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.

Signature:

_____	7/25/2023
Tucker Colvin	Date

Whole Blood Mycobacterial Growth Assays
And Tuberculosis Susceptibility

By

Tucker Colvin
Master of Public Health

Hubert Department of Global Health

Matthew Magee
Committee Co-Chair

Carlton A. Evans
Committee Co-Chair

Sumona Datta
Committee Member

Whole Blood Mycobacterial Growth Assays And Tuberculosis Susceptibility

By

Tucker Colvin

A.B.
Bowdoin College
2013

Thesis Committee Co-Chair: Matthew Magee, PhD
Thesis Committee Co-Chair: Carlton A. Evans, FRCP PhD MBE

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 Global Health
2023

Abstract

Whole Blood Mycobacterial Growth Assays
And Tuberculosis Susceptibility
By Tucker Colvin

Background: Developing a valid and simple biomarker test to predict tuberculosis (TB) susceptibility is necessary for reversing recent trends and ending TB. Whole blood mycobacterial growth assays (WBMGA) have demonstrated potential to be used in low resource settings in order to direct screening and preventative care to those most at risk. While WBMGA have been linked to determinants of TB, this thesis explores the association between WBMGA results and the risk of incipient TB disease.

Method: A secondary analysis was conducted with the permission of Innovation for Health and Development (IFHAD) in Lima, Peru on an anonymous, unlinked dataset. Participants were healthy, TB-negative adults who were close contacts of TB index cases.

Results: Mycobacterial growth in blood relative to plasma at or below the median log difference in luminosity of 0.53 relative light units (RLU) was significantly associated with a 3.0 (95% confidence interval= 1.0, 8.0) times increase in risk of developing TB disease within six months (p=0.031).

Conclusions: The results from this analysis demonstrate the potential of WBMGA to be used as a predictive tool of TB susceptibility within six months. Future research is necessary to explore the implementation of WBMGA for use in point-of-care risk evaluation and its potential contribution to the global campaign against TB.

Whole Blood Mycobacterial Growth Assays And Tuberculosis Susceptibility

By

Tucker Colvin

A.B.
Bowdoin College
2013

Thesis Committee Co-Chair: Matthew Magee, PhD
Thesis Committee Co-Chair: Carlton A. Evans, FRCP PhD MBE

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 Global Health
2023

Acknowledgements

I would like to thank my committee members and the incredible team at IFHAD.

Table of Contents

Introduction

An estimated 10.6 million people became ill from tuberculosis (TB) in 2022, a 4.5% increase from the previous year (WHO, 2022). Despite a global partnership and commitment to end TB by 2030, TB is on track to remain the leading cause of death from a single infectious agent for the near future (WHO, 2022). In order to reverse this deleterious trend, new diagnostic, preventative, and control strategies are required. With nearly one quarter of the world estimated to have latent TB infection (LTBI), there is an urgent need to identify and categorize which individuals are most at risk of developing active TB disease to allocate resources early when they matter most (Harries et al., 2020).

In order to identify those most at risk, a simple and affordable biomarker test is necessary to predict an individual's risk of TB disease. One viable candidate is whole blood mycobacterial growth assays (WBMGA), which measure the *in vitro* growth of mycobacteria in blood samples (Bok et al., 2021). Instead of focusing on a single biomarker, they act as functional assays which assess the combined effects of multiple factors that affect growth of mycobacteria *in vitro* (Bok et al., 2021). As both an affordable and simple test of an individual's immune response to mycobacteria, WBMGA potentially offer a means for low-resources healthcare systems to target screening and preventative measures at those who need them most.

The purpose of this analysis is to estimate the association between WBMGA and the risk of developing TB in the desert shanty towns of Ventanilla on the outskirts of Lima, Peru among close contacts of recently diagnosed TB cases. My research question is whether mycobacterial growth *in vitro* is associated with an individual's risk of developing TB *in vivo*. My null hypothesis is that mycobacterial growth *in vitro* is not associated with an individual's risk of developing TB *in vivo*.

A secondary analysis was conducted using an anonymous, unlinked dataset of close contacts of newly diagnosed index TB cases in Ventanilla collected by the Innovation for Health and Development (IFHAD) research group from 2002-2016. The purpose of the analysis was to determine if mycobacterial growth in the WBMGA correlates with known risk factors for TB disease and has a relationship with susceptibility to TB disease. Through this analysis, I hope to explore the possibility of TB susceptibility as a stable phenotype that can be measured and become a potentially valuable tool in the global effort to control and ultimately end TB.

Definition of terms

- **BCG** - Bacillus Calmette–Guérin vaccine
- **Biomarker (biological marker)** - an indicator of a physiological, pathological, pharmacological process ore response (Parida et al., 2014)
- **Biosignature** – A combination of indicators
- **BMI**- Body mass index
- **Correlate of protection** - a measurable sign of host response to an infectious agent indicating resistance or susceptibility to developing disease (Parida et al., 2014)
- **HAART**- Highly active antiretroviral therapy
- **IFHAD** - Innovation for Health and Development group in Lima, Peru
- **IGRA** - IFNγ release assays
- **LMICs**- Low- and middle-income countries
- **LTBI**- Latent tuberculosis infection
- **MOTT** - mycobacteria other than tuberculosis
- **PPD**- Positive purified protein derivative

- **RLU**- Relative light units
- **Surrogates of disease** - A biomarker or biosignature that is statistically associated with and pathophysiologically related to a clinical outcome (Parida et al., 2014)
- **TST** - Tuberculin skin test

Literature Review

In this literature review, the bibliographic databases PubMed and Google Scholar were searched for articles relating to TB risk, biomarkers, and WBMGA. The eligibility criteria were that articles must be in English, published since 2000, and be relevant to TB biomarkers (Table 1). Seven articles were included in the review, which were separated into five topic categories: determining who is at risk of TB, the importance of biomarkers in disease control and eradication, the importance of cost and ease of use in TB control, the search for TB biomarkers, and the potential of WBMGA (Table 2).

