Imperial College London

Sputum viability microscopy for tuberculosis

Dr Sumona Datta

Imperial College London, UK
Department of Infectious Disease
PhD

2021

ACKNOWLEDGEMENTS

""We know very little, and yet it is astonishing that we know so much, and still more astonishing that so little knowledge can give us so much power."

Bertrand Russell

The work presented in this thesis is the culmination of investigations, experiments and much thought over the span of many, many, many years. Consequently in reflection of this, there are many, many, many people to thank.

Firstly, I would like to thank all the study participants who took the time to provide samples and answer various questionnaires despite the stigma they experience regularly due to TB. Your answers and insights inspire me to continue searching for solutions that may be of some benefit.

I would like to thank the Innovation for Health and Development (IFHAD) team in Peru, past and present. Particular appreciation goes to: Teresa Valencia, Eric Ramos, Keren Alvarado, Willy Quino, Rosario Montoya, Rosario Sosa, Luz Quevedo Cruz, Pilar Tapia, Carlos Pozo, Marie Haro, Jessica Franco, Enrique Flores, Adelina Lozano, Carolina Tovar, Janett Quispe, Silvia Carrera, Maribel Rivero, Jonathan Gomez, Lourdes Bocanegra, Matthew Saunders, Louis Grandjean, Laura Martin and Jonathan Sherman. All of whom were involved in the development of FDA microscopy and/or the data collection. There are then the members or associates of the IFHAD team who were essential to my pastoral care: Tom Wingfield, Lena Shah and Luz Antonia Cardenas.

I would like to thank all other colleagues, family and friends who I have not mentioned by name but took the time to think and discuss my work and aspirations, hopefully you know who you are. Regardless of whether your comments were with excitement or apathy, it helped shaped my thinking and the work presented here.

I am deeply grateful to the Wellcome Trust Institutional Strategic Support Fund (ISSF) at Imperial College London and The Wellcome Trust for thinking my career and ideas were worth investment. Imperial College London have always been considerate and flexible and for that I thank Tammy Barrett, Jacob Lee and Hayley Kendall. I would like to end this paragraph by thanking Mike Levin. Your support and kindness from my ISSF interview and then as my Imperial College London based supervisor is greatly valued.

To Bob Gilman, your enthusiasm for my work was motivating and your care for me a huge support. You are a true mentor and a rare find.

To my parents, Rina and Satya, who believe in me completely. Thank you for your ongoing encouragement as you nod to every idea and read every abstract along the way.

To Carlton, the reason I ever started this particular journey. Your curiosity and optimism helped me persevere and see the importance of every finding.

Finally, to Atlas, whose arrival urged me to finish this chapter of my life and gave me renewed excitement for new ventures.

STATEMENT OF ORIGINALITY

I declare that the thesis titled "**Sputum viability microscopy for tuberculosis"** is my own original research. No part of this thesis has been previously submitted for any other degree. All work is my own and any other sources of information are appropriately referenced.

My role in the original research presented in this thesis is summarised below:

Research chapter	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Conception and design	No	No	No	Yes	Yes	Yes	Yes
Acquisition of data	No	Led	No	Yes	Led	Led	Led
Analysis	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interpretation	Yes	Yes	Yes	Yes	Yes	Yes	Yes

'No' indicates that I had no significant role because I was using pre-existing data.

'Yes' indicates that I had a primary, leading role, personally doing the great majority of this activity.

'Led' indicates that I had a primary, leading role, fulfilling the great majority of the co-ordination and management of this activity, as I supervised the majority of day-to-day collection of data, such as patient surveys and performing laboratory tests but I only personally did a small proportion of the actual day-to-day collection of data, such as patient surveys and performing laboratory tests.

Sumona Datta
Date: 16th February 2021

COPYRIGHT DECLARATION

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

Introduction. Sputum microscopy is the most frequently used tuberculosis laboratory test, but poorly assesses infectiousness and treatment response. The viability stain fluorescein diacetate (FDA) identifies metabolically active cells.

Objective. To optimise sputum microscopy for predicting tuberculosis infectiousness and treatment outcome.

Foundations. FDA sputum microscopy rapidly indicated the concentration of *Mycobacterium tuberculosis* that had not been sterilised by boiling and that was culturable.

A cohort of 35 patients had FDA microscopy performed on their pre-treatment sputum. Lower FDA microscopy results predicted higher risk of tuberculosis disease in their 209 household contacts (adjusted hazard ratio=3.9, $p=0.02$).

During the first 9 days of first-line therapy, these 35 patients had FDA microscopy on repeated sputum samples that rapidly differentiated patients with drug-susceptible tuberculosis from those with multi-drug resistant tuberculosis ($p<0.001$).

Refinement. To optimise sputum collection for programmatic use, a network meta-analysis of 23 studies demonstrated that pooled sputum, as used in the foundation studies, had similar performance to more easily collected, instructed on-demand "spot" sputum collection ($p=1.0$).

FDA microscopy was optimised for use in basic clinical laboratories. This replaced potentially biohazardous and expensive steps with more practicable alternatives.

Evaluation. These refined protocols were used to collect pre-treatment sputum samples from 978 patients. Higher mycobacterial concentrations assessed by culture, PCR, conventional or FDA microscopy all predicted greater risk of tuberculosis disease in their 4311 household contacts ($p<0.0001$), with concordance (C)-statistics between 58-62%.

Paired samples, pre-treatment and after 14 days of tuberculosis treatment, were collected from 384 patients. An alert FDA microscopy result at 14 days of treatment predicted adverse treatment outcomes (relative risk=3.0, $p<0.0001$), and performed better than currently used prognostic markers with 42% sensitivity and 85% specificity.

Conclusion. After methodological optimisation, FDA microscopy predicted patient infectiousness and treatment outcome with similar or greater reliability than currently used laboratory tests.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS 1

STATEMENT OF ORIGINALITY 2

COPYRIGHT DECLARATION 3

ABSTRACT 4

TABLE OF CONTENTS 5

TABLE OF ABBREVIATIONS 6

LIST OF TABLES 7

LIST OF FIGURES 8

INTRODUCTION 10

RESEARCH CHAPTER 1 15

Foundations – TB viability microscopy validation

RESEARCH CHAPTER 2 23

Foundations – Pre-treatment TB viability microscopy assessing infectiousness

RESEARCH CHAPTER 3 39

Foundations – TB viability microscopy assessing treatment response

RESEARCH CHAPTER 4 56

Refinement – Sputum collection optimisation for TB testing

RESEARCH CHAPTER 5 85

Refinement - TB viability microscopy optimisation

RESEARCH CHAPTER 6 109

Evaluation – Pre-treatment TB viability microscopy assessing infectiousness

RESEARCH CHAPTER 7 132

Evaluation – TB viability microscopy assessing treatment outcome

GENERAL DISCUSSION 152

CONCLUSION 157

REFERENCES 158

APPENDIX 171

TABLE OF ABBREVIATIONS

Abbreviation	**Definitions**
AFB	Acid-fast bacilli in microscopy
BMI	Body mass index
CFU	Colony forming units in quantitative culture
CT	Cycle times in polymerase chain reaction
DNA	Deoxyribonucleic acid
FDA	Fluorescein diacetate
HIV	Human immunodeficiency virus
HR	Hazard ratio
IQR	Interquartile range
LED	Light-emitting diode
Log	Base-10 logarithm
MDR	Multi drug resistance
MGIT	Mycobacteria growth indicator tube
MODS	Microscopic observation drug susceptibility test
NPV	Negative predictive value
NTP	National TB programme
OR	Odds ratio
PBS	Phosphate buffer solution
PCR	Polymerase chain reaction
PPV	Positive predictive value
RNA	Ribonucleic acid
SD	Standard deviation
SOP	Standard operating procedure
TB	Tuberculosis
TEMA	Tetrazolium microplate assay
TTP	Time to positivity in culture
WHO	World Health Organisation
ZN	Ziehl-Neelsen
95% CI	95% confidence intervals

LIST OF TABLES

Table 1.1 Linear regressions

Table 2.1 Relationship between patient characteristics and laboratory results
Table 2.2 Characteristics predicting secondary TB disease in contacts
Table 2.3 Patient FDA microscopy predicting secondary TB disease in contacts

Table 3.1 Baseline patient and laboratory data
Table 3.2 Regression analysis of laboratory results during early treatment (data from triplicate microscopy slides)
Table 3.3 Regression analysis of laboratory results during early treatment (data from the first microscopy slide)

Table 4.1 Study demographics
Table 4.2 Summary of the main findings of pairwise and network meta-analyses
Table 4.3 Sputum collection instructions
Table 4.4 Meta-regression to explain heterogeneity between studies comparing early morning versus spot sputum collection
Table 4.5 Metaregression to explain variability within reference group results
Table 4.6 Study quality as assessed by the QUADAS-2 checklist
Table 4.7 Diagnostic yield of GeneXpert versus sputum microscopy
Table 4.8 Comparison of the number of publications in 2012-2017

Table 5.1 Published FDA protocols
Table 5.2 Optimisation study protocols and results
Table 5.3 Validation study
Table 5.4 Cost analysis

Table 6.1 Mean values of each test by demographics and TB disease type
Table 6.2 Factor analysis to identify potential subpopulations
Table 6.3 Cox proportional hazards model
Table 6.4 Test performance to predict secondary TB

Table 7.1 Definition of treatment groups
Table 7.2 Patient demographic characteristics and treatment
Table 7.3 Treatment outcome by number of key drugs prescribed during the interval between the Day 0 and the Day 14 sputum collections
Table 7.4 Day 14 laboratory results and reported symptoms
Table 7.5 Test performance in predicting adverse outcome

LIST OF FIGURES

Figure 0.1 TB diagnostics routinely available in Callao, Peru

Figure 1.1 Line graphs

Figure 2.1 Study flow chart
Figure 2.2 Distribution of patients' pre-treatment sputum results
Figure 2.3 Time-to-event analyses
Figure 2.4 Time-to-event analysis of the interaction between FDA microscopy and culture results

Figure 3.1 Difference between quantitative FDA microscopy and culture results (data from triplicate microscopy slides)
Figure 3.2 Treatment response for patients with non-MDR TB (data from triplicate microscopy slides)
Figure 3.3 Treatment response for patients with MDR TB (data from triplicate microscopy slides)
Figure 3.4 Comparison of results for first microscopy slides versus geometric means
Figure 3.5 Difference between quantitative FDA microscopy and culture results (data from the first microscopy slide)
Figure 3.6 Treatment response for patients with non-MDR TB (data from the first microscopy slide)
Figure 3.7 Treatment response for patients with MDR TB (data from the first microscopy slide)

Figure 4.1 Study flow chart
Figure 4.2 Forest plot for sputum microscopy
Figure 4.3 Forest plot for mycobacterial culture
Figure 4.4 Network meta-analysis
Figure 4.5 Bar chart of mycobacterial load
Figure 4.6 Funnel plots

Figure 5.1 Cording colonies
Figure 5.2 Optimisation study distribution of quantitative results
Figure 5.3 Optimisation study comparison with conventional sputum microscopy and culture
Figure 5.4 Photographs of each FDA microscopy protocol
Figure 5.5 Validation study
Figure 5.6 Storage study

Figure 6.1 English translated instructions used for sputum sample collection
Figure 6.2 Time to positivity versus colony-forming units in thin layer agar
Figure 6.3 Study flow chart
Figure 6.4 Distribution of results

Figure 6.5 Comparison of results from each test
Figure 6.6 Time-to-event analyses
Figure 6.7 Time-to-event analysis of the interaction between FDA microscopy and culture results
Figure 6.8 Permutations simulation results

Figure 7.1 Sample pathway
Figure 7.2 Study flow chart
Figure 7.3 Pie chart of treatment outcomes
Figure 7.4 Venn Diagram of Day 14 laboratory results
Figure 7.5 Bar graph of adverse outcomes by laboratory and clinical results
Figure 7.6 Contingency table results
Figure 7.7 Graphical display of post-test probabilities of an adverse treatment outcome

Introduction

The World Health Organization (WHO) estimates that there were 10 million cases of tuberculosis (TB) disease, causing 1.4 million deaths in 2019.[1] Globally, TB elimination is hampered by inadequate laboratory testing, and increasing rates of multi-drug resistant (MDR)-TB constitutes a public health crisis.[1] TB laboratory tests have 4 main clinical objectives: (a) diagnosing TB; (b) identifying drug-resistant TB; (c) assessing treatment response; and (d) predicting infectiousness.

TB drug-susceptibility testing, especially gold standard culture-based testing has limited availability globally because of inadequate laboratory infrastructure and skills. Furthermore, the vast majority of TB occurs in resource-constrained settings where it is currently difficult to afford culture-based or molecular tests. Therefore, globally the only test that many patients with TB have access to is sputum microscopy.[2–5] This test needs only basic expertise, equipment and provides same-day results. At the turn of the millennium, the WHO "STOP TB" plan included and by 2015 met the global target of 1 microscopy centre per 100,000 people.[6] However, conventional TB sputum microscopy using either Ziehl-Neelsen (ZN) or auramine stains has limitations for the 4 main objectives of TB testing described above because it is: (a) insensitive; (b) cannot identify drug-resistant TB; (c) cannot reliably assess early treatment response; and (d) poorly predicts infectiousness.[7]

The presence and state of *Mycobacterium tuberculosis* (*M. tuberculosis*) in sputum may be assessed in various ways. Conventional acid-fast microscopy with ZN or auramine staining are the most used techniques and detect the presence of *M. tuberculosis* bacteria. However, they have poor sensitivity and are not specific to *M. tuberculosis*, staining other mycobacteria and acid-fast organisms. Various polymerase chain reaction (PCR) based tests detect the presence of *M. tuberculosis* deoxyribonucleic acid (DNA), usually with higher sensitivity than microscopy. However, both conventional acid-fast microscopy and PCR techniques detect *M. tuberculosis* whether it is viable (a term used here to indicate that it is culturable, as demonstrated by ability to replicate under certain conditions) or is non-viable (e.g. has had chemical sterilisation or been killed by boiling or by antibiotics). Thus, neither conventional acid-fast microscopy nor DNA-based PCR methods can discriminate viable versus non-viable *M. tuberculosis* bacilli. The presence of non-viable *M. tuberculosis* in sputum causes some patients to have persistently positive sputum microscopy or PCR results despite adequately responding to several months of tuberculosis treatment. These persistently positive laboratory tests throughout adequate treatment may inappropriately cause patients to have their treatment modified and/or extended and

may also cause lengthy restrictive infection control precautions.[8,9] Thus, the inability of these rapid tests to differentiate viable versus non-viable *M. tuberculosis* hampers their role in assessing treatment response and guiding infection control measures.

Fluorescein diacetate (FDA) is a vital stain that generally causes metabolically active cells to fluoresce because non-specific esterase in the cytoplasm must be present to hydrolyse the stain to a fluorescent form.[10] FDA was first demonstrated to indicate esterase activity in mammalian cells by Rotmann and Papermaster, followed by Medzon and Brady's study demonstrating its application in numerous bacteria in 1969.[11,12] Since then it has been used in various fields of biology to assess cell viability, especially for organisms that are difficult to culture. For example, FDA has been used to monitor treatment response in patients receiving therapy for leprosy, because *M. lepra* is difficult to culture.[10,13] In the 1980s, FDA was demonstrated to stain *M. tuberculosis*, and was subsequently shown with clinical specimens to predict within 1 hour the results of TB culture that would usually be available weeks later.[14–16]

GLOBAL OBJECTIVE: To optimise sputum microscopy for predicting TB infectiousness and treatment outcome.

HYPOTHESIS 1: FDA microscopy quantitative results can assess the concentration of culturable *M. tuberculosis* in sputum.

The findings of this experiment are discussed in Research Chapter 1 and confirmed that FDA microscopy results predicted both quantitative TB culture results and the proportion of unsterilised *M. tuberculosis* within a sample. Consequently, hypotheses regarding the potential clinical utility of FDA microscopy were generated.

Household contacts of a cohort of newly diagnosed patients in a foundation study were followed up for 6 years to detect secondary TB disease in contacts to assess:

HYPOTHESIS 2: FDA sputum microscopy in pre-treatment samples can assess patient infectiousness.

The findings of this study testing hypothesis 2 are discussed in Research Chapter 2, which was published in the Journal of Infectious Diseases.[17]

In 2006, a group in Bangladesh reported that sputum microscopy with FDA could be used to identify culture-confirmed treatment failure in patients who had positive conventional sputum microscopy results after 2 months of first-line treatment.[18] They then implemented

in 4 regional laboratories sputum FDA microscopy for patients suspected of having failed first-line treatment. This predicted multi-drug resistant (MDR) TB with 93% accuracy, and led to 23% more patients switching to appropriate second-line treatment earlier than with the previous use of reference laboratory culture results.[19] Newly diagnosed patients recruited in the foundation study provided sputum at multiple time points to assess:

HYPOTHESIS 3: FDA sputum microscopy can predict whether patients are responding to therapy early in treatment.

Hypothesis 3 was assessed and described in Research Chapter 3, that was published in the journal Clinical Infectious Diseases.[20]

The Wellcome Trust awarded a Clinical Research Training Fellowship to further evaluate FDA microscopy. Prior to starting this evaluation, the protocol for sputum collection was refined. During the foundation study, sputum was collected over a 12-hour period, which may not be feasible in the resource-constrained settings where most TB occurs.[21] To make sputum collection for FDA microscopy more practical, a systematic review and meta-analysis was done to assess:

HYPOTHESIS 4: Sputum collection methods can be optimised to increase the positivity rate of microscopy.

The findings of this study testing hypothesis 4 are discussed in Research Chapter 4, and were published in the journal Lancet Global Health.[22]

Until recently FDA microscopy used fluorescence that required expensive mercury or halogen light sources, which previously limited feasibility for routine use in clinical settings. Furthermore, in the foundation study described in Research Chapters 2 and 3, the FDA staining protocol used sputum decontamination with centrifugation. Few microscopy centres in resource-constrained settings would be able to carry out this procedure, especially with the WHO-recommended high centrifugal force, hermetically sealed centrifuges and biosafety level-III laboratory facilities.[23,24] There have been various FDA microscopy protocols published, which were evaluated to test:

HYPOTHESIS 5: The FDA microscopy protocol can be simplified and optimised for routine use in microscopy centres in resource-constrained settings.

The results of the experiments testing hypothesis 5 are discussed in Research Chapter 5, and led to the publication of a standard operating procedure (SOP) for FDA microscopy in the journal PLOS One.[25]

Research Chapter 6 and 7 used the optimised sputum collection technique and FDA protocol described above to focus on the main clinical evaluation of FDA microscopy to assess if:

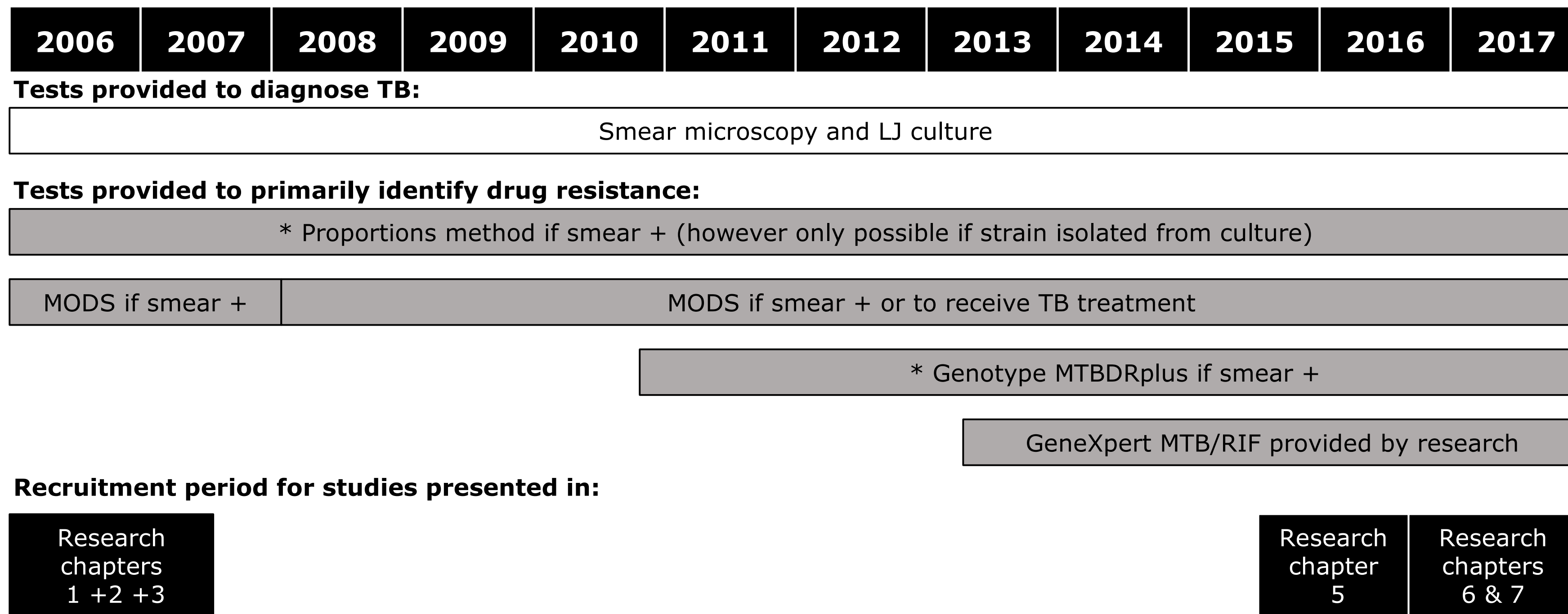
HYPOTHESIS 6: FDA microscopy performed under operational conditions can identify patients most likely to transmit TB; and

HYPOTHESIS 7: FDA microscopy early in treatment can be used to predict adverse clinical outcome versus long-term cure from TB.

The final Chapter concludes this body of work by integrating all the findings and discussing their relevance and potential areas for future research.

All the research presented was carried out in Callao, Perú. Since 1990, the Peruvian National TB Programme (NTP) has continually strengthened their healthcare service delivery particularly their diagnostics in the states of Lima and Callao (Figure 0.1). During 2006-2007, the NTP adopted the Microscopic Observational Drug Susceptibility (MODS) assay in regional laboratories that had biosafety level 3 facilities. MODS is a liquid culture technique providing integrated TB diagnosis with direct isoniazid and rifampicin drug susceptibility testing that takes up to 21 days.[26,27] Peru is the only NTP in the world that has adopted this technology for routine clinical care. In 2013 all research participants in our study setting were provided rapid PCR testing with the Cepheid GeneXpert MTB/RIF test, from which clinicians received clinically important results approximately 24 hours after sample reception. Other new diagnostics have also been adopted by the Peruvian NTP. However, these were generally only used in the national reference laboratory for positive cultures from selected patients so had little impact on the time to start treatment for drug resistant TB, in contrast to the universal provision of MODS and GeneXpert.[28] Thus, the diagnostic landscape for rapidly identifying MDR-TB has changed throughout the course of this research, as summarised in Figure 0.1.

Figure 0.1 TB diagnostics routinely available in Callao, Peru. Note *=tests provided at the national reference laboratory, LJ=Lowenstein-Jensen, and +=positive.



RESEARCH CHAPTER 1

Foundations – TB viability microscopy validation

Background

There are no appropriate-technology rapid tests for quantifying culturable *M. tuberculosis* in sputum. The most commonly used method in TB research is quantitative culture, which takes weeks and is labour intensive. There are 3 main culture-based approaches used to quantify the amount of viable *M. tuberculosis* in a sample, as detailed below.

1. Colony counts. Counting the number of colonies visible after a defined interval, usually 4- or 6-weeks incubation, may be considered the reference standard approach. If this is done only with the undiluted sample then high concentrations of colonies may coalesce, impairing accuracy. Thus, colony counts are often performed concurrently on undiluted sample and serial dilutions of the sample, to increase accuracy. However, this approach is labour intensive to replicate e.g. triplicate cultures of each concentration are generally required for accuracy, increasing expense.

2. Time to positivity. This approach utilises the fact that shorter time from culture inoculation until culture positivity (termed time to positivity, TTP) predicts higher colony counts and is frequently used in liquid culture methods. TTP is generally easier and simpler to perform than colony counts, so is a popular approach for quantifying the amount of viable *M. tuberculosis* in a sample. First, colony count experiments are performed on multiple samples under standardised conditions, noting the time taken for each undiluted culture to become culture positive. These data are used to determine the mathematical association between TTP and colony counts under these experimental conditions. Then, the TTP of subsequent samples can be used to estimate colony counts and hence the amount of viable *M. tuberculosis* in a sample.[29–31] A disadvantage of this approach is that either automated equipment must be used to determine the TTP, or cultures must be checked for positivity frequently, e.g. daily, which is also labour intensive.

3. Most probable number. For this approach, serial (usually 10-fold) dilutions of the sample are inoculated into multiple (usually triplicate small-volume broth) cultures. The most probable number of bacteria in the sample is statistically estimated using the principle of extinction dilution from the proportion of cultures of each dilution that become culture-positive by a defined interval. This approach is frequently used to quantify bacteria in water and food and has also been reported for *M. tuberculosis.*[32]

All the above approaches are technically demanding using infrastructure that is infrequently available in low and middle income countries and they provide results several weeks after sample processing, which are often out of date for clinical decision making.[33] It should also be noted that all of these approaches directly or indirectly estimate the concentrations of colony forming units (CFU) of bacteria. *M. tuberculosis* usually exists in sputum samples in clumps or chains, so the concentration of *M. tuberculosis* bacteria in sputum is generally much greater than the concentration of CFU that is determined by the above techniques.

Rapid molecular tests that quantify ribonucleic acid (RNA) from *M. tuberculosis* are possible but not feasible in most settings.[34] Quantitative reverse transcription polymerase chain reaction (qRT-PCR) tests have prohibitive costs and are technically challenging due to the instability of RNA, sensitivity to inhibitors, and variable interpretation of real time data.[35] Furthermore, qRT-PCR does not directly assess the number of living cells within the sample but measures cellular activity. In contrast, staining cells to indicate their viability with the use of microscopy may allow quantification of individual living cells. As described in the introduction, FDA may be a suitable candidate viability stain that can be safely used in resource-constrained settings.

The objective of this laboratory experiment was to validate the use of TB viability microscopy with FDA for identifying and quantifying culturable *M. tuberculosis*.

Methods

Participants were adult patients newly diagnosed with pulmonary tuberculosis and who had not received any anti-mycobacterial treatment and who gave informed written consent to participate. The study had approval from ethics committees including Asociacion Benefica PRISMA, Peru and Imperial College London UK.

Samples. A pre-treatment sputum sample was collected from patients who had recent acid-fast microscopy results with a range of positivity: one patient with smear grade +, another with sputum smear grade ++ and a third patient with sputum smear grade +++.[36] These sputum samples were processed concurrently for comparison with an aliquot of reference strain H37RV *M. tuberculosis.* This pure tuberculosis strain was diluted in phosphate buffered saline (PBS) at pH 6.8 to an optical density corresponding to the McFarland grade 1, indicating a TB concentration of approximately 300,000 CFU/μl.[37]

Sputum processing. Sputum was decontaminated by briefly vortexing 2 ml of sputum with an equal volume of 4% sodium hydroxide containing 2.9% sodium citrate and 0.5% N-acetyl-L-cysteine, as described.[37] After 20 minutes at room temperature excess PBS at pH 6.8 was added and the mixture was centrifuged for 15 minutes at 3,000 gravities.[37] The pellet was re-suspended in PBS at pH 6.8 to the original volume of 2 ml. The H37RV strain and decontaminated pellet suspensions were split into equal aliquots: one kept at ambient temperature and the other sterilised by immersion of the tube in boiling water for 1 hour.[38] The aliquots were mixed in a range of proportions to create TB suspensions that were 0%, 90%, 99% and 100% heat-sterilised. All underwent triplicate acid-fast microscopy, triplicate FDA microscopy, and triplicate quantitative cultures.

Slide preparation. Reagents were obtained from Fisher Scientific, USA, except where otherwise stated. To improve adherence of sputum to slides, the decontaminated samples for microscopy were mixed with 10 μl of 2.0% bovine serum albumin. To standardise smear density, sputa were smeared over areas marked on the underside of each slide. All smears were then air dried at room temperature and heat fixed by passing the slide through a flame. For acid-fast staining, 10 μl decontaminated sputum was smeared over a 2 cm^2 area. For FDA staining, 100 μl decontaminated sputum was smeared within a 1 cm^2 square, so the sputum density was 20-times greater than for acid-fast staining.

FDA staining. A fresh working solution of 20 μg/ml FDA was prepared daily by dissolving 10 μl of a 5 mg/ml FDA stock solution (Sigma) in 2.5 ml of 40% acetone in PBS at pH 6.8. A 1 cm^2 square of Whatman grade 3 filter paper was soaked in this working solution and placed on the dried sputum sample and incubated at 37ºC for 20 minutes. The filter

paper was discarded, the slide allowed to dry in the dark for 10 minutes and microscopy was then performed immediately with a Nikon Eclipse E600 fluorescent source (excitation filter 480nm) microscope that had a mercury bulb using 100-times magnification.[39]

Conventional acid-fast staining. Smears were flooded with 0.1% auramine for 15 minutes, decolourised with 0.5% hydrochloric acid-ethanol (acid-alcohol) solution for 2 minutes, rinsed with distilled water, flooded with 0.5% potassium permanganate for 30 seconds, dried in the dark and microscopy performed within 6 hours using 100-times magnification.

Cultures were done with the Microscopic-Observation Drug-Susceptibility (MODS) technique,[26,27] adapted to provide quantitative results in 24-well culture plates, as described.[33] A 1:10 dilution of 50 µl decontaminated sputum in 450 µl MODS culture broth (Middlebrook 7H9 liquid culture medium with enrichment and antimicrobial supplements[33]) to a final volume of 500 µl was mixed by pipetting. Then 50 µl of this was mixed into another well to a final volume of 500 µl making a 1:100 dilution. This was repeated again to make a 1: 1,000 dilution. Cultures were incubated in un-supplemented air at 37ºC. The sealed cultures were examined 3-times a week with an inverted light microscope using a 40-times magnification for cording colonies to determine time to positivity.[33]

The threshold for positivity was calculated to be the concentration of *M. tuberculosis* in the original sample corresponding to a positive result for microscopy or culture. The threshold of positivity for microscopy was ≥10 staining bacteria per 100 consecutive microscopy fields, according to policy during the time of this study in Peru.[40] The threshold of positivity for culture was >1 colony forming unit seen.

Analysis. Staff were blinded to other data from the same sample. Concentrations of bacteria or CFU visualised per ml of sample were calculated from manual counts. These concentrations were exponentially distributed. All concentrations are summarised as geometric means with 95% confidence intervals (95% CI). For all analyses of concentrations, zero values were changed to the mid-point between zero and the detection threshold and then concentrations were transformed to their base-10 logarithm (log). Spearman's test was used to calculate non-parametric correlation coefficients (rho). Linear regression used random effects to adjust for between-sample variation. All analyses were 2-tailed, performed with STATA version 11 (StataCorp, Texas, USA). All data are reported to 2 significant figures.

Results

As shown in Figure 1.1, the proportion of the sample that had been heat-sterilised correlated with the concentrations of bacilli visualised in FDA microscopy and CFU in quantitative culture (rho=0.84, $p<0.001$ and rho=0.72, $p<0.001$, respectively). All 100% heat-sterilised samples were culture-negative, and 11 out of 12 slides had <10 bacilli per hundred fields visualised by FDA microscopy. All 100% heat-sterilised samples were strongly auramine microscopy positive.

As shown in Table 1.1, linear regression demonstrated that with each 90% increase in the proportion of heat-sterilised sputum, mean FDA positive bacilli/ml reduced by 84% (95% CI=71-91%, $p<0.001$) and mean CFU/ml in culture reduced by 92% (95% CI=72-98%, $p<0.001$). In contrast, heat sterilisation only had a small effect on auramine microscopy results. Specifically, for every 90% increase in the proportion of heat-sterilised sputum, auramine positive bacilli/ml reduced by 12% (95% CI=4.1-19%, $p=0.04$), which was less than was demonstrated in both FDA microscopy and quantitative culture (both $p<0.001$, Figure 1.1).

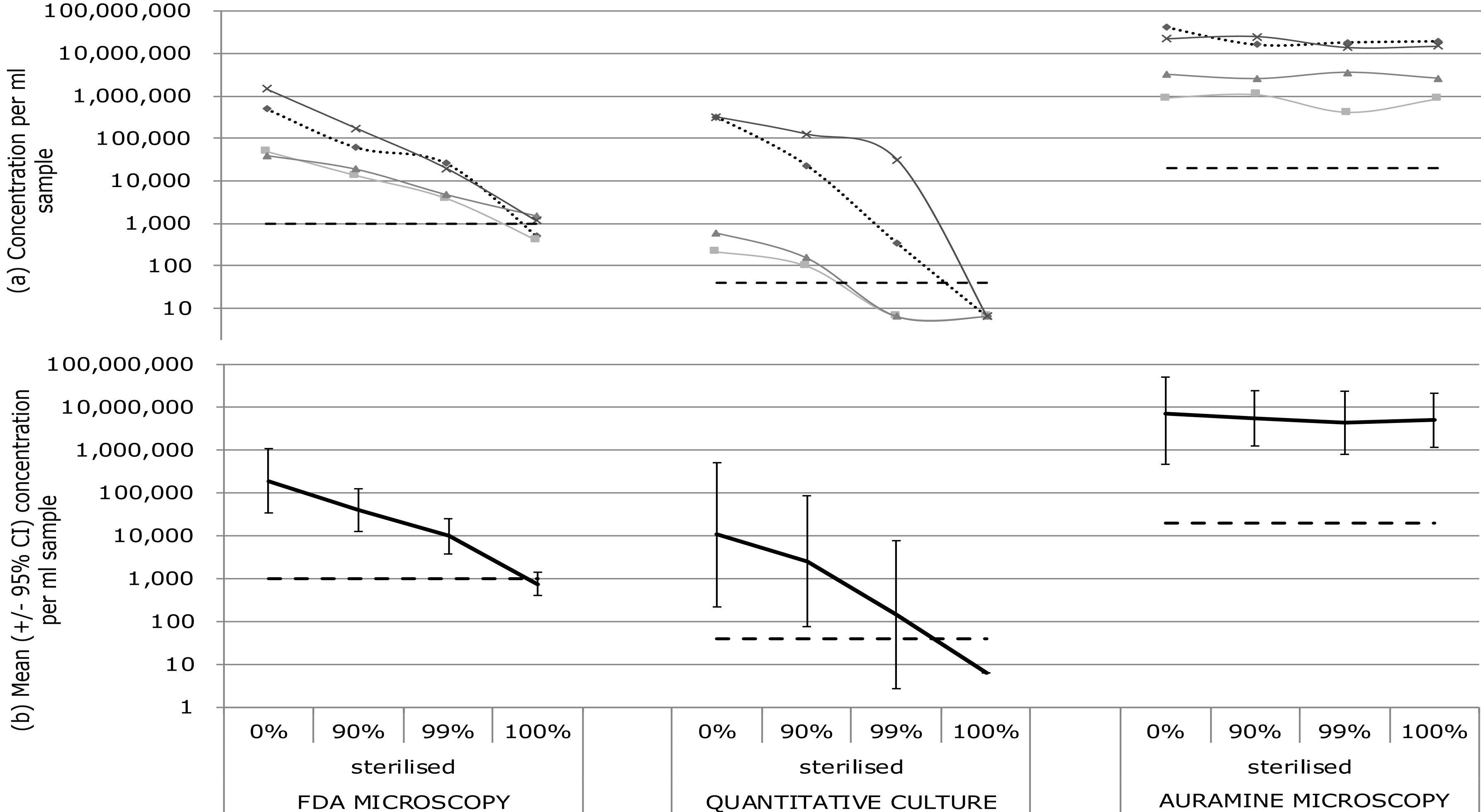
Table 1. 1 Linear regressions

Regression analysis demonstrating the change in *M. tuberculosis* fluorescein diacetate (FDA) microscopy, quantitative culture and auramine microscopy results for aliquots with increasing proportions that were heat sterilised. Linear regression with random effects was used to adjust for sample variation. Values in bold have associations with $p<0.05$. Concentration calculations were made from the geometric mean of the triplicate microscopy and quantitative culture data and were log transformed prior to regression analysis (see Methods). The coefficients indicate the differences in base-10 logarithms of concentrations compared to 0% sterilised samples. Note: log=base 10 logarithm; 95% CI=95% confidence interval.

	Difference in log concentration from the unsterilised sample								
	FDA microscopy			Quantitative culture			Auramine microscopy		
% of sample sterilised	Coefficient	95% CI	p value	Coefficient	95% CI	p value	Coefficient	95% CI	p value
90	**-0.68**	**-1.00 to -0.37**	**<0.001**	**-0.62**	**-1.00 to -0.21**	**0.003**	-0.11	-0.35 to 0.14	0.4
99	**-1.3**	**-1.7 to -0.85**	**<0.001**	**-1.9**	**-2.8 to -0.95**	**<0.001**	**-0.21**	**-0.41 to -0.0089**	**0.04**
100	**-2.4**	**-3.3 to -1.5**	**<0.001**	**-3.2**	**-5.1 to -1.3**	**<0.001**	**-0.15**	**-0.30 to -0.014**	**0.03**

Figure 1. 1 Line graphs

Graphs showing the geometric mean concentration per millilitre (y-axis) of triplicate counts of fluorescein diacetate (FDA) microscopy-positive bacteria, colony forming units in quantitative culture and auramine microscopy-positive bacteria in unsterilised, 90% sterilised, 99% sterilised and 100% sterilised samples for: (a) 3 sputum and 1 laboratory strain sample (dotted line); (b) the geometric mean data for these 4 samples with 95% confidence intervals (95% CI). The x-axis shows percentage sterilisation. The horizontal broken line indicates the cut off for a positive result (see Methods).



Discussion

This experiment confirmed that the concentration of FDA microscopy-positive bacilli correlated with the concentration of culturable *M. tuberculosis* in samples, and both indicated the proportion of the sputum that had been heat-sterilised. FDA microscopy took less than 1 hour and was less resource-intensive, whereas quantitative culture took weeks to provide results and required more laboratory equipment and expertise. Quantitative conventional acid-fast microscopy results with auramine staining were only slightly affected by proportion of heat-sterilisation.

Boiling for more than 30 minutes has been shown to kill *M. tuberculosis* by lysis, which is in keeping with the negative culture results shown in this study for the 100% sterilised samples.[41] Clinical samples that were 100% sterilised had 1-3 bacilli visualised with FDA microscopy. This may be explained by artefacts rather than mycobacteria being visualised, incomplete sterilisation caused by poor penetration of heat in large volume samples or large clumps of mycobacteria.[42,43] Alternatively, these low concentrations of FDA-positive cells seen on FDA microscopy of boiled samples may indicate false-positive FDA microscopy results for sterilised cells, possibly because of residual non-specific esterase activity persisting in cells that had been killed by boiling. Even though *M. tuberculosis* was killed by boiling, its structure was retained as demonstrated by bacilli in conventional acid-fast microscopy having similar morphology and concentrations for unsterilised versus boiled samples. *M. tuberculosis* can also be killed by chemical and mechanical lysis, which have also been shown to impact both quantitative culture and FDA results.[41,44]

Mycobacterial growth is complex and dependent on its physiological state, which may permit culture in some conditions and not in others.[45] Differential growth or "viable but not culturable" dormant *M. tuberculosis* has been observed, and is thought to be an important population regarding TB latency and treatment relapse.[46] The experiment described here is not a model for studying populations of *M. tuberculosis* that are on the spectrum between fully metabolically active and sterilised, and cannot determine the ability of FDA microscopy to identify viable but not culturable *M. tuberculosis*. However, we can conclude that FDA microscopy does not generally visualise killed *M. tuberculosis* with retained morphology and can within 1 hour of sample processing provide results that accurately predict quantitative culture results that become available weeks later.

Research Chapter 2

Foundations – Pre-treatment TB viability microscopy assessing infectiousness

Background

Patients with pulmonary TB expectorate sputum that contains heterogeneous subpopulations of *M. tuberculosis,*[47] which has unknown implications for assessing patient infectiousness.[48] Currently, predictions of TB infectiousness are largely based on conventional acid-fast sputum microscopy that are used to guide decisions about infection control measures and to prioritise which patients' contacts should receive screening, preventative therapy, and surveillance.[49,50] As demonstrated in Research Chapter 1, conventional acid-fast microscopy stains *M. tuberculosis* bacilli whether they are culturable or sterilised. Furthermore, *in vivo* and *in vitro* experiments have suggested that *M. tuberculosis* bacteria have a slowly metabolizing phenotype produced by growth under stress-inducing conditions that may in fact be most infectious, and are difficult to detect.[51–54]

The hypothesis of this study is that FDA microscopy of pre-treatment sputum would identify the most infectious patients. To test this hypothesis, newly diagnosed patient's pre-treatment sputum was assessed with FDA microscopy, and their household contacts were followed-up and screened for TB disease.

Methods

Ethics. All participants gave informed written consent. The internationally accredited ethics committee of the Universidad Peruana Cayetano Heredia approved the study. TB-related care for patients and their close contacts including diagnostic testing, treatment and preventative therapy is provided free of direct charge in Peru and was not delayed or modified by study participation.

Background TB notification rates. Recruitment took place in shantytowns in Callao, Peru, where the average TB notification rate in these participating clinics during the study period was 194 cases/100,000 people/year. According to local practice, we estimated that the background case notification rate in these communities was 20% higher i.e. 233 cases per 100,000 people/year because some cases were treated by other providers.[55–57]

Patients and Samples. Patients were interviewed at recruitment from 2006 to 2007. Sputum was collected over 12 hours in patients' homes at room temperature, transported at 4°C and decontaminated within 24 hours by mixing 2 ml sputum with 2 ml of combined 4% sodium hydroxide, 2.9% sodium citrate, 0.5% N-acetyl-L-cysteine.[37] Reagents were obtained from Sigma, USA. After 20 minutes, excess PBS (pH 6.8) was added, centrifuged (15 minutes, 3,000 gravities), the pellet re-suspended in 2 ml PBS, immediately smeared onto microscope slides with 10 µl bovine serum albumin for adhesion and dried at room temperature. Samples then underwent FDA microscopy (see below), quantitative culture and acid-fast microscopy, each in triplicate. This standardised decontamination protocol was used and the exact smear area marked on slides with a hydrophobic pen to reduce processing variability between the triplicate smears.

FDA microscopy. 100 µl decontaminated sputum was smeared over 1 cm^2. Note that this was 20-times greater density on the microscopy slide than for acid-fast microscopy. FDA staining used a stock solution of 5 mg FDA that had been dissolved in 1 ml of a solution of 40% acetone in PBS. This stock solution was diluted daily to 20 µg/ml in the same solution of 40% acetone in PBS and soaked onto a 1 cm^2 square of Whatman grade 3 filter paper that was incubated on the smear at 37ºC for 20 minutes.[39] The paper was discarded, the slide allowed to dry in the dark for 10 minutes and microscopy performed immediately. Concurrent non-viable counter-staining was not used.

Conventional acid-fast microscopy. 10 µl decontaminated sputum was smeared over 2 cm^2, air-dried at room temperature, heat-fixed by passing through a flame and auramine stained by flooding with 0.1% auramine (for 15 minutes); 0.5% acid-alcohol decolorizing

solution (for 2 minutes); rinsing; flooding with 0.5% potassium permanganate (for 2 minutes); rinsing; drying; storing in the dark and performing microscopy within 6 hours.[37]

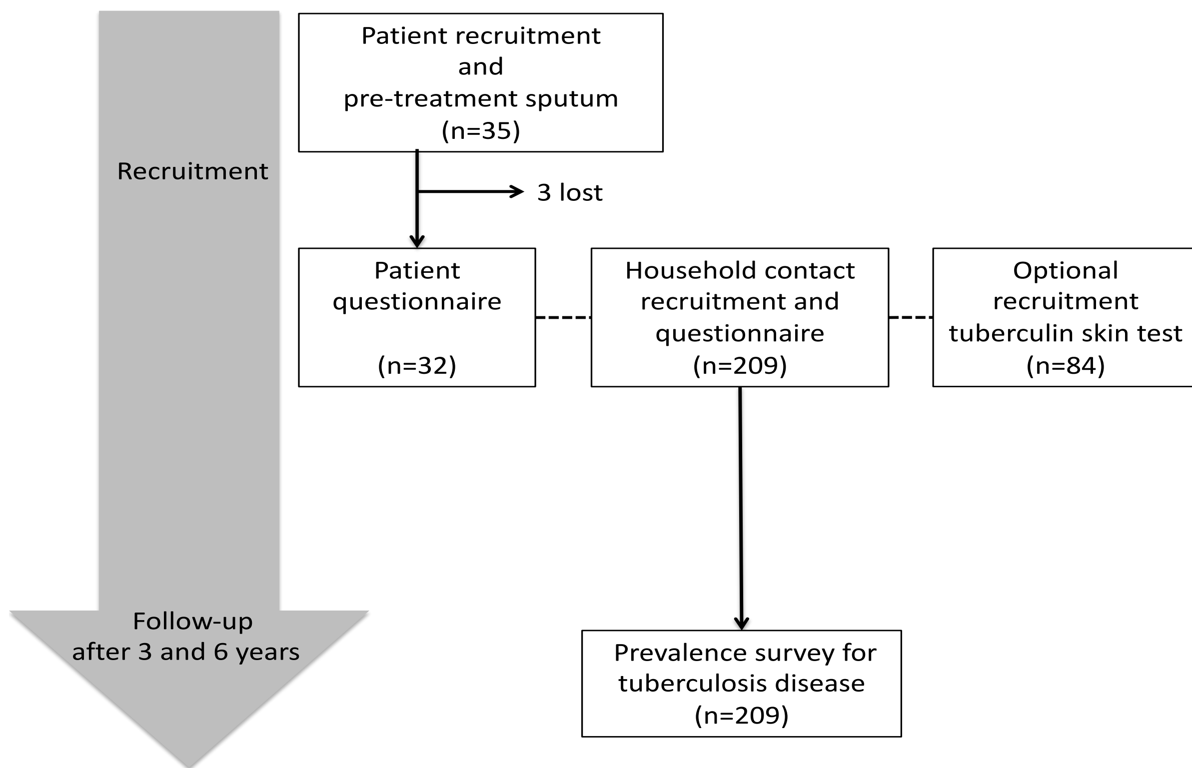
Microscopy. Visible bacteria were counted in 100-consecutive microscopy fields at 1,000-times magnification (constituting an area of approximately 2 mm^2).[58] Concentrations of stained bacteria per ml of sample were calculated and are termed 'FDA microscopy' and 'auramine microscopy'.

Cultures used the microscopic-observation drug-susceptibility (MODS) technique[26,27] adapted to be quantitative, as described.[33] A 1:10 dilution of 50 µl decontaminated sputum in 450 µl culture broth[26] was mixed by pipetting, 50 µl was mixed into a 1:100 dilution well, and then into a 1:1,000 dilution well. Cultures were sealed in un-supplemented room air, incubated at 37ºC and examined with an inverted microscope for cording colonies to determine time to positivity. This is labour intensive so was performed 3-times weekly,[33] thus TTP data was poorly discriminatory compared with if it had been performed continuously, as in the Mycobacteria growth indicator tube (MGIT) commercial technique, or by manual observation more frequently. Consequently the primary quantification of *M. tuberculosis* growth was CFU as described,[33] divided by the sample volume (ml) inoculated into that culture, i.e. CFU/ml, termed 'quantitative culture'. Positive cultures were confirmed to be *M. tuberculosis* with the Capilia assay (Tauns, Tokyo, Japan).

Drug-susceptibility testing. Isoniazid and rifampicin susceptibility of pre-treatment samples were analysed with the standard non-quantitative MODS technique[26,27] and the tetrazolium microplate assay (TEMA).[59]

Patients' contacts were invited to participate if they spent ≥6 hours/week in the patient's household in the 2 weeks prior to the patient's diagnosis.[60] Patients and contacts were interviewed to record symptom and demographic data, they were weighed, had their height measured and for adults (≥18 years old), body mass index (BMI) was calculated. Socioeconomic status was assessed by a composite household index in arbitrary units derived by principal component analysis from 13 variables characterising education, housing conditions, basic services and assets, as described.[60,61] Radiographs were not routinely performed so data were not available.[40] The study flow diagram is shown in Figure 2.1.

Figure 2. 1 Study flow chart



Contact TB preventative therapy. Isoniazid preventative therapy was provided for 6 months free of charge by the Peruvian national TB program (NTP) to contacts aged <15 years old, once the NTP had excluded active TB disease, regardless of any tuberculin skin-test results.[40] Older contacts, especially those aged 15-19 years old, were also eligible to receive preventative therapy at the discretion of the NTP clinician.[40]

Contact tuberculin skin testing. When the patient with newly diagnosed TB was commencing therapy, contacts aged ≥15 years in their household were asked at the time of patient recruitment to undergo screening for asymptomatic latent TB infection with tuberculin skin-testing, as described.[62,63] However, to encourage participation our study did not require tuberculin skin-testing.

Contact TB prevalence surveys. At the time of recruitment during 2006-2007, and again during follow-up visits approximately 3 and 6 years later until 2013, participants were asked whether they had been diagnosed with TB, as illustrated in Figure 2.1. Self-reported TB episodes were confirmed by checking NTP records. At these recruitment and follow-up visits, contacts were also asked whether they had symptoms suggestive of TB disease. Contacts who reported currently having cough, fever, night sweats or weight loss were requested to provide a sputum sample that we tested with conventional acid-fast microscopy[37] and the MODS liquid culture technique.[26,27]

Statistical analysis. Concentrations of bacilli and CFU per ml of sputum were exponentially distributed and transformed to their base-10 log value. Tests were 2-tailed and were performed with a 95% CI. Time-to-event Cox Proportional Hazards regression analysis was used to analyse the hazard ratio (HR) of TB diagnosis in contacts with estimates adjusted for household clustering, censused at the time of last follow-up. For this time-to-event regression analysis, secondary TB was a rare outcome and independent variables did not have Gaussian distributions, so all data were analysed as dichotomous variables above or below the median value for contacts, labelled 'high' or 'low' respectively. All univariable analyses, including univariable time-to-event regression analyses were performed with non-imputed data. To prevent missing data in more than one variable from reducing the statistical power of multivariable analysis, the primary multivariable analyses were performed with imputed data. Data was imputed by extrapolating results from corresponding laboratory results, or by using the median response from the participants' questionnaire data. Two multivariable sensitivity analyses were also done: 1. excluding imputed data; and 2. censoring follow-up after 5 years.

Results

Laboratory characterisation. In pre-treatment samples, quantitative culture CFU results were associated with the results of both FDA microscopy (coefficient=0.66 R^2=0.46 p<0.001) and auramine microscopy (coefficient=1.1 R^2=0.47 p<0.001, Table 2.1). TTP in days was significantly associated with quantitative culture results and with FDA microscopy results (both p≤0.01, Table 2.1).

FDA microscopy results in pre-treatment sputum. The median concentration of FDA microscopy-positive bacilli/ml was 119,140 (interquartile range (IQR)=46,781-386,158, Figure 2.2). This FDA microscopy-positive bacilli concentration was median 5.1% (IQR=2.4%-11%) of the concentration of auramine microscopy-positive bacilli/ml (2,069,160 IQR=1,358,430-3,734,268 Figure 2.2). The concentration of CFU/ml was median 40,000 (IQR=9,381-80,000, Figure 2.2). This CFU concentration was median 1.8% (IQR=0.42%-2.2%) of the concentration/ml of auramine microscopy-positive bacilli. The between-patient variability of concentrations of FDA-stained bacilli and CFU were both more than double the between-patient variability of concentrations of auramine stained bacilli (IQR of the log-concentrations 0.92 and 0.93 versus 0.44, respectively, Figure 2.2).

Figure 2. 2 Distribution of patients' pre-treatment sputum results
Box plots showing the distribution of concentrations per ml in sputum of: bacilli positive with auramine microscopy; bacilli positive with fluorescein diacetate (FDA) microscopy; and colony forming units (CFU) in quantitative culture.

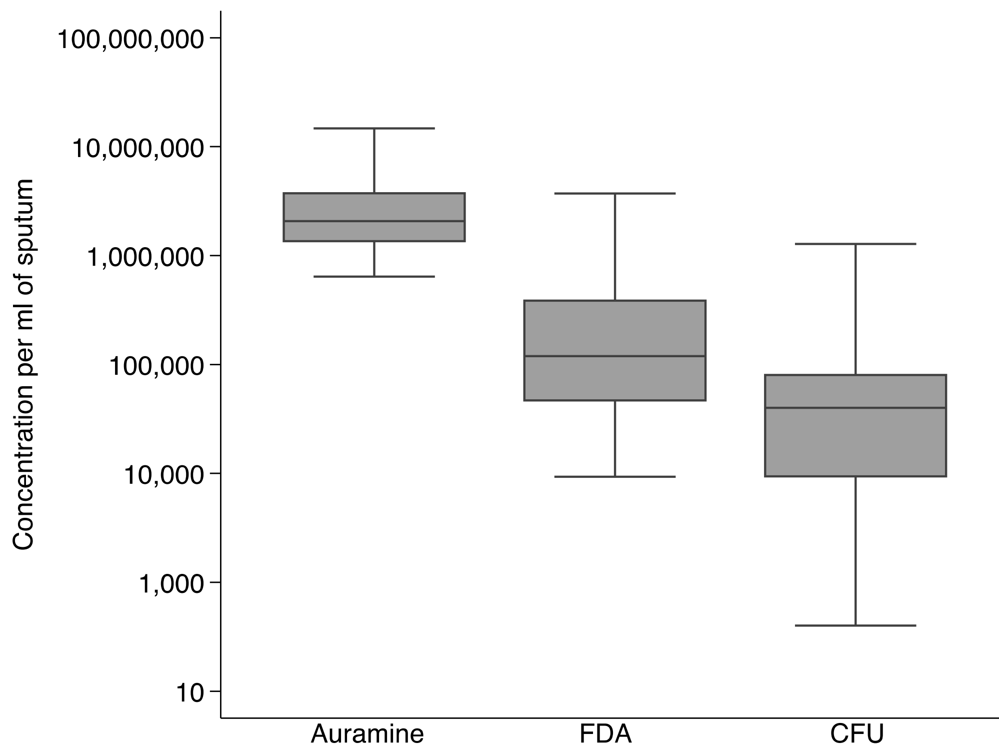


Table 2. 1 Relationship between patient characteristics and laboratory results

Linear regressions to predict quantitative results of auramine microscopy, fluorescein diacetate (FDA) microscopy and quantitative culture.

