small (n=207) population-based study conducted along the U. S. Mexico border, Oren, et.al, found a discordance/kappa of 26.1%/0.39, (Oren et al., 2015). In another (larger) population-based study which examined TST – QFT-GIT agreement in the National Health and Nutrition Examination (NHANES) Survey, Ghassemieh, et.al, determined a discordance/kappa of 3.0%/0.27 among U. S.-born and 18.4%/0.38 among foreign-born in this large, population-based sample of 6,064 subjects (Ghassemieh et al., 2016). In this NHANES study, QFT-GIT and TST were performed on the same day, as was the case in this present study.

TST Boosting

The potential for an initial TST to boost the response measured with a subsequent TST has been documented in many studies (Menzies et al., 1994; Menzies, 1999; Hobby, Holman, Iseman, & Jones, 1973; Murthy et al., 2013; Isler, Rivest, Mason, & Brassard, 2013; Teixeira et al., 2008; Salles et al., 2007; Al Mazrou, 2004; Kraut, Coodin, Plessis, & McLean, 2004; Besser et al., 2001; Habiban, Momeni, & Amiri, 2013). In each of these studies, the percentage of persons converting from an initial negative TST to a subsequent positive TST (i.e., % boosting to positive) was determined. Proportions with boosting ranged from 5.1% to 13.2%: [5.6%, n=1,961 students (Menzies et al., 1994)]; [8.3%, n=322 hospital workers (Hobby et al., 1973)]; [8.4%, n=764 medical students (Teixeira et al., 2008)]; [6.0%, n=455 TB contacts (Salles et al., 2007)]; [5.4%, n=65 dialysis patients (Habiban et al., 2013)] [13.2%, n=1,098 Indian adolescents (Murthy et al., 2013)]; [5.1%, n=256 employees of services for homeless (Isler et al., 2013)]; [12.0%, n= 236 nurses (Al Mazrou, 2004)]; [6.6%, n=698 HCW (Kraut et al., 2004)]. Factors found to be associated with boosting in these studies included prior BCG vaccination and older age. Although most studies were performed on populations of healthy people, none were restricted to individuals having a prior positive TST, as is done this present study.

The potential for TST to boost the response in a subsequent QFT-GIT has been assessed in several studies. In a 2009 review, van Zyl-Smit, et al. (van Zyl-Smit et al., 2009) identified five studies (Perry et al., 2008; Leyten et al., 2007; Richeldi, Bergamini, & Vaienti, 2008; van Zyl-Smit et al., 2009; Baker, Thomas, Stauffer, Peterson, & Tsukayama, 2009) in which the boosting effect of TST on QFT-GIT was

assessed. In these studies, the proportion of subjects with initial QFT-GIT negative to subsequent QFT-GIT positive boosting varied widely. Proportions with boosting ranged from 1.5% to 68.0%: [8.0%, n=26 HCW (van Zyl-Smit et al., 2009)], [68.0%, n=114 immigrants/refugees (Baker et al., 2009)], [6.0%, n=63 infectious disease cohort (Perry et al., 2008)], [1.5%, n=81 pediatric TB contacts (Richeldi et al., 2008)], and [5.0%, n=66 mixed TST- and TST+ (Leyten et al., 2007)]. In two of these studies, boosting was assessed at 3 days following TST administration, but was not observed at this short interval (Leyten et al., 2007; van Zyl-Smit et al., 2009). Several studies since this review have also been conducted (O'Shea et al., 2014; Esmail et al., 2016; Sauzullo et al., 2011; Ritz et al., 2011). Sauzullo, et al. found no evidence of boosting when QFT-GIT was administered at 1, 2, 4, and 6 weeks after the TST, and it should be noted that their population of TST negatives were also BCG negative (Sauzullo et al., 2011). Similarly, in a population of 16 BCG negative subjects with no history of TB exposure, Ritz, et al. found no boosting of QFT-GIT when administered 6 and 10 weeks after the initial TST and QFT-GIT (Ritz et al., 2011). In a study of 166 Nepalese military recruits recently arrived in the UK, O'Shea, et al. found that 9.5% boosted from QFT-GIT negative to positive 7 days after TST (O'Shea et al., 2014). A recent study by Esmail, et al. performed in 22 HIV-1-infected adults from South Africa, found that 40.9% boosted from QFT-GIT negative to QFT-GIT positive, with a median of 62 days after the TST (Esmail et al., 2016).

Chapter III: Methods

Study Population

The subjects in this analysis were healthy employees recruited in 2010 at the CDC (Atlanta, GA) and Lackland Air Force Base (USAF, San Antonio, TX) as part of a larger experimental study investigating QFT-GIT reproducibility (Whitworth et al., 2012; Whitworth et al., 2014). Inclusion criteria were age ≥ 18 years and a self-reported history of a prior-positive TST (documentation was not required as it was reasoned that those recruited would rarely have this). Exclusion criteria included a TST in the past 3 years or a history of an adverse reaction to TST (e.g., blistering, scarring, or anaphylaxis). Requiring a self-reported prior-positive TST avoided ethical concerns regarding treatment of LTBI when conflicting test results were encountered. This was expected also to increase the proportion of subjects with positive test results, LTBI, and previously treated MTB infection. Prior unpublished assessments in similar cohorts found that 40% to 50% of persons with self-reported prior-positive TST results were positive by QFT-GIT as compared to 3% for the general U. S. population (G. Mazurek, personal communication, and Ghassemieh et al., 2016). All subjects provided written informed consent and completed a detailed study questionnaire. The CDC and Wilford Hall Medical Center (USAF) human subjects institutional review boards approved the study.

Analysis and Study Design

For the parent study, subjects had multiple QFT-GITs and three TST performed over a period of two weeks. Figure 3 illustrates the ten-day portion of the parent study that involved TSTs and which pertains to this sub-study. For this sub-study, five tests were performed: three TSTs and two QFT-GITs. On day 1, blood was collected for the first QFT-GIT. Immediately following this, PPD was injected for two TSTs, one on the left arm and one on the right arm. On day 3, transverse induration was measure at the PPD injection sites on the left and right arm. On day 8, subjects returned and had blood collected for a second QFT-GIT, again followed by injection of PPD for a third TST in either the left or the right arm. A third TST was not performed if either of the first two TSTs produced induration ≥ 20 mm. On day 10, transverse induration was measured at the third PPD injection site. For the assessment of within-subject TST variability, results of TSTs applied on day 1 on the right and left arm were compared (#1). Assessment of TST - QFT-GIT agreement consisted of two comparisons of day 1 tests: left-arm TST vs. QFT-GIT and right-arm TST vs. QFT-GIT (#2). Assessment of TST boosting of a subsequent TST also required two comparisons: TST applied on day 1 on the right arm vs. TST applied on day 8 on either arm, and TST applied on day 1 on the left arm vs. TST applied on day 8 on either arm. To examine TST boosting of a subsequent QFT-GIT, results of the day 1 QFT-GIT (with blood collected prior to PPD injection for TST) were compared with results from the day 8 QFT-GIT (with blood collected seven days after TSTs were applied to the left and right arms).

Test Methods

TST was performed using the Mantoux method (Lee & Holzman, 2002) to inject intradermally 0.1 ml (5 TU) of Tubersol (PPD, Connaught Laboratories, Inc, Toronto, Ontario, Canada) following American Thoracic Society (ATS)/CDC guidelines (American Thoracic Society & Centers for Disease Control, 2000). Transverse induration at the at the TST placement site was measured 48 to 72 hours after PPD injection by trained healthcare workers according to ATS/CDC guidelines (American Thoracic Society & Centers for Disease Control, 2000). An induration diameter ≥ 10 mm was interpreted as a positive test result. Results of the first two TSTs (right and left arm) were read by different readers.

QFT-GIT was performed according to manufacturer's package insert (Cellestis Limited, 2010) and as previously described (Powell, III, Whitworth, Bernardo, Moonan, & Mazurek, 2011). Blood was collected into three specially designed tubes that contained: a) heparin alone (Nil tube); b) heparin, dextrose, and phytohemagglutinin A (PHA, Mitogen tube); or c) heparin, dextrose, and a cocktail of peptides representing ESAT-6, CFP-10, and TB7.7 (TB Antigen tube). Tubes were first mixed so the entire inner surface of the tubes was coated with blood and then incubated within 12 hours of collection for 16 to 24 hours at 37˚C prior to harvesting plasma. The tubes were centrifuged, and the IFN-γ concentration in the plasma was measured by ELISA. ELISAs were performed on a Triturus automated ELISA workstation (Grifols USA, Miami, FL), using eight IFN-γ calibrators (8, 4, 2, 1, 0.5, 0.25, 0.125, and 0 IU/mL) in duplicate to create standard curves. IFN-γ concentration was calculated from a standard curve using software developed at the

CDC, and test results interpreted as indicated in the Cellestis package insert and CDC guidelines (Mazurek et al., 2010; Cellestis Limited, 2010). TB Response was defined as the IFN-γ concentration in plasma from TB antigen-stimulated blood minus the IFN-γ concentration in unstimulated (Nil) blood. Mitogen Response was defined as the IFN-γ concentration in plasma from mitogen-stimulated blood minus Nil. A TB Response value ≥ 0.35 IU/mL was considered as a positive test result. TB Responses were considered "indeterminate" if: 1) Nil ≤ 0.7 IU/mL and TB Response < 0.35 IU/mL and Mitogen Response < 0.5 IU/mL or 2) Nil > 0.7 IU/mL and TB Response < 50% of Nil (Powell, III et al., 2011; Cellestis Limited, 2010).

Statistical Analysis Methods

Comparisons of both qualitative (positive/negative test interpretation) and quantitative (numeric test values) test results were performed using paired-sample methods. For assessment 3 (TST – QFT-GIT agreement) a quantitative comparison was not performed as the test measurement units were not the same (mm induration diameter vs. IU/mL of IFN-γ). For qualitative comparisons, 2 x 2 tables were created, and percent agreement, percent discordant, and kappa were calculated. For comparisons of day 1 and day 8 (boosting, assessments 3 and 4) test results, the percent of initial negatives that boosted to positive was calculated, and differences in proportions were assessed with McNemar's test. For quantitative comparisons, means and medians of test results were determined, and distributions of numeric test results were compared using the Wilcoxon signed-rank test.