Table 1: Literature review search terms

	PubMed Keywords	PubMed MeSH Terms
Concept 1: Tuberculosis	tuberculosis, mycobacterium tuberculosis, mycobacterium infections, Koch's disease, TB	Tuberculosis, Mycobacterium tuberculosis, Mycobacterium infections
Concept 2: WBMGA	Whole Blood Mycobacterial Growth Assay, WBMGA, recombinant luminescent mycobacterial, BCG lux, mycobacterial growth inhibition tube, MGIT	Assay

Excluded Concepts	None	None
Qualifiers		
Age	All	All
Language	English	English
Date range	2000-present	2000-present

Table 2: Literature review articles by category

Category	Author and Year	Title
Determining who is at risk of TB	(Saunders et al., 2017)	A score to predict and stratify risk of tuberculosis in adult contacts of tuberculosis index cases: a prospective derivation and external validation cohort study
Importance of biomarkers in disease control and eradication	(Bonassi & Au, 2002)	Biomarkers in molecular epidemiology studies for health risk prediction
Importance of cost and ease of use in TB control	(Brown et al., 2021)	Implementation of GeneXpert for TB testing in low- and middle-income countries: a systematic review
The search for TB biomarkers	(Jacobsen et al., 2008)	Novel strategies to identify biomarkers in TB
	(Parida & Kaufmann, 2010)	The quest for biomarkers in TB
The potential of WBMGA	(Bok et al., 2021)	Whole blood mycobacterial growth assays for assessing human TB susceptibility: a systematic review and meta-Analysis
	(Eisen et al., 2013)	Effects of ascent to high altitude on human antimycobacterial immunity

Determining who is at risk of TB

Low- and middle-income countries (LMICs) carry the largest burden of TB incidence and control efforts (WHO, 2022). With limited resources to allocate to prevention efforts, it is important that these healthcare systems be able to target screening and preventative interventions to those most at risk. The WHO has identified several risk factors that raise the probability of

developing TB (WHO, 2022). These factors include HIV, undernutrition, diabetes, smoking, and alcohol consumption.

There are several other methods to estimate TB disease risk. One such system was developed by Saunders et al. (2017) using a score assigned in adult contacts of TB index cases to predict risk of TB. Adult contacts of index TB cases, aged 15 and older, were recruited in the desert shanty towns of Ventanilla, Peru from 2002-2006. Contacts were followed until February 2016. Using a Cox proportional hazards model, Saunders et al. identified risk factors and derived a score to classify contacts as low, medium, and high risk. This score was then validated in the urban community of Callao, Peru, just south of Ventanilla, from 2014-2015.

In total, 2017 contacts were identified from 715 index cases with a median follow-up of 10.7 years. 178 (9%) of contacts developed TB within the 19,147 person-years of follow-up. The risk factors used to derive the predictive score were body-mass index (BMI), previous TB, age, sustained exposure to the index case, the index case being in a male patient, lower community household socioeconomic position, indoor air pollution, previous TB among household members, and living in a poorly-ventilated household with a low number of windows per room. Overall, 27% of all contacts were classified as high risk. This group accounted for 60% of the contacts who later developed TB.

This study provided an adaptable method for predicting risk and targeting prevention in a LMIC setting. Saunders et al. acknowledged the limitations of the study including the reliance on self-report from index cases and contacts, potentially underestimating the prevalence of certain risk factors. And while the study offers a useful tool in initial contact investigations, there remains a need to have a standard biomarker with which to predict TB susceptibility.

Importance of biomarkers in disease control and eradication

Bonassi and Au (2001) highlighted the need for validated biomarkers to predict health risks in their review. Among the issues they identified are the stability of the biomarker with respect to time after the exposure, the interpretation of altered levels of a predictive biomarker at the individual level, and the need for validation of the biomarker.

Bonassi and Au also highlighted the limited size of study groups as a major limitation to identifying and validating biomarkers in epidemiological studies. The barriers to building appropriately sized samples include the low frequency of biological events and the high cost of required assays. Despite these issues and limitations, the authors predicted the move towards more personalized and targeted prevention of disease.

Importance of cost and ease of use in TB control

One of the barriers to establishing and implementing predictive biomarker surveillance for TB is cost. This financial barrier includes facilities, equipment, training, and staffing. An optimal biomarker is one that can be used on the ground in point-of-care testing which is affordable and simple to use (Lucas & Gaudieri, 2013). Brown et al. (2021) examined these barriers in the context of implementing GeneXpert testing in LMICs in their systematic review. They conducted a qualitative review of peer-reviewed articles published between 2010 and 2020. Eleven studies were included in the review.

Barriers to implementation that the articles identified included patient-level factors, human resources, material resources, service implementation, service coordination, and technical operations. The high cost of purchasing, using, and maintaining the machine led to the most common occurring barrier, the hub-and-spoke model. This model involves the use of a central

laboratory where community TB samples are sent and from which results are disseminated. This contrasts with the point-of-care model where testing is done near the point of collection in community health centers. In comparing the two models, Brown et al. cited that the hub-and-spoke model is associated with longer time frames for delivery of laboratory results which creates another barrier to proper engagement in the treatment pathway.

The authors acknowledged the limitations of their review, such as limiting their search to three databases and literature published in English. And while the article focused on GeneXpert, it highlights several important lessons for the development of a cost effective and accessible predictive biomarker test for TB. Specifically, there is a need for a point-of-care risk assessment method to accelerate and direct preventative interventions.

The search for TB biomarkers

The need for a predictive TB biomarker that is both affordable and simple is well established and has been the focus of intensive research. Jacobsen et al. (2008) focused their 2008 review on the need for additional biomarkers beyond IFNγ in order to reliably correlate with protection and susceptibility. With less than 10% of TB infected individuals developing active TB disease, it is imperative to be able to identify this subset to control and end this global epidemic.

Jacobsen et al. examined the available TB biomarkers including tuberculin skin tests (TST), IGRAs, and multifunctional T-cells. TST's usefulness as a biomarker is complicated by its low specificity due to cross-reaction with BCG vaccination, exposure to mycobacteria other than TB (MOTT), and its three-day wait period. It also has low sensitivity with immunosuppressed individuals such as those with HIV co-infections. IGRAs, on the other hand, can provide results overnight with higher sensitivity and specificity. However, analysis of IGRAs alone is

insufficient as a predictor of protection. Jacobsen et al. highlighted promising results from multifunctional T-cells for predicting protective immunity against TB, but stated that they are limited in their usefulness in the same manner as the other biomarkers described immediately above. This limitation is that they rely on a single organism to predict the course of a complex disease like TB. As a solution, Jacobsen et al. proposed using a systems level approach, what they termed "biosignatures" (Jacobsen et al., 2008). These biosignatures, they argued, have the potential to distinguish between different stages of infection, predict disease susceptibility, as well as treatment outcome. 15 years later, the need for such a biomarker for use at point-of-care is just as strong.

Parita and Kaufman (2010) in their review seconded the need for moving away from single markers to biosignatures as correlates of protection against TB. They continued to classify a clinically useful biomarker as fulfilling three criteria: it must "provide accurate, repeated measurements at reasonable cost and with a short turnaround time, provide information not available from clinical assessment, [and] assist in medical decision making" (Parida et al., 2014).

Parita and Kaufman proposed several methods of measuring these biomarkers including transcriptomics, epigenetics, proteomics, and metabolomics. They highlighted whole-blood bactericidal assays as a marker of the sterilizing effects of various drugs against mycobacteria. Through the combination of several methods and initiatives, they argued that immune markers will find increasing use and relevance in the field.