				Auramine microscopy		FDA microscopy		Quantitative culture	
Variable	**Units**	**Value**	**n**	**Coefficient**	**p value**	**Coefficient**	**p value**	**Coefficient**	**p value**
Patient laboratory characteristics									
Auramine microscopy	Median bacilli/ml (IQR)	2,069,160 (1,358430-3,734268)	35	NA	NA	**1.3**	**<0.001**	**1.1**	**<0.001**
FDA microscopy	Median bacilli/ml (IQR)	119,140 (46,781-386,158)	35	**0.49**	**<0.001**	NA	NA	**0.66**	**<0.001**
Quantitative culture	Median CFU/ml (IQR)	40,000 (9,381-80,000)	33	**0.43**	**<0.001**	**0.69**	**<0.001**	NA	NA
Time to positive culture	Median days (IQR)	7 (6-10)	34	-0.059	0.06	**-0.13**	**0.01**	**-0.14**	**0.005**
Multi-drug resistant TB	n (%)	4 (12%)	35	-0.015	1	-0.53	0.2	-0.47	0.3
Patient symptoms									
Productive cough	n with symptom (%)	29 (90%)	32	0.13	0.7	0.65	0.2	0.54	0.4
Haemoptysis	n (%)	17 (53%)	32	0.16	0.4	0.31	0.3	0.18	0.6
Cough duration before care seeking	Median days (IQR)	30 (20-60)	23	-0.0058	1	0.04	0.9	-0.39	0.4
Fever	n with symptom (%)	22 (69%)	32	**0.41**	**0.04**	0.54	0.1	0.47	0.2
Night sweats	n with symptom (%)	23 (72%)	32	**0.65**	**0.001**	**0.88**	**0.008**	**0.89**	**0.02**
Constitutional symptoms: fevers or night sweats	n with symptom (%)	24 (75%)	32	**0.67**	**0.001**	**0.95**	**0.006**	**1.03**	**0.007**
Patient demographics									
Age	Median years (IQR)	26 (22-35)	35	0.0049	0.6	0.0014	0.9	0.0066	0.6
Male sex	n (%)	20 (57%)	35	0.081	0.7	0.15	0.6	0.46	0.1
Body mass index	Mean kg/m^2 (SD)	20.8 (2.7)	32	-0.019	0.6	0.017	0.8	-0.046	0.4

Note: Predictors were calculated by univariable linear regression analysis. Values in bold font have associations with $p \leq 0.05$. Additionally, 2 patients reported previous TB diagnosis, 2 patients did not have a BCG scar, and these variables were not associated with the results of auramine microscopy, FDA microscopy or quantitative culture (all $p > 0.2$). Only 1 patient reported having a coexistent chronic respiratory illness, 1 patient reported substance abuse and no patients had a history of HIV, diabetes, regular smoking, or heavy alcohol use. All microscopy and quantitative culture data were log transformed prior to regression analysis. Thus, the coefficients indicate the differences in log concentrations. TB, tuberculosis; SD, standard deviation; n, number; IQR, interquartile range; log, base-10 logarithm; NA, not applicable. The column labelled 'n' indicates the number of patients with available data.

Patient characteristics. The patient characteristics and their associations with the results of sputum laboratory testing are shown in Table 2.1. Constitutional symptoms (night sweats and/or fevers) were associated with higher quantitative results of: auramine microscopy; FDA microscopy; and culture CFU (all $p \leq 0.04$, Table 2.1). No patients had evidence of extra-pulmonary or disseminated TB.

TB infection in contacts. At the time of patient recruitment, 209 of the patients' contacts were also recruited (Figure 2.1). These contacts were 50% (105/209) male, with a median age of 20 years (IQR=10-32), and mean BMI of 25 kg/m^2 (SD 5.2) in adults. None of the contacts reported a history of human immunodeficiency virus (HIV) or diabetes. Only 40% (84/209) of the contacts underwent tuberculin skin-testing. This was because 38% (80/209) of the contacts were ineligible because they were aged less than 15 years old and 35% (45/129) of the remaining contacts declined. 70% (59/84) of tuberculin skin-test results were positive. Positivity was not associated with any laboratory or clinical characteristics (data not shown), but confidence intervals were wide because so few participants had a tuberculin skin-test.

TB disease in contacts; univariable analysis. TB disease was diagnosed in 6.2% (13/209) contacts during follow-up that lasted median 6.3 years (IQR=6.2–6.5). This was significantly higher than the background community rate (HR=4.0 95% CI=1.9–8.4, $p<0.001$, Figure 2.3a). All the patient variables in Table 2.1, and the contact characteristics were tested for associations with secondary TB disease in contacts using univariable Cox regression analysis. FDA microscopy-positive concentrations of bacilli/ml were analysed around the median value as high (median=386,158, IQR=163,343-600,796) or low (median=21,741, IQR=9,312-52,862). All variables with biologically plausible or statistical evidence ($p<0.1$) of associations with secondary TB in contacts are shown in Table 2.2. Contacts of patients with low FDA microscopy results were 4-times more likely to have TB disease (unadjusted HR=3.8, 95% CI=1.1-13, $p=0.03$, Table 2.3 and Figure 2.3b). Specifically, 2.7% (95% CI=0.57%-7.8%, 3/110) of contacts exposed to patients with high FDA microscopy results developed TB disease during follow-up, which was not significantly different from the background community rate ($p=0.4$, Figure 2.3a). In contrast, 10% (95% CI=5.0%-18%, 10/99) of contacts exposed to patients with the low FDA microscopy results developed TB disease during follow-up (Figure 2.3b).

TB disease was more likely in contacts of patients who reported having a productive cough and in households with poorer than median socioeconomic score (both $p \leq 0.03$, Table 2.2). Preventative therapy was taken by 9.1% (19/209) of contacts, all aged <15 years, none of whom developed TB disease. Constitutional symptoms (fever and/or night sweats) in

patients was associated with FDA microscopy results, and also known to impact diagnostic delay which may influence TB transmission,[64,65] but including either or both variables in all multivariable regression analyses did not affect the significant association between lower than median FDA microscopy results and increased secondary TB. TB disease in contacts was not associated with high versus low auramine microscopy results (p=0.7, Table 2.2), but all patient samples were strongly auramine microscopy positive. Although in univariate analysis, quantitative culture CFU did not predict TB disease (p=0.2, Table 2.2), the association of FDA microscopy and secondary TB disease in contacts was modified by the patient's corresponding quantitative culture CFU results, (p<0.001, Figure 2.4). Specifically, contacts of patients with high CFU concentrations and low FDA microscopy results had the highest risk of developing TB disease (unadjusted HR=7.4, 95% CI=2.7-20, p<0.0001, Figure 2.4).

TB disease in contacts; adjusted multivariable regression analyses. For the primary multivariable Cox regression analysis shown in Table 2.3, five missing data points were imputed. Two patients had missing CFU data because of fungal overgrowth, and high/low CFU categorization was imputed from TTP data and auramine microscopy results. Three patients did not report whether they had productive cough, so this was imputed as the median response.

The association between low FDA microscopy results predicting increased risk of TB disease in patients' contacts maintained statistical significance after adjusting for quantitative culture CFU results (HR=8.1, 95% CI=3.1-21, p<0.001), disease severity indicated by productive cough (HR=4.5, 95% CI=1.4-15, p=0.01), socioeconomic score (HR=4.2, 95% CI=1.4-13, p=0.01), preventative therapy (HR=3.8, 95% CI=1.0-14, p=0.04), and also after adjusting for all these variables (HR=5.5, 95% CI=1.8-17, p=0.03), see Table 2.3.

Sensitivity analyses that excluded all imputed data or had follow-up censused after 5 years (that excluded the last episode of secondary TB in contacts that occurred 5.7 years after index case diagnosis and could have been unrelated to that exposure) had almost identical results to the analyses above (Table 2.3).

Post hoc exploratory analysis of TTP data considered patient cultures as faster or slower than the median value of 7 days. Faster TTP predicted a significantly increased risk of secondary TB in contacts (HR=4.5, 95% CI=1.1-19, p=0.04). Lower than median FDA microscopy-positive concentrations predicted an increased risk of secondary TB disease in contacts independently of TTP (HR=3.9, 95% CI=1.1-13, p=0.03).

Table 2. 2 Characteristics predicting secondary TB disease in contacts

	Patient n	Contact n	Hazard Ratio	95% CI	p value
Patient sputum laboratory characteristics					
Auramine microscopy above median (versus below median)	35	209	0.75	0.21 – 2.6	0.6
FDA microscopy above median (versus below median)	35	209	**0.26**	**0.08 - 0.90**	**0.03**
Quantitative culture above median (versus below median)	33	201	2.4	0.74 – 7.5	0.2
MDR (versus non-MDR)	35	209	2.6	0.39 - 17	0.3
Patient disease characteristics					
Productive cough in patient (versus no productive cough)	32	203	**high**	**high**	**<0.001**
Haemoptysis in study patient (versus no haemoptysis)	32	203	3.9	0.85 – 18	0.08
Fever in study patient (versus no fever)	32	203	4.7	0.81 – 27	0.09
Cough duration before care seeking above median (versus below median)	23	133	0.44	0.10 – 2.0	0.3
Household characteristics					
Socioeconomic score poorer than median (versus less poor)	35	209	**3.8**	**1.1 – 13**	**0.03**
Number of contacts/household above median 5 (versus less than median)	35	209	1.0	0.27 – 4.0	0.1
Contact characteristics					
Household contact completed preventative therapy (versus not completed)	-	209	**low**	**low**	**<0.001**
Male sex (versus female)	-	209	0.61	0.21 – 1.8	0.4
Age younger than 15 years (versus above 15 years)	-	209	0.29	0.10 – 1.3	0.1
Age above median 20 years (versus below median)	-	209	1.8	0.71 – 4.5	0.2
BMI above median 25 kg/m2 (versus below median) for adults	-	96	1.3	0.34 – 5.1	0.7

Note: All the variables in Table 2.1 and the contact characteristics were tested for associations predicting secondary TB disease in contacts by univariable Cox regression analysis. Imputation of missing data was only performed for multivariable analyses, so the analyses in this table only used original, non-imputed data (see Methods). See Table 3.3 for multivariable regression analysis. Values in bold font have associations with $p \leq 0.05$. FDA, fluorescein diacetate; MDR, multi-drug resistant; TB, tuberculosis; CI, confidence interval. The columns labelled 'n' indicate the number of patients with available data. The contact characteristics are also reported here stratified for contacts of patients with high versus low FDA microscopy results: completed preventative therapy 11/110 (10%) versus 8/99 (8.9%); male sex 54/110 (49%) versus 51/99 (52%); age <15 years 42/110 (38%) versus 38/99 (38%); age >median 20 years 54/110 (49%) versus 44/99 (44%); BMI >median 25 kg/m^2 25/50 (50%) versus 23/46 (50%).

Figure 2. 3 Time-to-event analyses

Kaplan-Meier curves for patients' household contacts being diagnosed with TB disease. In Graph A, this is compared with the background community rate of 233 new cases of TB per 100,000 people per year. In Graph B, this is analysed according to whether the patients' pre-treatment sputum fluorescein diacetate (FDA) microscopy result was above (high) or below (low) the median. The p-values indicate the results of Cox regression analysis (see text)

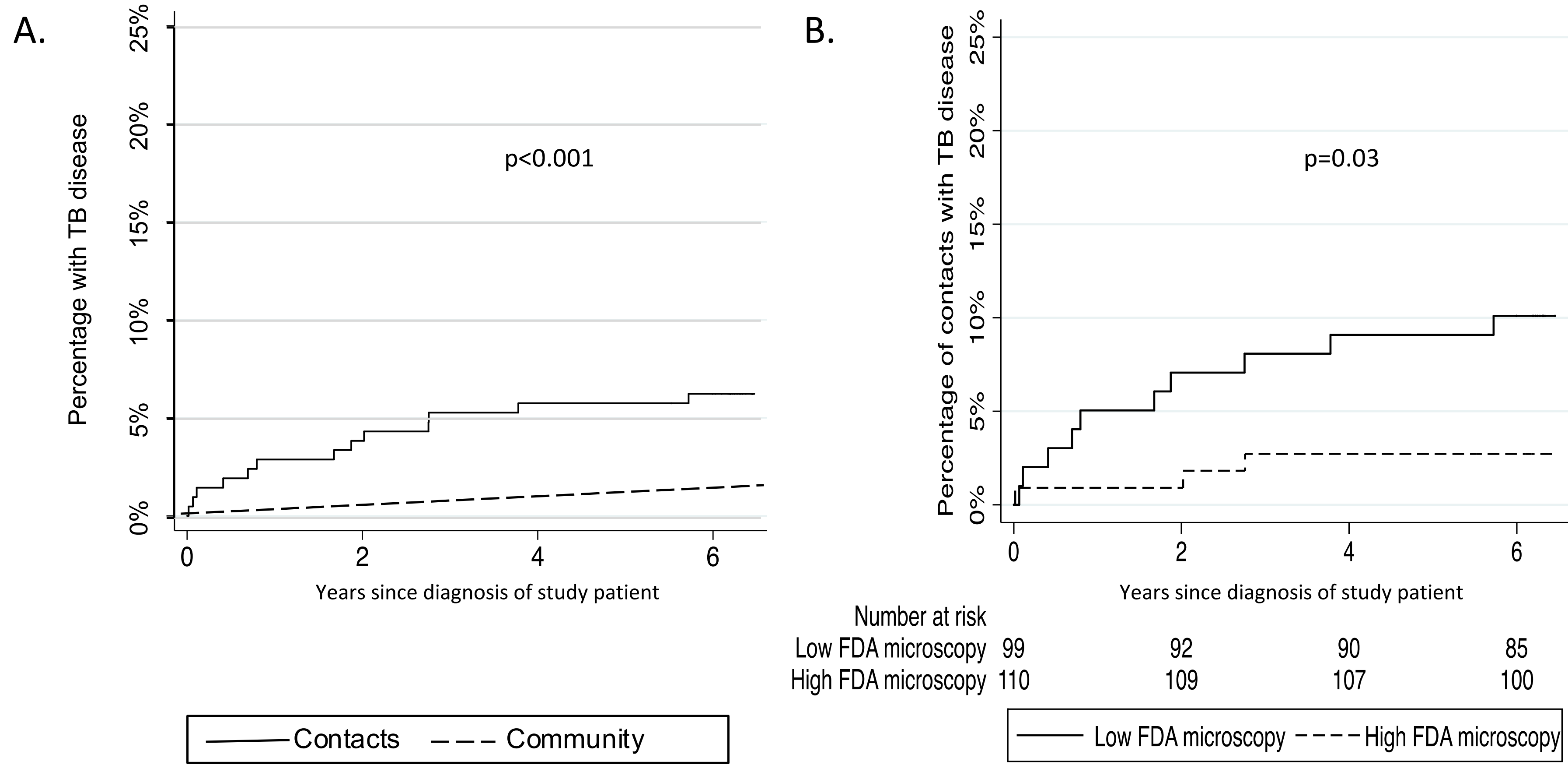


Figure 2. 4 Time-to-event analysis of the interaction between FDA microscopy and culture results
Kaplan-Meier time to event curves for patients' household contacts being diagnosed with TB disease analysed according to whether the study patients' pre-treatment sputum fluorescein diacetate (FDA) microscopy result was above (high) or below (low) the median concentration. In Graph A, this is shown for patients with quantitative culture colony forming unit (CFU) results below the median concentration. In Graph B, this is shown for patients with quantitative culture CFU results above the median concentration. The p-values indicate the results of Cox regression analysis (see text).

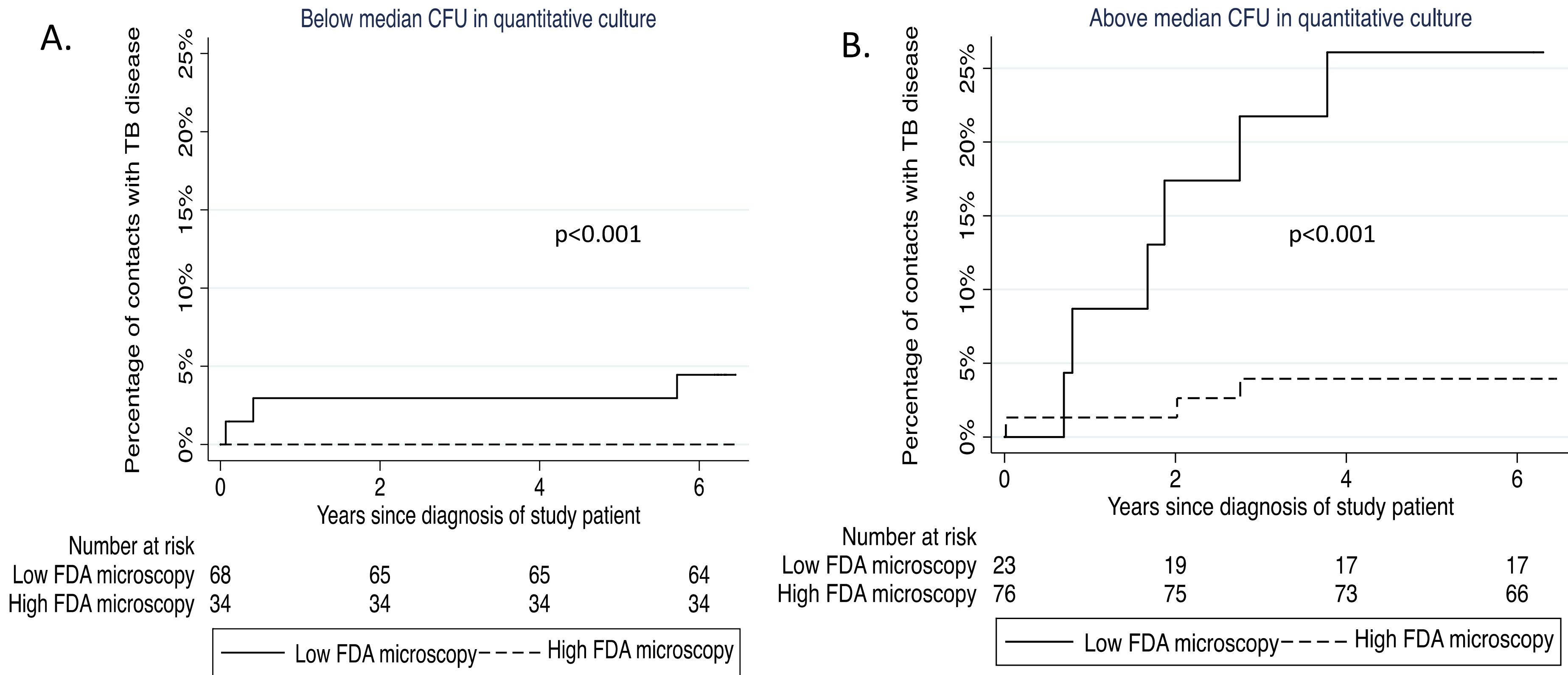


Table 2. 3 Patient FDA microscopy predicting secondary TB disease in contacts

	Primary analysis including imputed data (n=35 patients) (n=209 contacts)			Sensitivity analysis 1 excluding imputed data (n=35 to 32 patients) (n=209 to 201 contacts)			Sensitivity analysis 2 censused at 5 years (n=35 patients) (n=209 contacts)		
	HR	**95% CI**	**p value**	**HR**	**95% CI**	**p value**	**HR**	**95% CI**	**p value**
FDA low (versus high), unadjusted analysis	3.8	1.1 - 13	0.03	3.8	1.1 - 13	0.03	3.4	0.96 - 12	0.059
FDA low (versus high) adjusted for:									
Higher than median quantitative culture	8.1	3.1 - 21	<0.001	8.2	3.1 - 21	<0.001	7.8	3.0 - 20	<0.001
Productive cough in patient	4.5	1.4 - 15	0.01	3.9	1.1 - 14	0.03	4.1	1.2 - 14	0.03
Poorer than median socioeconomic score	4.2	1.4 - 13	0.01	4.2	1.4 - 13	0.01	3.9	1.2 - 12	0.02
Preventative therapy completed	3.8	1.0 - 14	0.04	3.8	1.0 - 14	0.04	3.3	0.9 - 13	0.07
Quantitative culture, productive cough, Socioeconomic score and preventative therapy	5.5	1.8 - 17	0.003	5.3	1.8 - 16	0.003	4.6	1.6 - 13	0.005
Quantitative culture, productive cough and preventative therapy	7.1	2.5 - 20	<0.001	6.7	2.4 - 18	<0.001	6.8	2.4 - 19	<0.001

Note: The results were calculated by multivariable Cox regression analysis adjusting the association of FDA microscopy with the variables that in univariable analysis significantly predicted secondary TB disease in contacts (shown in Table 2.2). High versus low auramine microscopy results were not associated in univariable analysis with secondary TB (p=0.6) and if despite this auramine microscopy results were included in the adjusted analyses then the pattern of significance was unchanged. The presence versus absence of MDR TB was not associated in univariable analysis with secondary TB (p=0.3) and if despite this MDR TB was included in the adjusted analyses then the pattern of significance was unchanged. The primary multivariable analysis included five imputed data for quantitative culture and productive cough (see text). Sensitivity analysis 1 excluded the imputed data. Sensitivity analysis 2 censused all follow-up at five years, causing one late case of secondary TB to be excluded. FDA, fluorescein diacetate; HR, hazard ratio; 95% CI, 95% confidence intervals.

Discussion

In strongly microscopy-positive sputum samples collected before TB treatment, the proportions of FDA-stained and culturable *M. tuberculosis* bacteria varied considerably between patients. These findings are clinically significant because patients with low FDA microscopy results were most infectious, as demonstrated by the finding that their household contacts were considerably more likely to develop TB disease, including when adjusted for confounding variables and in sensitivity analyses.

Sputum from patients with pulmonary TB is known to contain heterogeneous subpopulations of *M. tuberculosis* bacteria.[30,47,66] Culture techniques used for diagnosing TB do not generally identify individual subpopulations because results are determined by the overall quantity of bacteria as well as the fastest growing subpopulation. A subpopulation of particular interest are persister cells, a term used in the context of chemotherapy to describe *M. tuberculosis* that persists in sputum in a slowly-metabolizing, non-replicating state.[46] In pre-treatment sputum, most *M. tuberculosis* bacteria are believed to be in a persister-like state, as they cannot be cultured using conventional techniques and contain lipid bodies, both features of this phenotypic variant.[30,47,67] Furthermore *M. tuberculosis* in persister-like states has been demonstrated not to fluoresce with FDA, unless resuscitation-promoting factors are added.[68] Non-fluorescence with FDA could be explained by loss of cell membrane integrity, changes in the cell wall effecting FDA penetration, or low esterase activity when in a slow metabolizing state.[11,46]

We found that household contacts of patients who had low concentrations of FDA microscopy-positive *M. tuberculosis* in their sputum were approximately 4-times more likely to develop TB disease. *M. tuberculosis* persister-like states develop in response to stressful environments such as nutrient deficiency or hypoxia,[69] and persister-like phenotypes of *M. tuberculosis* are thought to be better adapted to surviving harsh environments such as during airborne transmission.[47] This increased infectiousness of *M. tuberculosis* in persister-like states has been demonstrated by growing *M. tuberculosis* under hypoxic or nutrient stress. Compared with *M. tuberculosis* grown under conventional aerobic conditions, *in vivo* this caused 10-times greater airborne infectiousness to guinea pigs and an altered immune response that increased pathogenicity to mice; and *in vitro* this increased invasion of epithelial cells and alveolar macrophages.[51–54] We hypothesize that in our study, which only included strongly sputum microscopy-positive TB, patients with low concentrations of FDA stained *M. tuberculosis* in their sputum had high concentrations of persister-like mycobacteria that were better adapted to airborne transmission, so were more infectious. This hypothesis warrants specific research, in

addition to confirmation of the association with infectiousness in larger studies and with other patient groups, as will be described in Research Chapter 6.

We found that the risk of secondary TB disease was higher for contacts of patients with high quantitative culture results, whether indicated by high CFU (in adjusted analysis) or culture TTP (in univariable analysis), both indicators of mycobacterial load. Generally, smear positivity is also associated with greater infectiousness, but a meta-analysis of risk factors for TB transmission required thousands of participants to demonstrate that patients with concentrations of *M. tuberculosis* in their sputum sufficient for positive acid-fast microscopy had a 1.4-times (95% CI=1.2-1.6) increased risk of causing TB disease in their contacts compared with acid-fast microscopy-negative patients.[70] This finding does not contradict the results of our study, which included only strongly microscopy-positive patients. In contrast to the 1.4-times greater TB infectiousness in acid-fast microscopy-positive versus microscopy-negative patients in this meta-analysis, our results suggest that FDA microscopy results had a considerably larger effect than the results of conventional microscopy, CFU, TTP or other potentially contributory variables. In fact, our study demonstrates that contacts of patients who were conventional microscopy-positive but had high concentrations of fluorescing bacteria in FDA microscopy had no higher risk then the background rate in their community.

Contacts of patients with pulmonary TB disease are at high risk of developing TB disease and are often provided with preventative therapy. This policy is consistent with the apparent protective effect of preventative therapy in our study. However in most high TB-prevalence settings, resources are insufficient for comprehensive contact tracing within and outside of the home.[71] Provision of preventative therapy to all household contacts may be appropriate in some settings. However, a low-cost tool, such as FDA microscopy, may identify the most infectious individuals and allow better prioritised contact tracing for TB disease screening and prevention than the current use of conventional sputum microscopy results.

In addition to studying infectiousness to cause TB disease in contacts, we also aimed to assess infectiousness to cause asymptomatic latent TB infection in contacts. However, most of the contacts in our study were too young to be eligible for or declined tuberculin skin-testing. This low uptake of tuberculin skin-testing reduced the statistical power so much that we were unable to meaningfully test for associations with asymptomatic latent TB infection. Furthermore, concurrently we discovered that the majority of healthy adults in these communities were already tuberculin skin-test positive, even without living with a known TB patient.[61,72] Thus, living with a patient newly diagnosed with pulmonary TB

increased the risk of tuberculin skin-test positivity by less than 50%.[61,72] Therefore, to adequately test the hypothesis that FDA-negative *M. tuberculosis* in sputum also predicts infectiousness for causing asymptomatic latent TB infection would require much larger future studies, ideally using the more specific interferon-gamma release assays in a setting with a low background rate of tuberculosis infection.[72–74]

To reduce the risk of contamination in culture and false-positive FDA microscopy results due to bacteria other than *M tuberculosis* within the sample, we decontaminated all sputum samples, which is known to kill some *M. tuberculosis*.[44] Whether decontamination selectively kills subpopulations with specific FDA-staining, culturability or infectiousness is unknown and warrants future research. However, this would be expected to affect all samples because the same protocol was followed for all the specimens processed and should also similarly affect the quantitative culture results. Therefore, decontamination could not explain the contrasting results demonstrated between FDA microscopy and quantitative culture predicting infectiousness. Another limitation is that the *M. tuberculosis* TB strains were not available to use molecular fingerprinting to confirm whether any of the secondary TB cases in contacts were caused by TB patients other than the index case recruited to our study.[75,76] However, our conclusions were supported by a sensitivity analysis excluding the last episode of secondary TB in case that late case was unrelated to the index patient.

In conclusion, the inexpensive, rapid, low-technology technique of FDA microscopy revealed that only a small and variable sub-population of *M. tuberculosis* in sputum was FDA microscopy-positive. This study provides evidence that amongst strongly sputum microscopy-positive patients, those with low concentrations of FDA microscopy-positive *M. tuberculosis* were most infectious to cause TB disease in their contacts. This supports the hypothesis that a subpopulation of *M. tuberculosis* that do not stain with FDA, such as slowly metabolising persister-like state, are most infectious to transmit TB disease. Thus, FDA microscopy may identify a high-risk group of TB patients whose contacts particularly need interventions to prevent and screen for secondary TB disease

RESEARCH CHAPTER 3

Foundations – TB viability microscopy assessing treatment response

Background

Quantitative culture has demonstrated that concentrations of culturable *M. tuberculosis* in patients' sputum decline rapidly during the first days of adequate TB treatment.[77] Furthermore, quantifying early treatment response may allow the early identification of patients whose treatment is later discovered to have been inadequate,[77–79] potentially allowing drug-susceptibility testing to be provided sooner to patients most likely to benefit. This is important because although providing universal drug-susceptibility testing is a policy priority, globally most patients with TB are treated without testing to assess whether their *M. tuberculosis* is resistant or susceptible to the drugs they are administered.[80] Assessing early treatment response may help detect not only drug-resistance, but also other causes of incipient treatment failure including malabsorption and poor adherence, which can be difficult to detect. Currently the detection of failing TB treatment is often delayed, risking morbidity, mortality, and contagion.[81]

FDA microscopy of sputum has been used to predict TB culture positivity[15,39] and FDA with flow cytometry to determine drug-susceptibility.[82–84] We hypothesised that sputum FDA microscopy would be associated with the concentration of culturable *M. tuberculosis*, indicating whether patients are responding to TB treatment. To test this hypothesis, in this foundation study we compared FDA microscopy results with culture and conventional acid-fast microscopy during early treatment.

Methods

Ethics. All participants gave informed written consent. The internationally accredited ethics committee of the Universidad Peruana Cayetano Heredia approved the study. All clinically relevant results were provided to participants and their physicians in collaboration with the Peruvian Ministry of Health. This research had no role in patient care, which was free of direct charges for all patients with suspected TB and was not delayed or modified by participation in this research. Patients received empiric first-line TB therapy with clinic-based direct observation of every dose from the TB program, according to Peruvian policy.[40,56]

Inclusion criteria were unselected adults diagnosed with sputum microscopy-positive TB disease in shantytowns in Callao, Peru.[60] All sputa were collected over 12 hours in order to reduce random variation between samples and patients had to be available to collect their first baseline sputum prior to commencing treatment.

Specimens. Single sputum samples were collected pre-treatment ('day 0'), and on treatment days 3, 6, and 9. Drug-susceptibility results became available and could influence treatment only after these samples were collected.

Patients were recruited from 2006 to 2007 as part of the study described in Research Chapter 2. Radiography was not a routine part of TB management and was not studied.[56] We made follow-up visits every 3 years until 2013 to ask whether TB had recurred, screen ex-patients for respiratory symptoms and offer TB testing with sputum culture to symptomatic ex-patients. We categorised treatment outcome as 'good' (cured or completed treatment without treatment failure as assessed by the TB program, and with no recurrence) or 'adverse' (treatment failure, suspension, death during treatment or recurrence).[85,86]

Specimen collection and processing has been described in Research Chapter 2.

Analysis. Recruitment took place over 12 months, so recruitment and follow-up samples from different patients were analysed concurrently. Staff were blinded to all other clinical and laboratory data. According to Peruvian policy for conventional acid-fast microscopy at that time, all microscopy was considered positive if ≥10 objects with the appearance of *M. tuberculosis* were seen.[40] The international MODS protocol determined that >1 colonies indicated a positive MODS result, because single colonies often result from cross-contamination.[87] Concentration per ml of sputum were calculated from the bacterial and

colony counts, which were then transformed to their base-10 log value (after zero values were transformed to the mid-point between zero and the detection threshold) and are reported as geometric means with 95% CI. Parametric data were summarised as means with standard deviation (SD). Non-parametric data were summarised as median with IQR. Linear regression used random effects to adjust for between-patient variations. The study sample size was resource-limited and treatment response power calculations were not done. All analyses were 2-tailed, done with STATA version 11 (StataCorp, Texas, USA).

Results

Pre-treatment sputum was available from 35 patients whose characteristics are shown in Table 3.1. All participants had culture-confirmed TB with positive auramine microscopy results. All pre-treatment sputum had FDA-positive bacteria visualised; and 94% (33/35) had FDA-positive bacilli above the threshold for positivity.

Table 3. 1 Baseline patient and laboratory data
Patient and laboratory data are shown at the time of recruitment ('day 0') and on days 3, 6 and 9 of treatment. Note: SD=standard deviation; n=number of participants with that characteristic; N=number of participants with available data; IQR=interquartile range; TEMA=tetrazolium microplate assay; denominators vary because of 11% missing samples and 10% contaminated quantitative cultures.

	Day 0	Day 3	Day 6	Day 9
Laboratory results				
Samples collected, N	35	34	30	25
FDA microscopy positive, % (n/N)	94% (33/35)	76% (26/34)	53% (16/30)	24% (6/25)
Culture positive, % (n/N)	100% (33/33)*+	94% (29/31)*	81% (21/26)*	64% (14/22)*
Auramine microscopy positive, % (n/N)	100% (35/35)	97% (33/34)	97% (29/30)	88% (22/25)
Multi-drug resistant, % (n/N)	12% (4/35)			
Isoniazid mono-resistant % (n/N)	5.60% (2/35)			
Rifampicin mono-resistant % (n/N)	0.0% (0/35)			
Ethambutol resistant in TEMA % (n/N)	37% (13/35)			
Streptomycin resistant in TEMA % (n/N)	34% (12/35)			
Ciprofloxacin resistant in TEMA % (n/N)	0.0% (0/35)			
Capreomycin resistant in TEMA % (n/N)	2.9% (1/35)			
Patient Characteristics				
Age, median years (IQR)	26 (22-35)			
Sex, % males (n/N)	57% (20/35)			
Body mass index, mean (SD)	21 (2.7)			
Poverty indicator: median food spending US$/week (IQR)	6.3 (4.0-8.7)			
Bacillus Calmette–Guérin, % with scar (n/N)	94% (30/35)			
Past tuberculosis diagnosis, % (n/N)	6.30% (2/35)			
HIV, % (n/N)	0.0% (0/35)			

* 12 quantitative culture results were unavailable because of contamination of the cultures.
[+] All pre-treatment (day 0) samples were culture-positive in conventional non-quantitative MODS.

***All sputum* samples** had interpretable microscopy results. Sputa were available for 89% (124/140) of the intended samples on days 0, 3, 6 and 9. Quantitative cultures provided interpretable results for 90% (112/124) of samples (Table 3.1). FDA microscopy bacilli counts had moderate agreement with CFU in quantitative culture (Figure 3.1a) such that 76% of results were within ±1 log, 96% within ±2 log and there was good correlation (rho=0.85, p<0.001). Similarly, 78% of the changes in FDA microscopy during each interval of 3 days treatment (baseline 0 days to 3 days, 3 to 6 days, and 6 to 9 days follow-up) were within ±1 log of the changes in quantitative culture (Figure 3.1b). TTP correlated with both CFU in quantitative culture (rho=-0.58 p<0.0001) and FDA microscopy (rho=-0.50 p<0.0001). Drug-susceptibility results were 100% concordant between the MODS and TEMA assays.

Non-MDR TB (Figure 3.2). 89% (31/35) patients had non-MDR TB, including 2 patients with isoniazid mono-resistance. Compared with pre-treatment, mean FDA microscopy results for patients with non-MDR TB after 3, 6 and 9 days of treatment reduced 11-fold, 43-fold and 250-fold, respectively (all p<0.001). By treatment day 9, FDA microscopy counts reduced >10-fold in all patients with non-MDR TB and became negative for 79% (19/24). Quantitative culture mean CFU results for patients with non-MDR TB after 3, 6 and 9 days of treatment reduced 15-fold, 46-fold and 190-fold, respectively compared with pre-treatment (all p<0.001). By treatment day 9, quantitative culture had reduced CFU on average >10-fold in 95% (19/20) patients with non-MDR TB and became negative for 38% (8/21). Auramine microscopy mean results fell on average <5-fold throughout early treatment, so by treatment day 9, auramine microscopy had reduced >10-fold in only 25% (6/24) patients with non-MDR TB and became negative in only 13% (3/24). Regression analysis (Table 3.2) demonstrated that in patients with non-MDR TB, FDA microscopy counts approximately halved daily (p<0.001), as did quantitative culture (p<0.001). Auramine microscopy changed little, reducing only by 1.2-fold daily (p=0.04).

MDR TB (Figure 3.3). In the 11% (4/35) patients with MDR TB there were no significant changes in FDA microscopy (p=0.4) or quantitative culture (p=0.6) during treatment (Table 3.2), but only 12 samples were available.

Non-MDR TB versus MDR TB (Table 3.2). Changes in FDA microscopy and quantitative culture during treatment differed significantly for patients with non-MDR TB versus MDR TB (both p<0.001). FDA microscopy changes were significantly greater for non-MDR TB than MDR TB after 3-, 6- and 9-days treatment (all p<0.001). Quantitative culture changes were significantly greater for non-MDR TB versus MDR TB after 6- and 9-days treatment (p=0.001 and p=0.04, respectively). Auramine microscopy was similar comparing patients

with non-MDR TB versus MDR TB overall and after 3-, 6- and 9-days treatment ($p=0.6$, $p=0.2$, $p=0.08$, $p=0.3$ respectively).

Proportion FDA stained and culturable. For every decontaminated sputum sample, per ml concentrations were greater for acid-fast microscopy than FDA microscopy and quantitative culture (both $p<0.001$). Proportion FDA stained was calculated by dividing concentrations/ml sputum of FDA microscopy-positive bacteria by concentrations/ ml sputum of acid-fast microscopy-positive bacteria. Similarly, proportion-culturable was calculated by dividing the quantitative culture CFU/ml sputum by concentrations of acid-fast microscopy-positive bacteria/ml sputum. For patients with non-MDR TB (Figure 3.2), proportion FDA stained and proportion-culturable fell during treatment (both $p<0.001$). Consequently, for patients with non-MDR TB by treatment day 9 there had been >10-fold fall in proportion FDA stained for 96% (23/24) patients; and proportion-culturable for 80% (16/20) patients. In contrast, for the few patients with MDR TB (Figure 3.3), there was no significant change in proportion FDA stained or proportion-culturable during treatment (both $p>0.4$).

Microscopy reproducibility (Figure 3.4). Linear regression demonstrated that variation between triplicate slides accounted for no more than 5% of the variation between microscopy results (rho=0.95). Consequently, 100% of first slides predicted geometric mean results of triplicate (i.e. first, second and third) slides within ±1 log for FDA microscopy (Figure 3.4a) and 99% for acid-fast microscopy (Figure 3.4b). The results of all the above analyses were little changed by using only first slide results (Table 3.4 and Figure 3.5 – 3.7).

Clinical. All patients had outcome assessed at the time of treatment completion. All 31 patients with non-MDR TB were cured and all but one of them were followed-up for ≥3 years with no recurrence of disease. All 4 patients with MDR TB had adverse outcomes: 1 patient had treatment suspended and the 3 other patients died. These deaths occurred 22 days, 10 and 11 months after initially commencing therapy.

Figure 3. 1 Difference between quantitative FDA microscopy and culture results (data from triplicate microscopy slides)
Histograms showing for each sputum sample the difference between (A) quantitative fluorescein diacetate (FDA) microscopy results minus quantitative culture colony forming units (CFU) results; and (B) the change in quantitative FDA microscopy results minus the change in quantitative culture CFU results during each interval of 3 days of treatment (baseline 0 days to 3 days, 3 to 6 days, and 6 to 9 days follow-up). All results are shown on a log scale and log=base-10 logarithm. 76% to 78% of results agreed within ±1 log and, 96% to 97% agreed within ±2 log. All quantitative FDA microscopy results are the geometric mean of data from triplicate microscopy slides prepared from each sputum sample.

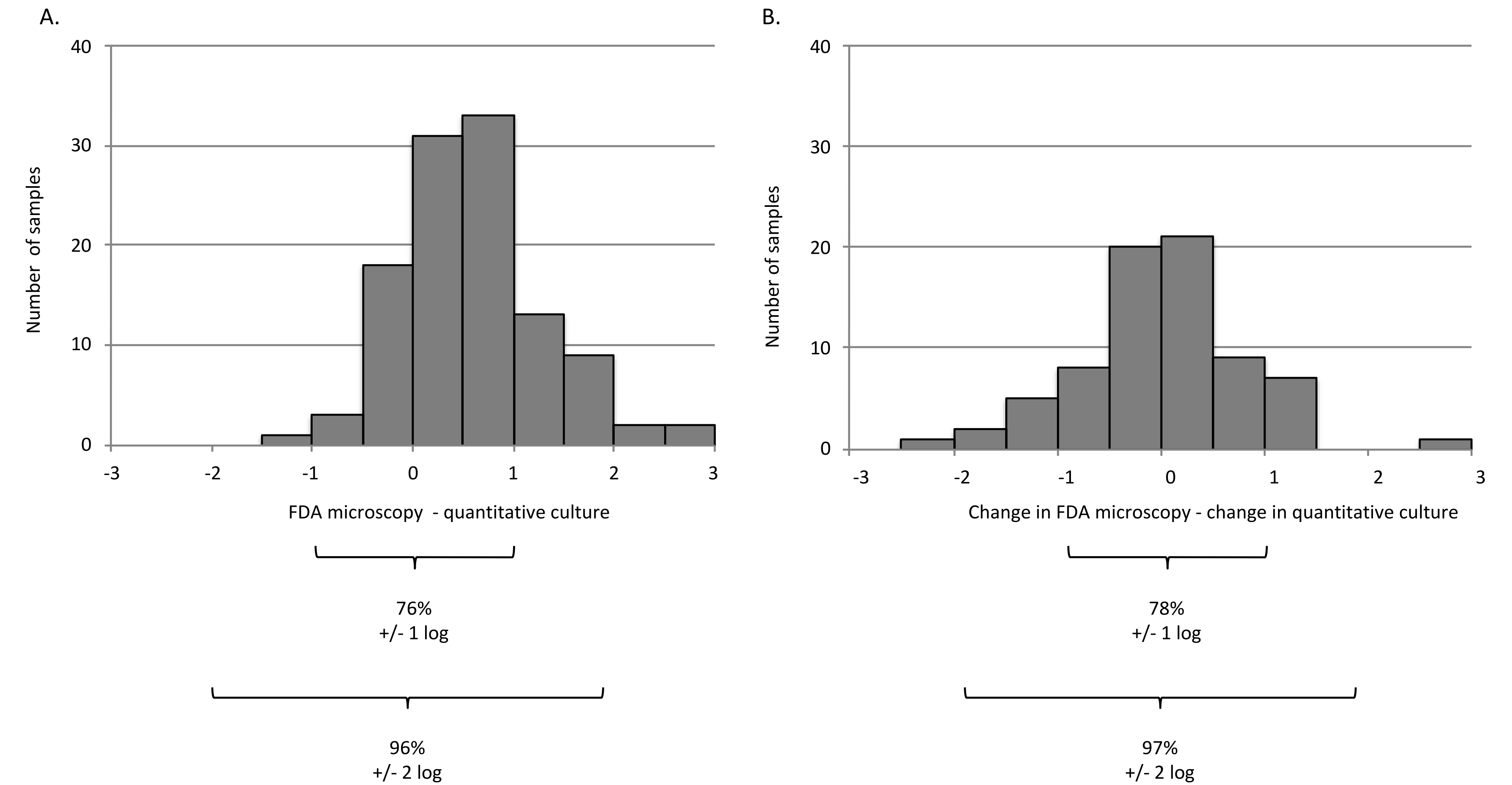
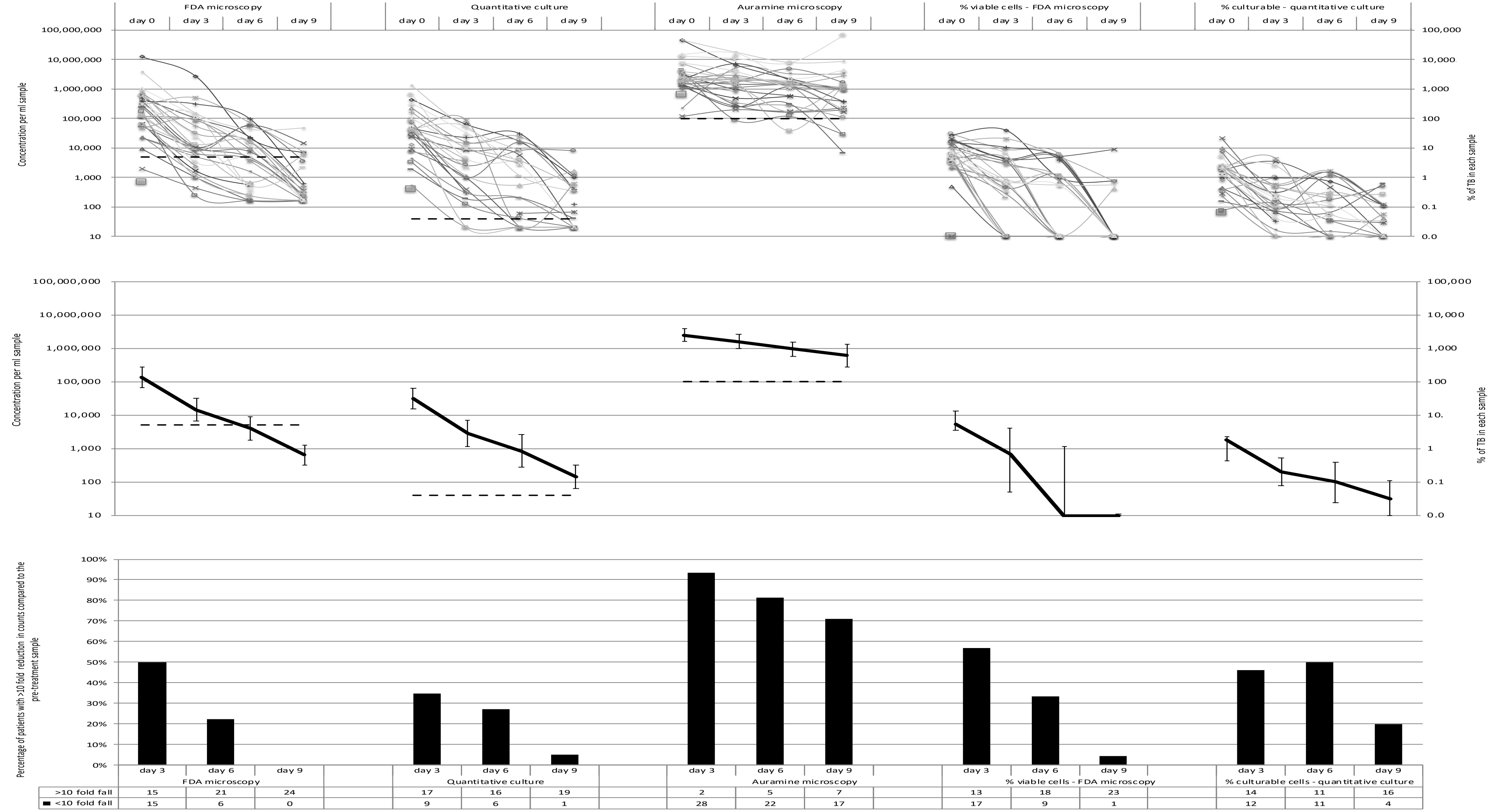


Figure 3. 2 Treatment response for 31 patients with non-MDR TB (data from triplicate microscopy slides)
Showing results from days 0, 3, 6 and 9 of TB treatment for: (a) each patient; (b) all patients; and (c) % of patients with <10-fold reduction. The horizontal axis shows days of treatment. Dashed lines indicate cut-offs for positivity. Proportion fluorescein diacetate (FDA) stained was calculated by dividing concentrations of FDA positive bacteria by concentrations of auramine microscopy-positive bacteria. Similarly, proportion-culturable was calculated by dividing quantitative culture colony forming units results by concentrations of auramine microscopy-positive bacteria. Microscopy results are geometric means of triplicate identical slides from each sample.



	FDA microscopy day 3	FDA microscopy day 6	FDA microscopy day 9	Quantitative culture day 3	Quantitative culture day 6	Quantitative culture day 9	Auramine microscopy day 3	Auramine microscopy day 6	Auramine microscopy day 9	% viable cells - FDA microscopy day 3	% viable cells - FDA microscopy day 6	% viable cells - FDA microscopy day 9	% culturable cells - quantitative culture day 3	% culturable cells - quantitative culture day 6	% culturable cells - quantitative culture day 9
>10 fold fall	15	21	24	17	16	19	2	5	7	13	18	23	14	11	16
■ <10 fold fall	15	6	0	9	6	1	28	22	17	17	9	1	12	11	4

Table 3. 2 Regression analysis of laboratory results during early treatment (data from triplicate microscopy slides)
This demonstrates the effects of multi-drug resistant tuberculosis (MDR TB) and days of treatment on laboratory results. Note 95% CI=95% confidence intervals; log=base-10 logarithm; CFU=colony forming units

	FDA microscopy			**Quantitative culture**			**Auramine microscopy**		
	Coefficient (difference in log bacilli concentration)	**95% CI**	**p value**	**Coefficient** (difference in log CFU concentration)	**95% CI**	**p value**	**Coefficient** (difference in log bacilli concentration)	**95% CI**	**p value**
(i) Day 0 (pre-treatment): difference in concentration for MDR TB versus non-MDR TB	-0.48	-1.2 to 0.27	0.2	-0.39	-1.3 to 0.55	0.4	-0.073	-0.62 to 0.48	0.8
(ii) Non-MDR TB: daily change in concentration during treatment	**-0.27**	**-0.29 to -0.25**	**<0.001**	**-0.27**	**-0.29 to -0.25**	**<0.001**	**-0.070**	**-0.085 to -0.056**	**<0.001**
(iii) MDR TB: daily change in concentration during treatment	-0.028	-0.10 to 0.044	0.4	-0.019	-0.095 to 0.057	0.6	**-0.056**	**-0.11 to 0.0033**	**0.04**

Note: there was very strong evidence for an interaction between MDR TB and daily change in concentration for both FDA microscopy and quantitative culture (both $p<0.001$), but no evidence of an interaction for auramine microscopy ($p=0.6$). Values in bold font have associations with $p\leq0.05$.

Figure 3. 3 Treatment response for 4 patients with MDR TB (data from triplicate microscopy slides)
Showing results from days 0, 3, 6 and 9 of TB treatment for: (a) each patient; and (b) all patients. The horizontal axis shows days of treatment. Dashed lines indicate cut-offs for positivity. Proportion fluorescein diacetate (FDA) stained was calculated by dividing concentrations of FDA microscopy-positive bacteria by concentrations of auramine microscopy-positive bacteria. Similarly, proportion-culturable was calculated by dividing concentrations of colony-forming units in quantitative culture by concentrations of auramine microscopy-positive bacteria. Microscopy results are geometric means of triplicate identical slides from each sample.

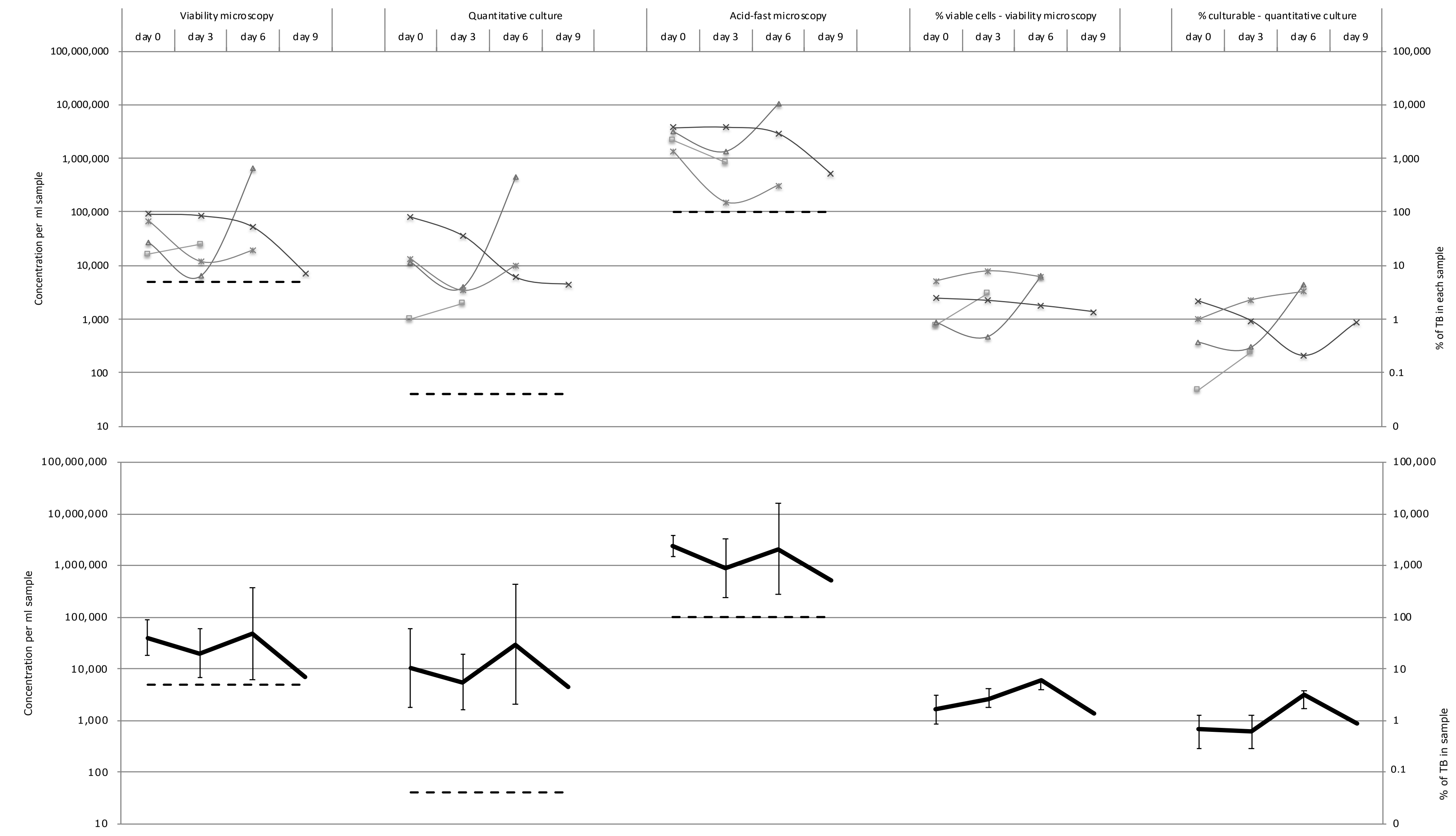


Figure 3. 4 Comparison of results for first microscopy slides versus geometric means
Histograms showing for each sputum sample the difference between (A) quantitative fluorescein diacetate (FDA) microscopy results; and (B) the auramine microscopy results for first microscopy slides versus geometric means of the results for the first, second and third slides. All results are shown on a log scale and log= base-10 logarithm. 100% of results agreed within ±1 log.