Multivariate logistic regression was used to determine factors that were associated with the outcome of boosting to positive (assessments 3 and 4) among those that initially tested negative. Models were refined using stepwise backward elimination and alpha of 0.05 for covariate retention. Covariates in the full model included age (as categories), sex, race/ethnicity, birth region, prior TB exposure, prior TB treatment, prior LTBI therapy, BCG history, lived outside US for > 1 year, year of previous positive TST (as categories), study site (CDC or USAF), TST placer, TST reader.

Epi Info for Windows, v 7.0 (CDC) (Centers for Disease Control and Prevention, 2016) and MS Excel (Microsoft) were used to calculate frequencies and perform 2 x 2 table analyses. Kappa was calculated using the online calculator “QuickCalcs: Quantify Agreement with Kappa” (GraphPad Software, 2017). Significance testing for the differences in numeric test result distribution was performed using the online calculator “Wilcoxon Signed-Rank Test Calculator” (Social Science Statistics, 2017). Tests for differences in proportions were calculated using McNemar’s test in Open Epi (Dean, Sullivan, & Soe, 2013). SAS, v9.4 (SAS, Cary, NC) was used to create various analysis variables and to perform multivariate logistic regression.

Chapter IV: Results

Subject Characteristics

One hundred fifty-eight people consented to participate in the study and all provided evaluable data (97 were enrolled at CDC and 61 were enrolled at USAF). Subject characteristics are shown in Table 1. One subject had a TST as

far back as the 1950s, but the majority (75%) had a TST since 1990. Most (72%) had received therapy for LTBI, and a smaller number (2.5%) had received treatment for active TB. Twenty percent reported a previous vaccination with BCG, and 38% reported an exposure to someone with TB.

Test Results

A summary of test results for the five tests is shown in Table 2. Of note is the lower number of analyzable results for the day 8 TST, a consequence of not repeating the TST when a day 1 TST reaction was ≥ 20 mm. For the TSTs, 51.6% and 48.7% were positive with the (first) right and left arm TSTs, respectively, and 58.1% were positive with the second TST given a week later. For the two QFT-GITs, 20.8% positive with the first test increased to 46.4% positive with the second test. Frequencies of the test result values for the five tests are shown in Figure 4 for the 3 TSTs and Figure 5 for the 2 QFT-GITs. For the two QFT-GITs (Figure 4) there was a larger proportion of subjects with TB response values < 0.1 IU/mL for the day 1 QFT-GIT (95/149, 63.8%) as compared to the day 2 QFT-GIT (70/153, 45.8%).

Objective I: Assessment of Within-Subject TST Variability

Results of the qualitative assessment of within-subject TST variability are shown in Table 3 (Assessment I). Day 1 (left arm vs. right arm) TST test agreement was 84.4%, discordance was 13.6%, and kappa was 0.73. Of the 154 subjects available for this comparison, 51.9 % were positive with the left-arm TST vs. 48.7% positive with the right-arm TST, $p = 0.28$. The percentage of left-arm

positive/right-arm negative discordants was higher than the percentage of left-arm negative/right-arm positive discordants, but not significantly (8.4% vs. 5.2%, p = 0.28). Results of the quantitative assessment of within-subject TST variability are shown in Table 4 (Assessment I). Although the mean/median for the left arm TST (9.55/10.50) were slightly higher than for the right arm TST (9.07/9.00), the distributions of induration size were not significantly different (p = 0.10).

Objective II: Assessment of TST - QFT-GIT Agreement

Results of the qualitative assessment of day 1 TST vs. day 1 QFT-GIT agreement are shown in Table 3 (Assessment II). For comparison of the left-arm TST vs. QFT-GIT, agreement, discordance, and kappa were 55.4%, 44.6%, and 0.11, respectively. While 20.9% were positive with QFT-GIT, 50.7% were positive with TST, p < 0.0001. For the comparison of the right-arm TST vs. QFT-GIT, agreement, discordance, and kappa were 61.2%, 38.8%, 0.20, respectively. While 21.1% were positive with QFT-GIT, 47.6% were positive with TST, p < 0.0001. The percentage of QFT-GIT negative/TST positive discordants (37.2% and 32.7%) was significantly greater than the percentage of QFT-GIT positive/TST negative discordants (7.4% and 6.1%, both p-values < 0.0001).

Objective III: Assessment of TST Boosting of a Subsequent TST

Results of the qualitative assessment of TST boosting of a subsequent TST are shown in Table 3 (Assessment III) and yielded similar results for both left and right arm day 1 (1st) TSTs. Both indicate boosting. For the 1st (left arm) TST compared to the 2nd (any arm) TST one week later, 27.6% were discordant and

kappa = 0.46. There were 42.3% positive with 1st TST vs. 58.5% positive with 2nd TST, p < 0.0001. Of the 71 negative with the 1st TST, 38.0% were positive with the 2nd TST one week later (i.e., 38% boosted to positive). There was a higher percentage of 1st test negative/2nd test positive discordants than 1st test positive/2nd test negative discordants (22.0% vs. 5.7%, p < 0.0001). For the 1st (right arm) TST compared to the 2nd (any arm) TST one week later, 29.3% were discordant and kappa = 0.44. There were 37.4% positive with 1st TST vs. 58.5% positive with 2nd TST, p < 0.0001. Of the 77 negative with the 1st TST, 40.3% were positive with 2nd TST one week later (i.e., 40.3% boosted to positive) There was a higher percentage of 1st test negative/2nd test positive discordants than 1st test positive/2nd test negative discordants (25.2% vs. 4.1%, p < 0.0001).

Quantitative results for TST boosting of a 2nd TST are shown in Table 4 (Assessment III). Significant distribution differences for 1st and 2nd TSTs were seen for both right and left arms, with 2nd TST distribution locations much greater than 1st TST locations, thus indicating boosting of the 2nd TST. In multivariate analyses, a history of BCG was significantly associated with boosting to positive: OR 95% CI = 13.29 (1.77-99.58) for left arm and 7.00 (1.32-37.01) for right arm, (reference = no BCG history).

Objective IV: Assessment of TST Boosting of a Subsequent QFT-GIT

Qualitative results for TST boosting of a subsequent QFT-GIT are shown in Table 3 (Assessment IV). For the 1st QFT-GIT compared to the 2nd QFT-GIT one week later, 27.4% were discordant and kappa = 0.43. There were 21.2 % positive with 1st QFT-GIT vs. 47.3% positive with 2nd QFT-GIT, p < 0.0001. Of the 115

negative with the 1st QFT-GIT, 33.5% were positive with 2nd QFT-GIT one week later (i.e., 33.5% boosted) There was a higher percentage of 1st test negative/2nd test positive discordants than 1st test positive/2nd test negative discordants (26.7% vs. 0.7%, $p < 0.0001$). Quantitative results for TST boosting of a 2nd QFT-GIT are shown in Table 4 (Assessment IV). These results also show a large and highly significant increase if distribution location with the 2nd QFT-GIT compared to the 1st QFT-GIT, thus indicating boosting of the 2nd QFT-GIT. In multivariate analyses, a recent past positive TST (2000-2006, most recent) was significantly associated with boosting of the 2nd QFT-GIT: OR 95% CI = 9.29 (1.67-51.35), (reference = earliest, 1950-1989).

Chapter V: Discussion

Introduction

The main goal of the parent study was to assess factors affecting QFT-GIT variability, the results of which have been published. The study had a TST component as well, and the data regarding the TST had yet to be analyzed. An analysis of the TST component of the parent study, described in this thesis, was therefore performed using existing data from the parent study. The goals of this analysis have been successfully realized, and the methods described in this analysis, a mixture of qualitative and quantitative comparisons of paired test result data, were sufficient to answer these questions.

Summary and Findings

While making a correct diagnosis of LTBI is critical importance, available tests are limited in their accuracy, as evidenced by higher than expected variability. This is a continuation of analyses examining variability of tests for LTBI. Variability is of concern for both TST and IGRAs such as QFT-GIT. Variability in quantitative test result may cause qualitative variability in test interpretation without a change in infection status. Test variability can affect test accuracy. Incorrectly diagnosing a true negative as positive could result in unnecessary treatment, while incorrectly diagnosing a true positive as negative could result in a missed opportunity to prevent TB.

Another concern with these two tests is how well their results agree. Many health departments and healthcare facilities are transitioning from the older TST to the newer IGRAs, using them in tandem, or using them interchangeably (since in certain cases, the use of either one is recommended). For this reason, a precise assessment of the amount of agreement between these two tests is important.

TST boosting is an immunologic phenomenon that can occur when a previously MTB-infected person’s capability to mount an immune response to MTB antigens, weakened over time, reawakens following the injection of MTB antigens with a TST. This reaction (called “boosting”) can be quantitated by either a subsequent TST or an IGRA, such as the QFT-GIT. Although this has been assessed on other studies, there is some uncertainty as to how soon this reaction occurs in QFT-GIT following TST placement. In addition, the

determination of this phenomenon and its extent, using both TST and QFT-GIT as the subsequent tests in the same population and at the same time, had yet to be performed. For each of these three concerns (within-subject variability, agreement, and TST boosting) their assessment in a population of non-diseased individuals with a past-positive TST (and hence, a history of LTBI infection) had also not yet been determined.

The assessment of within-subject TST variability (Assessment I), comparing left and right arm TSTs, yielded discordance and kappa of 13.6% and 0.73, respectively, indicating good agreement. The discordance and kappa measures are slightly lower than those found in the previous study by Chaparas, et al., which used a similar methodology (17.0% and 0.84) (Chaparas et al., 1985).