The potential of WBMGA

One such test that fits all the criteria mentioned in the previous sections is WBMGA. Bok et al. (2021) conducted a systematic review and meta-analysis evaluating the evidence that *in vitro*

results from WBMGA can predict *in vivo* TB susceptibility. Inclusion criteria were that the paper must be a peer-reviewed, English-language publications that described either cross-sectional, case-control, or cohort studies using WBMGA to study mycobacterial growth in human blood in relation to risk of TB infection, TB disease, or an established or possible TB risk factors. Quality of the studies was assessed with a quality assessment tool from the National Heart, Lung, and Blood Institute. No studies were found to directly assess if WBMGA results predicted TB susceptibility. However, 15 studies were included that assessed the association of WBMGA results with established and possible correlates of TB. Due to the lack of participant level data and the variety of statistical methods, Bok et al. conducted a meta-analysis to calculate the weighted means of mycobacterial growth.

The 15 studies included were separated into two categories. In the first category are studies assessing factors which are established as decreasing TB susceptibility. These factors include BCG vaccination, vitamin D, altitude, and HIV negativity/therapy. The second category was for studies assessing factors that likely affect TB susceptibility. These factors include TB infection, TB disease, and parasitism. Meta-analysis was used to derive relative mycobacterial growth ratios for BCG vaccination, TB infection, TB disease, and HIV infection. Parasitism, vitamin D, and altitude were deemed not amenable to meta-analysis.

The results of the meta-analysis showed that BCG vaccination resulted in a significant reduction in mycobacterial growth six months after vaccination. Mycobacterial growth was less in TB-infected populations compared with TB-uninfected populations. Mycobacterial growth was also significantly less for patients with TB disease when compared with uninfected individuals. Finally, mycobacterial growth was significantly less in immunocompetent (HIV-

negative and HIV-positive individuals receiving highly active antiretroviral therapy, HAART)) individuals when compared with untreated individuals with HIV-infection.

The indirect evidence assessed by Bok et al. in their review showed that less mycobacterial growth *in vitro* was usually associated with risk reducing factors *in vivo*. One area where this hypothesis was challenged was in the two studies that showed less mycobacterial growth in WBMGA in individuals with LTBI. One possible explanation offered by Bok et al. is that mycobacterial infection in the blood sample donor may provoke an immune response inhibiting mycobacterial growth in WBMGA. This WBMGA result, they continue, highlights the spectrum of LTBI and a possible use of WBMGA to stratify risk of LTBI advancing to active TB.

Bok et al. lay out the limitations of this review, specifically the lack of direct evidence to answer their research question as well as the heterogeneity of research methods that hampered direct comparisons of mycobacterial growth rates. They also state that half of the studies lacked a control group, which limited the precision of the meta-analysis.

WBMGA's research utility was demonstrated by Eisen et al. (2013) in assessing the effects of altitude on mycobacteria and antimycobacterial immunity. Blood samples were collected from 15 healthy, low-altitude adults from Lima, Peru (less than 100 meters above sea level) without a history of visits to high altitude or tuberculin skin test in the previous 6 months. These 15 participants then ascended to the high-altitude environment of Cusco, Peru (approximately 3,400 meters above sea level) to have a second blood sample taken. Blood samples were also collected from 47 permanent high-altitude residents of Cusco.

A minimum of 1.5 ml of blood was collected from each participant, a portion of which was spun down to make negative-control plasma due to its tuberculostatic quality. Luminescent *mycobacterium bovis* (BCG, Montreal strain) was mixed with samples to make blood assays and

negative-control assays for each participant. A positive control broth was also analyzed with each batch of approximately 10 participants. Luminescence readings were taken using an automated portable luminometer at baseline and at 96 hours.

Eisen et al. found in both low and high altitude that mycobacteria grew abundantly in positive-control culture broth, minimally in negative-control plasma, and moderately in whole blood. Table 3 shows the growth of each group of assays, the relative growth of whole blood to culture broth, as well as the relative growth of whole blood to plasma. Both the culture broth and the plasma showed increases in mycobacterial growth after ascent to high altitude with no relative difference to each other. At low altitude, mycobacteria growth in whole blood was similar to that in culture broth with the former only being 0.17 logs (1.5 times) less. After ascent to high altitude, mycobacterial growth in whole blood was restricted to 0.62 logs (4 times) the growth in the culture broth. Whole blood for the long-term Cusco residents also showed mycobacterial growth was restricted 0.65 logs (4 times) when compared to the culture broth. A similar degree of restriction was seen in whole blood relative to plasma (Figure 1). There was no apparent difference found in mycobacterial growth when comparing Lima residents in high altitude to long-term Cusco residents.

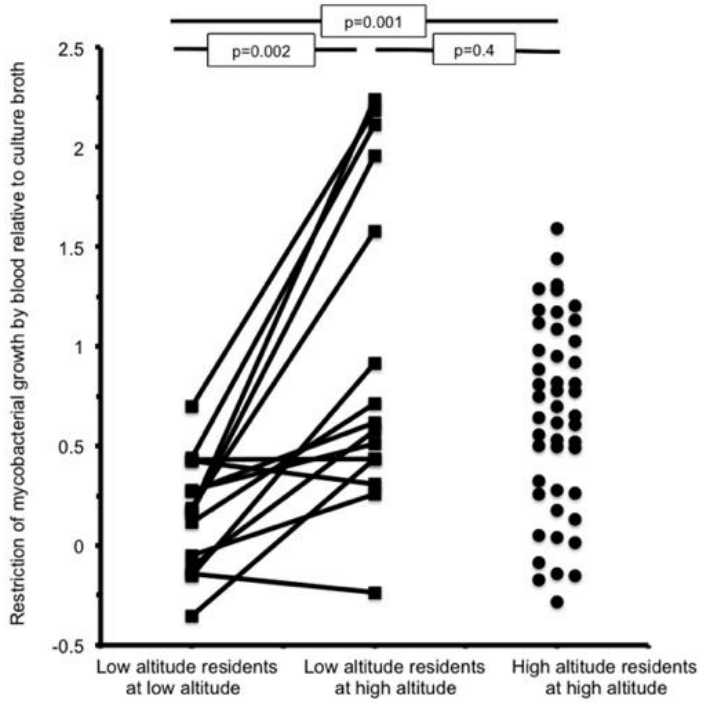
Table 3: Altitude effects on luminescence indicating mycobacterial growth* in blood, positive-control culture broth and negative-control plasma (Eisen et al., 2013)

