Evaluating TB treatment responses by [18F]FDG-PET/CT imaging

By

Dr Stephanus Theron Malherbe

Dissertation presented for the degree of

Doctor of Philosophy (Molecular Biology) in the

Faculty of Medicine and Health Sciences at

Stellenbosch University.

Supervisor: Prof Gerhard Walzl

Co-supervisors: Prof James Warwick

Prof Katharina Ronacher

December 2016

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Stephanus Theron Malherbe

August 2016

Abstract

BACKGROUND

Tuberculosis (TB) presents a massive health care problem around the world and the rate of unfavourable outcomes after TB treatment remains unacceptably high. The absence of a gold standard to determine when antibiotics have induced sterilising cure confounds the development of new approaches to treat pulmonary tuberculosis (PTB).

Positron Emission Tomography/Computerised Tomography (PET/CT) is well established in the staging, treatment planning and response assessment of cancer. In animal models, ^{18}F-FDG PET/CT has been used to accurately describe disease progression after infection and response to treatment in Pulmonary Tuberculosis (PTB) and human studies with small sample sizes have shown PET/CT to be promising in monitoring the effect of treatment using simple descriptive techniques.

PET uptake intensity is influenced by numerous patient and equipment factors. Reproducible segmentation and quantification of lesions become particularly important in diseases with heterogenic morphology, vague lesion borders and multi-focal distribution throughout an organ or system, such as PTB.

OBJECTIVE

To discover if the functional and anatomical information gained by FDG PET/CT scans on patients with PTB before, during and after treatment can be applied to provide insight into the dynamics of *Mycobacterium tuberculosis* (MTB) versus host interaction, aid treatment response assessment and facilitate the discovery of biomarkers to monitor treatment response.

METHODS

We recruited 99 newly diagnosed adult, HIV-negative, PTB patients and performed PET/CT scans at diagnosis, month 1 and month 6 of treatment, as well as 1 year later. We collected clinical specimens for microbiological testing and biomarker discovery. Scans were evaluated qualitatively and we developed and implemented an automated technique to standardise uptake, segment lung lesions and quantify the scan information. We compared the qualitative and quantitative data to clinical and microbiological outcomes.

RESULTS

We detected PET/CT imaging response patterns consistent with active disease plus the presence of MTB mRNA in sputum and bronchoalveolar lavage fluid in a substantial

proportion of PTB patients after standard treatment and one year later, including patients with a durable cure and others who later developed recurrent disease. However, quantification of PET/CT parameters provided results that correlated very well with microbiological outcomes. A large cavity volume was the best prognostic indicator of failure, while a suboptimal reduction in total glycolytic activity was the best prognostic indicator of recurrent disease. Quantified results also showed promise to serve as reference in discovering novel biomarkers of treatment response.

CONCLUSION

The presence of MTB mRNA in the context of non-resolving and intensifying lesions on PET/CT might indicate ongoing transcription. Thus, that even apparently curative PTB treatment may not eradicate all organisms in most patients, and points to an important complementary role for the immune response in maintaining a disease-free state. The correlation of quantified scan characteristics with microbiological outcomes, provided insight into factors that influence successful treatment. The promising implementation of these parameters in biomarker discovery suggests that PET/CT might be a useful tool that leads to new, improved ways to monitor treatment response.

Opsomming

Agtergrond

Tuberkulose veroorsaak ‘n massiewe gesondheidsprobleem regoor die wêreld en die insidensie van slegte uitkomste na die afloop van behandeling is onaanvaarbaar hoog. Die gebrek aan ‘n goue standaard om vas te stel of antibiotika steriliserende genesing bereik het, belemmer die ontwikkeling van nuwe metodes om tuberkulose te behandel.

Positron Emmissie Topografie/Rekenaar Topografie (PET/RT) is ‘n gevestigde modaliteit in die diagnose, stadiëring en assesering van repons tot behandeling van kanker. In diere studies is ^{18}F-FDG-PET/RT reeds gebruik om die progressie ná infeksie en die reaksie tot behandeling van pulmonale tuberkulose (PTB) akkuraat aan te toon. In voorlopige menslike studies het PET/RT ook belofte getoon in monitering van die effek van behandeling deur middel van eenvoudige beskrywende tegnieke.

Die intensiteit van PET opname word beïnvloed deur verskeie eienskappe van die pasiënt en die toerusting. Die konsekwente segmentasie en kwantifisering van letsels is van uiterste belang in siektes met heterogene morfologie, onduidelike grense en ‘n wye verspreiding deur ‘n orgaansisteem - soos in geval van PTB.

Doelwitte

Om te ontdek of die funksionele en anatomiese inligting wat verkry word deur PET/RT beeldings ondersoeke op pasiënte voor, gedurende en na behandeling toegepas kan word om insig te gee in die dinamiese interaksies tussen Mycobacterium tuberculosis (MTB) en die gasheer.

Metodes

Ons het 99 nuut gediagnoseerde volwasse, HIV-negatiewe pasiënte met pulmonale TB gewerf en PET/CT skanderings uitgevoer by diagnose, asook maand 1 en maand 6 van behandeling; en 1 jaar na die afloop van behandeling. Ons het kliniese monsters versamel vir mikrobiologiese toetse en biomerker ontdekking. Ons het die beelde kwalitatief asseseer, asook ‘n ge-outomatiseerde metode om letsels te segmenteer en kwantifiseer ontwikkel en geïmplementeer. Die kwalitatiewe en kwantitatiewe uitslae is met kliniese en mikrobiologiese uitkomstes vergelyk.

Resultate

Ons het PET/RT respons patrone waargeneem wat inpas met aktiewe letsels, asook die teenwoordigheid van MTB mRNA in sputum en bronchoalveolêre spoeling vloeistof, in ‘n

substansiële getal pasiënte na standard TB behandeling en‘n jaar later - insluitende pasiënte wat genees is. Die gekwantifiseerde resultate het wel baie goed met die mikrobiologiese uitkomstes gekorreleer. ‘n Groot kaviteit volume was die beste prognostiese aanwyser van gefaalde behandeling en onvoldoende vermindering van die totale glikolitiese aktiwiteit was die beste aanwyser van herhalende PTB na behandeling. Gekwantifiseerde resultate het ook belofte getoon as ‘n hulpmiddel in die ontdekking van nuwe biomerkers in perifere bloed van behandelings respons.

Gevolgtrekking

Die teenwoordigheid van MTB mRNA in die konteks van letsels wat nie opklaar nie en selfs intensifiseer op PET/RT, kan voortdurende transkripsie aandui. Dit kan voorstel dat selfs ooglopend genesende PTB behandeling nie al die organismes in die gasheer eradikeer nie; en ‘n belangrike komplimentêre rol vir die van die immuunrespons om ‘n siekte-vrye toestand te handhaaf. Die korrelasie van gekwanifiseerde beelding karakteristieke met mikrobiologiese uitkomstes en die belowende aanwending in biomerker ontdekking, stel voor dat dit tot potensiële hulp kan wees om behandelingsresponse beter te moniteer.

Acknowledgements

Sincere gratitude to supervisors for their advice, guidance, mentorship and many hours of discussing challenges, results, analysis and content.

For funding, I would, firstly like to thank the South African National Research Fund and the Medical Research Council's Clinician Scholarship Program for support.

The Catalysis Foundation for Health (Grant OPP51919) and the Division of Intramural Research, National Institute of Allergy and Infectious Diseases provided funding of the main study.

A special thanks to our participants for their willingness to take part in this study.

Further, I acknowledge the staff at the Stellenbosch University Immunology Research Group (especially Srs. Abrahams, McAnda and Van Zyl), the International TB Research Centre Seoul, the Western Cape Academic PET/CT Centre, Tygerberg Academic Hospital's Nuclear Medicine Department and Pulmonology Unit, as well as managers and health care providers from the City of Cape Town Health Department.

For statistical analysis, I thank Gerard Tromp, Tom Peppard, Yookwan Noh and Lori Dodd.

Further advice, guidance, analysis and help were provided by Jill Winter (Catalysis Foundation for Health), Patrick Dupont (Leuven University), André Loxton, Elizna Maasdorp (Stellenbosch University), Ilse Kant, Clif Barry (NIH), Ray Chen (NIH), Laura Via (NIH), David Alland (Rutgers-New Jersey Medical School), Shubhada Shenai (Rutgers-New Jersey Medical School), Gary Schoolnik (Stanford University), Greg Dolganov (Stanford University).

I look forward to future collaborations with all the above.

In a last and special word of thanks, I would greatly like to thank my wife, Chané for support and encouragement, as well as the rest of my family and my parents for an upbringing that allowed problem solving, discipline, and critical thinking.

Contents

1. Introduction 1

1.2.1 The challenge of tuberculosis 4

TB incidence and treatment outcome 4

Clinical cure, sterilising cure and post-TB lung sequelae 5

1.2.2 ^{18}F-Fluorodeoxyglucose positron emission tomography and computed tomography (PET/CT) 8

Background 8

Animal PET/CT studies 9

Human PET/CT studies 11

PET/CT lesion segmentation 12

1.3.1 Research question 14

1.3.2 Study objectives 14

2. Methods and procedures 15

2.1 Patient recruitment and follow-up 15

2.2 Treatment regimen 17

2.3 Recruitment and follow-up, South-Korea 17

2.4 18F-FDG PET/CT scanner protocol 18

2.5 Qualitative analysis of PET/CT images 18

2.6 Quantitative analysis of PET/CT images 19

Pre-processing 19

Lung masks 19

Background reference 19

Image Quantification 20

2.7 MTB RNA extraction and detection in M6 sputum 26

2.8 RNA in End of treatment bronchoalveolar lavage samples 27

Multiplex Real Time RT-PCR Transcriptional analysis of Sputum and BAL specimens 27

Extraction of total RNA from BAL 27

Statistical analysis 29

3. Results: Persistent lesions and MTB mRNA detected after treatment 31
3.1 Patient demographics and treatment outcome 31
3.2 Month 6 PET/CT findings in 99 South African patients 34
3.3 PET/CT findings in 14 South Korean patients 36
3.4 PET/CT findings 1 year after completion of treatment 38
3.5 DNA and mRNA in M6 sputum 40
3.6 RNA in end of treatment bronchoalveolar lavage samples 42
3.7 Statistical Associations between grouped variables 43
3.7.1 Scan response patterns and clinical outcome 43
3.7.2 M6 Sputum Xpert and outcome 45
3.7.3 M6 Sputum mRNA vs Outcome & Diagnosis 45
3.7.4 Broncho Alveolar lavage mRNA vs Outcome & diagnosis 46
3.7.5 Adherence, Smoking, Previous PTB vs outcome 47
3.7.8 Gene Xpert and M6 PET/CT scan 48
3.7.8 M6 Sputum mRNA and M6 PET/CT scan 49
3.7.9 Adherence and M6 PET/CT scan 49
3.7.10 M6 Xpert and M6 mRNA positivity 50
3.7.11 Poor adherence and Xpert positivity 51
3.7.12 mRNA Copy number rank correlation 52
4. Discussion: Persistent lesions and MTB mRNA detected after treatment 55
Summary 55
Significance of lesion activity after treatment 55
Significance of MTB DNA and mRNA detection after treatment 56
Strengths and limitations 57
Conclusion 57
5. Results: Operator independent technique to quantify widespread complex tuberculous lung lesions on PET/CT. 59
5.1 Application of quantification technique 59
Time requirements 59

Segmentation accuracy 59
Pilot case profiles 62
5.2 Quantified scan results from diagnosis, month 1 and month 6. 64
PET/CT characteristics in relation to time to culture negativity group. 64
Correlation between imaging parameters 70
Prognostic factors of failure 74
Prognostic factors of recurrence 78
Criteria for unfavourable outcomes 80
5.3 Quantified scan results: EOT + 1y. 82
Residual lesions 82
Correlations between M6 and EOT + 1 year findings. 85
Differences between recurrence and maintained cure. 87
5.4 Statistical associations with quantified PET/CT parameters. 88
Association with Microbiological markers 88
Association with clinical factors 91
6. Discussion: Operator independent technique to quantify widespread complex tuberculous lung lesions on PET/CT. 93
6.1 The use of the quantification technique 93
Summary 93
Novel aspects of quantification technique 93
Strengths and limitations of technique 93
Conclusion 94
6.2 Quantified PET/CT results. 95
Summary of main findings 95
Comparison with previous literature. 96
Study Limitations 98
Implications 99
Future research 100
Conclusions 101

7. Using quantified PET/CT variables to aid biomarker discovery 102
7.1 Lesion burden and host inflammatory cytokines 102
7.2 Lesion burden and host gene expression. 106
7.3 Discussion: Biomarker discovery 107
8. Concluding remarks 108
Realisation of study objectives. 108
Dissemination 109
Personal statement 110
Appendices 1
Appendix A Patient demographics, outcome and scan response pattern. 1
Appendix B EOT + 1y scan response patterns per patient. 3
Appendix C South Korea patient demographics, outcome and scan response. 4
Appendix D Sputum MTB mRNA copy numbers and clinical status. 5
Appendix E MTB mRNA copy numbers in Bronchoalveolar lavage and clinical status .. 7
Appendix F Primer sequences for MTB mRNA 9

List of Figures

Figure 1: Flow chart of study design. 16
Figure 2 Example of Volumes of interest (in red) required for quantification. 20
Figure 3 Range, median, 25th and 75th percentiles for percentage of total lung volume classified as FDG-avid at different z-score thresholds. 22
Figure 4 PET images 2 patients after treatment for pulmonary tuberculosis. 23
Figure 5 Scatter-plots of PET/CT scans from three healthy controls and two PTB patients during standard treatment. 25
Figure 6 Frequency of strain clades cultured from sputum at Dx for South-African cohort. 33
Figure 7 PET/CT images at Dx, M1 and M6 for representative cured patients after 6 months of standard treatment. 35
Figure 8 PET/CT images for representative cases at Dx and M6 from South-Korean cohort. 36
Figure 9 Clinical outcome and associated M6 PET/CT scan findings. 37
Figure 10 Dx, M6 and EOT + 1y PET/CTs for 3 representative patients after 6 months of treatment. 39
Figure 11 Clinical outcome and associated PET/CT scan findings 1 year after the end of successful PTB treatment (EOT + 1y). 40
Figure 12 Box and whisker plots of the transformed PC1 of mRNA transcript PCR data 41
Figure 13 Box and whisker plots of the transformed PC1 of mRNA transcript PCR data 43
Figure 14 Trend association between clinical outcome and scan response pattern. 44
Figure 15 Loess regression of gene expression rank of cured patients on the rank of the failed patient. 53
Figure 16 Ordinary least squares linear regression of mean rank of genes of cured on rank of failed case. 54
Figure 17 Comparison of PET/CT parameters between scans from 5 controls and 5 PTB patients at diagnosis. 61
Figure 18 case-profiles of 5 PTB patients showing auto-segmented PET/CT parameters at Dx, M1 and M6. 63
Figure 19 Mean (±SE) Log10 transformed values of principal PET and CT parameters over time by time to negativity group and recurrent cases. 65
Figure 20 Mean (±SE) Log10 transformed values of PET parameters over time by time to negativity group. 67
Figure 21 Mean (±SE) Log10 transformed values of CT parameters at Dx, M1 and M6 by time to negativity group. 68

Figure 22 Mean (±SE) Log10 transformed values of combination PET/CT parameters at Dx, M1 and M6 by time to negativity group. 69
Figure 23 Correlations matrix of quantified scan parameters at Diagnosis. 70
Figure 24 Correlation matrix of quantified scan parameters at M1. 71
Figure 25 Correlation matrix of quantified scan parameters at M6. 72
Figure 26 Correlation between values of imaging parameters and TTN 73
Figure 27 Correlation between imaging parameters' change from baseline and TTN 74
Figure 28 Example of PET/CT scan of failed treatment case at Dx, M1 and M6. 75
Figure 29 Receiver operating characteristic (ROC) curves of cavity volume predicting failed treatment 76
Figure 30 Box and whisker plot of M6 cavity wall thickness, showing the median, 25^{th} & 75^{th} percentile (Box) and range (whiskers). 77
Figure 31 Examples of cavity lesions on M6 CT images. 78
Figure 32 Mean ± SE ± 1.96xSE Change in TGAI for cured and recurrent groups 79
Figure 33 Mean ± SE ± 1.96 x SE Change in $TGAI_{com}$ for cured and recurrent groups 80
Figure 34 Box and whisker plot showing median, quartiles and range for TGAI at Dx, M6 and EOT + 1y. 83
Figure 35 Box and whisker plot for Cavity volume at Dx, M6 and EOT + 1y. 84
Figure 36 Box and whisker plot for MLV and V_{total} at Dx, M6 and EOT + 1y. 85
Figure 37 Dx, M1, M6 and EOT + 1y PET/CTs for 3 representative cases. 86
Figure 38 Scatterplots to show correlation of PET/CT scan parameters at M6 and EOT + 1y. 87
Figure 39 Association between M6 Xpert positivity, M6 TGAI and TGAI change from Dx. 88
Figure 40 Correlation between the number of MTB mRNA transcripts detected in EOT BAL and M6 cavity volume. 89
Figure 41 Scatterplot of correlation between TTP and $TGAI_{com}$. 90
Figure 42 Association between grouped strain clades and change in cavity volume from Dx to M1&M6. 91
Figure 43 Association with between a history of previous PTB episode(s) and TGAI at Dx and M6. 92
Figure 44 ROC curves to show accuracy of host-protein markers to indicate low TGAI at Dx. 103
Figure 45 The host proteins used in the 12-marker mathematical model. 104
Figure 46 ROC curves to show accuracy of host-protein markers to indicate high TGAI at Dx. 105

List of Tables

Table 1: Clinical and microbiological parameters for PTB patients 32
Table 2 Summary of evidence supporting diagnosis of PTB recurrence. 33
Table 3 Cross tabulation of clinical outcome and scan response pattern. 44
Table 4 Association between scan response pattern and collapsed clinical outcome 44
Table 5 Association between clinical outcome and M6 Xpert result. 45
Table 6 Association between clinical status and sputum MTB mRNA. 45
Table 7 Association between collapsed clinical status compared to sputum MTB mRNA. 46
Table 8 Association between clinical TB status and BAL MTB mRNA. 46
Table 9 Association between collapsed clinical status and BAL MTB mRNA 47
Table 10 Association between clinical outcome and whether patient took >80% of doses . 47
Table 11 Association between smoking status and collapsed clinical outcome 48
Table 12 Association between collapsed clinical outcome and previous PTB episodes 48
Table 13 Association between grading of most intense M6 PET lesion and M6 sputum Xpert. 48
Table 14 Association between M6 PET/CT response pattern and M6 sputum Xpert. 48
Table 15 Association between grading of most intense M6 PET lesion and M6 sputum Xpert 49
Table 16 Association between grading of most intense PET lesion and whether patient took > 80% of TB doses. 49
Table 17 Association between adherence groups and scan response pattern. 50
Table 18 Association between adherence groups and M6 scan response pattern. 50
Table 19 Association between M6 sputum MTB mRNA and Xpert positivity. 50
Table 20 Association between M6 Sputum Xpert positivity and whether patient took >80% of doses. 51
Table 21 Association between M6 Sputum MTB mRNA positivity and whether patient took >80% of doses. 51
Table 22 Ranked mRNA transcript, on copy number in BAL from failed treatment case. 52
Table 23 Summary of Contingency table statistics for scan parameters. 80
Table 24 Association between mixed scan response patterns at M6 and EOT + 1y. 85
Table 25 Student's T-test results for difference in M6 PET/CT parameters and M6 Xpert positivity. 88
Table 26 T-test results testing for difference in M6 PET/CT parameters and grouped MTB strains. 90
Table 27 T-test results testing for difference in M6 PET/CT parameters and history of previous PTB. 91

List of Abbreviations

AFB	Acid-fast bacilli
AUC	Area under the curve
BAL	Bronchoalveolar lavage
CFU	Colony forming units
CXR	Chest X-Ray
Dx	Diagnosis
DNA	Deoxyribonucleic acid
DS	Drug-sensitive
EMB	Ethambutol
EOT	End of treatment
EOT + 1y	1 year after end of treatment
EOT + 2y	2 years after end of treatment
FDG	18F-Fluorodyoxyglucose
HRZE	Standard 4 drug anti-TB treatment
HU	Hounsfield Units
INH	Isoniazid
kBq	Kilobequerel
kg	Kilogram
LCC	Low copy clade
M1	Month 1
M6	Month6
MBq	Megabequerel
MGIT	Mycobacterial Growth Indicator tube
ml	millilitres
mRNA	messenger RNA
Zmean	Mean standardised intensity
MTB	Mycobacterium tuberculosis
MTV	Metabolic tumour volume
NL	Normal Lung/Background lung reference
NPV	Negative predictive value
OLD	Other than TB lung diseases
PCR	Polymerase chain reaction
PET/CT	Positron emission tomography/computed tomography
PHC	Primary Healthcare clinics
PPV	Positive predictive value
PTB	Pulmonary tuberculosis
PZA	Pyrazinamide
QFN	Quantiferon TB Gold Test™
qRT-PCR	Real-time PCR
RIF	Rifampicin
RNA	Ribonucleic acid
ROC	Receiver operating characteristic
rRNA	ribosomal RNA
RT	Reverse transcriptase
SUV	Standardised uptake value
TB	Tuberculosis

TGA	Total glycolytic activity
TGAI	Total glycolytic activity index
TGAIabn	Total glycolytic activity index in high density areas
TGAIcav	Total glycolytic activity index of cavity volume.
TLV	Total lung volume
TTP	Time to liquid culture positivity in days
Vabn	Volume with high density (> -500 HU) and high intensity (> Z-score 7)
Vhard	Hard lesions volume (> -100 HU)
Vlow	Hypodense lesions (< -950 HU)
Vmedium	Medium lesion volume (-300HU to -100 HU)
VOI	Volume of interest
Vsoft	Soft lesion volume (-500 HU to -300 HU)
W12	Week 12
W24	Week 24
W24+	After week 24
W4	week 4
W8	Week 8
WHO	World Health Organisation
XDR	extensively drug resistant
Xpert	GeneXpert MTB/Rif assay
Z	Z-score/standardised score

1. Introduction

1.1 Context and overview of research project.

I was given the opportunity to become part of the '*Mycobacterium tuberculosis (MTB) biomarkers for diagnosis and cure*' project. The project was a collaborative effort between the Stellenbosch University Immunology Research Group, the Tuberculosis Research Section of the National Institute of Allergy and Infectious Diseases, National Institutes of Health, the Catalysis Foundation for Health, the Center for Emerging Pathogens Rutgers-New Jersey Medical School, the Center for Infectious Disease Research and the International Tuberculosis Research Center, Seoul, South Korea. Funding was provided by the Catalysis Foundation for Health.

The aim of the study was to recruit newly diagnosed patients with pulmonary tuberculosis in Cape Town and Seoul and observe them throughout treatment. Serial collection of clinical data and established microbiological tests would be carried out. 18F-FDG PET/CT scans were to be performed at baseline and during treatment. Since PET/CT scans have shown promise as a sensitive tool to monitor treatment response, the objective was to use the results in conjunction with the collected clinical and microbiological data to serve as a surrogate for MTB load. In addition, to help identify patients who failed to reach sterilising cure, in spite of sputum culture negativity. The combined surrogate would then be used to aid the discovery of novel biomarkers in collected body fluids: to correlate with MTB load, distinguish cure from unfavourable outcomes and shed light on underlying scientific mechanisms.

We successfully completed recruitment and follow-up in Cape Town, however, due logistical problems, recruitment targets could not be met in Seoul.

I came on board during the recruitment phase of the study. As a study clinician, I managed the nursing staff, supervised participant recruitment and assisted in procedures, like PET/CT scans and bronchoscopies. I also reviewed test results from a clinical perspective and quickly got involved in scientific analysis. The focus of my attention was drawn to the surprising PET/CT images: the different lesion responses in the same pair of lungs and the persistence of lesions that appear active toward the end of treatment – even in cases that achieved clinical cure. The nuclear physicians reporting the scans were baffled and research questions begged to be answered: 1) What are the implications of the inflammatory lesions at the end of treatment? 2) Can specific scan characteristics be found that correspond to the clinical and microbiological information? 3) Can these characteristics

be measured in a reproducible manner to serve as a reference for further biomarker discovery?

To look for answers to the first question we applied for an amendment to the study protocol to perform bronchoscopies with alveolar lavage at the end of treatment and an additional PET/CT scan 1 year after the end of treatment on the subgroup of the participants that have not yet passed this time-point. The bronchoalveolar lavage specimens were analysed for any evidence of persistent viable MTB, in spite of sputum culture negativity. The additional scans were evaluated to observe the natural history of the persistent lesions seen at the end of treatment.

I would like to provide an outline of this dissertation.

The literature review, which follows this introduction, would firstly provide an overview of why it is important to find improved biomarkers. It gives a summary of tuberculosis epidemiology and some detail about the rate of unsuccessful treatment and recurrence – emphasising the need for improved tuberculosis therapy. I also discuss the shortfalls of tests that are currently used in programmatic healthcare and clinical trials to monitor the response to treatment, as well as the uncertainty that still exists when it comes to defining sterilising cure. The second section relates to PET/CT scans. I first provide some background regarding the technology and most common applications. Subsequently, I discuss very interesting results from PET/CT studies in animal TB models, as well as available literature on PET/CT scans in human patients with tuberculosis. I discuss standard tools used to report PET/CT findings as reproducibly as possible, in addition to recently described techniques to segment lesions with limited user input, plus methods to normalise and quantify PET/CT scans.

The last part of the introductory chapter 1 contains the research question and aims of the project.

The second chapter contains the methods and procedures regarding the recruitment, follow-up and medical management of participants, as well as research tests performed. In this chapter, I describe how the scans were evaluated qualitatively and include a detailed description of how the operator independent technique to quantify scan images were developed and applied.

The main results are separated into two chapters, each with a following chapter dedicated to a discussion of the respective results provided. In chapter 3, I provide the participants' demographical data and clinical outcomes, before documenting dynamic scan response patterns observed qualitatively and results of in depth microbiological sampling and testing

after standard treatment. The last section of chapter 3 is dedicated to extensive statistical testing exploring any meaningful or confounding associations between clinical profiles, microbiological test results and scan response patterns. Chapter 5 is dedicated to detailed results of the scan quantification by firstly discussing the performance of the user independent technique. The subsequent section in the chapter explores the various quantified scan characteristics and how well different characteristics correlate (alone or in combinations) to clinical profiles and treatment outcomes at the different time-points.