The assessment of TST- QFT-GIT agreement yielded discordance/kappa of 44.6%/0.11 for left arm TST and 38.8%/0.20 for right arm TST, thus indicating poor agreement between the two types of tests. These measures were generally similar to those found in other TST vs. QFT-GIT agreement studies using other healthy populations. The kappa estimates were slightly higher than the pooled estimate of 0.35 found by the Ayubi meta-analysis (Ayubi et al., 2015).

TST boosting, of both a subsequent TST and a subsequent QFT-GIT, was demonstrated. For the assessment of TST boosting on a subsequent TST (Assessment III), of those with a negative initial TST, 38% (left arm TST) and 40% (right arm TST) converted to a positive TST when the TST was administered a week later. These boosting rates are higher than the range of rates seen in the literature (5.1% to 13.2%). One likely reason for this is that all of

the subjects in this study had a prior positive TST, which was not the case in any of the populations used in the previous studies. For the assessment of TST boosting on a subsequent QFT-GIT (Assessment IV), of those with a negative initial QFT-GIT, 34% converted to a positive QFT-GIT when the QFT-GIT was administered a week after the initial QFT-GIT and TST. This boosting rate is within the wide range seen in the previous studies assessing boosting of QFT-GIT. These rates of boosting were greater than within-subject TST variability of 13.6% found in this study and greater than QFT-GIT within subject variability of 5.0% found in a previous analysis (Whitworth et al., 2014), $p < 0.0001$.

In the multivariate analysis, BCG was found to be significantly associated with boosting of TST, but not QFT-GIT. This may be due to PPD boosting of cross-reactive immune response initiated by BCG. However, since BCG lacks ESAT-6 and CFP-10, vaccination does not elicit a primary immune response to these antigens and a subsequent TST would not boost the response to these antigens. Thus, a PPD injection might boost a TST response initiated by BCG, but not boost the ESAT-6 and CFP-10 response measured by QFT-GIT. In contrast, injection of PPD for TST following treatment and resolution of TB or LTBI, could boost subsequent TST and QFT-GIT responses.

Limitations

A limitation of this study is its non-generalizability. The study population was selected based on a prior positive TST rather than a selected sample that is representative of the U. S. population. Therefore, all participants had a history of LTBI. As stated, this selection was done to enrich the population of subjects with

those that had a high probability of a positive QFT-GIT result, and additionally, to ensure that there was sufficient number of subjects that could be assessed for boosting, since the boosting phenomenon requires a prior infection. This requirement of the participants to have a prior positive TST was necessary to do since the LTBI rate among the U. S. population is relatively low, and therefore, a random sample from the general population would not have adequately provided the necessary amount of those with a history of infection. Another possible limitation of this analysis was the use of a single TST cutoff point of 10 mm. It was beyond the scope of the present analysis to have performed a parallel assessment using a 15 mm cutoff point, which is normally used for low-risk populations. It would have been interesting to see what difference the use of this lower cutoff point would have made in the findings.

Implications

PPD injection for a TST may boost subsequent TST and QFT-GIT responses and increase the number of people testing positive by either test. Positive TST and QFT-GIT results following a TST may be due to boosting, and may not indicate new or existing MTB infection. This may inflate estimates of infection prevalence. For example, in an attempt to identify converters who are thought to be newly infected, repeat TST is advised for contacts who initially test negative. However, some of those testing positive with repeat testing may have converted due to boosting and not due to a new MTB infection. Some guidelines and many programs recommend initial screening with TST and confirmation of positive

results with an IGRA. However, for some people, a follow-up positive IGRA may be due to boosting.

The results of this study could be most applicable to populations in high-burden settings, such as high-burden countries like South Africa, where the LTBI prevalence can be as high as 89% (Ncayiyana et al., 2016; Stop TB Partnership, 2017; Wood et al., 2010). Information of this kind can be used to inform public health officials who regularly use the TST in health departments and screening programs, and especially in LMIC settings with high background TB infection rates, where the TST is still recommended (World Health Organization, 2011). The results from this analysis provide a valuable addition to the existing sparse literature on within-subject TST variability, of which only two studies using multiple TSTs performed in the same person exist (Chaparas et al., 1985; Furcolow et al., 1967). These results also add to the existing literature on TST - QFT-GIT agreement in non-diseased populations with high background TB infection rates. The agreement findings from this analysis are, as well, beneficial in these settings, especially where the TST and QFT-GIT are used interchangeably, in order to provide caregivers a reliable estimate of how often these tests can disagree. The demonstration of TST boosting of a subsequent QFT-GIT also provides a beneficial and informative addition to the existing literature on TST boosting of both TST and QFT-GIT responses, and especially provides a much-needed confirmation of this effect in QFT-GIT, which was demonstrated seven days after TST administration.

Recommendations for Future Studies

Recently, a new type of TST has been developed (Aggerbeck et al., 2013; Bergstedt et al., 2010). The C-Tb is identical to the TST, except that instead of using PPD, C-Tb uses the same antigens used in IGRAs (ESAT-6 and CFP-10), resulting in improved specificity. It thus combines simplicity of skin test with high specificity of the IGRA. C-Tb has been shown to have a high concordance with QFT-GIT (approximately 95%) and, although not currently being used yet, has the potential to eventually replace the TST (Abubakar, Jackson, & Rangaka, 2017; Hoff et al., 2016; Ruhwald et al., 2017). Future studies, examining the same assessments that are described in this analysis and using the same study design will, therefore, be necessary using not only this new TST, but also the new QuantiFERON-TB Gold Plus. Of particular interest will be to see how, if any, boosting is caused by the new C-Tb test.

Conclusions

Results from this analysis of within-subject TST variability, TST – QFT-GIT agreement, and TST boosting show good agreement between two TSTs placed simultaneously in the same individual, but poor agreement when the TST is compared to the QFT-GIT. Boosting, caused by a TST on a subsequent TST and QFT-GIT, was also demonstrated. The presence of a history of a BCG vaccine was significantly associated with the likelihood of TST boosting to positive. The findings from this analysis, performed on subjects with a prior positive TST, are in keeping with findings from the same assessments in previous studies.

Figures

Figure 1. TST Boosting.

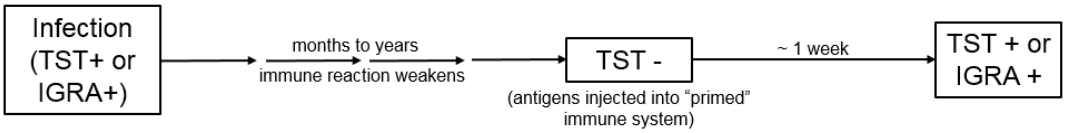


Figure 2. Two-Step TST Procedure.

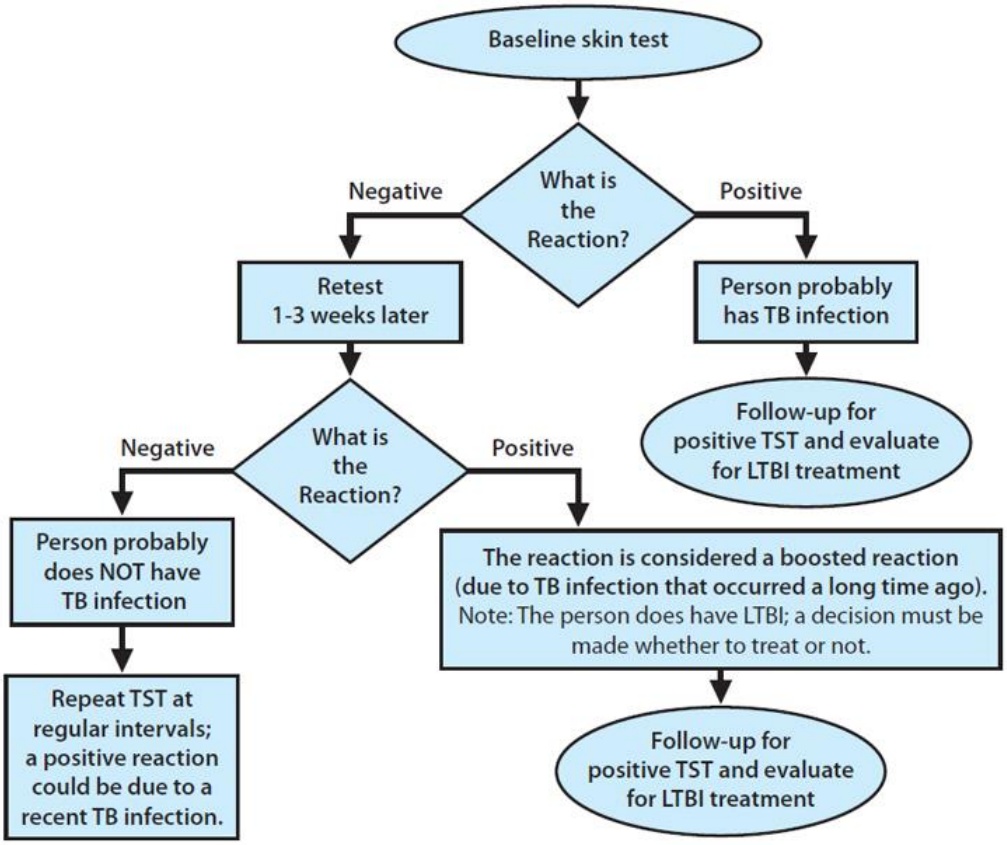


Figure courtesy of the Centers for Disease Control and Prevention, 2016. https://www.cdc.gov/tb/topic/testing/healthcareworkers.htm

Figure 3. Sub-study design.