	Mycobacterial growth			p-values		
	Low altitude residents at low altitude	Low altitude residents at high altitude	High altitude residents at high altitude	Low altitude residents: change on ascent to high altitude	At high altitude: low altitude residents versus high altitude residents	Low altitude residents at low altitude versus high altitude residents at high altitude
	(n=15)	(n=15)	(n=47)			
(a) MYCOBACTERIAL GROWTH						
Culture broth	1.1	1.9	1.5	<0.001	<0.001	<0.001
	[0.70, 1.1]	[1.9, 2.3]	[1.5, 1.7]			
Whole blood	0.86	1.2	0.95	0.3	0.7	0.5
	[0.53, 1.2]	[0.30, 1.4]	[0.48, 1.4]			
Plasma	-0.55	0.40	0.094	0.004	0.9	<0.001
	[-0.64, -0.37]	[-0.59, 0.70]	[-0.15, 0.44]			
Culture broth relative to plasma	1.5	1.4	1.5	0.2	0.3	0.8
	[1.3, 1.7]	[1.2, 2.8]	[1.2, 1.7]			
(b) ANTIMYCOBACTERIAL IMMUNITY						
Whole blood relative to culture broth	0.17	0.62	0.65	0.002	0.4	0.001
	[-0.10, 0.43]	[0.43, 2.0]	[0.26, 1.0]			
Whole blood relative to plasma	1.4	0.76	0.79	0.01	1.0	0.002
	[1.1, 1.4]	[0.50, 1.0]	[0.46, 1.2]			

*. Median [interquartile range] increase in luminescence measured in $\log_{10}$ relative light units in each medium. Unpaired data comparisons were made with Wilcoxon rank sum test and paired data comparisons were made with the Wilcoxon signed-rank test.

doi: 10.1371/journal.pone.0074220.t002

Figure 1: Restriction of mycobacterial growth by blood relative to culture broth (Eisen et al., 2013)



Eisen et al. acknowledged the limitations of their study to include a small sample size and the lack of adjustments for demographic differences in the Lima and Cusco residents. They asserted, however, that there was no association found between differences in demographic characteristics (such as age) and the results of the assay. The main strength of this assay is that it shows the feasibility of WBMGA in low-resource settings. Creating and reading the assay required only portable equipment such as a hand-held incubator, pipettes, and a briefcase sized luminometer. Quantifying mycobacterial growth was simplified through the use of genetically modified luminescent mycobacteria. Finally, the assay required only 1.5 ml of blood from subjects, removing a major barrier to participation.

The Eisen et al. study focused on antimycobacterial immunity in relation to altitude because altitude is known to influence the risk of TB disease. Eisen and colleagues studied

mycobacterial growth in whole blood, which may indicate antimycobacterial immunity but can also be influenced by diverse other factors that influence mycobacterial growth including environmental conditions in which the WBMGA was performed. Consequently, to more directly assess human antimycobacterial immunity, Eisen at al. also considered WBMGA assay results as the difference between mycobacterial growth in whole blood relative to concurrent mycobacterial growth in negative control plasma samples derived from each blood sample. The relative merits of these two approaches have not been well characterized.

Objectives

It is easy to imagine the potential public health impact if there were an established association between mycobacterial growth in WBMGA and risk of TB disease. The risk of progression to TB disease may be considered in terms of the risk of the recent exposure to TB: (1) immediately, over the next six months progressing to so-called early or 'incipient' TB; or (2) causing asymptomatic LTBI that later progresses to cause subsequent symptomatic incident TB disease. To test for such potential associations would require a prospective laboratory study determining WBMGA results, followed by an epidemiological study following the participants in the laboratory study to determine who developed symptomatic TB disease. As the foundation for this thesis, I was fortunate to collaborate with researchers at IFHAD who have generated such a dataset with both WBMGA and long-term TB epidemiological follow-up data. I was not involved in the generation of this dataset.

In this thesis, I present my analysis and interpretation of those data with the following objectives: (1) to determine whether WBMGA results are associated with the risk of incipient and/or subsequent incident TB disease; (2) to determine what factors affect this association; (3)

to determine which WBMGA results (whole blood alone or in relation to growth in negative control plasma or positive control broth) provide the most potential value as a biomarker for further study.

My hypothesis is that WBMGA results relative to concurrent mycobacterial growth in negative control plasma (as studied by Eisen et al.) will be positively associated with risk of incipient TB disease with minimal influence by operational variables but with a strong relationship to TB-associated factors.

Methods

Sample population

The anonymous and unlinked dataset used in this analysis was provided by IFHAD. First index cases, who did not contribute to this dataset, had been newly diagnosed with infectious (sputum smear microscopy positive) pulmonary TB by the 16 local community health centers serving all of Ventanilla, Lima Peru. In collaboration with the Peruvian Ministry of Health and those community health centers, the IFHAD research team visited each newly-diagnosed index case's household to recruit close contacts. This study concerned those close contacts, for whom inclusion criteria were age 15 years or older and close contact with a newly-diagnosed index case defined as being in the same house for more than six hours per week in the two weeks prior to the index case TB diagnosis. Exclusion criteria for these close contacts were a current diagnosis and/or treatment for TB disease or unwillingness or inability to give informed written consent. These participants were recruited from 2002-2006 and followed up until 2016.

Participants completed a baseline nurse assessment of TB risk factors. For the first six months after enrollment, research staff conducted follow-up visits every two weeks for the first two months and once a month for the following four months. During these visits, free TB testing was offered for contacts with a productive cough lasting more than two weeks. After six months, passive TB disease surveillance continued to be offered free of direct charges by the local health system in that anyone presenting to these health centers reporting symptoms suggestive of TB was offered free testing for TB disease. Additionally, the IFHAD research team offered free TB disease testing during follow-up household visits every four years (Saunders et al., 2017).

Immediately after recruiting each close contact who gave informed written consent, study staff collected a single blood sample that was used for a WBMGA. Some participants also provided subsequent follow-up samples. Using the same method that is described in the Eisen et al. study, a minimum of 1.5 ml of blood was collected from each participant and first separated into two aliquots, one of which was used as whole blood and the other equal aliquot was centrifuged to generate plasma for concurrent negative-control assays. The blood and plasma from each sample were both mixed with luminescent *mycobacterium bovis* (BCG, Montreal strain) for the WBMGA. As an additional positive-control, mycobacterial culture broth was created through the serial passage of BCG mycobacteria. The positive control broth was analyzed with each batch of approximately 10 participant samples. All samples were read at baseline and 96 hours for each participant using an automated luminometer.

Measuring mycobacterial growth

The first step of analysis was to evaluate mycobacterial growth in each medium (i.e. in blood, negative-control plasma and positive control culture broth). All measurements were made

in quadruplet and all luminometer readings were first converted to their base-10 logarithm. In samples where the mycobacterial luminescence was below the limit of detection, the midpoint of zero and the lower limit of detection was used for conversion to the base-10 logarithm. The mean luminescence was calculated at baseline and was also calculated for the readings 96 hours later. The mycobacterial growth in each WBMGA was then calculated as the mean luminescence at 96 hours minus the mean luminescence at baseline. A correlation analysis was then completed to determine which variables were the most relevant to the study.