Chapter 7 is a relatively short chapter providing some preliminary results to serve as examples of how the quantified scan parameters are already being used as a reference for the discovery of novel biomarkers in peripheral blood.

In chapter 8, I wrote a final wrap-up discussion and concluding remarks.

I hope you find this work interesting.

Sincerely

Stephanus

1.2 Literature review

1.2.1 The challenge of tuberculosis

TB incidence and treatment outcome

Tuberculosis (TB), with an estimated 9,6 million new TB cases and 1,2 million TB related deaths in 2014 as well as increasing numbers of cases with drug-resistant strains, presents a massive health care problem around the world, particularly in low and middle-income countries like South Africa.[1]

The current standard treatment period of 6 months for drug-sensitive TB was determined by acceptable rates of treatment failure and disease recurrence after discontinuation of chemotherapy.[2] In literature, the reported rate of unfavourable outcomes varies considerably, however it is usually disconcerting.[3] Unfavourable outcomes include failure to convert to sputum culture negative, treatment default and disease recurrence, which could be due to either endogenous relapse or exogenous reinfection. Clinical trials utilise a strict inclusion and exclusion criteria, adherence counselling and monitoring. In recently published multi-national trials failed treatment rates were relatively low and ranged from 1.3%–2.4%. Relapse rates within 12 – 24 months after successful standard treatment were somewhat higher at 2.6%,[4] 3%[5] and 5%.[6] Combined data from trials in Uganda, however found unsuccessful treatment outcomes (failure and default) in 8.5% of participants and a recurrence rate of 10%.[7] Unsuccessful treatment reached 19% in a group of South-African miners and the relapse rate 9.7%[8]

The outcome reported in programmatic conditions around the world, is considerably worse. According to the World Health Organisation's (WHO) Global TB Report 2015,[1] the global rate of unsuccessful treatment in HIV-negative patients with drug-sensitive (DS) TB is 12%. Of the 22 countries that carry 80% of the world's TB burden, 17 had unsuccessful treatment rates higher than 10%. Five of these countries (including South Africa), as well as the greater African region, European region and the Region of Americas all report unsuccessful treatment rates greater than 20%. India, with the world's largest absolute TB burden, and with low HIV co-infection rates (4%), reports unsuccessful treatment in 12 % of cases. In addition, 18.5% of notified cases were recurrences. Russia and Brazil, also listed as high burden countries, reported that respectively 36.1% and 16.1% of notified cases were recurrences. Although longitudinal data are not available in these reports, estimation based on other publications would suggest that more than half of the recurrent cases previously completed treatment within the previous 2 years.[9–11]

A recently published, population-based cohort study, conducted in Cape Town,[9] retrospectively followed up 2359 acid-fast bacilli (AFB) smear positive pulmonary tuberculosis (PTB) patients and found successful treatment in only 79.2%. Of the 1869 patients (which included only 100 HIV co-infected cases) that successfully completed treatment, retreatment episodes were documented in 309 (16.5%), of which 286 were confirmed by AFB smear or culture. Smear positive recurrence was detected in 203 cases, of which only 10 were HIV positive. The median time to retreatment in these cases was 21 months. The TB prevalence in Cape Town is sadly amongst the worst in the world,[12] with the interesting note that the HIV co-infection rate is lower than in other high TB prevalence areas in South Africa.

Thus, there exists discrepancy between trials and observational studies and national health programs, which may be due to the strict inclusion and exclusion criteria, as well as more active adherence counselling and monitoring.

Clinical cure, sterilising cure and post-TB lung sequelae

The absence of a fail-proof gold standard to determine when antibiotic treatment has induced sterilising cure in PTB confounds new approaches for measuring treatment response. Although it has been suggested that in trials, follow-up after treatment could be shortened due the most frequent timing of relapse, the standard follow-up period for phase 3 TB treatment trials is 18 – 24 months after treatment completion[13]. This greatly contributes to the costly and time-consuming nature of trials testing new TB treatment approaches.

The current treatment response measurement recommended by the WHO is still based on sputum AFB smear.[14] AFB smear positivity is only 61% - 70% when compared to MTB culture at baseline.[15] A meta-analysis, found smear positivity at month 2 and month 5 a poor predictor of outcome.[16] The pooled sensitivity for a positive month 2 smear to predict treatment failure was 57% and 24% to predict recurrence [positive predictive value (PPV) of 9% and 10%] with a specificity of respectively 81% and 85% [negative predictive value (NPV) of 98% and 93%]. A positive month 2 sputum culture performed only slightly better, found to be 40% sensitive and 85% specific for relapse. Month 2 culture conversion in liquid media has also been suggested as a surrogate marker for favourable response to treatment, but has not been proven to be accurate enough.[13,17]

Shortening treatment is also a major aim of anti-tuberculosis drug development.[18] Four trials[4–6] introduced an experimental 4-month treatment arm, which yielded significantly increased relapse rates, even though sputum culture conversion rates were higher within the first 2 months of treatment.[4–6,19,20] The first trial selected low-risk patients, based on a criteria

requiring absence of cavities on baseline chest x-ray and sputum culture conversion by month 2.[19] The three other trials tested treatment regimens, respectively containing: Gatifloxacin,[6] Rifapentine with Moxifloxacin,[5] or Moxifloxacin[4] to replace components of standard therapy. These regimens appeared promising in pre-clinical studies to induce more rapid MTB killing and culture negativity. The increased relapse rates, in spite of earlier culture negativity, strongly suggests that *Mycobacterium tuberculosis* (MTB) can persist in lung tissue for months to years after culture negativity has been achieved through antibiotics and the important role that MTB sub-populations (unculturable in sputum) plays in TB treatment relapse. It also suggests a need for better understanding the scientific mechanism of a favourable clinical response to TB treatment.[20]

In contrast to sputum culture conversion, lung lesions seen on chest X-ray (CXR) or computed tomography (CT) often persist long after the end of successful treatment.[21,22] A retrospective Korean study analysed CT scans before and after PTB treatment in 66 patients and found a reduction in lesion size in 61 (92%) of the cohort. Thirty-three patients (50%), however, still had at least one lesion with features consistent with active disease. None of the patients relapsed in the 15 month follow-up period, thus no causal link between residual lesions and relapse was found.[21] A chest x-ray severity grading score showed 80% reduction during standard TB treatment, but not complete resolution in most cases.[22] The score did show promise as a clinical outcome predictor, but this could not be validated in a separate study.[23]

Similar to residual lung lesions, persistent symptoms and signs after microbiologically confirmed cure are also common, due to post tuberculosis lung impairment.[24–26] Lung impairment after TB disease could be due to a restrictive pattern (loss of forced vital capacity), an obstructive pattern (decreased forced expiratory volume), or a combination of both. Lung impairment is present to varying degrees in most cases after the successful treatment of pulmonary tuberculosis.[27–29] Both these patterns of functional impairment increase incrementally after repeated episodes of PTB.[28] The Post tuberculosis pulmonary impairment contributes roughly 75% of TB-related disability burden.[30]

Another tool used for TB diagnosis, the polymerase chain reaction-based, GeneXpert MTB/Rif assay (Xpert) remains positive after treatment in roughly 30% of clinically cured patients.[31,32] This is believed to be due to the relative stability of MTB DNA in alive and dead cells.

The persistence of lung lesions, clinical symptoms and MTB DNA after clinical cure, is in contrast to sputum culture negativity before sterilising cure. This contributes to uncertainty

when defining sterilising cure and indicate a great need for improved methods to measure treatment response.

The need for a protracted treatment duration with multiple anti-TB agents, can in part, be explained by the presence of sub-populations of slow-replicating MTB that are naturally tolerant to all TB drugs.[33] Anti-tuberculosis therapy introduces a selection process, and surviving MTB are progressively slower to replicate and more tolerant to drugs. Slow-replicating MTB are also more difficult to grow in culture mediums, which can manifest as persistent, viable, non-culturable MTB. This has been demonstrated in mice after 14 weeks of isoniazid and pyrazinamide treatment when MTB mRNA transcripts (shown to be absent in dead bacteria *in vitro*) were detected in culture negative tissue from all mice necropsied at this time-point. High-dose steroid-induced bacteriological relapse occurred in 21 of 23 mice that were followed up[34]. The presence of mRNA could suggest transcriptionally active bacteria due to its very short half-life (on average 9.5 minutes for MTB *in vitro*), although stabilisation of transcripts occurs when exposing the intact MTB to stresses (cold shock and hypoxia) and therefore the *in vivo* half-life is not known.[35]

The retention of low-abundance mRNA and a relative increase of non-coding RNA in dormant MTB have been shown during *in vitro* persistence induced by potassium depletion and Rifampicin exposure and might play a role in early reactiviation.[36] Persistent bacteria have, to date, not been observed in humans as they have been in mice and *in vitro* culture.

Slow replicating, resilient MTB are also selected by stress induced by the immune response. It has been shown that the MTB cultured from sputum, mostly originate from the wall surface of cavities that communicate with the airways. On histology, this area harbours the most abundant bacterial load, as well as the most actively replicating populations of bacteria. Other lesion types harbour much lower numbers of bacteria and those that are present in these lesions, have shown an altered metabolic state which make them more resistant to growth in culture.[37] These bacteria may play an important role in endogenous relapse.

Several different MTB virulence factors have been identified.[38] They include: 1) promotors of growth and metabolism, such as enzymes of lipid pathways, cell wall mediators, regulatory systems. 2) factors that protect against the stressors of host defence, such as nitrosative and oxidative stress, phagosome inhibition and latency. 3) specific virulence compounds, like proteases, secretion systems and sigma factors. These factors varies between strains and between different phenotypes, contributing to different manifestations of TB pathology and levels of drug tolerance.[39] The Beijing strain is described as an especially virulent strain, It predominates geographical areas (including the Western Cape) and has been associated with outbreaks world-wide.[40] MTB virulence factors also play a role in treatment

response and the Beijing strain has been linked to a higher risk of persistent culture positivity after treatment (failure).[41,42]

In addition to bacterial factors, multiple patient related factors influence severity of TB disease and treatment response. Poor adherence, cavitatory or extensive lung lesions and impaired immunity (e.g. HIV and diabetes) are all associated with a higher incidence of treatment failure and relapse. [7,8,17,43]

An adequate response to treatment, firstly require adherence.[43] Apart from adherence, multiple factors could lead to substandard drug absorption, which are also independently associated with an unfavourable outcome.[44,45] In an animal model[46] and a study on (human) spinal TB,[47] drug concentrations were shown to vary widely, influenced by the type of lesion and the individual drug.

Conditions leading to impaired immunity, such as HIV and diabetes, are also independent risk factors for disseminated disease and an unfavourable treatment outcome – in spite of a tendency to be associated with less cavitation and a higher incidence of sputum culture negative TB.[8,9,17,43] This further highlights the pitfalls of using sputum culture as a measure of treatment response and suggests an important complimentary role of the immune system in a successful response to treatment.

1.2.2 ^{18}F-Fluorodeoxyglucose positron emission tomography and computed tomography (PET/CT)

Background

Positron Emission Tomography/Computerised Tomography (PET/CT) is well established in the staging, treatment planning and response assessment of cancer and various inflammatory and infectious diseases.[48,49] ^{18}F-fluorodeoxyglucose (FDG) is the most commonly used PET tracer. It reflects glucose metabolism and shows increased uptake in areas of inflammation. PET scanners measure the radiopharmaceutical concentration in tissue (kBq/ml).[50–52]

Uptake intensity in tissue is variable and influenced by numerous patient and equipment factors, which is why the lesion- to- background ratio is often a more robust measure.[53] Several semi-quantitative measurements of intensity have been developed, of which, Standardised Uptake Value (SUV), is most commonly used and compensates for variation in body size, injected activity and decay. It uses the formula:

$$SUV = \frac{Tracer\ concentration\ (\frac{kBq}{ml})}{Injected\ dose\ (MBq)\ .\ patient's\ weight\ (kg)}$$

Segmentation of the Metabolic Tumour Volume (MTV), i.e. the anatomical volume of the FDG-avid lesion, is essential for the generation of more comprehensive parameters. The simplest method to quantify a lesion uses the most intense voxel in the lesion (SUV_{max}). This, however, is prone to variability and disregards information from the remainder of the lesion. The MTV allows the mean (SUV_{mean}) to be calculated. Total Glycolytic Activity (TGA) is the product of the mean intensity (SUV_{mean}) and the volume of an area of interest. [50,51,54–56]

SUV, however, are also influenced by most of the same factors influencing uptake intensity, such as body composition, with an overestimation of intensity occurring in obese patients.[53,57] Using the body surface area or the lean body mass to calculate the SUV have been proposed to compensate for this. A high blood glucose concentration reduce FDG uptake in tissue and needs to be well controlled prior to FDG administration. Time between administration and image acquisition should also be optimised and standardised to improve consistency in results. The body's surface and other anatomical structures cause scatter and attenuation of the radio-signal derived from FDG. This is usually partially corrected by computer algorithms using anatomical data from the CT scan during reconstruction of the images.[58] Many different algorithms are used for attenuation correction and reconstruction of the PET images, adding to inter-scanner variability.

Animal PET/CT studies

In mouse, rabbit and primate models, FDG PET/CT has been used to accurately display the progression of disease after exposure to Mycobacterium tuberculosis (MTB) and the response to treatment.[59–62]

Via et al. (2012) provided a detailed documentation of the dynamics of lung lesions on PET-CT in rabbits after infection with MTB and also after treatment initiation.[61] After infection with MTB by aerosol, serial PET/CT's were performed and animals necropsied at various time-points. In tissue, the CFU (colony forming units) count raised within the first 2 weeks, while granulomatous lesions became apparent on histology by 4–5 weeks. The bacterial load also peaked at this stage. After that time-point, the CFU count dropped somewhat and stabilised, as the animals entered the chronic phase of infection. On PET/CT, a mild, homogenous increase in background lung uptake was noticed and FDG-avid lung nodules were seen from week 4. After week 4, the progression of individual lesions were found to be markedly different. Some continued to get bigger, others formed cavities and some resolved. If a lesion was in the process of resolving, the individual lesions showed decreasing PET intensity, followed by a reduction in density on CT before ultimate volume reduction. After

the increase in FDG-avidity and CT density of lesions during the acute phase, the chronic phase of infection coincided with a stabilisation of the SUVmax. The maximum CT density in the lung and the percentage of lung volume affected also stabilised, but dynamic changes were still found in individual lesions.

On rifampicin or isoniazid treatment, a significant and progressive reduction in lesions' FDG-avidity was seen from week 1, with a sharp decrease towards week 4 and some stabilisation toward week 8. This was closely followed by a reduction in CT density and ultimate volume loss. In untreated rabbits there were a trend for some correlation between FDG-avidity in lesions and bacterial numbers, while a significant correlation was found between FDG-avidity (measured in SUVmax) and both pathology and bacterial numbers.

Similar responses were shown in MTB infected cynomolgus macaques in collective works by Lin, Coleman, Flynn at al.[59,62–64] Additionally, this also provided a model of latency[64]: some animals progressed to active disease, while others managed to suppress the initial infection and maintain latency. The course of latent disease was characterised by some FDG-avid granulomas (visualised as small nodules on scan) and increased lymph node intensity within 3 weeks of infection. The nodes and the granulomas in latent disease typically did not progress and resolved over time. Progress to active disease was characterised by comparatively more nodules at week 3, which showed a rapid increase in number, size and uptake intensity over the course of the following 4 to 5 weeks. At that stage, the lymph node intensity in this group also tended to be higher. The course of individual lesions varied in terms of size and SUV intensity within the same host, while the central tendency of lesion size, SUV intensity and numbers within a host was associated with outcome. This report also noted the 'percolator' phenomenon in one primate, referring to a state between latency and active disease, characterised by steady state of individual lesion changes, without progress to overt deterioration in clinical state of the host. Macaques showed a similar response to anti-TB treatment compared to rabbits.[59] In untreated subjects, the SUV intensity of lesions weakly correlated to the M. tuberculosis CFU count, while a good correlation with size and MTB load were found.[62] On treatment, the reduction in SUV showed a negative correlation with the activity of the chemotherapy used and also correlated with the CFU counts.[59] On mono-therapy, there were individual lesions that increased in size and SUV and residual activity in apparently sterilised lesions were also noted in some cases.[59]

There are also some other notable animal models found in literature. A difference in the disease progression and complications were noticed on scans of Marmosets infected with MTB strains with different virulence factors.[65] In a mouse model, PET/CT was also able to detect the development of relapse before microbiological evidence.[66]

Human PET/CT studies

There are relatively few reports available in literature regarding the potential use of PET/CT scans in tuberculosis.

PET/CT has been used to determine the risk of malignancy in solitary pulmonary nodules. The use of PET/CT has reduced the number of unnecessary thoracotomies.[67] PET/CT is also used to evaluate mediastinal lymph nodes in staging non-small-cell lung cancer, to provide a prognosis and guide treatment.[68] In both indications, however, the accuracy decreased in TB endemic areas, due to the overlap between the range of lesion PET intensities associated with malignancies and TB, thus introducing the need for adjusted diagnostic thresholds and reducing the benefit of PET/CT.[67–69]Human studies with small sample sizes have shown promise in using PET/CT as a tool to monitor TB treatment response. A prospective study found decreased uptake intensity after 1 month treatment of pulmonary (n=10) and extra-pulmonary (n=10) TB.[70] The authors reported a 31% median decline in the SUVmax of each individual's most intense lesion (range 2%—84%).

Due to the invasive procedures required in getting a tissue specimen for culture and the longer treatment duration required (12—18 months), spinal tuberculosis has sparked an interest in the possible use of PET to monitor treatment response. A small prospective study showed a steady decline in lesion uptake intensity (measured in SUVmax) after 6, 12 and 18 months of treatment for spinal TB (n=18).[71] The mean improvement in the first 6 months were 57% and at month 18, the SUVmax ranged between 1.1 and 2.4. A Japanese study reported the PET/CT findings of patients with MTB and Mycobacterium Avium Complex disease over the course of treatment. Scan findings and responses in the two groups were very similar. Treatment duration for the MTB group were 12 months and in the eight MTB patients that a had follow up scan after the end of treatment, no lesions with significantly increased uptake were found.[72] A European study on 35 patients with pulmonary or extra-pulmonary TB reported persistent PET/CT lesions in 15 and progressive lesions in four patients after a mean treatment duration of 16.1 months.[73] In patients with TB and HIV co-infection PET/CT has shown promise to predict treatment response based on the amount of lymph-node bastions involved at baseline.[74]

In a trial, linezolid was added to the therapy of patients with extensively drug-resistant (XDR) PTB and PET/CT scans were performed at baseline, month 6 and either month 1 (n=4), 2 (n=4) or 3 (n=5). Whole lung analysis with fixed threshold were performed on the scans and showed a decrease in total glycolytic activity (TGA) and Gross Tumour Volume (GTV) over the treatment period, when compared to patients not receiving the additional chemotherapy. The most marked improvement was noted within the first month.[60]

Another trial tracked the quantified changes from baseline on PET/CT after 2months and CT after 6 months of treatment in 28 patients with multi-drug resistant (MDR) TB.[75] The investigators made use of whole lung quantification of PET, using fixed thresholds; and reader scores to describe the CT. Quantified measures of PET images showed less inter-reader variability than CT reader scores. The reduction in total glycolytic activity over 2 months accurately predicted outcome, while CT volume changes at month 6 were also predictive. This study also throuroughly documented the density ranges associated with different lesion morphology on CT. Nodules and tree-in-bud lesions were usually between -500 and -200 Hounsfield Units (HU). Bronchial thickening was from -300 to 100 HU and more established lesions, such as consolidation; cavity walls, fibrosis and collapse usually showed densities between -100 and 100 HU.

PET/CT lesion segmentation

During segmentation, the borders of the metabolic tumour volume can be delineated visually or using various semi-automated techniques including thresholding or gradient-based boundary techniques. However no single technique proven optimal for all applications.[53] Lesion delineation in most cases is still performed manually, based on visual interpretation of PET or CT images. This is prone to inter- and intra-operator variation, especially for PET, due to the difficulty of delineating anatomical boundaries using functional information, in addition to low resolution and inherent noise.[55–57,76–80]

Multiple methods are used and proposed to decrease variation in lesion segmentation. These include using reference values to normalise the lesion to background uptake intensity by comparison to liver or mediastinal blood pool uptake.[81–83] A five-point grading system, referred to as the Deauville score, is commonly used as a standard aid in describing lesion intensity in clinical reporting and density: 1) No uptake above background. 2) Lesion intensity equal or less than the mediastinal blood pool. 2) More uptake than the mediastinum, but equal or less than the liver. Uptake moderately increased compared to the liver. 5) Uptake markedly increased compared to the liver at any site.

Reproducible segmentation and quantification become particularly important in diseases with heterogenic morphology, vague borders and multi-focal distribution throughout an organ or system, such as TB or Sarcoidosis.[60,75] This is especially important when accurate tracking of changes over time is required. To account for the spatially complex lesions associated with pulmonary tuberculosis, computer aided segmentation, based on affinity propagation, has shown high accuracy and increased efficiency when tested in PET images of TB infected rabbits.[84]

Evaluating the accuracy of lesion segmentation techniques can also prove to be challenging. A ground truth (objective gold standard segmentation), to measure the segmented volume against must be established. The gold standard is histology – although it can be difficult to compare the size of a tissue sample after surgical removal to a scan before removal. In the absence of histology, the technique could be measured by testing it on a phantom scan, but this could not always be compared to real-life physiology. Another option is creating a surrogate ground truth, by getting one or more experts to delineate the boundaries of the lesion. If more than one expert is employed, the different segmentations is statistically combined.

The most commonly used segmentation technique relies on thresholding.[53] Fixed thresholding delineates the voxels above an assigned intensity value. Adaptive thresholding uses an assigned scan specific background reference and one of various analytical expressions to determine the edges of the segment. Region-based segmentation usually requires the user to define a point in the lesion and the technique delineates the boundaries according to the gradient. More complex computer learning-based methods are also used. The 'fuzzy-locally-adaptive Bayesian' method has shown to be robust and reproducible. This uses the spatial relationship of the 'fuzzy' (difficult to classify) areas to the lesion or background segments to choose to which class it belongs. However, it has also show variable accuracy in heterogenic lesions.[85]

A phantom study compared 6 different proposed auto-segmentation methods and found adaptive thresholding techniques using a reference normal background volume of interest (VOI) in the formulas to improve accuracy.[80]

Some work has also concentrated on the quantification of CT. CT images, acquired with PET provide anatomical information. CT scanners measure the radio-opacity, or density of anatomical structures and lesions in Hounsfield Units (HU). It is not a functional scan and is less prone to variability. Joint-segmentation, which implies using information from both modalities, of fused PET and CT images have shown improved robustness.[86] Only a few studies have evaluated densometric quanitification of CT scans in diffuse lung disease. [87–90] It appeared to be reproducible and correlate well with other disease markers. It was used to successfully track changes over time in patients with idiopathic pulmonary fibrosis.[88] In interstitial lung disease, quantitated CT indexes correlated well with lung function tests.[90]

1.3 Research Justification

1.3.1 Research question

Can the functional and anatomical information gained by 18F-FDG PET/CT scans on patients with PTB before, during and after treatment be applied to provide insight into the dynamics of MTB versus host interaction, aid clinical treatment response monitoring and facilitate the discovery of biomarkers to monitor treatment response?

1.3.2 Study objectives

- Recruitment and follow of patients newly diagnosed with sputum positive pulmonary tuberculosis.
- Perform 18F-FDG PET/CT scans at diagnosis, after 1 month and after 6 months of standard anti-tuberculosis treatment and collect clinical specimens for biomarker discovery.
- Define clinical outcome categories based on established tests, namely Chest X-rays, sputum culture microbiology and symptom monitoring.
- Qualitative description of the scans.
- Development of technique to quantify scan information.
- Use developed technique to quantify scans.
- Comparison and correlation of qualitative and quantitative PET/CT data to defined clinical outcomes to identify positive and negative prognostic scan characteristics.
- Comparison and correlation of biomarkers in clinical samples to most accurate PET/CT characteristics to aid discovery, understanding and ultimate implementation of these markers.

2. Methods and procedures

2.1 Patient recruitment and follow-up

Ethical approval was obtained for all study related activities from the Stellenbosch University Human Research Ethics Committee (registration number N10/01/013), as well as City Of Cape Town Health department and Tygerberg Hospital management. In total, 131 HIV-uninfected adults with newly diagnosed pulmonary tuberculosis (PTB), confirmed by sputum culture, were recruited after informed consent between April 2010 and April 2013 at Primary Healthcare Clinics (PHC's) in Cape Town. Of these patients, 32 were excluded for reasons including: missed scan visits due to technical reasons (7), patient decision (8), sputum culture negativity at diagnosis (8) and severe co-morbidity (3). We did not consider patients who received TB treatment within the previous 12 months.

Sputum specimens for MGIT culture with speciation and additional clinical information were collected at Day 0, week 1, 4, 8, 12, and 24 (month 6) for all participants and an additional culture, collected at week 20 for the last 64 patients who completed the study. Of the 99 completing study time-points, 85 patients were also available for a formal follow-up study visit 1 year after the end of treatment, while the other cured patients were screened by telephonic interview and review of medical records and results. The screen was repeated for all participants, 2 years after the end of treatment. Patients' clinical outcomes were classified as cured if they proved and maintained sputum culture negativity by month 6 (M6), failed treatment if they're M6 culture was still positive, un-evaluable if contamination caused uncertainty in outcome, and recurrent if they were restarted on TB treatment by healthcare providers in the 2 years after initial treatment completion. Sputum and scans were obtained 6 months after initiation of treatment, while BAL fluid was always collected after the end of treatment. Bronchoscopies were performed on PTB patient who did not yet pass this time point by the time ethical approval was granted for this procedure. Adequate volume BAL fluid for RNA extraction was obtained for 15 of these patients.

Sixty-six community controls and 13 controls diagnosed with infective or inflammatory lung diseases other than PTB (other lung disease; OLD) were recruited from the same community. Out of this group, 20 community and 5 OLD consecutively enrolled controls were included in sputum RNA analysis. OLD controls included four cases of pneumonia and one asthma exacerbation case. Active tuberculosis was excluded for all controls based on negative sputum culture and a chest X-ray not suggestive of PTB. All controls were HIV-negative and not on any systemic corticosteroid treatment. Figure 1 shows a diagram of recruitment, follow-up and procedures.