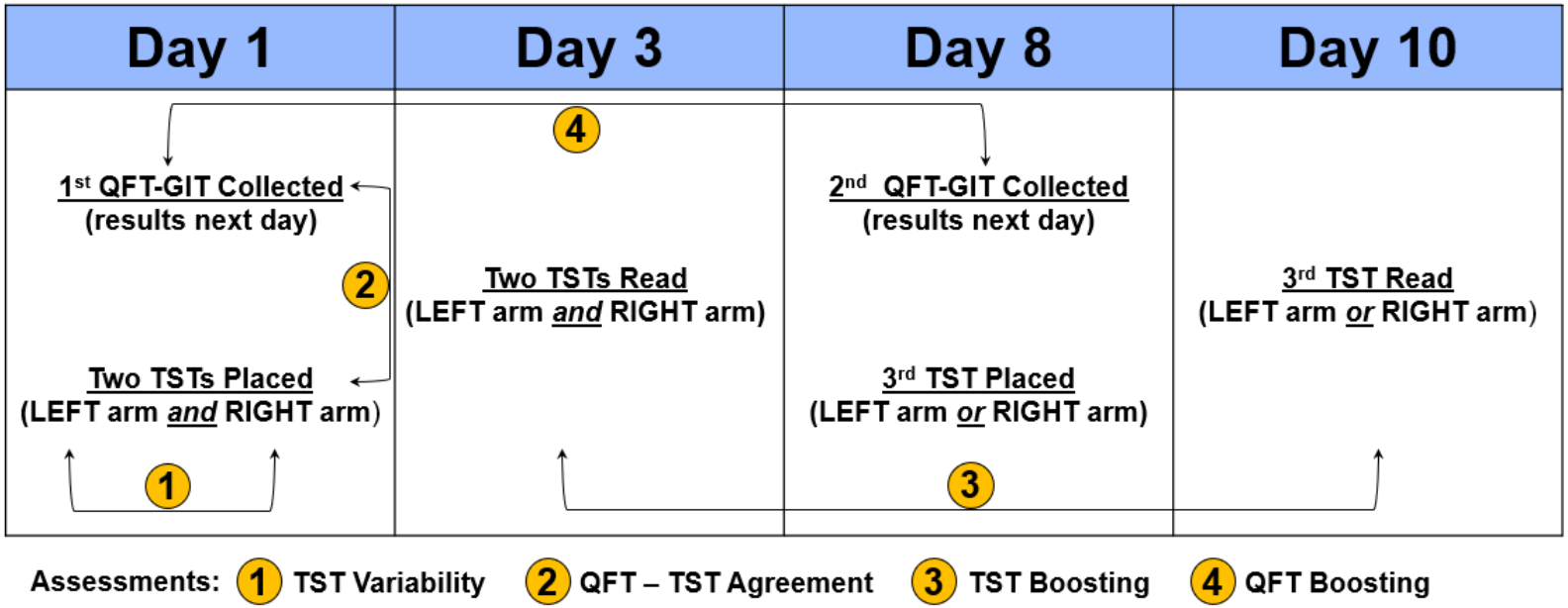


Figure 4. Frequencies of measured TST induration size.

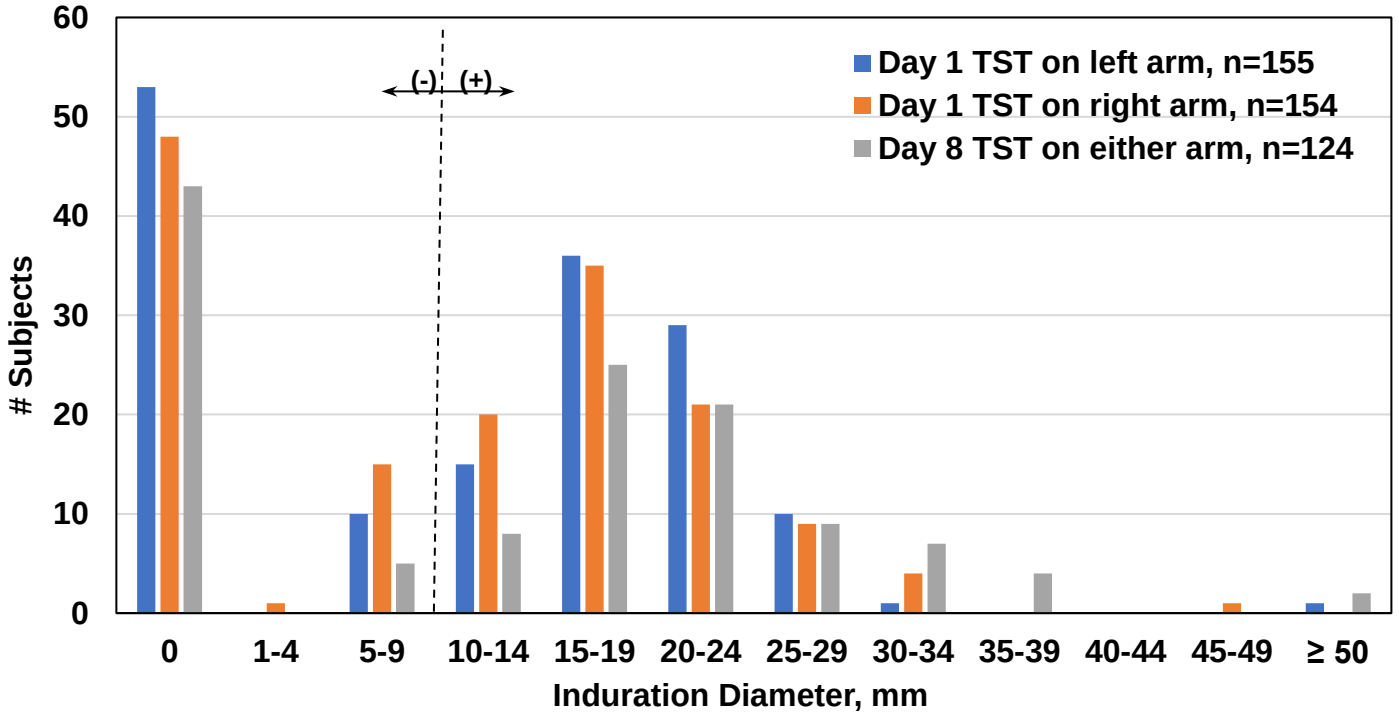
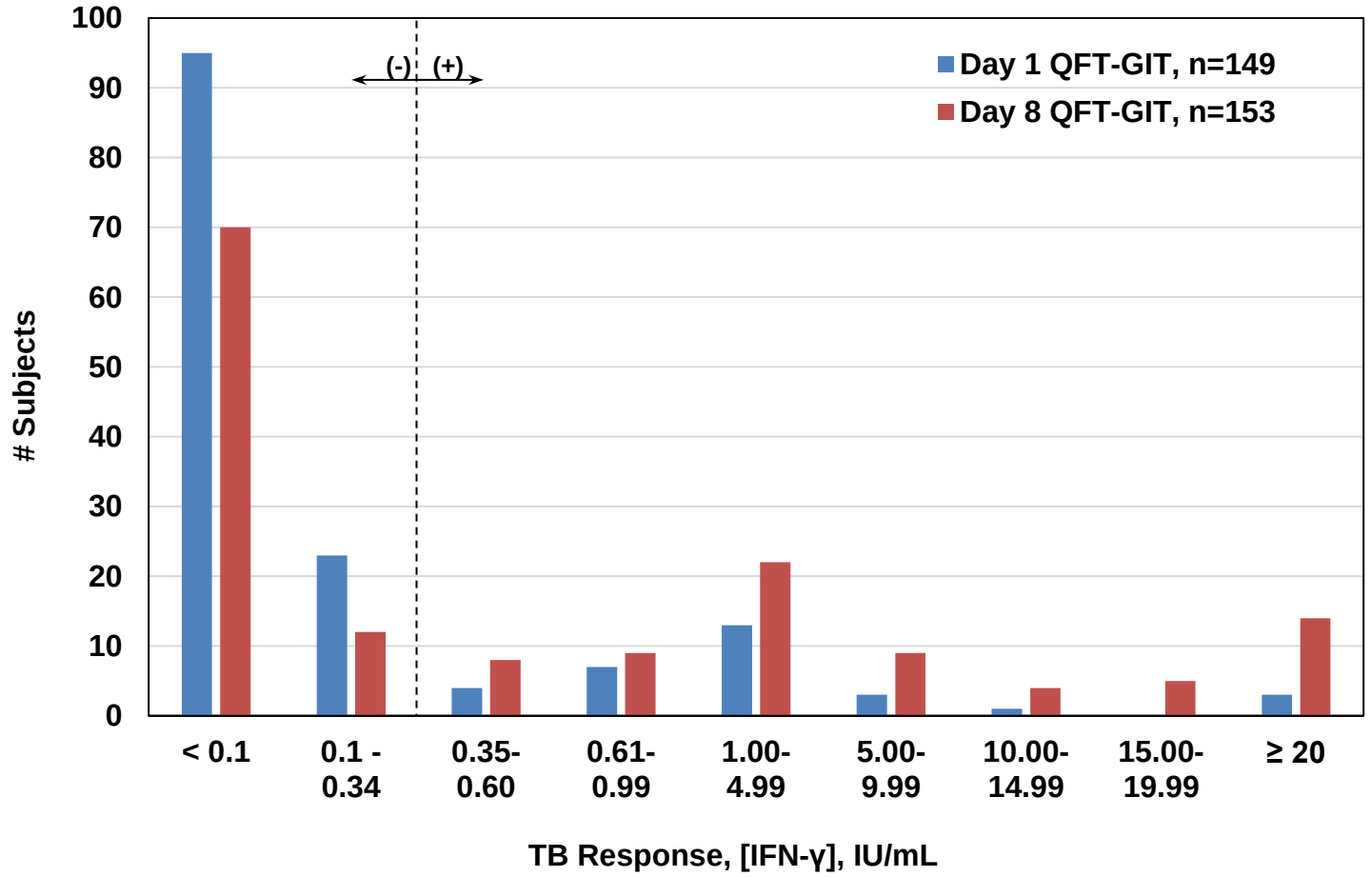


Figure 5. Frequencies of measured TB Responses from QFT-GIT.



Tables

Table 1. Subject characteristics, N = 158 (97 CDC, 61 USAF).