Assessing the relationship of WBMGA results and operational variables

In order to test the utility of WBMGA as a predictor of TB susceptibility, operational variables were examined to test for associations with the assay results. The operational variables assessed were processing delays and mycobacterial strain characteristics. Processing delays were considered as a dichotomous variable: whether the sample was analyzed the same day as it was collected or the next day. Mycobacterial strain characteristics were also considered as a dichotomous variable: whether the mycobacterial stocks that were used to inoculate the WBMGA over the next 96 hours grew in broth cultures (as expected) or occasionally (for unknown reasons) failed to grow (i.e. failed to cause increasing luminescence over the 96 hours after inoculation).

Assessing the relationship of WBMGA results and variables associated with TB

The next step in examining WBMGA's association with TB susceptibility was to analyze other variables that might be involved. The first of these were demographic variables: age and sex. Also included were variables indicating history or previous contact with TB: previous TB

diagnosis, positive purified protein derivative (PPD) TST result suggestive of LTBI, presence of BCG scar(s) indicating past administration of intradermal BCG vaccination(s). Finally, body mass index (BMI) was included.

Incorporating long-term TB epidemiological follow-up data

Next, long-term TB epidemiological follow-up data was incorporated to see if the assay results predicted if participants went on to get TB disease. The outcomes of interest were incipient TB, which was defined as receiving a TB diagnosis within six months of enrollment, and incident TB, which was defined as receiving a TB diagnosis more than six months after enrollment.

Analysis Strategy

I performed the data analysis using Stata/BE 17.0. In the first part of the analysis, a generalized least squares (GLS) random effects model was used to estimate the relationship between operational and TB-associated variables with mycobacterial growth in WBMGAs as the dependent variable. This analysis was then rerun using mycobacterial growth in blood relative to plasma as the dependent variable in order to determine which assay results had the greater association and thus more potential as a biomarker. A two-tailed p-value of less than 0.05 was considered statistically significant. Significant results were selected as covariates for survival analysis.

Next, I conducted survival analyses using a Cox proportional hazards model. The outcomes of interest were developing incipient TB and incident TB. Incipient TB was limited to TB diagnoses within six months of enrollment. Participants not diagnosed with TB within that

window were right censored. Incident TB will be limited to TB diagnosis after six months from enrollment. Participants diagnosed with TB before six months were left censored. Participants not diagnosed with TB before the end of the available follow-up period (February, 2016) were right censored. The independent variables for incipient TB and incident TB were mycobacterial growth in blood and mycobacterial growth in blood relative to plasma. Each of these independent variables were considered as dichotomous variables: growth less than or equal to the median or growth above the median. In order to incorporate the multiple blood samples from some subjects, a variance estimator (in Stata: vce(cluster)) was used in the survival analysis.

If variables had a p-value less than 0.05, then they remained in the Cox model. Proportional hazards assumptions were assessed using a goodness of fit with Schoenfeld residuals with an alpha of 0.05.

Ethical considerations

The original research that generated this dataset had ethical approval in Peru, the United States, and the United Kingdom. All participants gave informed written consent. The secondary analysis that I have done for this thesis was conducted with the approval and in consultation with IFHAD. I did not participate in the collection or generation of the dataset. All data used was anonymous and unlinked. Emory IRB was consulted and it was determined that this thesis did not require IRB review.

Results

Participants

293 close contacts of index TB cases identified by the 16 local community health centers serving all of Ventanilla, Lima Peru were enrolled in the study between 2002-2006. 26 subjects were excluded from the analysis due to incomplete data (Table 4). 672 blood samples were collected from the remaining 266 participants. These samples represent baseline samples from all participants as well as additional samples from 167 participants approximately two weeks and/or six months after enrollment.

Table 4: Subjects included in analysis

Enrolled	Missing luminescent data	Baseline blood draw took place after secondary TB diagnosis	Missing enrollment date	Total included in analysis
293	3	1	23	266