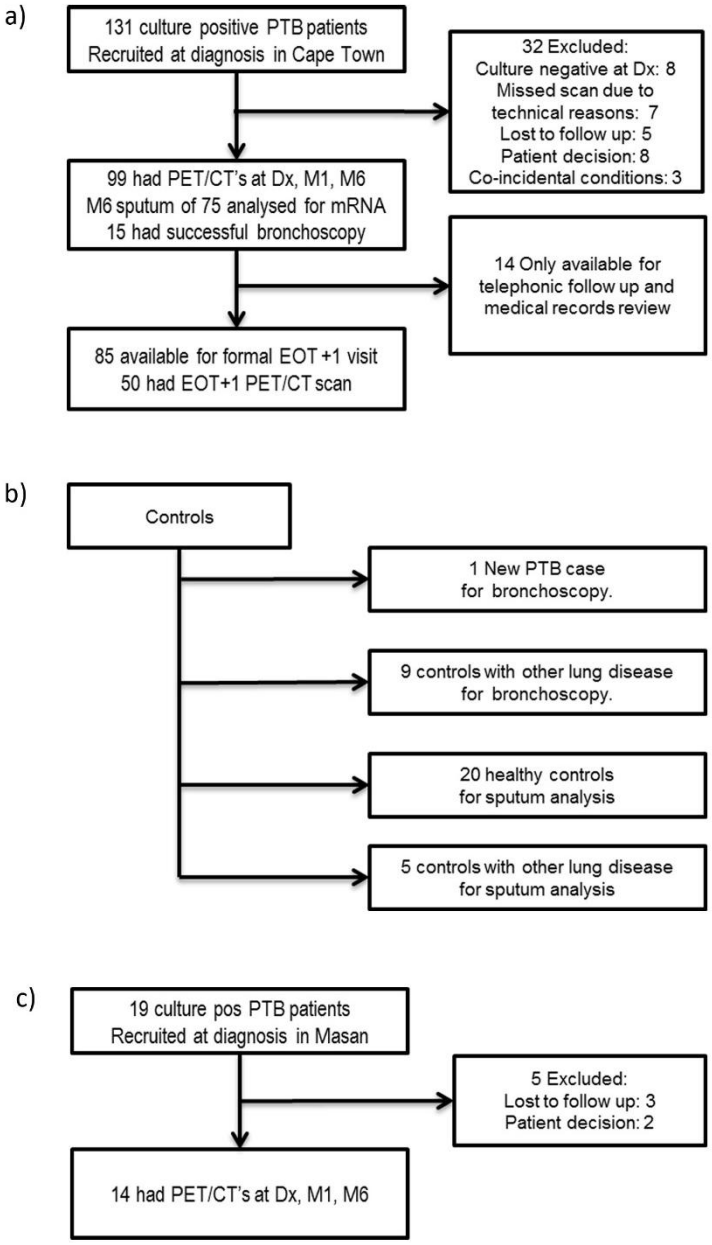


Figure 1: Flow chart of study design.

(a) In, South-Africa, 99 sputum culture positive Pulmonary Tuberculosis (PTB) patients underwent 18F-FDG PET-CT scans at diagnosis (Dx), month 1 (M1) and month 6 (M6) of treatment. Fifty patients also had PET-CT scans 1 year after the end of treatment (EOT + 1y). QRT-PCR assays for MTB mRNA were performed on month 6 sputum from 75 patients and end of treatment bronchoalveolar lavage (BAL) samples of 15 patients. (b) QRT-PCR assays were repeated on sputum from 20 community controls, and 5 controls with lung disease other than PTB; as well as BAL samples of 10 controls undergoing diagnostic bronchoscopies for suspected lung cancer. (c) PET-CT scans were also performed on 14 sputum culture positive PTB patients from South Korea at Dx, M1 and M6.

2.2 Treatment regimen

All patients were subjected to directly observed treatment (DOT) at local healthcare clinics.

Seventy-two patients received a standard treatment regimen, consisting of a two months intensive phase with daily fixed-dose combination tablets (HRZE) consisting of isoniazid (INH), rifampicin (RIF), ethambutol (EMB) and pyrazinamide (PZA) followed by a 4 month continuation phase of daily INH and RIF (currently the standard of care according to the WHO guidelines).[14] Twenty-three patients received an extended regimen: 3 months initiation phase of HRZE, with additional streptomycin in the first month (19) or without (4), and a 5 months continuation phase. In 22 patients, the indication for the extended regimen was previous episode(s) of PTB and for the remaining patient, delayed smear conversion. The local guidelines were amended during the study period to standard 6 months of HRZE treatment for all drug-susceptible cases, resulting in seven patients with previous episode(s) not receiving extended treatment.

Two patients had INH mono-resistance and received 6 months of HRZE with added ofloxacin and 12 months of HRZE respectively. Two patients with MDR received individualised regimens, containing kanamycin and terizidone among others.

Clinical outcome after 6 months of treatment was classified according to time to maintained sputum culture negativity (TTN). Patients were classified as: cured if they proved and maintained sputum culture negativity by M6; failed treatment if their M6 culture was still positive; un-evaluable if contamination caused uncertainty in outcome; and recurrent if they were restarted on TB treatment in the 2 years after initial treatment completion. None of the un-evaluable group was diagnosed by healthcare providers with recurrent PTB before EOT + 2y.

2.3 Recruitment and follow-up, South-Korea

All participants signed informed consent (Ethical approval from National Medical Centre Institutional Review board: IRB00008343). 19 culture positive PTB patients were recruited at diagnosis in Masan, South Korea. Three patients were lost to follow-up and 2 decided to leave the study. The 14 patients that completed Dx, M1 and M6 scans included 3 with MDR strains and 2 with Rifampicin mono-resistance. Patients with DS-PTB were treated with standard HRZE for 6 months, Rifampicin resistant cases were treated for 12 months and MDR cases were treated for 18 months after culture conversion, on individualised regimens.

Patients were followed up for 2 years after treatment.

2.4 18F-FDG PET/CT scanner protocol

A one-week window was allowed for the Dx and M1 scans, while a 4-week window was allowed for M6 and EOT + 1y year scans. Scans were performed at Panorama Mediclinic or Tygerberg Hospital, using a Siemens Biograph or Philips Gemini scanner respectively, both of which were equipped with a 16 slice CT scanner. Patients fasted for at least 6 hours before FDG administration, while being encouraged to drink plenty of water. According to body weight 185-259MBq of ^{18}F -FDG was administered intravenously 60 min before scan. Scans at baseline were performed from the base of skull to the upper thigh and from the neck to the upper abdomen (whole lung) at M1, M6 and EOT + 1y. PET images were reconstructed to 4x4x4 mm voxels using an iterative algorithm. The CT scan parameters were set at 120kV, 100mAs without dose modulation with 1.17x1.17 mm pixels and a 3 mm slice thickness.

2.5 Qualitative analysis of PET/CT images

CT and PET images were evaluated in collaboration and consensus with a radiologist, a nuclear physician and a pulmonologist. Patient clinical outcomes were not known at the time of reading. Lesion morphology was characterised and measured based on CT images. PET uptake intensity of prominent lesions was measured using the mean and maximum standardised uptake value (SUV) and compared across time-points, using the uptake in the right lobe of the liver and the mediastinal aortic blood pool, as reference values to account for variability in background uptake.[50,54–56,77–79,81]

The most intense lesion of each scan was ranked based on its mean standard uptake value (SUVmean). The method was adapted from the Deauville grading system that used the mediastinal blood pool and liver as references.[81–83] a) None: No increased uptake compared to surrounding lung tissue. b) Minimal: Uptake more intense than surrounding lung tissue, but less intense than mean (SUV) of 10 mm sphere in mediastinal blood pool. c) Mild: Uptake more intense than mediastinal blood pool but less than 1 standard deviation (SD) below the mean SUV of a 30 mm sphere in right liver lobe. d) Moderate: Uptake similar to liver (between 1 SD below mean SUV and 2 SD above mean SUV). e) High: Uptake more intense than 2 SD above mean liver SUV and less than double the mean liver SUV. f) Very high: Intensity higher than double the mean liver SUV.

Scans were then grouped into response patterns by comparing each M6 lesion to the corresponding M1 lesion to summarise the central trend of lesion progression: 1) Resolved, 2) Improved and 3) Mixed. Resolved response pattern refers to minimal or no increased FDG uptake compared to surrounding healthy tissue, regardless of structural abnormalities

on CT. Improved scans had decreased intensity of all lesions compared to the baseline scan, but still one or more lesion(s) with increased uptake compared to surrounding lung tissue and reference structures. Mixed responses showed at least one new FDG-avid lesion or at least one lesion with increased FDG uptake (intensified) compared to the baseline scan.

2.6 Quantitative analysis of PET/CT images

Pre-processing

Images were exported from the PET/CT workstation in DICOM format and converted to ANALYZE format using MRIConvert.[91] To ensure reproducibility and facilitate direct comparison, each patient's baseline and follow-up scans were co-registered, using Statistical Parametric Mapping (SPM 8)[92] with Matlab 2013b (Mathworks Inc.). The baseline CT scan was used as the reference volume, and spatial transformations were applied to the follow-up CT and PET scans. CT images were re-sliced to the coarser voxel matrix of the corresponding PET scan. This allowed direct voxel based comparison of each patient's PET and CT images.

Most of the source code for the Matlab script was created Patrick Dupont and Ilse Kant and adjustments were made by James Warwick. I made further adaptations to the script and optimised the application procedure.

Lung masks

For each series of co-registered CT studies, a volume of interest (VOI) was generated as a lung mask (Figure 2a), using MRICro version 1.39[93]. The lung mask was created on the overlaid CT images, using MRICro's 3-dimensional region-growing, gradient-based tool that allows the user to set the origin, radius, and differences from origin and edge gradient. This was combined with manual correction, since dense lesions extending into the pleura were often excluded from the VOI. The lung mask excluded lung hila and main pulmonary vessels, but included smaller vessels. A VOI of organs around the lungs leading that affected motion misregistration (usually liver, spleen and mediastinum), was created with the region-grow tool on the three overlaid PET scans and deleted from the lung mask (Figure 2b). The lung mask was then transformed to a binary image using ImageJ[94] (Figure 2c).

Background reference

Two areas of normal lung parenchyma (NL) were selected as a reference to standardise uptake in each remaining lung voxels (Figure 2d). Each NL volume consisted of spheres of 15 mm to 25 mm in diameter that visually appeared lesion-free all on co-registered scans.

To enhance the representivity of the sample, the spheres were selected in opposite lungs or in some cases different ipsilateral lobes, depending on the distribution and extent of lung pathology. The spheres were selected to exclude uptake attributed to structures surrounding the lung. Images were then transformed to binary format using ImageJ.[94]

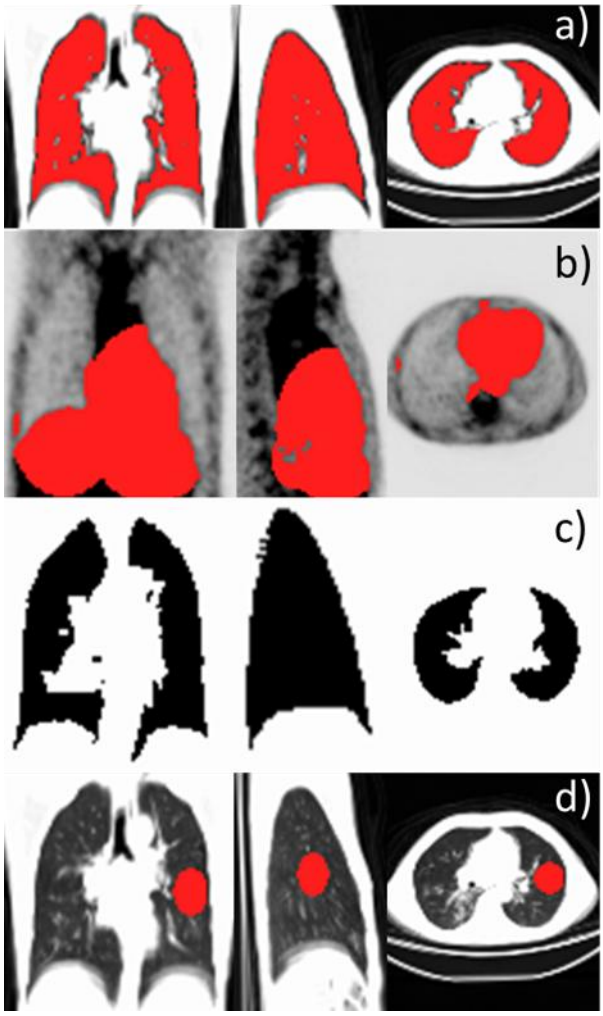


Figure 2 Example of Volumes of interest (in red) required for quantification.

Image Quantification

The pre-processed series of PET/CT images were quantified using a MATLAB script incorporating SPM source code, developed in-house. The script quantified disease burden from the PET images. Using the mean and the standard deviation of PET counts within the NL volume for each study, all voxel counts within the lung mask were assigned a Z-score based on the formula below:

$$Z = \frac{counts - \mu_{NL}}{\sigma_{NL}}$$

, in which μ_{NL} and σ_{NL} are the mean and standard deviation of PET counts within the normal volume for each study.

The Z-score provided a statistical way to standardise relative intensity of FDG uptake throughout the lungs. All lung voxels exceeding this Z-score threshold were defined as part of FDG-avid lesions. Images of segmented lesion volumes were then exported to view alongside the original images for visual quality assessment.

To determine the optimal Z-score threshold, we first processed a test set of 15 scans from control patients, using ascending Z-score thresholds, to minimise false positive segmentation. These scans were obtained from patients undergoing PET/CT scans for non-pulmonary clinical indications, but with lungs that were visually lesion-free on PET and CT. As expected, at a low Z-score threshold, the volumes segmented as abnormal in these healthy lungs were high, but decreased to a value close to zero for $Z \geq 8$ (Figure 3). To minimise false negative findings, ascending Z-score thresholds were also tested on scans from five PTB patients that had residual lesions with minimal or mild intensity and complex morphology. Minimal intensity was defined as visually more intense than normal background lung parenchyma, but less intense than the mediastinal blood pool and mild intensity as more intense than mediastinal blood-pool, but less than the right lobe of the liver.[83] A Z-score threshold of 8 delineated all lesions with minimal FDG-avidity, while a Z-score of nine delineated all lesions with mild FDG-avidity, but failed to detect some lesions with minimal FDG avidity (Figure 4). A cut-off of $Z = 8$ was thus chosen subsequently, based on its low false positive and false negative rates.

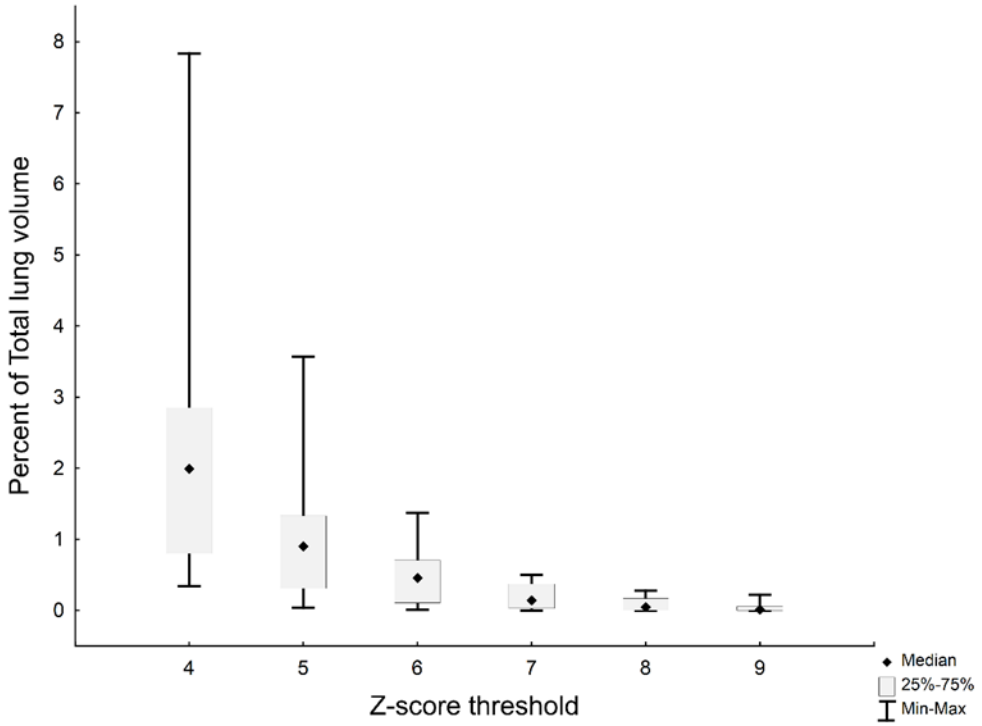


Figure 3 Range, median, 25th and 75th percentiles for percentage of total lung volume classified as FDG-avid at different z-score thresholds.

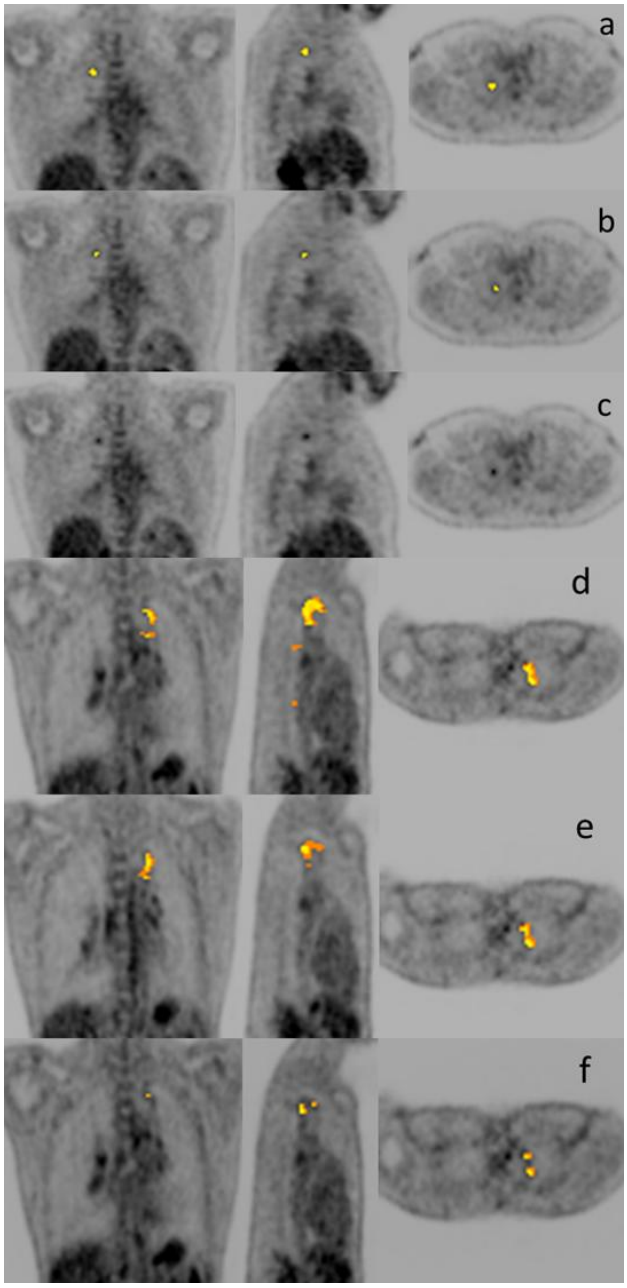


Figure 4 PET images 2 patients after treatment for pulmonary tuberculosis.

Coronal (left), saggital (middle) and transverse views. The first patient (a-c) has a residual nodule in the right upper lobe showing minimally increased FDG-avidity. The second patient (d-f) has a complex lesion in the left upper lobe with mild FDG-avidity. The auto-delineated metabolic lesion volume is overlaid using ascending Z-score thresholds, respectively at 7 (a,d) 8 (b,e) and 9 (c,f).

Concurrently with the PET segmentation and using the same lung mask, the MATLAB script also segmented the CT images into 5 categories based on the density of each voxel compared to set values obtained from literature [75,89]: 1) Low density (<-950HU), attributed to cavitation or extremely hyper-inflated lung tissue (V_{low}). 2) Normal density, between -950HU and -500HU. 3) Soft lesions (V_{soft}), from -500HU to -300HU, usually tree-in-bud lesions or

nodules, but can also including normal medium to large vasculature. 4) Hard lesions (V_{hard}), above -100HU, are usually due to consolidation, cavity walls, bronchial thickening, or calcified fibrosis. 5) Medium lesions (V_{medium}) from -300HU to -100 HU. This category should include very little normal lung tissue and usually consists of nodular infiltrates, but may also include hard lesions in early progression or partial resolution.

Visual checks of the accuracy of CT lesion delineation were performed, based on these fixed density thresholds. For lesions with increased density (V_{soft}, V_{medium}, V_{hard}), the segmented areas correlated well to lesion morphology. However, using -950HU as the upper limit for low-density lesions was not specific enough for cavitation and the segmented areas in some cases included bullae, bronchiectasis and severe emphysema. This necessitated an additional step to measure the volumes of individual cavities for each scan manually. This was done using the MRICRo's 3D region-growing tool.

After segmentation, the program quantified the following parameters for each scan: 1) Total lung volume (TLV) (ml). 2) Metabolic lesion volume (MLV): the total volume of FDG-avid lung lesions, analogous to metabolic tumour volume in oncology patients. 3) The mean intensity in the MLV divided by the mean NL volume (Z_{mean}). 4) Total glycolytic activity index (TGAI): the product of Z_{mean}. 4) Volumes of each abnormal density category on CT, i.e. V_{low}, V_{soft}, V_{medium}, and V_{hard}. 5) Volume with both increased metabolism on PET and abnormal density on CT (MLV_{abN}).

Re-slicing of the CT to the corresponding PET voxels allowed the program to compare CT density and PET intensity per voxel directly. This is facilitated by the generation of CT density vs standardised PET uptake by the script. Figure 5 shows different CT images with the auto-delineated metabolic lesion volume as an overlay and a scatterplot representing the voxels held within the lung mask.

The cured patient's values (Figure 5b) move closer to normal at follow-up, while for the patient who failed treatment (Figure 5c), the pattern remains grossly abnormal on the density and intensity axes.

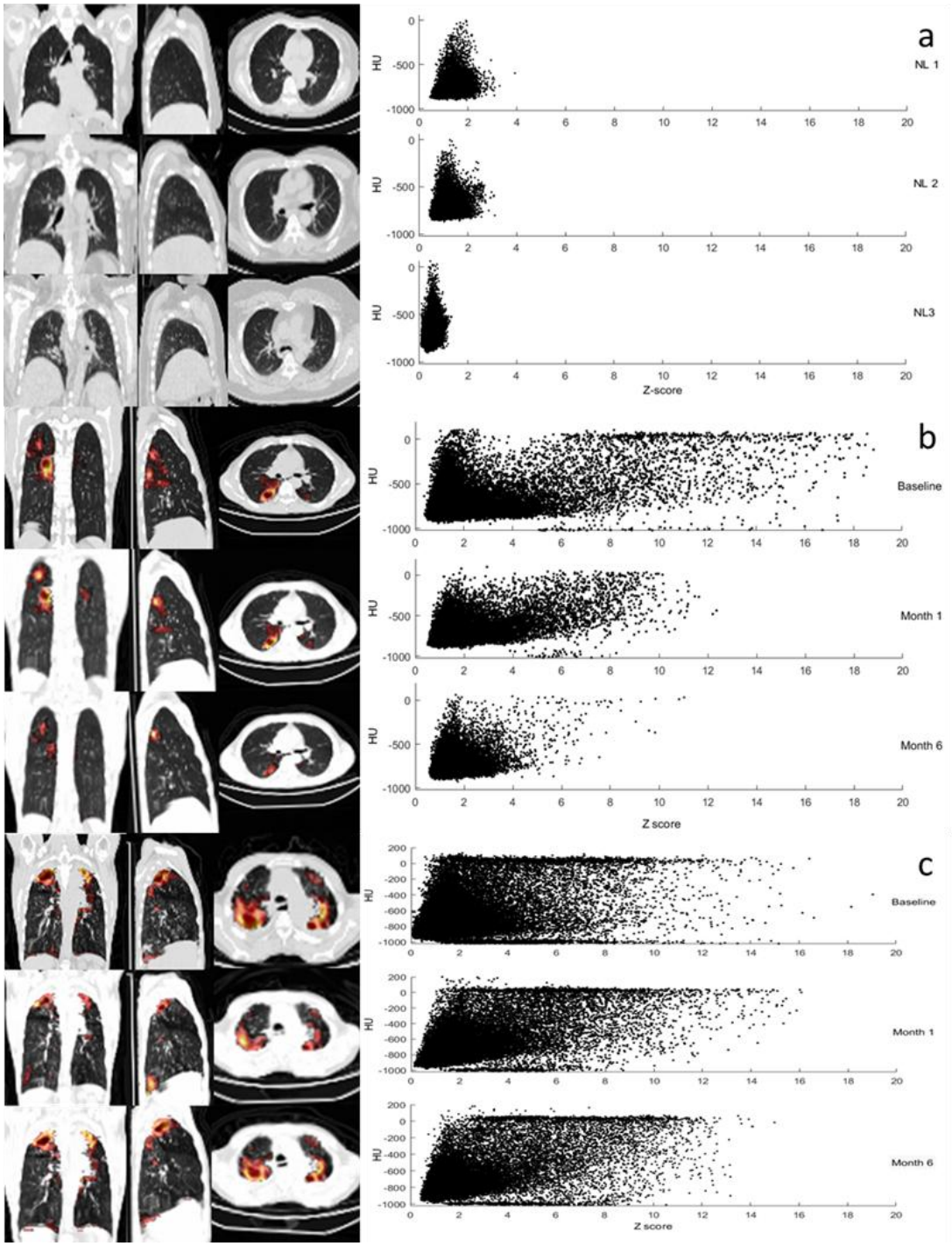


Figure 5 Scatter-plots of PET/CT scans from three healthy controls and two PTB patients during standard treatment.

Coronal (left), sagittal (middle) and transverse (right) views. Y-axis shows CT density (HU), X-axis shows PET uptake (Z-score). a) Lesion-free lungs from three control patients. b) Dx, M1, and M6 for a patient clinically responding well to TB treatment and showing improvement, but not resolution after anti-TB treatment. c) Dx, M1, and M6 scans for a patient not clinically responding well to treatment and showing very little change after treatment.

2.7 MTB RNA extraction and detection in M6 sputum

Sputum samples were sent to David Alland's Laboratory in Rutgers New Jersey Medical School for MTB mRNA extraction and detection, as well as Xpert assays. Shubhada Shenai designed and optimised the primers and conducted the experiments.