Age, years (mean, median, range)	45.2, 46, 24 to 74
Time since last positive TST, years (mean, median, range)	16.7, 13.5, 4 to 52

	Category	n (%)
Age Categories, years		
	20 - 29	16 (10.1%)
	30 - 39	35 (22.2%)
	40 - 49	45 (28.5%)
	50 - 59	42 (26.6%)
	≥ 60	20 (12.7%)
Gender		
	M	70 (44.3%)
	F	88 (55.7%)
Race/Ethnicity		
	White, non-Hispanic	79 (50.0%)
	Black, non-Hispanic	37 (23.4%)
	Hispanic	15 (9.5%)
	Asian/Pacific	18 (11.4%)
	Native American	1 (0.6%)
	Other	8 (5.1%)
Year of Last Positive TST		
	1950 - 1959	1 (0.6%)
	1960 - 1969	8 (5.1%)
	1970 - 1979	12 (7.6%)
	1980 - 1989	18 (11.4%)
	1990 - 1999	65 (41.1%)
	2000 - 2006	54 (34.2%)
Received Therapy for TB		
	Yes	4 (2.5%)
	No	154 (97.5%)

Table 1, continued.

	Category	n (%)
Received Therapy for LTBI		
	Yes	114 (72.2%)
	No	43 (27.2%)
	Unknown	1 (0.6%)
Exposure to Active TB		
	Yes	60 (38.0%)
	No	74 (47.1%)
	Unknown	24 (15.2%)
BCG Vaccine		
	Yes	32 (20.3%)
	No	112 (70.9%)
	Unknown	14 (8.9%)
Region of Birth		
	United States and Canada	113 (71.5%)
	Asia	14 (8.9%)
	Central America/Caribbean	12 (7.6%)
	Africa	7 (4.4%)
	Europe/Russia	4 (2.5%)
	Pacific	3 (1.9%)
	Southeast Asia	2 (1.3%)
	Middle East	2 (1.3%)
	South America	1 (0.6%)
Years Lived Outside USA		
	1 - 10	63 (39.9%)
	11 - 20	14 (8.9%)
	21 - 30	12 (7.6%)
	31 - 35	3 (1.9%)
	None	66 (41.8%)
Lived/worked/volunteered in homeless shelter, jail, or drug rehab unit		
	Yes	46 (29.1%)
	No	112 (70.9%)
Lived/worked/volunteered > 1 month in a hospital or nursing home		
	Yes	109 (69.0%)
	No	49 (31.0%)

Table 2. Characteristics of the five tests.

	N	Minimum	Maximum	Mean	Median	N (%) Positive
First (day 1) TST (Left Arm)	155	0	49	9.50	10.00	80 (51.6%)
First (day 1) TST (Right Arm)	154	0	45	9.10	9.00	75 (48.7%)
Second (day 8) TST (Either Arm)	124	0	50	11.60	11.50	72 (58.1%)
First (day 1) QFT-GIT	149	-0.59	26.63	0.96	0.30	31 (20.8%)
Second (day 8) QFT-GIT	153	-0.14	37.74	4.39	0.19	71 (46.4%)

Table 3. Qualitative results (summary of 2 x 2 table analyses).

Assessment	Comparisons: Test 1	Comparisons: Test 2	n	Concordant: Cell A (Test 1 + / Test 2 +)	Concordant: Cell D (Test 1 - / Test 2 -)	Discordant: Cell B (Test 1 + / Test 2 -)	Discordant: Cell C (Test 1 - / Test 2 +)	% Discordant: Total	% Discordant: Test 1 + / Test 2 -	% Discordant: Test 1 - / Test 2 +	Kappa	Comments
I. Within-Subject TST Variability	TST, left arm	TST, right arm	154	67	66	13	8	13.6	8.4	5.2	0.73	51.9 % (80/154) positive with left arm vs. 48.7% (75/154) positive with right arm, p = 0.28
II. TST – QFT-GIT Agreement	QFT-GIT	TST, left arm	148	20	62	11	55	44.6	7.4	37.2	0.11	20.9% (31/148) positive with QFT-GIT vs. 50.7% (75/148) positive with TST (left arm), p < 0.0001
	QFT-GIT	TST, right arm	147	22	68	9	48	38.8	6.1	32.7	0.20	21.1% (31/147) positive with QFT-GIT vs. 47.6% (70/147) positive with TST (right arm), p < 0.0001
III. TST Boosting of a Subsequent TST	TST, left arm	TST, either arm (1 week later)	123	45	44	7	27	27.6	5.7	22.0	0.46	42.3% (52/123) positive with 1st TST (left arm) vs. 58.5% (72/123) positive with 2nd TST, one week later (either arm), p < 0.0001 Boosting: Of 71 negative with 1st TST (left arm), 38.0% (27/71/) became positive with 2nd TST, one week later (either arm).
	TST, right arm	TST, either arm (1 week later)	123	41	46	5	31	29.3	4.1	25.2	0.44	37.4% (46/123) positive with 1st TST (right arm) vs. 58.5% (72/123) positive with 2nd TST, one week later (either arm), p < 0.0001 Boosting: Of 77 negative with 1st TST (right arm), 40.3% (31/77/) became positive with 2nd TST, one week later (either arm).
IV. TST Boosting of a Subsequent QFT-GIT	QFT-GIT	QFT-GIT (1 week later and 1 week after TST)	146	30	76	1	39	27.4	0.7	26.7	0.43	21.2 % (31/146) positive with 1st QFT-GIT vs. 47.3% (69/146) positive with 2nd QFT-GIT, one week later, p < 0.0001 Boosting: Of 115 negative with 1st QFT-GIT, 33.5% (39/115) became positive with 2nd QFT-GIT, one week later.

Table 4. Quantitative results.

Assessment**	n Compared	Test Results Compared	Mean	Median	P***
I. Within-Subject TST Variability	154	Day 1 TST on left arm	9.55*	10.50	0.10
		Day 1 TST on right arm	9.07	9.00	
III. TST Boosting of a Subsequent TST	123	Day 1 TST on left arm	7.33	7.00	$<10^{-5}$
		Day 8 TST on either arm	11.70	12.00	
	123	Day 1 TST on right arm	6.44	5.00	$<10^{-10}$
		Day 8 TST on either arm	11.70	12.00	
IV. TST Boosting of a Subsequent QFT-GIT	147	Day 1 QFT-GIT	0.97	0.03	$<10^{-10}$
		Day 8 QFT-GIT	4.34	0.22	

* Values are induration diameter (mm) for TST and IFN-γ concentration (IU/mL) for QFT-GIT.
** Quantitative Assessment II not performed because of different type/scale of test values (i.e., mm vs. IU/mL).
*** Wilcoxon signed-rank test.

References

Abubakar, I., Jackson, C., & Rangaka, M. X. (2017). C-Tb: a latent tuberculosis skin test for the 21st century? *Lancet Respir.Med., 5,* 236-237.

Aggerbeck, H., Giemza, R., Joshi, P., Tingskov, P. N., Hoff, S. T., Boyle, J. et al. (2013). Randomised clinical trial investigating the specificity of a novel skin test (C-Tb) for diagnosis of M. tuberculosis infection. *PLoS.ONE., 8,* e64215.

Ahmad, S. (2011). Pathogenesis, immunology, and diagnosis of latent Mycobacterium tuberculosis infection. *Clin.Dev.Immunol., 2011,* 814943.

Al Mazrou, A. M. (2004). Booster effect of two-step tuberculin skin testing among hospital employees from areas with a high prevalence of tuberculosis. *Infection Control and Hospital Epidemiology, 25,* 1117-1120.

American Thoracic Society & Centers for Disease Control (2000). Diagnostic standards and classification of tuberculosis in adults and children. *Am.J.Respir.Crit Care Med., 161,* 1376-1395.

American Thoracic Society & Centers for Disease Control and Prevention (2000). Targeted tuberculin testing and treatment of latent tuberculosis infection. *Am.J.Respir.Crit Care Med., 161,* S221-S247.

Antonucci, G., Girardi, E., Raviglione, M. C., & Ippolito, G. (1995). Risk factors for tuberculosis in HIV-infected persons. *JAMA, 274,* 143-148.

Ayubi, E., Doosti-Irani, A., & Mostafavi, E. (2015). Do the tuberculin skin test and the QuantiFERON-TB Gold in-tube test agree in detecting latent tuberculosis among high-risk contacts? A systematic review and meta-analysis. *Epidemiol.Health, 37,* e2015043.

Ayubi, E., Doosti-Irani, A., Sanjari, M. A., Khazaei, S., Mansori, K., Safiri, S. et al. (2017). Comparison of QuantiFERON-TB Gold In-Tube (QFT-GIT) and tuberculin skin test (TST) for diagnosis of latent tuberculosis in haemodialysis (HD) patients: a meta-analysis of kappa estimates. *Epidemiology and Infection,* 1-10.

Baker, C. A., Thomas, W., Stauffer, W. M., Peterson, P. K., & Tsukayama, D. T. (2009). Serial testing of refugees for latent tuberculosis using the QuantiFERON-gold in-tube: effects of an antecedent tuberculin skin test. *American Journal of Tropical Medicine and Hygiene, 80,* 628-633.

Bass, J. A. & Serio, R. A. (1981). The use of repeat skin tests to eliminate the booster phenomenon in serial tuberculin testing. *The American review of respiratory disease, 123,* 394-396.

Bearman, J. E. (1964). A study of variability in tuberculin test reading. *Am Rev Respir Dis, 90,* 913-919.

Belknap, R. & Daley, C. L. (2014). Interferon-gamma release assays. *Clin.Lab Med., 34,* 337-349.

Bergstedt, W., Tingskov, P. N., Thierry-Carstensen, B., Hoff, S. T., Aggerbeck, H., Thomsen, V. O. et al. (2010). First-in-man open clinical trial of a combined rdESAT-6 and rCFP-10 tuberculosis specific skin test reagent. *PLoS.ONE., 5,* e11277.