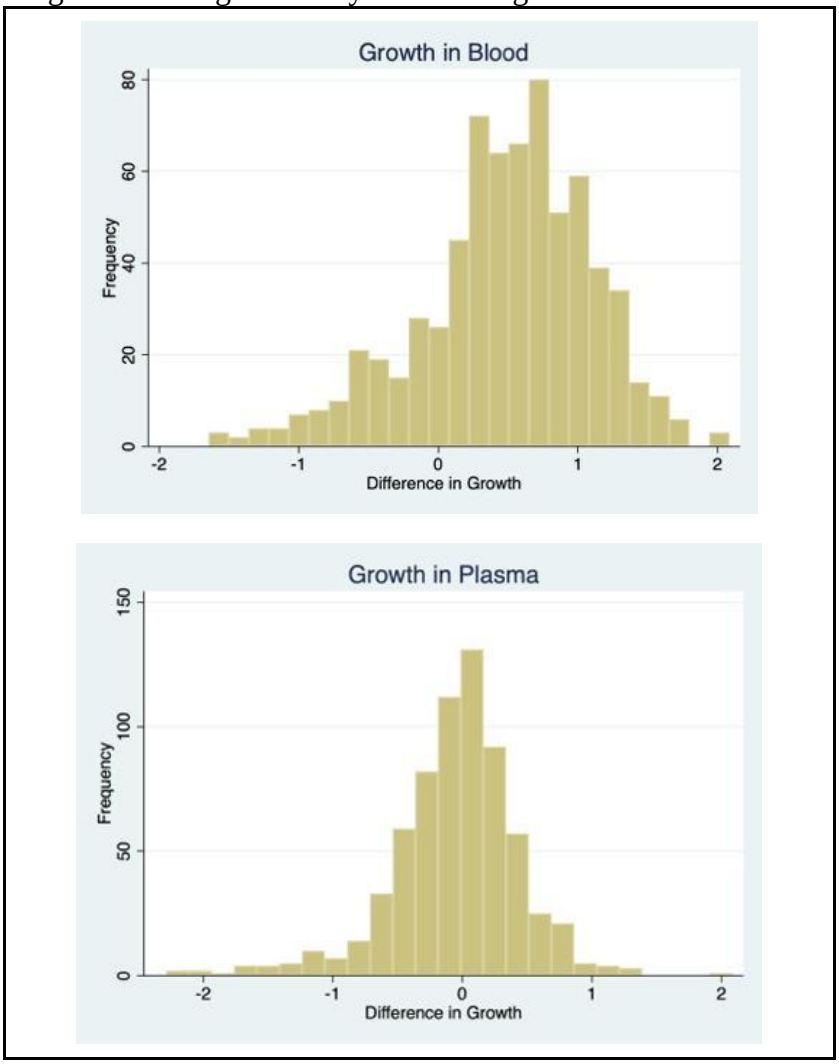
Mycobacterial growth in each medium

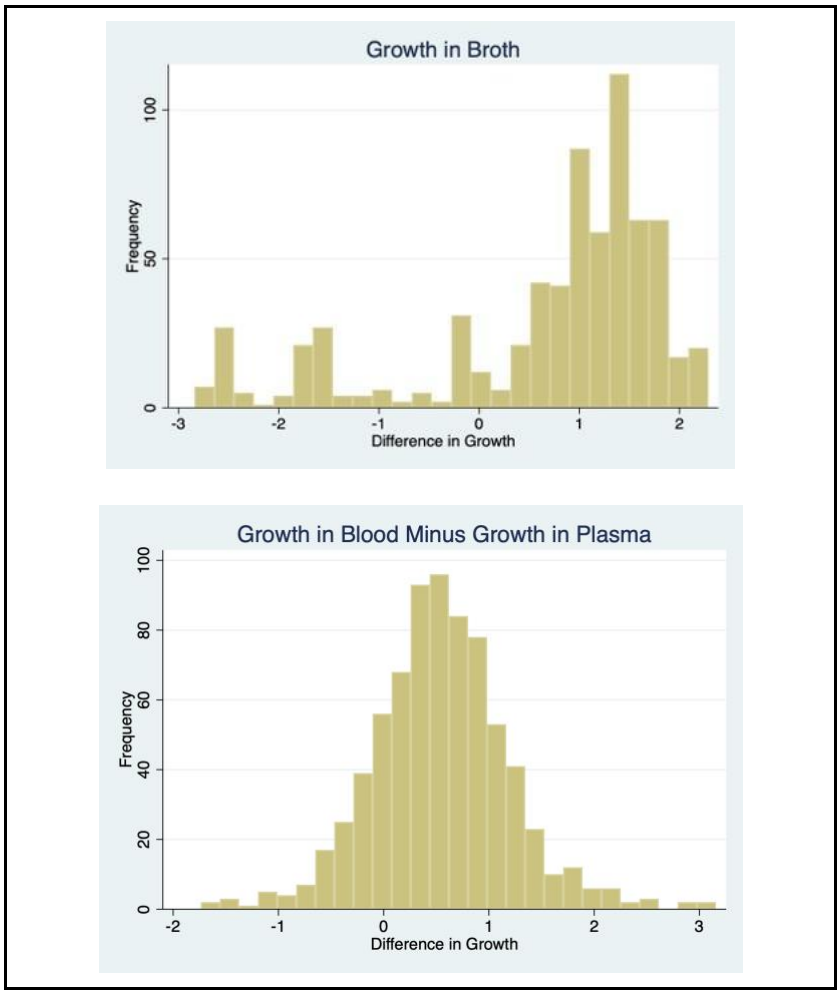
Table 5 shows the mycobacterial growth for each medium while Figure 2 shows the distribution of growth for each medium. Using the difference in luminescence at 96 hours relative to baseline, mycobacteria grew the most in positive culture broth (mean=0.670 relative light units, RLU), the least in negative culture plasma (mean=-0.054 RLU), and intermittently in blood (mean=0.480 RLU). The histogram for growth in positive culture broth has the most left skewness.

Table 5: Mycobacterial luminescence in RLU for each medium (mean values indicate the mean of base 10 logarithm transformed quadruplet readings)

Medium	Baseline *mean (SD)*	Baseline *median (IQR)*	96 hours *mean (SD)*	96 hours *median (IQR)*	Growth: 96 hour minus baseline *mean (SD)*	Growth: 96 hour minus baseline *median (IQR)*	Growth: difference relative to plasma *mean (SD)*	Growth: difference relative to plasma *median (IQR)*
Blood	0.169 (0.441)	0.229 (0.429)	0.649 (0.758)	0.761 (0.874)	0.480 (0.627)	0.544 (0.761)	0.532 (0.640)	0.538 (0.782)
Plasma	1.127 (0.506)	1.220 (0.368)	1.073 (0.716)	1.164 (0.786)	-0.054 (0.497)	-0.009 (0.505)		
Broth	1.081 (0.459)	1.161 (0.352)	1.750 (1.267)	2.156 (1.335)	0.670 (1.261)	1.071 (1.042)		

Figure 2: Histogram of mycobacterial growth in each medium

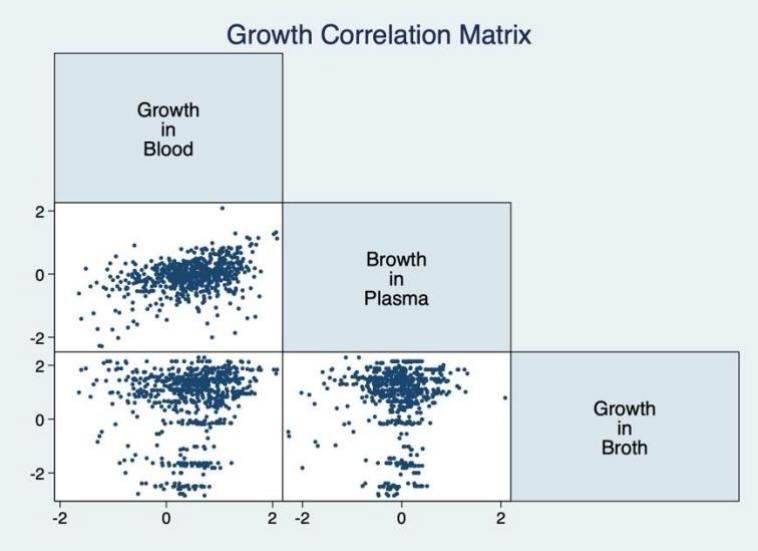




Selecting a comparison for growth in blood

A correlation analysis was conducted in order to test if the association between mycobacterial growth in blood and risk of secondary TB could be improved when measured relative to the negative control plasma or the positive control broth. A scatterplot of mycobacterial growth shows the highest correlation between blood and plasma (Figure 3). Based on this assessment, the difference of mycobacterial growth in blood relative to plasma was selected as a comparison to the results of the mycobacterial growth in blood.

Figure 3: Scatterplot of log difference in mycobacterial growth for each medium



WBMGA association with operational, demographic, and TB-associated variables

Table 6 displays the results from the secondary analysis of demographic, operational, and TB-associated variables, and their relationship between growth in blood and growth in blood relative to plasma. The only statistically significant relationship was between delays in process and mycobacterial growth in blood relative to plasma: for every 1 unit of change in mycobacterial growth in blood relative to plasma, there was a -.0197 unit of change in processing time with a p-value of 0.000.

Table 6: Results of examining the relationship of operational, demographic, and TB-associated variables. Associations with $p<0.05$ are shown in bold font.