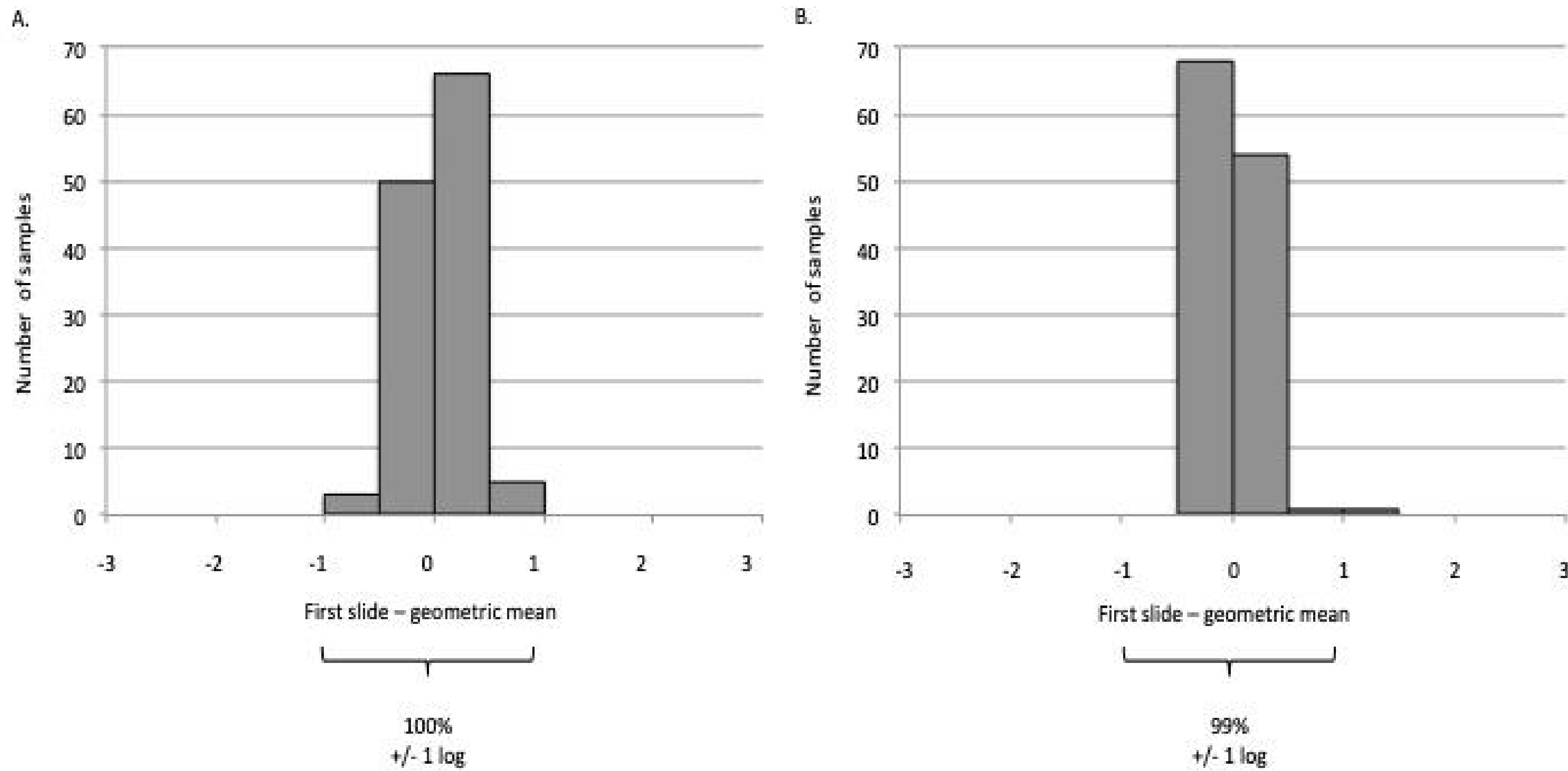


Table 3. 3 Regression analysis of laboratory results during early treatment (data from the first microscopy slide)
This demonstrates the effects of multi-drug resistant tuberculosis (MDR TB) and days of treatment on laboratory results. Note 95% CI=95% confidence interval; log=base-10 logarithm; CFU=colony forming units.

	FDA microscopy (first slide)			**Quantitative Culture**			**Auramine microscopy (first slide)**		
	Coefficient (difference in log bacilli concentration)	**95% CI**	**p value**	**Coefficient** (difference in log CFU concentration)	**95% CI**	**p value**	**Coefficient** (difference in log bacilli concentration)	**95% CI**	**p value**
(i) Day 0 (pre-treatment): difference in concentration for MDR TB versus non-MDR TB	-0.45	-1.3 to 0.39	0.3	-0.40	1.4 to 0.62	0.4	-0.16	-0.81 to 0.49	0.6
(ii) Non-MDR TB: daily change in concentration during treatment	**-0.26**	**-0.29 to -0.22**	**<0.001**	**-0.27**	**-0.31 to -0.22**	**<0.001**	**-0.072**	**-0.10 to -0.043**	**<0.001**
(iii) MDR TB: daily change in concentration during treatment	-0.044	-0.17 to 0.086	0.5	-0.0010	-0.16 to 0.14	0.9	-0.034	-0.14 to 0.074	0.5

Note: there was very strong evidence for an interaction between MDR TB and daily change in concentration for both FDA microscopy and quantitative culture (p=0.002 and p=0.001 respectively), but no evidence of an interaction for auramine microscopy (p=0.5).

Figure 3. 5 Difference between quantitative FDA microscopy and culture results (data from the first microscopy slide)
Histograms showing for each sputum sample the difference between (A) quantitative fluorescein diacetate (FDA) microscopy results minus quantitative culture colony forming units (CFU)results; and (B) the change in quantitative FDA microscopy results minus the change in quantitative culture CFU results during each interval of 3 days of treatment (baseline 0 days to 3 days, 3 to 6 days, and 6 to 9 days follow-up). All results are shown on a base-10 logarithm scale and log= base-10 logarithm. 68% to 79% of results agreed within ±1 log and, 95% to 96% agreed within ±2 log. All quantitative FDA microscopy results are the data from only the single first slide prepared from each sputum sample.

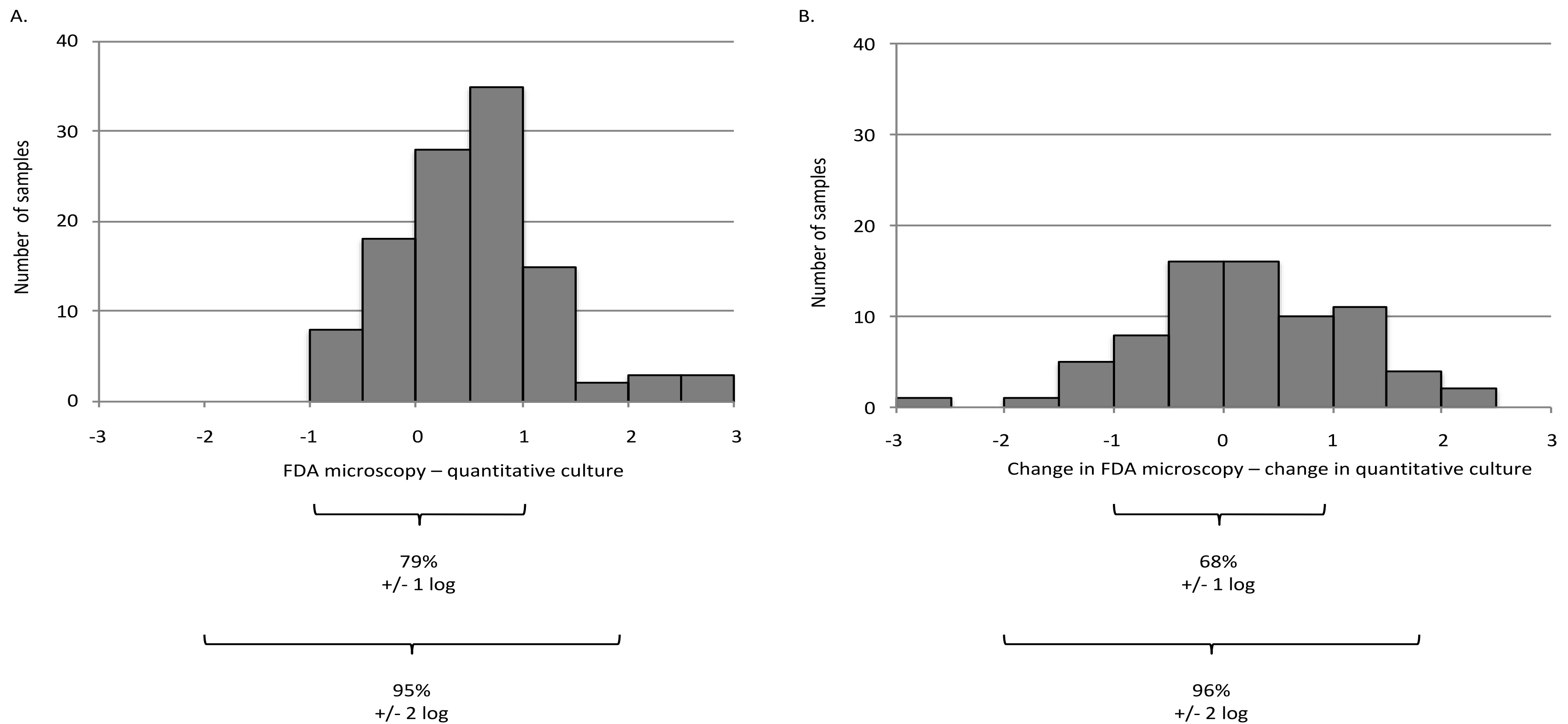
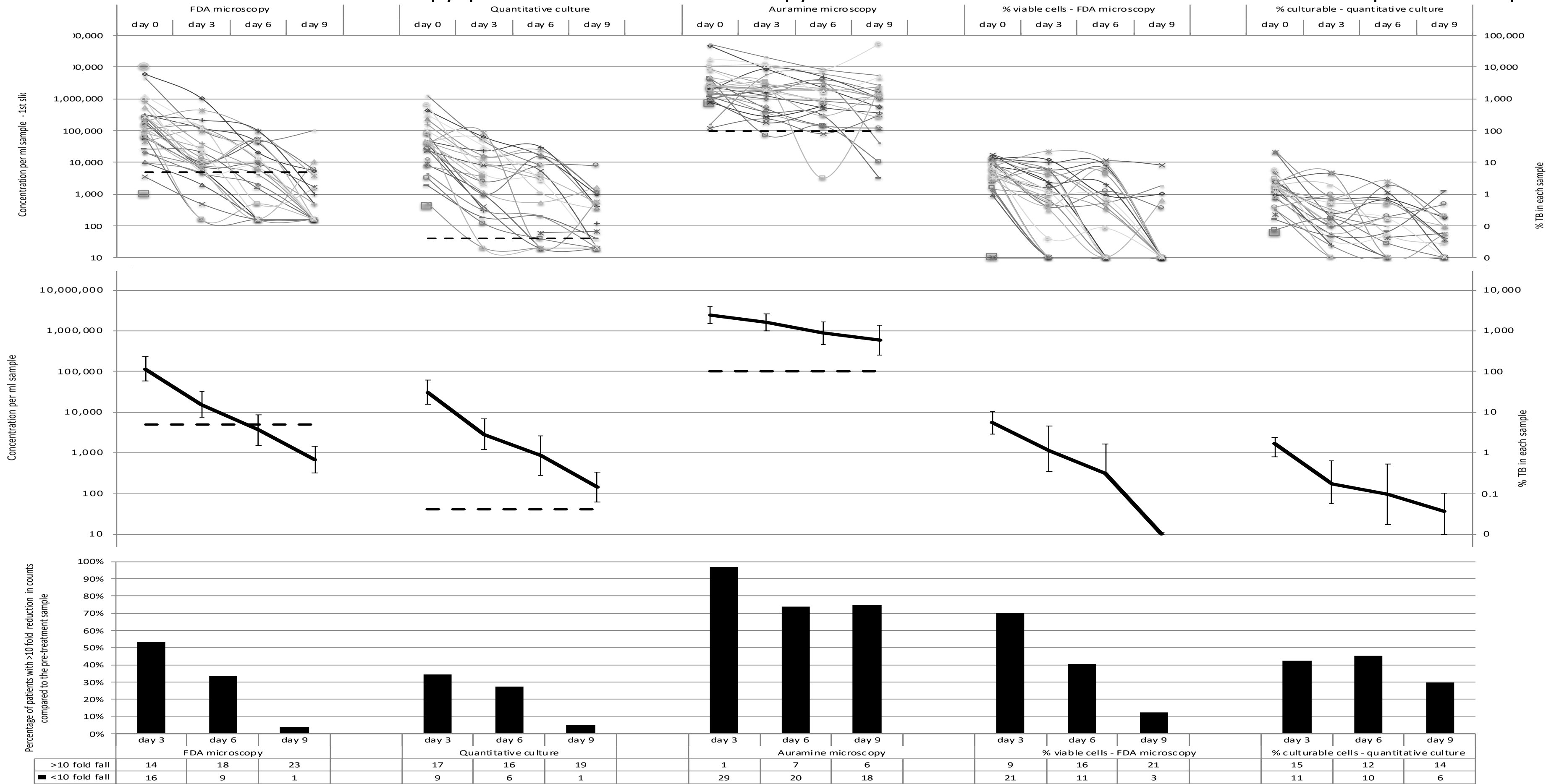


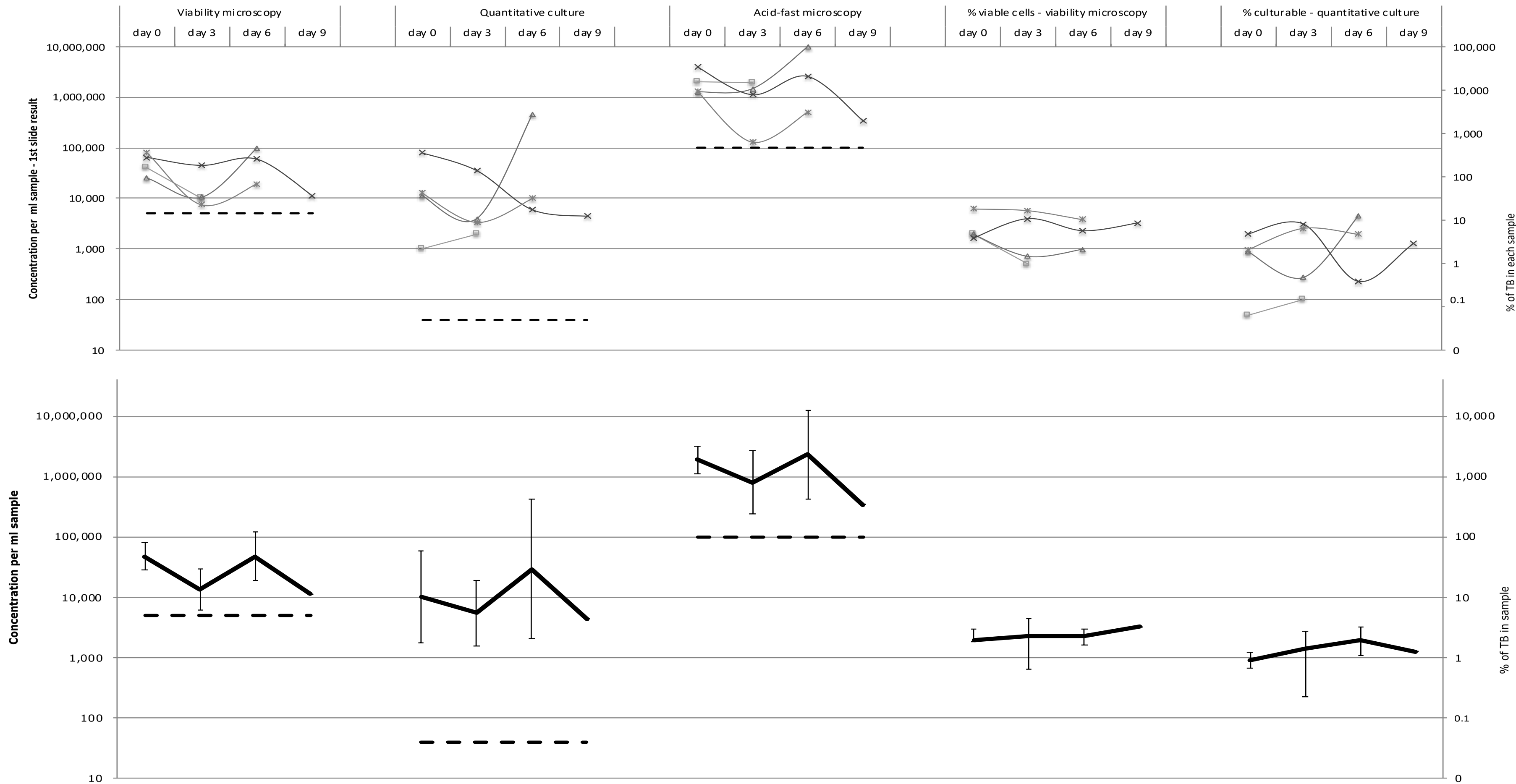
Figure 3. 6 Treatment response for 31 patients with non-MDR TB (data from the first microscopy slide)
Showing results from days 0, 3, 6 and 9 of TB treatment for: (a) each patient; (b) all patients; and (c) % of patients with <10-fold reduction. The horizontal axis shows days of treatment. Dashed lines indicate cut-offs for positivity. Proportion fluorescein diacetate (FDA) stained was calculated by dividing concentrations of FDA microscopy-positive bacteria by concentrations of auramine microscopy-positive bacteria. Similarly, proportion-culturable was calculated by dividing concentrations of colony forming units in quantitative culture by concentrations of auramine microscopy-positive bacteria. Microscopy results are from the first slide from each sputum sample.



	FDA microscopy day 3	FDA microscopy day 6	FDA microscopy day 9		Quantitative culture day 3	Quantitative culture day 6	Quantitative culture day 9		Auramine microscopy day 3	Auramine microscopy day 6	Auramine microscopy day 9		% viable cells - FDA microscopy day 3	% viable cells - FDA microscopy day 6	% viable cells - FDA microscopy day 9		% culturable cells - quantitative culture day 3	% culturable cells - quantitative culture day 6	% culturable cells - quantitative culture day 9
>10 fold fall	14	18	23		17	16	19		1	7	6		9	16	21		15	12	14
■ <10 fold fall	16	9	1		9	6	1		29	20	18		21	11	3		11	10	6

Figure 3. 7 Treatment response for 4 patients with MDR TB (data from the first microscopy slide)

Showing results from days 0, 3, 6 and 9 of TB treatment for: (a) each patient; and (b) all patients. The horizontal axis shows days of treatment. Dashed lines indicate cut-offs for positivity. Proportion FDA stained was calculated by dividing concentrations of fluorescein diacetate (FDA) microscopy-positive bacteria by concentrations of auramine microscopy-positive bacteria. Similarly, proportion-culturable was calculated by dividing concentrations of colony forming units in quantitative culture by concentrations of auramine microscopy-positive bacteria. Microscopy results are for the first slide from each sputum sample.



Discussion

Over the first 9 days of TB treatment, FDA microscopy predicted changes in culturable *M. tuberculosis* concentrations. FDA microscopy and quantitative culture changes appeared to separate into 2 distinct patterns during early treatment depending on whether patients had MDR TB. For patients with non-MDR TB, after 3 days treatment mean FDA microscopy and quantitative culture results fell by more than 90% and by day 9 by 99.6%. This concurs with early-bactericidal activity and serial sputum colony count studies of TB treatment response which reported exponential isoniazid-mediated decline in culturable *M. tuberculosis* during the first days of treatment, followed by slower rifampicin-mediated killing.[77,79,88,89] Acid-fast microscopy with auramine staining changed little during early treatment so diverged from FDA microscopy and quantitative culture as treatment continued.[78] In contrast to these results for patients with non-MDR TB, FDA microscopy and quantitative culture for the few patients with MDR TB differed significantly because they did not decrease during early treatment.

Studies in Bangladesh and Thailand reported that "positive-versus-negative" FDA microscopy after 2 months treatment identified patients needing drug-susceptibility testing for probable MDR TB.[18,90,91] Here we demonstrate the value of a quantitative FDA microscopy technique to screen patients with TB for poor treatment response and MDR TB within 9 days of starting treatment. Patients with MDR TB have worse outcomes and FDA microscopy during early treatment may allow the limited capacity for drug-susceptibility testing to be provided to the patients most likely to have MDR TB until resources may become sufficient for universal drug-susceptibility testing. This has the potential to improve patient outcomes and reduce MDR TB dissemination.[50]

FDA microscopy was analysed not only as positive concentrations, but also as the proportion of FDA-stained bacilli, calculated as the FDA microscopy result divided by the acid-fast microscopy results. We included this approach in case low-quality sputum specimens (or saliva submitted instead of sputum) caused misleading results, but this did not occur in this study that used sputum samples collected at home over 12-hour periods. These results suggest that the simpler assessment of the concentration of FDA microscopy-positive *M. tuberculosis* was as reliable, more consistent between samples and easier to analyse than proportion FDA-stained. To reduce the potential effect of random variation between microscopy slides, all microscopy tests were done in triplicate from each specimen. However, this increased workload and demonstrated that results from only the first microscopy slide gave similar results, perhaps because sputum homogenization and processing were carefully standardised to reduce slide-to-slide variations and the risk of bias comparing results between different techniques. Therefore, in future studies, only a single slide would be required for each microscopy technique from each sample.

Study limitations include that microscopy was only used in patients with initially acid-fast microscopy-positive TB. Filters or centrifugation can concentrate *M. tuberculosis* from sputum and may reduce this limitation.[92] Furthermore in this study, quantitative culture analysis used only crude colony counts. However, for samples with abundant growth, crude colony counts are likely to underestimate true mycobacterial concentrations because colonies coalesce, so are hard to count. In the future, high CFU counts can be replaced with estimations from the culture time to positive (TTP) result, using an equation calculated from experiments determining their relationship. The research protocol requirement for patients to collect sputum for 12 hours before commencing treatment slowed recruitment because TB treatment was usually started immediately following diagnosis. This logistical limitation was unrelated to patient characteristics, so participants were locally representative. Other limitations included missing samples when patients could not be located; the lack of pyrazinamide susceptibility testing; and some uninterpretable contaminated cultures. The culture contamination rate was within the range of previous studies using MODS but was higher than average,[93] perhaps because samples were collected over a longer than typical interval (12 hours). Despite these limitations, the study size was sufficient to demonstrate statistically significant differences between patients with non-MDR TB versus MDR TB.

Building on the results of this foundation study, the main study described in Research Chapter 7 is considerably larger, extending the results to assess long-term cure and the potential programmatic value of these findings. Importantly, it includes more patients with MDR TB and with HIV co-infection whilst assessing single-slide FDA microscopy of "spot" sputum samples both prior to and again 14 days after treatment initiation.

Research Chapter 4

Refinement – Sputum collection optimisation for TB testing

Background

One of the key challenges to global TB control is correctly diagnosing TB disease, and the WHO prioritises improving TB diagnostic guidelines and tests.[80] However, the sensitivity of a diagnostic test depends on the quality of the sputum samples obtained,[94,95] which has been the subject of much less attention than new diagnostic products.

Microscopy is inexpensive and is the most used TB laboratory test globally, but is only likely to be positive if the concentration of acid-fast *M. tuberculosis* bacilli exceeds 10,000 per ml of sputum.[78] TB culture techniques generally have greater sensitivity than microscopy, and PCR generally has intermediate sensitivity between that of culture and microscopy. However, these more expensive tests can only diagnose TB in samples containing sufficient concentrations of *M. tuberculosis*. Thus, poor quality sputum samples can lead to missed TB diagnoses for all tests. Inevitably, diagnostic sputum samples contain mixtures of respiratory secretions from diseased lung, that may contain *M. tuberculosis,* and healthy airway tract together with variable amounts of saliva. Consequently, the positivity of laboratory tests often vary between samples from the same patient,[58] so more than one sputum sample are usually tested from each person with suspected TB.[96,97] To mitigate within-patient variability between single sputum samples during the foundation study, sputa were collected over 12-hour periods (see Results Chapters 2 and 3, Methods section). However, the challenging logistics of collecting these samples contributed to attrition of participants during the follow-up sample collection at 3, 6 and 9 days after starting treatment.

Recommended sputum collection methods vary globally[36,98] and their relative merits for TB diagnosis are poorly characterised. Reviews have been carried out to guide the number of sputum samples required, and how many days they should be collected over.[2,96] However, no published reviews or meta-analyses assessing how best to collect sputum samples to improve TB diagnosis in community settings could be found.

A recent systematic review evaluated the effect of two pre-collection interventions (instructions and mouth washing) on indirect indicators of sputum quality, such as sputum viscosity.[99] The heterogeneity and paucity of data provided no definitive conclusion, and questioned the validity of using these indirect sputum quality indicators to predict diagnostic yield.[99] Other studies have assessed the utility of techniques such as sputum induction and bronchoscopy which are considered as 'invasive', with limited applicability in the resource-

constrained settings where most TB occurs because of biosafety, equipment availability and human capacity.[100] Therefore, the aim of this study was to systematically review and complete a meta-analysis of non-invasive sputum collection method effects on TB diagnostic yield for persons at least 12 years old suspected of having pulmonary TB. Participants aged less than 12 years were not included because both spontaneous expectoration of sputum and laboratory confirmation of TB are uncommon in children.[101]

Methods

Search strategy. The systematic review and meta-analysis were carried out according to a protocol based on international standards that was developed before data collection commenced.[102–104] The following databases were searched until 14/04/2017: PubMed, Google Scholar, ProQuest, Web of Science, CINAHL and EMBASE. The search terms used were "tuberculosis/TB", "sputum", "collection or clearance or submission or acquisition", "technique or guide or method" and "diagnosis". Additionally, references cited by these publications and relevant articles identified by the authors were searched by hand.

Inclusion criteria. Full text, peer-reviewed articles in English were selected that described experimental, case control or cohort studies comparing the effects of any two non-invasive sputum collection methods on the diagnostic yield of laboratory tests for pulmonary TB in humans who were at least 12 years old. Studies were included that also evaluated invasive collection methods, but only the comparison of non-invasive sputum collection methods was considered. Studies were included regardless of publication date and whether participants had co-infection with HIV. These inclusion criteria did not include studies that only assessed sputum quality or culture contamination rates, without assessing diagnostic yield.[99]

Exclusion criteria. In order to ensure that previous systematic review findings were not repeated, studies were excluded if they investigated only: the effects of sputum storage; collecting samples on the same versus different days; or the incremental yield of second and third sputum samples.[2,96]

Definitions. 'Spot' sputum collection referred to rapid, on-demand sputum collection during a single consultation, regardless of the time of day, usually at a health centre. For this research, 'pooled' sputum collection referred to sputum that was pooled from each spontaneous expectoration into the same sputum container over a period of several hours. 'Early-morning' sputum collection referred to sputum expectorated in the morning after waking. The current WHO guideline is to collect two spot sputum collections during the same day,[105] and is widely practised because this approach reduces the risk that patients may be lost to follow-up during testing.[21] In contrast, pooled and early-morning sputum collection require additional clinic visits: first to provide a sputum container and request the collection, then another for the sputum to be received for testing.

Reference standard. There is no gold standard sputum collection method, so the reference standard used in each study was considered to be the local standard (or 'control') sputum collection method for TB diagnosis in the reference group. In most studies, the reference group collected spot sputum samples.

Study selection. The inclusion/exclusion criteria were applied, and studies were grouped into those comparing the following aspects of sputum collection: (A) sputum collection time and duration; and (B) sputum collection instructions and techniques.

Data extraction. Data was extracted in duplicate using a standardised form. Some studies compared more than two sputum collection methods, concurrently providing data for different research questions. Discrepancies were resolved by discussion between the authors.

Data analysis. All analyses were carried out using STATA version 13 (STATA Corporation, College Station, TX, USA). All diagnostic yield data were calculated at the sputum-sample level, except where indicated otherwise (because the authors only reported the data at individual-case level). Positivity for each test was defined by the authors of each study. Diagnostic yields of sputum collection methods were analysed to calculate the odds ratio (OR) of the likelihood of a positive laboratory result, with 95% CI, displayed in forest plots. The threshold for statistical significance was $p<0.05$. All analyses were carried out separately for each of the laboratory tests used to diagnose TB (microscopy, culture, and PCR).

Pairwise meta-analyses were done with a random-effects model when there were 2 or more studies that investigated the same pre-defined research question with the same laboratory test using the DerSimonian and Laird method, because significant heterogeneity was expected.[106] Heterogeneity was assessed visually by forest plots, and analytically by I^2 and Cochrane Q test.

Meta-regression analysis using a logistic regression model of the study effect sizes was performed if there was statistically significant heterogeneity in the pairwise meta-analysis. Furthermore, to evaluate other factors that contributed to the likelihood of a positive TB diagnostic test, another meta-regression was performed to evaluate the study or sputum collection characteristics, including components of the instructions given that affected the odds of a positive TB diagnostic test in the study reference groups. Missing data for the meta-regression were imputed using the average result of the other studies, as shown in Table 4.1. When national TB incidence or HIV co-infection rates were not reported by a study, the published WHO or Centres for Disease Control (USA) data from the closest available year for the country where the study took place were used.[107–109]

Network meta-analysis data analysis. When pairwise meta-analyses revealed more than one method with significantly different diagnostic yield to the reference group, a

Bayesian network meta-analysis was used to identify the method with the best overall diagnostic yield[110] using the 'mvmeta' and 'network' STATA commands.[111]

Publication bias was assessed with funnel plots and rank Egger's test.[112]

Methodological quality of studies was assessed with the QUADAS-2 checklist because analysing diagnostic yield is comparable to diagnostic accuracy studies.[113]

Table 4.1 Study demographics

Note N=number, ZN=Ziehl-Neelsen, RCT=randomised control trial, HIV=human immunodeficiency virus, NS= not specified, NR=not reported, confirmed TB= microbiological confirmation of TB diagnosis, deco=decontamination, cent=centrifugation, neg=negative, RCT= randomised control trial, and 100k=100,000.

First author, year	Country	Study design	Microscopy stain	Processing prior to microscopy	Culture method	Participant eligibility	Participant (N)	% Female	Sample (N)	Recruited during treatment	National TB prevalence (per 100k)	TB HIV co-infection[b]
Mpagama, 2012	Tanzania	within patient	ZN	Deco and cent	Not done	Suspected TB	50	28%	150	0%	609	16%*
Andrews, 1959	India	within patient	Fluorescent	None	Solid	Confirmed TB	348	31%	1392	NR	465	0%
Majumdar, 1962	India	within patient	ZN	None	Solid	Confirmed TB	76	NR	222	42%	465	0%
Warren, 2000	USA	pre-post	Fluorescent	Deco and cent	Solid	Confirmed TB	65	NR	221	32%	7.9	15%
Abdel Aziz, 1985	Egypt	within patient	ZN	None	Not done	Suspected TB	90	NR	360	NR	82	0.01%
Krasnow, 1969	USA	within patient	Not done	Not done	Solid	Suspected TB	261	NR	1336	NR	19	0%
Kestle, 1967	USA	within patient	Not done	Not done	Solid	Suspected TB	183	NR	366	NR	23	0%
Ssengooba, 2012	Uganda	within patient	ZN	NS	Liquid	Suspected TB	1862	43%	3724	NR	170	50%
Pande, 1974	India	within patient	ZN	None	Solid	Confirmed TB	160	38%	320	38%	465	0%
Geldenhuys, 2014	South Africa	within patient	Fluorescent	Deco and cent	Liquid	Suspected TB	600	50%	1068	NR	696	22%*
Schoch, 2007	Switzerland	within patient	ZN	Cent only	Solid and liquid	Suspected TB	101	25%	167	0%	206	5%
Khan, 2007[a]	Pakistan	within patient	ZN	NS	Not done	Suspected TB	1520	NR	2819	0%	384	0.40%

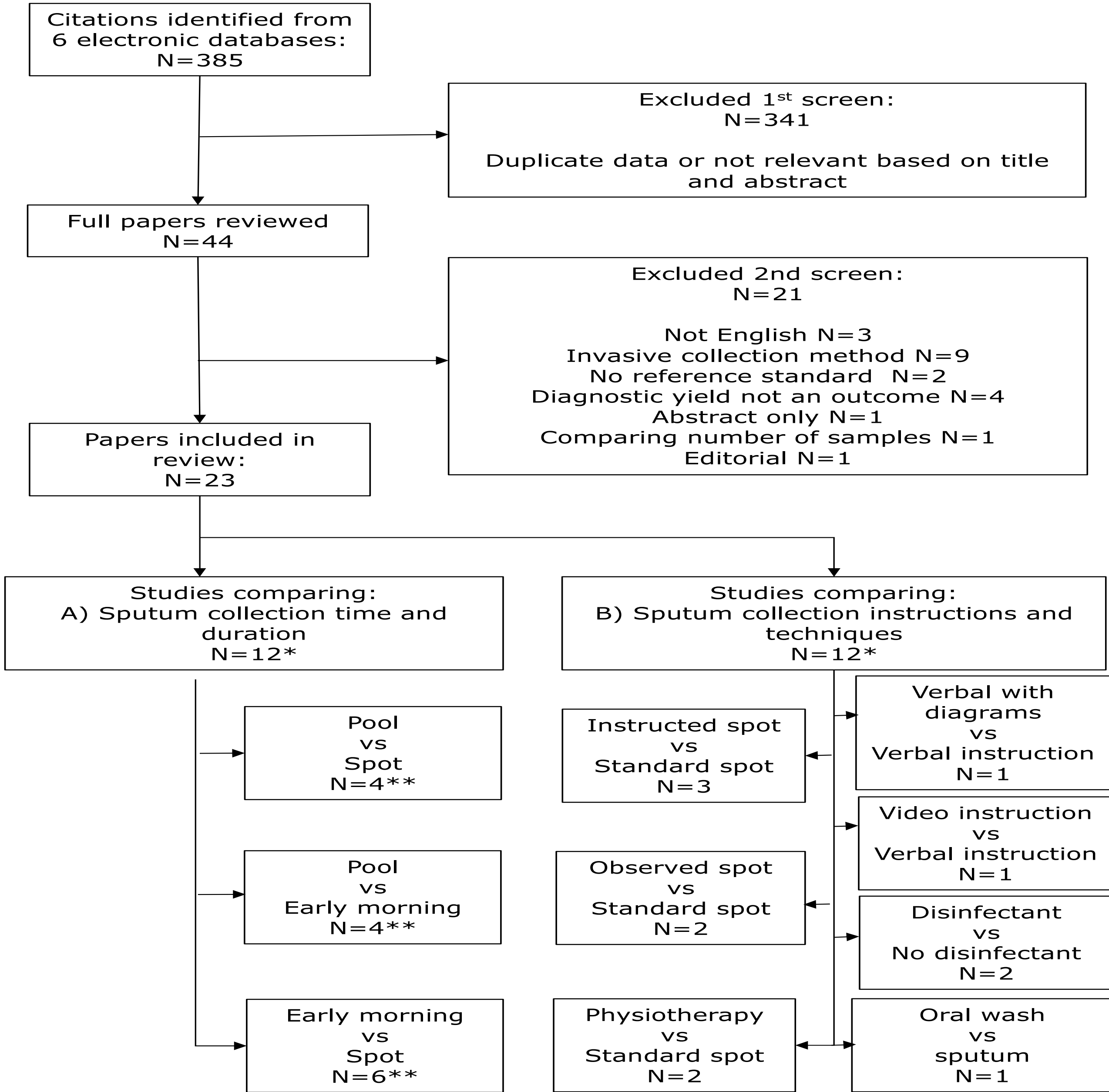
Khan, 2007[a]	Pakistan	RCT	ZN	NS	Not done	Suspected TB	3055	49%	5632	0%	384	0.40%
Alisjahbana, 2005	Indonesia	RCT	ZN	NS	Not done	Suspected TB	174	45%	505	0%	775	2.70%
Sakundarno, 2009	Indonesia	pre-post	NS	NS	Not done	Suspected TB	508	NR	1168	NR	715	4.60%
Mohamed, 2014	Sudan	RCT	ZN	NS	Not done	Suspected TB	328	NR	656	NR	151	3.40%
Maciel, 2009	Brazil	case control	Not done	Not done	Solid and liquid	Started TB treatment	119	45%	119	100%	58	20%
Bell, 2009	Malawi	within patient	ZN	Deco and cent	Solid	Suspected TB smear-neg	111	NR	220	0%	368	68%*
Souza, 2007	Brazil	within patient	ZN	Deco and cent	Solid	Suspected TB	132	NR	264	12%	58	100%*
Lee, 2013	South Korea	RCT	NS	NS	NS	Suspected TB	77	NR	228	0%	106	0.40%
Mhalu, 2015	Tanzania	RCT	Fluorescent	NS	Not done	Suspected TB	200	47%	200	NR	528	25%
Kalema, 2012	Uganda	RCT	ZN	None	Solid	Suspected TB	220	52%	440	0%	170	80%*
Peres, 2011	Brazil	RCT	Not done	Not done	Solid and liquid	Suspected TB	120	NR	240	NR	57	17%
Davis, 2009	Uganda	within patient	PCR	Not done	Not done	Suspected TB	127	NR	254	0%	108	46%*

FOOTNOTE. In addition to the variables shown, the following data were extracted from each study: research question, number of sputum samples, sputum collection characteristics, exact instructions given and method, sample transport details, sputum processing, laboratory methods, sample volume. Laboratory test results were extracted as: diagnostic yield (positive results as a proportion of all positive and negative results); and for microscopy also the concentration of acid-fast bacilli visualised.
a=This study assessed instructed spot vs uninstructed spot sputum collection, but also carried out a sub-analysis comparing instructed early morning vs. instructed spot sputum collection.
b=National data was used in this variable, unless indicated by an asterisk.
*=Data reported within the study

Results

Study characteristics are shown in Table 4.1. Eligibility criteria were fulfilled by 23 studies published from 1959 to 2017 that involved 8,967 participants who provided 19,252 sputum samples (Figure 4.1). A within-patient design in which participants each provided sputum for both the intervention and reference groups was used by 57% (13/23) studies; and 83% (19/23) took place in low or middle-income countries.

Figure 4.1 Study flow chart



FOOTNOTE.
* = Findings from one study contributed to both (A) and (B).
**= One study compared Pool Vs Spot Vs Early morning and therefore contributed to all meta-analyses in (A).

Diagnostic testing by sputum microscopy was assessed in 18 studies, 67% (12/18) of which used ZN staining with light microscopy, 22% (4/18) fluorescent microscopy and 11%

(2/18) did not provide this information. TB culture was assessed in 14 studies: 64% (9/14) used solid culture, 14% (2/14) used liquid culture, and 21% (3/14) used both. PCR was assessed in one study.

Main findings. A summary of all the main findings are shown in Table 4.2, and these results are described in detail below.

Table 4.2 Summary of the main findings of pairwise and network meta-analyses
Note: statistically significant findings are in bold font, N=number and * indicates comparison versus uninstructed spot.

MICROSCOPY	Pool vs. Spot	Instructed spot vs Uninstructed spot	Early morning vs Spot	Pool vs Early morning
Pairwise	**3 studies**	**6 studies**	6 studies	2 studies
N of samples	**1764**	**8445**	8248	510
OR (95% CI)	**1.6 (1.3 - 1.9)**	**1.6 (1.3 - 2.0)**	1.5 (0.9 - 2.6)	2.1 (0.9 - 4.2)
p value	**p<0.001**	**p<0.001**	p=0.2	p=0.07
Heterogeneity I^2, p value	0.0%, p=0.5	25%, p=0.3	84%, p<0.001	45%, p-0.2
Pairwise excl. outlier			5 studies	
N of samples			7928	
OR (95% CI)			1.1 (0.95 - 1.3)	
p value			p=0.2	
Heterogeneity I^2, p value			0.0%, p=0.7	
Network				
OR (95% CI)	**1.8 (1.1 -3.0)***	**1.8 (1.1 - 2.8)**	1.4 (0.9 - 2.1) *	1.3 (0.7 - 2.4)
p value	**p=0.03**	**p=0.01**	p=0.2	p=0.4
CULTURE				
Pairwise	**2 studies**	3 studies	4 studies	2 studies
N of samples	**1614**	603	5279	1702
OR (95% CI)	**1.7 (1.2 - 2.4)**	1.1 (0.1 - 1.7)	1.4 (0.9 - 2.4)	1.2 (0.6 - 2.2)
p value	**p=0.01**	p=0.6	p=0.2	p=0.6
Heterogeneity I^2, p value	0.0%, p=1.0	0.0%, p=0.6	50%, p=0.1	81%, p=0.02

(A) Sputum collection time and duration were assessed by studies that compared: pooled versus spot sputum collection (n=4); pooled versus early morning sputum collection (n=4), and early morning versus spot sputum collection (n=6). Findings from one study were included in all three comparisons. Pooled sputum was collected: overnight; for 24 hours; or for 72 hours (Table 4.3).

Table 4.3 Sputum collection instructions
Full description of the instructions provided to participants in each arm of the studies included in this review. Note. n=number of samples collected, NS=not specified, ml=millilitre and physio=physiotherapy

	Comparison group			Reference group		
First author, year	**Sample type**	**n**	**Instructions given**	**Sample type**	**n**	**Instructions given**
Mpagama, 2012[114]	Pool (Overnight)	1	"Collect in a wide mouth container"	Spot or Early morning	1	"Collect in a wide mouth container"
Andrews, 1959[115]	Pool (Overnight /24 hour)	2	"Asked to fill half of the container with sputa or all saliva if scanty"	Spot	2	"Produced either sputum or saliva on demand within a minute or 2 between 07:30 - 13:00"
Majumdar, 1962[116]	Pool (24 hour)	1 or 2	"Aided by health care worker"	Spot	1 or 2	"Asked to produce sputum"
Warren, 2000[117]	Pool (5 ml)	NS	"Accepted samples with more than 5ml of sputum. If less than 5 ml, sample kept for 48 hours and when another sample arrived pooled them to 5 ml"	Spot	NS	NS
Abdel-Aziz, 1985[118]	Pool (24 hour)	1	"Collect in a wide mouth container. Advised to collect true sputum from the depth of their chest"	Early morning	3	"Collect in a wide mouth container. Advised to collect true sputum from the depth of their chest"
Krasnow, 1969[119]	Pool (24 hour)	2 or 3	Not specified	Early morning	2 or 3	Not specified
Kestle, 1967[120]	Pool (24-72hour)	1	"Collect sample not exceeding 40-50 ml"	Early morning	1	"Collect any sputum raised after waking not exceeding 10 ml"
Ssengooba, 2012[121]	Early morning	1	Not specified	Spot	1	Not specified
Pande, 1974[122]	Early morning	1	"Produce sample in the early hours within a matter of few minutes before cleaning mouth"	Spot	1	"Collect sputum or saliva in a wide mouth container"
Geldenhuys, 2014[123]	Early morning	1	"Asked to take deep breath in and hold for 1 second then cough deeply and vigorously when breathing out. Advised to collect thick mucoid sputum, not saliva or nasal secretions"	Spot	1	"Asked to take deep breath in and hold for 1 second then cough deeply and vigorously when breathing out. Advised to collect thick mucoid sputum, not saliva or nasal secretions"

Schoch, 2007[124]	Early morning	1	"One hour was allowed for spontaneous expectoration of sputum"	Spot	1	"One hour was allowed for the production of spontaneous sputum"
Khan, 2007[125]	Spot and Early morning	2	"Take 3 deep breaths, then cough deeply, aim for 5 ml. Emphasised the importance of sputum rather than saliva providing visual difference"	Spot and Early morning	2	No instructions given
Alisjahbana, 2005[126]	Spot and Early morning	3	"Explained the importance of sputum examination and instructed patients on how to produce adequate samples. A picture of good quality sputum (purulent and thick) was shown and they were specifically instructed to produce at least 5 ml of sputum"	Spot and Early morning	3	"Give 3 samples to the laboratory spot, morning, and spot"
Sakundarno, 2009[127]	Spot and Early morning	3	"Healthcare workers were trained about advising patients to collect samples greater 3 ml of the colour and consistency of sputum"	Spot and Early morning	3	"Healthcare workers were not trained"
Mohamed, 2014[128]	Spot and Early morning	2	"Advised to rinse the mouth, clear the nose and throat, try a hot drink, place both hands in the waist. Take 2-3 deep breaths and cough. Participants were observed"	Spot and Early morning	2	Not specified
Maciel, 2009[129]	Early morning	1	"Advised to use a toothbrush to remove food residue prior to collecting 10 ml. Participants were observed"	Early morning	1	"Advised to collect first morning sample after fasting the previous night"
Bell, 2009[100]	Physio-assisted spot	1	"Deep breaths were taken whilst the physiotherapist carried out cupped clapping on chest, shake for 5 minutes and then asked to blow with pursed lips"	Spot	1	"Collect self-expectorated sputum sample under direct observation"
Souza, 2007[130]	Physio-assisted spot	1	"This manoeuvre is based on slow expiration from the functional residual capacity with the glottis open and the patient in lateral decubitus. The	Spot	1	Not specified

			physiotherapist stands behind the patient and uses the right hand to exert counter-supportive pressure on the superolateral rib cage. With the open palm of the left hand, the physiotherapist exerts counter- supportive pressure on the opposite lateral-costal wall. This technique was performed 12 times for approximately 20 minutes per session."			
Lee, 2013[131]	Early morning	3	"After rinsing the mouth collect expectorated sputum from the lungs after productive cough in morning." Instructions given verbally with diagrams.	Early morning	3	"After rinsing the mouth collect expectorated sputum from the lungs after productive cough in morning." Instructions given verbally.
Mhalu, 2015[132]	Early morning	1	4 min video in Swahili explaining that the collection required secretions from bronchial tree, and not saliva, demonstrating both. Advised to inhale and exhale deeply 3 times and cough after the last breath. This should be done on waking after inhaling steam. Also demonstrates the volume.	Early morning	1	“Sit or stand in an open space, inhale deeply 2-3 times, breathe hard each time and cough as hard as possible after the last breath.”
Kalema, 2012[133]	Spot	2	"Prior to sputum expectoration, patients performed a 60 second oral rinse with chlorhexidine followed in immediate succession by another 60 second rinse with nystatin. Participants were observed"	Spot	2	"Emphasised the importance of sputum rather than saliva providing visual difference. Advised to take 3 deep breaths, then cough deeply, aim for 5 ml. Participants were observed. For the early morning samples advised to do it on awakening. "

Peres, 2011[134]	Early morning	1	"Advised to rinse their mouths with 10 ml of a commercially available mouthwash solution for 1 minute before sample collection"	Early morning	1	"Advised to washed hands with soap and water and rinsed their mouths with tap water before collecting the specimens"
Davis, 2009[135]	Spot	1	"For the collection of oral wash specimens, subjects were instructed to cough vigorously 5 times, then gargle 10 ml of sterile saline for 60 seconds"	Spot	1	Not specified

Figures 4.2 and 4.3 demonstrate that compared with spot sputum collection, pooled collection significantly increased the odds of a positive result in sputum microscopy (OR=1.6, 95% CI=1.3-1.9, $p<0.001$) and culture (OR=1.7, 95% CI=1.2-2.4, p=0.01). These study results had no significant heterogeneity (p=0.5 and p=1.0, respectively). One study compared spot sputum collection versus pooled collection until a volume of 5 ml sputum had been collected, which was excluded because this collection method was so different from the other studies.[117]

Pooled sputum collection tended to have higher odds of a positive microscopy result than early morning sputum collection (OR=2.1, 95% CI=0.9-4.2, p=0.07). This was not the case for culture results (0R=1.2, 95% CI=0.6-2.2, P=0.6). However, one of these two studies only reported individual-case level data.[119] This pairwise meta-analysis had significant heterogeneity (I^2=81%, p=0.02), as there were only two studies, a meta-regression could not be carried out.

Studies comparing spot versus early-morning sputum collection had the largest total sample size and demonstrated no significant difference in the diagnostic yield for microscopy (8,248 samples, OR=1.5 95% CI=0.9-2.6, p=0.2) or culture (5,279 samples, OR=1.4 95% CI=0.9-2.4, p=0.2). There was significant heterogeneity in the effect size within the studies in the pairwise meta-analysis for sputum microscopy (I^2=84%, $p<0.001$), and not for culture (I^2=50%, p=0.2). Meta-regression (Table 4.4) revealed that all heterogeneity between the studies using microscopy was explained by one study that included only participants with microbiologically confirmed TB (coefficient=1.5, 95% CI=0.74-2.3, p=0.006), whereas all other studies included participants with suspected TB. Excluding this study did not affect the result (OR=1.1, 95% CI=0.95-1.3, P=0.2), but eliminated heterogeneity (I^2=0.0%, p=0.7).

(B) Sputum collection instructions and techniques were assessed by reviewing studies that compared: instruction prior to sputum collection versus none (n=3); instruction provided during observation by a healthcare professional versus none (n=2); instruction provided whilst a physiotherapist assisted versus unassisted expectoration (n=2); verbal instructions with diagrams versus verbal instructions (n=1); video versus verbal instructions (n=1); disinfectant mouthwash prior to sputum collection versus none (n=2); and mouthwash used as a diagnostic sample versus sputum (n=1).

Table 4.4 Meta-regression to explain heterogeneity between studies comparing early morning versus spot sputum collection Meta-regression assessing factors that contributed to the heterogeneity in odds ratios for a positive versus a negative sputum microscopy result within studies that compared early morning versus spot sputum collection. The meta-analysis of these 6 studies had I-squared 84% ($p<0.01$). Variables that effected heterogeneity with statistical significance ($P<0.05$) is indicated by bold font, and the residual heterogeneity describes the I^2 if the independent variable was equal to 0.

Independent variable	Coefficient	95% confidence interval	p value	Residual I^2 heterogeneity
TB prevalence (per 100,000 people)	-0.0009	-0.0059 - 0.0041	0.6	87%
TB/HIV coinfection (%)	-0.0088	-0.067 - 0.50	0.7	86%
Female participants (%)	0.01	-0.10 - 0.12	0.8	87%
Recruited only microbiologically confirmed TB cases	**1.5**	**0.7 - 2.3**	**0.006**	**0%**
Studies carried out after the change in “Smear positive” definition	-0.63	-2.3 – 1.0	0.4	85%
Participants recruited during treatment (%)	**0.04**	**0.02 - 0.06**	**0.006**	**0%**
Fluorescent microscopy used	-0.51	-2.8 - 1.8	0.6	86%
Instructions given to study either or both arms	-0.53	-2.3 - 1.2	0.4	78%

Instructions given for sputum collection (for either intervention or reference groups) varied greatly between all studies and are shown in Table 4.3: 35% (8/23) of the studies did not mention the instructions, if any, given to the reference group. The advice given in the intervention groups to produce sputa included: spontaneous expectoration; taking numerous deep breaths and coughing; holding breath for a second and coughing; drinking a hot drink prior to collection; or physiotherapy-assisted pressure on the chest accompanied by slow expiration. One study found that compared to verbal instructions there was no difference in the odds of positive microscopy if the instructions were given with diagrams (p=0.6), whilst another study found video instructions to be better (OR=4.3, 95% CI=2.3-7.9, p<0.001).

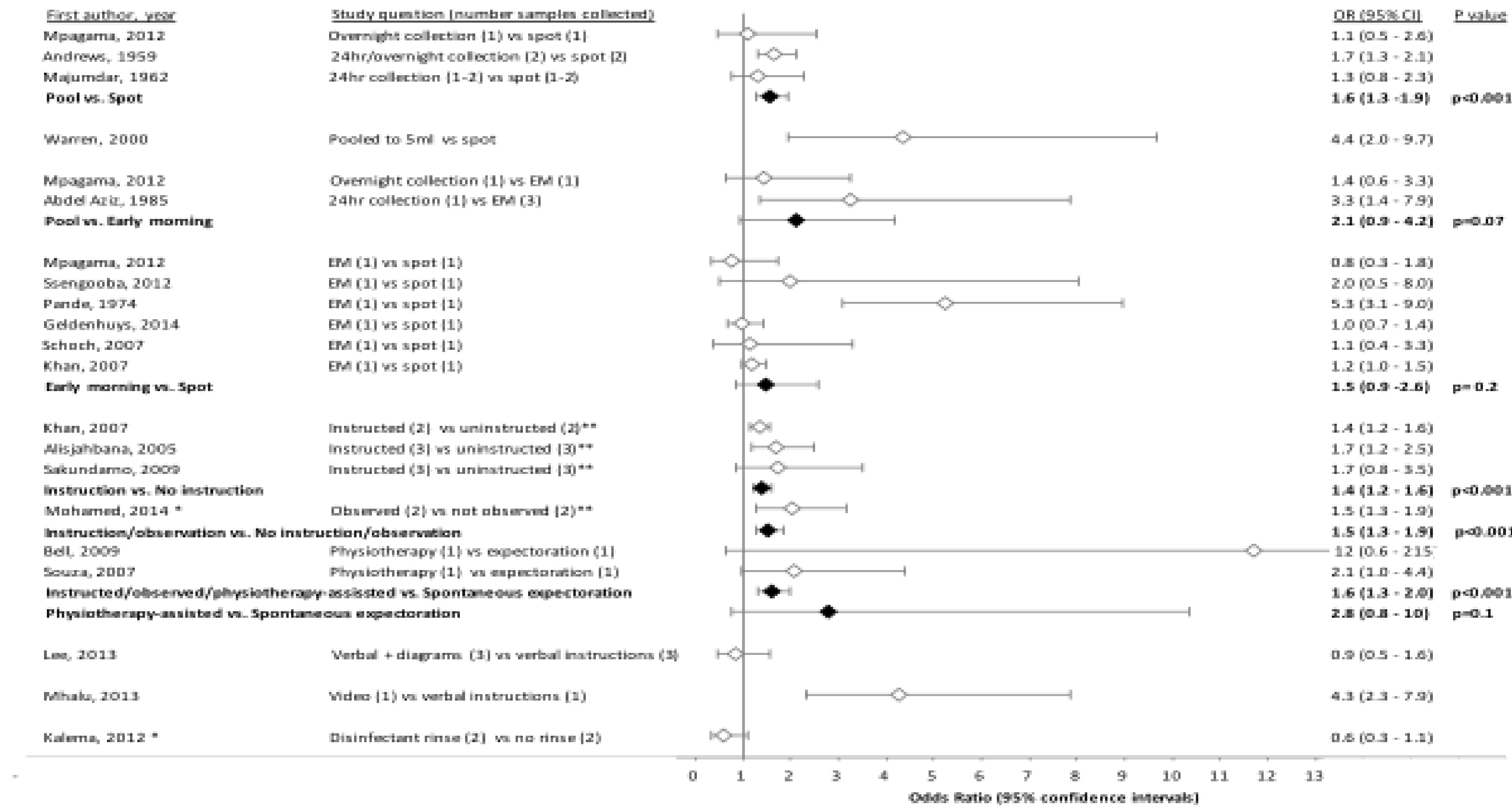
Pairwise meta-analyses demonstrated that instructions on how to produce sputum and/or emphasising the difference between sputum and saliva (verbally or visually) increased the odds of positive microscopy (OR=1.4, 95% CI=1.2-1.6, p<0.001, Figure 4.2), with no significant heterogeneity (p=0.4). None of these studies used culture for diagnosis. There were not enough studies for meta-analysis to investigate whether observing patients whilst collecting sputum affected the diagnostic yield of microscopy or culture. Compared to standard spot sputum collection, physiotherapy-assisted collection did not significantly affect microscopy or culture diagnostic yield (Figures 4.2 and 4.3), although these were small studies. If these interventions are considered as differing levels of guidance to assist participants to produce spot or early-morning sputum, then the odds of a positive microscopy result increased cumulatively: 1.4-times (95% CI=1.2-1.6, p<0.001) by instructions; 1.5-times (95% CI=1.3-1.9, p<0.001) by instructions and/or observing the participant; and 1.6-times (95% CI=1.3-2.0, p<0.001) by instruction, observation and/or physiotherapy (Figure 4.2), with no significant heterogeneity (p=0.2). Too few studies have been published for a meaningful equivalent analysis for mycobacterial culture (Figure 4.3).