For transcription, a QuantiFast Multiplex RT-PCR kit (QIAGEN) with a primer designed to be specific for each MTB target gene was used. We selected genes that were previously shown to have increased expression under conditions of stress and persistence.[95,96] Two ml sputum samples were collected and added to 2 ml of Trizol (Invitrogen), frozen and stored at −80°C. Samples were not processed in a blinded fashion, but clinical outcomes were unknown at the time of analysis. For RNA extraction, samples were thawed and spun down at 3500 g for 25 min. The supernatant was discarded and the pellet was re-suspended in 1 ml Trizol reagent. The entire content of the tube was transferred into new 2 ml screw cap micro-centrifuge tubes containing 0.5 ml of zirconia-silica beads (diameter, 0.1 mm). Bacteria were disrupted in a Bead Beater (Fast Prep Cell Disrupter, FP120) using two 45 second pulses at maximum speed of 6.5 m/sec, incubated on ice for 5 min, followed by the addition of 250 µl of chloroform and vigorous mixing for 15 s and then a 2–3 min room temperature incubation. Samples were then centrifuged at 12, 000 g for 5 min and the upper clear phase transferred carefully to a new tube and mixed with an equal volume of 70% ethanol, applied to an RNeasy mini column (QIAGEN) and processed according to the manufacturer's recommendations.

Fourteen mRNA targets (*85B, dosR, hspX (acr), pstS1, carD, nuoB, tgS1, TB8.4, TB31.7, acpM, icL, prcA sigA, rrnAP1*) were reverse transcribed, using a QuantiFast Multiplex RT-PCR kit (QIAGEN) and a primer specific for each target gene. The conditions for Reverse Transcription PCR were 50°C for 50 min, followed by 95 °C for 2 min. The amount of cDNA produced was then quantified by real-time PCR with the corresponding molecular beacon. The 10 µl PCR reaction mixture consisted of 1X PCR buffer, 250 µM deoxynucleoside triphosphates, 4 mM $MgCl_2$, 0.5 µM each primer, 4ng/µl molecular beacon and 0.03U/µl Jumpstart Taq polymerase (Sigma-Aldrich). To normalise the individual reactions 6-carboxy-x-rhodamin (ROX) was always included as passive reference dye. PCRs were performed in 384-well microtiter plates in an ABI 7900 prism (Applied Biosystems, Foster City, CA) according to the following parameters: initial denaturation at 95°C for 1 min, followed by 50 cycles of: denaturation at 95°C for 30 seconds, annealing at 58°C for 30 seconds, and extension at 72°C for 15 seconds. PCR conditions were identical for all assays. The fluorescence was recorded during the annealing step of the assay. The quantity of specific target DNA was determined from the cycle threshold (CT) value with

reference to a standard curve of genomic DNA. The copy numbers of target standards used ranged from 1 to 10^6 genomic copies per reaction (i.e. 10 fg to 10 ng DNA from CDC 1551 strains). The lower limit of detection for each of the assays was 10 fg, which is equivalent to 1-5 copies of cDNA. qRT-PCR reactions were performed in triplicate. The mean of each triplicate are used in calculations. The *sigA* target showed cross reactivity with non-tuberculosis mycobacterial RNA at high concentrations, in addition to low-level positivity in four community controls and one control with other lung disease, and was subsequently dropped from further analysis.

Primer sequences are listed in Appendix F.

2.8 RNA in End of treatment bronchoalveolar lavage samples

Culture from bronchoalveolar lavage BAL fluid is generally more sensitive than culture from sputum[97,98], although the quality of the sample can be variable. An additional sputum sample is often taken after the procedure, as the irritation of the airways leads to increased sputum expectoration and a high-quality sputum sample.[99] Bronchoscopies were performed on 15 of the PTB patients within 3 months after the end of their standard treatment (EOT) and BAL fluid and post bronchoscopy sputa obtained. MGIT cultures on bronchoscopy samples were allowed extended growth periods (12 weeks). BAL fluid was also collected from 10 patients suspected of having cancer and undergoing diagnostics bronchoscopies, as controls. We added 15 ml of fresh BAL samples to 30 ml Trizol and stored it at -80°C until RNA analysis.

Multiplex Real Time RT-PCR Transcriptional analysis of Sputum and BAL specimens

Fluorescence-based qRT-PCR is accepted as a gold standard for gene expression profiling due to its high specificity, sensitivity, accuracy and large dynamic range.

Extraction of total RNA from BAL

BAL samples were sent, in a blinded fashion, for MTB RNA analysis via two-step qRT-PCR with validated TaqMan assays developed at Stanford.[100,101] The tests were conducted by Gregory Dolganov and Tran Van in Gary Schoolnik's laboratory at the Department of Microbiology and Immunology, Stanford University School of Medicine.

Total MTB RNA isolated from BAL was reverse transcribed using random primers and pre-amplified in a controlled multiplex PCR with 743 MTB-specific primer sets targeting corresponding mRNA transcripts as described previously.[102] Finally, we used only 24 gene-specific TaqMan assays for individual gene transcript quantification of multiplex PCR products targeting the genes shown as differentially regulated in stress or persistence based

on available literature. [95,96] These genes included *recF, eccD3, menA, pabC, accD3, iprB, Rv1421, gabD2, Rv1910c, hspX, lipX, fadD9, tgs1, fadE34, Rv3675, sodA, trxB2, rpsJ, rplV, sigI, Rv1255c, sigB, sigH, rpsK*.[95]

Samples were thawed and total RNA was extracted using a glass matrix tube for cell lysis (Lysing Matrix B, Q Biogene) in a FastPrep FP120 instrument (Bio 101, Thermo Savant) with a speed setting of six for three iterations of 30 s, with cooling on ice for 1 min after each iteration. After processing, chloroform was added, followed by separation of the aqueous and organic layers. The aqueous phase containing the RNA was removed to a fresh tube and the RNA was precipitated overnight at -20°C after addition of glycoblue (Ambion), 0.1 volume of 5 M ammonium acetate, and an equal volume of isopropanol were added. The resulting RNA pellet was re-suspended in 50 µl of RNase free H2O and cleaned by three rounds of the RNeasy® mini column system, interspersed with off-column DNase treatments (Promega RQ1 DNase).[95]

Genome Expression Profiling of MTB was performed using Two-Step Multiplex RT-PCR. The protocol below consists of three parts: 1) First strand cDNA synthesis and Controlled Multiplex Pre-Amplification of cDNAs; 2) Preparation of Primer and Probe Sets; and 3) Individual qRT-PCR (Taqman) quantification of Amplified cDNAs using LightCycler480.

Each RNA sample (5 µl) was taken into two separate first strand cDNA synthesis reactions (RT+ and RT-) to control for DNA contamination. An additional water control (zero RNA) was also added. To each sample, 0.5 µl Exo-resistant Random Primer (Fermentas S0181), 1 µl 10mM dNTPs (Fermentas R0193) and 3.5 µl Nuclease Free Water (Ambion AM9938) were added for a total of 10 µl. This mix was incubated for 3 minutes at 70 °C in a thermal cycler, and then placed on ice.

During this incubation, two cocktails were prepared: one containing reverse transcriptase (RT+) and one without (RT-). Each RT+ cocktail contained 4 µl 5X Maxima RT Buffer (Fermentas EP0741), 0.5 µl Ribolock RNase-Inhibitor (Fermentas EO0382), 0.5 µl Maxima RT enzyme (Fermentas EP0741), and 3.0 µl Nuclease Free Water for a total of 10 µl. The RT- cocktail was the same, except water was substituted for the RT enzyme. These cocktails were scaled up for multiple samples and 10 µl aliquots were added to each RT+ or RT- sample prepared above. The samples were mixed gently and then gently centrifuged for 2 min. They were incubated at 50 °C 1 hour, 95 °C 2 minutes (to inactivate the reverse transcriptase), then stored at 4 °C. Maxima Reverse Transcriptase (RT) possesses an RNA and DNA-dependent polymerase activity as well as RNase H activity.

qRT-PCR on pre-amplified material: All outflanking primers and TaqMan probe sets had been validated in multiplex PCR pre-amplification for linearity of amplification using all the genes used in each pre-amplification cocktail. We also validated all individual TaqMan assays from our collection for sensitivity and linearity before we started using them in gene expression profiling. A complete database with all available validated TaqMan primer sets used for pre-amplification available at ftp://smd-ftp.stanford.edu/tbdb/rtpcr/taqman_oligos.fa. Sequences and design of specific PCR primer/probe sets shown in Appendix F.

Primer and probe qRT-PCR sets for each gene consist of a forward primer (TMF), a FAM/BHQ-labelled probe (TMP), and a reverse primer (TMR). These were ordered from Biosearch at a 100 uM concentration and a Taqman Mix was prepared for each gene. Twenty-seven µl of the forward primer, 27 µl of the reverse primer, and 9 µl of the probe were mixed in 1737 µl of Nuclease Free water for a total of 1800 µl. This resulted in a dilution of the forward and reverse primer to 1500 nM and the probe to 500 nM. We used 2 µl of this mix in a 10 µl qRT-PCR, resulting in a final reaction concentration of 300 nM for each primer and 100 nM for the probe.

Two different water controls were run throughout the process: one starting at the cDNA step and one starting at the amplification step. All RT- controls were negative, but water controls for pabC, accD3, lipX, Rv1255c and sigB showed some reactivity and these targets were excluded from further analysis. Additionally, two samples of 104 gene copy number H37Rv genomic DNA were amplified in each amplification mix and run on each plate.

Genomic Equivalent DNA calculation

We isolated genomic DNA by buffer extraction from the interphase after MTB lysis with Trizol. DNA was pre- amplified with 24 genes and individual TaqMan assays quantified the amplicons. Median Ct was calculated for 24 genes for each sample and compared to genomic DNA reference to calculate GE.[100]

Statistical analysis

As this was an observational study, we did not apply randomisation. The calculated sample size of the cohort was intended to provide a representative range of biological reactions during a favourable response to PTB treatment. We performed bronchoscopies and EOT +1y PET/CT scans on all consenting patients who have not passed the time point when ethical approval was obtained for this amendment to the protocol. Analysis was either performed blinded (BAL) or outcomes were not known at the time (M6 sputum mRNA and Xpert; M6 PET/CT scans).

For analysis in chapter 3: 1) Clinical and microbiological parameters of failed treatment and recurrent outcome groups were compared to the cured group using a two-tailed, unpaired Student's t-test – that I performed with Statistica[tm] Version 12. 2) Associations between categorical variables were evaluated using the Fisher Exact Test for independence, the Chi-squared Test for trend in proportions and the Asymptotic independence test in R version 3.2.2 by Prof Gerard Tromp. 3) Principal component analysis was performed by Prof Tromp, using the singular value decomposition approach (prcomp in R version 3.2.2), to derive a limited set of variables (principle components; PC) that captured the variance in the mRNA values. Matrices of Ct values for each mRNA species in each patient were used as input for the analysis. The first PC captured more than 80% of the mRNA variance in the analyses and was used to represent the mRNA data for each individual with a single value. Figure 12 and Figure 13 contains box and whisker plots of the arcsinh-transformed first principal component of mRNA transcript PCR data.

For analysis in chapter 5: 1) Tom Peppard and Yookwan Noh performed statistical analysis and created graphical representations regarding correlations between scan parameters and time to culture negativity groups. The PET and CT values were log transformed and subjected to analysis of variance. 2) Tom and Yookwan also created receiver operating curves for a failed treatment outcome and the associated Mann-Whitney U-test P-values, the distance to unity and the level of informedness. 3) Correlation matrices were created by Tom Peppard. 4) Optimal thresholds for prognostic criteria of pooled unfavourable outcome were set after I performed analysis of distribution (Statistica), receiver operating curves (Prism Graphpad[tm] software) and also performed some rounding of values. 5) I entered the results of the criteria of contingency tables and calculated sensitivity, specificity, negative predictive value, positive predictive value and relative risk using Prism Graphpad software. 6) I calculated associations and between PET/CT variables and clinical and microbiological factors and prepared graphical representations in Statistica[tm] version 13. Test performed include Fisher Exact test for categorical variables, correlation coefficients for continuous variables and two-tailed, independent student's T-tests for grouped continuous variables. We also report the results for Levene's test for equal variance. Since most of the variables showed a mild positive skew, I also performed Mann-Whitney U-test for independent samples. As a rule, there was good cohesion between the parametric and non-parametric testing.

3. Results: Persistent lesions and MTB mRNA detected after treatment

3.1 Patient demographics and treatment outcome

Ninety-nine HIV-uninfected, non-diabetic, adult patients with PTB were recruited at diagnosis and followed up during treatment in Cape Town, South Africa. This included 39 patients with one or more previous episode of PTB, more than one year prior to recruitment. Ninety-five patients had DS strains. Of these, 72 received a standard, 6-month regimen, while 23 received an extended regimen of 8 months [22 due to previous episode(s) of TB and one due to delayed smear conversion]. Two patients had INH mono resistant TB and were treated for six and 12 months respectively and two patients with multi-drug resistant (MDR) TB were treated for 2 years on individualised regimens. Month 6 (M6) was selected as a fixed time-point for PET/CT and sputum collection, consistent with current WHO treatment guidelines.[14]

Of the 99 patients, 8 remained sputum culture positive at M6 and failed treatment, 76 culture converted and maintained cure and 12 initially culture converted, but were diagnosed with recurrent (due to relapse or reinfection) PTB within 2 years after treatment completion (EOT + 2y). Initial treatment outcome was un-evaluable (UE) for three patients due to contaminated cultures. All extended treatment patients maintained an unchanged sputum culture status between M6 and EOT. Fourteen patients took fewer than approximately 80% of their treatment dosages during the 6-month period, which is regarded as poor adherence in most clinical trial designs.[4,6] The failed treatment group included 4 patients with poor treatment adherence and 1 with MDR disease. Most patients (96%) reported symptomatic improvement during treatment, but residual symptoms or signs were often encountered at the end of treatment. The most frequent were coughing (47), dyspnoea (21) and low body-mass index (24).

Table 1 provides a summary of patient demographics, treatment adherence, as well as culture and GeneXpert MTB/Rif (Xpert) results.

Table 1: Clinical and microbiological parameters for PTB patients

	All (n=99)		Cured (76)		Failed (8)		Recurrence (12)	
Age	31	(17 -66)	32	(17-64)	29	(18-66)	34	(21-52)
Male (n)	60	(60%)	49	(82%)	5	(63%)	8	(67%)
Previous PTB	39	(39%)	20	(27%)	4	(50%)	4	(33%)
Poor adherence	14	(10%)	9	(12%)	4*	(50%)	1	(8%)

MDR	2	(2%)	1	(1%)	1	(12.5%)	0	(0%)
M6 Xpert positive (n)	30	(30%)	16	(22%)	7	(88%)	6	(50%)
Smoker (n)	77	(78%)	56	(74%)	6	(75%)	12	(100%)
Dx Body Mass Index	18	(13.1 -42)	18	(13.1-23)	17	(16.7-23)	19*	(16.3-42)
M6 Body Mass Index	18	(14.5 - 43)	19	(14.5-25)	19	(17.2-26)	19*	(17.5-43)
Time to Negativity (weeks)	8	(0.2-24)	8	(1-24)	>24*		8	(0.3-12)

Summary of median (range) or number (%) of clinical and microbiological parameters for PTB patients. The asterisk (*) in the Failed and Recurrence columns indicate a significant difference from cured values (two-tailed, non-paired, Student's T-test; P-value < 0.05). Poor adherence refers to patients that missed more than 20% of their treatment (note: only three of these patients met WHO criteria of treatment default).

Of the recurrent PTB patients, two were culture confirmed; five were confirmed by both Xpert with supporting acid fast bacillus (AFB) testing for smear positivity by direct microscopy; three were Xpert negative at month 6, but converted back to positive; and two remained Xpert positive for more than 6 months, but complained of increased symptoms. None of the un-evaluable group was diagnosed by healthcare providers with recurrent PTB before EOT + 2y.

Table 2 shows a summary of the time from end of treatment (EOT) to re-diagnosis and the supporting results - the asterix indicates results at the time of re-diagnosis.

Table 2 Summary of evidence supporting diagnosis of PTB recurrence.

ID	M6 Study culture	M6 Study Xpert	Healthcare MGIT after EOT?	Healthcare Xpert after EOT	Healthcare smear after EOT	EOT + 1y Study culture	EOT + 1y Xpert	Months to recurrence	Summary
S22	Neg	Neg	Positive	n/a	n/a	Positive*	Neg	15	Late, Definite
S34	Neg	Pos	n/a	Pos*	Neg	Contaminated	Neg	9	Early, Moderate evidence
S88	Neg	Neg	n/a	Pos*	Neg	Negative*	Pos*	12	Early, Strong evidence
S93	Neg	Neg	n/a	Pos*	N/a	Negative	Neg	18	Late, Strong evidence
S95	Neg	Pos	n/a	Pos*	Neg	Negative	Pos	10	Early, Moderate evidence
S101	Neg	Pos	Contaminated	Pos*	Pos*	Contaminated	Neg	17	Late, Definite
S112	Neg	Neg	n/a	Pos*	Pos*	Contaminated	Pos	3	Early, Definite
S130	Neg	Neg	n/a	Pos*	Neg	Negative	Neg	17	Late, Strong evidence
S137	Neg	Pos	n/a	Pos*	Pos*	Negative	Pos	5	Early, Definite
S140	Neg	Pos	n/a	Pos*	Pos*	Negative	Neg	5	Early, Definite
S142	Neg	Pos	n/a	Pos*	Pos*	Negative	Pos	17	Late, Definite
S152	Neg	Neg	Positive	Pos*	Neg	Positive	Pos	3	Early, Definite

Spoligotyping was performed on sputum cultures at diagnosis. Spoligotyping was unsuccessful in 10 patients. The Beijing strain clade was, by far the most prevalent strain and found in the culture of 51 (57%) of PTB patients (Figure 6). This was followed by LAM and LCC strains.

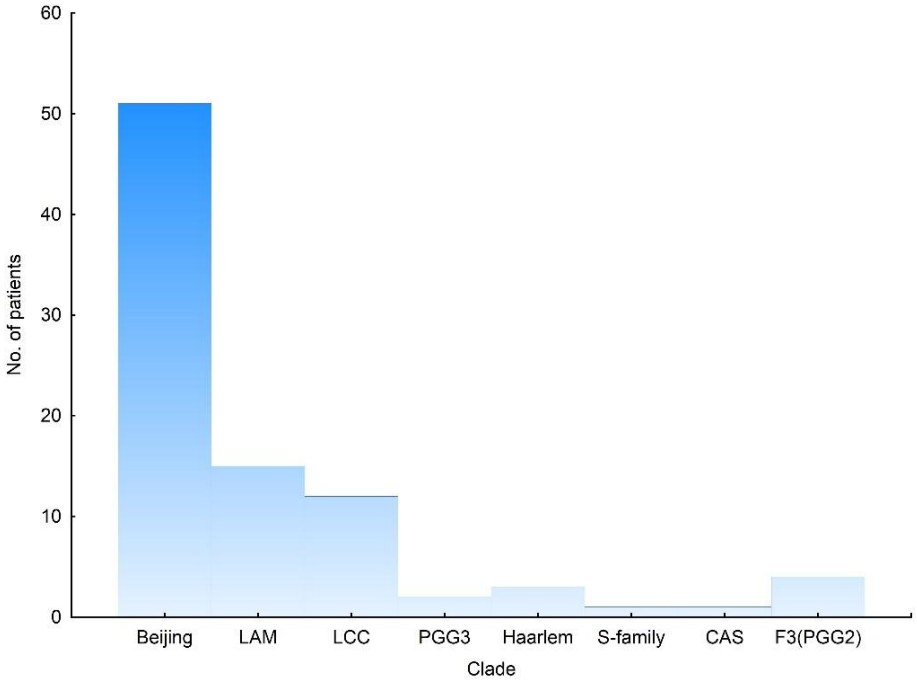


Figure 6 Frequency of strain clades cultured from sputum at Dx for South-African cohort.

Spoligotyping was also performed at M6 for patients that failed treatment. In 6 of the 8 cases identical strains were found, while the remaining 2 patients had different strains at Dx

and M6. This may be due to reinfection during the treatment phase, or concurrent infection with different strains from Dx – previously reported in 19% of patients in Cape Town.[103]

We also performed qualitative analysis of PET/CT images from 14 HIV-negative patients with PTB in Masan, South Korea. They included three MDR cases and two with mono-resistance to rifampicin. Twelve were cured (culture converted within 6 months), one had un-evaluable outcome and one was diagnosed with culture confirmed recurrence after initial cure. In this case a strain was cultured on sputum different to the strain at diagnosis, suggesting reinfection. Study design and research procedures are detailed in Figure 1c.

3.2 Month 6 PET/CT findings in 99 South African patients

All patients had ^{18}F-FDG PET/CT lung scans at Dx, month 1 (M1), and M6 of TB treatment.

Comparing M6 to Dx scans we identified three distinct response patterns: 1) Resolved, 2) Improved and 3) Mixed. A resolved response pattern refers to minimal or no increased FDG uptake compared to surrounding healthy tissue, regardless of structural abnormalities on CT. Improved scans had decreased intensity of all lesions compared to the baseline scan, but still one or more lesion(s) with increased uptake compared to background and reference structures. Mixed responses showed at least one new FDG-avid lesion or at least one lesion with increased FDG uptake (intensified) compared to the baseline scan. Representative cases are shown in Figure 7.

Surprisingly, only 14 (14%) patients had a resolved pattern on their M6 scan (Figure 7a) and in 51 (52%) we found an improved response on M6 scan (Figure 7b). A mixed response was seen in 34 (34%) patients (Figure 7c), of which 14 had both new and intensified lesion(s), 16 had only intensified lesion(s) and 4 had only new FDG-avid lesion(s).

The morphology associated with the most intense lesion of each mixed and improved M6 scan included features suggestive of active PTB, such as cavities (found in 26 cases), patchy consolidation (in 22), complex lesions involving consolidation with cavitation (16), nodular infiltrates (17), enlarged hilar lymph nodes (3) and pleural based infiltrates (1). Treatment outcome was associated with scan response pattern, (Figure 9a; $p < 0.01$) with a mixed response found in all failed patients. Neither a mixed response nor a high maximum lesion intensity, however, was specific to a poor outcome and 21 (28%) of cured patients had a mixed response, while 55 (72%) still had M6 lesions with moderate to very high intensity, which would be compatible with the intensity range found at Dx.

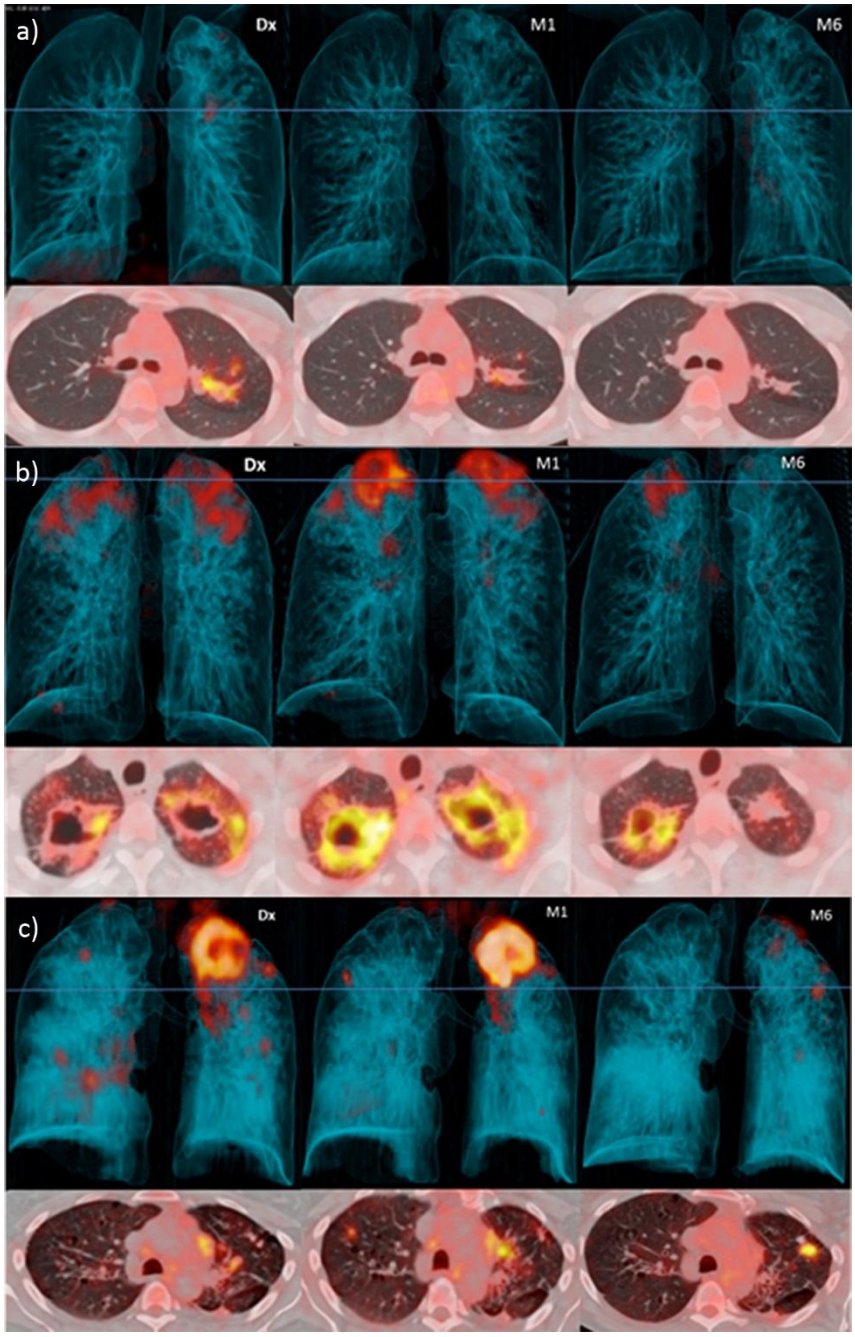


Figure 7 PET/CT images at Dx, M1 and M6 for representative cured patients after 6 months of standard treatment.

3-dimensional anterior (top panels) and transverse views, at the level of blue lines, (lower panels). (a) Resolved scan, without residual abnormal FDG uptake, some structural abnormalities remain on CT scan. (b) Improved scan, where all lesions have improved since diagnosis, but abnormal CT lesions and increased uptake persist. This example includes thick-walled cavities with high FDG uptake. (c) Mixed scan response with either new FDG-avid lesions or increased intensity of some lesions. This example shows a large new FDG-avid nodule with high intensity uptake in the left lung upper lobe.