Besser, R. E., Pakiz, B., Schulte, J. M., Alvarado, S., Zell, E. R., Kenyon, T. A. et al. (2001). Risk factors for positive mantoux tuberculin skin tests in children in San Diego, California: evidence for boosting and possible foodborne transmission. *Pediatrics, 108,* 305-310.

Bouros, D., Maltezakis, G., Tzanakis, N., Tzortzaki, E., & Siafakas, N. (1992). The role of inexperience in measuring tuberculin skin reaction (Mantoux test) by the pen or palpation technique. *Respir.Med., 86,* 219-223.

Bozkanat, E., Kaya, H., Sezer, O., Caliskan, T., Kilic, E., Ciftci, F. et al. (2016). Comparison of tuberculin skin test and quantiferon-TB gold in tube test for diagnosis of latent tuberculosis infection in health care workers: A cross sectional study. *J.Pak.Med.Assoc., 66,* 270-274.

Brock, I., Munk, M. E., Kok-Jensen, A., Andersen, P. (2001) Performance of whole blood IFN-gamma test for tuberculosis diagnosis based on PPD or the specific antigens ESAT-6 and CFP-10. *Int J Tuberc Lung Dis*. 5(5):462-7

Cellestis Limited (2010). QuantiFERON®-TB Gold In-Tube Package Insert [ver 2010-01 for US]. [On-line]. Available:

http://www.cellestis.com/IRM/Content/pdf/QuantiFeron%20US%20VerG_Jan2010%20NO%20TRIMS.pdf

Centers for Disease Control and Prevention (2000). Targeted tuberculin testing and treatment of latent tuberculosis infection. *MMWR Recomm.Rep., 49,* 1-51.

Centers for Disease Control and Prevention (2011a). Latent Tuberculosis Infection: A Guide for Primary Health Care Providers. [On-line]. Available: http://www.cdc.gov/tb/publications/LTBI/default.htm

Centers for Disease Control and Prevention (2011b). Recommendations for use of an isoniazid-rifapentine regimen with direct observation to treat latent Mycobacterium tuberculosis infection. *MMWR.Morbidity and Mortality Weekly Report, 60,* 1650-1653.

Centers for Disease Control and Prevention (2013). *Core Curriculum on Tuberculosis: What the Clinician Should Know*. Atlanta: U. S. Department of Health & Human Services; Public Health Services.

Centers for Disease Control and Prevention (2014). Diagnosis of Latent TB Infection. [On-line]. Available: https://www.cdc.gov/tb/publications/ltbi/diagnosis.htm

Centers for Disease Control and Prevention (2016). Epi Info. [On-line]. Available: https://www.cdc.gov/epiinfo/index.html

Centers for Disease Control and Prevention (2017). Reported Tuberculosis in the United States, 2015. [On-line]. Available: https://www.cdc.gov/tb/statistics/reports/2015/

Chaparas, S. D., Vandiviere, H. M., Melvin, I., Koch, G., & Becker, C. (1985). Tuberculin test. Variability with the Mantoux procedure. *Am.Rev.Respir.Dis., 132,* 175-177.

Chkhartishvili, N., Kempker, R. R., Dvali, N., Abashidze, L., Sharavdze, L., Gabunia, P. et al. (2013). Poor agreement between interferon-gamma release assays and the tuberculin skin test among HIV-infected individuals in the country of Georgia. *BMC.Infect.Dis., 13,* 513.

Choi, J. C., Shin, J. W., Kim, J. Y., Park, I. W., Choi, B. W., & Lee, M. K. (2008). The effect of previous tuberculin skin test on the follow-up examination of whole-blood interferon-gamma assay in the screening for latent tuberculosis infection. *Chest, 133,* 1415-1420.

Comstock, G. W., Livesay, V. T., & Woolpert, S. F. (1974). The prognosis of a positive tuberculin reaction in childhood and adolescence. *American Journal of Epidemiology, 99,* 131-138.

Dagnew, A. F., Hussein, J., Abebe, M., Zewdie, M., Mihret, A., Bedru, A. et al. (2012). Diagnosis of latent tuberculosis infection in healthy young adults in a country with high tuberculosis burden and BCG vaccination at birth. *BMC.Res.Notes, 5,* 415.

Dean, A. G., Sullivan, K. M., & Soe, M. M. (2013). OpenEpi: Open Source Epidemiologic Statistics for Public Health. [On-line]. Available: www.OpenEpi.com

DeKeyser, E., DeKeyser, F., & DeBaets, F. (2014). Tuberculin skin test versus interferon-gamma release assays for the diagnosis of tuberculosis infection. *Acta Clinica Belgica, 69,* 358-366.

Denkinger, C. M., Dheda, K., & Pai, M. (2011). Guidelines on interferon-gamma release assays for tuberculosis infection: concordance, discordance or confusion? *Clin.Microbiol.Infect., 17,* 806-814.

Dewan, P. K., Grinsdale, J., Liska, S., Wong, E., Fallstad, R., & Kawamura, L. M. (2006). Feasibility, acceptability, and cost of tuberculosis testing by whole-blood interferon-gamma assay. *BMC.Infect.Dis., 6,* 47.

Dheda, K., van Zyl-Smit, R. N., Sechi, L. A., Badri, M., Meldau, R., Meldau, S. et al. (2009). Utility of quantitative T cell responses versus unstimulated IFN-{gamma} for the diagnosis of pleural tuberculosis. *European Respiratory Journal*.

Doosti-Irani, A., Ayubi, E., & Mostafavi, E. (2016). Tuberculin and QuantiFERON-TB-Gold tests for latent tuberculosis: a meta-analysis. *Occup.Med.(Lond), 66,* 437-445.

Edathodu, J., Varghese, B., Alrajhi, A. A., Shoukri, M., Nazmi, A., Elgamal, H. et al. (2017). Diagnostic potential of interferon-gamma release assay to detect

latent tuberculosis infection in kidney transplant recipients. *Transpl.Infect.Dis., 19*.

Edwards, P. Q. & Edwards, L. B. (1960). Story of the tuberculin skin test from an epidemiologic viewpoint. *Am Rev Respir Dis, 81,* 1-47.

Erdtmann, F. J., Dixon, K. E., & Llewellyn, C. H. (1974). Skin testing for tuberculosis. Antigen and observer variability. *JAMA, 228,* 479-481.

Esmail, H., Barry, C. E., III, Young, D. B., & Wilkinson, R. J. (2014). The ongoing challenge of latent tuberculosis. *Philos.Trans.R.Soc.Lond B Biol.Sci., 369,* 20130437.

Esmail, H., Thienemann, F., Oni, T., Goliath, R., Wilkinson, K. A., & Wilkinson, R. J. (2016). QuantiFERON conversion following tuberculin administration is common in HIV infection and relates to baseline response. *BMC.Infect.Dis., 16,* 545.

Furcolow, M. L., Watson, K. A., Charron, T., & Lowe, J. (1967). A comparison of the tine and Mono-Vacc tests with the intradermal tuberculin test. *Am.Rev.Respir.Dis., 96,* 1009-1027.

Geldenhuys, H., Verver, S., Surtie, S., Hatherill, M., van, L. F., Kafaar, F. et al. (2010). The tuberculin skin test: a comparison of ruler and calliper readings. *Int.J.Tuberc.Lung Dis., 14,* 1266-1271.

Getahun, H., Matteelli, A., Chaisson, R. E., & Raviglione, M. (2015). Latent Mycobacterium tuberculosis infection. *N.Engl.J.Med., 372,* 2127-2135.

Ghassemieh, B. J., Attia, E. F., Koelle, D. M., Mancuso, J. D., Narita, M., & Horne, D. J. (2016). Latent Tuberculosis Infection Test Agreement in the National Health and Nutrition Examination Survey. *Am.J.Respir.Crit Care Med., 194,* 493-500.

GraphPad Software (2017). QuickCalcs: Quantify agreement with kappa. [On-line]. Available: http://www.graphpad.com/quickcalcs/kappa1.cfm

Habiban, R., Momeni, A., & Amiri, M. (2013). Tuberculin skin test, booster phenomenon, and related changes in hemodialysis patients in Shahrekord, Iran. *Turkish Journal of Medical Sciences, 44,* 1-4.

Hefzy, E. M., Wegdan, A. A., Elhefny, R. A., & Nasser, S. H. (2016). Predictors of low prevalence of latent tuberculosis infection among Egyptian health care workers at intensive care and bronchoscopy units. *GMS.Hyg.Infect.Control, 11,* Doc22.

Hobby, G. L., Holman, A. P., Iseman, M. D., & Jones, J. (1973). Enumeration of tubercle bacilli in sputum of patients with pulmonary tuberculosis. *Antimicrobial Agents and Chemotherapy, 4,* 94-104.

Hoff, S. T., Peter, J. G., Theron, G., Pascoe, M., Tingskov, P. N., Aggerbeck, H. et al. (2016). Sensitivity of C-Tb: a novel RD-1-specific skin test

for the diagnosis of tuberculosis infection. *European Respiratory Journal, 47,* 919-928.

Horsburgh, C. R., Jr. (2004). Priorities for the treatment of latent tuberculosis infection in the United States. *N.Engl.J.Med., 350,* 2060-2067.

Howley, M. M., Painter, J. A., Katz, D. J., Graviss, E. A., Reves, R., Beavers, S. F. et al. (2015). Evaluation of QuantiFERON-TB gold in-tube and tuberculin skin tests among immigrant children being screened for latent tuberculosis infection. *Pediatric Infectious Disease Journal, 34,* 35-39.

Huebner, R. E., Schein, M. F., & Bass, J. B. J. (1993). The tuberculin skin test. *Clin.Infect.Dis., 17,* 968-975.

Huebner, R. E., Schein, M. F., Hall, C. A., & Barnes, S. A. (1994). Delayed-type hypersensitivity anergy in human immunodeficiency virus- infected persons screened for infection with *Mycobacterium tuberculosis*. *Clinical Infectious Diseases, 19,* 26-32.