Mouth washing with disinfectant was evaluated by two studies and did not demonstrably affect the odds of a positive result in microscopy or culture (Figures 4.2 and 4.3). One of these studies was analysed with individual-case level data.[133]

The study that assessed PCR demonstrated that although TB could be diagnosed by testing oral rinse samples, this technique had lower odds of a positive PCR result than spot sputum collection (OR=0.56, 95% CI=0.34-0.93, p<0.05). This study was not included in any meta-analysis or following meta-regression because it used a different laboratory test and collection method to any other study.

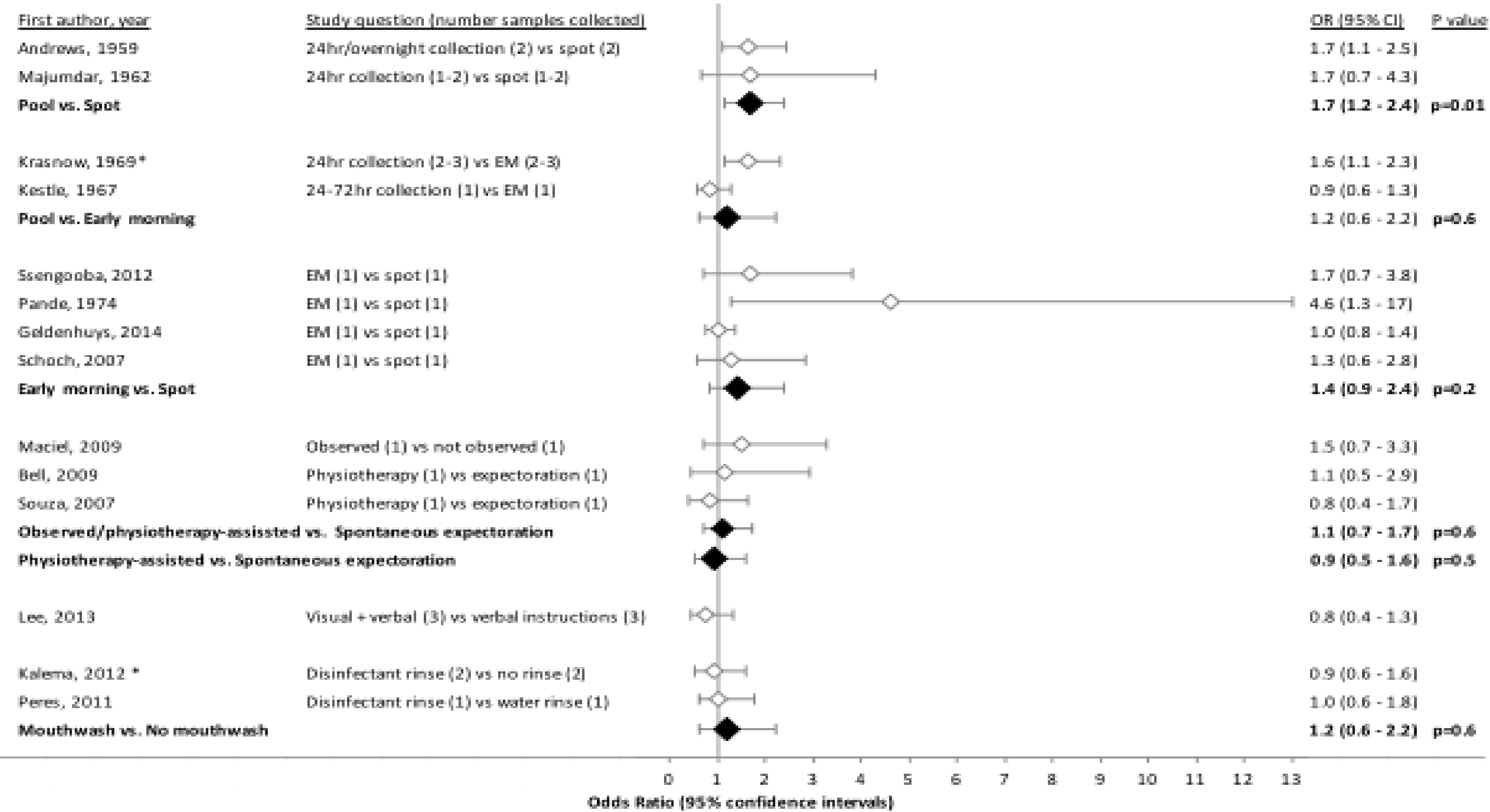
Figure 4.2 Forest plot for sputum microscopy

Forest plot demonstrating both individual study (white diamonds) and combined meta-analysis (black diamonds and bold type face) results for a positive sputum microscopy. Note. hr=hour, EM=early morning, +=and, vs=versus, OR=odds ratio and 95% CI= 95% confidence intervals.



FOOTNOTE. * Indicates that OR was calculated from individual-level data, not sputum sample-level data. ** Indicates that the samples collected in both comparison and reference group were a combination of spot and early morning samples.

Figure 4.3 Forest plot for mycobacterial culture
Forest plot demonstrating both individual study (white diamonds) and combined meta-analysis (black diamonds and bold type face) results for a positive mycobacterial culture. Note. hr=hour, EM=early morning, +=and, vs=versus, OR=odds ratio and 95% CI= 95% confidence intervals.



FOOTNOTE. * Indicates that OR was calculated from individual-level data, not sputum sample-level data.

Integrated analyses of (A) sputum collection time and duration and (B) sputum collection instruction and techniques.

Meta-regression of factors predicting diagnostic yield in the reference group was possible because all studies used either spot and/or early-morning sputum collection as their reference group collection method. This demonstrated that the collection of early-morning sputum collection did not affect the variation in odds of positive microscopy or culture in the reference group (p=0.8 and p=0.7 respectively, Table 4.5). The only factor that increased the odds of a positive microscopy or culture test in the reference group was if the study only recruited individuals with microbiologically confirmed TB. Increasing prevalence of HIV co-infection, and studies carried out in recent years decreased the odds of a positive culture result. More recent studies were also associated with increased odds of positive microscopy (Table 4.5).

Network meta-analysis was only possible for the results of sputum microscopy in 15 studies. These assessed: uninstructed spot; pooled; instructed (including observed and physiotherapy-assisted) spot and early morning sputum collection. Four studies assessing the value of instructions, indicated in Figure 4.2, requested a combination of spot and early morning samples in both comparison and reference groups. However, because the pairwise meta-analysis comparing early-morning with spot sputum demonstrated no significant difference in diagnostic yield, these studies in the network meta-analysis were considered as comparing uninstructed versus instructed spot samples. Compared to uninstructed spot collection, the network meta-analysis demonstrated that the odds of positive sputum microscopy were increased by pooled sputum samples (OR=1.8, 95% CI=1.1–3.0, p=0.03) or by instructed spot sputum collection (OR=1.8, 95% CI=1.2–2.8, p=0.01, Figure 3.4). The network estimated that there was no difference in the odds of positive sputum microscopy between pooled versus instructed spot collection (OR=0.99, 95% CI=0.50–2.0, p=1.0, Figure 4.4). Pooled (47%) and instructed spot collections (46%) had similar likelihood of being the best method to obtain a positive microscopy test within the network.

Microscopy grade. Meta-analysis of the grade[36] of positivity of diagnostic test results was only possible for the effect of pooled sputum collection on sputum microscopy, because this was the only research question addressed by at least 2 studies. Combining results of these 2 studies demonstrated that pooled sputum collection significantly increased both the positivity and the grade of positivity of sputum microscopy (Figure 4.5).

Table 4.5 Metaregression to explain variability within reference group results
Meta-regression assessing factors that contributed to the variation demonstrated in the likelihood (odds) of a positive diagnostic test within the reference groups of each study. Note. Not applicable results are represented by NA, and 95% CI=95% confidence intervals. Statistical significance ($p<0.05$) is indicated by bold font

	Sputum microscopy (n=20 studies)			**Culture (n=14 studies)**		
	Coefficient	**95% CI**	**p value**	**Coefficient**	**95% CI**	**p value**
National TB prevalence (100,000 people)	0.00020	-0.0038 - 0.0042	0.9	0.0023	-0.0035 - 0.0081	0.4
TB HIV co-infection (%)	-0.033	-0.072 - 0.007	0.1	**-0.031**	**-0.063 - 0.0006**	**0.05**
Year	**-0.051**	**-0.10 – 0.003**	**0.04**	**-0.068**	**-0.012 -0.03**	**0.002**
Studies carried out after the change in “Smear positive” definition	-0.31	-2.3 – 1.7	0.7	NA	NA	NA
Female participants (%)	-0.069	-0.20 - 0.060	0.3	-0.057	-0.24 - 0.13	0.5
Recruited only microbiologically confirmed TB cases	**2.2**	**-0.088 - 4.2**	**0.04**	**3.5**	**1.4 - 5.6**	**0.004**
Participants recruited during treatment (%)	0.048	-0.018 - 0.11	0.1	0.017	-0.027 - 0.061	0.4
Fluorescent microscopy used	1.0	-1.3 - 3.3	0.4	NA	NA	NA
Sample concentration prior to microscopy	-0.51	-2.1 – 2.0	1.0	NA	NA	NA
Liquid culture used	NA	NA	NA	**-2.3**	**-4.4 – 0.12**	**0.04**
Breathing manoeuvre to aid sputum production	-0.11	-2.8 – 2.5	0.9	-0.32	-3.8 - 3.2	0.8
Instructed on sputum quality	-0.21	-2.2 - 1.8	0.8	-0.88	-3.5 - 1.8	0.5
Early-morning samples collected	0.45	-1.5 - 2.4	0.6	0.42	- 2.2 - 3.0	0.7
Number of samples collected per patient	0.53	-0.66 - 1.7	0.4	0.90	-0.92 - 2.7	0.3

FOOTNOTE. The meta-regression analysed the odds of positive sputum microscopy or culture in the reference group. It included the variables shown above and: research question, type of instructions, and sputum processing methods. The effect of liquid culture on the odds of a positive culture result lost its association when the year of the study was included in the meta-regression (Coefficient=-0.82, 95% confidence interval=-2.9, 1.2 and p=0.4). Note that n indicates the number of analyses versus a reference standard that were included in this meta-regression, which is different from the number of publications (see Figure 4.1).

Figure 4.4 Network meta-analysis
Forest plot demonstrating the results of the network meta-analysis with the outcome being the odds ratio for a positive sputum microscopy result. Note 95% CI=95% confidence interval; vs= versus.

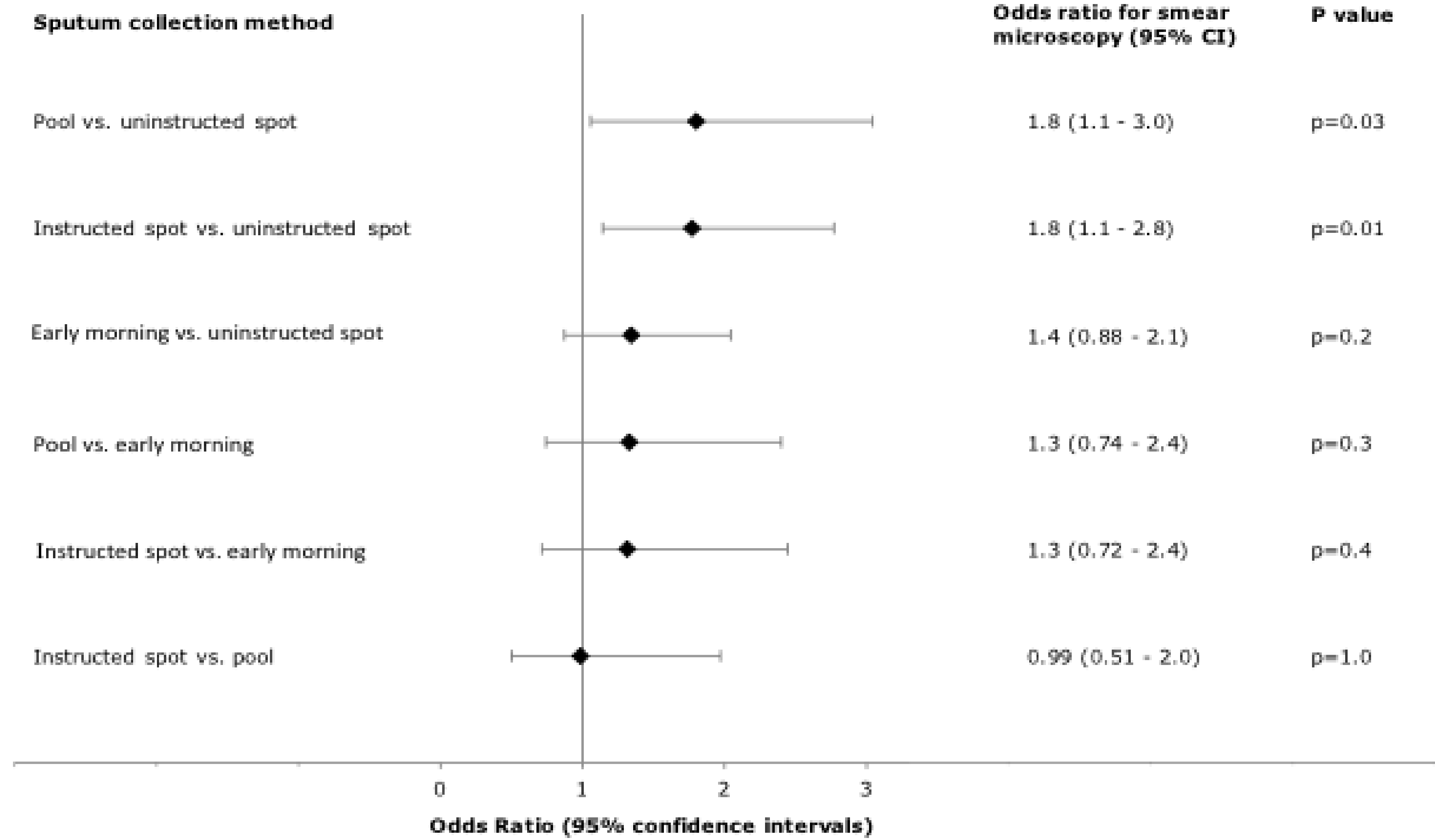
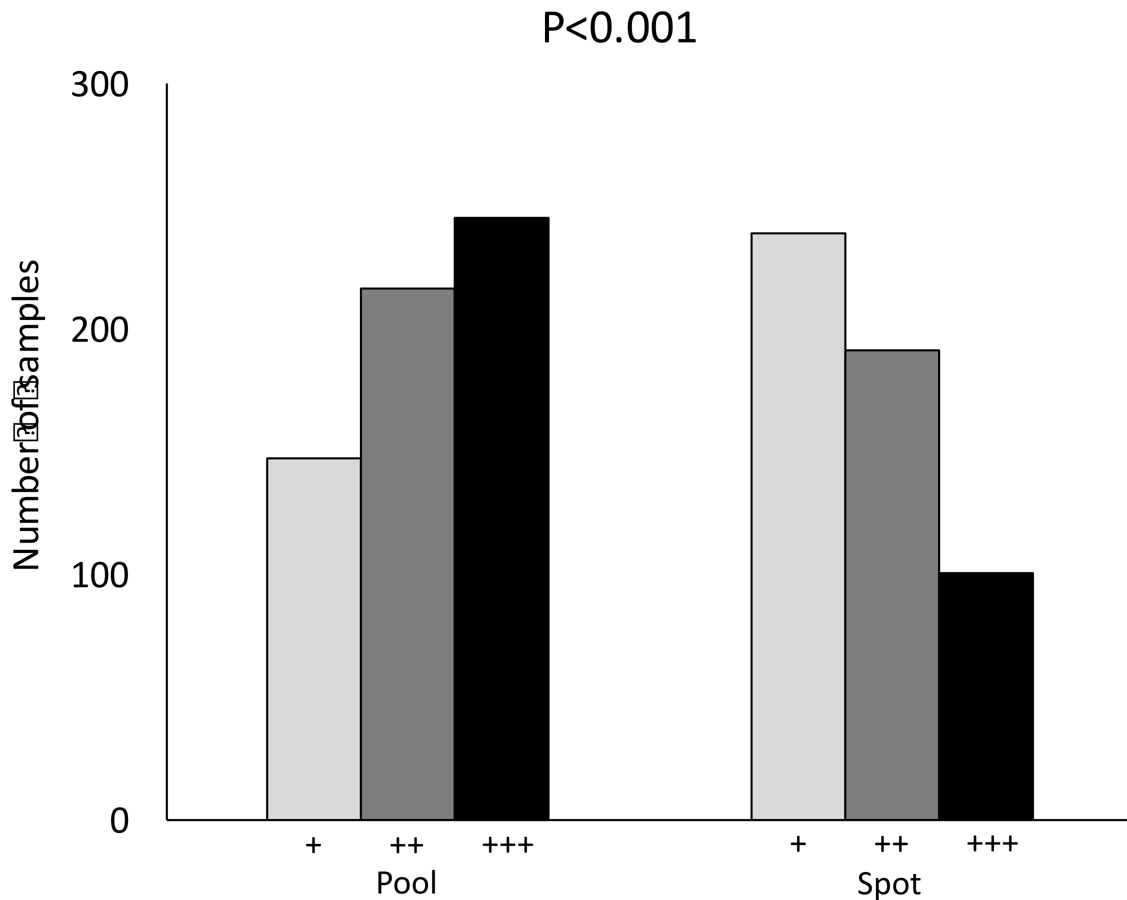


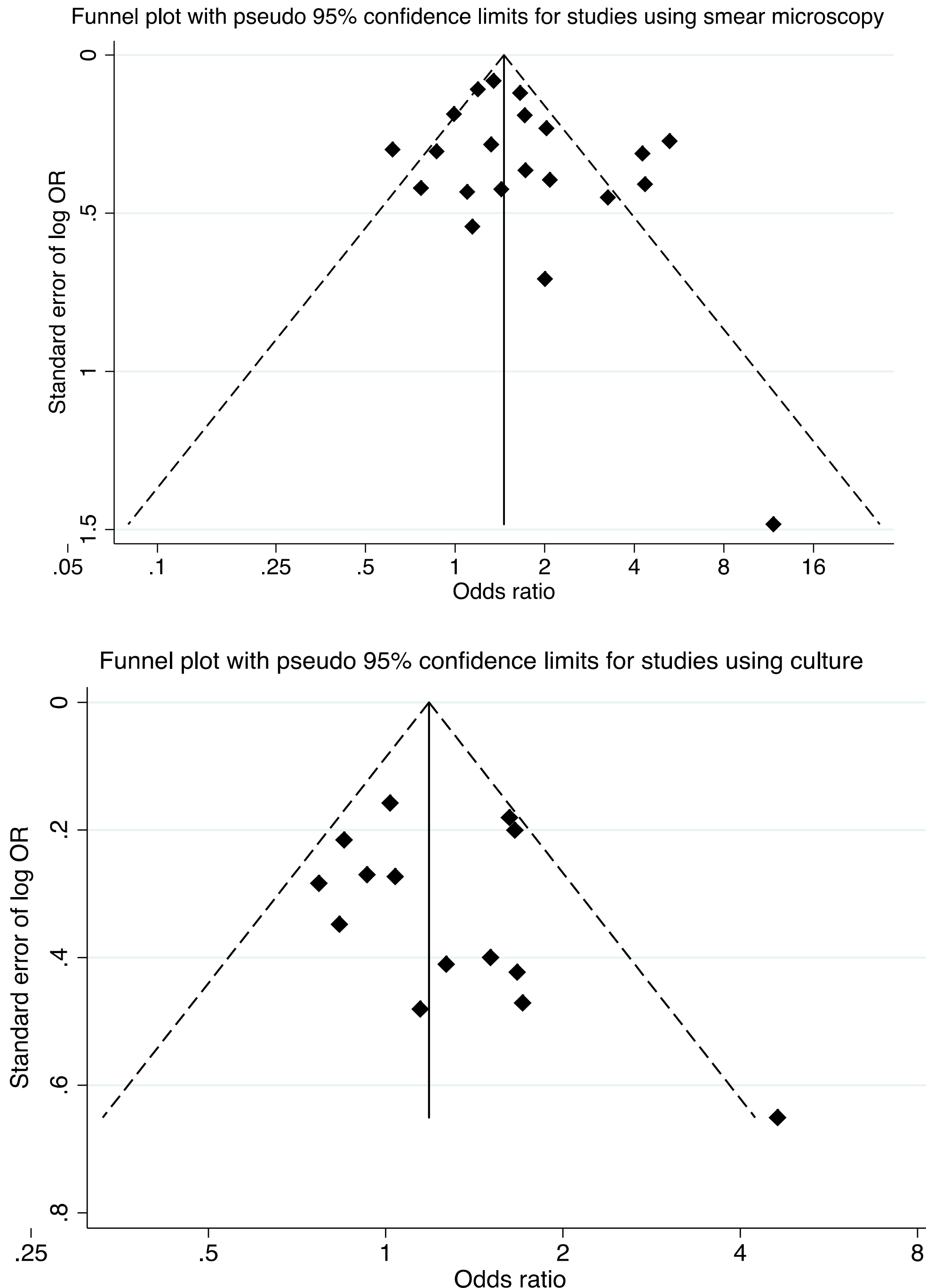
Figure 4.5 Bar chart of mycobacterial load
Bar chart comparing the difference in mycobacterial load within sputum microscopy positive samples between pool and spot sputum collection. This analysis involved combining the results reported by Andrews et al.1959 and Majumdar et al. 1962, which demonstrated 75% (n/N=608/807) positive sputum microscopy results in pooled collection compared to 66% (530/807) in spot sputum collections.[115,116] Smear grade is indicated by the number of "+" signs.



FOOTNOTE: As only aggregate data for these studies were published, matched analysis was not possible. Therefore, the p value shown in the figure was calculated from a Wilcoxon-Mann-Whitney rank sum test.

Publication bias. Funnel tests did not show any asymmetry, and no small-study effects were demonstrated with Egger's test for microscopy (p=0.4) or culture (p=0.7, Figure 4.6).

Figure 4.6 Funnel plots
Funnel plots assessing publication bias in studies investigating the effects of sputum collection methods on the diagnostic yield of sputum microscopy (top) and culture (bottom).



Methodological quality. Potential bias was identified for 65% (15/23) of studies (Table 4.6). The main sources of bias in the studies were: participant selection different from target populations; and fixed order of collection techniques, which may have introduced a systematic error.

Table 4.6 Study quality as assessed by the QUADAS-2 checklist

First Author, Year	Could bias have been introduced to the study by the:			
	(a) selection of patients	(b) interpretation of the comparison sputum collection method	(c) interpretation of the control sputum collection method	(d) patient flow
Mpagama, 2012[114]	low	low	low	high
Andrews, 1959[115]	high	low	low	low
Majumdar, 1962[116]	high	unclear	unclear	unclear
Warren, 2000[117]	high	high	low	high
Abdel Aziz, 1985[118]	unclear	high	low	high
Krasnow, 1969[119]	low	low	low	unclear
Kestle, 1967[120]	unclear	low	low	low
Ssengooba, 2012[121]	high	unclear	unclear	low
Pande, 1974[122]	low	low	low	high
Geldenhuys, 2014[123]	high	low	low	high
Schoch, 2007[124]	low	low	low	high
Khan, 2007[125]	low	low	low	low
Alisjahbana, 2005[126]	low	unclear	unclear	low
Sakundarno,2009[127]	high	high	high	high
Mohamed, 2014[128]	low	high	high	low
Maciel, 2009[129]	high	low	low	high
Bell, 2009[100]	low	high	low	high
Souza, 2007[130]	low	low	low	high
Lee, 2013[136]	low	low	low	low
Mhalu, 2015[132]	low	high	unclear	low
Kalema, 2012[133]	low	low	low	low
Peres, 2011[134]	low	low	low	low
Davis, 2009[135]	low	low	low	low

Discussion

This review demonstrated that compared to standard 'spot' sputum collection, providing instructions or collecting a 'pooled' sputum sample increased the odds of TB diagnosis in sputum microscopy and culture by 1.6 to 1.8 times. It is unsurprising that diagnosing pulmonary TB depends on the quality of the sputum sample tested. Remarkably these simple, inexpensive strategies had comparable impact on diagnostic yield than did introducing the relatively expensive GeneXpert MTB/RIF test, which in the largest published studies increased the odds of diagnosing TB by only 1.3 to 1.5 times[137–139] (Table 4.7). Whilst accepting the usefulness of molecular tests in detecting drug resistance and in reducing variability of interpretation by laboratory technicians, there seems much more interest in new tests compared to strengthening existing methods. For example, during the 5 years before the publication of this review[22], a search of the PUBMED database suggests that for TB diagnosis, 79 studies have focused on the impact of GeneXpert MTB/RIF test, whereas only 7 studies have focused on non-invasive sputum collection techniques (Table 4.8). Although these are different aspects of the diagnostic pathway, and PUBMED-indexed publications do not summarise all TB-related publications, these comparisons suggest that greater attention and resources should be invested in how sputum is collected, matching the allure and marketing associated with new diagnostic commercial products.[4]

In 2007, the WHO-recommended case definition of sputum microscopy positivity changed to include cases with the presence of at least one acid-fast bacillus in at least one sputum sample.[140] Although most studies did not define the threshold of positivity used in sputum microscopy, we assumed that six studies carried out after 2007 would be affected by this change. The studies varied in the number of samples collected from each participant, so positivity in this analysis was, where possible, defined for each sample not individual case, and odds ratios were calculated instead of sensitivity and specificity. Furthermore, our meta-regression analyses demonstrated that this change in case definition of smear positivity did not contribute to the heterogeneity within the study effect sizes, nor the likelihood of a positive test in the reference groups. It should be noted that a positive sputum microscopy does not equate to a diagnosis of tuberculosis, as false positive arise because acid-fastness is not unique to *M. tuberculosis*. Thus culture, which is unaffected by this change in definition for positive sputum microscopy, is the gold standard for confirming TB and demonstrated similar results to microscopy in our meta-analyses.

Table 4.7 Diagnostic yield of GeneXpert versus sputum microscopy
A Table demonstrating the results from the largest, international studies comparing GeneXpert versus sputum microscopy, and a meta-analysis that calculated the incremental yield of GeneXpert compared to sputum microscopy. Note 95% CI= 95% confidence interval.

First author, year	Study site	Study type	Participant	Number of participants	Percentage GeneXpert positive (n)	% Sputum microscopy positive (n)	Odds ratio (95% CI)
Boehme, 2010[137]	International multi-site	within patient	Adults suspected of pulmonary TB	1730	39% (675)	33% (567)	1.3 (1.1 - 1.5)
Boehme, 2011[138]	International multi-site	within patient	Adults suspected of pulmonary TB	6069	15% (933)	12% (699)	1.4 (1.3 - 1.6)
Steingart, 2014[139]	Hypothetical 2.5% prevalence	meta-analysis	Adults suspected of pulmonary TB	1000	2% (22)	2% (16)	1.4 (0.72 - 2.7)
Steingart, 2014[139]	Hypothetical 5.0% prevalence	meta-analysis	Adults suspected of pulmonary TB	1000	4% (44)	3% (33)	1.4 (0.85 - 2.1)
Steingart, 2014[139]	Hypothetical 10% prevalence	meta-analysis	Adults suspected of pulmonary TB	1000	9% (88)	7% (65)	1.4 (1.0 - 1.9)
Steingart, 2014[139]	Hypothetical 30% prevalence	meta-analysis	Adults suspected of pulmonary TB	1000	26% (264)	20% (195)	1.5 (1.2 - 1.8)

Table 4.8 Comparison of the number of publications in 2012-2017
Table demonstrating the number of peer-reviewed publications involving the impact of non-invasive sputum collection methods or GeneXpert on pulmonary TB diagnosis in the 5 years prior to publication of this systematic review in June 2017.

Theme associated with the diagnosis of pulmonary TB	PubMed search terms	Number of PubMed results in last 5 years	Number of publications relevant to theme
Non-invasive sputum collection methods	((pulmonary tuberculosis [Title/Abstract] AND diagnosis [Title/Abstract]) AND sputum [Title/Abstract]) AND (clearance [Title/Abstract] OR collection [Title/Abstract] OR submission [Title/Abstract] OR acquisition [Title/Abstract]))	23	7
GeneXpert	(pulmonary tuberculosis [Title/Abstract] AND diagnosis [Title/Abstract]) AND (Xpert [Title/Abstract] OR GeneXpert [Title/Abstract])	100	79

Two sputum samples are sufficient for TB testing, because approximately 97% of sputum microscopy-positive TB cases are identified by the first 2 samples provided.[96] A meta-analysis concluded that these 2 samples could be collected during the same-day with no decrease in diagnoses.[2] Globally, 16% (95% CI 13-18%) of individuals suspected of having TB are estimated to be lost to follow-up before starting treatment, so ideally a spot diagnostic specimen should be collected from patients during the same-day as their first assessment.[21] Providing instruction before or during a spot sputum collection considerably increased TB diagnoses. Specifically, the evidence in this review demonstrates that explaining (verbally or visually) the difference between saliva and sputum, providing guidance on how to produce sputum, and/or advising breathing exercises or physiotherapy assistance almost doubled the odds of diagnosing TB. Although the instructions in these studies included common elements, as demonstrated in Table 4.3, further research is needed to determine what instructions optimally increase diagnostic yield.

Same-day, single-visit TB diagnosis is not always possible for reasons including: sputum could not be produced, perceptions of superior early morning collection or due to health system constraints delaying testing and results dissemination.[5] Therefore people being tested for suspected TB are usually asked to return for follow-up appointments before TB diagnosis can be achieved. Under these circumstances, our analysis demonstrates that TB diagnoses could be increased if patients were asked to pool sputum spontaneously expectorated over several hours. Requiring another visit does have costs, risks loss to follow-up, and may not increase the number of people on treatment as many have their treatment started empirically.[4,141] Yet, enabling laboratory confirmation of TB with drug susceptibility results is cost-effective due to the avoidance of inappropriate therapy, associated catastrophic costs, and generation of drug resistance, although this needs to be specifically studied.[56,141,142] Pooled sputum collection increased not only the proportion of positive results, but also how strongly positive microscopy results were, most probably due to increasing the likelihood of collecting sputum from the lower respiratory tract, relative to contents of the upper airways and saliva. We have commenced research to test whether the increased TB diagnoses associated with pooled sputum over several hours are additive to the advantages provided by instructing spot sputum collection. This is particularly important because the comparisons between pooled and spot samples were most influenced by the one study that requested sputum or saliva, which may have influenced the study effect size.[115] Pooling sputum collection need not increase the risk of TB transmission if patients are advised to complete sputum collection in well-ventilated open-air areas, away from staff and other people.[143,144]

Patients being tested for TB are often asked to complete early-morning sputum collection, consistent with some international recommendations.[36] Early morning sputum collections

have been thought to be superior to standard spot sputum collection, perhaps because of the diurnal variation in sputum volume and symptoms in other inflammatory lung diseases, together with the assumption that poor lung clearance whilst sleeping leads to increased expectoration of sputum on waking.[145–147] However, our analyses demonstrate that early morning sputum collection had no significant effect on diagnostic yield, in keeping with findings that 2 spot samples had equivalent diagnostic accuracy as spot plus early morning samples.[2] Thus, sputum collection should be instructed spot or pooled collections, that may be done at any time.

Strengths of this systematic review and meta-analysis include: broad inclusion criteria (reducing selection bias and capturing more research); the absence of major small-study bias (suggested by funnel-plot analysis); pairwise meta-analysis of subgroups with a random-effects model (to control for heterogeneity); and large sample size (most studies were adequately powered and meta-analysis included >19,000 samples). Limitations include: the paucity of good-quality studies, especially those involving culture or PCR; that some studies evaluated combined interventions and the cumulative effect on diagnostic yield could not be measured; and the potential fragility of network meta-analysis.[110,148] However, these limitations were ameliorated by the broadly consistent findings of: pairwise meta-analyses and network meta-analysis; research assessing both microscopy and culture positivity. Furthermore, to evaluate the true impact of these collection methods in the real-world, routine programmatic evaluations are required.

The key implications of this systematic review, pairwise and network meta-analyses are that (1) pooled sputum collection or providing instructions prior to a spot collection equally increases the chances of a positive TB diagnostic test compared to the global standard uninstructed spot; and (2) that sputum may be collected at any time, not necessarily in the early morning. These inexpensive, simple interventions warrant further research and policy emphasis because their effects are potentially comparable to much more expensive interventions such as investing in new diagnostic products.

RESEARCH CHAPTER 5

Refinement - TB viability microscopy optimisation

Background

The results of the foundation study presented in Research Chapters 2 and 3 demonstrated that the concentration of fluorescing bacilli visualised with FDA microscopy in pre-treatment samples may identify the most infectious patients; and how these concentrations changed during the first days of anti-TB therapy predicted treatment response.[20,85] However, FDA microscopy uses fluorescence that until recently required expensive mercury or halogen light sources (as for the foundation study), limiting feasibility for routine use in clinical settings. Affordable light-emitting diode (LED) microscopes are now widely available and for sputum microscopy the WHO encourages their use with auramine staining in place of conventional light microscopy with ZN staining.[149] Consequently, LED fluorescence microscopes have widespread and increasing availability, which may allow FDA microscopy to have greater clinical applicability. It therefore seemed desirable to validate FDA microscopy for use with LED microscopes.

In addition to optimising the type of microscope used for FDA microscopy, there is a need to optimise the protocol used for FDA staining prior to microscopy. Review of the published protocols for FDA microscopy for TB identifies diverse methods used, as shown in Table 5.1.[14,15,19,20] There is no published evidence to guide which of these protocols should be used. Variability in methods may lead to misleading interpretations due to false-positive or false-negative results.[90] Additionally, some protocols involve sputum processing with centrifuge decontamination, which is a barrier to implementation because centrifugation may be biohazardous and centrifuges with sealed rotors suitable for use for TB diagnosis are expensive and have limited availability.

The aim of the following experiments was to: clarify the optimum FDA microscopy protocol that is simple and safe; refine the staining method; assess whether storage conditions of sputum or FDA working solutions affect results; and determine the cost of this protocol. Consequently, a SOP for FDA microscopy was developed, which can simply and safely provide reproducible results in resource constrained clinical settings.

Table 5.1 Published FDA protocols
A comparison of fluorescein diacetate (FDA) protocols published in English for staining *Mycobacterium tuberculosis*. Note: NS=not specified; °C=degrees centigrade; mins=minutes; CPC=cetylpyridinium chloride; NAOH-NALC=sodium hydroxide with N-acetyl cysteine and "-"=not done.

	Jarnigin [14]	**Datta** [20]	**Salim** [18]	**Schramm** [90]	**Van Deun** [19]
Stock solution					
Dissolvent	acetone	acetone	NS	acetone	NS
FDA concentration (mg/ml)	5	5	5	25	0.5
Storage temperature (°C)	4	-20	-20	-20	-20
Maximum storage (days)	56	730	730	730	730
Working solution					
Dissolvent	Dubos albumin broth	40% acetone	acetone	acetone	acetone
FDA concentration (mg/ml)	0.5	0.02	0.05	0.5	0.02
Storage temperature (°C)	NS	4	-20	NS	-20
Surfactant added for storage	Tween 80	-	Tween 80	Tween 80	Tween 80
Maximum storage (days)	NS	1	7	NS	7
Sputum sample processing					
Preservative for transport	used	-	CPC	-	-
Decontamination	the culture	4% NAOH-NALC	-	4% NAOH-NALC	-
Centrifugation	suspension	3000 x g	-	NS	-
Staining					
Staining area (cm^2)	1	1	NS	1	NS
Sample volume (drops)	1	3	NS	3	NS

Slide fixation	flame	serum albumin and flame	-	-	-
FDA application	filter paper	filter paper	NS	filter paper	NS
FDA incubation time (mins)	30	20	30	30	30
Acid-alcohol step	-	-	1% for 1-2 mins	1% for 3 mins	0.5% for 3 mins
Sterilisation step	-	-	5% phenol for 10 mins	5% phenol for 10 mins	5% phenol for 10 mins
Quenching step	-	-	-	-	0.5% KMnO4 for 1 min
Coverslip	yes, with glycerol	-	-	-	-
Reading					
Microscope description	BD-12 primary filter	Nikon Eclipse E600 mercury	Mercury vapor system	Olympus CX21 LED	FluoroLED LED 450nm
Magnification (times)	450	1000 (oil)	1000 (oil)	1000 (oil)	200

Methods

Ethics approvals included Imperial College London and the Peruvian Ministry of Health DIRESA Callao. This research was done with the collaboration of the Peruvian national TB program.

Setting. Patient sputum samples were collected from all consecutive unselected consenting patients diagnosed with pulmonary TB in community health centres in the peri-urban shantytowns of Callao, Peru. The study involved 3 phases:

1. optimisation study;
2. validation study; and
3. storage study.

In both the optimisation study and the storage study, control samples were collected from asymptomatic, healthy individuals.

Inclusion and exclusion criteria are detailed below for each experiment.

Samples. Sputum samples were processed even if they appeared to be salivary. All sputum samples were collected and transported to the local research laboratory at ambient temperature and then stored at 4°C and processed within 24-hours of arrival, usually within 72 hours of expectoration.

Slides. All slides were prepared by smearing sputum over standard glass microscope slides that had been cleaned with 95% alcohol.

Measurements. To increase operational relevance, we measured liquids (including sputum) as drops from disposable transfer pipettes, which we found to have an average volume of 40 µl per drop. Microscopy using a 100x objective with oil immersion (i.e. 1,000x overall magnification) to examine 100 high power fields assesses an area of 2 mm^2.[58] For clarity, we therefore report the volume of sputum smeared per mm^2 of the area of the glass microscopy slide.

Sputum smears. 'Thin smears' were prepared at a typical density used for conventional sputum microscopy by smearing 1 drop over a 2 cm^2 area i.e. 0.2 $\mu l/mm^2$. 'Thick smears' were made with a higher density than is usually used for sputum microscopy by smearing 3 drops over an area of 1 cm^2 i.e. 1.2 $\mu l/mm^2$, as previously used for FDA microscopy, see Table 5.1.[20,85,90]

Fixation. All sputum smears were then heat fixed to the slide by passing each slide through a flame 3 times.

Microscopy. All stained slides were dried, protected from light and for fluorescence microscopy were read with Zeiss Primo Star iLED (Heidenheim, Germany) microscopes using the 100x objective with oil immersion without using microscopy cover slips.

Blinding. All slides prepared in the same way were 'shuffled' before staining to ensure that the order in which slides were prepared neither influenced the order in which they were processed, nor the protocol used. Microscopy was performed by multiple laboratory biologists and technicians who were unaware of both the clinical status of the patient and the results of other tests.

Optimisation study.

Inclusion criteria. This initial optimisation study used sputum samples that were either from: selected patients who had already been determined to have a positive acid-fast microscopy result but had not yet received any TB therapy; or healthy negative control participants with neither symptoms nor suspicion of TB disease, during 6 May until 20 July 2015.

Exclusion criteria were lack of informed written consent or inability to produce a sample.

Sputum processing. Reagents were obtained from Thermo Fisher Scientific (MA, USA) except where otherwise stated. Sputum samples were divided into aliquots that were: processed with centrifuge decontamination; or left unprocessed at room temperature to be directly smeared onto slides.

Centrifuge decontamination was done as described.[37] Briefly, the 2 ml sputum aliquot for decontamination was briefly vortexed with an equal volume of 2% sodium hydroxide containing 2.9% sodium citrate and 0.5% N-acetyl-L-cysteine. After 20 minutes at room temperature, excess PBS (pH 6.8, Sigma-Aldrich, MO, USA) was added, and the mixture was centrifuged for 15 minutes at 3,000 gravities in a Thermo Fisher Scientific centrifuge that had sealed rotors to increase biosafety for the laboratory personnel.[37] The pellet was re-suspended in PBS to a final volume of 2 ml. Both decontaminated and unprocessed aliquots were then used for slide preparation and culture inoculation.

Solutions. A stock solution of 5 mg/ml FDA (Sigma-Aldrich, MO, USA) in acetone was stored at -20°C. A fresh working solution of 20 µg/ml FDA was prepared daily by dissolving

10 µl of stock solution in 2.5 ml of 40% acetone in PBS. Standard acid-alcohol (AA) solution used in TB fluorescent microscopy was prepared by adding 0.5% hydrochloric acid to 96% ethanol.[36] Phenol solution was prepared by adding 5% phenol (Merck, NJ, USA) to distilled water. Potassium permanganate solution was prepared by dissolving 0.5% $KMnO_4$ (Merck, NJ, USA) in distilled water.[36]

Smears. A thick smear was prepared from the centrifuge-decontaminated aliquot and 2 thick smears and 5 thin smears were prepared from the direct aliquot. Aliquots of each sample were processed concurrently with all of the following FDA staining protocols, as described below, in Box 5.1 and in Table 5.2.

FDA staining protocol A (Centrifuge-thick) used thick smears from the centrifuge-decontaminated aliquot. FDA staining was done by covering the smear with a 1 cm^2 square of Whatman grade 3 filter paper that was soaked with FDA working solution and incubated at 37ºC for 20 minutes, after which the filter paper was discarded, excess FDA was tapped off the slide, and left to dry prior to microscopy. As shown in Table 5.1, we had used and published this protocol previously.[20]

FDA staining protocol B (Direct-thick) assessed modifying protocol (A) only by using a thick smear of unprocessed instead of centrifuge-decontaminated sputum.

FDA staining protocol C (Direct-thick-AA) was the same as protocol (B) except that after FDA staining, before being left to dry, an AA decolourisation step was added. For this, AA solution was flooded over the entire slide for 2 minutes and then rinsed with distilled water.

The remaining protocols used thin smears. For thin smears, all FDA staining were done by flooding FDA working solution onto the slide (without filter paper), incubating at 37ºC for 30 minutes and then rinsing with distilled water.

FDA staining protocol D (Direct-thin-AA) was stained with FDA followed by an AA step.

FDA staining protocol E (Direct-thin-AA-phenol) was same as protocol (D) but after AA for 3 minutes, the slide was flooded with phenol solution for 10 minutes and then rinsed with distilled water.

Box 5.1. Protocols for fluorescein diacetate (FDA) microscopy in the optimisation experiment.

FDA staining protocol A (Centrifuge-thick) was processed as published [20], by applying 3 drops of centrifuge-decontaminated sputum to a slide to make a smear of approximately 1 cm^2.

FDA staining protocol B (Direct-thick). Using cleaned slides, 3 drops of unprocessed sputum were smeared over an area of approximately 1 cm^2. The slides were then dried and protected from ultraviolet light. When slides were dry, they were passed over a flame 3 times. A 1 cm^2 square of Whatman grade 3 filter paper was placed on top of the smear, and 13-15 drops of freshly prepared FDA working solution at a concentration of 20 µg/ml FDA was applied to cover the filter paper. Slides were then incubated for 30 minutes at 37°C. After removing the slides from the incubator, excess liquid was tapped off.

FDA staining protocol C (Direct-thick-AA). Same as protocol B, except with the following added steps: Slides were then rinsed with distilled water; afterwards 0.5% acid-alcohol (AA) was flooded onto the slides and left for 3 minutes; slides were then rinsed with distilled water.

FDA staining protocol D (Direct-thin-AA). Using cleaned slides, 1 drop of unprocessed sputum was smeared over an area of approximately 2 cm^2. The slides were then dried and protected from ultraviolet light. When slides were dry, they were passed over a flame 3 times. 13-15 drops of freshly prepared FDA working solution at a concentration of 20 µg/ml FDA was applied to the sample to cover the smear. Slides were then incubated for 30 minutes at 37°C. After removing the slides from the incubator, excess liquid was tapped off. Slides were then rinsed with distilled water. Afterwards 0.5% AA was flooded onto the slides and left for 3 minutes. Slides were then rinsed with distilled water.

FDA staining protocol E (Direct-thin-AA-phenol) Same as protocol D, except with the following steps: phenol at 5% concentration was applied to slides and left for 10 minutes; slides were then rinsed with distilled water.

FDA staining protocol F (Direct-thin-AA-phenol-$KMnO_4$30s). Same as protocol E, except with the following added steps**;** Potassium permanganate at a concentration of 0.5% was applied to slides and left for 30 seconds; slides were then rinsed with distilled water.

FDA staining protocol G (Direct-thin-AA-phenol-$KMnO_4$60s). Same as protocol E, except with the following added steps**;** potassium permanganate at a concentration of 0.5% was applied to slides and left for 60 seconds; slides were then rinsed with distilled water.

Reading slides

After staining, all the slides were left to dry in the dark.
Slides were then read within 4 hours of staining
Using the fluorescent light source and 100x objective with oil immersion on the Zeiss Primo Star iLED microscope (Heidenheim, Germany), the number of bacilli visible in 100 fields was recorded.

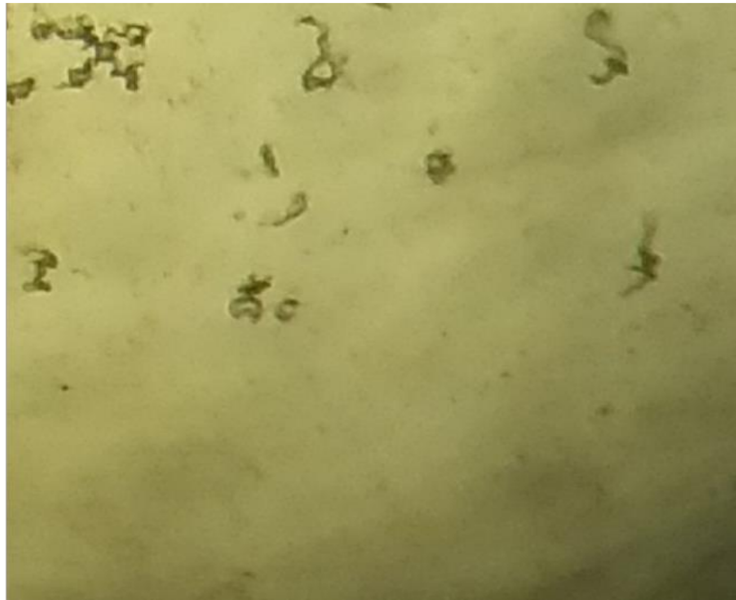
FDA staining protocol F (Direct-thin-AA-phenol-$KMnO_4$30s) was the same as protocol (E), followed by applying potassium permanganate solution for 30 seconds and then rinsing with distilled water.

FDA staining protocol G (Direct-thin-AA-phenol-$KMnO_4$60s) was the same as protocol (F), except that the potassium permanganate solution was applied for 60 instead of 30 seconds. As shown in Table 5.1, this protocol has been evaluated previously [19].

Conventional acid-fast staining*.* Smears were flooded with 0.1% auramine for 15 minutes, decolourised with AA for 2 minutes, rinsed with distilled water, flooded with 0.5% potassium permanganate for 30 seconds and then rinsed with distilled water.

Culture*.* Decontaminated and direct aliquots of sputum samples were inoculated for quantitative culture results using Middlebrook 7H9 culture broth supplemented with glycerol, casitone and the standard oleic acid, albumin, dextrose and catalase (OADC) growth supplement according to the manufacturer's instructions. To reduce the risk of bacterial or fungal overgrowth contamination, the culture medium was additionally supplemented with Selectatab (Mast Group, Bootle, UK) according to the manufacturer's instructions, plus 0.25% carbendazim as described.[26,27,150] The cultures were performed to provide quantitative results in 24-well culture plates, as described,[33] preparing a 1:10 dilution by adding 1 drop (approximately 40 µl) of either the unprocessed or decontaminated sputum to a well containing 9 drops (final volume approximately 400 µl) of supplemented Middlebrook 7H9 culture broth. Then 1 drop of this suspension was mixed into another well containing 9 drops of supplemented Middlebrook 7H9 culture broth making a 1:100 dilution and repeated to make 1:1,000, 1:10,000 and 1:100,000 dilutions. Cultures were incubated in un-supplemented air at 37ºC. The cultures were sealed in a Ziploc bag, and examined 3-times per week for 6 weeks with an inverted microscope using a 4x objective, final magnification 40x, to detect growth of *M. tuberculosis*. Colonies were counted at day 42 of growth, with a colony being defined as a single cell or a clump of cells with the characteristic cording pattern of *M. tuberculosis,* as is shown in Figure 5.1*,* which is easily distinguishable from other bacteria and filamentous hyphae of fungi. All cultures were done in duplicate.

Figure 5.1 Cording colonies
Photograph demonstrating the characteristic cording colony of *Mycobacterium tuberculosis.* There are 11 colonies in this image.



Validation study.

Inclusion criteria. FDA microscopy was introduced into routine laboratory work from 18 November 2015 until January 2016 and applied to all sputum samples from consecutive unselected patients, whether they were about to commence or were already receiving TB treatment, without knowledge of their acid-fast microscopy results.

Exclusion criteria were lack of informed written consent or inability to produce a sample.

Slide preparation. Three thin smears were prepared from each unprocessed sputum sample.

FDA staining. The 2 FDA protocols described above that produced the best results in the optimisation study were re-evaluated in the validation study.

Conventional acid-fast staining*.* For operational reasons, the third thin smear was processed for acid-fast microscopy, as described [36,151], the results of which are not reported here.

Storage study.

Inclusion criteria*.* Sputum was collected from a randomly selected patient who was already known to their health centre to have strongly sputum smear-positive TB and who

had not yet commenced TB treatment, in May 2016. On the same day, a negative control sputum sample was collected from a healthy participant with neither symptoms nor suspicion of TB disease.

Exclusion criteria were lack of informed written consent or inability to produce a sample.

Storage conditions. All sputa and slides were stored protected from light for up to 4 weeks. Direct thin smears were made from each fresh sample and these slides stored at room temperature. The remaining volume of sputa were divided into equal aliquots that were stored at 4°C and at room temperature.

FDA staining. Twice a week, direct thin smears were prepared from the patient and the control sputum that were stored at 4°C and at room temperature. Together with patient and control stored slides, these were stained using the optimum FDA protocol identified in the previous studies described above. All of these procedures were performed in duplicate: 1 slide from each pair of slides was stained using FDA working solution that was prepared daily as described above; the other duplicate from each pair of slides was stained using FDA working solution that had been prepared on the first day of the experiment and stored at room temperature, protected from direct light.

Cost analysis.

All materials and reagents were procured locally in Peru, except for FDA that was obtained in the United Kingdom. An inventory of all reagents was kept, and the volume of reagent used for FDA microscopy recorded. With this data, the cost of FDA microscopy using the optimum staining protocol was calculated. Labour costs and equipment such as microscope, glassware, Bunsen burners, drying racks and distilled water were not included in the cost analysis because these were sunk costs i.e. would have been purchased by a laboratory that was already performing TB fluorescent microscopy with auramine staining, as recommended by the WHO.[149]

Analysis.

Bacterial counts. The number of stained bacteria visible in 100 consecutive microscopy fields was counted. In culture the number of CFU in each serial dilution was recorded. When required, these counts were transformed to the concentration per ml of sample, which was calculated from the volume of sample smeared on the slide or inoculated and diluted for culture.[33] Concentration calculations were used when comparing microscopy results with CFU counts in culture, or microscopy protocols that used different volumes of sputa. When there were duplicate results, geometric means were calculated. Positive

microscopy was considered if there was ≥ 1 bacillus seen per 100 high powered fields in microscopy due to changes in global policy.[140] and a positive culture defined as > 1 CFU per well, according to local practice and the MODS protocol.[87]

Quality assessment. During each reading a subjective assessment of quality was made based on the following criteria: background contrast; how easy it was to focus the slide; how bright the bacilli were; and how easy it was to identify bacilli. A score was allocated to each answer, the best being 100 and the worst 0. A total quality score was then calculated as the sum of the individual scores for background, brightness, identification and focus. Therefore, the best score that could be achieved was 400, and the worst was 0.