3.3 PET/CT findings in 14 South Korean patients

All 14 patients had PET/CT scans at Dx, M1 and M6. We applied the same criteria as above and found 7 mixed (50%), 6 improved and only 1 resolved M6 scans. Eleven had moderate to very high intensity lesions at M6. Figure 9b shows a summary of scan findings; detailed demographical data in Appendix C and representative images in Figure 8.

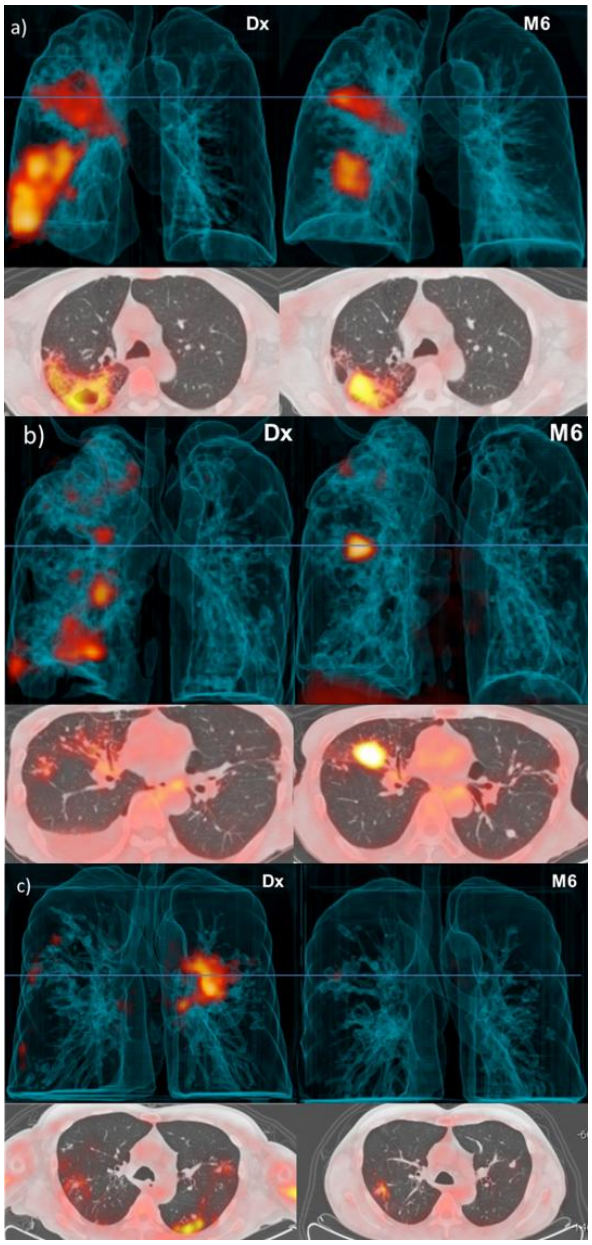
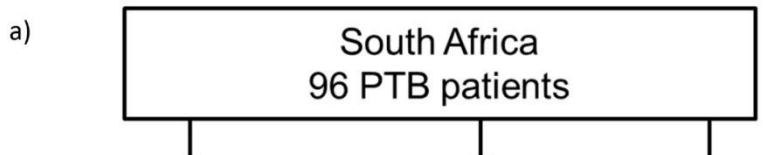


Figure 8 PET/CT images for representative cases at Dx and M6 from South-Korean cohort.

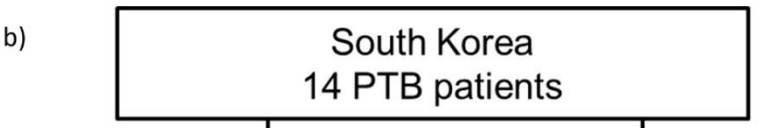
3-dimensional anterior (top panels) and transverse views, at the level of blue lines (lower panels), of Cases included in this group are all diagnosed with drug-sensitive TB strains and cured (culture negative) after 6 months of standard treatment a) Mixed response pattern with old cavity, becoming consolidated and more intense. b) Mixed response pattern with new patch of consolidation. c) Improved scan with nodular infiltrates becoming smaller, but still maintaining moderate intensity.

a)



Max lesion intensity	Mixed 21 (28%)	Improved 43 (57%)	Resolved 12 (16%)	Mixed 8 (100%)	Mixed 4 (33%)	Improved 7 (58%)	Resolved 1 (8%)
None (5)	-	-	3	-	-	-	1
Minimal (9)	-	-	9	-	-	-	0
Mild (11)	0	9	-	0	0	2	-
Moderate (23)	6	13	-	2	1	0	-
High (41)	12	19	-	3	2	4	-
Very High (10)	3	2	-	3	1	1	-

b)



Max lesion intensity	Mixed 7 (64%)	Improved 4 (36%)	Resolved 0	Mixed 0	Improved 2 (66%)	Resolved 1 (33%)
None (0)	-	-	0	-	-	0
Minimal (1)	-	-	0	-	-	1
Mild (2)	0	1	-	0	1	-
Moderate (4)	1	2	-	0	1	-
High (1)	0	1	-	0	0	-
Very High (6)	6	0	-	0	0	-

Figure 9 Clinical outcome and associated M6 PET/CT scan findings.

(a) South-Africa: Clinical outcome and associated M6 scan response pattern. Mixed scan responses were more likely to have an unfavourable outcome (Fisher exact test, two-sided: P = 0.002). Data per patient in Appendix A. (b) South Korea: Patients grouped into those diagnosed with MDR strains or with drug-sensitive or rifampicin mono-resistant (Rif-mono) strains. All MDR cases culture converted within 6 months, while the DS/Rif-mono group includes nine cured, one un-evaluable and one recurrent patient. Data per patient in Appendix C.

3.4 PET/CT findings 1 year after completion of treatment

In view of the findings in the M6 scans we added a fourth scan, 1 year after the end of treatment (EOT + 1y) for 50 South-African patients who had not yet passed this time-point and who were cured at M6. Eight of these 50 patients were diagnosed with recurrent disease by healthcare providers within two years of treatment completion, five before the EOT + 1y scan and three subsequently. The other 42 maintained cure.

When the EOT + 1y scans were compared to the M6 scans, there was improvement in size and intensity of most residual lesions. However, only 32% of EOT + 1y scans were completely resolved. The remaining 68% had significant residual lesion(s), half of which had improvement of all lesions and the other 34% a mixed lesion response compared to the M6 scan. Morphology of new FDG-avid lesions at EOT + 1y included nodular infiltrates (found in four cases), hilar lymph nodes (in 1 case), cavitation (2), consolidation (2), or lesions with combined morphology (3). Residual M6 lesions showing similar or more intense FDG uptake at EOT + 1y included consolidation (2), cavitation (4) and nodules (2). Examples of EOT + 1y scan progression can be seen in Figure 10. All three patients who developed recurrent PTB after EOT + 1y had mixed scan outcomes at this time-point, while none with resolved scans was diagnosed with recurrence (Figure 11).

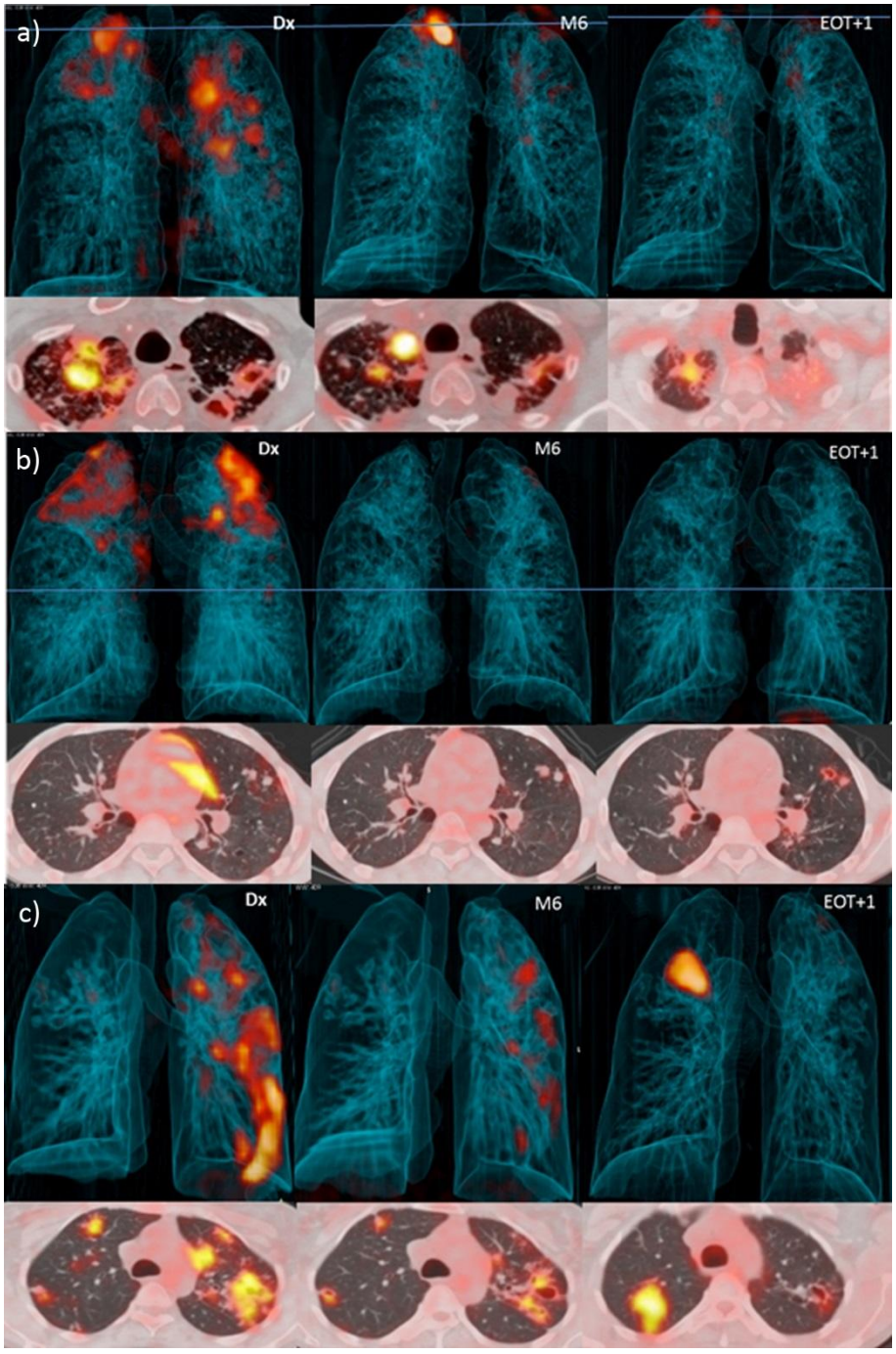
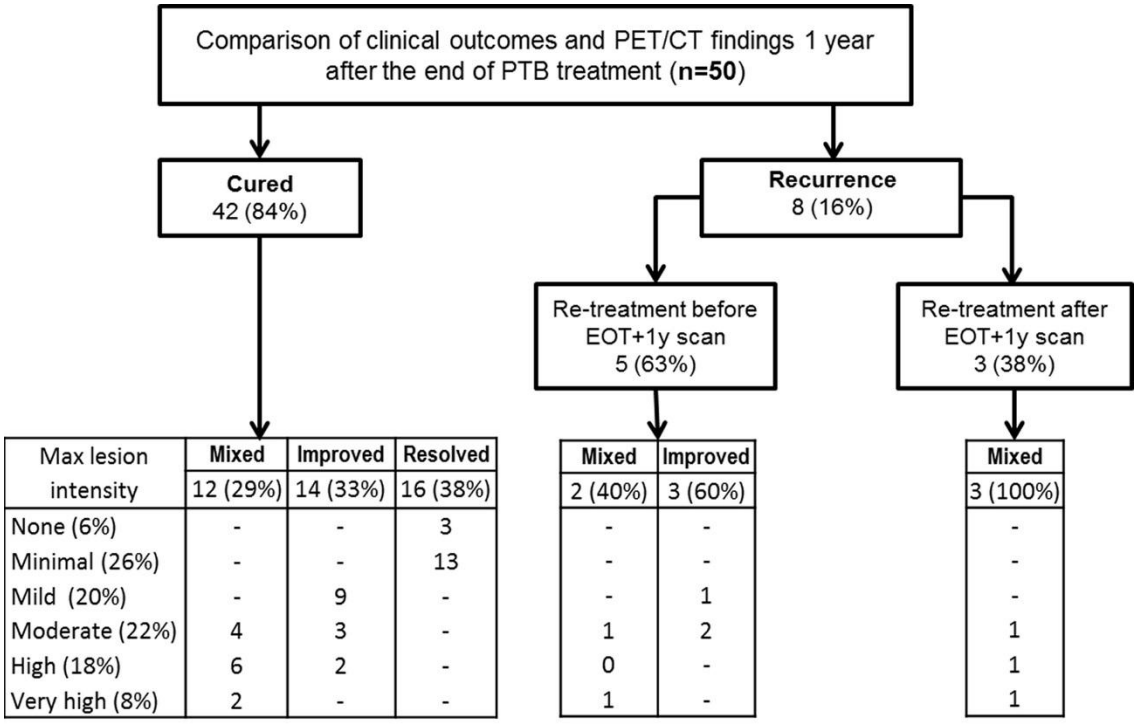


Figure 10 Dx, M6 and EOT + 1y PET/CTs for 3 representative patients after 6 months of treatment.

3-dimensional anterior and transverse slices at the level of horizontal blue line. a) A large new FDG-avid nodule in right lung apex lesion develops by M6 but, improves over the next year, with some residual increased FDG uptake. b) Bilateral upper lobe cavities improve during treatment, but nodules with normal FDG-uptake at M6 develop new cavitation and increased uptake by EOT + 1y. c) Multiple lesions with continuing improvement of FDG-avidity at M6 and EOT + 1y but with a large new area of consolidation at EOT + 1y. This patient was Gene Xpert positive, but culture negative at the time of EOT + 1y scan, but diagnosed with recurrent disease 6 months later (18 months after treatment completion), while patients a) and b) maintained cure.



Max lesion intensity	Mixed	Improved	Resolved
	12 (29%)	14 (33%)	16 (38%)
None (6%)	-	-	3
Minimal (26%)	-	-	13
Mild (20%)	-	9	-
Moderate (22%)	4	3	-
High (18%)	6	2	-
Very high (8%)	2	-	-

Mixed	Improved
2 (40%)	3 (60%)
-	-
-	-
-	1
1	2
0	-
1	-

Mixed
3 (100%)
-
-
-
1
1
1

Figure 11 Clinical outcome and associated PET/CT scan findings 1 year after the end of successful PTB treatment (EOT + 1y).

Patients are grouped based on whether they remained free from PTB or diagnosed with recurrent disease within 2 years after the end of treatment. The patients with a recurrent PTB episode are also divided into those that received re-treatment before or after the EOT + 1y scan. EOT + 1y scans were compared to corresponding M6 scans and categorised as Mixed, Improved or Resolved. Lesion intensity ranking categories are shown in bottom row of boxes for the respective response patterns. Data per patient supplied in Appendix B. Due to rounding of percentage values totals may equate to 99% or 101%.

3.5 DNA and mRNA in M6 sputum

Patients with an Xpert positive M6 sputum were more likely to have an unfavourable clinical outcome (7/8 failed treatment, 6/12 recurrent and 16/65 cured; $p < 0.001$) and M6 PET/CT lesions with high to very high intensity ($P = 0.04$).

mRNA detection assays were performed on M6 sputum from 75 consecutive participants (60 cured, 4 failed treatment, 9 recurrent and 2 un-evaluable) who produced adequate volume sputum, as well as on sputum from 20 community controls and 5 controls with lung disease other than TB (OLD). 13 MTB-specific mRNA targets [*85B, dosR, hspX (acr), pstS1, carD, nuoB, tgS1, TB8.4, TB31.7, acpM, icL, prcA and rrnAP1*] were assayed.

Among the 75 patients' M6 sputum, at least one MTB mRNA target was detected in 29 (39%): 22 (37%) of the cured, 4 (100%) of the failed treatment, 2 (22%) of the 9 recurrent

PTB and 1 of the un-evaluable outcome patients. The most frequently detected transcripts were *hspX* (*Acr*), *acpM* and *rrn*, respectively present in 20, 17 and 12 patients' M6 sputum. *Tgs1* and *acpM* transcripts were detectable at low levels in one community control each.

We performed principal component analysis (PCA) to reduce the dimensions of all transcript levels into one variable that indicated mRNA presence. PCA results in Figure 12 illustrate that all cases with failed treatment and 22 cured patients still had detectable MTB mRNA in their M6 sputum. Two community controls and no controls with other lung disease had low levels of mRNA transcripts (p = 0.001). There was a trend for patients with very high M6 PET/CT lesion intensity to have MTB mRNA in M6 sputum and patients with no significant M6 scan uptake to be sputum mRNA-free (P = 0.08).

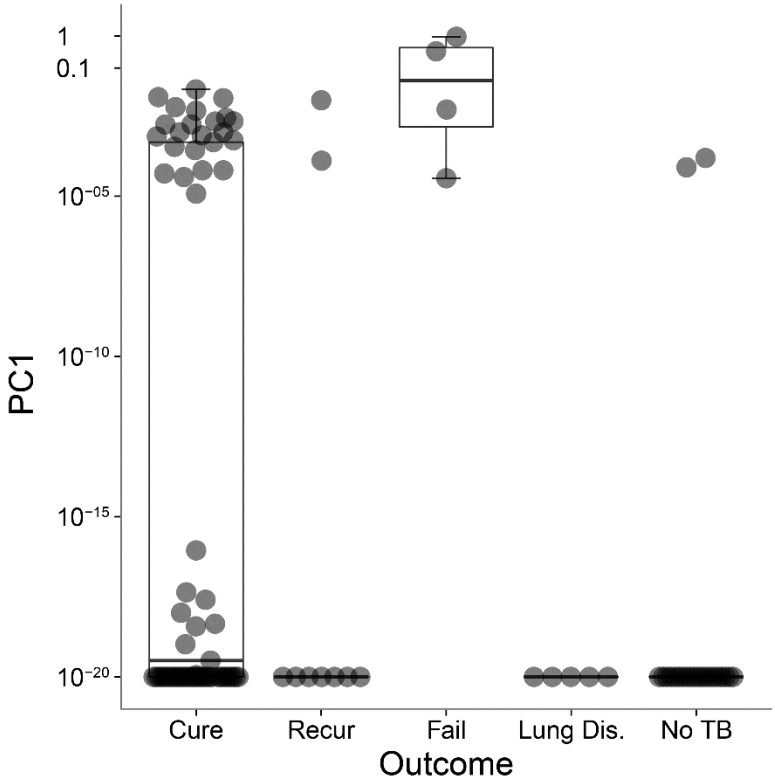


Figure 12 Box and whisker plots of the transformed PC1 of mRNA transcript PCR data

Principal component one (arcsinh); derived from principal component analysis (transcript copy numbers for each patient supplied in Appendix D), demonstrating separation by detectable MTB mRNA as captured in PC1. M6 sputum from 75 PTB cases by clinical category; 22 of 60 cured cases, all 4 failed treatment cases and 2 of the 9 recurrent cases in the group have detectable MTB mRNA in their sputum. None of the other lung disease participants and two of 20 healthy community controls have detectable MTB mRNA (p < 0.01).

3.6 RNA in end of treatment bronchoalveolar lavage samples

We obtained bronchoalveolar lavage (BAL) from the most affected lung segment identified on PET/CT as well as post-bronchoscopy sputa from 15 consecutive, consenting patients with DS PTB within 3 months after EOT. Treatment duration was 6 months for this group. We also collected BAL fluid in 10 control patients undergoing bronchoscopy as part of clinical workup for suspected lung cancer. The 15 EOT patients included 13 cured (M6 culture negative), 1 recurrent (M6 culture negative, but converted back 3 months after treatment) and 1 failed treatment (M6 culture positive) case. The 10 controls included one patient in whom new active PTB was subsequently diagnosed on BAL culture (although sputum culture was negative). All other BAL and post-bronchoscopy sputa were negative for MTB in liquid culture. Cancer was confirmed in 5 controls, while 3 controls had bacterial infections and one had interstitial fibrosis. Five of these controls had a positive Quantiferon® blood test (suggesting latent TB infection). Xpert was positive on BAL or post-bronchoscopy sputum at EOT for 10 of the 15 PTB cases and the new PTB case and MTB DNA was detected in all EOT BAL samples by qRT-PCR.

Blinded mRNA detection assays were performed on all BAL samples. 19 MTB-specific mRNA targets (*recF, eccD3, menA, iprB, Rv1421, gabD2, Rv1910c, hspX, fadD9, tgs1, fadE34, Rv3675, sodA, trxB2, rpsJ, rplV, sigI, sigH and rpsK*) previously shown to be differentially regulated in stress or persistence[95,96] were assayed.

A median of 8 mRNA transcripts were positive in the BAL from EOT cases (range 1-19); one MTB mRNA transcript was detected in BAL from 3 of the 5 Quantiferon® positive control cases (EOT vs controls $p < 0.001$). The target most frequently detected was *hspX*, present in the treatment failure and new diagnosis cases, as well as in 14 of the cured patients, but none of the controls. The newly diagnosed PTB case (from the control set) had 5 transcripts. We used PCA to summarise mRNA positivity (Figure 13 shows principal component 1 in and Appendix D provides transcript values per patient).

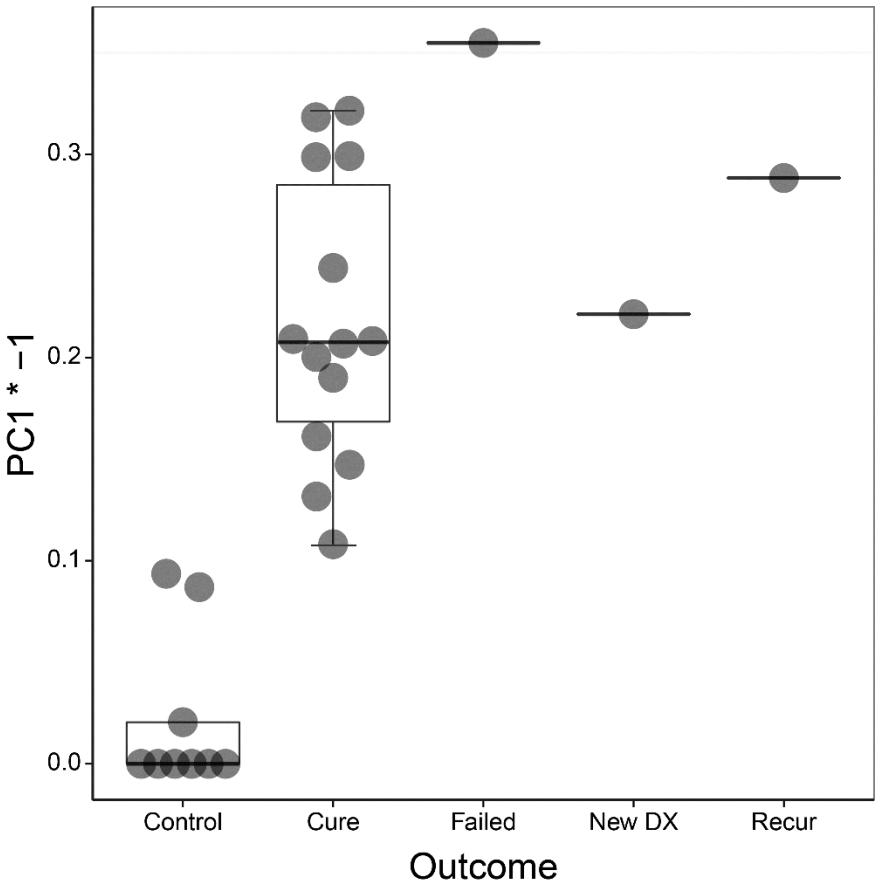


Figure 13 Box and whisker plots of the transformed PC1 of mRNA transcript PCR data

Principal component one (arcsinh); derived from principal component analysis demonstrating separation by detectable MTB mRNA as captured in PC1 for End of treatment (EOT) bronchoalveolar lavage fluid by clinical category. The new PTB case and 15 EOT cases (1 failed treatment, 1 recurrent and 13 cases that maintained cure), all still have significantly detectable mRNA. Three of the 6 Quantiferon® positive controls show inconclusive levels of MTB mRNA, whereas the Quantiferon® negative controls have no detectable MTB mRNA (P < 0.001). mRNA values per case in Appendix D.

3.7 Statistical Associations between grouped variables

The aim of the analysis was to test which clinical, imaging and microbiological factors had either meaningful or confounding associations.

3.7.1 Scan response patterns and clinical outcome

Table 3 Cross tabulation of clinical outcome and scan response pattern

	Mixed	Improved	Resolved
Cure	21	43	12
Recur	4	7	1
Fail	8	0	0

Fisher's Exact Test: **p-value = 0.002061**

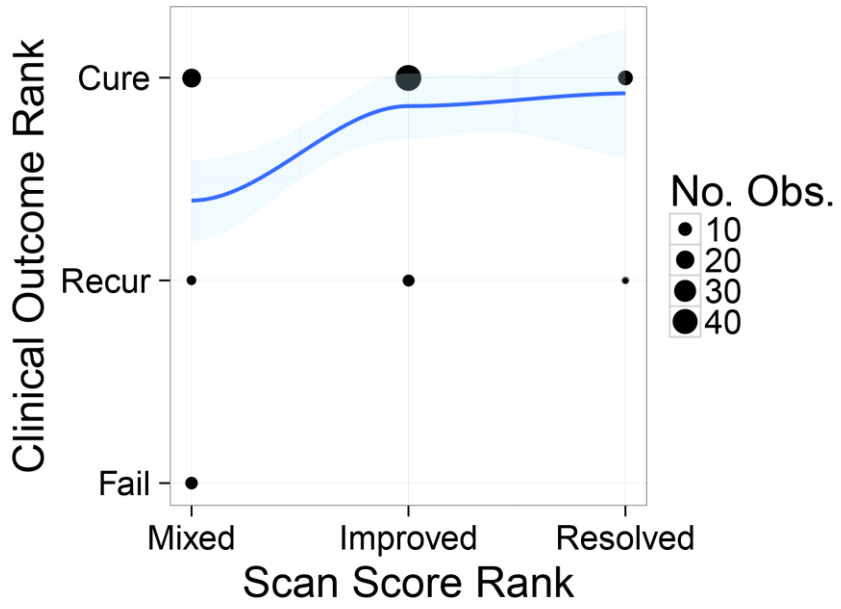


Figure 14 Association between clinical outcome and scan response pattern.