Isler, M. A., Rivest, P., Mason, J., & Brassard, P. (2013). Screening employees of services for homeless individuals in Montreal for tuberculosis infection. *J.Infect.Public Health, 6,* 209-215.

Jambaldorj, E., Han, M., Jeong, J. C., Koo, T. Y., Min, S. I., Song, E. Y. et al. (2017). Poor predictability of QuantiFERON-TB assay in recipients and donors for tuberculosis development after kidney transplantation in an intermediate-TB-burden country. *BMC.Nephrol., 18,* 88.

Jensen, P. A., Lambert, L. A., Iademarco, M. F., Ridzon, R., & Centers for Disease Control and Prevention (2005). Guidelines for preventing the transmission of *Mycobacterium tuberculosis* in health-care settings, 2005. *MMWR Recomm.Rep., 54,* 1-141.

Judson, F. N. & Feldman, R. A. (1974). Mycobacterial skin tests in humans 12 years after infection with *Mycobacterium marinum*. *Am.Rev.Respir.Dis., 109,* 544-547.

Kahwati, L. C., Feltner, C., Halpern, M., Woodell, C. L., Boland, E., Amick, H. R. et al. (2016). Screening for Latent Tuberculosis Infection in Adults: An Evidence Review for the U. S. Preventive Services Task Force. *Agency for Healthcare Research and Quality, 14-05212-EF-1* .

Khazraiyan, H., Liaei, Z. A., Koochak, H. E., Ardalan, F. A., Ahmadinejad, Z., & Soltani, A. (2016). Utility of QuantiFERON-TB Gold In-Tube Test in the Diagnosis of Latent TB in HIV-Positive Patients in a Medium-TB Burden Country. *J.Int.Assoc.Provid.AIDS Care, 15,* 101-106.

Kraut, A., Coodin, M., Plessis, R., & McLean, D. (2004). Predictors of positive tuberculin skin test (TST) results after 2-step TST among health care workers in Manitoba, Canada. *Clin.Infect.Dis., 39,* e113-e118.

Kussen, G. M., Dalla-Costa, L. M., Rossoni, A., & Raboni, S. M. (2016). Interferon-gamma release assay versus tuberculin skin test for latent tuberculosis infection among HIV patients in Brazil. *Braz.J.Infect.Dis., 20,* 69-75.

Lamberti, M., Uccello, R., Monaco, M. G., Muoio, M., Feola, D., Sannolo, N. et al. (2015). Tuberculin skin test and Quantiferon test agreement and influencing factors in tuberculosis screening of healthcare workers: a systematic review and meta-analysis. *J.Occup.Med.Toxicol., 10,* 2.

Lee, E. & Holzman, R. S. (2002). Evolution and current use of the tuberculin test. *Clin.Infect.Dis., 34,* 365-370.

Lee, J. H., Sohn, H. S., Chun, J. H., Kim, H. A., Suh, C. H., Lee, Y. W. et al. (2014). Poor agreement between QuantiFERON-TB Gold test and tuberculin skin test results for the diagnosis of latent tuberculosis infection in rheumatoid arthritis patients and healthy controls. *Korean Journal of Internal Medicine, 29,* 76-84.

Leung, C. C., Chan, K., Yam, W. C., Lee, M. P., Chan, C. K., Wong, K. H. et al. (2016). Poor agreement between diagnostic tests for latent tuberculosis infection among HIV-infected persons in Hong Kong. *Respirology., 21,* 1322-1329.

Leyten, E. M., Prins, C., Bossink, A. W., Thijsen, S., Ottenhoff, T. H., van Dissel, J. T. et al. (2007). Effect of tuberculin skin testing on a *Mycobacterium tuberculosis*-specific interferon-gamma assay. *European Respiratory Journal, 29,* 1212-1216.

Longfield, J. N., Margileth, A. M., Golden, S. M., Lazoritz, S., Bohan, J. S., & Cruess, D. F. (1984). Interobserver and method variability in tuberculin skin testing. *Pediatr.Infect.Dis., 3,* 323-326.

Lonnroth, K., Castro, K. G., Chakaya, J. M., Chauhan, L. S., Floyd, K., Glaziou, P. et al. (2010). Tuberculosis control and elimination 2010-50: cure, care, and social development. *Lancet, 375,* 1814-1829.

Mamishi, S., Pourakbari, B., Marjani, M., & Mahmoudi, S. (2014). Diagnosis of latent tuberculosis infection among immunodeficient individuals: review of concordance between interferon-gamma release assays and the tuberculin skin test. *British Journal of Biomedical Science, 71,* 115-124.

Mancuso, J. D., Diffenderfer, J. M., Ghassemieh, B. J., Horne, D. J., & Kao, T. C. (2016). The Prevalence of Latent Tuberculosis Infection in the United States. *Am.J.Respir.Crit Care Med., 194,* 501-509.

Mancuso, J. D., Mazurek, G. H., Tribble, D., Olsen, C., Aronson, N. E., Geiter, L. et al. (2012). Discordance among commercially available diagnostics for latent tuberculosis infection. *Am.J.Respir.Crit.Care Med., 185,* 427-434.

Masoumi Asl, H. M. M., Alborzi, A. M., Pourabbas, B., & Kalani, M. M. (2015). QuantiFERON-TB Gold and Tuberculin Skin Test for the Diagnosis of Latent Tuberculosis Infection in Children. *Iran J.Med.Sci., 40,* 411-417.

Mathad, J. S., Bhosale, R., Balasubramanian, U., Kanade, S., Mave, V., Suryavanshi, N. et al. (2016). Quantitative IFN-gamma and IL-2 Response

Associated with Latent Tuberculosis Test Discordance in HIV-infected Pregnant Women. *Am.J.Respir.Crit Care Med., 193,* 1421-1428.

Mazurek, G. H., Jereb, J., Vernon, A., LoBue, P., Goldberg, S., Castro, K. et al. (2010). Updated guidelines for using Interferon Gamma Release Assays to detect Mycobacterium tuberculosis infection - United States, 2010. *MMWR Recomm.Rep., 59,* 1-25.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochem.Med.(Zagreb.), 22,* 276-282.

Mehta, B., Zapantis, E., Petryna, O., & Efthimiou, P. (2015). Screening Optimization of Latent Tuberculosis Infection in Rheumatoid Arthritis Patients. *Arthritis, 2015,* 569620.

Menzies, D. (1999). Interpretation of repeated tuberculin tests. Boosting, conversion, and reversion. *Am.J.Respir.Crit Care Med., 159,* 15-21.

Menzies, R., Vissandjee, B., Rocher, I., & St Germain, Y. (1994). The booster effect in two-step tuberculin testing among young adults in Montreal [see comments]. *Ann.Intern.Med., 120,* 190-198.

Mostafavi, E., Nasehi, M., Hashemi, S. A., Esmaeili, S., Ghaderi, E., Sharafi, S. et al. (2016). Comparison of the tuberculin skin test and the QuantiFERON-TB Gold test in detecting latent tuberculosis in health care workers in Iran. *Epidemiol.Health, 38,* e2016032.

Murthy, M., Selvam, S., Jesuraj, N., Bennett, S., Doherty, M., Grewal, H. M. et al. (2013). Two-step tuberculin skin testing in school-going adolescents with initial 0-4 millimeter responses in a high tuberculosis prevalence setting in South India. *PLoS ONE, 8,* e71470.

Nardell, E., McInnis, B., Thomas, B., & Weidhaas, S. (1986). Exogenous reinfection with tuberculosis in a shelter for the homeless. *The New England journal of medicine, 315,* 1570-1575.

Ncayiyana, J. R., Bassett, J., West, N., Westreich, D., Musenge, E., Emch, M. et al. (2016). Prevalence of latent tuberculosis infection and predictive factors in an urban informal settlement in Johannesburg, South Africa: a cross-sectional study. *BMC.Infect.Dis., 16,* 661.

Nicas, M., Nazaroff, W. W., & Hubbard, A. (2005). Toward understanding the risk of secondary airborne infection: emission of respirable pathogens. *J.Occup.Environ.Hyg., 2,* 143-154.

Nienhaus, A., Schablon, A., Costa, J. T., & Diel, R. (2011). Systematic review of cost and cost-effectiveness of different TB-screening strategies. *BMC.Health Serv.Res., 11,* 247.

O'Shea, M. K., Fletcher, T. E., Beeching, N. J., Dedicoat, M., Spence, D., McShane, H. et al. (2014). Tuberculin skin testing and treatment modulates interferon-gamma release assay results for latent tuberculosis in migrants. *PLoS ONE, 9,* e97366.

Oren, E., Alatorre-Izaguirre, G., Vargas-Villarreal, J., Moreno-Trevino, M. G., Garcialuna-Martinez, J., & Gonzalez-Salazar, F. (2015). Interferon Gamma-Based Detection of Latent Tuberculosis Infection in the Border States of Nuevo Leon and Tamaulipas, Mexico. *Front Public Health, 3,* 220.

Pai, M. & Behr, M. (2016). Latent Mycobacterium tuberculosis Infection and Interferon-Gamma Release Assays. *Microbiol.Spectr., 4*.

Pai, M., Behr, M. A., Dowdy, D., Dheda, K., Divangahi, M., Boehme, C. C. et al. (2016). Tuberculosis. *Nature Reviews Disease Primers, 2,* 16076.

Pape, J. W., Jean, S. S., Ho, J. L., Hafner, A., & Johnson, W. D. J. (1993). Effect of isoniazid prophylaxis on incidence of active tuberculosis and progression of HIV infection [see comments]. *Lancet, 342,* 268-272.

Perez-Stable, E. J. & Slutkin, G. (1985). A demonstration of lack of variability among six tuberculin skin test readers. *American Journal of Public Health, 75,* 1341-1343.

Perry, S., Sanchez, L., Yang, S., Agarwal, Z., Hurst, P., & Parsonnet, J. (2008). Reproducibility of QuantiFERON-TB gold in-tube assay. *Clin.Vaccine Immunol., 15,* 425-432.