Statistics. As all counts and concentrations were exponentially distributed, results were transformed to their base-10 log for analysis. Because the log of zero values cannot be calculated, before analysis zero values were transformed to the midpoint between zero and the detection threshold. Tests were 2-tailed and were performed with a 95% CI. Data with normal distributions were summarised as means with the SD and non-parametric data were summarised as median and IQR. When paired data were used to compare results for the same sample processed by different protocols, the paired Student's t-test was used when the data were normally distributed, and the Wilcoxon signed rank test was used for non-parametric data. Repeated measures were considered for all regression analyses that were used to assess factors that impacted microscopy results and a random effects term was used to adjust for inter-sample variation. When assessing agreement and repeatability between protocols, the limits of agreement method was used, as described by Bland and Altmann.[152]

Results

Optimisation study

Quantitative assessment. There were 9 acid-fast microscopy positive patients and 2 healthy controls who provided samples that were used to make 80 slides and 264 culture wells in this study. All samples from the healthy controls were culture and acid-fast microscopy negative. No cultures failed due to bacterial or fungal overgrowth. Cultures from centrifuge-decontaminated sputa had 5.6% (1/18) false-negative results, whereas directly tested sputa had no false-negative results.

Figure 5.2 shows the concentration of CFU/ml in unprocessed sputum culture versus CFU/ml in corresponding decontaminated sputum culture, and bacilli/ml in acid-fast microscopy and FDA microscopy protocols (A-G). Compared to direct sputum culture CFU/ml results, the centrifuge-decontaminated sputum samples had median 10-times less CFU/ml (IQR=1.8-40 times, p=0.01). Compared to direct sputum culture CFU/ml results, conventional acid-fast microscopy had a median 12-times (IQR 3.6-63 times) more bacilli/ml (p=0.01).

In patient samples, 13 slides had false-negative results, of which 11 were stained with protocol A-C. Bacilli/ml results in protocols A-C were lower than other protocols (Figure 5.2) and were also significantly less than the CFU/ml in the corresponding direct sputum culture (all $p \leq 0.05$, Figure 5.3b). However, protocols D-G had higher bacilli/ml results that had better agreement with the direct sputum culture results (Figure 5.3b). Specifically, 46% (13/28, 95% CI=28-66%) of bacilli/ml results in FDA protocols D-G were within +/- 1 logarithm of the corresponding CFU/ml in unprocessed sputa, and 89% (25/28, 95% CI=72-98%) within +/- 2 logarithms.

Linear regression of the bacilli/ml results for patient samples demonstrated that the factors that predicted higher bacilli/ml in FDA microscopy were: if the protocol used thin (1 drop) versus thick (3 drops) smears ($p<0.001$); and if the slides were allocated a higher score in the quality assessment ($p<0.0001$). These 2 variables explained 70% of the within sample variability. The phenol step, which was added for biosafety, did not impact the quantitative results (p=0.5).

Slide quality assessment. Several smears prepared from centrifuge-decontaminated sputum in protocol A were inadvertently partially removed from the slide when the filter paper used during staining was removed. For the protocols using thick smears from direct sputum, identifying bacilli was difficult because they were often obscured by cells and other sputum contents. These factors appeared to contribute to FDA protocols A-C having

significantly lower positivity rates in patient samples and lower concentrations of visualised bacilli/ml compared to the CFU/ml in unprocessed sputum culture (Figure 5.2 and 5.3, all $p \leq 0.05$).

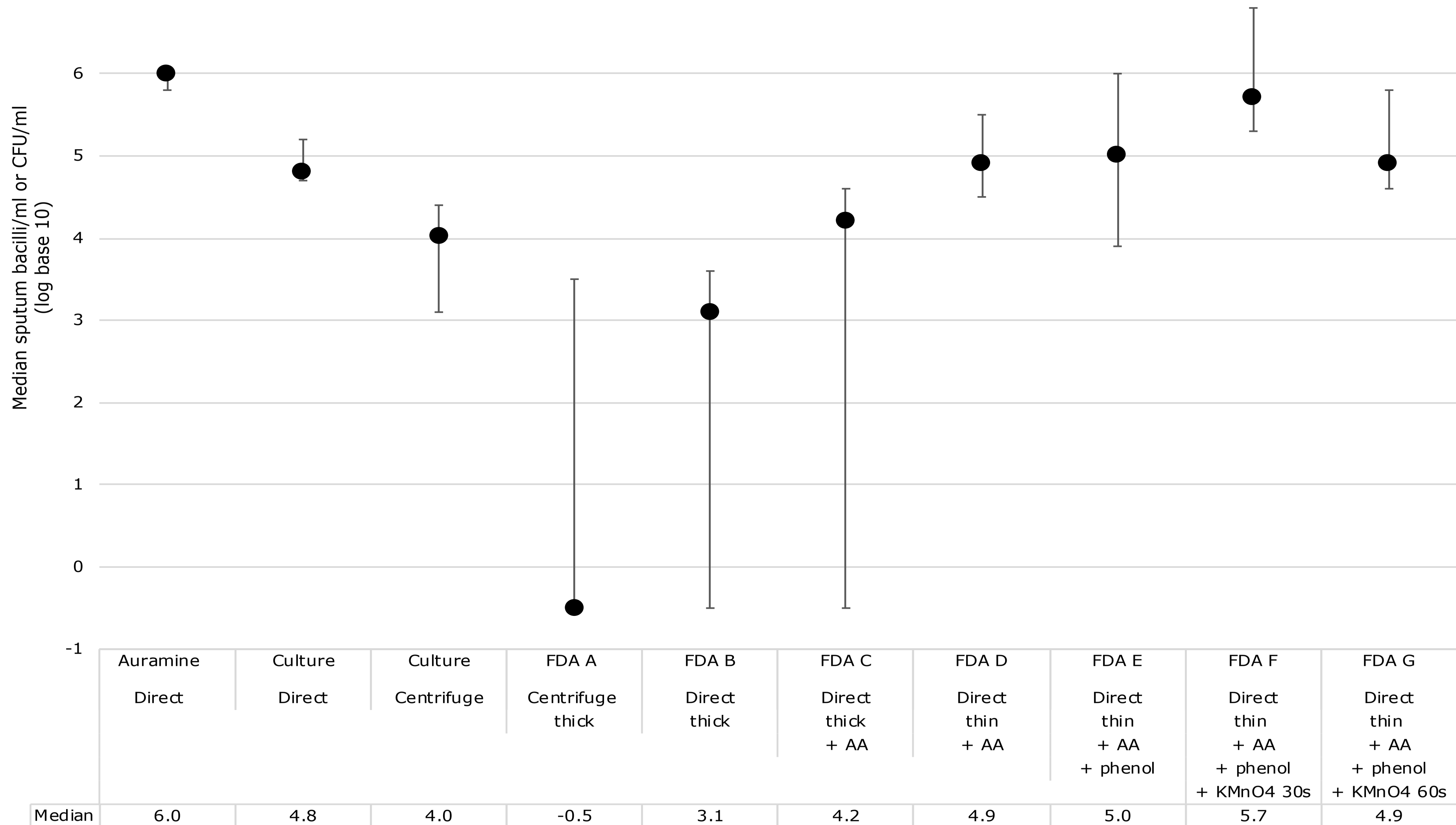
Protocols D, E and G had similar counts (Figure 5.2), but they differed in quality and therefore ease in reading slides (Table 5.2 and Figure 5.4). Compared to FDA protocol D, protocol G (which applied potassium permanganate for 60 seconds) had improved background contrast ($p=0.005$) and bacilli were easily identifiable. However, slides stained with Protocol G had focusing difficulties and bacilli were less bright ($p=0.05$). Therefore, FDA protocol F was later introduced (with 30 seconds of potassium permanganate), which tended to have better background contrast than protocol D ($p=0.08$), without any reported focusing difficulties. As shown in Table 5.2, FDA protocol F and G produced the best quality slides.

Table 5.2 Optimisation study protocols and results
A table demonstrating the different protocols of fluorescein diacetate (FDA) microscopy, the number samples, and quality assessment. Please see methods and Box 4.1 for full explanation of the protocols. Note. NaOH=sodium hydroxide, AA=acid-alcohol, $KMnO_4$=potassium permanganate, sec=seconds, min=minutes and IQR=interquartile range.

	FDA PROTOCOLS	**A** Centrifuge-thick **Datta**[20]	**B** Direct-thick	**C** Direct-thick-AA	**D** Direct-thin-AA	**E** Direct-thin-AA-phenol	**F** Direct-thin-AA-phenol-$KMnO_4$30s	**G** Direct-thin-AA-phenol-$KMnO_4$60s **VanDeun**[19]
Staining method	Decontamination	2% NaOH	no	no	no	no	no	no
	Drops of Sputum (μl)	3 (120)	3 (120)	3 (120)	1 (40)	1 (40)	1 (40)	1 (40)
	20 μg/ml FDA incubation (min)	20	20	20	20	30	30	30
	0.5% Acid-alcohol (min)	-	-	2	2	3	3	3
	5% phenol (min)	-	-	-	-	10	10	10
	KMnO4 (sec)	-	-	-	-	-	30	60
General results	Number of samples, including 2 negative controls for each technique	11	11	11	9	11	5	11
Quality markers	Background score median (IQR)*	67 (33-100)	24 (0-33)	67 (33-67)	33 (33-67)	67 (33-67)	67 (67-67)	100 (100-100)
	Bacillary brightness score median (IQR)*	40 (0-60)	60 (0-80)	60 (0-80)	60 (40-80)	80 (40-100)	60 (20-60)	40 (40-60)
	Bacillary identification score median (IQR)*	33 (0-33)	33 (0-33)	33 (0-33)	33 (33-100)	66 (33-100)	100 (33-100)	100 (67-100)
	Easy to focus, % (n)	73% (8)	92% (10)	64% (7)	89% (8)	82% (9)	80% (4)	56% (6)
	Total quality score median (IQR)**	200 (200-240)	206 (100-247)	213 (133 - 240)	240 (206-280)	260 (240-333)	327 (253-326)	273 (240-306)

Footnote. * the score was made from Likert-type scales and transformed to a score out of 100, where 100 was the best and 0 was the worst.
** the total quality score was a sum of the score for background, brightness, identification and focus, therefore the maximum and best score that could be achieved was 400.The median here refers only to the patient samples.

Figure 5.2 Optimisation study distribution of quantitative results
This graph demonstrates the median bacilli or colony forming units (CFU) concentration per ml of sputum in each culture and microscopy technique. Error bars indicate the interquartile range. All microscopy and quantitative culture concentration data were transformed to base-10 logarithmic (log) values. Note. FDA=fluorescein diacetate, AA=acid-alcohol, KMnO4=potassium permanganate, s=seconds.



	Auramine Direct	Culture Direct	Culture Centrifuge	FDA A Centrifuge thick	FDA B Direct thick	FDA C Direct thick + AA	FDA D Direct thin + AA	FDA E Direct thin + AA + phenol	FDA F Direct thin + AA + phenol + KMnO4 30s	FDA G Direct thin + AA + phenol + KMnO4 60s
Median	6.0	4.8	4.0	-0.5	3.1	4.2	4.9	5.0	5.7	4.9

Figure 5.3 Optimisation study comparison with conventional sputum microscopy and culture

Bar graphs comparing the quantitative sensitivity of fluorescein diacetate (FDA) microscopy protocols A-G: (A) to identify acid-fast bacilli (AFB), shown as the difference in logarithmic base 10 counts between FDA bacilli/ml of sample versus AFB bacilli/ml in conventional microscopy with auramine staining (B) to identify colony-forming units (CFU), shown as the difference in logarithmic base 10 counts between FDA bacilli/ml of sample versus CFU/ml in quantitative culture. Error bars indicate 95% confidence intervals.

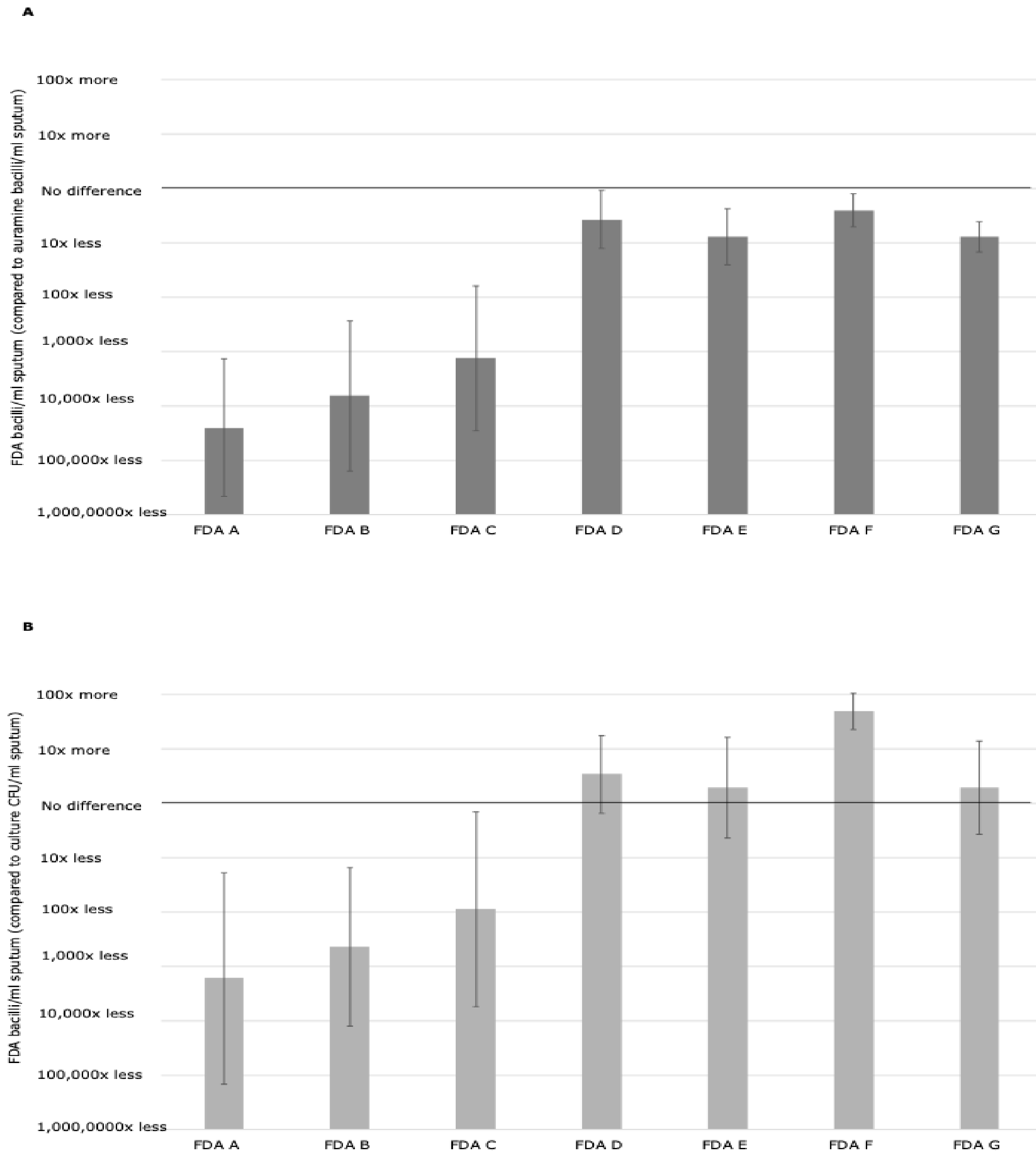
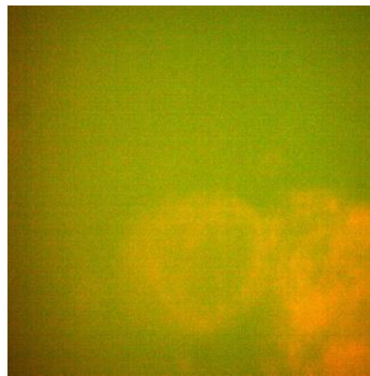
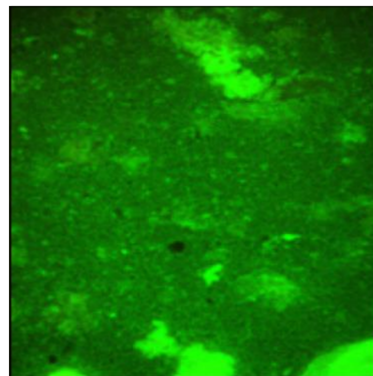


Figure 5.4 Photographs of each FDA microscopy protocol
Photographs of the same sample after staining with different FDA microscopy protocols.

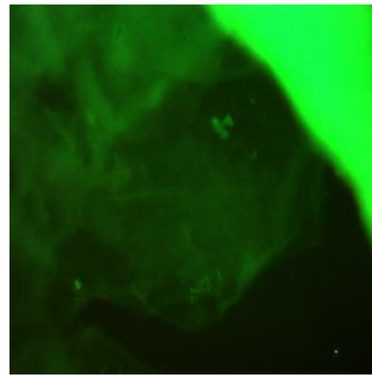
FDA A



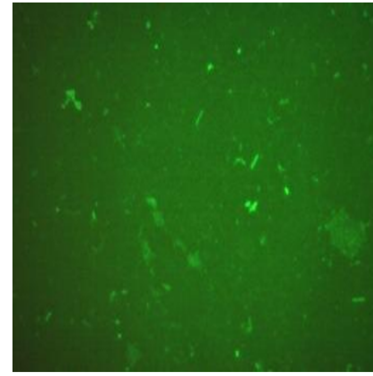
FDA B



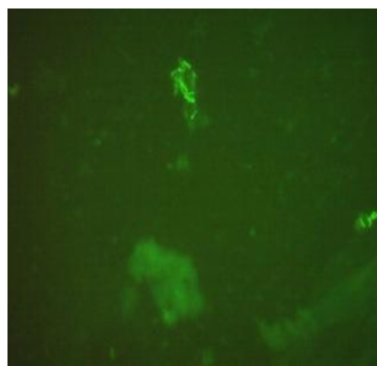
FDA C



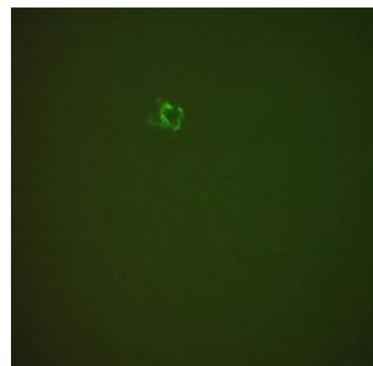
FDA D



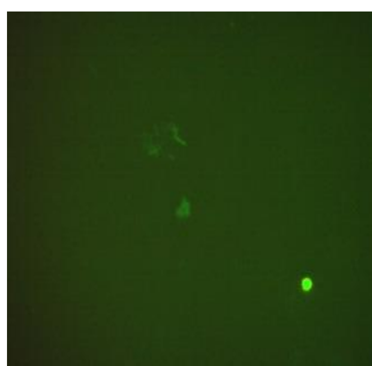
FDA E



FDA F



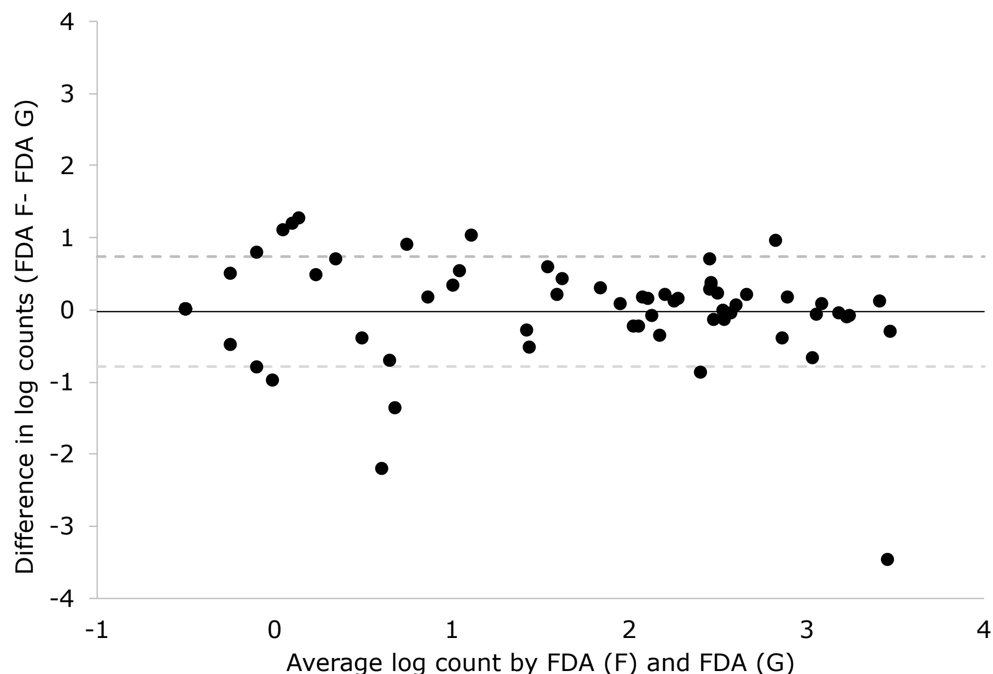
FDA G



Validation study

Quantitative assessment. 206 fresh sputa were collected from 200 patients, to prepare 412 slides comparing FDA microscopy protocols F versus G. Sputum was provided by patients with a median age of 28 years (range=5-82, IQR=21-44) and 61% (122/200) were male. Sputum was collected after treatment initiation in 35% (72/206), and 50% (102/206) had a positive conventional acid-fast smear-microscopy result. There were 122 positive FDA microscopy slides, and the number of bacilli visualised by microscopy were similar between the 2 protocols (p=0.4). Figure 5.5 demonstrates the high level of agreement between counts in FDA protocol F versus G, with 97% (199/206, 95% CI= 93-99%) of results differing by less than +/- 1 logarithm. Figure 5.5 shows that the agreement between counts tended to be higher in samples with higher bacillary load.

Figure 5.5 Validation study
A Bland-Altman plot demonstrating agreement between the bacilli count per 100 high powered fields in fluorescein diacetate (FDA) microscopy protocol F versus FDA microscopy protocol G. The difference between these 2 protocols is that the latter involves potassium permanganate quenching for 30 seconds longer. All count data were transformed to their logarithmic base-10 (log) value. The mean difference is -0.007 log (solid line), and the limits of agreement (dotted lines) are -0.70 to 0.71 log.



Slide quality assessment. Protocol F and G produced similar quality slides (p=0.3) with median quality score 254 (IQR=208-292). The time required to count the number of bacilli in 100 high powered microscopy fields was similar for protocol F versus protocol G (p=0.2) and overall was median 10 minutes (IQR=9-13). Regression analysis in Table 5.3 demonstrates that in FDA microscopy-positive samples (N=122), sputum that did not

contain blood had higher quality FDA microscopy slides (odds ratio 6.7, 95% CI=1.3-34, p=0.02). However, there were only 6 blood-stained sputum samples in this study.

Table 5.3 Validation study
Table demonstrating the factors that improved the quality of slides if fluorescein diacetate (FDA) microscopy was positive (n=122). The quality score was transformed to a binary variable, above and below the median score, and logistic regression with random effects was used to adjust for inter-sample variation showing the odds ratio for a higher quality slide.

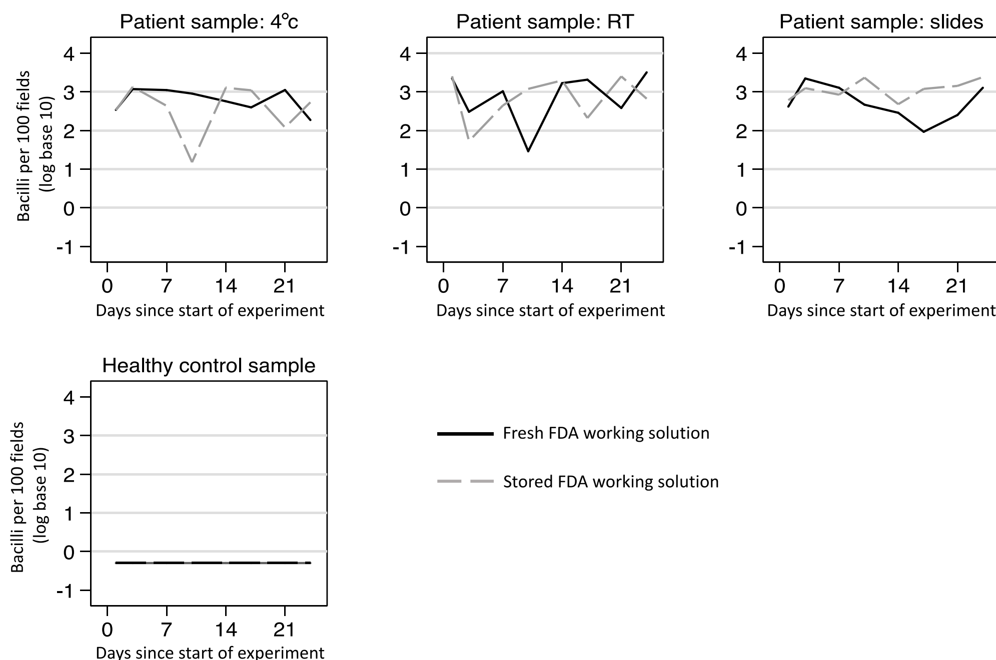
		Univariate regression analysis		
Variable		Odds ratio	95% CI	p value
FDA protocol F versus FDA protocol G		1.4	0.57 - 3.5	0.5
Acid-fast microscopy (grade), ++/+++ % (n)	64% (78)	1.4	0.74 - 2.7	0.3
Sputum had a salivary consistency, % (n)	42% (51)	0.52	0.15 – 1.8	0.3
No blood present in sputum, % (n)	90% (110)	13.8	0.97 - 194	0.05
Delay before processing, median days (IQR)	1 (0-3)	0.93	0.72 - 1.2	0.6
Delay in reading slide, median hours (IQR)	3.6 (3.1 – 4.4)	1.5	0.98 – 2.8	0.2
Rifampicin resistance, % (n)	13% (16)	0.95	0.15 – 6.0	1.0

Storage study

Quantitative assessment. For this study FDA protocol F was used to prepare 80 slides: 64 from a single patient sputum sample that was conventional acid-fast microscopy +++ smear grade, and 16 from a single healthy control sample. All slides from the healthy negative control sample had negative microscopy results. All slides from the patient with TB had positive results, with FDA microscopy results having a mean log-count per 100 fields of 2.8 (SD=0.40) on the first day and 2.7 (SD=0.53) over the next 4 weeks (Figure 5.6). Neither the way the samples were stored, nor the length of storage impacted the microscopy results (all p>0.2).

Quality assessment. The quality of slides for patient samples were high, with median score 300 (IQR=245-320) out of 400 and this did not change with storage time. Slide quality was more stable if the positive sputum sample was stored on slides and stained with stored FDA working solution than sputum stored at room temperature stained with fresh FDA working solution (p=0.03).

Figure 5.6 Storage study
Graphs showing the number of bacilli per 100 fields at 8 time-points during the 4-week experiment, comparing fresh versus stored fluorescein diacetate (FDA) working solution in each patient sample storage condition and healthy control sample. Note: RT=room temperature.



Cost analysis

The material and reagent costs for FDA microscopy using protocol F cost USD $0.02 more than conventional acid-fast microscopy with auramine staining. For a batch of 40 slides, FDA microscopy costs USD $0.05 per slide (Table 5.4).

Table 5.4 Cost analysis
Table demonstrating the reagents required and costs for fluorescein diacetate (FDA) microscopy using the optimum FDA staining protocol, Protocol F. All reagents except FDA were procured in Peru, and the suppliers are specified in the methods section. Note. USD=United States dollar, mg=milligrams and ml=millilitres.

	Reagent (unit)	Quantity	Price (USD)
FDA stock solution	FDA (mg)	5	$ 0.03
	acetone (ml)	1	$ 0.05
	Total cost		**$ 0.08**
FDA working solution for 40 slides	FDA stock solution (ml)	0.1	$ 0.01
	Acetone (ml)	10	$ 0.50
	Phosphate buffered saline (ml)	15	$ 0.01
	Total cost		**$ 0.52**
FDA staining for 40 slides	Slide	40	$ 0.69
	FDA working solution (ml)	24	$ 0.52
	0.5% acid alcohol (ml)	24	$ 0.05
	5% phenol (ml)	24	$ 0.19
	0.5% KMn04 (ml)	24	$ 0.02
	Transfer pipette	5	$ 0.55
	Total cost		**$ 2.02**
Total cost per slide			**$ 0.05**

Discussion

We compared, optimised, and assessed the reproducibility, optimal logistics and costs of sputum FDA microscopy. This enables us to recommend a simple, safe and inexpensive protocol, as shown in the final SOP in Box 5.2, which we recommend for use in future research and clinical practice. Importantly, this protocol obviates centrifuge-decontamination and uses low-cost LED microscopy, improving cost, feasibility and biosafety.

Sputum processing with centrifuge-decontamination is required for most mycobacterial culture methods to reduce cultures becoming unreadable because of overgrowth (often termed contamination) by other non-mycobacterial bacteria and fungi, which are abundant in sputum. Our and other studies have shown that the great majority of culturable *M. tuberculosis* in sputum is killed and/or discarded during centrifuge-decontamination.[44] It has also been proposed that sputum processing with centrifuge-decontamination prior to microscopy may increase diagnostic sensitivity by homogenizing sputum and breaking up clumps of *M. tuberculosis,*[153] but our results do not support this. Centrifugation requires costly equipment and often the centrifuges that are used with unsealed rotors are biohazardous because they generate infectious aerosols.[23,24] Centrifuge-decontamination may also selectively kill a specific phenotype of *M. tuberculosis,* distorting findings important to monitoring treatment response in both quantitative culture and FDA microscopy.[150] It is therefore of considerable operational importance that we found that FDA microscopy had optimal quality and reproducibility when performed directly on unprocessed sputum, without centrifuge-decontamination. Similar to conventional acid-fast microscopy, our proposed FDA protocol uses acid-alcohol to discriminate mycobacteria from non-acid-fast organisms, obviating centrifuge-decontamination.

There is some evidence that conventional sputum microscopy is relatively safe for laboratory workers.[154] However, FDA is not toxic to cells, so 5% phenol was applied after staining with FDA to ensure that the slides were sterile because there is abundant evidence that phenolics kill *M. tuberculosis* even when dried onto surfaces.[155–157] We found that applying phenol to slides after staining with FDA had no adverse effects on FDA microscopy sensitivity or quality.

We measured the quality of slides in each protocol, because good quality slides facilitate reading and reduce the time required by a microscopist for this process. We found that decontaminated sputum pellets were inadequately fixed onto the standard glass slides that are commonly used in clinical laboratories, resulting in false-negative results. We also found that thick direct sputum smears produced very poor-quality slides because sputum from patients contained other cells and extra-cellular material that obscured visualization

of mycobacteria. Thin direct sputum smears with the use of 30 seconds of potassium permanganate helped quench fluorescence from background material sufficiently to allow the mycobacteria to be optimally visualised. Consequently, the proposed FDA microscopy protocol required a median 10 minutes to read 100 high-powered fields of patient sputum smear. As microscopists gain more experience this slide reading time may reduce.

Preparing FDA working solution daily is time consuming and potentially wasteful. The storage study established that FDA working solution can be made once every 4 weeks without affecting microscopy results, increasing efficiency. Quality assurance systems are an integral part of providing a reliable TB laboratory service, and have been shown to have a positive impact when implemented.[158] Quality assurance of conventional sputum microscopy is done both internally by the preparation and regular reading of positive and negative controls, and externally, for example by the blinded staining and reading of centrally prepared slides.[159,160] This study demonstrates that quality assurance of FDA microscopy can be carried out in a similar manner, as the storage of positive control slides for up to 4 weeks did not affect results. There was no deterioration in any measures after 4 weeks storage and future research may demonstrate how much longer the reagents and slides may be stored.

FDA microscopy materials were calculated to cost only USD$0.05 per slide, which is USD$0.02 more than for auramine stained conventional acid-fast microscopy. Similar to all sputum microscopy techniques, FDA microscopy cannot guide management in paucibacillary disease. However, this low cost implies that FDA microscopy may be a cost-effective tool to answer clinically relevant questions and formal cost-effectiveness studies are warranted.

Similar to conventional acid-fast microscopy, a limitation of the current FDA microscopy protocol is that it will not be able to differentiate *M. tuberculosis* from other acid-fast bacilli, such as non-tuberculous mycobacteria. We did not specifically test the specificity of FDA microscopy on non-mycobacteria microorganisms, but we assume that the acid-fast step in our protocol would generally prevent them from staining, as is the case for other acid-fast stains. This hypothesis is supported by our finding that all slides prepared from sputum from healthy control participants were consistently negative with the chosen FDA microscopy protocol. Another limitation is that although microscopists were blinded, in the optimisation study there may have been clues from the appearance of slides that could have differentiated some protocols (e.g., thick versus thin smears). However, this potential limitation was reduced by taking photos of the microscopy views, which were reviewed by other team members. Finally, the same microscopist read all slides prepared that day; but to remove any systematic bias all FDA slides were shuffled before reading, and the number

of bacilli counted in 100 high-powered fields started from the first field in focus without knowledge of the conventional acid-fast microscopy result.

In conclusion, these experiments have optimised and demonstrated the reproducibility of a simple and relatively safe FDA microscopy protocol. This is novel because it provides the first published SOP recommended for clinical and research laboratories. Furthermore, by demonstrating that centrifuge-decontamination is an unnecessary step, this evidence-based protocol reduces barriers to implementation, especially in resource-constraint settings where FDA microscopy may have most value.

Box 5.2. Final standard operating procedure for fluorescein diacetate (FDA) solution preparation, staining and microscopy.

A. Stock solution

Mix 5 mg FDA in 1 ml acetone in a tube.
Store at -20°C. Can keep up to 2 years.

B. Working solution (25 ml) – enough for 40 slides

Put 10 ml of acetone into a clean glass or non-polystyrene tube (because acetone reacts with polystyrene)
Add 15 ml of phosphate buffer solution, pH 6.8
Cover tube with foil to protect from light
Add 0.1 ml of FDA stock solution.
Shake very well before use - this is necessary as the FDA stock solution is a suspension and separates within minutes.
Store in the dark at room temperature for up to 4 weeks.

C. Slide cleaning

1. Remove slides from box
2. Place slides in alcohol for at least 1 hour to clean the slides. We have found this to be particularly important.
3. Removes slides from solution and place on a clean area e.g. a new sheet of aluminium foil.
4. Wipe away excess solution with a clean, lint free cloth or tissue, for example the type of tissue used for cleaning lenses.
5. Do not touch the area where the sample will be applied.

D. Smear preparation

Use cleaned slides
Apply 1 drop of sample from a disposable transfer pipette to slide and make a smear of approximately 2 cm^2 area.
Leave slides to dry slides. If using a slide warmer, do not use a temperature more than 40°C.
Do no expose to ultraviolet light.

E. Staining protocol

When the slide is dry, pass over flame 3 times to fix the sample to the slide.
Shake FDA working solution before use – this is necessary as the FDA stock solution is a suspension and separates within minutes.
Apply 13-15 drops of FDA from a disposable transfer pipette onto samples to cover the smear.
Incubate slides for 30 minutes at 37°C.
Remove slides from incubator and remove excess liquid by tapping the slide.
Rinse GENTLY with distilled water.
Apply 0.5% acid-alcohol for 3 minutes.
Rinse GENTLY with distilled water.
Apply 5% phenol to slide for 10 minutes.
Rinse GENTLY and WELL with distilled water. Very important.
Apply 0.5% potassium permanganate for 30 seconds.
Rinse gently with distilled water.
Dry in dark place

F. Reading protocol with Zeiss Primo Star iLED microscope (Heidenheim, Germany)

Read slides within 0.5 to 5 hours of staining
Use the microscope's fluorescence light
Start with the 40x objective to locate the field of focus
Do not move the platform, add a drop of oil, and change the objective to 100x. MAKE SURE NOT TO GET OIL ON THE 40X OBJECTIVE.
Now only use the fine focus to focus the slide and start counting the number of bacilli visible with the 100X objective with oil.
When the slides are not being read, protect them from light.
Start each batch by reading the positive control slide.

RESEARCH CHAPTER 6

Evaluation – Pre-treatment TB viability microscopy assessing infectiousness

Background

There is evidence that improved diagnosis and prompt, successful treatment of patients with TB will not be sufficient to control the global TB epidemic and achieve the WHO End TB Strategy objective to reduce TB incidence by 90% by 2035.[161,162] Consequently, there is renewed interest in improving contact tracing and the provision of preventive TB therapy.[163,164] With an estimated quarter of the world population having latent TB infection, widespread treatment of latent tuberculosis infection may be an unfeasible task for most NTP. Therefore, guidelines have been proposed for low and middle income countries to focus contact investigations on people recently exposed to smear positive pulmonary TB.[165] However, even this approach may not currently be achievable in resource-constrained settings, especially where most treated patients are sputum microscopy positive.

Research Chapter 2 described a small study of 35 patients and their 209 contacts who were followed up for 6 years. This demonstrated that contacts of index patients who had less than the median concentrations of FDA-stained bacilli in their pre-treatment pooled sputum were more likely to develop secondary TB than those with higher than median concentrations of FDA-stained bacilli in sputum. This study also suggested that there was a particular *M. tuberculosis* population that grew in culture but did not stain with FDA that was the most infectious. These findings warranted a larger confirmatory study. Prior to starting this larger study, research described in Research Chapters 4 and 5 aimed to make FDA microscopy more feasible under operational conditions by optimising sputum collection and FDA microscopy staining.

The main objectives for this study were to confirm if: FDA microscopy can identify the most infectious index cases; and microbiological tools can identify an infectious subpopulation of *M. tuberculosis*.

Methods

Ethics. The study was approved by ethics committees including Asociacion Benefica PRISMA, Peru, Imperial College London Research Ethics Committee, UK and the local Peruvian Ministry of Health DIRESA Callao, Peru. The clinical significance of FDA microscopy results was uncertain, so these results were not shared with healthcare providers and therefore did not have any role in the care provided.

Patients. The study setting was 32 participating health clinics in the shantytowns and urban settlements in Callao, Peru. Recruitment took place from November 10, 2015 to December 14, 2017 and all consecutive patients were assessed for eligibility to participate. Inclusion criteria were commencing treatment for pulmonary TB and having received ≤1 day of treatment. Exclusion criteria were unwillingness or inability to give informed written consent to participate.

Procedures. For eligible participants who gave informed written consent, all sputum was collected as spot sputum samples using the instructions shown in Figure 6.1. Most patients provided sputum prior to or on the day of starting treatment. In the rare event of treatment not starting within 2 weeks of sample collection, a repeat sample was collected 14 days after the first sample. At the time of recruitment, patient demographics, symptom duration, co-morbidities, and TB treatment history were recorded.

Phases of the study. In the first phase of the study, all sputum samples were included in the study. However, FDA microscopy results were the focus of this research project and as is reported in detail below, it was noted that FDA microscopy results were rarely positive for samples that had negative results in conventional sputum microscopy using ZN staining. Consequently, to increase efficiency in research laboratory workload, in the second phase of the study from July 13, 2016 only sputum samples that had a positive result for ZN sputum microscopy were included in the study.

Contacts were defined as individuals who reported sharing the same household space with the patient for ≥6 hours/week for 2 weeks prior to the patient's initial treatment start date (or if treatment initiation was delayed, prior to the date of diagnosis). Contacts' name, national identification number, age and sex were recorded at recruitment. Almost all contacts had sputum testing for TB as a part of initial screening at the time of patient recruitment. All contacts who later reported developing symptoms suggestive of TB also then underwent microbiological testing

for TB disease. This in all cases included ZN sputum microscopy and in selected cases (those with a high level of clinical suspicion of TB disease, all with positive microscopy results and all those commencing treatment) also culture for *M. tuberculosis*. Secondary TB was defined as a case of TB in a contact whose symptoms started after the index patient's cough commenced (or in the absence of cough, the other first symptom attributed to TB). In order to identify these cases of secondary TB, data was collected from the local TB program and local research database identifying all newly diagnosed patients in the participating health posts from 1st January 2013 until 10th June 2020.

Sample processing. Fresh sputum was collected at room temperature, transported at 4°C and processed within 24 hours of arrival at the local research laboratory, which was usually less than 72 hours after expectoration. For the few samples that were less than 5 ml in volume, PBS (pH 6.8) was added to make the sample a final volume of 5 ml to ensure that there was sufficient sample for all tests.

Slides were prepared by smearing 1 drop from a disposable transfer pipette (approximately 40 µl volume) of unprocessed sputum over an area of 2 cm^2 on glass microscope slides that had been previously cleaned with 95% alcohol. Smears were dried on a slide warmer at 40°C for 1 hour, and then heat fixed by passing through a flame 3 times.

Conventional acid-fast staining. Smears were stained with a standard ZN protocol.[36] In summary, smears were flooded with carbol fuchsin and heated until steam was produced. This slide heating step was repeated another 2 times; the slide was rinsed with distilled water, decolourised by acid-alcohol for 3 minutes; rinsed with distilled water, flooded with methylene blue for 1 minute and then rinsed with distilled water.

FDA staining. Smears were stained using the SOP for FDA microscopy stated in Box 5.2. In summary, smears were covered with 20 µg/ml FDA working solution and left to incubate at 37°C for 30 minutes, excess fluid was then removed, and slides were rinsed with distilled water. Then 0.5% acid-alcohol solution was applied and left for 3 minutes, slides were rinsed well with distilled water before 0.5% potassium permanganate was applied for 30 seconds and rinsed with distilled water.

Figure 6.1 English translated instructions used for sputum sample collection

COLLECTING SPUTUM FOR TB TESTING

During this collection please make sure you are in designated área for sputum collection or in an open space with no one near you.

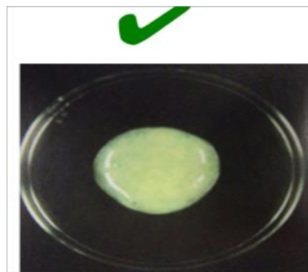
Take a deep breath in

Hold your breath for a few seconds

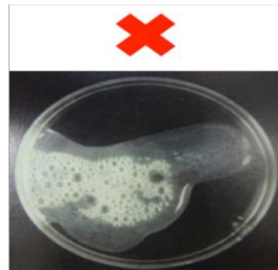
Cough rigourously until you can feel your sputum/phlegm moving from deep within your lungs

Deposit your sputum in the tube provided

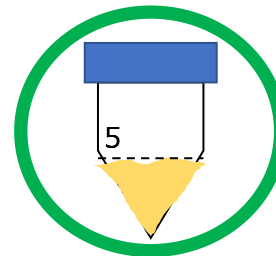
- **Repeat steps 1-4 at least 5 times**
- **Take at least 5 minutes for the whole process**



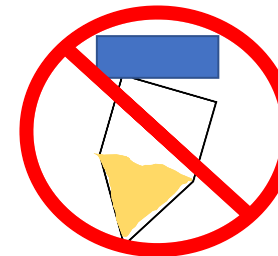
We need sputum (phlegm) with or without blood from your lungs



We do not want saliva or any nasal secretions



We need at least 5ml (a teaspoon full)



Make sure the lid is on tightly

Microscopy. All slides were stored in a cool place protected from light. Slides were read within 6 hours of staining using a Zeiss Primo Star iLED microscope.[166] Normal light was used for microscopy reading the slides stained with ZN. As explained above, for all samples until July 13, 2016 and after that date only for samples with a positive ZN microscopy result, the slide stained for FDA microscopy was read using florescent light with the 455 nm LED. All slides were read using the 100x objective with oil immersion without using microscopy cover slips. The number of stained bacteria visible in 100 consecutive microscopy fields were counted. A positive result was defined as 1 or more bacilli visible, as per current WHO recommendations.[140]

Quantitative culture was done using the Colour Test method, which is a thin-layer agar culture technique.[167,168] One ml of unprocessed sputum was added to 2 ml of trisodium phosphate decontaminant for 15 min. Two drops (with a disposable transfer pipette, approximately 80 µl) of the sputum-trisodium phosphate mixture were then applied to a petri dish quadrant prepared with 4 ml selective agar media (Middlebrook 7H11 solid culture with added carbendazim, Selectatab [Mast diagnostics] and a redox colorimetric indicator, 2,3-diphenyl-5-thienyl-(2)-tetrazolium chloride [STC][92]). Sealed culture plates were incubated at 37°C in un-supplemented air, examined 3-times a week first by naked eye, then if this revealed any colour change (which is suggestive of microbial growth) also with 40× magnification. The dates of decontamination and of first evidence of microbial growth were noted to calculate the TTP. The number of CFU with the characteristic cording morphology of *M. tuberculosis* was counted 42 days after decontamination and inoculation. If any colony with cording morphology was present then confirmation of the presence of *M. tuberculosis* was carried out with the Capilia assay (Tauns, Tokyo, Japan). Positive cultures were sub-cultured and analysed with TEMA, which is a WHO-approved test for identifying rifampicin and isoniazid resistance.[59]

PCR. Cepheid (Sunnyvale, California, USA) GeneXpert PCR testing was carried out concurrently with processing for microscopy and culture according to the manufacturer's instructions. If a patient had had a valid GeneXpert result within 2 weeks and had still not started treatment during the repeated sample collection, then the PCR was not repeated. GeneXpert cycle times (CT) for individual probes were extracted with user written software created by SAVICS International (Brussels, Belgium).

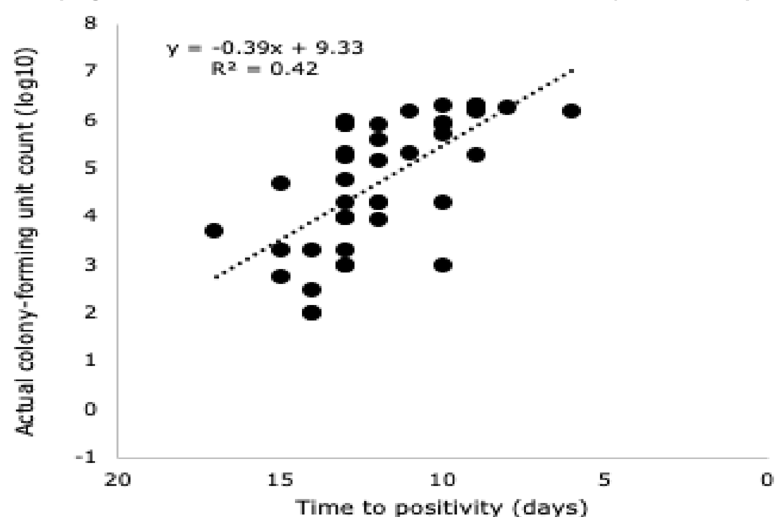
Sample dilutions. The volume of each sample of sputum that was tested for the presence of *M. tuberculosis* was calculated to be approximately 0.4 µl for ZN microscopy, 0.4 µl for FDA microscopy, 27 µl for culture and 667 µl for PCR (although the proportion of this volume actually tested by this commercial equipment is not reported by the manufacturer). Assuming that one *M. tuberculosis* bacteria would be detectable if it were present in each test, then the lowest expected concentration threshold of detection of *M. tuberculosis* in sputum samples would therefore be 2500 bacilli/ml for ZN microscopy, 2500 bacilli/ml for FDA microscopy, 38 CFU/ml for culture and reportedly 100 CFU/ml for PCR.[169,170] Thus, PCR and culture may be expected to detect lower concentrations of *M. tuberculosis* than ZN and FDA microscopy.

Mycobacterial load. Counts of bacteria visualised per 100 microscopy fields and CFU in culture were adjusted if the sample had been diluted with PBS. These data were then used to calculate bacilli and CFU concentrations per ml of sputum and transformed to base-10 log values prior to analysis. Zero values were transformed to 0.32, so that the log values equalled -0.5. CT results from negative PCR were transformed to 39 because GeneXpert runs a maximum of 38 cycles.[171] CFU are difficult to count when there is high mycobacterial load because of coalescing colonies. As described in Research Chapter 2, CFU has a good correlation with TTP, and therefore a composite CFU variable was made for analysis, where CFU counts above 200 were replaced with estimated CFU derived from TTP. This value was derived from a laboratory experiment carried out prior to analysis with 60 cultures of sputum from treatment-naive unselected, consecutively recruited study participants. Serial dilutions (1:1, 1:10, 1:100 and 1:1000) of each sample were cultured as described above. Figure 6.2 shows the association between TTP and log CFU. Consequently, CFU counts over 200 were replaced with the following equation log (CFU)=9.33-0.39*TTP, only if this value was higher that the recorded count. Mycobacterial load from PCR was calculated as the lowest CT recorded in probes A-E, which are targets for *M. tuberculosis*. For samples where individual probe CT data were not available, the semi quantitative result was extracted from Cepheid software to impute the mean CT value for the corresponding group ("negative" = 39, "very low" = 29.6, "low" = 24.3, "medium" = 18.5 and "high" = 14.2).

Power. Using data from Research Chapter 2, in order to evaluate a minimum 3% difference of secondary TB rates between those with higher than median versus lower than median FDA microscopy results with a power of 80% (alpha = 0.05),

4000 contacts were required. The median number of contacts per index patient in the study setting is 4, so this sample size required recruitment of 1000 patients.

Figure 6.2 Time to positivity versus colony-forming units in thin layer agar Scatterplot of the relationship between time to positivity (TTP) in days versus the number of colony-forming units (CFU) transformed to the logarithmic base 10 number in culture (log10 CFU = 9.33 - 0.39*TTP, R^2=0.42, p<0.0001).



General statistical analysis. Data were analysed with Stata version 13 (StataCorp, College Station, TX). Tests were 2-tailed, and proportions were calculated with their 95% CI. Paired analysis was used to compare results from the same sample by paired Student's t-test if data were normally distributed, and the Wilcoxon signed rank test for non-parametric data. Exploratory factor analysis with oblique rotation was used to assess variance and covariance between each microbiological test and identify mycobacterial populations as latent variables.

Time-to-event analysis. Secondary TB was analysed with a Cox proportional hazards model using a Huber White sandwich estimator to provide robust standard variance adjusting for household clustering. To compare the strength of association of each test in predicting secondary TB, log transformed bacilli/ml in ZN and FDA microscopy, log transformed CFU/ml in quantitative culture and 1 minus CT data in PCR were converted to Z scores and the hazard ratios (HR) and Harrell's concordance (C)-statistic were evaluated.[172] Previous data in this source population demonstrated that approximately 30% of secondary TB in household contacts were caused by transmission from sources outside of the house.[75,173] In the absence of strain typing to confirm patient-contact pairs, we used permutations in 10,0000 simulations to evaluate the variation in HR estimates when a random third of secondary TB cases are excluded from analysis.

Results

Participants. As shown in the study flow chart (Figure 6.3), during the study period 1030 pre-treatment samples were collected prior to initiation of a course of TB treatment from a total of 978 patients. The median age of patients was 29 years (IQR=21, 45) and 63% (612/978) were male. At recruitment, HIV co-infection was self-reported in 3.6% (35/974) of patients, and 7.2% (70/974) stated having type 2 diabetes mellitus. Extrapulmonary TB with no clinical or microbiological evidence of pulmonary involvement was diagnosed in 3.3% (34/1030) of these participants.

Samples. At the time of sample collection, 24% (245/1028) of patients reported that they had previously received treatment for TB at any time in their lives, with 1.8% (19/1030) of patients receiving some form of TB treatment in the 2 months prior to the current sample collection. During the first phase of the study, 46% (236/510) samples had negative ZN microscopy, of which 5.1% (12/236) were positive in FDA microscopy. In contrast as shown in Figure 6.3, FDA microscopy was positive in 78% (615/794) of samples with positive ZN microscopy and 61% (627/1030) of all samples. There were 2.6% (27/1030) quantitative cultures that were positive for *M. tuberculosis* but had significant contamination, and their CFU data were excluded from analysis. CT data from PCR was not available for 1.1% (11/1030) of samples due to failed analysis ('error' result) or the test not being performed because a recent result was available for the patient.

Mycobacterial load. The distribution of bacilli and CFU per ml of sputum and CT data are shown in Figure 6.4. Although culture detected lower concentrations of *M. tuberculosis* than ZN and FDA microscopy, in the 77% (794/1030) sputum samples that were ZN microscopy positive, ZN results were 0.49 log (SD=1.4) bacilli/ml higher than the paired CFU/ml in culture ($p<0.001$), and 0.61 log (SD=0.79) bacilli/ml higher than the paired FDA microscopy result ($p<0.001$); whilst FDA microscopy results were 0.12 log (SD 1.6) bacilli/ml lower than the paired CFU/ml in culture ($p=0.01$). Considering the 55% (567/1030) sputum samples that had positive quantitative results in ZN microscopy, FDA microscopy and culture, mean ZN results were 0.32 log (SD=0.98) bacilli/ml higher than the paired CFU/ml in culture ($p<0.001$), and 0.37 log (SD=0.63) bacilli/ml higher than the paired FDA microscopy result ($p<0.001$); whilst mean FDA microscopy results were 0.052 log (SD 1.0) bacilli/ml lower than the paired CFU/ml in culture ($p=0.06$)

Clinical associations (Table 6.1). Children aged between 0-15 years had lower mycobacterial load than participants aged >15 years ($p<0.001$). Participants with

indicators of more severe TB disease i.e., symptomatic disease for more than 6 weeks; constitutional symptoms (fever and/or night sweats); or a frequent cough (more than once per hour) had higher mycobacterial load in sputum (ZN and FDA microscopy, culture and PCR all $p<0.001$). Adults with a lower BMI tended to have higher mycobacterial load, which was significant for PCR. HIV coinfection, diabetes or drug resistance were not associated with any significant difference in mycobacterial load.

Agreement. The results of ZN microscopy bacilli/ml, FDA microscopy bacilli/ml, culture CFU/ml and PCR CT were all significantly correlated with one another (all $p<0.0001$, Figure 6.5). The strongest correlations were between ZN versus FDA (rho=0.84) and ZN versus CT (rho=-0.86). As shown in Figure 6.4, several samples had negative or low counts in one test but very high counts in another.

Factor analysis results are shown in Table 6.2. Factor 1 accounted for 79% of the total variability and was described best by microscopy results in both ZN and FDA. Factor 2 accounted for 12% of total variation in results and was influenced most by CFU results in quantitative culture. Both factors accounted for 92-93% of variation in microscopy and culture results, and 86% of the variation in CT results from PCR.

Figure 6.3 Study flow chart

Note ZN=Ziehl-Neelsen and microbiologically positive sample or microbiologically confirmed TB= evidence of a positive result in ZN microscopy, culture or PCR in the same sample or patient respectively.