The size of the dots represent observation frequencies. The blue line shows the linear fit and the 95th confidence interval.

To perform the proportion trend test, we pooled failed and recurrent treatment groups into unfavourable.

Table 4 Association between scan response pattern and collapsed clinical outcome

	Cured	Unfavourable
Mixed	21	12
Improved	43	7
Resolved	12	1

Chi-squared Test for Trend: **p-value = 0.009**

Conclusion:

The data in the contingency tables (Table 3,Table 4) are clearly not independent by the Fisher exact test ($P < 0.004$ with or without the unevaluable class). There is a good correlation (trend; Figure 14) between scan score and clinical outcome, with a $P < 0.01$ for the Cochran-Armitage trend test when the clinical outcome is collapsed into favourable and unfavourable outcomes.

The same trend is shown in the graphics where clinical outcome and scan score are ranked ordinally. The linear fit for the graphic data is inappropriate (diagnostics show a poor fit for the QQ-plot, many points with high leverage and large distance); nevertheless, the graphics show the weighted trends more effectively than the unweighted proportions against the ranks of scan scores.

3.7.2 M6 Sputum Xpert and outcome

Table 5 Association between clinical outcome and M6 Xpert result.

	Xpert Negative	Xpert Positive
Cure	59	16
Recur	6	6
Fail	1	7

Fisher's Exact Test: **p-value = 0.0002**

There is a clear association between M6 xpert positivity and an unfavourable outcome (Table 5). A positive M6 Xpert implied a relative risk of unfavourable outcome of 4.2, with a sensitivity of 0.65 and a specificity of 0.79.

3.7.3 M6 Sputum mRNA vs Outcome & Diagnosis

Table 6 Association between clinical status and sputum MTB mRNA.

	Positive	Negative
Fail	4	0
Recur	2	7
Cure	23	29
Community Controls	2	18
Other Lung Disease Controls	0	5

Fisher's Exact Test: **p-value = 0.0006**

While the mRNA status and clinical status are not independent of each other, the distribution of the data (Table 6) is not what is expected given the marginal totals. If the data were distributed randomly (e.g., equal chance of being positive or negative) the P-value would be unremarkable.

The data do not permit testing for a trend as there is no reasonable ordering of diagnostic classes. Collapsing all TB into one class and all controls into another (Table 7), the trend test will approximate the Fisher test, i.e., a test of non-random distribution.

Table 7 Association between collapsed clinical status compared to sputum MTB mRNA.

	Positive	Negative
EOT TB	29	36
Control	2	23

Asymptotic General Independence Test: **p-value = 0.001.**

Fisher's Exact Test for Count Data: **p-value = 0.001**.

Detection of MTB mRNA was significantly more likely in patients at end of treatment, compared to controls.

3.7.4 Broncho Alveolar lavage mRNA vs Outcome & diagnosis

Considering the RNA in BAL, ordinal arrangement of diagnosis make somewhat more sense. One could consider the worst condition to be failure, next worst newly diagnosed and untreated TB, then end-of-treatment TB (successfully treated recently ill with TB), and then controls. Among controls, there is a distinction between latently infected controls and Mtb-free controls. A surrogate marker is Quantiferon status, positive indicating exposure to TB and likely latently infected. This is shown in Table 8.

Table 8 Association between clinical TB status and BAL MTB mRNA.

	mRNA Positive	mRNA Negative
Fail	1	0
New TB	1	0
EOT TB	14	0
QFN+ Other Lung Disease	3	3
QFN- Other Lung Disease	0	3

The Cochran-Armitage trend test then yields the following:

Asymptotic General Independence Test: **p-value = 0.0028**

Using a linear models implementation of the Cochran-Armitage trend test:

Chi-squared Test for Trend in Proportions: **p-value = 0.0003**

We also tested the trend for all TB (collapsing the single failure and newly diagnosed TB into the EOT TB) against the controls stratified by quantiferon status (Table 9). This reduced the degrees of freedom.

Table 9 Association between collapsed clinical status and BAL MTB mRNA

	mRNA Positive	mRNA Negative
NEW & EOT PTB	16	0
QFN+ Other Lung Disease	3	3
QFN- Other Lung Disease	0	3

Using a linear models implementation of the Cochran-Armitage trend test:

Asymptotic General Independence Test: **p-value = 0.00032**

Note that using this implementation we can collapse groups by assigning them the same "score" or category.

Chi-squared Test for Trend in Proportions: **p-value= 4.2 x 10^{-5}**

Fisher's Exact Test: **p-value = 0.0002**

3.7.5 Adherence, Smoking, Previous PTB vs outcome

Table 10 Association between clinical outcome and whether patient took >80% of doses

	Good adherence	Poor adherence
Cure	69	7
UE	1	2
Fail	4	4
Recurrent	11	1

The Pearson's Chi squared test for independence: **P-value = 0.0066**

This showed that the data (Table 10) is not independent, however, no non-linear trend or correlation were found.

When we collapsed the Outcomes into favourable (cured) and Unfavourable (Failed and Recurrent) and excluded Unevaluable, no association was found.

Fisher Exact test: **P-Value = 0.12006**

Table 11 Association between smoking status and collapsed clinical outcome

	Smoker	Non-Smoker
Cure	56	20
Unfavourable	18	2

Fisher Exact Test: **P-Value 0.147**

No significant association was found between smoking status and clinical outcome (Table 11). Note that non-smokers were underrepresented in our PTB patient cohort.

Table 12 Association between collapsed clinical outcome and previous PTB episodes

	1st episode	Previous PTB
Cure	54	22
Unfavourable	12	8

Fisher Exact Test: **P-Value = 0.418**

No significant association was found between clinical outcome and previous PTB episodes (Table 12).

3.7.8 Gene Xpert and M6 PET/CT scan

Table 13 Association between grading of most intense M6 PET lesion and M6 sputum Xpert.

	Xpert Negative	Xpert Positive
Very High	5	5
Moderate - High	52	22
Minimal	11	3

Fisher's Exact Test: **p-value = 0.3031**

Table 14 Association between M6 PET/CT response pattern and M6 sputum Xpert.

	Xpert Negative	Xpert Positive
Mixed	19	14
Improved	38	13
Resolved	11	3

Fisher's Exact Test: **p-value = 0.1963.**

No association was found between the grading of most intense M6 PET lesion and M6 sputum XPert positivity (Table 13), or between M6 PET/CT scan response pattern and Xpert positivity (Table 14).

3.7.8 M6 Sputum mRNA and M6 PET/CT scan

Table 15 Association between grading of most intense M6 PET lesion and M6 sputum Xpert

	Xpert Negative	Xpert Positive
Very High	2	5
Moderate - High	37	23
Minimal	7	1

Fisher's Exact Test p-value = 0.08077

We noted a trend for patients with very high intensity lesions to be more likely to have M6 sputum positive for XPert and mRNA and for patients with resolved scans to have no mRNA in M6 sputum (Table 15).

3.7.9 Adherence and M6 PET/CT scan

Table 16 Association between grading of most intense PET lesion and whether patient took > 80% of TB doses.

	Good adherence	Poor adherence
Very High	8	2
High	37	4
Moderate	19	4
Mild	10	1
Minimal	9	1
None	2	2

Fisher's Exact Test: p-value = 0.3248

Chi-squared Test for Trend in Proportions: p-value = 0.4279

There is no trend suggesting that adherence leads to detectably more intense lesions (Table 16).

Table 17 Association between adherence groups and scan response pattern.

Scan response pattern	Good adherence	Poor adherence
Mixed	26	8
Improved	48	3
Resolved	11	3

Fisher's Exact Test: **p-value = 0.03621**

Chi-squared Test for Trend in Proportions: p-value = 0.3472

There was no trend suggesting that adherence leads to detectably different scan outcomes (Table 17); although the groups seem to differ (non-adherence is not independently distributed among the groups). The Spearman correlation is 0.115, weak to no correlation.

Table 18 Association between adherence groups and M6 scan response pattern.

Scan response pattern	Good adherence	Poor adherence
Mixed	26	8
Improved/Resolved	59	6

Fisher's Exact Test: p-value = 0.06986

There was no significant evidence for association and a weak Spearman correlation of 0.19 for adherence resulting in better (improved/resolved) outcomes (Table 18).

3.7.10 M6 Xpert and M6 mRNA positivity

Table 19 Association between M6 sputum MTB mRNA and Xpert positivity.

	Xpert Negative	Xpert Positive
mRNA Negative	35	10
mRNA Positive	21	8

Fisher's Exact Test: p-value = 0.78198

There is no significant association between positivity of mRNA and Xpert in M6 sputum (Table 19).

3.7.11 Poor adherence and Xpert positivity

Table 20 Association between M6 Sputum Xpert positivity and whether patient took >80% of doses.

	Good adherence	Poor Adherence
Xpert Negative	62	6
Xpert Positive	22	8

Fisher's Exact Test: **p-value = 0.02882**

Chi-squared Test for Trend in Proportions: **p-value = 0.01999**

Spearman correlation: 0.235008

We found a signicant trend suggesting that poor adherence leads to M6 sputum Xpert positivity, in spite of the weak Spearman correlation of 0.24. The trend is supported by the Fisher test showing an association (Table 20).

Table 21 Association between M6 Sputum MTB mRNA positivity and whether patient took >80% of doses.

	Good adherence	Poor adherence
mRNA Negative	40	6
mRNA Positive	26	3

Fisher's Exact Test: p-value = 1

Chi-squared Test for Trend in Proportions: p-value = 0.7262

There was no evidence of non-independence between adherence and Mtb RNA presence (Table 21) and there is essentially no correlation (Spearman correlation of 0.04).

3.7.12 mRNA Copy number rank correlation

Table 22 Ranked mRNA transcript, on copy number in BAL from failed treatment case.

	S163	S139	S141	S144	S145	S146	S150	S153	S154	S155	S159	S161	S162	Mean
Mtb Gene	**Fail**	Cure	Cure	Cure	Cure	Cure	Cure	Cure	Cure	Cure	Cure	Cure	Cure	Cure
hspX	1	5		4	3	8	1	5	1	1	3	1	4	4.58
sodA	2			1	7						1			15
rpsK	3	2	5		1	4			5			4	5	10.08
rpsJ	4	1	6	7	2	2		2	4		2	2	9	6.25
lprB	5	4	2	3	12	1		3		2	6		6	8
tgs1	6			2		5					4		3	13.83
eccD3	7	7		5	8	9					7	5		12.92
trxB2	8	6	1	9		3					5		11	12.42
Rv3675	9	3	3	6	9			1	2		8	3	1	7.75
rplV	10		7	11	11	11							13	15.5
sigH	11				5	12								17.25
fadD9	12					10			3				12	16.33
gabD2	13					7							8	17.08
Rv1421	14			10	6								7	16.17
Rv1910c	15				10			4						17
sigI	16				4								2	16.33
menA	17		4	8		6							10	15
fadE34	18													19
recF	19					13		6					14	17

Since the copy number data were exponential (base 2), it was necessary to transform them to obtain a robust estimate of the relationship between copy number in the treatment failure (sputum and BAL culture positive) and in the cures. The RNA copy number data were ranked in descending order (most frequent RNA species set to 1) within each individual where 0 copies (undetected) was replaced by the highest rank (Table 22). Rank order in this sense can be considered equivalent to quantiles. The data were analysed in two ways: 1) Loess regression of all data points (ranks of cured being dependent variables and ranks of the treatment failure the independent variable), and 2) linear regression of the mean of the ranks of cured on the ranks of the treatment failure.

The Loess regression showed that there is a correlation and that although non-linear (Figure 15), it can be reasonably approximated by a line. Linear regression has the advantage that is more readily interpreted and provides a measure of the strength of the correlation (Figure 16). The Loess estimate corresponds to a locally weighted mean, and we used the mean of ranks for the linear regression. The fit was analysed for diagnostics of fit (QQ-plot and influence plot; data not shown). The diagnostics suggested a good fit with a few genes

having sizeable residuals (hspX, sodA, and Rv3675), but still fell close to the expected QQ line; and some genes being influential (hspX, sodA, and rpsJ), although only sodA had a Cook's distance barely more than 0.5 at high leverage (> 0.15).

We estimated the correlation using the Pearson and Spearman methods, as well as using the result of ordinary least squares regression. The Pearson correlation was 0.754 and the Spearman correlation was 0.774. The regression estimate (slope) was 0.5664 (SE 0.1198) and the p-value of the regression fit was 0.0001947 (F-statistic: 22.35 on 1 and 17 DF).

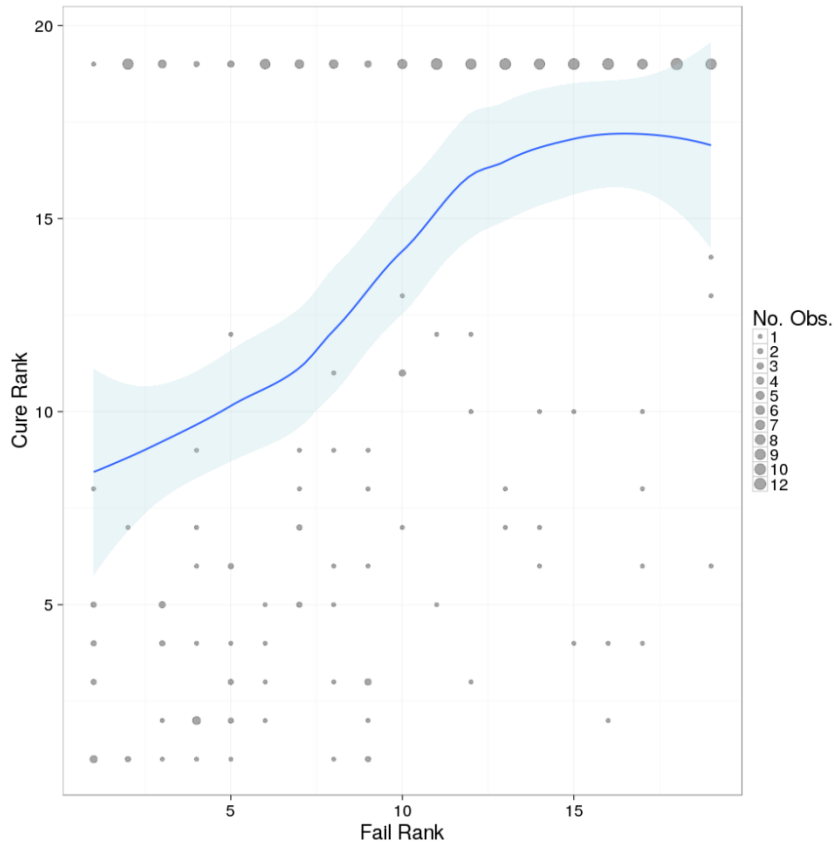


Figure 15 Loess regression of gene expression rank of cured patients on the rank of the failed patient.

Since the data are integers several observations can occur with the same coordinates; the size of the dots is proportional to the number of coincident observations. The Loess fit suggests a somewhat linear relationship, i.e., a correlation; the 95% confidence interval is shown in light blue shading.

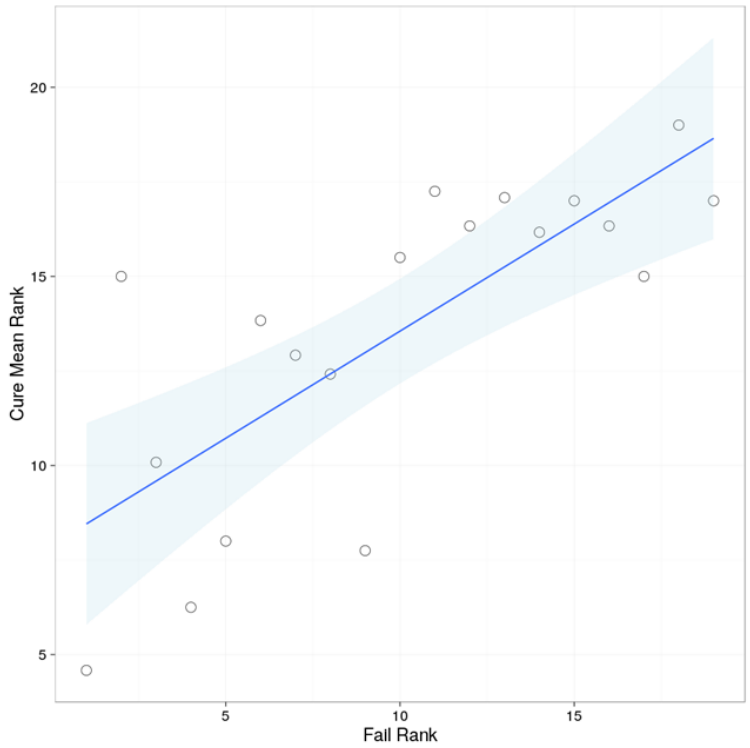


Figure 16 Ordinary least squares linear regression of mean rank of genes of cured on rank of failed case.

The fit and 95% confidence interval (light blue shading) are shown. The fit is a reasonable approximation of the Loess regression fit.

The genes with the highest expression in the failed treatment samples were also the most abundant mRNA's in the cured patients. We speculate that the most abundant mRNA's per bacterium are most easily detected when cell counts are low.

3.7.13 Summary of associations.

Significant associations included: 1) Clinical outcome was correlation with M6 scan response pattern and M6 sputum Xpert status. 2) Clinical status (new PTB diagnosis, PTB after treatment, latent TB or no TB) correlated well with the presence of MTB mRNA in sputum and BAL fluid. 3) Poor adherence was association with a failed treatment outcome and M6 sputum Xpert positivity.

Comparisons where no significant association was found included: 1) Clinical outcome was not associated with smoking status or previous PTB episodes. 2) Neither the PET/CT response pattern, nor the lesion grading was correlated with M6 sputum Xpert positivity or adherence. 3) There was no association between Xpert positivity and mRNA detection in M6 sputum.

4. Discussion: Persistent lesions and MTB mRNA detected after treatment

Summary

We found ongoing inflammation, as detected by PET/CT, in the majority and the presence of MTB mRNA and DNA in respiratory samples from a substantial number of cured South African PTB patients at the end of treatment. Similar PET/CT response patterns were seen in a smaller South-Korean cohort, for whom RNA data were unavailable.

Significance of lesion activity after treatment

Our study confirms the persistence of TB lung lesions previously described on CT and CXR.[21,22] We also frequently found residual symptoms or signs after treatment, similar to other reports on post-tuberculosis lung sequelae.[24–29,104] In addition, although most lesions improved and scan response patterns showed an association with clinical outcomes, we found that the majority of patients still had lesions with an ongoing inflammatory response on PET and roughly a third of patients showed intensified or new lesions. These dynamic patterns were found after anti-TB treatment and even a year later, regardless of drug sensitivity, sustained culture conversion or clinical cure. A European study on 35 patients with pulmonary or extra-pulmonary TB reported persistent PET/CT lesions in 15 and progressive lesions in 4 patients after a mean treatment duration of 16.1 months,[73] while a Japanese study reported resolution of all lesions in 8 patients with DS-PTB after 12 months of treatment.[72] It's not clear whether the favourable scan outcome in the Japanese cohort is related to extent of disease, treatment duration, lower infection pressure or better living standards and health.

FDG uptake only reflects the increased metabolic activity associated with inflammation and does not necessarily imply active MTB infection. After the successful treatment of other infective lung pathogens, however, radiological lesions typically resolve within 6 weeks[105] and relative intensity on PET has shown correlation with MTB bacterial load in animal lesions.[61] PTB related pathology is the most concordant with our PET/CT findings, considering the underlying morphology of individual lesions and the absence of clinical findings supporting other diagnosis. Persistent antigens, associated with dead MTB, might lead to increased FDG-avidity after treatment, but it is unlikely to cause very high intensity, intensified or new lesions. PET/CT imaging provides insight into the local host response and further research is required to unravel the interaction with MTB.

The individual fate of TB lung lesions have been demonstrated fairly consistently in reports of animal models: in Cynomolgus Macaques,[62,63,106] as well as in rabbits[61] and to some extent in mice.[66] The similarities found between the clinical cases in our study, the first report of such comprehensively investigated patients, and the respective animal models add significantly to the importance of our findings. Our data provide support for the use of these different animal models in evaluating TB radiological response. The demonstration of similar responses in humans and the animal models suggest that these animal models can indeed be used to build scientific hypotheses to explain our findings.

What makes our findings even more interesting when compared to the animal models are:

i) The spontaneous progression and resolution in of individual lesions on PET/CT in animal models were mainly described prior to anti-TB treatment, not during or after anti-TB treatment.

ii) The longest treatment period in these models was 12 weeks, vs 24 weeks and EOT + 1y for our clinical cases.

iii) No increase in intensity or size of individual lesions were shown in animals after 8-12 weeks on standard 4 drug anti-TB treatment, only on monotherapy.[59] On 4 drug therapy all lesions improved.

iv) CFU numbers in animal lesions did correlate to reduction in PET uptake intensity.

v) Considering results from another mice model experiment,[34] that demonstrated MTB mRNA transcripts in culture negative lesions taken from mice after 14 weeks of dual anti-TB treatment, followed by bacteriological relapse after high-dose steroid treatment in 21 of 23 mice that were followed up, it is reasonable to propose that culture even culture negative lesions may harbor viable bacteria.

<u>Significance of MTB DNA and mRNA detection after treatment</u>

Our results confirm the high Xpert positivity rate previously reported in culture-negative sputum samples and noted an even higher rate in bronchoscopy samples.[31,32] In addition, we detected MTB mRNA in 35% of culture-negative sputum samples at M6 and in all BAL fluid samples taken from 15 PTB patients at EOT. Alpha-crytallin (*hspX*), linked to long-term MTB persistence,[107] was present in 20 (27%) of the M6 sputum samples and 14 (93%) of the EOT BAL samples, while absent in M6 sputum and BAL of all controls. The detection of MTB mRNA in 2 separate types of respiratory samples suggests either ongoing transcription, based on the short half-life of mRNA,[14,15] or persistence of stabilised mRNA as

seen in non-replicating MTB *in vitro*.[36] Although persistent MTB DNA in intact, non-viable bacteria may explain Xpert positivity, the same is unlikely to apply to MTB mRNA.

Strengths and limitations

This is the largest study examining PET/CT scans and mRNA TB patients at the end of treatment; however, it is a relatively small study, with a small number of unfavourable outcomes and without differentiation between relapse and re-infection. We performed spoligotyping on sputum culture at diagnosis for most cases, but sputum culture was not available at time of re-diagnosis in most recurrent cases. Other studies in Cape Town reported similar recurrence rates and if onset of recurrence occurred within 2 years of EOT, found roughly 66% due to relapse.[9,10] We could not establish whether the positivity of some mRNA transcripts in the sputum of two controls could be the result of latent or subclinical MTB infection or transient MTB excretion after recent exposure;[108,109] or eliminate the possibility that even pre-validated probes could show cross-reactivity. However, in a blinded test, the EOT patients scored more MTB transcriptional signals, more frequently than the community controls and most probes did not show any positivity in the controls. The extent to which the stochastic detection of low levels of MTB mRNA in individuals in communities with a high TB-burden represents true detection of MTB in individuals without manifestations of clinical TB has yet to be determined.

The paradox that some samples were Xpert negative, but mRNA positive could be explained by the different sensitivities of the assays: the mRNA isolation procedures concentrate specimens prior to the assays and usually there are many copies of an mRNA species per genome copy. We also observed Xpert-positive, mRNA -specimens, possibly due to greater stability of DNA. In BAL samples, however, we observed a good concordance between mRNA and DNA detection, since the BAL samples were from targeted sampling of the affected lobe, processed immediately and concentrated. The presence of DNA and mRNA in all BAL samples, suggest that both are usually present in parenchyma, but may be variable in sputum. We found an excellent correlation between MTB mRNA in BAL and clinical TB status ($p < 0.001$; as noted in statistical associations section 5.7.4).

Conclusion

Our study was exploratory and the patients were investigated intensively, allowing complementary results from different tests. In summary, we describe a marked heterogeneity in treatment responses of individual PET/CT lesions and the presence of MTB mRNA in sputum and BAL of a substantial number of cured patients at end of treatment. Although MTB mRNA stability remains a possible explanation for its presence in intact

bacteria, we suggest that viable MTB, with the potential to elicit a host response, often persists even after clinically curative treatment. The higher relapse rates in patient groups with impaired immunity[8,17,43,110] support the concept that a competent immune response plays an important complementary role in the ultimate control of remaining bacteria after antibiotic treatment. The nature of host-pathogen interaction and association with poor treatment outcome and post-tuberculosis lung impairment need to be investigated in future studies. Sterilising drug or host-directed therapies and improved treatment response markers are likely needed for the successful development of improved or shortened PTB treatment strategies.

5. Results: Operator independent technique to quantify widespread complex tuberculous lung lesions on PET/CT.

5.1 Application of quantification technique

As a proof of concept, the quantification methodology described in Section 2.6, page 19, was tested to quantify disease burden on PET/CT scans from five controls, and the first 5 patients participating in the South-African PTB cohort. The controls were from the same communities and had contact with PTB patients, but were sputum culture negative for mycobacterium tuberculosis and had no active lesions visible on PET/CT scan. PTB cases were all diagnosed with drug-sensitive TB strains and were HIV uninfected. Patients underwent scans at time-points within 1 week from initiation of treatment (Dx), after 1 month of treatment (M1) and after 6 months of treatment (M6), the duration of standard treatment. At the end of treatment, four of the PTB patients were classified as cured by healthcare providers in charge of treatment, while one was still sputum culture positive and diagnosed with a failed treatment outcome.

Time requirements

The quantification methodology was successfully implemented in all controls and patients. Creating the masks required a working knowledge of lung anatomy, while the other steps required basic computer literacy and a mid-range personal computer. User input required for quantification could be divided into 1) file management, including the selection, indexing and formatting of image files; 2) creating VOI's for lung masks, background references, as well as mediastinum and liver; 3) computation time, which included running the co-registration algorithm and the quantification calculations.

Time required to quantify the scans ranged from roughly 10 minutes for a single lesion-free lung scan (roughly 4 minutes for file management, 2 minutes computation and 4 minutes creating masks) to 45 minutes for a 3 time-points series of extensively diseased lungs (roughly 8 minutes file management, 12 minutes computation and 25 minutes creating masks). Where applicable, cavity volume were easily measured with 3D MRICro region grow tool on CT, taking less than 1 minute per scan.