	Growth in blood			Growth in blood minus growth in plasma		
	Coefficient	95% Confidence interval (CI)	P-value	Coefficient	95% CI	P-value
Operational variables						
Broth grew [vs not]	0.108	-0.077, 0.223	0.067	0.067	-0.053, 0.187	0.275

Same day assay [vs next day]	-0.097	-0.194, 0.002	0.054	**-0.197**	**-0.299, -0.096**	**0.000**
Demographic variables						
Age in years *median (IQR, N)*	0.001	-0.002, 0.0042	0.443	0.002	-0.001, 0.005	0.182
Male	0.025	-0.005, 0.055	0.103	0.000	-0.047, 0.047	0.996
TB-associated variables						
PPD positive	0.091	-0.004, 0.192	0.060	0.061	-0.043, 0.166	0.247
BCG scar present	-0.113	-0.260, 0.033	0.130	-0.053	-0.207, 0.101	0.500
BMI, mean (sd, N)	-0.003	-0.014, 0.007	0.542	-0.004	-0.015, 0.008	0.516
Previous TB, % (n/N)	-0.069	-0.167, 0.028	0.163	-0.050	-0.152, 0.051	0.332

Outcome of interest: Secondary TB

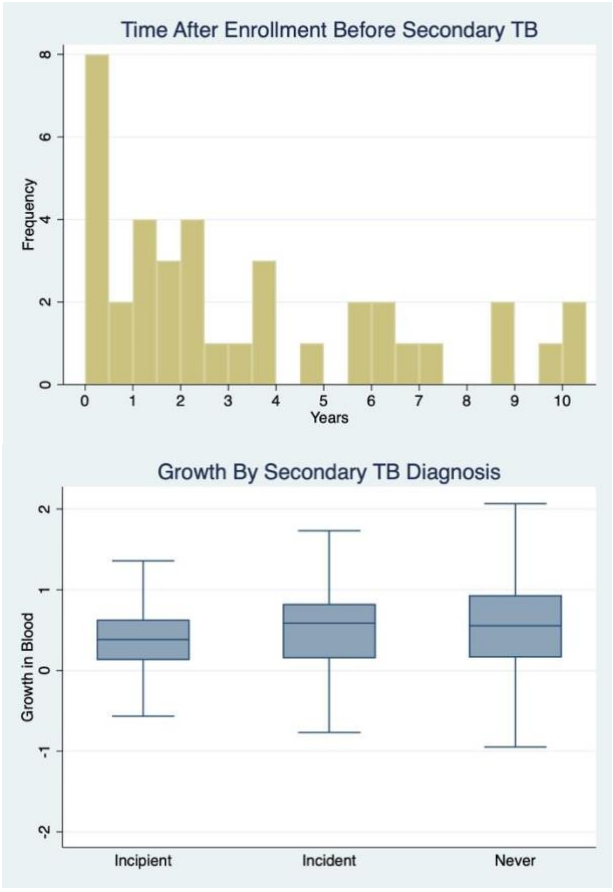
The long-term TB epidemiological follow-up data revealed a total of 38 secondary cases of TB per 5,689.95 person years in participants as of February, 2016 (Table 6). The distribution of secondary TB cases during the follow up period can be seen in Figure 4. The highest incidence of new TB cases occurred in the first six months after close contact with index TB cases.

Table 7: Number of individuals who developed secondary TB per person-time

Time	Number of participants who developed secondary TB	Person-years*	Rate per 100,000 person-years
Diagnosed within first 6 months (incipient)	8	103	7,767
Diagnosed after 6 months (incident)	30	1,633	1,837
Total	38	1,736	2,304

*Person-time was calculated in years as the difference between date of enrollment and date of secondary TB diagnosis

Figure 4: Secondary TB distribution



In using a GLS random effects model to examine the association between mycobacterial growth and the TB risk, no statistically significant association was found (Table 8).

Table 8: Result of examining the relationship of risk of TB disease with mycobacterial growth in blood and blood relative to plasma.

	Growth in blood			Growth in blood minus growth in plasma		
TB Disease Risk	Coefficient	95% CI	P-value	Coefficient	95% CI	P-value
Incipient TB, %	-0.183	-0.521, 0.154	0.287	-0.215	-0.546, 0.117	0.204
Incident TB, %	-0.046	-0.218, 0.126	0.598	-0.064	-0.233, 0.105	0.457

Survival Analysis

A survival analysis was done to calculate hazard ratios (HR) for incipient TB. Binary variables were used to describe mycobacterial growth in blood as well as growth in blood relative to growth in plasma, in both cases using the median growth as the division. When added to the Cox model, the process delay variable was not statistically significant so was removed. Mycobacterial growth in blood relative to plasma at or below the median log difference in luminosity of 0.53 RLU was significantly associated with a 3.0 (95% CI=1.1, 8.0) times increase in risk of developing TB disease within 6 months (p=0.031, Table 8). The hazard ratios for mycobacterial growth in blood with incipient TB and both hazard ratios for incident TB were not found to be statistically significant.

Table 9: Survival analysis for mycobacterial growth in blood and growth in blood relative to plasma as dichotomous variables

	Growth in Blood Below Median				Growth in Blood - Growth in Plasma Below Median			
Time-to-event	Number of Events	Hazard Ratio	95% CI	P- Value	Number of Events	Hazard Ratio	95% CI	P- Value
To predict incipient TB	20 events	1.633	0.947, 2.8918	0.078	**20 events**	**2.979**	**1.103, 8.044**	**0.031**
To predict incident TB	103 events	0.936	0.606, 1.447	0.767	103 events	0.909	0.583, 1.417	0.672

Proportional hazards assumption for incipient TB

The results of the proportional hazards assumption test using Schoenfeld residuals for incipient TB are displayed in Table 9. Both results had a p-value greater than the alpha of 0.05, so the proportional hazards assumptions appear not to have been violated.

Table 10: Testing the proportional hazards assumption using Schoenfeld residuals test for blood and blood relative to plasma

	Growth in Blood			Growth in Blood - Growth in Plasma		
	Chi^2	df	P-Value	Chi^2	df	P- Value
To predict incipient TB	0.220	**1**	**0.636**	**0.100**	**1**	**0.752**

Discussion

This thesis had three main objectives: (1) to determine whether WBMGA results were associated with the risk of incipient and/or subsequent incident TB disease; (2) to determine what factors affected this association; (3) to determine which WBMGA results (whole blood alone or in relation to growth in negative control plasma or positive control broth) provided the most potential as a biomarker for further study.