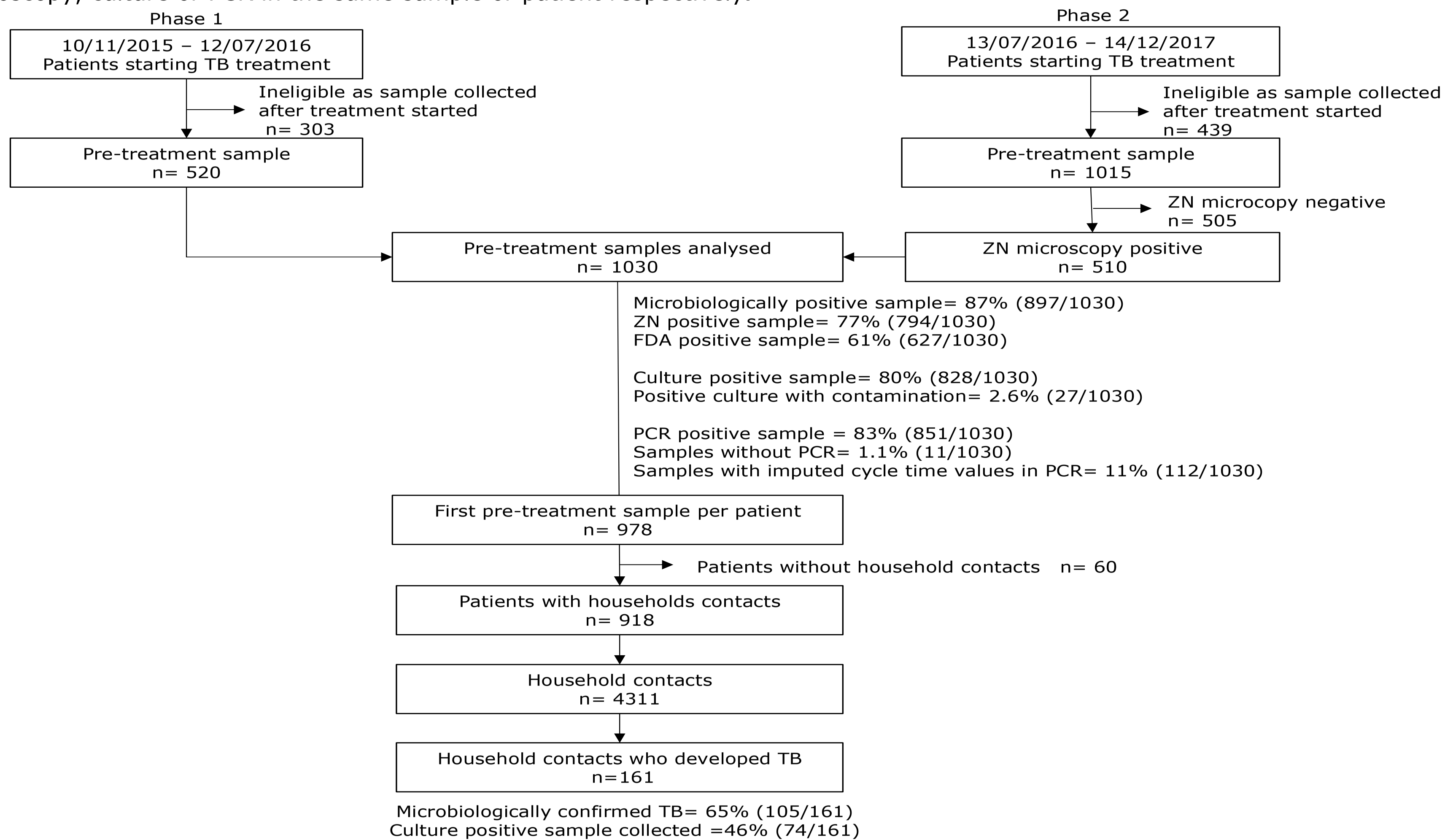


Figure 6.4 Distribution of results
Histograms demonstrating the distribution of results in Ziehl-Neelsen (ZN) microscopy, fluorescein diacetate microscopy (FDA), colony forming units (CFU) in quantitative culture and cycle times in PCR with GeneXpert. Below the X-axis is the indication of the range of concentrations in each microscopy, culture and PCR grade. Note that positive cultures with <1000 CFU/ml, equivalent to 3 base-10 logarithms (log), were classified as '-' in order to correspond to the equivalent concentrations for microscopy.

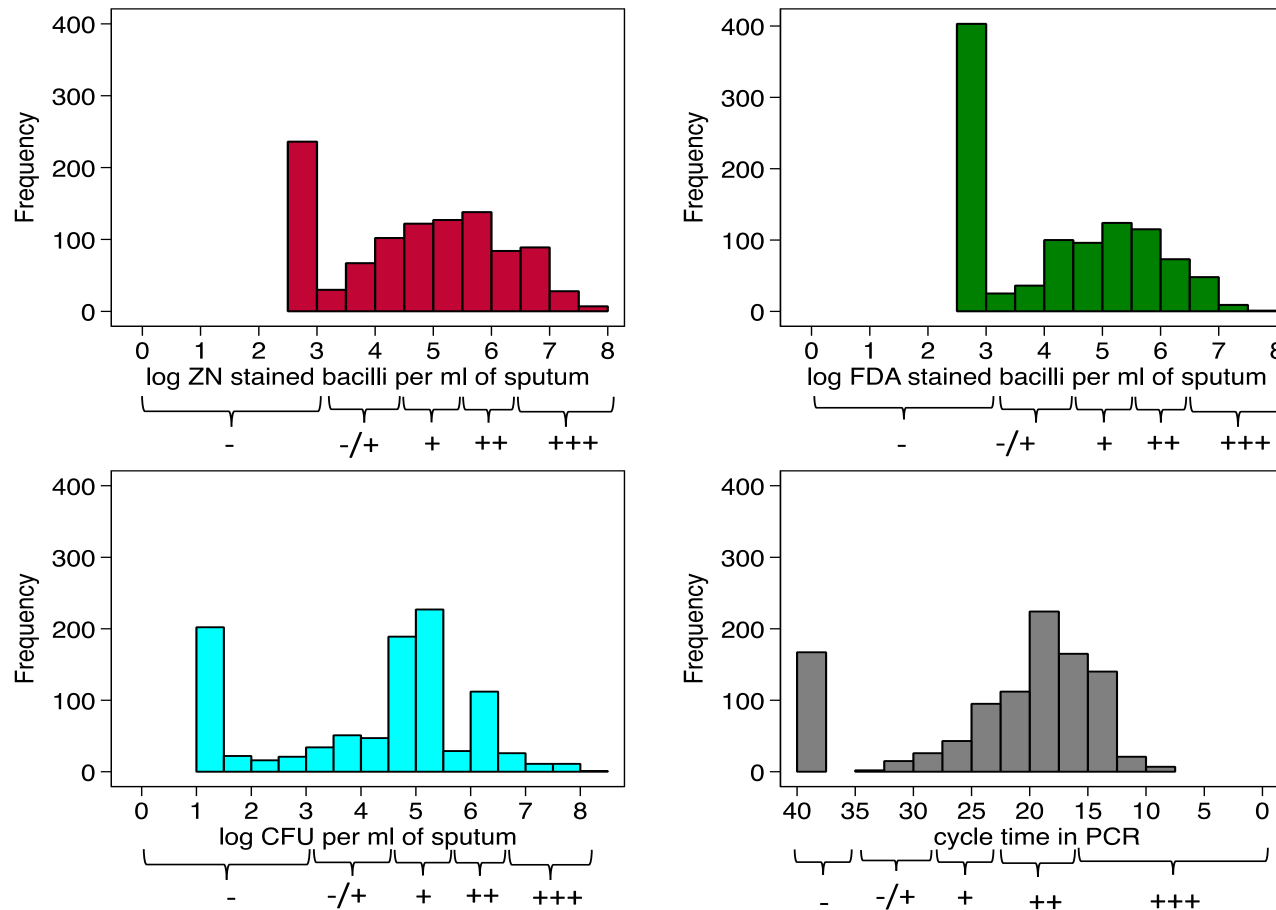


Table 6.1 Mean values of each test by demographics and TB disease type
Demographic data (sex, age, HIV, diabetes and body mass index [BMI]) only take into consideration the patient's first sample (N=978), whereas TB disease data (drug resistance, cough frequency, symptom duration and presence of constitutional symptoms) show results of all samples (N=1030). Bold typeface indicates statistically significant differences ($p<0.05$) between groups with the first category as the reference group. Negative microscopy with either Ziehl-Neelsen (ZN) or fluorescein diacetate (FDA) have base-10 logarithm (log) of 2.9 per ml, negative culture had log 1.1 colony forming units (CFU)/ml, and negative PCR were allocated a result of 39 cycle times (CT)

		n	log ZN/ml	log FDA/ml	log CFU/ml	CT
All samples		1030	4.7 (1.4)	4.3 (1.3)	4.1 (1.9)	22 (8.5)
Samples from patient	First patient sample	978	4.7 (1.4)	4.3 (1.3)	4.2 (1.8)	**22 (8.4)**
	Repeated sample from same patient	52	4.7 (1.4)	4.5 (1.4)	3.7 (2.1)	**26 (9.8)**
Sex	Female	366	**4.6 (1.4)**	**4.2 (1.3)**	**4.0 (1.9)**	**23 (9.1)**
	Male	612	**4.9 (1.3)**	**4.4 (1.3)**	**4.3 (1.8)**	**21 (7.9)**
Age	0-15 years	28	**3.9 (1.2)**	**3.7 (1.2)**	**2.8 (1.9)**	**28 (9.5)**
	15-29 years	465	**4.8 (1.3)**	**4.3 (1.3)**	**4.3 (1.8)**	**22 (8.1)**
	30-59 years	373	**4.8 (1.4)**	**4.4 (1.3)**	**4.2 (1.9)**	**22 (8.5)**
	>60 years	108	**4.7 (1.4)**	**4.2 (1.4)**	**4.0(1.9)**	**23 (8.7)**
Self-reported HIV	Positive	35	4.7 (1.4)	4.3 (1.3)	3.9 (1.9)	23 (9.4)
	Negative/ unknown	939	4.7 (1.4)	4.3 (1.3)	4.2 (1.8)	22 (8.3)
Self- reported diabetes	Positive	70	4.6 (1.2)	4.3 (1.2)	4.1 (1.8)	22(7.7)
	Negative/unknown	904	4.8 (1.4)	4.3 (1.3)	4.2 (1.9)	22 (8.5)
BMI in age ≥18 years	Under weight (<18.5)	131	5.0 (1.3)	4.4 (1.3)	4.4 (1.6)	**20 (6.5)**
	Normal weight (18.5-24.9)	584	4.8 (1.4)	4.4 (1.3)	4.2 (1.8)	**22 (8.2)**
	Overweight (25-29.9)	127	4.7 (1.4)	4.2 (1.3)	4.1 (2.0)	**23 (9.2)**
	Obese (≥30)	25	4.3 (1.2)	4.2 (1.3)	3.8 (1.9)	**26 (9.7)**
TB drug resistance*	Rifampicin susceptible	752	5.1 (1.2)	4.5 (1.3)	4.6 (1.5)	19 (5.1)
	Rifampicin resistance	123	5.1 (1.2)	4.6 (1.3)	4.5 (1.6)	20 (5.5)
Cough	None	79	**3.4 (0.93)**	**3.3 (0.80)**	**2.2 (1.5)**	**32 (8.5)**
	Infrequent (daily/weekly)	208	**4.5 (1.3)**	**4.1 (1.3)**	**3.8 (2.0)**	**24 (9.1)**
	Frequent (≥ hourly)	746	**5.0 (1.3)**	**4.5 (1.3)**	**4.4 (1.7)**	**21 (7.4)**
Symptom duration	0-44 days	480	**4.5 (1.4)**	**4.2 (1.3)**	**3.9 (2.0)**	**24 (9.2)**
	≥45 days	548	**4.9 (1.3)**	**4.4 (1.3)**	**4.4 (1.7)**	**21 (7.5)**
Constitutional symptoms	No fever/night sweats	224	**4.3 (1.3)**	**4.0 (1.2)**	**3.7 (1.9)**	**25 (9.4)**
	Fever/night sweats present	804	**4.9 (1.4)**	**4.4 (1.3)**	**4.3 (1.8)**	**21 (8.0)**

Footnote. * Excluding 154 samples without drug susceptibility results as both culture and PCR results were negative, failed or GeneXpert was not processed for that sample.

Figure 6.5. Comparison of results from each test
Dot plots of the distribution of: ZN-stained bacilli/ml by mycobacterial load in (A) quantitative culture, (B) FDA microscopy and (C) PCR; FDA-stained bacilli/ml by mycobacterial load in (D) quantitative culture, (E) ZN microscopy and (F) PCR; colony-forming units (CFU)/ml in quantitative culture by mycobacterial load in (G) FDA microscopy, (H) ZN microscopy and (I) PCR; and cycle times (CT) in PCR by mycobacterial load in (G) FDA microscopy, (H) ZN microscopy and (I) quantitative culture. Note boxes indicates the median and interquartile range. For the range of bacilli/ml, CFU/ml and CT in grades of mycobacterial load in microscopy, culture and PCR respectively, please see Figure 6.4

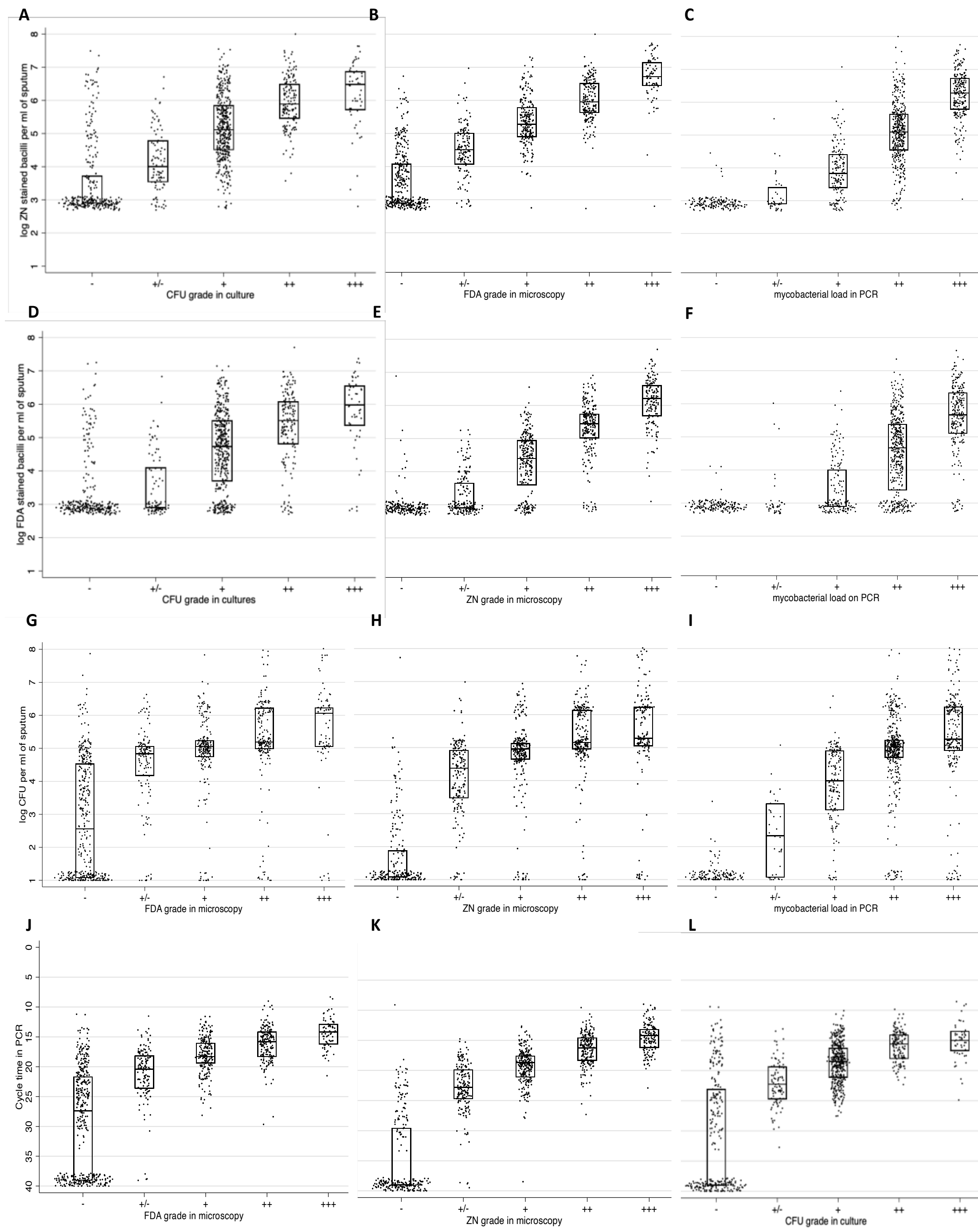


Table 6.2 Factor analysis to identify potential subpopulations
Factor loadings are shown after oblique rotation for results in Ziehl-Neelsen (ZN) and fluorescein diacetate (FDA) microscopy, colony forming units (CFU) in quantitative culture and cycle times (CT) in PCR.

			Rotated factor loadings			
	Eigenvalue	Proportion	ZN	FDA	CFU	CT
Factor 1	3.2	79%	0.78	1.0	-0.06	-0.35
Factor 2	0.49	12%	0.25	-0.08	1.0	-0.67
Uniqueness	-	-	0.08	0.07	0.07	0.14

Contacts. At the time of patient recruitment, 4311 contacts details were recorded. Contacts each reported spending at least 6 hours per week in the households of 918 index patients in the 2 weeks prior to the TB diagnosis of these index cases. The median age of contacts was 25 years (IQR=11, 45 years) and 51% were female (2187/4246). Contacts were followed-up for a median 3.5 (IQR=2.9-4.0) years.

Secondary TB. During follow-up, 3.7% (161/4311, 95% CI=3.2%, 4.3%) of contacts started treatment for TB disease. These cases of secondary TB occurred in 14% (132/918, 95% CI=12, 17%) of the index case households. Of these cases of secondary TB, 65% (105/161) had microbiological confirmation of TB with either ZN microscopy, culture or PCR and 46% (74/161) had a positive culture in our research sample. Our research sputum culture was also positive for 42% (67/161) of the index patients sharing the same household as these secondary cases.

Predictors of secondary TB in contacts. Table 6.3 and Figure 6.6 show that higher mycobacterial load in any sputum test significantly predicted a higher risk of contacts subsequently developing secondary TB, considering all secondary TB (all $p<0.001$) or only microbiologically confirmed secondary TB cases (all $p<0.001$). In univariable analysis the only index patient characteristics described in Table 6.1 that were associated with secondary TB were: reporting a frequent cough (HR=1.7, 95% CI=1.1, 2.7, p=0.003); the presence of constitutional symptoms (HR=2.0, 95% CI=1.2, 3.3, p=0.01) and rifampicin resistance (HR=0.55, 95% CI=0.30, 0.98, p=0.04). The most parsimonious multivariable regression model with the highest C-statistic for predicting secondary TB included a marker of mycobacterial load with self-reported cough frequency (Table 6.3). Mycobacterial load assessed by either ZN microscopy and PCR had higher C-statistics compared to FDA microscopy or quantitative culture, as both ZN and PCR results were able to discriminate 62% of secondary TB cases from contacts who did not develop secondary TB. ZN microscopy, FDA microscopy, culture and PCR positivity were also associated with 2-3 times increased hazard of secondary TB (all $p<0.01$), but all with poor discriminatory ability as C-statistics ranged between 56-59%.

Table 6.3 Cox proportional hazards model
Unadjusted hazard ratio (HR) for all secondary cases of TB, microbiologically confirmed secondary cases of TB, and HR adjusted for cough frequency for each Z score increase in Ziehl-Neelsen microscopy (ZN), fluorescein diacetate microscopy (FDA), colony-forming unit (CFU) in quantitative culture and cycle times (CT) in PCR. All results shown had p≤0.001

	Unadjusted analysis All TB (n=161 events)		Unadjusted analysis Confirmed TB (n=105 events)		Adjusted analysis All TB (n=161 events)	
	HR (95% CI)	C-statistic	HR (95% CI)	C-statistic	HR (95% CI)	C-statistic
ZN	1.5 (1.3, 1.8)	62%	1.9 (1.5, 2.4)	68%	1.5 (1.2, 1.8)	64%
FDA	1.4 (1.2, 1.7)	61%	1.7 (1.4, 2.1)	66%	1.4 (1.2, 1.7)	62%
CFU	1.4 (1.2, 1.7)	58%	1.6 (1.3, 2.0)	63%	1.3 (1.1, 1.6)	61%
CT	1.6 (1.3, 2.0)	62%	2.2 (1.5, 3.2)	68%	1.6 (1.2, 2.0)	63%

Footnote. Cough frequency (coughing hourly or more versus less than hourly) adjusted hazard ratio=1.6 (95% CI 1.1, 2.2) and p<0.02 in all analyses.

Sub-group analysis. Repeating the key analyses from the foundation study, a pre-defined sub-group analysis was performed to test the hypothesis that in index patients with higher than median CFU/ml in sputum quantitative culture and lower-then-median FDA microscopy results predicted a higher risk of household contacts developing secondary TB. The 50% (478/957) of index patients with ≥75000 CFU/ml of sputum had study follow-up data for 2178 household contacts, 4.4% (96/2178) of whom were treated for secondary TB disease during follow-up. The HR of secondary TB in contacts of the 50% (486/978) selected index patients with lower than median FDA microscopy results was 0.43 (95% CI 0.22, 0.84, p=0.01), significantly less than those with higher than median FDA microscopy result (2.2%, 11/490 versus 5.1%, 85/1682 respectively, Figure 6.7).

Simulation. Permutations accounting for 30% (48/161) of secondary TB cases in contacts that may have been infected from an outside source did not alter the positive association between increasing mycobacterial load and secondary TB on all the four microbiological tests. (Figure 6.8). The median HR estimates were similar to the unadjusted HRs shown in Table 6.3, and the C-statistic ranged between 57-66%.

Test performance. Table 6.4 demonstrates the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of each sputum test at

varying thresholds in predicting secondary TB in contacts. As shown, all had poor predictive power, including PPV which ranged between 4.1%-6.8%.

Figure 6.6 Time-to-event analyses

Kaplan-Meier curves demonstrating rates if secondary TB in contacts by grade of mycobacterial load in (a) Ziehl-Neelsen microscopy; (b) FDA microscopy; (c) quantitative culture; and (d) PCR. For the range of bacilli/ml, colony-forming units (CFU) /ml and cycle times (CT) in grades of mycobacterial load in microscopy, culture and PCR respectively, please see Figure 6.4

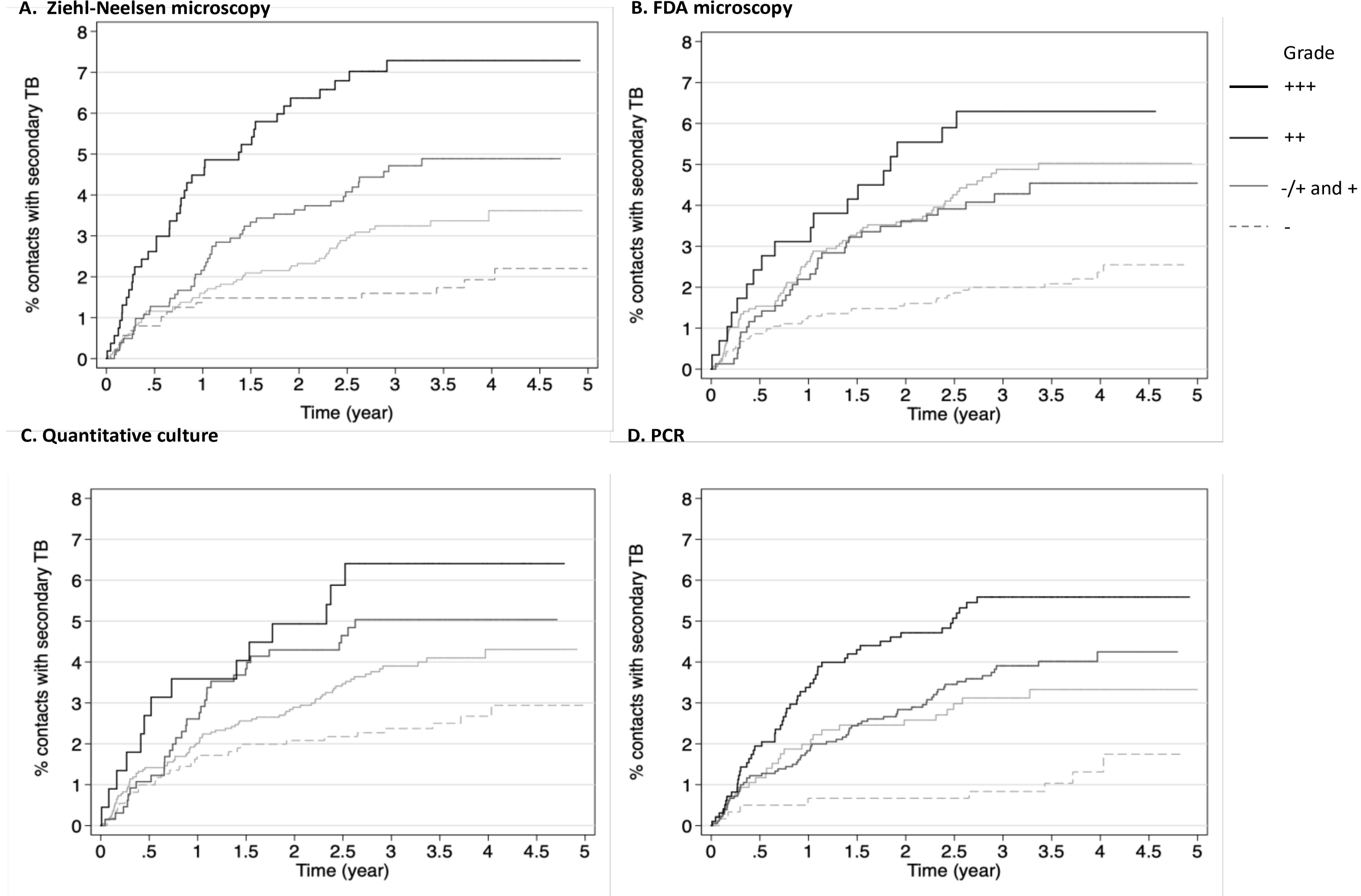


Figure 6.7 Time-to-event analysis of the interaction between FDA microscopy and culture results
Kaplan-Meier time to event curves for patients' household contacts being diagnosed with TB disease analysed according to whether the study patients' pre-treatment sputum fluorescein diacetate (FDA) microscopy result was above (high) or below (low) the median concentration. In Graph A, this is shown for patients with quantitative culture colony forming units (CFU) results below the median concentration. In Graph B, this is shown for patients with quantitative culture CFU results above the median concentration.

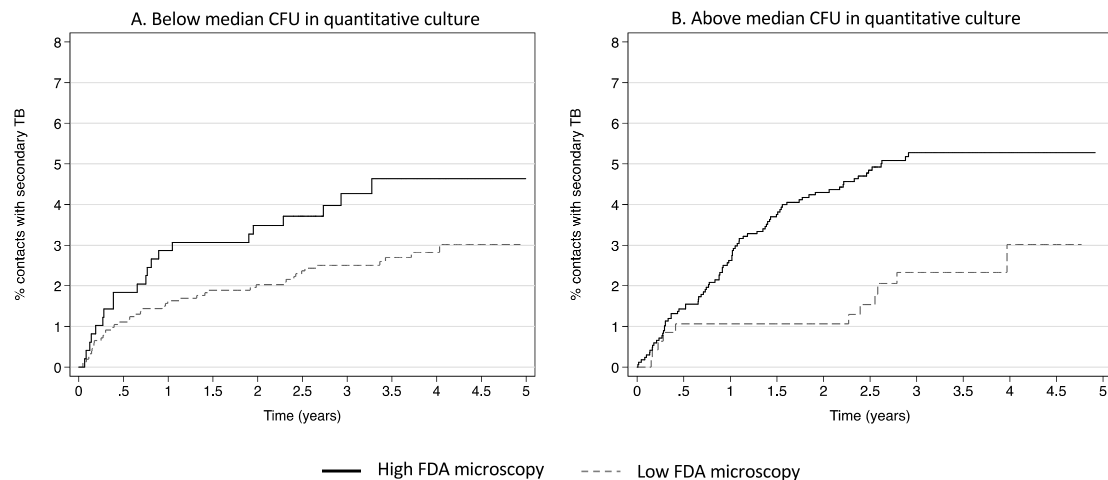


Figure 6.8 Permutations simulation results
Distribution of unadjusted hazard ratio (HR) to predict secondary TB in the permutations simulation of Z scores calculated from the results in:

(A) Ziehl-Neelsen microscopy (ZN);
(B) fluorescein diacetate microscopy (FDA);
(C) colony-forming units (CFU) in quantitative culture; and
(D) cycle times in PCR.

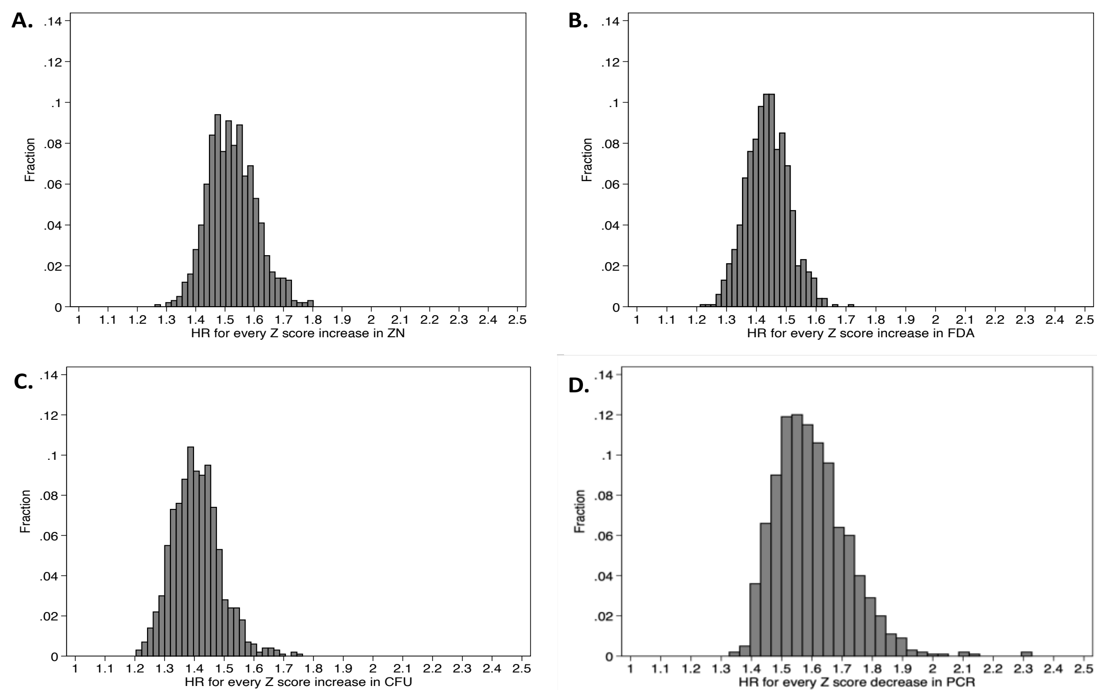


Table 6.4 Test performance

Sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for Ziehl-Neelsen (ZN) microscopy, fluorescein diacetate (FDA) microscopy, culture and PCR to predict secondary TB in contacts. For these calculations binary variables were made with varying thresholds: "positive" = positive versus negative test results; "++/+++" = mycobacterial grade ++ or +++ versus + or +/- or negative test results; and "+++" =mycobacterial grade +++ versus ++ or + or +/- or negative test results. For the range of bacilli/ml, colony-forming units (CFU) /ml and cycle times (CT) in grades of mycobacterial load in microscopy, culture and PCR respectively, please see Figure 6.4. Note. Confidence intervals are shown in parentheses and are estimates as they do not take clustering into consideration.

	Threshold	Sensitivity	Specificity	PPV	NPV
ZN	positive	89% (84%, 94%)	21% (20%, 23%)	4.2% (3.6%, 5.0%)	98% (97%, 99%)
	++/+++	53% (45%, 61%)	64% (63%, 65%)	5.4% (4.3%, 6.6%)	97% (97%, 98%)
	+++	24% (17%, 31%)	88% (86%, 88%)	6.8% (4.9%, 9.2%)	97% (96%, 97%)
FDA	positive	78% (70%, 84%)	39% (37%, 40%)	4.7% (3.9%, 5.6%)	99% (97%, 99%)
	++/+++	32% (25%, 40%)	75% (74%, 77%)	4.8% (3.6%, 6.2%)	97% (96%, 97%)
	+++	11% (6.8%, 17)	93% (93%, 94%)	6.2% (3.7%, 9.6%)	96% (96%, 97%)
Culture	positive	93% (88%, 97%)	18% (17%, 20%)	4.3% (3.6%, 5.0%)	99% (97%, 99%)
	++/+++	29% (22, 37%)	80% (78%, 81%)	5.2% (3.9%, 6.9%)	97% (96%, 97%)
	+++	8.8% (4.9%, 14%)	95% (94%, 96%)	6.3% (3.7%, 9.6%)	96% (96%, 97%)
PCR	positive	94% (89%, 97%)	15% (14%, 16%)	4.1% (3.4%, 4.7%)	98% (97%, 99%)
	++/+++	77% (69%, 83%)	34% (33%, 36%)	4.3% (3.6%, 5.1%)	98% (97%, 98%)
	+++	34% (26%, 41%)	77% (76%, 78%)	5.2% (4.0%, 6.8%)	97% (96%, 97%)

Discussion

This well powered study of 4311 contacts followed up for a median 3.5 years demonstrated that the mycobacterial load in pre-treatment sputum was associated with increased risk of secondary TB in household contacts. Mycobacterial load was assessed by bacilli counts in ZN and FDA microscopy, colony counts in solid culture and by CT in PCR, all of which were associated with symptom severity. These four microbiological tests had good agreement with each other, and factor analysis suggested they were identifying the same mycobacterial population in pre-treatment samples, with none detecting a more infectious sub-phenotype.

Evidence increasingly suggests that latent TB infection and active TB disease are on a spectrum with the latter generally having higher mycobacterial burden, detectable lung pathology and symptoms.[174,175] In this study we found that mycobacterial load in pre-treatment sputum was also associated with increased symptom severity such as: increased cough frequency, longer symptom duration, and the presence of constitutional symptoms. This is supported by previous findings discussed in Research Chapter 2, and other published studies.[176–178] Although the mechanisms by which TB disease cause symptoms such as cough in humans are incompletely characterised, contributing factors probably include the presence of mycobacteria, the chemicals they produce and the host immune response.[179,180] Therefore the association we demonstrated between symptom severity and mycobacterial load was expected.

FDA, ZN and culture identify different components of *M. tuberculosis* bacteria: FDA metabolic activity; conventional acid-fast microscopy the mycolic acid in the cell wall; and culture the ability to replicate.[11,181] Therefore, quantitative results of these different tests in pre-treatment sputa are likely to be correlated. The results of factor analysis suggested that in pre-treatment sputum, ZN microscopy, FDA microscopy, culture and PCR were predominantly identifying the same mycobacterial populations, which was best described by ZN and FDA microscopy counts. The second factor in this analysis accounted for 12% of the variability in results and was determined by quantitative culture and PCR results; most likely this second latent variable was identifying samples with lower concentrations of mycobacteria in sputum (that generally had false-negative microscopy results) rather than identifying a particular mycobacterial population.

The results of this study demonstrated that conventional microscopy with ZN detected higher concentrations of acid-fast bacilli than the colonies that were

cultured from the same samples, whereas culture had a more sensitive threshold of detection. To summarise these paradoxical findings, culture and PCR had more sensitive thresholds of detection probably because they evaluated 68 to 1668-times more sputum in each assay than microscopy. However, ZN microscopy consistently visualised higher concentrations of bacilli/ml of sputum than the CFU/ml that were cultured from the same sample, probably because many of the mycobacteria visualised within sputum were not culturable. FDA microscopy visualised similar concentrations of bacilli/ml of sputum to the CFU/ml that were cultured from the same sample, potentially because FDA microscopy visualised metabolically active, viable mycobacteria that are likely to be culturable. It should also be noted that mycobacteria are known to clump together due to their hydrophobic nature and therefore are not equally distributed within a sample, which can contribute to quantification discrepancies.[182]

The association between lower FDA microscopy bacilli counts and increased infectiousness as described in Research Chapter 2 was not supported by the current much larger study, which instead demonstrated the opposite. The foundation study findings in Chapter 2 led to the hypothesis that FDA microscopy was unable to stain a particular mycobacterial phenotype that was conditioned for more effective transmission. This difference in our findings may be explained by: changes in the current FDA staining protocol; or using instructed spot sputum collection rather than 12-hour pooled sample collections which may have added increased variation to mycobacterial load in each sample; or improvement in TB care and changes in patient and contact characteristics between the cohort in 2006 versus 2015-17.[40,151] However, it is most likely due to the 5% random chance of a type 1 error in the study reported in Research Chapter 2. A recent study using cough aerosols, demonstrated that there was no association between the number of "persister-like" dormant cells identified by red Nile staining in sputum microscopy and the number of CFU in cough aerosols collected from patients before they started treatment.[183]

Higher concentrations of mycobacteria in sputa identified by ZN or FDA microscopy, culture or PCR were associated with increased risk of secondary TB. The WHO recommend that in low- and middle-income countries where contact investigations of all TB-affected households are not considered feasible, contact investigations are prioritised for households of conventional sputum microscopy positive patients.[184] This may be explained by smear positivity identifying sputum that is directly more infectious because it contains higher concentration of *M. tuberculosis* and also indirectly because it is associated with higher cough frequency and aerosol

generation.[183] Quantitative assessment of mycobacterial load was more strongly predictive of secondary TB than a simple positive or negative test result. Yet, even quantitative microbiological markers, as shown in this study, only predicted the risk of secondary TB with moderate accuracy (with/or without the inclusion of cough frequency), so that using index characteristics such as mycobacterial load in a single sputum sample to identify households that need investigations would miss many contacts who would go on to develop secondary TB. This result is in keeping with increasing evidence highlighting the multifactorial nature of TB transmission which is not only dependent on mycobacterial load of an infectious index case, but the proximity, length of exposure and susceptibility of the contact.[185] Therefore, the use of multivariable risk scores, proximity scores or more reliable biomarkers in the contact require further attention than the current practice of providing care according to index case characteristics.[186–188]

In high TB prevalence settings, exposure to infectious TB can occur from multiple people, in addition to the household "index case".[75,76] TB strain typing, previously with restriction fragment length polymorphism (RFLP) and increasingly using whole genome sequencing, is performed to assess the likelihood that the *M. tuberculosis* in index-contact pairs are clonal strains. If these tests demonstrate that both strains in the pair are genomically similar and the household contact developed disease after the index case it is assumed that the contact was infected by the index case, but it could also mean that they were both infected by another person.[189] Furthermore, strains may become more genetically diverse with continuing disease in the same person.[190] Strain typing is normally carried out from a culture to prevent test failures and obtain clear results, thus generally requiring that both of the pairs are culture positive. Consequently, these techniques can usually only include index case-contact pairs that both have positive *M. tuberculosis* cultures. Culture-negative patients often include children and people with HIV coinfection or extrapulmonary disease, who are less likely to have microbiological confirmation of TB disease due to the paucibacillary nature of their illness and/or challenging sample collection.[101,191] A simulation approach as was used in this study offers a more inclusive strategy to adjust the estimated impact of epidemiological/microbiological markers on TB transmission, and can be complementary to analysis by strain typing.

The main limitation of this study was that most of the secondary TB was only identified if contacts had received a diagnosis of TB in the participating health posts or in our research laboratory. Therefore, the number of cases with secondary TB

may have been underestimated as those who were diagnosed in other regions or health systems and those who did access care were not accounted for. However we feel that this is a minor concern because: (1) in this setting few people moved out of the area; (2) this population infrequently used alternative health care systems for TB care (3) regular active case finding over 10-years detected a minority of all cases of secondary TB in household contacts of a sputum microscopy positive patient, most of whom were identified early in disease and may have presented to the health posts later in time.[192]

In conclusion, low concentrations of FDA microscopy-positive *M. tuberculosis* in sputum did not identify a subgroup of patients who were the most infectious. Although higher sputum mycobacterial load of index patients as assessed by ZN microscopy, FDA microscopy, culture or PCR was significantly associated with higher risk of secondary TB in household contacts, these results alone were not discriminatory enough to reliably predict which TB-affected households should be prioritised for contact investigations.

RESEARCH CHAPTER 7

Evaluation – TB viability microscopy assessing treatment outcome

Background

Research Chapter 3 described a foundation study involving 35 patients with pulmonary TB, which demonstrated that response to treatment could be assessed within 9 days using sputum FDA microscopy.[20] To make FDA microscopy suitable for use in clinical settings, especially resource-constrained settings, refinements were made to both the manner in which sputum was collected (Research Chapter 4) and to the FDA microscopy protocol (Research Chapter 5).[22,25]

Rates of drug resistance are believed to be increasing globally and it is now widely recommended that a sample from every patient starting TB therapy should undergo drug-susceptibility testing.[1] TB culture is the gold standard technique for identifying drug resistance, but is not currently feasible in many settings and usually takes several weeks to provide results. Since the foundation study was carried out there have been advances in diagnostics for drug-susceptibility testing, in particular widespread implementation of the GeneXpert cartridge-based nucleic acid amplification test produced by Cepheid (CA, USA).[193] GeneXpert is thought to have become the most frequently used test of rifampicin susceptibility and implementation has been associated with an increase in testing coverage from <20% of patients globally in 2013 to 61% in 2019.[194] The reasons for the rapid adoption of this tool include: simple processing; greater sensitivity for diagnosing TB compared to conventional sputum microscopy; and simultaneous identification of rifampicin resistance, which is a surrogate marker for MDR TB. However, the costs (of the equipment, consumables and maintenance) and requirements for stable electricity and temperature have hampered implementation.[195] Furthermore, introduction of GeneXpert to clinical settings has not had any discernible population-level impact on TB mortality nor morbidity.[196,197]

The currently widely available GeneXpert cartridges test for TB and rifampicin resistance, but not for isoniazid resistance. Isoniazid monoresistance occurs in an estimated 13% of TB cases globally, although the precise proportion is unknown because globally only a small proportion of patients with TB are tested for isoniazid resistance and drug-resistance survey coverage is incomplete.[1] Early bactericidal activity (EBA) studies have shown that treatment with only one drug that has rapid bactericidal activity was consistently sufficient to achieve a good microbiological response in the first few weeks of treatment.[198] However, patients treated in this way usually later experienced symptomatic and microbiological deterioration due

to a *M. tuberculosis* strain resistant to the monotherapy administered.[199–201] Consistent with this observation, meta-analyses have demonstrated that patients treated with first-line treatment for what was later found to be isoniazid mono-resistant TB disease had higher treatment failure and relapse rates than patients with fully drug-susceptible TB, highlighting the need for providing isoniazid susceptibility testing for patients with TB.[202]

Globally, most people notified as having pulmonary TB have a positive result from conventional acid-fast sputum microscopy.[1] The majority of these patients' sputum microscopy result convert to negative when tested after one month of treatment. The foundation study described in Research Chapter 3 demonstrated that there was only a minimal reduction in conventional acid-fast sputum microscopy results over the first 9 days of therapy, whereas concurrent FDA microscopy results were considerably reduced. However, the samples in the foundation study were: collected under optimal conditions with pooled, prolonged sputum collection; that were rapidly transported to the research laboratory; and processed using a protocol that included sputum concentration by centrifugation. These operational characteristics are likely to be unfeasible in the resource-constrained settings where most TB occurs globally.

The current study aimed to assess FDA microscopy under operational conditions that were expected to increase variability in the quality of the sputum samples and potentially the metabolic state of *M. tuberculosis* within those samples, both of which may be expected to affect FDA staining results. Therefore, in the present study we decided empirically to collect the treatment sample after 14 rather than 9 days treatment (as was done during the foundation study described in Research Chapter 3) in order to provide more time for treatment to potentially have a consistent affect upon FDA microscopy results. Furthermore, the exact counts of mycobacteria visualised by microscopy and the quantitative calculations described in Research Chapter 3 would be difficult to reproduce for a routine test in busy clinical laboratories. Therefore, in the current study we aimed to evaluate a simple binary result to assess response in FDA microscopy.

Research Chapter 3 demonstrated that FDA microscopy identified individuals who were not responding to inappropriate first-line therapy for drug-resistant TB (because their drug-resistance had not yet been diagnosed). The purpose of the current study was not only to validate the findings from the foundation study, but also to extend them. Drug resistance and incomplete adherence to treatment are frequent reasons for poor response to treatment. However, there are also other

causes for poor treatment response that are not readily identified, such as drug malabsorption, so assessing drug resistance and monitoring treatment adherence are inadequate for predicting treatment response.[203] Therefore, the primary objective of the current study was to evaluate the role of FDA microscopy performed in the first 2 weeks of TB treatment for predicting whether or not treatment achieved long-term TB cure.

Methods

Ethics. All participants were recruited within the PREVENT TB study (http://www.isrctn.com/ISRCTN17820976) and chose to provide informed written consent. The study was approved by international ethics committees including: the Asociacion Benefica PRISMA, Peru; Imperial College London Research Ethics Committee, UK; and the Peruvian Ministry of Health DIRESA Callao. The clinical significance of FDA microscopy results was uncertain, so these results were not shared with healthcare providers and therefore did not have a role in the care provided. All clinically relevant results of nationally approved drug susceptibility tests from research samples, such as GeneXpert results were provided to the participants and their healthcare provider. TB care was provided free of charge, and participation in the study did not delay treatment or clinical decision making.

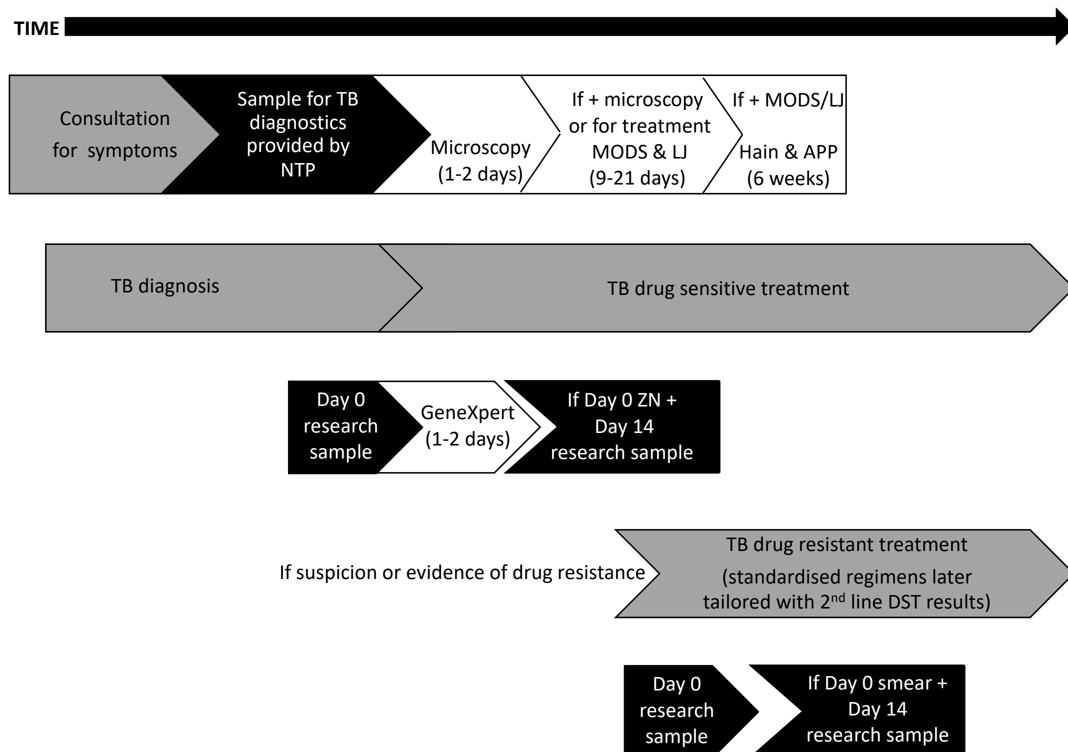
Participants. All patients diagnosed with TB in the 32 participating health clinics in the shantytowns and urban settlements in Callao, Peru between 13 July 2016 and 14 December 2017 were invited to participate in the PREVENT TB cohort study.[204] Patients who were known to have a positive conventional acid-fast sputum microscopy result and had not yet started treatment were then invited to participate in the current study.

Sample collection and interview. For each participant, a pre-treatment sputum sample was collected for this research in the days prior to, or on the day of treatment initiation that was referred to as the "Day 0" sample. Another research sample was then collected 14 days (+/- 5 days) after starting treatment, referred to as the "Day 14" sample. Sputa were collected using the "instructed spot" collection method, as shown in Figure 6.1. Sample processing, reading, and quantification for ZN, FDA, culture and PCR were as described in Research Chapter 6. PCR was not performed on Day 14 samples. At the time that each research sample was collected, participants were interviewed to record demographic data and symptoms. Figure 7.1 demonstrates the sample pathway and diagnostics provided in the study setting in the context of the research samples collected on Day 0 and Day 14.

Drug-susceptibility testing. All positive cultures from research samples underwent drug-susceptibility testing for isoniazid, rifampicin, ethambutol and ciprofloxacin with TEMA.[59] Additionally, as described in Research Chapter 6 and shown in Figure 7.1, Day 0 sputum samples were also tested for *M. tuberculosis* detection and rifampicin resistance using the GeneXpert cartridge-based PCR method.[205] The Peruvian NTP also usually routinely tested their diagnostic pre-

treatment sample with the MODS technique and if positive, *M. tuberculosis* strains underwent indirect drug-susceptibility testing with the GenoType MTBDR*plus* (Hain Lifescience GmbH, Germany) and the proportions method (Figure 7.1).[151]

Figure 7.1 Sample pathway
Samples collected are shown in black whereas the subsequent tests carried out on each sample are shown in white, with the turnaround time in parenthesise. Note. NTP=national TB programme, +=positive, MODS=microscopic observational drug susceptibility test, LJ=Lowenstein-Jenson, APP=proportions method, ZN=Ziehl-Neelsen and DST=drug susceptibility test.



TB treatment. The date of treatment initiation and the treatment scheme provided were noted when each research sample was collected, and this information was confirmed by referring to the NTP treatment card. Patients commenced empiric first-line TB treatment with isoniazid, rifampicin, ethambutol and pyrazinamide, unless they had proven or a high suspicion of drug resistance. The NTP provided treatment 6 days per week throughout the intensive phase of first-line treatment that lasted at least 8 weeks (unless serious adverse reactions occurred, patients interrupted their treatment for other usually psycho-socio-economic reasons or if drug-resistance was detected). Generally, only isoniazid and rifampicin drug-susceptibility results become available within the first 8 weeks of commencing therapy. Isoniazid mono-resistant TB (i.e. resistant to isoniazid but susceptible to rifampicin) was routinely treated by substituting levofloxacin for isoniazid. Rifampicin-resistant TB (whether rifampicin mono-resistant or MDR TB) was

treated with a fluoroquinolone, an aminoglycoside and 3 other drugs guided by drug susceptibility testing results.[151] All treatment was provided under daily (6-days per week) direct observation in local health centres.

Treatment groups. Rifampicin, isoniazid and fluoroquinolones have been shown to rapidly kill *M. tuberculosis* and are referred to here as "key drugs".[206,207] The adequacy of treatment was characterised as the number of these key drug that the patient received and their *M. tuberculosis* was susceptible to (Table 7.1).

Table 7.1 Definition of treatment groups. Note. N=number, H=isoniazid, R=rifampicin, Z-pyrazinamide, E=ethambutol, and FLQ=fluoroquinolone

Treatment scheme received	*Mycobacterium tuberculosis* susceptibility	N of key drugs		
		2	**1**	**0**
First-line H/R/Z/E	Susceptible to H and R	X		
	Susceptible to H, resistant to R		X	
	Susceptible to R, resistant to H		X	
	Resistant to R and H			X
H-resistant FLQ/R/Z/E	Susceptible to FLQ and R	X		
	Susceptible to FLQ, resistant to R		X	
	Resistant to FLQ and R			X
R-resistant FLQ + aminoglycoside	Susceptible to FLQ		X	
	Resistant to FLQ			X

Microbiological response assessment. For the samples that were less than 5 ml in volume, PBS (pH 6.8) was added to make the sample a final volume of 5 ml to ensure that there was sufficient sample for all tests. Therefore, to ensure that Day 0 and Day 14 results were comparable, all counts were adjusted for the volume diluted with PBS. As per WHO recommendations ≥1 bacilli per hundred fields in microscopy was considered a positive test in ZN and FDA microscopy, whereas the local standard of ≥10 CFU was used to indicate a positive culture in the context of monitoring treatment response.[151] An "alert" result was defined as a positive test at Day 14 with a less than a 2 smear grade reduction i.e. less than a 100-fold fall in bacilli or CFU counts. Conversely, an "no alert" result was defined as a Day 14 result that was negative or had reduced by 2 smear grades e.g. from '+++' at Day 0 to '+' at Day 14[36].

Treatment outcomes were determined by the Peruvian NTP who shared them with our research team, together with the results of laboratory tests they performed. Relapse of TB disease was identified from the NTP treatment records. This information was supplemented with a follow-up interview 2 years after patient recruitment at which patients were asked if they had recommenced TB treatment and were all (independently of if they had symptoms) asked to provide a sputum sample (or saliva if sputum was unavailable). These sputum samples were tested

with: conventional acid-fast sputum microscopy; and culture with thin-layer agar; and if positive in either, stored aliquots of samples were reassessed with the GeneXpert PCR test. Participants who had positive results confirming the presence of *M. tuberculosis* in their sputum sample were referred to the NTP for assessment and if were thought to have active TB disease was considered to have suffered a TB relapse, whether or not the patient commenced TB re-treatment.

Treatment outcomes. The analysis of treatment outcomes considered the following possible categories:

- Participants whose treatment was suspended either due to adverse reactions to treatment or more commonly due to loss to treatment follow-up (locally termed "abandoned treatment") were considered to have unknown treatment outcome and were excluded from treatment outcome analyses because this outcome was usually believed to result from psycho-socio-economic factors, not the adequacy of the drugs prescribed.
- Adverse outcome was defined as death during TB treatment, treatment failure, treatment escalation (due to drug resistance or inadequate clinical response), or recurrence of disease within 24 months of recruitment.
- The alternative to adverse treatment outcome was long term cure, which was defined as successful treatment (cure or treatment completion) without recurrence within 24 months of recruitment.