Segmentation accuracy

The auto-segmented MLV for each scan corresponded well to visual assessment of the PET scans. No false positive segmentation was noted on control scans. In the PTB lung scans, all auto-segmented MLV's corresponded to areas that appeared FDG avid. False positive segmentation of uptake attributable to the liver or heart was noted in one case, which

necessitated repeating the quantification after adjusting the lung mask. No visually FDG-avid lesions were missed by auto-segmentation. Figure 17 compares the results from control cases with baseline PTB scans. TGAI, MLV, V_{hard} and MLV_{abn}, values approached zero for all controls, while showing increased values for all PTB scans. The same, however, did not apply to V_{soft}, which showed no clear distinction between the groups. V_{medium} values were below 2% for all controls and above 2% for all PTB patients. A control case with markedly higher values for V_{soft} and V_{medium}, showed signs of chronic obstructive airways disease, which was the likely cause for abnormal lung densities.

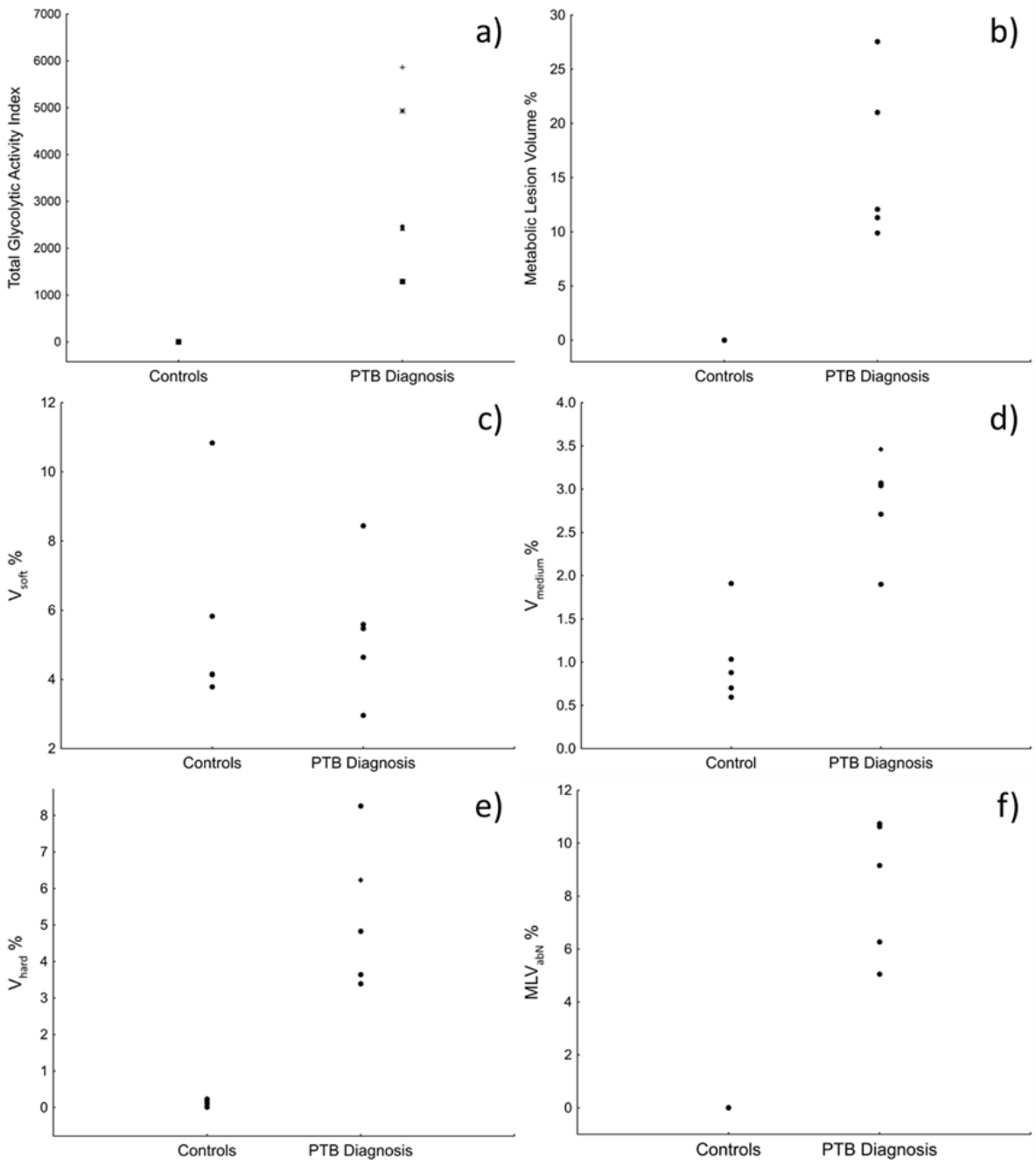


Figure 17 Comparison of PET/CT parameters between scans from 5 controls and 5 PTB patients at diagnosis.

a) Total Glycolytic Activity Index b) Metabolic lesion volume. c) Vsoft (-500 HU:-300 HU) CT lesion volume as a percentage of total lung volume. D) Vmedium (-300 HU:-100 HU) CT lesion volume as a percentage of total lung volume. E) Vhard (>-100 HU) CT lesion volume as a percentage of total lung volume. E) Area of lung that shows abnormal density (>-500 HU) and relative high uptake intensity as a percentage of total lung volume.

Pilot case profiles

Figure 18 demonstrates the dynamics of different scan parameters during treatment of the five PTB cases. The segmented TGAI (Figure 18a) of 3 out of 4 cured cases already show a partial improvement (30%-62%) at M1 and continue to improve markedly towards the end of treatment. The fourth cured patient shows a slight increase by M1, but improves significantly by M6. After 6 months of treatment all the cured patients showed marked improvement (80-99%), although only one patients showed nearly complete metabolic resolution. The patient that failed treatment remained stable over the first month, but deteriorated by M6. The MLV (Figure 18b) followed a very similar pattern during treatment.

The total lung volume with increased density ($V_{soft,\ medium,\ hard}$) followed a similar trend, although the reduction from Dx to M6 was less marked (41%-83%) and the failed treatment patient slightly decreased. In the breakdown of different density lesions (Figure 18c-e), all cured patients showed a marked volume decrease in hard lesions, already noted by M1 (35%-71%) and continued toward M6 (81%-98%). The volumes of medium and soft lesions were more variable and a relatively large residual volume at M6. Contributing factors to this might be conversion of voxels from V_{hard} to V_{medium} and V_{soft} during response to treatment, remaining post-TB fibrosis and a lower specificity at lower densities to pathological lesions.

After 1 month of treatment, in cured patients', MLV_{abN} (Figure 18f) decreased modestly in two and remained stable in the other two. The failed treatment patient showed a 13% increase. At M6, however, all cured patients showed marked decrease (82%-98%), while the failed treatment patient showed further increase.

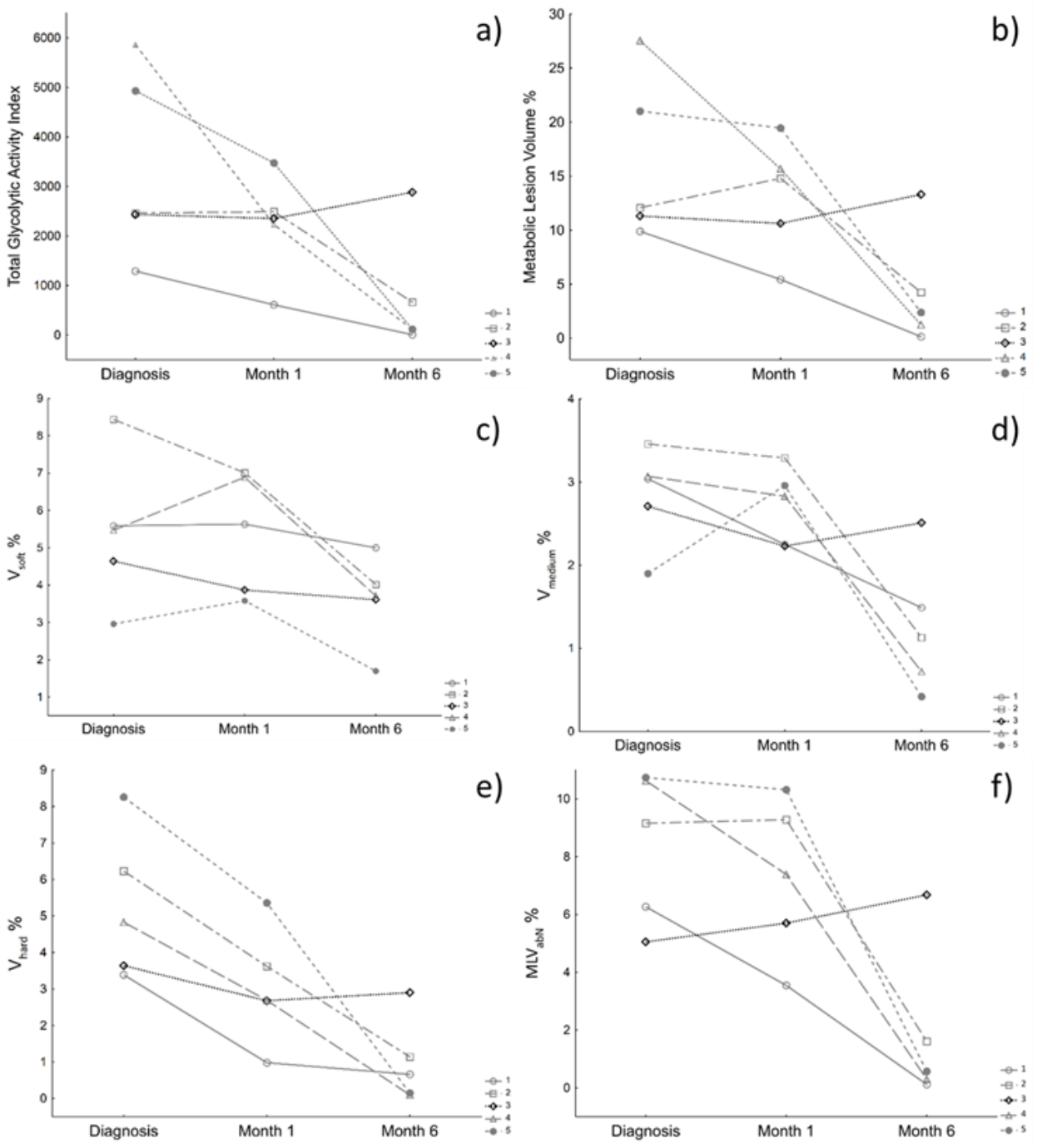


Figure 18 case-profiles of 5 PTB patients showing auto-segmented PET/CT parameters at Dx, M1 and M6.

Patient number 3 was the only patient with a failed treatment outcome (darker line). A) Total Glycolytic Activity Index B) Metabolic lesion volume . C) V_{soft} (-500 HU:-300 HU) CT lesion volume as a percentage of total lung volume. D) V_{medium} (-300 HU:-100 HU) CT lesion volume as a percentage of total lung volume. E) V_{hard} (>-100 HU) CT lesion volume as a percentage of total lung volume. E) Area of lung that shows abnormal density (>-500 HU) and relative high uptake intensity as a percentage of total lung volume.

5.2 Quantified scan results from diagnosis, month 1 and month 6.

Quantification was successfully performed on Dx, M1 and M6 scans from all 99 included participants in the South African cohort. Two additional challenges were encountered during application of the technique that was not apparent during the pilot or control cases. Firstly, Dx scans from two patients had such extensive disease, that it was impossible to identify visually lesion free areas of sufficient size to create two background references volumes. To solve this we created four smaller NL volumes to serve as background reference. Secondly, two other scans appeared to have some technical issues with the Dx scan that made the co-registration algorithm ineffective. This was resolved by co-registering Dx and M6 to the M1 scan. As previously discussed, cavity volume was measured separately in MRICro, using the 3D gradient based boundary tool.

In this section we discuss the various quantified characteristics which were analysed in different ways. The first analysis shows the relationship between changes in PET/CT metrics and time to sputum culture negativity. Subsequently, we analyse the correlation between different scan parameters. Finally, we discuss the risk of unfavourable outcomes entailed by different quantified PET/CT characteristics.

<u>PET/CT characteristics in relation to time to culture negativity group.</u>

Eighteen patients converted to sputum culture negative within 4 weeks, 39 by week 8, 22 by week 12, 9 by week 24 and 8 patients were still sputum culture positive after 24 weeks on treatment. Three patients were excluded from this classification after a panel of experts concluded that contaminated culture results prevented accurate assignment of a TTN. The mean value of all quantified scan parameters at each time-point were expressed in absolute values and proportional change from baseline.

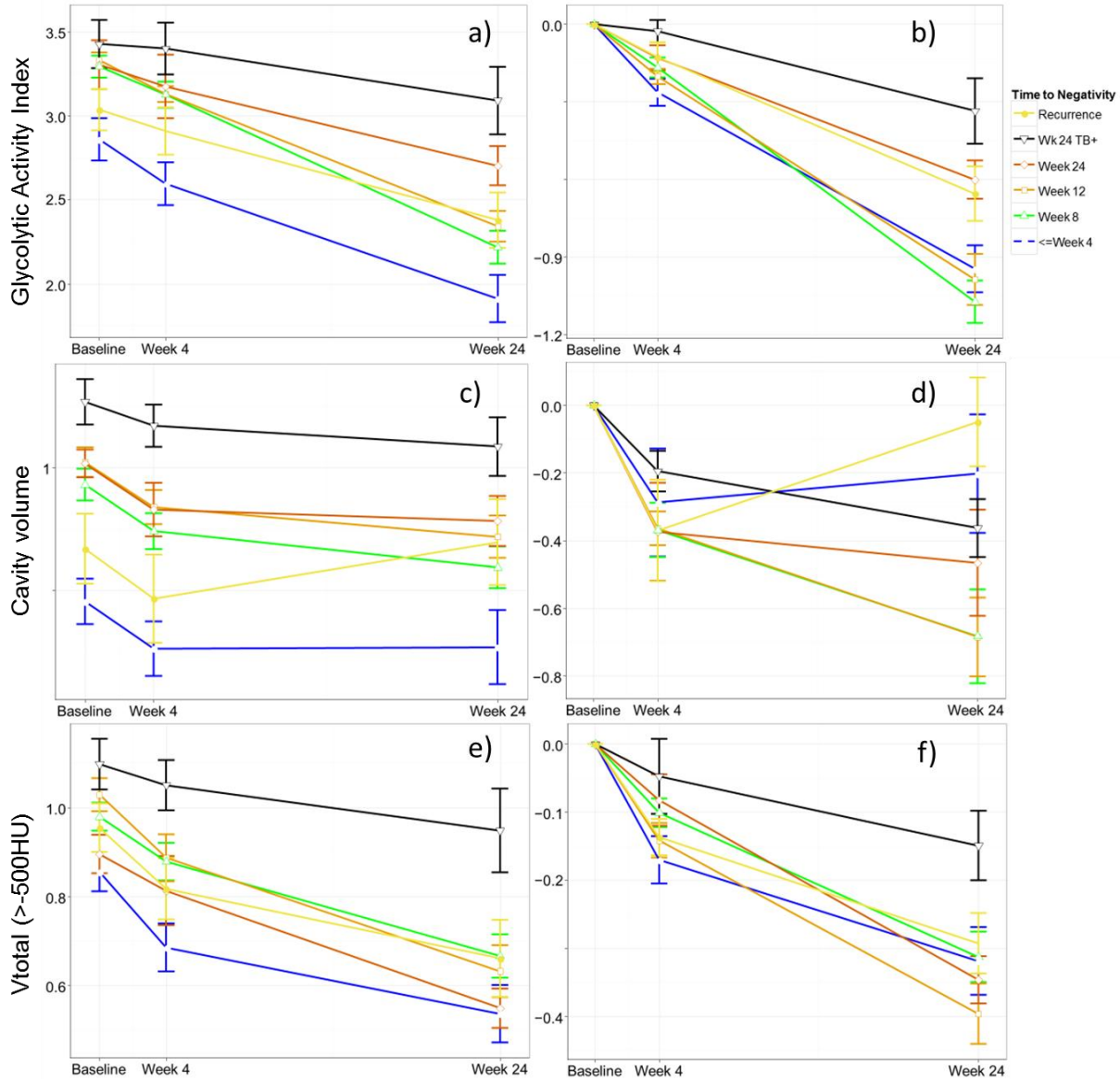


Figure 19 Mean (±SE) Log10 transformed values of principal PET and CT parameters over time by time to negativity group and recurrent cases.

Total glycolytic activity index in a) absolute values and b) change from baseline. Cavity volume in c) millilitres and d) change from baseline. Total high density CT lesions in e) percentage of lung volume and f) change from baseline.

TGAI, cavity volume and V_{total} could be considered the principal candidates to look for a measure to predict central trends for resolution or progression of lung pathology, since they hold the combined information from the main components: volume and uptake intensity on PET scan, total high-density volume on CT and total cavity volume on CT. A trend was evident for all three parameters, that and lower lesion burden was associated with earlier culture conversion and a higher burden with later conversion (Figure 19). This feature was most explicit for cavitation, where a significant difference between failed treatment cases and other groups were already seen at Dx.

Differentiation between the TTN groups tended to become more pronounced over time. This was confirmed by the graphs showing the change from baseline. It was less evident for cavity volume, where all TTN groups showed a rapid decline over the first month of treatment, but some stabilisation towards M6. However, stability from M1 – M6 was also seen in the W4 culture conversion group (fastest responders), which was likely influenced by the smaller initial cavity volume, resulting in earlier resolution with less room for improvement.

Unlike failed treatment cases, the recurrent patients did not have a large comparative baseline burden of disease. The recurrence group however, did show comparatively less improvement over time and even deteriorated in terms of cavitation volume.

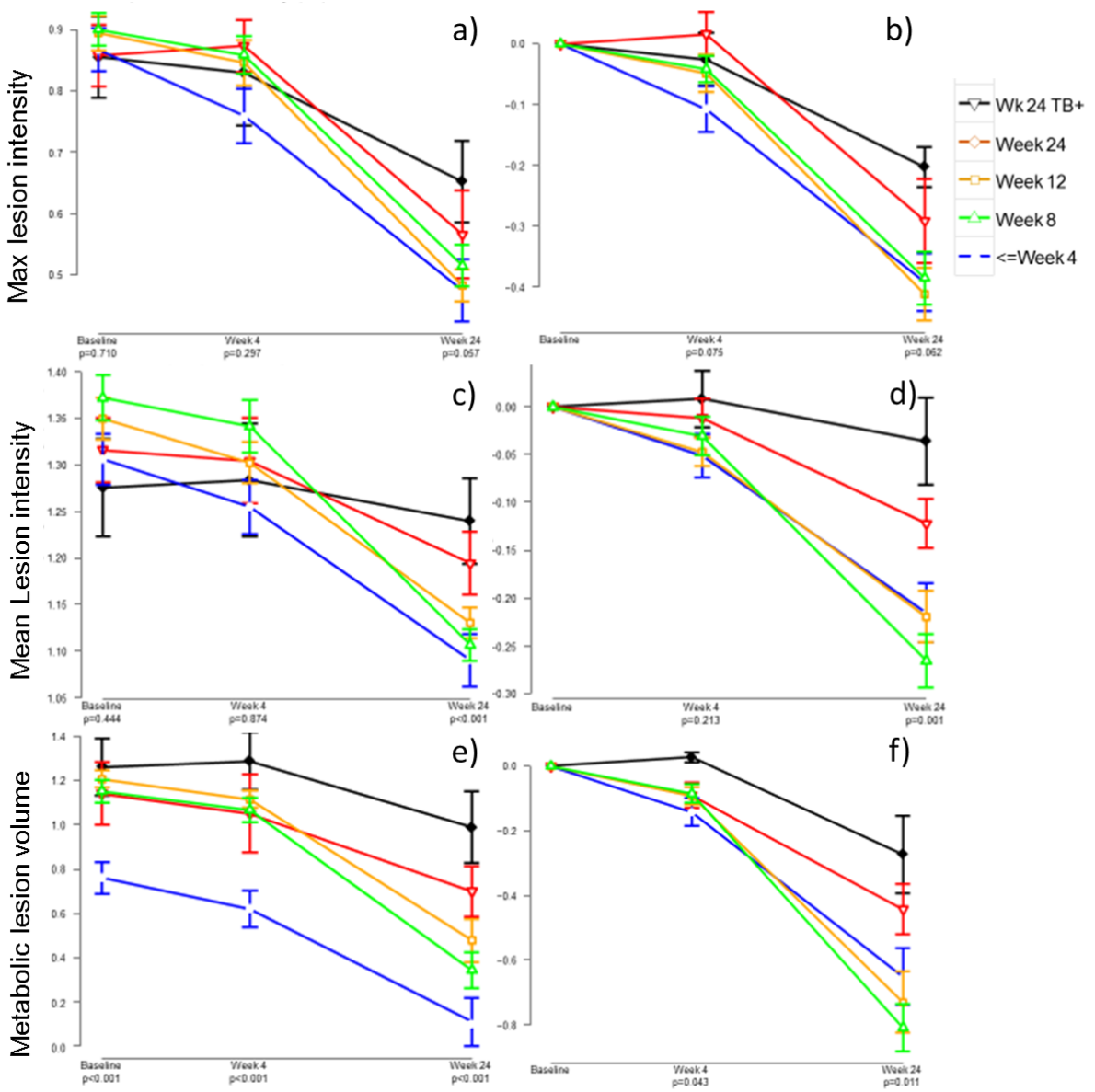


Figure 20 Mean (±SE) Log10 transformed values of PET parameters over time by time to negativity group

Maximum lesion intensity in a) SUVmax and b) change from baseline. Mean lesion intensity (Z_{mean}) in c) mean Z-score and d) change from baseline. Metabolic lesion volume in a) percentage of total lung volume and b) change from baseline.

Absolute values of maximum or mean lesion intensity did not show any prognostic or diagnostic potential. When expressed in terms of proportional change over time, however, both correlated very well with TTN groups (Figure 20). The SUVmax for the slow responders even showed a slight increase at M1. Metabolic lesion volume performed very similar to TGAI, but did not seem to differentiate the TTN groups quite as well.

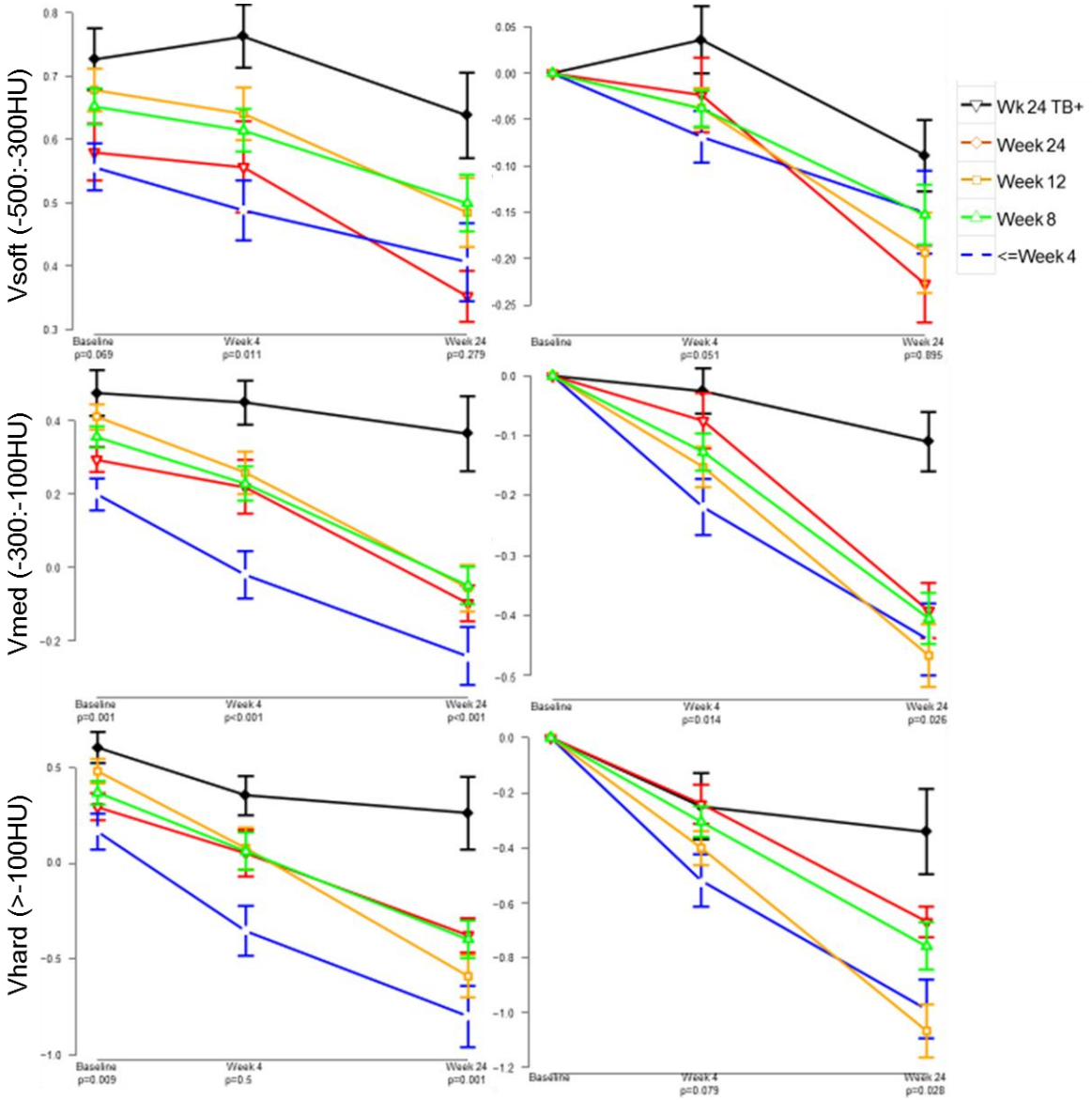


Figure 21 Mean (±SE) Log10 transformed values of CT parameters at Dx, M1 and M6 by time to negativity group.
Soft lesions (V_{soft}) in a) percentage of total lung volume and b) change from baseline. Medium lesions (V_{medium}) in a) percentage of total lung volume and b) change from baseline. Hard lesions in a) percentage of total lung volume and b) change from baseline.

In all the subdivisions of high-density lesions, failed treatment cases had significantly greater volume compared to cured cases. This was evident from as early as M1 and the gap between the groups became wider by M6 (Figure 21). V_{hard} showed a rapid decrease in volume in all groups. A rapid decrease by M1 is also noted for fast responders in V_{medium} and V_{soft}. For failed treatment cases, the mean for V_{medium} decreased only gradually and for V_{soft} an increase was seen in the first month.