Regression analysis using a random effects model showed no statistically significant association between mycobacterial growth in either blood or blood relative to plasma and secondary TB (incipient or incident). In order to investigate this future, I created a binary variable for mycobacterial growth and turned to survival analysis since it tends to have more statistical power for examining associations using binary outcomes (George et al., 2014). I felt this increase in statistical power might be necessary given the small sample size of only 8 subjects with incipient TB. Survival analysis showed a statistically significant hazard ratio for

mycobacterial growth in blood relative to plasma which was associated with a 3.0 (95% CI=1.1, 8.0) times increase in risk of developing TB disease within 6 months (p=0.031, Table 9). The 95% confidence interval of (1.103, 8.044) was quite large, but the result still supported the hypothesis that mycobacterial growth *in vitro* is associated with an individual's risk of developing incipient TB *in vivo.*

In examining the variables that might influence this association, the only significant result was between processing delays for mycobacterial growth in blood relative to plasma. This variable was not significant when included in the Cox survival model so was therefore removed. The results showed a larger than expected effect size for WBMGA with a hazard ratio of 3.0. It merits future research to see how a larger sample with a longer follow up period affects the statistical significance of WBMGA's relationship with risk of TB disease in the random effects model.

In terms of how best to analyze the WBMGA, the analysis of mycobacterial growth in blood had relatively similar results to the analysis of mycobacterial growth in blood relative to plasma. Mycobacterial growth in blood relative to plasma was statistically associated with processing delays (p=0.000) while growth in blood showed borderline significance of p=0.078 (Table 6). Mycobacterial growth in blood relative to plasma in the survival analysis predicted incipient TB with a statistically significant hazard ratio (p=0.031) whereas the hazard ratio for growth in blood approached but did not reach the threshold for statistical significance of p=0.05 (Table 6).

Given the overall goal of this thesis is to aid in the development of a biomarker test that can be used in the communities most at risk of TB disease, it is important to acknowledge despite the performance of mycobacterial growth in blood relative to plasma, separating plasma from

blood, performing concurrent assays in plasma as well as blood, and calculating results for blood minus plasma all increase costs, time required, and complexity. Therefore, I recommend that future research continue to analyze WBMGA only in blood in addition to using concurrent negative control plasma in order to prioritize simplicity and functionality in this near-patient relatively low-technology functional assay.

Limitations

There were several limitations to this analysis that future research into this topic might address. The first was the relatively small size of the study group with only 291 subjects included in the analysis. As previously stated by Bonassi and Au, this limited the power of analysis to find associations and validate WBMGA as a useful biomarker. I attempted to compensate for this limitation by creating a binary variable for WBMGA and using survival analysis which tends to have more statistical power. (George et al., 2014) Future research would benefit from increasing the study population and increasing the diversity of the populations studied.

Another limitation of the analysis was that the long-term TB epidemiological follow-up data was only collected until 2016. With only 38 confirmed cases of TB, updating the dataset to include diagnostic data from the more recent seven years would provide insight not only to WBMGA's association to incipient TB, but also potentially incident TB as well.

Additionally, this sample represented individuals at high risk of TB infection given their proximity to an index case. It would be beneficial to test for an association of WBMGA and TB susceptibility in a sample more reflective of the risks of the general public, although this would require a much larger study.

Conclusion

There is no perfect test to assess TB susceptibility, but this analysis shows that WBMGA deserves a place in the development of diagnostic and predictive tests. While we attempt to stay one step ahead of the increasing drug resistance of TB pharmacologically, expanding our understanding of risk would help to address the TB pandemic. Most excitingly, WBMGA offers the chance for low resource healthcare systems to more efficiently direct resources and services to those who are most at risk of TB disease. It is in these places that the battle to end TB is most important and this research suggests that the WBMGA has the potential to predict the risk of incipient TB disease, and hence contribute to focusing screening and preventive interventions on those individuals most likely to benefit from them.

References

Bok, J., Hofland, R. W., & Evans, C. A. (2021). Whole Blood Mycobacterial Growth Assays for Assessing Human Tuberculosis Susceptibility: A Systematic Review and Meta-Analysis. *Front Immunol, 12*, 641082. https://doi.org/10.3389/fimmu.2021.641082

Bonassi, S., & Au, W. W. (2002). Biomarkers in molecular epidemiology studies for health risk prediction. *Mutation Research/Reviews in Mutation Research, 511*(1), 73-86. https://doi.org/https://doi.org/10.1016/S1383-5742(02)00003-0

Brown, S., Leavy, J. E., & Jancey, J. (2021). Implementation of GeneXpert for TB Testing in Low- and Middle-Income Countries: A Systematic Review. *Global Health: Science and Practice, 9*(3), 698-710. https://doi.org/10.9745/ghsp-d-21-00121

Eisen, S., Pealing, L., Aldridge, R. W., Siedner, M. J., Necochea, A., Leybell, I., . . . Evans, C. A. (2013). Effects of Ascent to High Altitude on Human Antimycobacterial Immunity. *PLoS ONE*, *8*(9), e74220. https://doi.org/10.1371/journal.pone.0074220

George, B., Seals, S., & Aban, I. (2014). Survival analysis and regression models. *Journal of Nuclear Cardiology*, *21*(4), 686-694. https://doi.org/10.1007/s12350-014-9908-2

Harries, A. D., Kumar, A. M. V., Satyanarayana, S., Thekkur, P., Lin, Y., Dlodlo, R. A., . . . Zachariah, R. (2020). The Growing Importance of Tuberculosis Preventive Therapy and How Research and Innovation Can Enhance Its Implementation on the Ground. *Tropical Medicine and Infectious Disease*, *5*(2), 61. https://doi.org/10.3390/tropicalmed5020061

Jacobsen, M., Mattow, J., Repsilber, D., & Kaufmann, S. H. E. (2008). Novel strategies to identify biomarkers in tuberculosis. *389*(5), 487-495. https://doi.org/doi:10.1515/BC.2008.053

Lucas, M., & Gaudieri, S. (2013). The role of biomarkers in the prevention of globally important diseases. *Biomarkers in Medicine*, *7*(3), 331-332. https://doi.org/10.2217/bmm.13.54

Parida, A., Bairy, K. L., Chogtu, B., Magazine, R., & Vidyasagar, S. (2014). Comparison of Directly Observed Treatment Short Course (DOTS) with Self-Administered Therapy in Pulmonary Tuberculosis in Udupi District of Southern India. *J Clin Diagn Res*, *8*(8), Hc29-31. https://doi.org/10.7860/jcdr/2014/8865.4721

Parida, S. K., & Kaufmann, S. H. E. (2010). The quest for biomarkers in tuberculosis. *Drug Discovery Today*, *15*(3), 148-157. https://doi.org/https://doi.org/10.1016/j.drudis.2009.10.005

Saunders, M. J., Wingfield, T., Tovar, M. A., Baldwin, M. R., Datta, S., Zevallos, K., . . . Evans, C. A. (2017). A score to predict and stratify risk of tuberculosis in adult contacts of

tuberculosis index cases: a prospective derivation and external validation cohort study. *Lancet Infect Dis, 17*(11), 1190-1199. https://doi.org/10.1016/s1473-3099(17)30447-4

World Health Organization (WHO) (2022). *Global Tuberculosis Report 2022*. https://www.who.int/publications/i/item/9789240061729