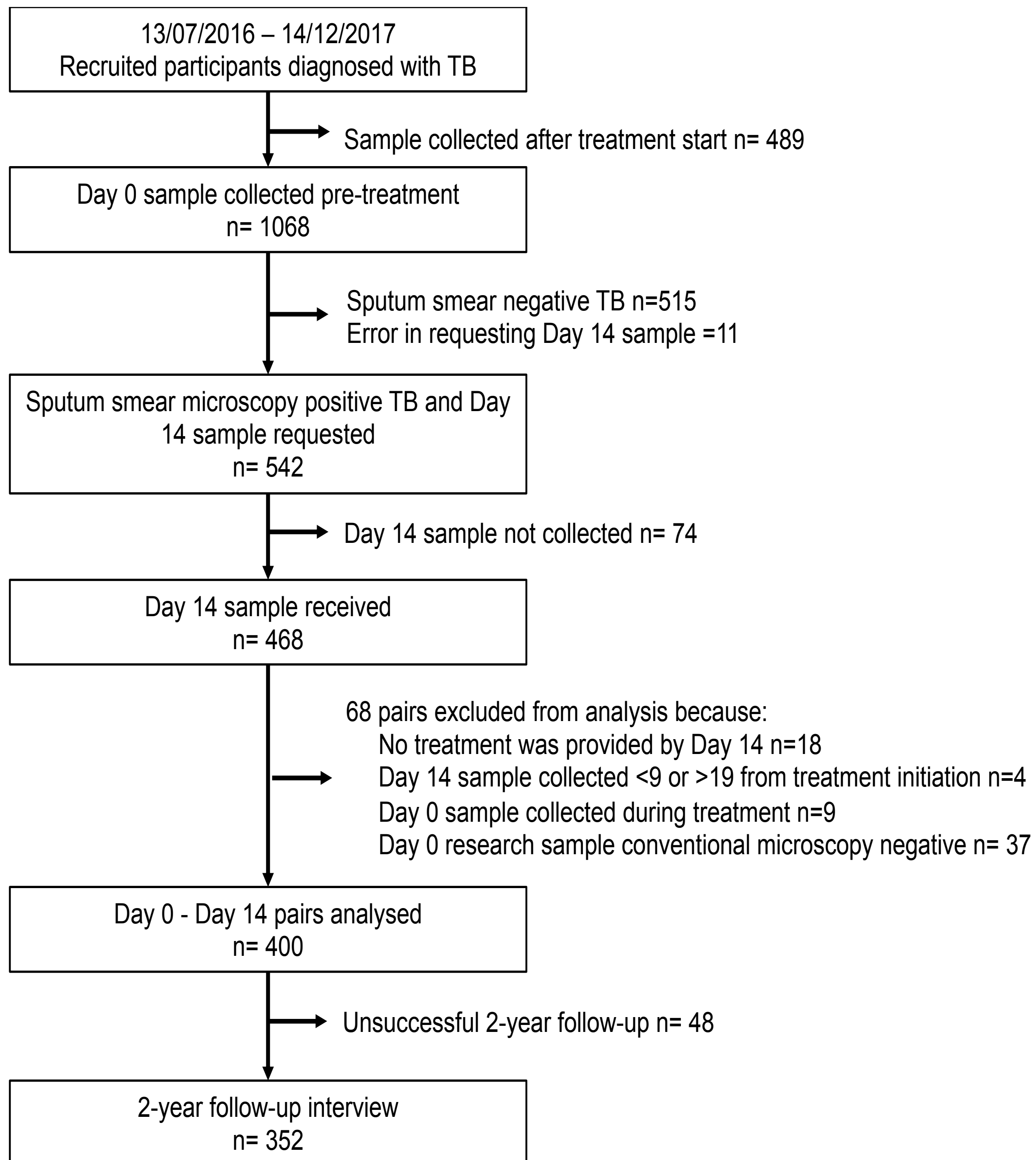
Analysis. Data were analysed with Stata version 13 (StataCorp, College Station, TX). Chi-squared analysis was used to compare binary variables, and 2 x 2 contingency tables were used to evaluate the performance of Day 14 markers to identify patients who went on to have an adverse treatment outcome.

Sample size. Using data from Research Chapter 3, it was assumed that 20% of patients with drug-susceptible TB disease would have an alert FDA microscopy result 14 days after treatment, whereas 50% of those with adverse outcome (20% of the study population) would have an alert FDA microscopy result. Therefore, for 80% power and alpha=0.05, 280 participants were required, which would give proportion estimates a precision of +/- 1%. Taking into consideration that 19% of the study population were expected to suspend treatment,[208] at least 350 participants were needed to be recruited into the study.

Results

Patients. Figure 7.2 shows the study flow chart. Over the study period, 382 participants recruited for this study provided 400 pairs of Day 0 and Day 14 sputa for analysis (because 18 participants had multiple treatment episodes or schemes). The characteristics of the study participants are shown in Table 3.2. All participants had TB disease confirmed by the presence of *M. tuberculosis* in culture and/or GeneXpert testing in sputum samples collected for this research or by the NTP.

Figure 7.2 Study flow chart



TB treatment schemes. Most sputum pairs were collected during first-line treatment (91%, 362/400), 1.3% (5/400) during isoniazid mono-resistant TB treatment, and 8.3% (33/400) during rifampicin resistant TB treatment. All participants had drug-susceptibility test results and Table 7.2 shows the number of sputum pairs collected categorised by the treatment scheme provided to the patient and the drug susceptibility of the *M. tuberculosis* strain. There were 81% (325/400) of sputum pairs collected when 2 active key drugs were provided, 15% (58/400) with 1 key active drug was provided and 4.3% (17/400) without any key drugs provided in the first 14 days of treatment

Table 7.2 Patient demographic characteristics and treatment
Information characterising participants and treatment provided during sample collection categorised by *Mycobacterium tuberculosis* susceptibility. Note. IQR=interquartile range, BMI=body mass index, BCG= Bacille Calmette-Guerin, H=isoniazid, R=rifampicin, Z=pyrazinamide, E=ethambutol, FLQ=fluoroquinolone and IV= intravenous or intramuscular aminoglycoside.

Participants (N=384)		
Age, median (IQR)		29 (22, 44)
Sex, % male (n)		66% (253)
Previous TB, % (n)		19% (71)
Self-reported HIV, % (n)		3.9% (15)
Self-reported diabetes mellitus		6.0% (23)
BMI in ≥18 years old, median (IQR)		22 (20, 24)
Previous vaccination with BCG, % with any visible BCG scar (n)		84% (319)
Day 0 and Day 14 sputum sample pairs (N=400)		
First-line treatment with H+R+Z+E	susceptible to H and R % (n)	80% (321)
	susceptible to R & resistant to H, % (n)	4.3% (17)
	susceptible to H & resistant to R, % (n)	2.2% (9)
	MDR/pre-XDR, % (n)	3.8% (15)
H-resistant treatment with R+FLQ+Z+E	susceptible to FLQ and R, % (n)	1.0% (4)
	susceptible to FLQ & resistant to R, % (n)	0.25% (1)
R-resistant treatment with FLQ+ IV + more	susceptible to FLQ % (n)	7.9% (31)
	resistant to FLQ, % (n)	0.50% (2)

Day 0. Overall, 92% (367/400) of Day 0 samples were culture-positive according to our definition of ≥10 CFU, 94% (375/400) had >1 CFU detected in culture, and 82% (326/400) were FDA microscopy-positive. Most participants at Day 0 reported a productive cough (86%, 341/399) and 79% (314/397) reporting a frequent cough (defined as coughing at least hourly).

TB treatment outcomes are shown in Figure 7.3. Treatment outcomes categorised by the number of key drugs prescribed during the interval between the Day 0 and the Day 14 sputum collections are shown in Table 7.3. People who did not complete treatment due to loss to follow-up (63/400) were excluded from treatment outcome analyses. There were 15% (52/337) adverse outcomes constituting: 4 deaths due to TB; 6 cases of disease recurrence after successful treatment (all of whom were prescribed first-line treatment for drug-susceptible TB); 7 treatment failures; and 35 patients who required treatment escalation. Most treatment escalation occurred because of resistance to isoniazid and/or rifampicin, as shown in Table 7.3. Thus, treatment escalation was most likely in patients who had evidence of resistance to one or more of the key drugs provided.

Figure 7.3 Pie chart of treatment outcomes
Participants who suspended treatment due to loss to follow-up (n=63) were excluded.

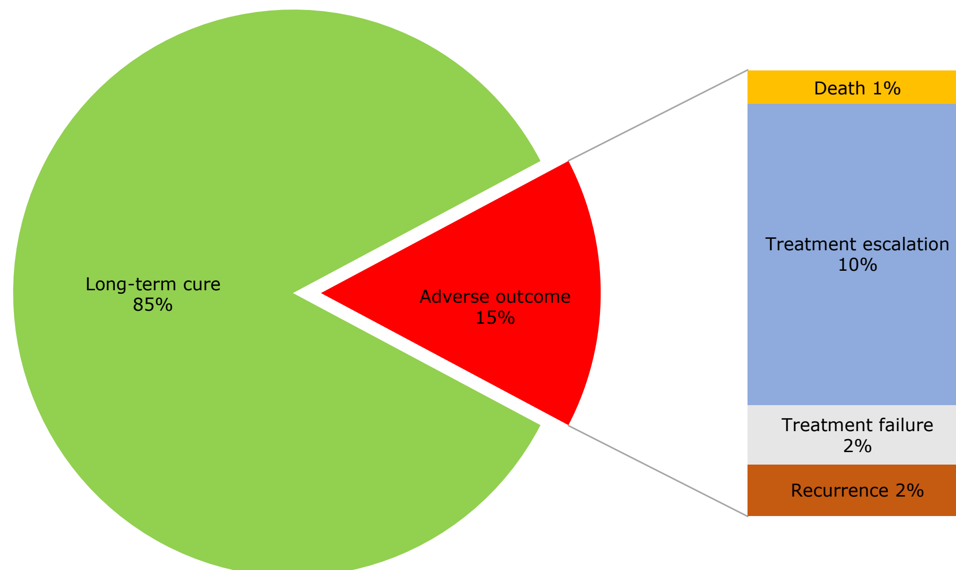


Table 7.3 Treatment outcome categorised by the number of key drugs prescribed during the interval between the Day 0 and the Day 14 sputum collections. People who did not complete treatment due to loss to follow-up (n=63) were excluded from this analysis. Note "key drugs" were considered to be rifampicin, isoniazid and a fluoroquinolone to which the patient's TB was susceptible; n=sample size; and percentages are for the data in each column.

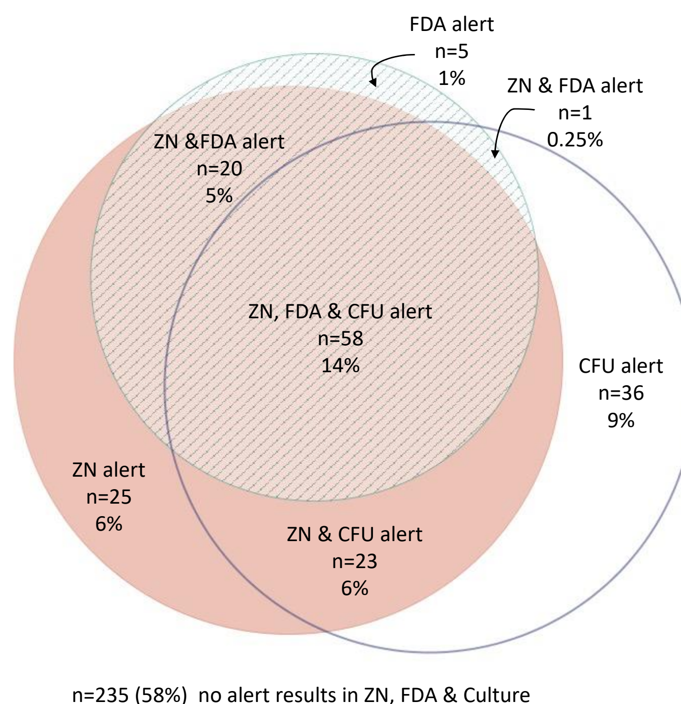
	2 key drugs n=276	1 key drug n=44	0 key drugs n=17
Long-term cure	94% (259)	57% (25)	5.9% (1)
Treatment escalation	1.1% (3)	39% (17)	88% (15)
Failed treatment	1.4% (4)	4.5% (2)	5.9% (1)
Died	1.4% (4)	0	0
Recurrence	2.2% (6)	0	0

Day 14. The median number of days of treatment received before the Day 14 sample collection was 14 days and 95% (379/400) of participants had received 13-16 days of treatment. Table 7.4 shows the proportion of positive results at Day 14 and the proportion of alert results at Day 14 in FDA microscopy (21%, 84/400), ZN microscopy (32%, 126/400) and culture (30%, 118/400). Figure 7.4 shows that 58% (235/400) of samples were negative in all tests, and out of the 165 samples that had any positive result, 62% (102/165) were positive in 2 or more tests. At the time of Day 14 sputum collection, 41% (156/379) of patients reported having a frequent cough, and 2.9% (11/385) stated that their symptoms had worsened.

Table 7.4 Day 14 laboratory results and reported symptoms (N=400)
For the definition of "alert" results in Ziehl-Neelsen (ZN) microscopy, fluorescein diacetate (FDA) microscopy and colony forming units (CFU) in culture please see Methods. Note 95% CI= 95% confidence intervals.

Day 14	n	% (95% CI)
Positive ZN microscopy	160	40% (35%, 45%)
ZN microscopy alert result	126	32% (27%, 36%)
Positive FDA microscopy	96	24% (20%, 28%)
FDA microscopy alert result	84	21% (17%, 25%)
Positive culture	173	43% (38%, 48%)
CFU in culture alert result	118	30% (25%, 34%)
Frequent cough (≥hourly)	156	41% (36%, 46%)
Worsened symptoms	11	2.9% (1.4%, 5.0%)

Figure 7.4 Venn diagram of day 14 laboratory results
For the definition of "alert" results in Ziehl-Neelsen (ZN) microscopy, fluorescein diacetate (FDA) microscopy and colony forming units (CFU) in culture please see Methods.



2-month NTP laboratory result. A total of 85% (306/362) of all participants received first-line treatment for more than 2 months and did not later become lost to follow-up. Within this group with analysable data, 3.2% (10/306) still had ZN microscopy-positive and 2.9% (9/306) culture-positive results from the 2-month sample collected by the NTP.

Predictors of adverse treatment outcome. A Day 14 alert result in FDA microscopy was associated with a 3.0-times relative risk of adverse outcome (95% CI=1.9, 4.9, p<0.0001, Figure 7.5). Adverse outcome was also associated with Day 14 alert results for ZN microscopy (relative risk=1.6, 95% CI=1.0, 2.7, p=0.05) but not for culture (p=0.1, Figure 7.5). Symptom assessment at Day 14 and 2-month NTP ZN microscopy and culture positivity in were also associated with adverse outcomes (Figure 7.5). As shown in the tables and figures, Day 14 symptom worsening and 2-month ZN microscopy or culture positivity had the highest relative risks in predicting adverse outcomes but were so rare that they would infrequently have the capacity to influence patient care and had such wide confidence intervals that their relative risks did not differ significantly from FDA microscopy, so they are not described further here.

Test performance to predict treatment outcome. Figure 7.6 and Table 7.5 show how each test performed for predicting adverse outcome. An alert result in FDA microscopy had the best accuracy (true positive and true negative results, 78%, 95% CI=73%, 82%) out of all Day 14 laboratory tests: with a specificity of 85% (95% CI=80%, 89%) and sensitivity of 42% (95% CI=29%, 57%). As shown in Figure 7.6, the low sensitivity of all Day 14 and 2-month tests would miss the majority of individuals who later had an adverse treatment outcome. Figure 7.7 shows that the positive and negative post-test probability of each of the Day 14 and 2-month markers were similar to the pre-test probability and therefore had limited capacity to provide clinical benefit.

Figure 7.5 Bar graph of adverse outcomes by laboratory and clinical results

Results are compared to a hypothetical test that would identify all adverse outcomes perfectly, as shown in the first bar and referred to as the "perfect predictor". The bars are labelled with the frequency of each adverse outcome and patients whose treatment outcome was loss to follow-up (locally termed "abandoned") were excluded from analysis. Relative risks, 95% confidence intervals (95% CI) and associated p values for each test are shown at the top of the graph. For the definition of an alert result in Ziehl-Neelsen (ZN) microscopy, fluorescein diacetate (FDA) microscopy, and colony forming units (CFU) in quantitative culture please see Methods. Note. n= number, which is reduced for symptoms by missing data and for 2-month ZN and culture results by including only those receiving first-line treatment, some of whom already having outcomes or being lost to follow-up by the end of 2 months (see text).

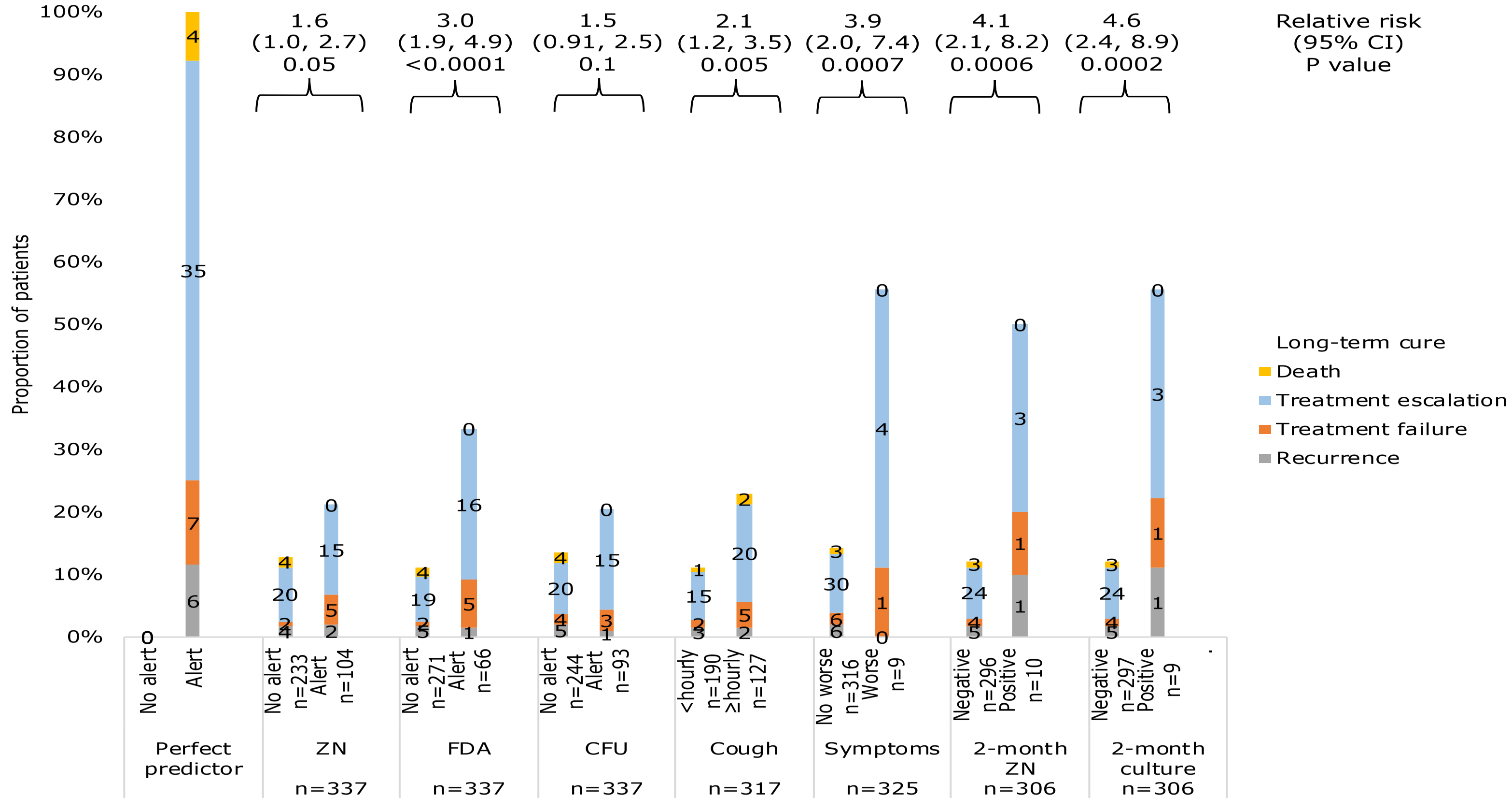
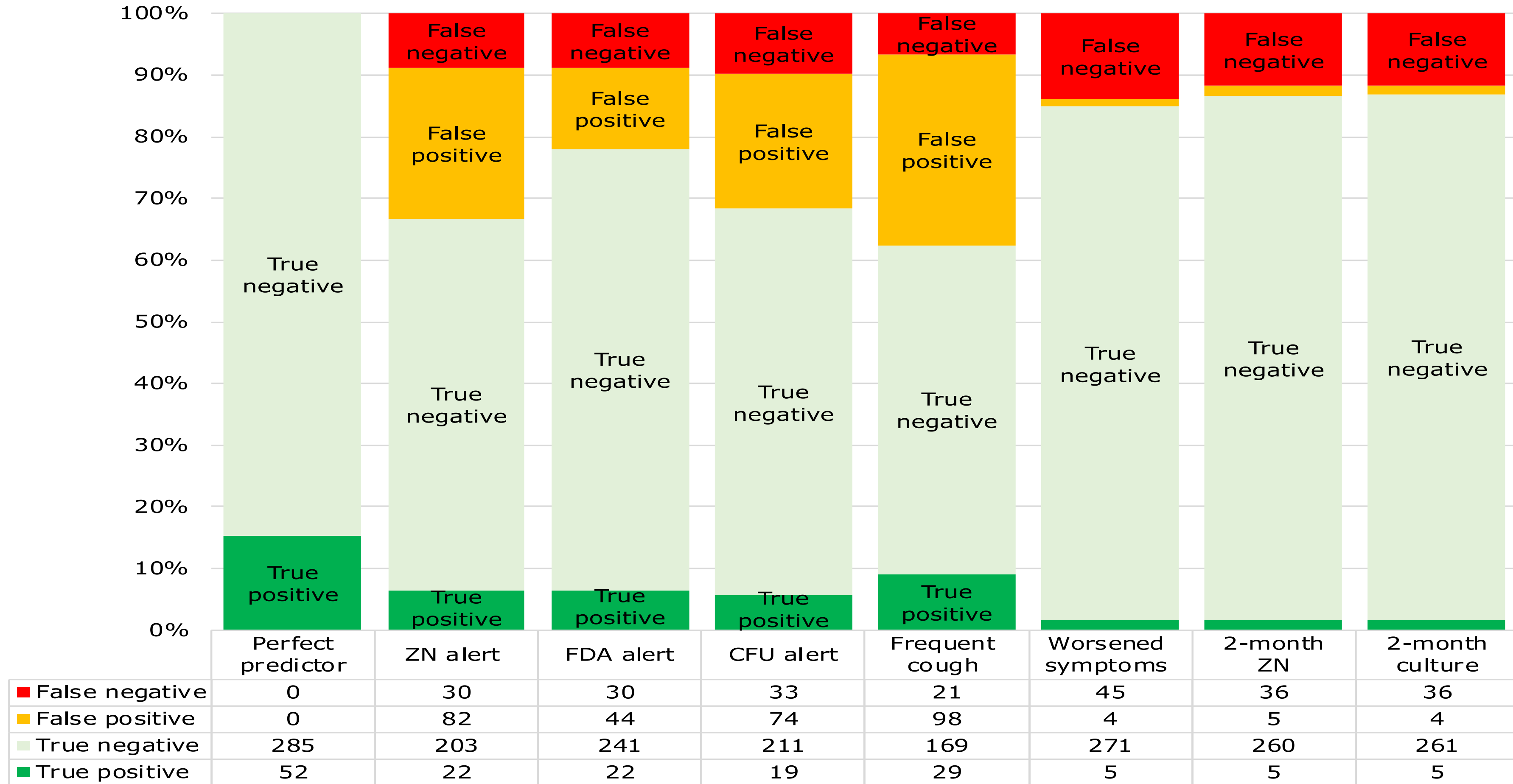


Figure 7.6 Contingency table results

Stacked plot of the percentage of true positive, false positive, true negative and false negative results for each test compared to a hypothetical test that would identify all adverse outcomes perfectly, as shown in the first stacked plot and referred to as the "perfect predictor". The data table below shows the frequencies of each from the 2 x 2 contingency tables. Note. Patients whose treatment outcome was loss to follow-up (locally termed "abandoned") were excluded from analysis.

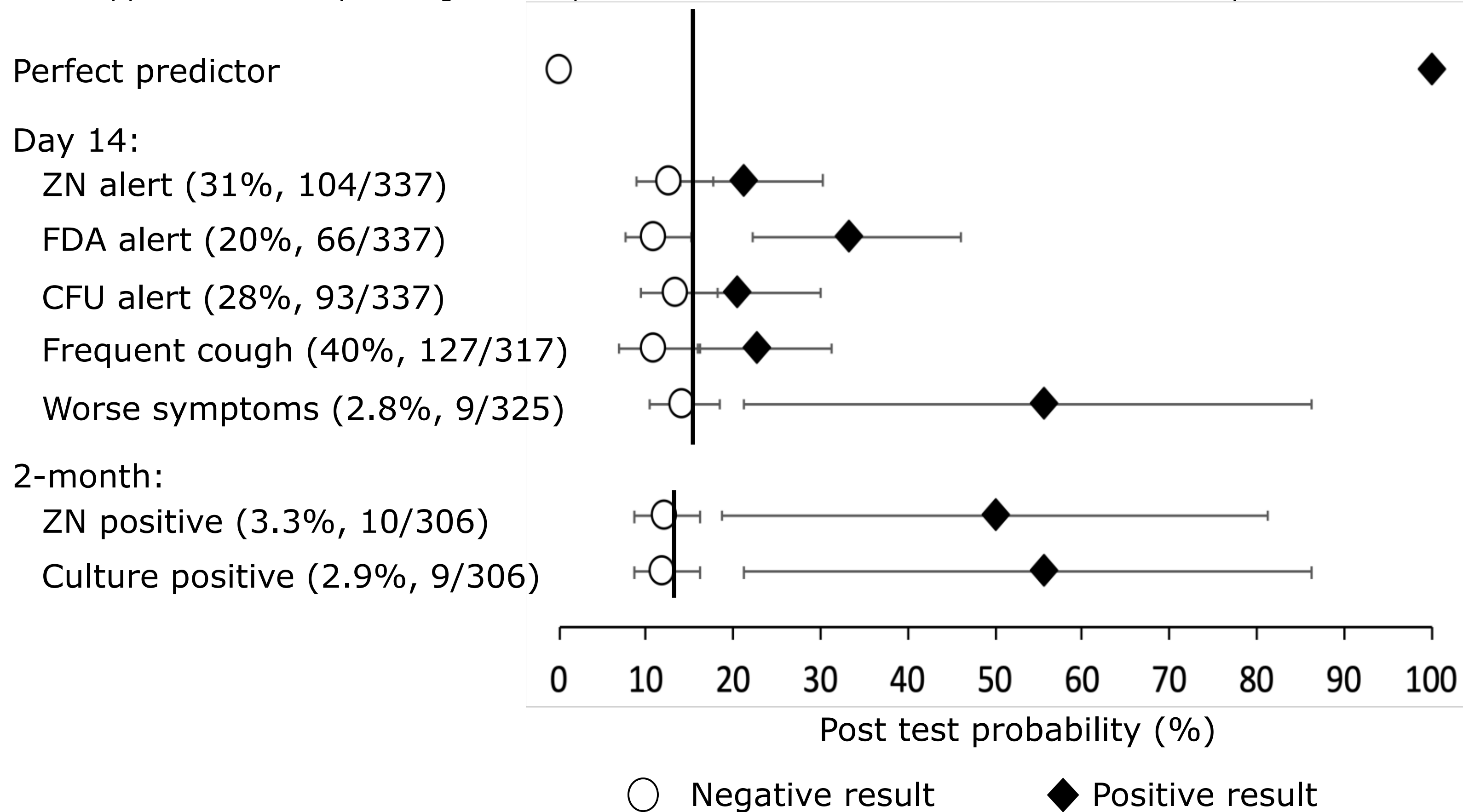


	Perfect predictor	ZN alert	FDA alert	CFU alert	Frequent cough	Worsened symptoms	2-month ZN	2-month culture
False negative	0	30	30	33	21	45	36	36
False positive	0	82	44	74	98	4	5	4
True negative	285	203	241	211	169	271	260	261
True positive	52	22	22	19	29	5	5	5

Table 7.5 Test performance for predicting adverse outcome
Adverse outcome is a composite outcome of: treatment escalation, death, treatment failure, and disease recurrence (n=52) versus long-term cure (n=285). Results are compared to a hypothetical test that would identify all adverse outcomes perfectly and referred to as the "perfect predictor". 95% confidence intervals are shown in parenthesise. Note. Patients whose treatment outcome was loss to follow-up (locally termed "abandoned") were excluded from analysis. ZN=Ziehl-Neelsen microscopy, FDA= fluorescein diacetate microscopy, CFU= colony forming unit in quantitative culture, PPV=positive predictive value, NPV=negative predictive value, LR+=positive likelihood ratio, and LR-=negative likelihood ratio. For the definition of "alert" results see Methods.

Day 14	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**LR+**	**LR-**	**Accuracy**
Perfect predictor	100%	100%	100%	100%	High	Low	100%
ZN alert	42% (29%, 57%)	71% (66%, 76%)	21% (14%, 30%)	87% (81%, 91%)	1.5 (1.0, 2.1)	0.81 (0.63, 1.0)	67% (61%, 72%)
FDA alert	42% (29%, 57%)	85% (80%, 89%)	33% (22%, 46%)	89% (85%, 92%)	2.7 (1.8, 4.2)	0.68 (0.54, 0.87)	78% (73%, 82)
CFU alert	37% (24%, 51%)	74% (69%, 79%)	20% (13%, 30%)	87% (82%, 91%)	1.4 (0.94, 2.1)	0.86 (0.69, 1.1)	77% (73%, 82%)
Frequent cough	58% (43%, 72%)	63% (59%, 69%)	23% (16%, 31%)	89% (84%, 93%)	1.6 (1.2, 2.1)	0.66 (0.47, 0.93)	62% (57%, 68%)
Worsened symptoms	10% (3.3%, 22%)	99% (96%, 99%)	56% (21%, 86%)	86% (81%, 89%)	6.9 (1.9, 25)	0.91 (0.83, 1.0)	85% (81%, 89%)
2-month							
ZN positive	12% (4.1%, 26%)	98% (96%, 99%)	50% (19%, 81%)	88% (83%, 91%)	6.5 (2.0, 21)	0.89 (0.80, 1.0)	87% (82%, 90%)
Culture positive	12% (4.1%, 26%)	99% (96%, 99%)	56% (21%, 86%)	88% (83%, 91%)	9.0 (2.5, 33)	0.89 (0.79, 1.0)	87% (83%, 91%)

Figure 7.7 Graphical display of post-test probabilities of an adverse treatment outcome
Results are compared to a hypothetical test that would identify all adverse outcomes perfectly and referred to as the "perfect predictor". The vertical line shows the pre-test probability (prevalence) of an adverse treatment outcome for all participants at 15% and for those receiving first-line treatment for ≥2 months at 13%. Note. Patients whose treatment outcome was loss to follow-up (locally termed "abandoned") were excluded from analysis. Error bars are 95% confidence intervals, ZN=Ziehl-Neelsen microscopy, FDA=fluorescein diacetate microscopy and CFU=colony forming units in quantitative culture. For definition of an "alert" result please see Methods.



Discussion

This study evaluated changes over the first 14 days of TB therapy in FDA sputum microscopy for predicting treatment outcome. Compared with previous research (including that described in Research Chapter 3), the current study used simplified sputum collection and FDA testing assessed under experimental conditions that were similar to routine clinical operations. An alert result in FDA microscopy 14 days after starting treatment was associated with a higher risk of adverse clinical outcome. This had greater potential clinical utility than the other microbiological and clinical tests assessed at Day 14 or at 2 months of treatment, which is noteworthy because microbiological conversion at 2 months is currently widely used in clinical practice. However, none of these tests performed well enough to be recommended as a stand-alone screening or diagnostic test for predicting adverse treatment outcome.

In this study we collected sputum pairs during all treatment schemes, including during rifampicin resistant TB treatment. There is evidence that microbiological response to second-line treatment for rifampicin resistant treatment is generally slower than for first-line treatment and treatment schemes for isoniazid resistant TB.[209–211] We found that in this setting where standardised treatment schemes were provided, Day 14 microbiological responses were influenced by the number of key drugs provided and resistance of *M. tuberculosis* to these key drugs leading to treatment escalation, rather than the actual treatment scheme itself. Thus, measuring microbiological response at Day 14 cannot reliably identify resistance to any one drug within a treatment scheme consisting of multiple drugs, and is in keeping with a recent study in Mali that demonstrated that microscopy positive results with FDA or ZN at 2 months of treatment performed poorly in identifying rifampicin resistance.[212] However the need for tests performed after weeks or months of treatment to identify those at risk of resistance is being reduced by the increasing availability of drug susceptibility results at the time treatment commences.

FDA microscopy results at Day 14 of treatment predicted adverse outcome. The most frequent adverse outcome was treatment escalation due to drug resistance. However, the relationship between FDA microscopy results and adverse outcome is likely to also be influenced by other causes of poor treatment response such as treatment adherence, which is an important factor that was not characterised in this study. Review of the participants' treatment cards revealed that adherence over the first 14 days was almost perfect, but these data are not currently available

for formal analysis. Other factors causing poor treatment response such as malabsorption, drug interactions, pharmacokinetics, drug quality, and persister mycobacterial populations can also contribute to poor treatment response, persistent illness and relapse even in participants on seemingly appropriate therapy, but are harder to identify.[213,214] These factors also vary with time, thus a single test early in treatment was previously thought to be unlikely to be able to predict outcomes months to years later. This was supported by evidence from drug trials that shortened treatment regimen for patients who had microbiological conversion at 2 months which was followed by frequent relapses in TB disease.[215] Remarkably, ZN microscopy at Day 14 significantly predicted adverse treatment outcome, contrasting with our previous findings.[20] This difference between our current versus previous research could be explained by differences in sample timing, sputum collection, sample storage, sample processing, or because these acid-fast stains stain different elements of mycobacteria (auramine DNA and RNA versus ZN staining mycolic acids in the cell wall).[216] These contrasting findings in acid-fast microscopy results during early treatment warrant further investigation.

By two months of treatment, continued conventional sputum microscopy or culture positivity i.e. lack of conversion to negative, are used in clinical practice to identify patients likely to have treatment failure or relapse of disease. Compared to other reports, in the present research there were very few positive conventional sputum microscopy or culture results at 2 months of treatment. This could be explained by Peru's excellent NTP, heterogeneity in study populations or due to patients who became lost to follow-up being excluded from treatment outcome analyses.[217] Despite this, the low sensitivity and poor post-test probability of 2-month sputum microscopy or culture conversion as a prognostic marker for adverse treatment outcome is in agreement with results from a meta-analysis.[217] Although evidence shows the poor performance as prognostic markers for adverse outcome, 2-month conventional sputum microscopy and culture conversion continue to be widely used in clinical practice, because they trigger clinical reassessment, medication review, additional laboratory testing; and serve as useful indicators of NTP performance.[218,219] However, 2-month conventional sputum microscopy and culture conversion misses the majority of cases that have an adverse treatment outcome, and focuses limited resources to those who do not require it. In comparison to 2-month conventional microscopy and culture conversion, FDA microscopy at Day 14 had higher sensitivity, acceptable specificity, and similar post-test probabilities. Self-reported cough frequency and symptoms worsening at Day 14 also predicted adverse outcome with similar performance to the laboratory tests. Thus, future

research may aim to create and validate a composite predictor combining FDA microscopy with symptoms at Day 14 to better identify those at risk of an adverse treatment outcome.

There is little evidence to guide clinicians when drug resistance has been excluded but TB disease is not improving despite seemingly appropriate treatment. Often in this scenario, first-line treatment with 2 key drugs may be prolonged or changed empirically to second-line treatment. Adjunctive corticosteroid use has been shown to increase clinical improvement within one month.[220] There is also interest in the use of doxycycline and metformin that have potential to limit or prevent lung damage.[221,222] Therefore, in conjunction with finding biomarkers to predict poor response to treatment, more research is required regarding what to do once these patients have been identified.

There are several limitations to this study. Firstly, the universal use of rapid drug susceptibility testing and prompt second-line treatment provision in Peru caused few patients to continue inappropriate first-line therapy for rifampicin resistant TB until the Day 14 sample was collected. This important improvement in clinical care limited the statistical power of this study to assess inadequate treatment response in this context. Furthermore, clinical outcomes such as death, treatment failure and disease recurrence were rare, preventing independent analysis of predictors of each type of adverse treatment outcome. We therefore propose evaluating FDA microscopy and symptoms assessment on a larger scale in diverse settings where rapid drug susceptibility testing is not universally provided, and/or adverse treatment outcomes are more common. Another limitation is that we were not able to carry out drug susceptibility testing for other drugs provided in the treatment regimens, particularly pyrazinamide. These drugs may have contributed to clinical response, although early bactericidal activity studies have shown that isoniazid and rifampicin are the main drivers of early mycobacterial killing in first-line treatment.[77] Lastly, this study only included patients with initially sputum microscopy positive results, and FDA microscopy was only performed on samples that were ZN microscopy positive because as shown in Research Chapter 6, 95% of ZN microscopy-negative samples are also negative in FDA microscopy. Although adults with sputum paucibacillary pulmonary TB have milder disease, sputum scarce individuals (especially those co-infected with HIV) do have a high risk of adverse treatment outcomes, and therefore sputum based strategies are not appropriate in this population.[176,223,224]

In conclusion, this study demonstrated that FDA microscopy predicted adverse treatment outcome but not sufficiently enough to be recommended for informing clinical decision making.

GENERAL DISCUSSION

Conventional sputum microscopy is the most frequently used laboratory test in TB care globally. However, this test poorly assesses infectiousness and early treatment response. FDA microscopy uses a viability stain, is relatively rapid, inexpensive and simple to perform and is believed to identify metabolically active cells. The objective of this body of work was to optimise sputum FDA microscopy and evaluate whether it provides clinically useful information complementary to the other tests that are widely used in TB care.

FDA microscopy is believed to identify on a unicellular level metabolically active cells (including mycobacteria), because only these cells contain the functioning esterases that are required to convert FDA to a fluorescent form visible by microscopy.[11] In contrast, culture quantifies replicating cells (including mycobacteria) that form colonies.[37] The proof-of-concept study presented in Research **Chapter 1** confirmed the hypothesis that quantitative results in FDA microscopy correlated with the proportion of sputum samples that had not been sterilised, and with the number of colonies derived from culture. This study suggested that FDA and culture were both identifying elements of "living" bacteria. In contrast, conventional acid-fast microscopy results with auramine staining were not affected by sterilisation. The aim of this study was to ensure that the laboratory techniques used were producing the expected results; similar findings have been reported before.[16,44]

Transmission studies, including those using the guinea pig model, have demonstrated that patient infectiousness is weakly associated with conventional sputum microscopy positivity.[50] Furthermore, a small minority of individuals (in one study less than 10% of patients) termed "super spreaders" are extremely infectious (in that study accounting for more than 95% of TB transmission).[50,177] It was hypothesised that positive FDA microscopy would identify the most infectious people. The foundation study described in Research **Chapter 2** recruited 35 patients with strongly positive conventional sputum microscopy results before TB treatment. Their pre-treatment sputum samples demonstrated considerable variability in the proportions of FDA-stained *M. tuberculosis* between patients. This observation was similar to studies reporting the variable proportion of "fat and lazy" (also referred to as "persister-like") *M. tuberculosis* phenotypes in pre-treatment sputum.[47] After a median 6-years of follow-up, contacts of patients with low FDA

microscopy results in pre-treatment sputum were more likely to develop secondary TB. Moreover, contacts of patients who had low FDA microscopy results and numerous CFU in pre-treatment sputum had the greatest risk of developing secondary TB. This finding was unexpected, the opposite of the research hypothesis, and could be explained if persister-like *M. tuberculosis* were most infectious and because of their low metabolic activity were not identified by FDA. Although statistically significant even after adjusting for potentially confounding variables, this finding was fragile because only 9 patients were associated with 13 cases of secondary TB disease in the 209 household contacts, so different results in only one TB-affected household had the potential to change the study conclusions. Indeed, this finding was inconsistent with the later, larger study described in Research Chapter 6 and discussed below. Furthermore a recent study in South Africa found no association between the number of persister-like *M. tuberculosis* present in sputa and culturable cough aerosols, which are a surrogate marker of infectiousness.[183]

The 35 patients recruited in the foundation study described in Research Chapter 2 also provided multiple serial sputum samples over the first 9 days of TB treatment for assessment with microscopy and culture. During this time (2006-2007) it was policy to only offer drug susceptibility testing to individuals who were at high-risk of drug resistance.[225,226] Therefore, another objective of this foundation study that is described in Research **Chapter 3** was to evaluate whether FDA microscopy early in treatment could screen for patients likely to not be responding to first-line treatment due to drug resistance. Although some patients remain sputum microscopy and culture positive a month or more despite adequate treatment,[227] this study demonstrated that for patients with drug susceptible TB, 9 days of adequate first-line treatment was usually enough to reduce by 99% quantitative counts in both FDA microscopy and culture. This early treatment response was absent in patients later found to have MDR TB. Similar to the proof-of-concept study described in Research Chapter 1, conventional microscopy with auramine staining changed only minimally during early first-line treatment for drug susceptible TB. The rate of decline in bacilli counts in FDA microscopy and culture during early treatment were similar to findings from early bactericidal activity and serial sputum colony count studies monitoring response to isoniazid and rifampicin.[77,79,88,89] These findings implied that quantitative FDA microscopy assessments could potentially be used as a biomarker to evaluate the efficacy of novel treatment regimens by providing rapid results and reducing overall costs.[218,228]

An important limitation of all sputum microscopy is its low diagnostic sensitivity, which can be increased by improving sample quality.[99] During the preparation for the operational evaluation of FDA microscopy it was apparent that no grade-A level evidence existed regarding the optimal method to collect sputum, even though the importance of sputum collection techniques was acknowledged and various guidelines existed.[36,98] Furthermore, the foundation study described in Research Chapters 2 and 3 utilised sputum collection techniques that were too complex, time-consuming and inconvenient for operational use. There was therefore a need to clarify optimal sputum collection techniques prior to further evaluation of FDA microscopy. The systematic review and pairwise meta-analysis presented in Research **Chapter 4** demonstrated that compared to the globally practised 'spot' sputum collection, either providing instructions prior to spot collections or collecting a 'pooled' sputum sample over hours increased the odds of diagnosing TB in microscopy and culture by 1.6 to 1.8 times. Without the need for another study, a network meta-analysis using published data collected in the systematic review suggested that there was no difference in yield between instructed spot sputum collections and pooled sputum collections (as had been done in the foundation study). This review also demonstrated that sputum could be collected at any time of day and that early morning sputum collection (which can be a barrier in the patient care pathway) was unnecessary.[21]

Previous FDA staining and microscopy protocols, as were used in the foundation study, presented multiple barriers for implementation as a routine test for laboratories in resource-constrained settings. Research **Chapter 5** describes attempts to evaluate and optimise FDA staining. Evaluation of the different published protocols for FDA microscopy with WHO-approved low-cost LED microscopes demonstrated that: decontamination and centrifugation could be replaced with an acid-alcohol step, reducing equipment and reagent costs whilst improving biosafety. Furthermore, optimisation with an additional biosafety measure with phenol could be used without impacting FDA-stained slide quality; and a fluorescent quencher improved FDA microscopy reading. Data presented in this thesis show that FDA microscopy using the proposed SOP with one slide provided reliable and reproducible results. The optimised FDA protocol used equipment and reagents that other than FDA itself are readily available in clinical laboratories, and materials would cost $0.05 USD per slide. To ensure that clinical laboratories provide reliable service, quality assurance for microscopy is generally carried out by blind reading of externally prepared slides.[160] The storage study

results described in Research Chapter 5 showed that quality assurance of FDA microscopy could be performed in a similar manner. With these refinements and assessments, FDA microscopy was ready for operational evaluation as a routine test in clinical sputum samples.

Data presented in Research **Chapter 6** indicated that the low sensitivity of sputum microscopy could be explained by the high concentrations of acid-fast bacilli necessary for a positive result because only a tiny volume of sputum is examined by microscopy, much less than is tested by culture or PCR. Concentrating samples with filters or centrifugation have been shown to improve the sensitivity of sputum microscopy, but also increase costs and biohazard.[92] In conventional sputum microscopy positive pre-treatment samples there were significantly more ZN-stained bacilli/ml compared to FDA-stained bacilli/ml (or compared to CFU/ml detected in culture). However independent of the differences in threshold of positivity and overall counts, there was a strong correlation between quantitative results in microscopy, culture and PCR suggesting that these tests were predominately identifying the same mycobacterial population in pre-treatment samples.

The operational evaluation of FDA microscopy described in Research **Chapter 6** involved 4311 contacts of patients diagnosed with TB who were followed up for a median 3.5 years. This study demonstrated that higher mycobacterial load in pre-treatment sputum assessed by ZN, FDA, culture or PCR all predicted increased risk of secondary TB in household contacts. This association between mycobacterial load and the risk of secondary TB is in keeping with results from other transmission studies,[229] whereas the fact that ZN or PCR quantitative results had the strongest association is new knowledge. The results of this study clarified that these laboratory tests at the time of diagnosis had poor discriminatory value and performance for predicting secondary TB in contacts, despite the strong statistical association. Therefore, with this data we propose that the current policy to prioritise contact investigations for sputum microscopy positive patients be re-considered and replaced with validated simple assessments that take multiple factors into consideration such as indicators of TB risk and contact proximity scores.[186,187]

Adverse treatment outcome (defined as treatment escalation, treatment failure, death or disease recurrence) was predicted by FDA microscopy results simplified to a binary variable indicating ‘no alert’ versus ‘alert’, 14 days after treatment commenced. Specifically, an ‘alert’ FDA microscopy result predicted treatment

outcome more reliably than Day 14 or 2-month conventional sputum microscopy or culture. The strong statistical associations demonstrated in Research **Chapter 7** for these early indicators of treatment response predicting adverse outcome were not sufficiently reliable to be used as clinical tools for identifying patients requiring additional care. Clinicians use 2-month microscopy or culture results to guide increased testing, treatment prolongation (although this policy was rescinded in the global guidelines published in 2010) or escalation to a second-line treatment regimen (not policy but common practice).[217,230] However, evidence suggests that this strategy may not improve cure rates (because most adverse treatment outcomes occur in patients who had 2-month conversion), and may lead to almost half of patients who did not have 2-month conversion to inappropriately receive additional tests, medications or longer treatment duration, which risk increased costs and impaired adherence, health and quality of life.[141,217,231]

FDA microscopy did strongly and consistently predict infectiousness and adverse treatment outcome, so has potential clinical utility if its performance is increased and/or it is used in conjunction with other clinical markers such as self-reported symptoms and cough monitoring. FDA microscopy predominately adds value to the interpretation of conventional sputum microscopy positive samples. Therefore, continuing investigations to improve the sensitivity of microscopy by sample concentration or automation allowing reading of more fields per slide may also improve FDA microscopy.[92,232] Furthermore, the clinical role of FDA microscopy (or any sputum-based technique identifying viable *M. tuberculosis*) is dependent on enhancing our knowledge of TB pathogenesis. In particular, *in vitro* sputum testing mainly assesses extracellular *M. tuberculosis* expelled from lung cavities and there is a need to increase understanding of how this indicates the status of diverse mainly intracellular sub-populations of *M. tuberculosis in vivo*.[233] With these potential improvements and increased knowledge, there may be a clinical role for FDA microscopy, warranting future evaluation in diverse TB endemic settings

CONCLUSION

This body of work investigated solutions to improve microscopy to better aid decision-making during TB care. This generated new evidence optimising sputum collection techniques for TB diagnosis and defined a simplified and safer SOP for FDA microscopy. These methodological improvements were used to demonstrate that FDA microscopy done in the first days of TB treatment predicted both patient infectiousness and treatment outcome with similar or greater reliability than currently used laboratory tests.

References

1 WHO | Global tuberculosis report 2019. *WHO* 2020.

2 Davis JL, Cattamanchi A, Cuevas LE, Hopewell PC, Steingart KR. Diagnostic accuracy of same-day microscopy versus standard microscopy for pulmonary tuberculosis: a systematic review and meta-analysis. *Lancet Infect Dis* 2013; **13**: 147–54.

3 Piatek a. S, Van Cleeff M, Alexander H, *et al.* GeneXpert for TB diagnosis: planned and purposeful implementation. *Glob Heal Sci Pract* 2013; **1**: 18–23.

4 Evans CA. GeneXpert--a game-changer for tuberculosis control? *PLoS Med* 2011; **8**: e1001064.

5 Kirwan DE, Gilman RH. Same-day diagnosis and treatment of tuberculosis. *Lancet Infect Dis* 2013; **13**: 102–4.

6 WHO | What services are there to diagnose TB? http://www.who.int/gho/tb/epidemic/diagnosis/en/ (accessed Sept 24, 2015).

7 Tostmann A, Kik S V, Kalisvaart NA, Sebek MM, Verver S, Boeree MJ. Tuberculosis Transmission by Patients with Smear- Negative Pulmonary Tuberculosis in a Large Cohort in The Netherlands. 2008; : 1135–42.

8 Rieder HL. Sputum smear conversion during directly observed treatment for tuberculosis. *Tuber Lung Dis* 1996; **77**: 124–9.

9 Kim TC, Blackman RS, Heatwole KM, Kim T, Rochester DF. Acid-fast bacilli in sputum smears of patients with pulmonary tuberculosis. Prevalence and significance of negative smears pretreatment and positive smears post-treatment. *Am Rev Respir Dis* 1984; **129**: 264–8.

10 Kvach J, Veras J. A fluorescent staining procedure for determining the viability of Mycobacterial cells. *Int J Lepr* 1982; **50**: 183.

11 Medzon EL, Brady ML. Direct measurement of acetylesterase in living protist cells. *J Bacteriol* 1969; **97**: 402–15.

12 Rotman B, Papermaster BW. Membrane properties of living mammalian cells as studied by enzymatic hydrolysis of fluorogenic esters. *Proc Natl Acad Sci U S A* 1966; **55**: 134–41.

13 Spark D. Macrophages offer test for resistant leprosy. *New Sci*; : 28.

14 Jarnagin JL, Luchsinger DW. The use of fluorescein diacetate and ethidium bromide as a stain for evaluating viability of mycobacteria. *Stain Technol* 1980; **55**: 253–8.

15 Harada S, Numata N. Application of FDA/EB staining for the detection of viable or non-viable mycobacteria in clinical specimens. *Kekkaku* 1992; **67**: 113–7.

16 Kinomoto M, Taekekawa M, Nakamura RM. Consistency in viabilities of mycobacteria detected by FDA/EB staining and colony-forming units on the medium. *Kekkaku* 1991; **66**: 485–8.

17 Datta S, Sherman JM, Tovar MA, *et al.* Sputum Microscopy With Fluorescein Diacetate Predicts Tuberculosis Infectiousness. *J Infect Dis* 2017; **216**: 514–24.

18 Salim AH, Aung KJM, Hossain MA, Deun A Van. Early and rapid microscopy-based diagnosis of true treatment failure and MDR-TB. *Int J TB Lung Dis* 2006; **10**: 1248–54.

19 Van Deun A, Maug AKJ, Hossain A, Gumusboga M, de Jong BC. Fluorescein diacetate vital staining allows earlier diagnosis of rifampicin-resistant tuberculosis. *Int J Tuberc Lung Dis* 2012; **16**: 1174–9.

20 Datta S, Sherman JM, Bravard M, Valencia T, Gilman RH, Evans CA. Clinical Evaluation of Tuberculosis Viability Microscopy for Assessing Treatment

Response. *Clin Infect Dis* 2014; **ciu 1153**. http://www.ncbi.nlm.nih.gov/pubmed/25537870.

21 MacPherson P, Houben RMGJ, Glynn JR, Corbett EL, Kranzer K. Pre-treatment loss to follow-up in tuberculosis patients in low- and lower-middle-income countries and high-burden countries: a systematic review and meta-analysis. *Bull World Health Organ* 2014; **92**: 126–38.

22 Datta S, Shah L, Gilman RH, Evans CA. Comparison of sputum collection methods for tuberculosis diagnosis: a systematic review and pairwise and network meta-analysis. *Lancet Glob Heal* 2017; **5**: e760–71.

23 Harper GJ. Contamination of the environment by special purpose centrifuges used in clinical laboratories. *J Clin Pathol* 1981; **34**: 1114–23.

24 Stern EL, Johnson JW, Vesley D, Halbert MM, Williams LE, Blume P. Aerosol Production Associated with Clinical Laboratory Procedures. *Am J Clin Pathol* 1974; **62**: 591–600.

25 Datta S, Alvarado K, Gilman RH, *et al.* Optimising fluorescein diacetate sputum smear microscopy for assessing patients with pulmonary tuberculosis. *PLoS One* 2019; **14**: e0214131.

26 Caviedes L, Lee TS, Gilman RH, *et al.* Rapid, efficient detection and drug susceptibility testing of Mycobacterium tuberculosis in sputum by microscopic observation of broth cultures. The Tuberculosis Working Group in Peru. *J Clin Microbiol* 2000; **38**: 1203–8.

27 Moore DAJ, Evans CAW, Gilman RH, *et al.* Microscopic-observation drug-susceptibility assay for the diagnosis of TB. *N Engl J Med* 2006; **355**: 1539–50.

28 Tovar M, Datta S, Saunders M, *et al.* Universalization of rapid MDR-TB testing is associated with decreased mortality. *Int J TB Lung Dis* 2017; **21**: S384-385.