Pouchot, J., Grasland, A., Collet, C., Coste, J., Esdaile, J. M., & Vinceneux, P. (1997). Reliability of tuberculin skin test measurement. *Ann.Intern.Med., 126,* 210-214.

Powell, R. D., III, Whitworth, W. C., Bernardo, J., Moonan, P. K., & Mazurek, G. H. (2011). Unusual Interferon Gamma Measurements with QuantiFERON-TB Gold and QuantiFERON-TB Gold In-Tube Tests. *PLoS ONE, 6,* e20061.

Public Health Agency of Canada (2017). The BCG World Atlas. [On-line]. Available: http://www.bcgatlas.org/index.php

QUIAGEN (2016a). QuantiFERON-TB Gold. [On-line]. Available: http://www.quantiferon.com/irm/content/quantiferon-tb-gold1.aspx?RID=300

QUIAGEN (2016b). QuantiFERON-TB Gold. [On-line]. Available: http://www.quantiferon.com/irm/content/quantiferon-tb-gold1.aspx?RID=300

Reechaipichitkul, W., Pimrin, W., Bourpoern, J., Prompinij, S., & Faksri, K. (2015). Evaluation of the QuantiFERON?-TB Gold In-Tube assay and tuberculin skin test for the diagnosis of Mycobacterium tuberculosis infection in northeastern Thailand. *Asian Pacific Journal of Allergy and Immunology, 33,* 236-244.

Richeldi, L., Bergamini, B. M., & Vaienti, F. (2008). Prior tuberculin skin testing does not boost QuantiFERON-TB results in paediatric contacts. *Eur Respir J, 32,* 524-525.

Ritz, N., Yau, C., Connell, T. G., Tebruegge, M., Leslie, D., & Curtis, N. (2011). Absence of interferon-gamma release assay conversion following tuberculin skin testing. *Int.J.Tuberc.Lung Dis., 15,* 767-769.

Rose, W., Read, S. E., Bitnun, A., Rea, E., Stephens, D., Pongsamart, W. et al. (2015). Relating Tuberculosis (TB) Contact Characteristics to QuantiFERON-TB-Gold and Tuberculin Skin Test Results in the Toronto Pediatric TB Clinic. *J.Pediatric.Infect.Dis.Soc., 4,* 96-103.

Ruhwald, M., Aggerbeck, H., Gallardo, R. V., Hoff, S. T., Villate, J. I., Borregaard, B. et al. (2017). Safety and efficacy of the C-Tb skin test to diagnose Mycobacterium tuberculosis infection, compared with an interferon gamma release assay and the tuberculin skin test: a phase 3, double-blind, randomised, controlled trial. *Lancet Respir.Med.*.

Salles, C. G., Ruffino-Netto, A., Lapa-e-Silva JR, Kritski, A. L., Cailleaux-Cesar, M., Queiroz-Mello, F. C. et al. (2007). The presence of a booster phenomenon among contacts of active pulmonary tuberculosis cases: a retrospective cohort. *BMC.Public Health, 7,* 38.

Sandul, A. L., Nwana, N., Holcombe, J. M., Lobato, M. N., Marks, S., Webb, R. et al. (2017). High Rate of Treatment Completion in Program Settings with 12-Dose Weekly Isoniazid and Rifapentine (3HP) for Latent Mycobacterium tuberculosis Infection. *Clin.Infect.Dis.*.

Sauzullo, I., Massetti, A. P., Mengoni, F., Rossi, R., Lichtner, M., Ajassa, C. et al. (2011). Influence of previous tuberculin skin test on serial IFN-gamma release assays. *Tuberculosis.(Edinb.), 91,* 322-326.

Selwyn, P. A., Sckell, B. M., Alcabes, P., Friedland, G. H., Klein, R. S., & Schoenbaum, E. E. (1992). High risk of active tuberculosis in HIV-infected drug users with cutaneous anergy. *JAMA, 268,* 504-509.

Shea, K. M., Kammerer, J. S., Winston, C. A., Navin, T. R., & Horsburgh, C. R., Jr. (2014). Estimated rate of reactivation of latent tuberculosis infection in the United States, overall and by population subgroup. *American Journal of Epidemiology, 179,* 216-225.

Shepardson, D. & MacKenzie, W. R. (2014). Update on cost-effectiveness of a 12-dose regimen for latent tuberculous infection at new rifapentine prices. *Int.J.Tuberc.Lung Dis., 18,* 751.

Shepardson, D., Marks, S. M., Chesson, H., Kerrigan, A., Holland, D. P., Scott, N. et al. (2013). Cost-effectiveness of a 12-dose regimen for treating latent tuberculous infection in the United States. *Int.J.Tuberc.Lung Dis., 17,* 1531-1537.

Sloot, R., Schim van der Loeff MF, Kouw, P. M., & Borgdorff, M. W. (2014). Risk of tuberculosis after recent exposure. A 10-year follow-up study of contacts in Amsterdam. *Am.J.Respir.Crit Care Med., 190,* 1044-1052.

Snider, D. E., Jr. (1982). The tuberculin skin test. *Am.Rev.Respir.Dis., 125,* 108-118.

Snider, D. E., Jr. (1985). Bacille Calmette-Guerin vaccinations and tuberculin skin tests. *JAMA, 253,* 3438-3439.

Social Science Statistics (2017). Wilcoxon Signed-Rank Test Calculator. [On-line]. Available: http://www.socscistatistics.com/tests/signedranks/Default2.aspx

Song, S. E., Yang, J., Lee, K. S., Kim, H., Kim, Y. M., Kim, S. et al. (2014). Comparison of the tuberculin skin test and interferon gamma release assay for the screening of tuberculosis in adolescents in close contact with tuberculosis TB patients. *PLoS ONE, 9,* e100267.

Stead, W. W. (1995). Management of health care workers after inadvertent exposure to tuberculosis: a guide for the use of preventive therapy. *Ann.Intern.Med., 122,* 906-912.

Sterling, T. R., Villarino, M. E., Borisov, A. S., Shang, N., Gordin, F., Bliven-Sizemore, E. et al. (2011). Three months of rifapentine and isoniazid for latent tuberculosis infection. *N.Engl.J.Med., 365,* 2155-2166.

Stop TB Partnership (2017). High Burden Countries. [On-line]. Available: http://www.stoptb.org/countries/tbdata.asp

Teixeira, E. G., Kritski, A., Ruffino-Netto, A., Steffen, R., Lapa E Silva JR, Belo, M. et al. (2008). Two-step tuberculin skin test and booster phenomenon prevalence among Brazilian medical students. *Int.J.Tuberc.Lung Dis., 12,* 1407-1413.

Thompson, N. J., Glassroth, J. L., Snider, D. E., Jr., & Farer, L. S. (1979). The booster phenomenon in serial tuberculin testing. *Am.Rev.Respir.Dis., 119,* 587-597.

Trajman, A., Steffen, R. E., & Menzies, D. (2013). Interferon-Gamma Release Assays versus Tuberculin Skin Testing for the Diagnosis of Latent Tuberculosis Infection: An Overview of the Evidence. *Pulm.Med., 2013,* 601737.

van Zyl-Smit, R. N., Pai, M., Peprah, K., Meldau, R., Kieck, J., Juritz, J. et al. (2009). Within-subject Variability and Boosting of T Cell IFN-{gamma} Responses Following Tuberculin Skin Testing. *Am.J.Respir.Crit Care Med., 180,* 49-58.

van Zyl-Smit, R. N., Zwerling, A., Dheda, K., & Pai, M. (2009). Within-subject variability of interferon-g assay results for tuberculosis and boosting effect of tuberculin skin testing: a systematic review. *PLoS.ONE., 4,* e8517.

Verhagen, L. M., Maes, M., Villalba, J. A., d'Alessandro, A., Rodriguez, L. P., Espana, M. F. et al. (2014). Agreement between QuantiFERON(R)-TB Gold In-Tube and the tuberculin skin test and predictors of positive test results in Warao Amerindian pediatric tuberculosis contacts. *BMC.Infect.Dis., 14,* 383.

Viera, A. J. & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Fam.Med, 37,* 360-363.

Villarino, M. E., Brennan, M. J., Nolan, C. M., Catanzaro, A., Lundergan, L. L., Bock, N. N. et al. (2000). Comparison testing of current (PPD-S1) and

proposed (PPD-S2) reference tuberculin standards. *Am.J.Respir.Crit Care Med., 161,* 1167-1171.

Villarino, M. E., Burman, W., Wang, Y. C., Lundergan, L., Catanzaro, A., Bock, N. et al. (1999). Comparable specificity of 2 commercial tuberculin reagents in persons at low risk for tuberculous infection. *JAMA, 281,* 169-171.

Whitworth, W. C., Goodwin, D. J., Racster, L., West, K. B., Chuke, S. O., Daniels, L. J. et al. (2014). Variability of the QuantiFERON(R)-TB Gold In-Tube Test Using Automated and Manual Methods. *PLoS ONE, 9,* e86721.

Whitworth, W. C., Hamilton, L. R., Goodwin, D. J., Barrera, C., West, K. B., Racster, L. et al. (2012). Within-Subject Interlaboratory Variability of QuantiFERON-TB Gold In-Tube Tests. *PLoS ONE, 7,* e43790.

Wood, R., Liang, H., Wu, H., Middelkoop, K., Oni, T., Rangaka, M. X. et al. (2010). Changing prevalence of tuberculosis infection with increasing age in high-burden townships in South Africa. *Int.J.Tuberc.Lung Dis., 14,* 406-412.

World Health Orgainzation (2017). Latent Tuberculosis Infection (LTBI). [On-line]. Available: http://www.who.int/tb/challenges/ltbi/en/

World Health Organization (2011). Use of tuberculosis interferon-gamma release assays (IGRAs) in low- and middle-income countries: policy statement. [On-line]. Available: http://www.who.int/tb/features_archive/igra_policy24oct/en/

World Health Organization (2016). Global Tuberculosis Control: WHO Report 2016. [On-line]. Available: http://www.who.int/tb/publications/global_report/en/