29 Diacon AH, Van Der Merwe L, Demers AM, Von Groote-Bidlingmaier F, Venter A, Donald PR. Time to positivity in liquid culture predicts colony forming unit counts of Mycobacterium tuberculosis in sputum specimens. *Tuberculosis* 2014; **94**: 148–51.

30 Dhillon J, Fourie PB, Mitchison DA. Persister populations of Mycobacterium tuberculosis in sputum that grow in liquid but not on solid culture media. *J Antimicrob Chemother* 2014; : 437–40.

31 Tillett HE. Most probable numbers of organisms: revised tables for the multiple tube method. 1987.

32 Vijay S, Nhung HN, Bao NLH, *et al.* Most-probable number based minimum duration of killing assay for determining the spectrum of rifampicin susceptibility in clinical M. tuberculosis isolates. *Antimicrob Agents Chemother* 2020.

33 Grandjean L, Martin L, Gilman RH, *et al.* Tuberculosis diagnosis and multidrug resistance testing by direct sputum culture in selective broth without decontamination or centrifugation. *J Clin Microbiol* 2008; **46**: 2339–44.

34 Desjardin LE, Perkins MD, Wolski K, *et al.* Measurement of sputum Mycobacterium tuberculosis messenger RNA as a surrogate for response to chemotherapy. *Am J Respir Crit Care Med* 1999; **160**: 203–10.

35 Bustin SA, Nolan T. Pitfalls of quantitative real- time reverse-transcription polymerase chain reaction. J. Biomol. Tech. 2004; **15**: 155–66.

36 Lumb R, Van Deun A, Bastian I, Fitz-gerald M. Laboratory diagnosis of tuberculosis by sputum microscopy - the handbook global edition. 2013 http://www.challengetb.org/publications/tools/lab/TB_Microscopy_Handbook_2013.pdf (accessed Nov 16, 2015).

37 Kent PT. Public health mycobacteriology: a guide for the level III laboratory. Atlanta, Ga: U.S. Dept. of Health and Human Services, Public Health Service, Centers for Disease Control, 1985 https://catalyst.library.jhu.edu/catalog/bib_1853570 (accessed Feb 12,

2014).
38 Blackwood KS, Burdz T V, Turenne CY, Sharma MK, Kabani AM, Wolfe JN. Viability testing of material derived from Mycobacterium tuberculosis prior to removal from a Containment Level-III Laboratory as part of a Laboratory Risk Assessment Program. *BMC Infect Dis* 2005; **5**: 4.
39 Kinomoto M. Development of slide-method to distinguish alive and dead Mycobacteria by flourescent staining. *Kekkaku* 1999; **74**: 599–609.
40 Ministerio de salud. Norma tecnica de salud para el control de la tuberculosis. Lima, Peru, 2006.
41 Shetty P, Ghosh D, Paul D. Thermal lysis and isothermal amplification of Mycobacterium tuberculosis H37Rv in one tube. *J Microbiol Methods* 2017; **143**: 1–5.
42 Sabiiti W, Azam K, Esmeraldo E, Bhatt N, Rachow A, Gillespie SH. Heat Inactivation Renders Sputum Safe and Preserves Mycobacterium tuberculosis RNA for Downstream Molecular Tests. *J Clin Microbiol* 2019; **57**. DOI:10.1128/JCM.01778-18.
43 Doig C, Seagar AL, Watt B, Forbes KJ. The efficacy of the heat killing of Mycobacterium tuberculosis. *J Clin Pathol* 2002; **55**: 778–9.
44 Burdz TVN, Wolfe J, Kabani A. Evaluation of sputum decontamination methods for Mycobacterium tuberculosis using viable colony counts and flow cytometry. *Diagn Microbiol Infect Dis* 2003; **47**: 503–9.
45 Barer MR, Harwood CR. Bacterial viability and culturability. Adv. Microb. Physiol. 1999; **41**: 93–137.
46 Zhang Y, Yew W, Barer MR. Targeting Persisters for Tuberculosis Control. *Antimicrob Agents Chemother* 2012; **56**: 2223–30.
47 Garton NJ, Waddell SJ, Sherratt AL, *et al.* Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum. *PLoS Med* 2008; **5**: e75.
48 Escombe AR, Oeser C, Gilman RH, *et al.* The detection of airborne transmission of tuberculosis from HIV-infected patients, using an in vivo air sampling model. *Clin Infect Dis* 2007; **44**: 1349–57.
49 Nahid P, Kim PS, Evans CA, *et al.* Clinical research and development of tuberculosis diagnostics: moving from silos to synergy. *J Infect Dis* 2012; **205 Suppl**: S159-68.
50 Escombe AR, Moore D a J, Gilman RH, *et al.* The infectiousness of tuberculosis patients coinfected with HIV. *PLoS Med* 2008; **5**: e188.
51 Bacon J, James BW, Wernisch L, *et al.* The influence of reduced oxygen availability on pathogenicity and gene expression in. *tuberculosis* 2004; : 205–17.
52 Ashiru OT, Pillay M, Sturm AW. Mycobacterium tuberculosis isolates grown under oxygen deprivation invade pulmonary epithelial cells. *Anaerobe* 2012; **18**: 471–4.
53 Li Y-J, Petrofsky M, Bermudez LE. Mycobacterium tuberculosis uptake by recipient host macrophages is influenced by environmental conditions in the granuloma of the infectious individual and is associated with impaired production of interleukin-12 and tumor necrosis factor alpha. *Infect Immun* 2002; **70**: 6223–30.
54 Dietrich J, Roy S, Rosenkrands I, *et al.* Differential influence of nutrient starved Mycobacterium tuberculosis on adaptive immunity results in progressive TB disease and pathology. *Infect Immun* 2015; **83**: IAI.01055-15.
55 Onifade DA, Bayer AM, Montoya R, *et al.* Gender-related factors influencing tuberculosis control in shantytowns: a qualitative study. *BMC Public Health* 2010; **10**: 381.
56 Kawai V, Soto G, Gilman RH, *et al.* Tuberculosis mortality, drug resistance, and infectiousness in patents with and without HIV infection in Peru. *Am J Trop Med Hyg* 2006; **75**: 1027–33.

57 Saito M, Bautista CT, Gilman RH, Bowering A, Levy MZ, Evans CA. The value of counting BCG scars for interpretation of tuberculin skin tests in a tuberculosis hyperendemic shantytown, Peru. *Int J Tuberc Lung Dis* 2004; **8**: 842–7.
58 Toman K. Toman's tuberculosis; case detection, treatment and monitoring - questions and answers, 2nd edn. Geneva, Switzerland: World Health Organisation, 2004.
59 Caviedes L, Delgado J, Gilman RH. Tetrazolium microplate assay as a rapid and inexpensive colorimetric method for determination of antibiotic susceptibility of Mycobacterium tuberculosis. *J Clin Microbiol* 2002; **40**: 1873–4.
60 Rocha C, Montoya R, Zevallos K, *et al.* The Innovative Socio-economic Interventions Against Tuberculosis (ISIAT) project: an operational assessment. *Int J Tuberc Lung Dis* 2011; **15 Suppl 2**: S50-7.
61 Martinez L, Arman A, Haveman N, *et al.* Changes in tuberculin skin test positivity over 20 years in periurban shantytowns in Lima, Peru. *Am J Trop Med Hyg* 2013; **89**: 507–15.
62 Pelly TF, Santillan CF, Gilman RH, *et al.* Tuberculosis skin testing, anergy and protein malnutrition in Peru. *Int J Tuberc Lung Dis* 2005; **9**: 977–84.
63 Zevallos K, Vergara KC, Vergara A, Vidal C, Garcia HH, Evans CA. Tuberculin skin-test reactions are unaffected by the severity of hyperendemic intestinal helminth infections and co-infections. *Am J Trop Med Hyg* 2010; **83**: 319–25.
64 Ford CM, Bayer AM, Gilman RH, *et al.* Factors Associated with Delayed Tuberculosis Test-seeking Behavior in the Peruvian Amazon. *Am J Trop Med Hyg* 2009; **81**: 1097–102.
65 Baldwin MR, Yori PP, Ford C, *et al.* Tuberculosis and nutrition: disease perceptions and health seeking behavior of household contacts in the Peruvian Amazon. *Int J Tuberc Lung Dis* 2004; **8**: 1484–91.
66 Ryan GJ, Hoff DR, Driver ER, *et al.* Multiple M. tuberculosis phenotypes in mouse and guinea pig lung tissue revealed by a dual-staining approach. *PLoS One* 2010; **5**: e11108.
67 Sharma S, Ryndak MB, Aggarwal AN, *et al.* Transcriptome analysis of mycobacteria in sputum samples of pulmonary tuberculosis patients. *PLoS One* 2017; **12**: e0173508.
68 Sun Z, Zhang Y. Spent culture supernatant of Mycobacterium tuberculosis H37Ra improves viability of aged cultures of this strain and allows small inocula to initiate growth. *J Bacteriol* 1999; **181**: 7626–8.
69 Wayne LG, Sohaskey CD. Nonreplicating persistence of mycobacterium tuberculosis. *Annu Rev Microbiol* 2001; **55**: 139–63.
70 Andersson N, Harris E, Mitchell S, Hamel C, Shea B. Risk factors associated with recent transmission of tuberculosis : systematic review and meta-analysis. *Int J TB Lung Dis* 2009; **13**: 17–26.
71 Saunders M, Datta S. Contact investigation: a priority for tuberculosis control programs. *Am J Respir Crit Care Med* 2016; **194**: 1049–51.
72 Sandhu G, Battaglia F, Ely BK, *et al.* Discriminating active from latent tuberculosis in patients presenting to community clinics. *PLoS One* 2012; **7**. DOI:10.1371/journal.pone.0038080.
73 Wingfield T, Schumacher SG, Sandhu G, *et al.* The seasonality of tuberculosis, sunlight, vitamin D, and household crowding. *J Infect Dis* 2014; **210**: 774–83.
74 Rao VB, Pelly TF, Gilman RH, *et al.* Zinc cream and reliability of tuberculosis skin testing. *Emerg Infect Dis* 2007; **13**: 1101–4.
75 Zevallos K, Alvarado J, Quino W, *et al.* Household TB clustering associated with TB reactivation and poverty. In: 40th International Union Against TB and Lung Disease - World Region. CDC Late Breaker. 2009: 3.
76 Brooks-Pollock E, Becerra MC, Goldstein E, Cohen T, Murray MB.

Epidemiologic inference from the distribution of tuberculosis cases in households in Lima, Peru. *J Infect Dis* 2011; **203**: 1582–9.

77 Jindani A, Doré CJ, Mitchison D a. Bactericidal and sterilizing activities of antituberculosis drugs during the first 14 days. *Am J Respir Crit Care Med* 2003; **167**: 1348–54.

78 Hobby GL, Holman AP, Iseman MD, Jones JM. Enumeration of Tubercle Bacilli in Sputum of Patients with Pulmonary Tuberculosis. *Antimicrob Agents Chemother* 1973; **4**: 94–104.

79 Davies GR, Brindle R, Khoo SH, Aarons LJ. Use of nonlinear mixed-effects analysis for improved precision of early pharmacodynamic measures in tuberculosis treatment. *Antimicrob Agents Chemother* 2006; **50**: 3154–6.

80 WHO | Global tuberculosis report 2018. *WHO* 2019. https://www.who.int/tb/publications/global_report/en/ (accessed June 12, 2019).

81 Escombe AR, Oeser C, Gilman RH, *et al.* The Detection of Airborne Transmission of Tuberculosis from HIV-Infected Patients , Using an In Vivo Air Sampling Model. *Clin Infect Dis* 2010; **44**: 1349–57.

82 Kirk SM, Schell RF, Moore a V, Callister SM, Mazurek GH. Flow cytometric testing of susceptibilities of Mycobacterium tuberculosis isolates to ethambutol, isoniazid, and rifampin in 24 hours. *J Clin Microbiol* 1998; **36**: 1568–73.

83 Moore A V, Kirk SM, Callister SM, Mazurek H, Schell RF, Mazurek GH. Safe Determination of Susceptibility of Mycobacterium tuberculosis to Antimycobacterial Agents by Flow Cytometry. *J Clin Microbiol* 1999; **37**: 479–83.

84 Norden M a, Kurzynski T a, Bownds SE, Callister SM, Schell RF. Rapid susceptibility testing of Mycobacterium tuberculosis (H37Ra) by flow cytometry. *J Clin Microbiol* 1995; **33**: 1231–7.

85 Sherman JM, Tovar M, Gilman RH, *et al.* Using treatment failure to screen for MDRTB is associated with TB recurrence, death and transmission. In: 55th annual conference of the American Society of Tropical Medicine & Hygiene. Atlanta, Ga, 2006: S312.

86 Datiko DG, Lindtjørn B. Tuberculosis recurrence in smear-positive patients cured under DOTS in southern Ethiopia: retrospective cohort study. *BMC Public Health* 2009; **9**: 348.

87 Moore DAJ, Caviedes L, Gilman RH, *et al.* Infrequent MODS TB culture cross-contamination in a high-burden resource-poor setting. *Diagn Microbiol Infect Dis* 2006; **56**: 35–43.

88 Chigutsa E, Patel K, Denti P, *et al.* A Time-to-Event Pharmacodynamic Model Describing Treatment Response in Patients with Pulmonary Tuberculosis Using Days to Positivity in Automated Liquid Mycobacterial Culture. *Antimicrob Agents Chemother* 2013; **57**: 789.

89 Bark CM, Gitta P, Johnson JL. Differences in EBA in the first two days of standard anti-tuberculosis treatment in different geographic regions. *Int J TB Lung Dis* 2014; **18**: 1006–7.

90 Schramm B, Hewison C, Bonte L, *et al.* Field evaluation of a simple fluorescence method for detection of viable Mycobacterium tuberculosis in sputum specimens during treatment follow-up. *J Clin Microbiol* 2012; **50**: 2788–90.

91 Deun A Van, Maug AKJ, Hossain A, Gumusboga M, Jong BC De. Fluorescein diacetate vital staining allows earlier diagnosis of rifampicin-resistant tuberculosis. *Int J TB Lung Dis* 2012; **16**: 1174–9.

92 Ramos E, Schumacher SG, Siedner M, *et al.* Optimizing tuberculosis testing for basic laboratories. *Am J Trop Med Hyg* 2010; **83**: 896–901.

93 Leung E, Minion J, Benedetti A, Pai M, Menzies D. Microcolony culture techniques for tuberculosis diagnosis : a systematic review. *Int J TB Lung Dis* 2011; **16**: 16–23.

94 Ramsay A, Bonnet M, Gagnidze L, Githui W, Varaine F, Guérin PJ. Sputum, sex and scanty smears: New case definition may reduce sex disparities in smear-positive tuberculosis. *Int J Tuberc Lung Dis* 2009; **13**: 613–9.
95 Banda HT, Harries AD, Boeree MJ, Nyirenda TE, Banerjee A, Salaniponi FML. Viability of stored sputum specimens for smear microscopy and culture. *Int J Tuberc Lung Dis* 2000; **4**: 272–4.
96 Mase SR, Ramsay a., Ng V, *et al.* Yield of serial sputum specimen examinations in the diagnosis of pulmonary tuberculosis: A systematic review. *Int J Tuberc Lung Dis* 2007; **11**: 485–95.
97 WHO | TB diagnostics and laboratory strengthening - WHO policy. http://www.who.int/tb/laboratory/policy_diagnosis_pulmonary_tb/en/ (accessed June 30, 2015).
98 centers of disease control. Find TB Resources - TB Education & Training - Display Materials. https://findtbresources.cdc.gov/index.cfm?FuseAction=DispMatl&MatlNbr=30212&SearchType=Intermediate&BTWFields=and&AllText=sputum collection&AllTextOper=and&Title=&TitleOper=and&Authors=&AuthorsOper=and&FormatCode=&FormatCodeOper=or&Publishers=&PublishersO (accessed Nov 16, 2015).
99 Ho J, Marks GB, Fox GJ. The impact of sputum quality on tuberculosis diagnosis : a systematic review. *Int J TB Lung Dis* 2015; **19**: 537–44.
100 Bell DJ, Dacombe R, Graham SM, *et al.* Simple measures are as effective as invasive techniques in the diagnosis of pulmonary tuberculosis in Malawi. *Int J Tuberc Lung Dis* 2009; **13**: 99–104.
101 Oberhelman RA, Soto-Castellares G, Gilman RH, *et al.* A Controlled Study of Tuberculosis Diagnosis in HIV-Infected and Uninfected Children in Peru. *PLoS One* 2015; **10**: e0120915.
102 Leeflang MMG, Deeks JJ, Takwoingi Y, Macaskill P. Cochrane diagnostic test accuracy reviews. *Syst Rev* 2013; **2**: 82.
103 Leeflang MMG, Deeks JJ, Gatsonis C, Bossuyt PMM. Systematic reviews of diagnostic test accuracy. *Ann Intern Med* 2008; **149**: 889–97.
104 Pai M. Systematic reviews of diagnostic test evaluations: what's behind the scenes? *Evid Based Med* 2004; **9**: 101–3.
105 World Health Organisation. Same-day-diagnosis of tuberculosis - Policy Statement. 2010 http://www.stoptb.org/wg/gli/assets/documents/WHO Policy Statement on Same-day-diagnosis by microscopy FINAL March 2010.pdf (accessed March 30, 2017).
106 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**: 177–88.
107 world health organization. WHO country data. 2015. http://www.who.int/tb/country/data/download/en/.
108 Jassal MS, Bishai WR. Epidemiology and Challenges to the Elimination of Global Tuberculosis. *Clin Infect Dis* 2010; **50**: S156–64.
109 Centers for disease control and prevention. CDC - Tuberculosis Cases and Case Rates per 1000,000 Population, 1953–2014. 2014. http://www.cdc.gov/tb/statistics/tbcases.htm.
110 Mills EJ, Thorlund K, Ioannidis JP a. Demystifying trial networks and network meta-analysis. *BMJ* 2013; **2914**: 10–5.
111 Chaimani A, Higgins JPT, Mavridis D, Spyridonos P, Salanti G. Graphical Tools for Network Meta-Analysis in STATA. *PLoS One* 2013; **8**. DOI:10.1371/journal.pone.0076654.
112 Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. .
113 Whiting PF, Rutjes AWS, Westwood ME, *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011; **155**: 529–36.
114 Mpagama SG, Mtabho C, Mwaigwisya S, *et al.* Comparison of Overnight

Pooled and Standard Sputum Collection Method for Patients with Suspected Pulmonary Tuberculosis in Northern Tanzania. *Tuberc Res Treat* 2012; **2012**: 1–5.
115 Andrews RH, Radhakrishna S. A comparison of two methods of sputum collection in the diagnosis of pulmonary tuberculosis. *Tubercle* 1959; **40**: 155–62.
116 Majumdar S, Kunte AB, Menon CRN. Evaluation of methods of sputum collection in the diagnosis of pulmonary tuberculosis. *Indian J Tuberc* 1962; **10**. https://scholar.google.co.uk/scholar?hl=en&q=Evaluation+of+methods+of sputum+collection in+the+diagnosis+of+pulmonary+tuberculosis+pande&btnG=&as_sdt=1%2C5&as_sdtp=.
117 Warren J, Bhattacharya M, De almeida Kleper N, Trakas K, Peterson LR. A minimum 5.0ml of sputum improves the sensitivity of acid-fast smear for Mycobacterium tuberculosis. *Am J Respir Crit Care Med* 2000; **161**: 1559–62.
118 Abdel Aziz A, Eissa H, Shoeb S, Yousef R. Comparative study of the differint methods of collection of sputum samples in tuberculosis patients. *J Egypt Public Health Assoc* 1985; **61**: 61–9.
119 Krasnow I, Wayne LG. Comparison of methods for tuberculosis bacteriology. *Appl Microbiol* 1969; **18**: 915–7.
120 Kestle DG, Kubica GP. Sputum collection for cultivation of mycobacteria: an early morning specimen or the 24-72 hour pool? *Tech Bull Regist Med Technol* 1966; **37**: 213–5.
121 Ssengooba W, Kateete DP, Wajja A, *et al.* An Early Morning Sputum Sample Is Necessary for the Diagnosis of Pulmonary Tuberculosis, Even with More Sensitive Techniques: A Prospective Cohort Study among Adolescent TB-Suspects in Uganda. *Tuberc Res Treat* 2012; **2012**: 6.
122 Pande DC, Gautam KD, Mehrotra ML, Misra J. A retrospective analysis of yield of cases by two methods of sputum collection in suspects of pulmonary tuberculosis among symptomatics attending a tuberculosis centre. *Indian J Tuberc* 1974; **21**: 191–5.
123 Geldenhuys HD, Whitelaw A, Tameris MD, *et al.* A controlled trial of sputum induction and routine collection methods for TB diagnosis in a South African community. *Eur J Clin Microbiol Infect Dis* 2014; **33**: 2259–66.
124 Schoch OD, Rieder P, Tueller C, *et al.* Diagnostic yield of sputum, induced sputum, and bronchoscopy after radiologic tuberculosis screening. *Am J Respir Crit Care Med* 2007; **175**: 80–6.
125 Khan MS, Dar O, Sismanidis C, Shah K, Godfrey-Faussett P. Improvement of tuberculosis case detection and reduction of discrepancies between men and women by simple sputum-submission instructions: a pragmatic randomised controlled trial. *Lancet* 2007; **369**: 1955–60.
126 Alisjahbana B, Van Crevel R, Danusantoso H, *et al.* Better patient instruction for sputum sampling can improve microscopic tuberculosis diagnosis. *Int J Tuberc Lung Dis* 2005; **9**: 814–7.
127 Sakundarno M, Nurjazuli N, Jati SP, *et al.* Insufficient quality of sputum submitted for tuberculosis diagnosis and associated factors, in Klaten district, Indonesia. *BMC Pulm Med* 2009; **9**: 16.
128 Mohamed EY, Abdalla SM, Ahmed M, Elimam AG, Abdalla M, Mansour AL. Diagnosis of pulmonary tuberculosis under Directly Observed Sputum Collection “ DOSC ” in a general Sudanese hospital. *Int J Pharm Med Res* 2014; **2**: 127–9.
129 Maciel E, do Prado T, Peres R, Palaci M, Johnon J, Dietze R. Guided sputum sample collection and culture contamination rates in the diagnosis of pulmonary TB. *J Brazilian Pneumol* 2009; **35**: 460–3.
130 Souza Pinto V, Bammann RH. Chest physiotherapy for collecting sputum samples from HIV-positive patients suspected of having tuberculosis. *Int J*

Tuberc Lung Dis 2007; **11**: 1302–7.
131 Lee YJ, Shin S, Roh EY, *et al.* The effectiveness of a brochure describing an acceptable method of sputum collection for tuberculosis testing. *Int J Tuberc Lung Dis* 2013; **17**: 1587–9.
132 Mhalu G, Hella J, Doulla B, *et al.* Do instructional videos on sputum submission result in increased tuberculosis case detection? A randomized controlled trial. *PLoS One* 2015; **10**: 1–15.
133 Kalema N, de n Boon S, Cattamanchi A, *et al.* Oral antimicrobial rinse to reduce mycobacterial culture contamination among tuberculosis suspects in uganda: A prospective study. *PLoS One* 2012; **7**: 1–5.
134 Peres RL, Palaci M, Loureiro RB, *et al.* Evaluation of oral antiseptic rinsing before sputum collection to reduce contamination of mycobacterial cultures. *J Clin Microbiol* 2011; **49**: 3058–60.
135 Davis JL, Huang L, Kovacs J a, *et al.* Polymerase chain reaction of secA1 on sputum or oral wash samples for the diagnosis of pulmonary tuberculosis. *Clin Infect Dis* 2009; **48**: 725–32.
136 Yoon S, Lee N, Yim J. Impact of sputum gross appearance and volume on smear positivity of pulmonary tuberculosis: a prospective cohort study. *BMC Infect Dis* 2012; **12**: 172.
137 Boehme CC, Nabeta P, Hillemann D, *et al.* Rapid molecular detection of tuberculosis and rifampin resistance. *N Engl J Med* 2010; **363**: 1005–15.
138 Boehme CC, Nicol MP, Nabeta P, *et al.* Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: A multicentre implementation study. *Lancet* 2011; **377**: 1495–505.
139 Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane database Syst Rev* 2014; : CD009593.
140 World Health Organisation. Definition of a new sputum smear-positive TB case. 2007. https://www.who.int/tb/areas-of-work/laboratory/policy_sputum_smearpositive_tb_case/en/.
141 Wingfield T, Boccia D, Tovar M, *et al.* Defining catastrophic costs and comparing their importance for adverse tuberculosis outcome with multi-drug resistance: a prospective cohort study, Peru. *PLoS Med* 2014; **11**: e1001675.
142 Oxlade O, Falzon D, Menzies D. The impact and cost-effectiveness of strategies to detect drug-resistant tuberculosis. *Eur Respir J* 2012; **39**: 626–34.
143 WHO Policy on TB Infection Control in Health-Care Facilities, Congregate Settings and Households. .
144 Escombe AR, Oeser CC, Gilman RH, *et al.* Natural Ventilation for the Prevention of Airborne Contagion. *PLoS Med* 2007; **4**: e68.
145 Popov TA, Shenkada MS, Tzoncheva A V, Pravtchanska MP, Mustakov TB, Dimitrov VD. Circadian changes in the sputum of asthmatic subjects and healthy controls. *World Allergy Organ J* 2008; **1**: 74–8.
146 Spengler CM, Shea SA. Endogenous circadian rhythm of pulmonary function in healthy humans. *Am J Respir Crit Care Med* 2000; **162**: 1038–46.
147 Panzer SE, Dodge AM, Kelly EAB, Jarjour NN. Circadian variation of sputum inflammatory cells in mild asthma. *J Allergy Clin Immunol* 2003; **111**: 308–12.
148 Jansen JP, Naci H. Is network meta-analysis as valid as standard pairwise meta-analysis? It all depends on the distribution of effect modifiers. *BMC Med* 2013; **11**: 159.
149 WHO | Fluorescent light-emitting diode (LED) microscopy for diagnosis of tuberculosis policy. http://www.who.int/tb/publications/2011/led_microscopy_diagnosis_97892

41501613/en/ (accessed Feb 20, 2014).
150 Sloan DJ, Corbett EL, Butterworth AE, *et al.* Optimizing outpatient serial sputum colony counting for studies of tuberculosis treatment in resource-poor settings. *J Clin Microbiol* 2012; **50**: 2315–20.
151 Ministerio de salud. Norma tecnica de salud para la atencion integral de las personas afectadas por tuberculosis. Lima, Peru, 2013 DOI:715-2013/MINSA.
152 Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; : 307–10.
153 Steingart KR, Ng V, Henry M, *et al.* Sputum processing methods to improve the sensitivity of smear microscopy for tuberculosis: a systematic review. *Lancet Infect Dis* 2006; **6**: 664–74.
154 Kim SJ, Lee SH, Kim IS, Kim HJ, Kim SK, Rieder HL. Risk of occupational tuberculosis in National Tuberculosis Programme laboratories in Korea. *Int J Tuberc Lung Dis* 2007; **11**: 138–42.
155 Best M, Sattar SA, Springthorpe VS, Kennedy ME. Efficacies of selected disinfectants against Mycobacterium tuberculosis. *J Clin Microbiol* 1990; **28**: 2234–9.
156 Rutala WA, Cole EC, Wannamaker NS, Weber DJ. Inactivation of Mycobacterium tuberculosis and Mycobacterium bovis by 14 hospital disinfectants. *Am J Med* 1991; **91**: 267S-271S.
157 Rutala WA, Weber DJ, Healthcare Infection Control Practices Advisory Committee. Guideline for Disinfection and Sterilization in Healthcare Facilities. 2008 DOI:10.1086/423182.
158 Olaru ID, Albert H, Zallet J, *et al.* Impact of quality improvement in tuberculosis laboratories in low- and lower-middle-income countries: a systematic review. *Int J Tuberc Lung Dis* 2007; **11**: 138–42.
159 Gilpin C, Kim SJ, Lumb R, Rieder HL, Van Deun A, Microscopy WG on SS. Critical appraisal of current recommendations and practices for tuberculosis sputum smear microscopy [Workshop Report]. http://www.ingentaconnect.com.ezp.welch.jhmi.edu/content/iuatld/ijtld/2007/00000011/00000009/art00004;jsessionid=1gwbdsrh6budn.x-ic-live-02# (accessed June 18, 2018).
160 Van Deun A, Portaels F. Limitations and requirements for quality control of sputum smear microscopy for acid-fast bacilli [Planning and Practice]. http://www.ingentaconnect.com/content/iuatld/ijtld/1998/00000002/00000009/art00011;jsessionid=1hsouk45t13cu.x-ic-live-02# (accessed June 18, 2018).
161 World Health Organization. The END TB strategy. 2015. http://www.who.int/tb/post2015_TBstrategy.pdf (accessed Dec 20, 2015).
162 Global strategy and targets for tuberculosis prevention, care and control after 2015 Report by the Secretariat. .
163 Keshavjee S, Amanullah F, Cattamanchi A, *et al.* Moving toward tuberculosis elimination critical issues for research in diagnostics and therapeutics for tuberculosis infection. Am. J. Respir. Crit. Care Med. 2019; **199**: 564–71.
164 Yuen CM, Seddon JA, Keshavjee S, Dodd PJ. Risk-benefit analysis of tuberculosis infection testing for household contact management in high-burden countries: a mathematical modelling study. *Lancet Glob Heal* 2020; **8**: e672–80.
165 WHO Library Cataloguing-in-Publication Data Recommendations for investigating contacts of persons with infectious tuberculosis in low-and middle-income countries. 2012 http://www.who.int/about/licensing/copyright_form/en/index.html (accessed Nov 13, 2020).
166 Stop TB Partnership | Global Drug Facility (GDF) - GDF Product Catalogue. http://www.stoptb.org/gdf/drugsupply/pc3.asp?PID=702 (accessed Sept 1,

2020).
167 Toit K, Mitchell S, Balabanova Y, *et al.* The Colour Test for drug susceptibility testing of Mycobacterium tuberculosis strains. *Int J Tuberc Lung Dis* 2012; **16**: 1113–8.
168 Wendell J, Hanson Herrera B, Ramos E, *et al.* The MDR/XDR-TB Color Test simultaneous culture and drug susceptibility results predict TB outcome. In: International Union against TB & Lung Disease- world conference. 2013: 17 (12);S271.
169 van Zyl-Smit RN, Binder A, Meldau R, *et al.* Comparison of quantitative techniques including Xpert MTB/RIF to evaluate mycobacterial burden. *PLoS One* 2011; **6**. DOI:10.1371/journal.pone.0028815.
170 Chakravorty S, Simmons AM, Rowneki M, *et al.* The new Xpert MTB/RIF ultra: Improving detection of Mycobacterium tuberculosis and resistance to Rifampin in an assay suitable for point-of-care testing. *MBio* 2017; **8**. DOI:10.1128/mBio.00812-17.
171 Lawn SD, Nicol MP. Xpert® MTB/RIF assay: Development, evaluation and implementation of a new rapid molecular diagnostic for tuberculosis and rifampicin resistance. Future Microbiol. 2011; **6**: 1067–82.
172 Pencina MJ, D'Agostino RB, Song L. Quantifying discrimination of Framingham risk functions with different survival C statistics. *Stat Med* 2012; **31**: 1543–53.
173 Ventanilla -Callao DH, Villegas KZ. "DISCORDANCIA ENTRE LAS CEPAS DEL CASO ÍNDICE DE TUBERCULOSIS Y LOS CASOS SECUNDARIOS DENTRO. Universidad Peruana Cayetano Heredia, 2017 http://repositorio.upch.edu.pe/handle/upch/3526 (accessed Nov 13, 2020).
174 Barry CE, Boshoff HI, Dartois V, *et al.* The spectrum of latent tuberculosis: Rethinking the biology and intervention strategies. Nat. Rev. Microbiol. 2009; **7**: 845–55.
175 Dheda K, Schwander SK, Zhu B, Van Zyl-Smit RN, Zhang Y. The immunology of tuberculosis: From bench to bedside. Respirology. 2010; **15**: 433–50.
176 Proaño A, Bravard MA, López JW, *et al.* Dynamics of Cough Frequency in Adults Undergoing Treatment for Pulmonary Tuberculosis. *Clin Infect Dis* 2017; **64**: 1174–81.
177 Fennelly KP, Jones-López EC, Ayakaka I, *et al.* Variability of Infectious Aerosols Produced during Coughing by Patients with Pulmonary Tuberculosis. *Am J Respir Crit Care Med* 2012; **186**: 450–7.
178 Hales CM, Heilig CM, Chaisson R, *et al.* The association between symptoms and microbiologically defined response to tuberculosis treatment. *Ann Am Thorac Soc* 2013; **10**: 18–25.
179 Ruhl CR, Pasko BL, Khan HS, *et al.* Mycobacterium tuberculosis Sulfolipid-1 Activates Nociceptive Neurons and Induces Cough. *Cell* 2020; **181**: 293-305.e11.
180 Turner RD. Cough in pulmonary tuberculosis: Existing knowledge and general insights. Pulm. Pharmacol. Ther. 2019; **55**: 89–94.
181 Bhatt A, Fujiwara N, Bhatt K, *et al.* Deletion of kasB in Mycobacterium tuberculosis causes loss of acid-fastness and subclinical latent tuberculosis in immunocompetent mice. *Proc Natl Acad Sci U S A* 2007; **104**: 5157–62.
182 Cattamanchi A, Huang L, Worodria W, *et al.* Integrated strategies to optimize sputum smear microscopy: A prospective observational study. *Am J Respir Crit Care Med* 2011; **183**: 547–51.
183 Theron G, Limberis J, Venter R, *et al.* Bacterial and host determinants of cough aerosol culture positivity in patients with drug-resistant versus drug-susceptible tuberculosis. *Nat Med* 2020; **26**: 1435–43.
184 WHO | Recommendations for investigating contacts of persons with infectious tuberculosis in low- and middle-income countries. *WHO* 2015.

185 Migliori GB, Ambrosio LD', Centis R, *et al.* Guiding Principles to Reduce Tuberculosis Transmission in the WHO European Region. 2018 http://www.euro.who.int/pubrequest (accessed Nov 28, 2020).
186 Saunders MJ, Wingfield T, Tovar MA, *et al.* 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* 2017; **17**: 1190–9.
187 Acuña-Villaorduña C, Jones-López EC, Fregona G, *et al.* Intensity of exposure to pulmonary tuberculosis determines risk of tuberculosis infection and disease. *Eur Respir J* 2018; **51**. DOI:10.1183/13993003.01578-2017.
188 Zak DE, Penn-Nicholson A, Scriba TJ, *et al.* A blood RNA signature for tuberculosis disease risk: a prospective cohort study. *Lancet* 2016; **387**: 2312–22.
189 Salipante SJ, SenGupta DJ, Cummings LA, Land TA, Hoogestraat DR, Cookson BT. Application of whole-genome sequencing for bacterial strain typing in molecular epidemiology. *J Clin Microbiol* 2015; **53**: 1072–9.
190 Walker TM, Ip CLC, Harrell RH, *et al.* Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: A retrospective observational study. *Lancet Infect Dis* 2013; **13**: 137–46.
191 Oberhelman RA, Soto-Castellares G, Gilman RH, *et al.* Diagnostic approaches for paediatric tuberculosis by use of different specimen types, culture methods, and PCR: a prospective case-control study. *Lancet Infect Dis* 2010; **10**: 612–20.
192 Saunders MJ, Tovar MA, Collier D, *et al.* Active and passive case-finding in tuberculosis-affected households in Peru: a 10-year prospective cohort study. *Lancet Infect Dis* 2019; **19**: 519–28.
193 Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database Syst Rev* 2014; : CD009593.
194 world health organization. Global tuberculosis report 2020. 2020. https://www.who.int/publications/i/item/9789240013131 (accessed Dec 31, 2020).
195 Genexpert Cepheid test - CB-NAAT - TBFacts. https://tbfacts.org/genexpert/ (accessed April 22, 2020).
196 Di Tanna GL, Raza Khaki A, Theron G, *et al.* Effect of Xpert MTB/RIF on clinical outcomes in routine care settings: individual patient data meta-analysis. *Lancet Glob Heal* 2019; **7**: e191–9.
197 Auld AF, Fielding KL, Gupta-Wright A, Lawn SD. Xpert MTB/RIF — why the lack of morbidity and mortality impact in intervention trials? *Trans R Soc Trop Med Hyg* 2016; **110**: 432–44.
198 Mitchison DA. The action of antituberculosis drugs in short-course chemotherapy. *Tubercle* 1985; **66**: 219–25.
199 Streptomycin treatment of pulmonary tuberculosis. *Br Med J* 1948; **2**: 769–82.
200 Streptomycin in Pulmonary Tuberculosis. *Br Med J* 1948; **2**: 790–1.
201 MITCHISON DA. Development of streptomycin resistant strains of tubercle bacilli in. *Thorax* 1950; **5**: 144–61.
202 Gegia M, Winters N, Benedetti A, van Soolingen D, Menzies D. Treatment of isoniazid-resistant tuberculosis with first-line drugs: a systematic review and meta-analysis. *Lancet Infect Dis* 2017; **17**: 223–34.
203 World Health Organisation. Treatment of tuberculosis guidelines. Fourth edition. 2009.
204 ISRCTN17820976. PREVENT TB: Improving determinants of TB cure, prevention & diagnosis. BMC Regist. 2016. https://doi.org/10.1186/ISRCTN17820976.
205 Automated Real-time Nucleic Acid Amplification Technology for Rapid and

Simultaneous Detection of Tuberculosis and Rifampicin Resistance: Xpert MTB/RIF System Policy Statement WHO Library Cataloguing-in-Publication Data Policy statement: automated real-ti. 2011 www.who.int (accessed April 12, 2020).
206 Pletz MWR, De Roux A, Roth A, Neumann KH, Mauch H, Lode H. Early Bactericidal Activity of Moxifloxacin in Treatment of Pulmonary Tuberculosis: A Prospective, Randomized Study. *Antimicrob Agents Chemother* 2004; **48**: 780–2.
207 Sirgel F a, Donald PR, Odhiambo J, *et al.* A multicentre study of the early bactericidal activity of anti-tuberculosis drugs. *J Antimicrob Chemother* 2000; **45**: 859–70.
208 Wingfield T, Tovar MA, Huff D, *et al.* A randomized controlled study of socioeconomic support to enhance tuberculosis prevention and treatment, Peru. *Bull World Heal Organ* 2017; **95**: 270–80.
209 Kurbatova E, Gammino VM, Bayona J, *et al.* Predictors of sputum culture conversion among patients treated for multidrug-resistant tuberculosis. *Int J Tuberc Lung Dis* 2012; **16**: 1335–43.
210 MC S, D B, M K, *et al.* Time to Sputum Culture Conversion and Treatment Outcomes Among Patients with Isoniazid-Resistant Tuberculosis in Atlanta, Georgia. *Clin Infect Dis* 2017; **65**: 1862-1871.
211 Kurbatova E V, Cegielski JP, Lienhardt C, *et al.* Sputum culture conversion as a prognostic marker for end-of-treatment outcome in patients with multidrug-resistant tuberculosis: a secondary analysis of data from two observational cohort studies. *Lancet Respir Med* 2015; **3**: 201–9.
212 Diarra B, Decroo T, Somboro A, *et al.* Fluorescein diacetate and rapid molecular testing for the early identification of rifampicin resistance in Mali. Int. J. Tuberc. Lung Dis. 2020; **24**: 763–9.
213 Lenaerts A, Barry CE, Dartois V. Heterogeneity in tuberculosis pathology, microenvironments and therapeutic responses. *Immunol Rev* 2015; **264**: 288–307.
214 Dartois V, E. Barry C. Clinical Pharmacology and Lesion Penetrating Properties of Second- and Third-Line Antituberculous Agents Used in the Management of Multidrug-Resistant (MDR) and Extensively-Drug Resistant (XDR) Tuberculosis. *Curr Clin Pharmacol* 2010; **5**: 96–114.
215 Johnson JL, Hadad DJ, Dietze R, *et al.* Shortening treatment in adults with noncavitary tuberculosis and 2-month culture conversion. *Am J Respir Crit Care Med* 2009; **180**: 558–63.
216 Hänscheid T, Ribeiro CM, Shapiro HM, Perlmutter NG. Fluorescence microscopy for tuberculosis diagnosis. *Lancet Infect Dis* 2007; **7**: 236–7.
217 Horne DJ, Royce SE, Gooze L, *et al.* Sputum monitoring during tuberculosis treatment for predicting outcome: systematic review and meta-analysis. *Lancet Infect Dis* 2010; **10**: 387–94.
218 Wallis RS, Doherty TM, Onyebujoh P, *et al.* Biomarkers for tuberculosis disease activity, cure, and relapse. *Lancet Infect Dis* 2009; **9**: 162–72.
219 Horne DJ, Royce SE, Gooze L, *et al.* Sputum monitoring during tuberculosis treatment for predicting outcome: systematic review and meta-analysis. Lancet Infect. Dis. 2010; **10**: 387–94.
220 Critchley JA, Orton LC, Pearson F. Adjunctive steroid therapy for managing pulmonary tuberculosis. Cochrane Database Syst. Rev. 2014; **2014**. DOI:10.1002/14651858.CD011370.
221 Walker NF, Clark SO, Oni T, *et al.* Doxycycline and HIV infection suppress tuberculosis-induced matrix metalloproteinases. *Am J Respir Crit Care Med* 2012; **185**: 989–97.
222 Degner NR, Wang J-Y, Golub JE, Karakousis PC. Metformin Use Reverses the Increased Mortality Associated With Diabetes Mellitus During Tuberculosis Treatment. *Clin Infect Dis* 2018; **66**: 198–205.
223 Sabur NF, Esmail A, Brar MS, Dheda K. Diagnosing tuberculosis in

hospitalized HIV-infected individuals who cannot produce sputum: Is urine lipoarabinomannan testing the answer? *BMC Infect Dis* 2017; **17**: 803.
224 Shah NS, Cavanaugh JS, Pratt R, *et al.* Epidemiology of smear-negative pulmonary tuberculosis in the United States, 1993-2008. *Int J Tuberc Lung Dis* 2012; **16**: 1234–40.
225 World Health Organisation. Treatment of tuberculosis: guidelines for National Programmes. 2003.
226 Martin LJ, Roper MH, Grandjean L, *et al.* Rationing tests for drug-resistant tuberculosis - who are we prepared to miss? *BMC Med* 2016; **14**: 30.
227 Fitzwater SP, Caviedes L, Gilman RH, *et al.* Prolonged infectiousness of tuberculosis patients in a directly observed therapy short-course program with standardized therapy. *Clin Infect Dis* 2010; **51**: 371–8.
228 Rockwood N, du Bruyn E, Morris T, Wilkinson RJ. Assessment of treatment response in tuberculosis. *Expert Rev Respir Med* 2016; **10**: 643–54.
229 Turner RD, Chiu C, Churchyard GJ, *et al.* Tuberculosis Infectiousness and Host Susceptibility. *J Infect Dis* 2017; **216**: S636–43.
230 World Health Organization. Guidelines for treatment of drug-susceptible tuberculosis and patient care (2017 update). 2017 https://www.who.int/tb/publications/2017/dstb_guidance_2017/en/.
231 Datta S, Gilman RH, Montoya R, *et al.* Quality of life, tuberculosis and treatment outcome; a case-control and nested cohort study. *Eur Respir J* 2020; : 1900495.
232 Chew R, Calderón C, Schumacher SG, *et al.* Evaluation of bleach-sedimentation for sterilising and concentrating Mycobacterium tuberculosis in sputum specimens. *BMC Infect Dis* 2011; **11**: 269.
233 Grosset J. Mycobacterium tuberculosis in the Extracellular Compartment: an Underestimated Adversary. *Antimicrob Agents Chemother* 2003; **47**: 833–6.

APPENDIX

Publication list during doctoral studies

Key first-author publications related to this thesis

Datta S, Sherman JM, Tovar MA, Bravard MA, Valencia T, Montoya R, Quino W, D'Arcy N, Ramos ES, Gilman RH, Evans CA. Sputum microscopy with fluorescein diacetate predicts TB infectiousness. J Infectious Disease. 2017 May 16. doi: 10.1093/infdis/jix229 https://academic.oup.com/jid/article/216/5/514/3828235

Datta S, Sherman J, Bravard M, Valencia T, Gilman RH, Evans CA. Clinical evaluation of sputum viability staining for assessing TB treatment response. Clinical Infectious Diseases; 2015 Apr 15;60(8):1186-95 https://academic.oup.com/cid/article/60/8/1186/2462595

Datta S, Shah L, Gilman RH, Evans CA. Comparison of sputum collection methods for TB diagnosis: a systematic review and pairwise and network meta-analysis. Lancet Global Health. 2017 Jun 15. S2214-109X(17)30201-2. doi: 10.1016/S2214-109X(17)30201-2. http://www.thelancet.com/journals/langlo/article/PIIS2214-109X(17)30201-2/fulltext

Datta S, Alvarado K, Gilman RH, Valencia T, Aparicio C, Ramos ES, Montoya R, Evans CA. Optimising fluorescein diacetate sputum smear microscopy for assessing patients with pulmonary tuberculosis. PLoS One. 2019 Apr 30;14(4):e0214131. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0214131

Other first-author publications

Datta S, Gilman RH, Montoya R, Quevedo Cruz L, Valencia T, Saunders MJ, Evans CA. Quality of life, tuberculosis and treatment outcome; a case-control and nested cohort study European Respiratory Journal –2020;1900495. https://erj.ersjournals.com/content/early/2020/04/20/13993003.00495-2019

Datta S, Evans CA. Healthy survival after tuberculosis [EDITORIAL]. The Lancet Infectious Diseases. 2019 Oct;19(10):1045-1047. https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(19)30387-1/fulltext

Datta S, Gonzales-Huerta L, Evans CA
The potential for testing stool to reduce tuberculosis missed diagnoses and misdiagnoses [EDITORIAL]. Am J Trop Med Hyg. 2018 Aug;99(2) http://www.ajtmh.org/docserver/fulltext/14761645/99/2/tpmd180507

Datta S, Saunders MJ, Tovar MA, Evans CA. Improving TB diagnosis: Better tests or better healthcare? [EDITORIAL] PLoS Medicine. 2017 Oct 17;14(10):e1002406. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5644978/

Other publications

Allwood B, van der Zalm B, Amaral AFS, Byrne A, **Datta S**, Egere U, Evans CA, Evans DM, Gray DM, Hoddinott G, Ivanova O, Jones R, Makanda G, Marx FM, Meghji J, Mpagama S, Pasipanodya JG, Rachow A, Schoeman I, Shaw J, Stek C, van Kampen S, von Delft D, Walker NF, Wallis RS, Mortimer K. Post-tuberculosis lung health: perspectives from the First International Symposium. Int J TB and Lung Dis. Volume 24, Number 8, 1 August 2020, pp. 820-828(9) https://www.ingentaconnect.com/content/iuatld/ijtld/2020/00000024/00000008/art00011

Saunders MJ, Wingfield T, Tovar MA, Herlihy N, Rocha C, Zevallos K, Montoya R, Ramos E, **Datta S,** Evans CA. Mobile phone interventions for tuberculosis should ensure access to mobile phones to enhance equity - A prospective, observational cohort study in Peruvian shantytowns. Trop Med Int Health. 2018 Jun 4. doi: 10.1111/tmi.13087 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6174992/

Wingfield T, Tovar MA, **Datta S,** Saunders MJ, Evans CA. Addressing social determinants to end TB [EDITORIAL]. Lancet. 2018 Mar 24;391(10126):1129-1132. https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(18)30484-7/fulltext

Saunders MJ, Tovar MA, **Datta S,** Evans BEW, Wingfield T, Evans CA. Pragmatic tuberculosis prevention policies for primary care in low- and middle-income countries. [EDITORIAL]Eur Respir J. 2018 Mar 22;51(3) https://erj.ersjournals.com/content/51/3/1800315.long

Saunders MJ, Wingfield T, Tovar MA, Baldwin MR, **Datta S,** Zevallos K, Montoya R, Valencia TR, Friedland JS, Moulton LH, Gilman RH, Evans CA. A score to predict and stratify risk of TB in adult contacts of TB index cases: a prospective derivation and external validation cohort study. Lancet Infectious Diseases. 2017 Nov;17(11):1190-1199.http://www.thelancet.com/journals/laninf/article/PIIS1473-3099(17)30447-4/fulltext

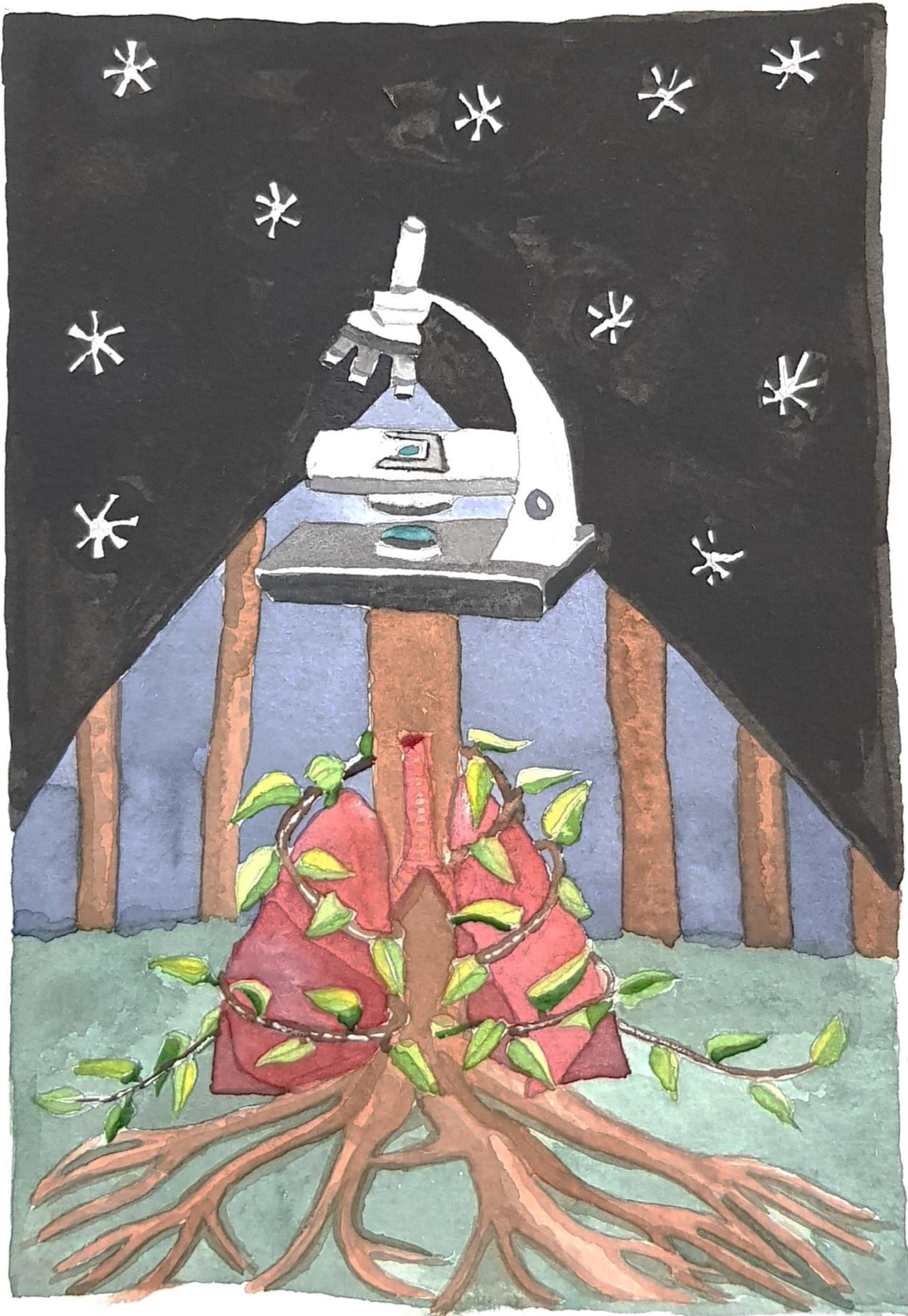
Proaño A, Bravard MA, López JW, Lee GO, Bui D, **Datta S,** Comina G, Zimic M, Coronel J, Caviedes L, Cabrera JL, Salas A, Ticona E, Vu NM, Kirwan DE, Loader MI, Friedland JS, Moore DAJ, Evans CA, Tracey BH, Gilman RH; Tuberculosis Working Group in Peru. Dynamics of Cough Frequency in Adults Undergoing Treatment for Pulmonary TB. Clinical Infectious Diseases. 2017 May 1;64(9):1174-1181. https://academic.oup.com/cid/article/64/9/1174/2957273

Wingfield T, Tovar MA, Huff D, Boccia D, Montoya R, Ramos E, **Datta S,** Saunders MJ, Lewis JJ, Gilman RH, Evans CA.A randomized controlled study of socioeconomic support to enhance TB prevention and treatment, Peru. Bulletin World Health Organisation. 2017 Apr 1;95(4):270-280. doi: 10.2471/BLT.16.170167. Epub 2017 Feb 9.https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5407248/

Saunders MJ, **Datta S.** Contact Investigation: A Priority for TB Control Programs [EDITORIAL]. American Journal Respiratory Critical Care Medicine. 2016 Nov 1;194(9):1049-1051. https://www.atsjournals.org/doi/abs/10.1164/rccm.201605-1007ED?url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org&rfr_dat=cr_pub%3Dpubmed

Wingfield T, Tovar MA, Huff D, Boccia D, Saunders MJ, **Datta S,** Montoya R, Ramos E, Lewis JJ, Gilman RH, Evans C.Beyond pills and tests: addressing the social determinants of TB [EDITORIAL].Clinical Medicine (London). 2016 Dec;16(Suppl 6):s79-s91. http://www.clinmed.rcpjournal.org/content/16/Suppl_6/s79.long

Oberhelman RA, Soto-Castellares G, Gilman RH, Castillo ME, Kolevic L, Delpino T, Saito M, Salazar-Lindo E, Negron E, Montenegro, Laguna-Torres VA, Maurtua-Neumann P**, Datta S,** Evans CA. A Controlled Study of TB Diagnosis in HIV-Infected and Uninfected Children in Peru. PLoS One. 2015 Apr 30;10(4):e0120915. https://www.ncbi.nlm.nih.gov/pubmed/25927526


Sumona 2021