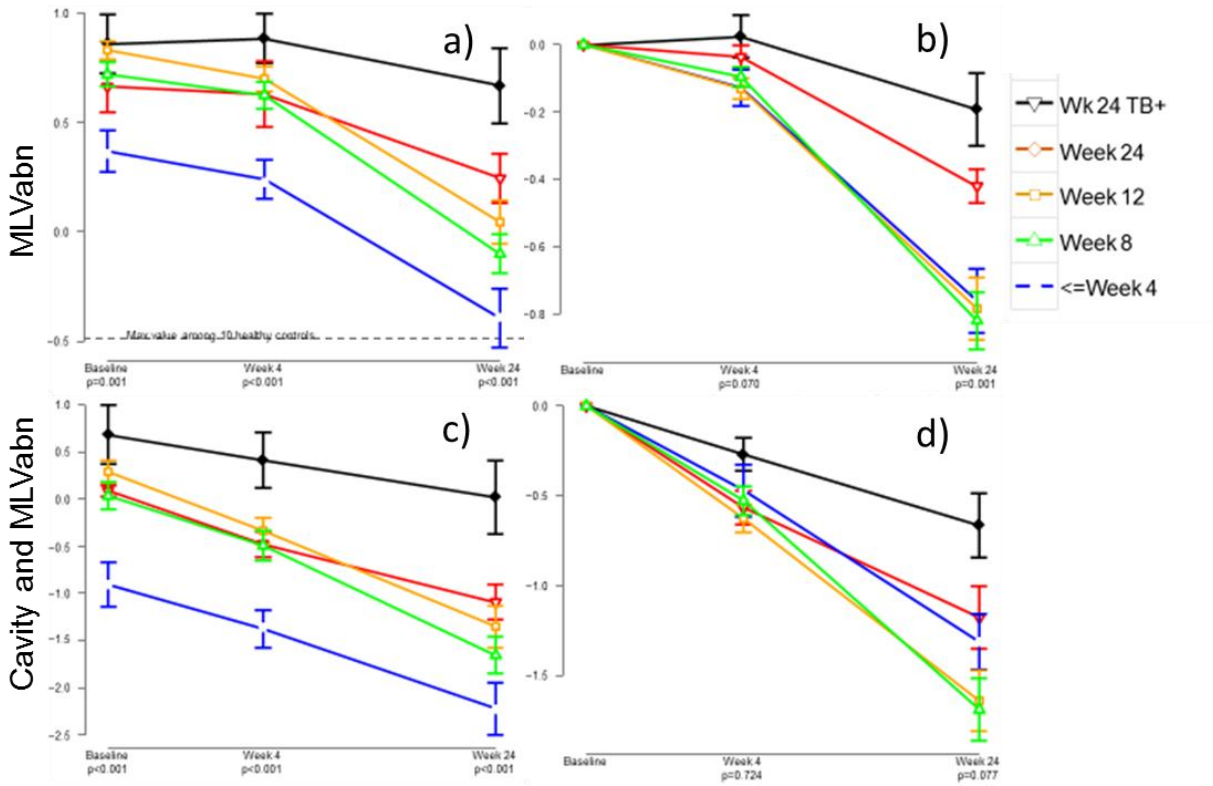


Figure 22 Mean (±SE) Log10 transformed values of combination PET/CT parameters at Dx, M1 and M6 by time to negativity group.
Volume with increased density and metabolic activity (MLV_{abn}) in a) percentage of lung volume and b) change from baseline. Weighted mean of cavity volume combined with MLV_{abn} in c) total and d) change from baseline.

Since MLV_{abn} is a combination PET and CT parameters, we expected it to show improved diagnostic accuracy and a very high specificity for diagnostic. While it did correlate to TTN groups (Figure 22), neither the absolute values nor the change from baseline seemed to perform better than the separated components. When a weighted cavity volume was added, however, the graph seemed to better differentiate the fast responders (TTN ≤ 4weeks) and failed treatment cases across time-points.

Correlation between imaging parameters

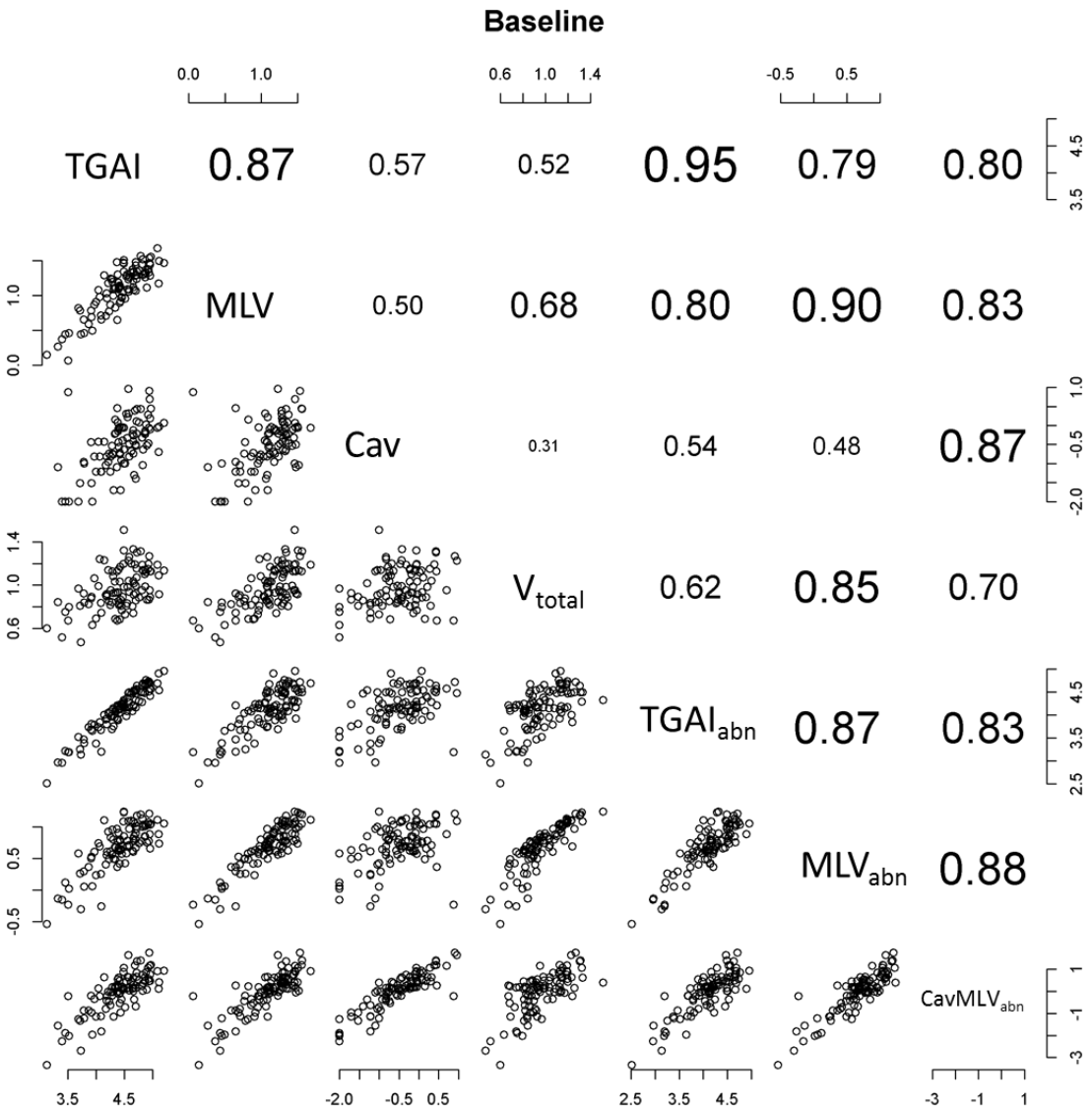


Figure 23 Correlations matrix of quantified scan parameters at Diagnosis.

Corresponding scatterplots left and below variables and correlation coefficient right and above.

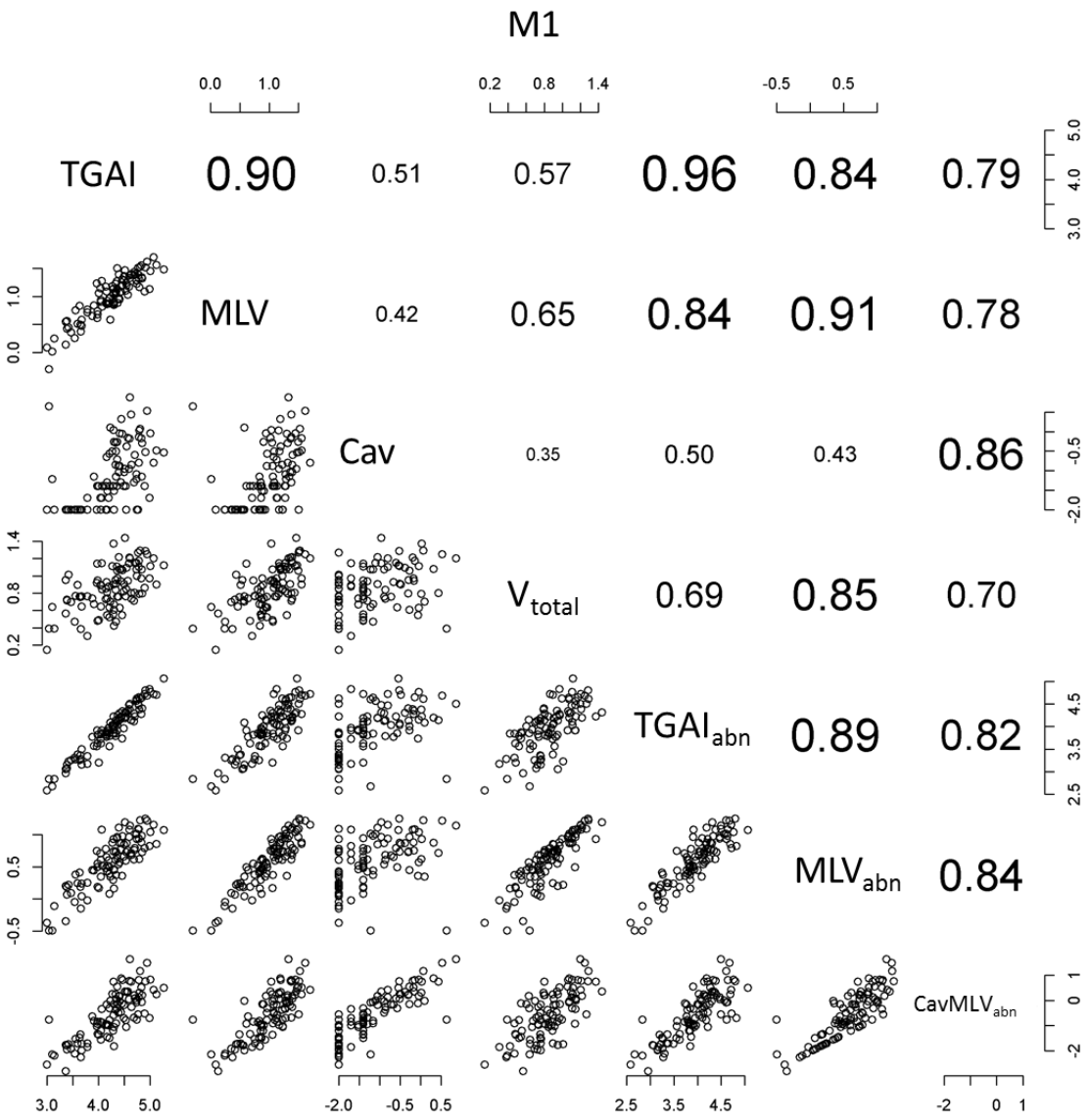


Figure 24 Correlation matrix of quantified scan parameters at M1.

Corresponding scatterplots left and below variables and correlation coefficient right and above.

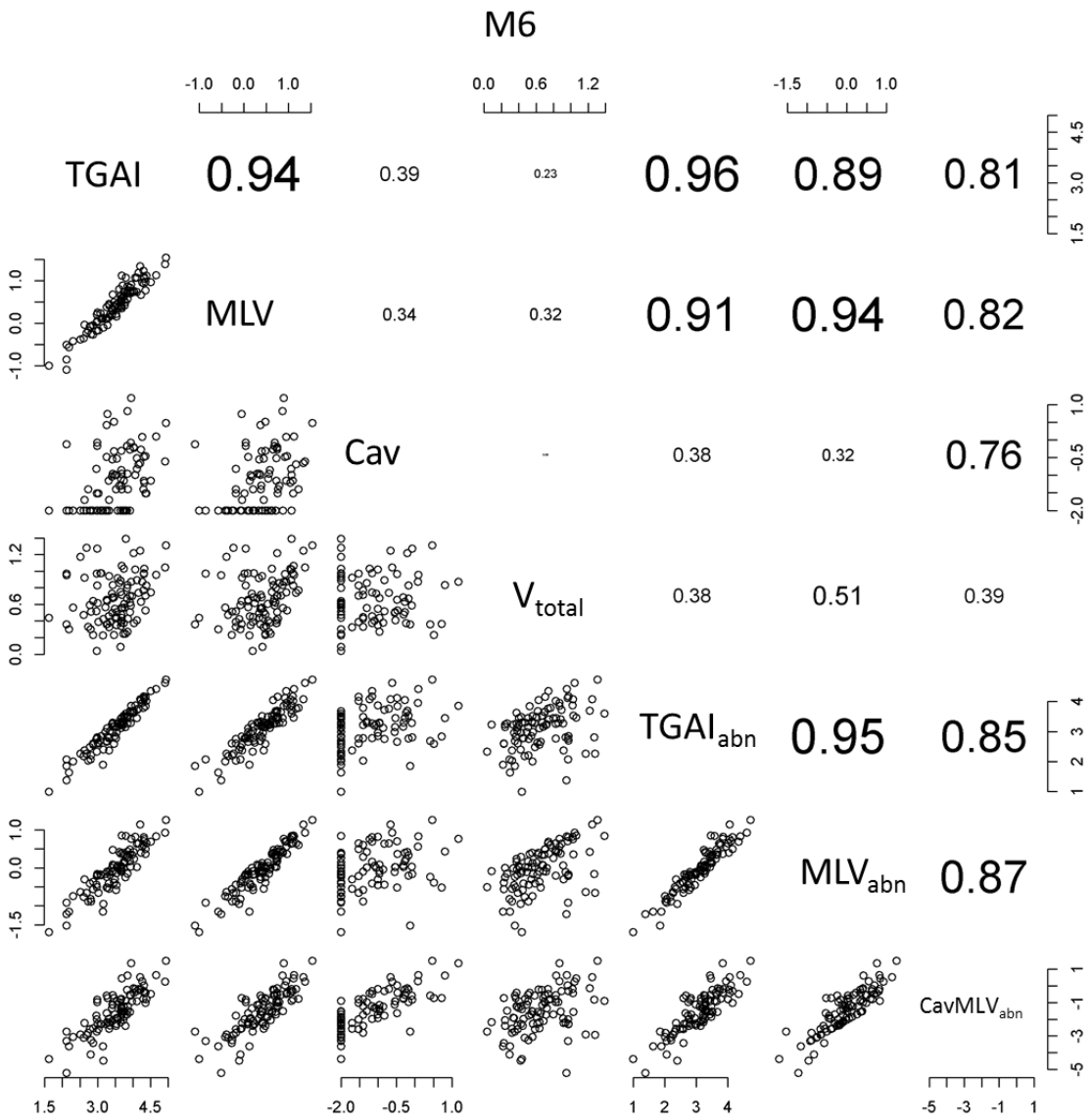


Figure 25 Correlation matrix of quantified scan parameters at M6.

Corresponding scatterplots left and below variables and correlation coefficient right and above.

We assembled correlation matrices of the principal scan parameters during treatment (Figure 23, Figure 24, Figure 25). TGAI and MLV were strongly correlated at all time-points. This was in keeping with the similar performance in PET TTN group analysis. V_{total} showed a moderate correlation with TGAI and slightly stronger correlation with MLV at Dx, but the relationships gets weaker towards M6. Cavity volume was the most independent variable at all time-points. Combined parameters (MLV_{abn} Weighted cavity volume/MLV_{abn}) showed strong correlation with related components, as expected.

Lung cavities on PET/CT has distinct characteristics when compared to other PET/CT parameters. While they form part of a pathological process and therefor represent diseased

lung, the lack of perfusion renders it devoid of FDG uptake. Thus, the potential for improved specificity by the addition of the PET component is lost. This also is responsible for the independence from other scan parameters. We tried to optimise the accuracy of cavity volume as measure of disease burden, by assigning the Z_{mean} from the rest of the lung to the cavity volume. The theory behind this was that if a low inflammatory response existed in the cavity wall and other lesions around the cavity, then the cavity is more likely to be inactive. Cavities serving as a reservoir of active MTB, on the other hand, would likely influence the surrounding tissue and lesions, in a pro-inflammatory manner. The resulting formula was:

Cavity volume (ml) x Z_{mean} = Total glycolytic activity index in cavity volume ($TGAI_{cav}$).

$TGAI_{cav}$ could also then be added to TGAI, to create a combined PET/CT parameter ($TGAI_{com}$) that include information from almost all PET/CT parameters without the need for more complex statistical conversions.

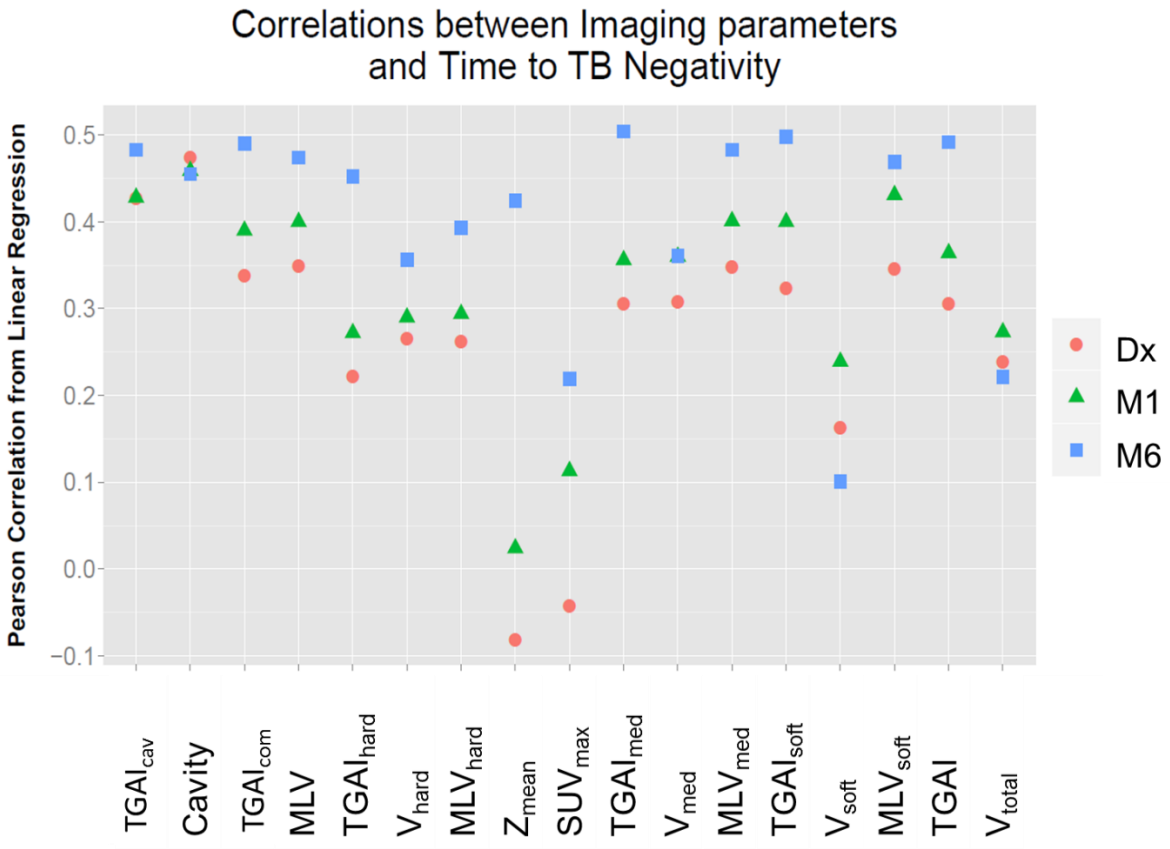


Figure 26 Correlation between values of imaging parameters and TTN

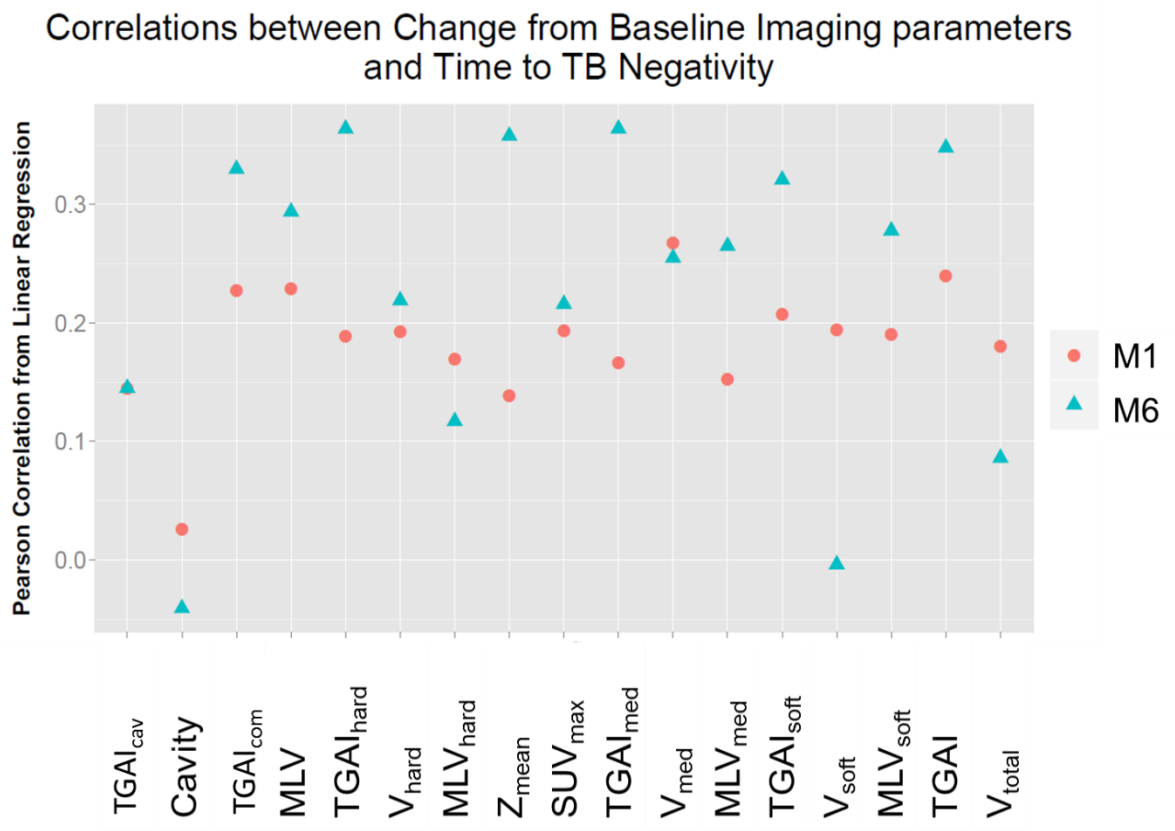


Figure 27 Correlation between imaging parameters' change from baseline and TTN

In Figure 26 and Figure 27 we compare the correlations between the various imaging parameters and TTN groups. Absolute values of cavity related parameters performed the most consistently across time-points – Cavity volume slightly better early in treatment and $TGAI_{cav}$ slightly better at M6. TGAI performed the most consistent across time-points in terms of change from baseline and $TGAI_{com}$ performed strongly in both analyses.

Prognostic factors of failure

As expected from TTN correlation results, Cavity volume was the best predictor of treatment failure.

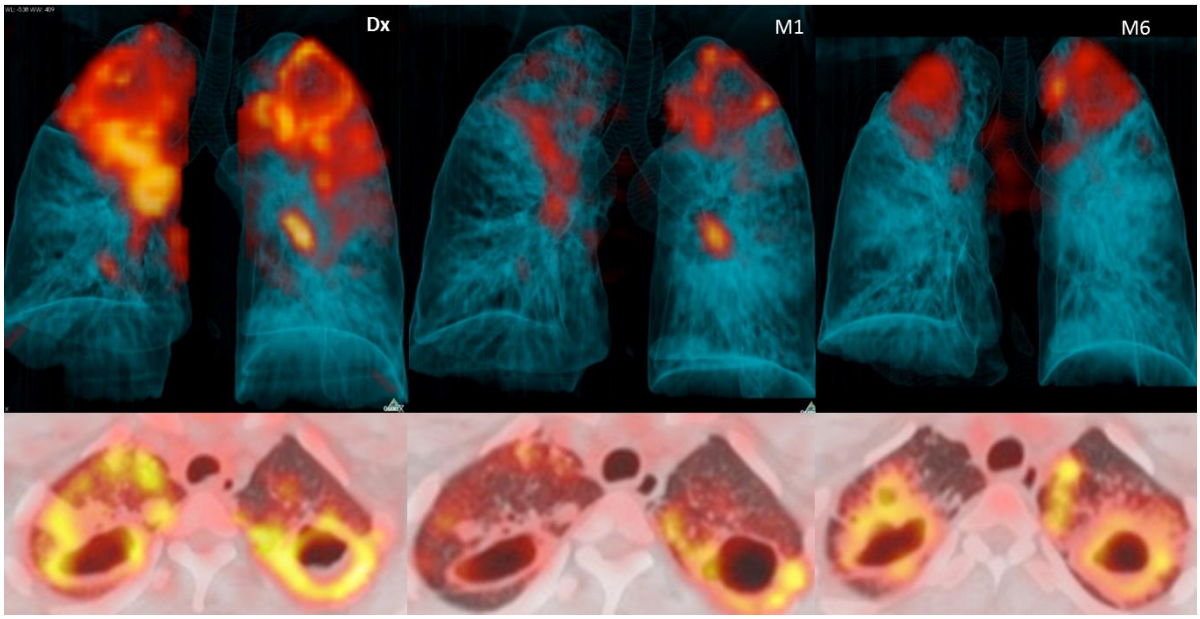


Figure 28 Example of PET/CT scan of failed treatment case at Dx, M1 and M6.

Large, bilateral upper lobe cavities with thick, dense walls with intense FDG-uptake can be seen throughout treatment.

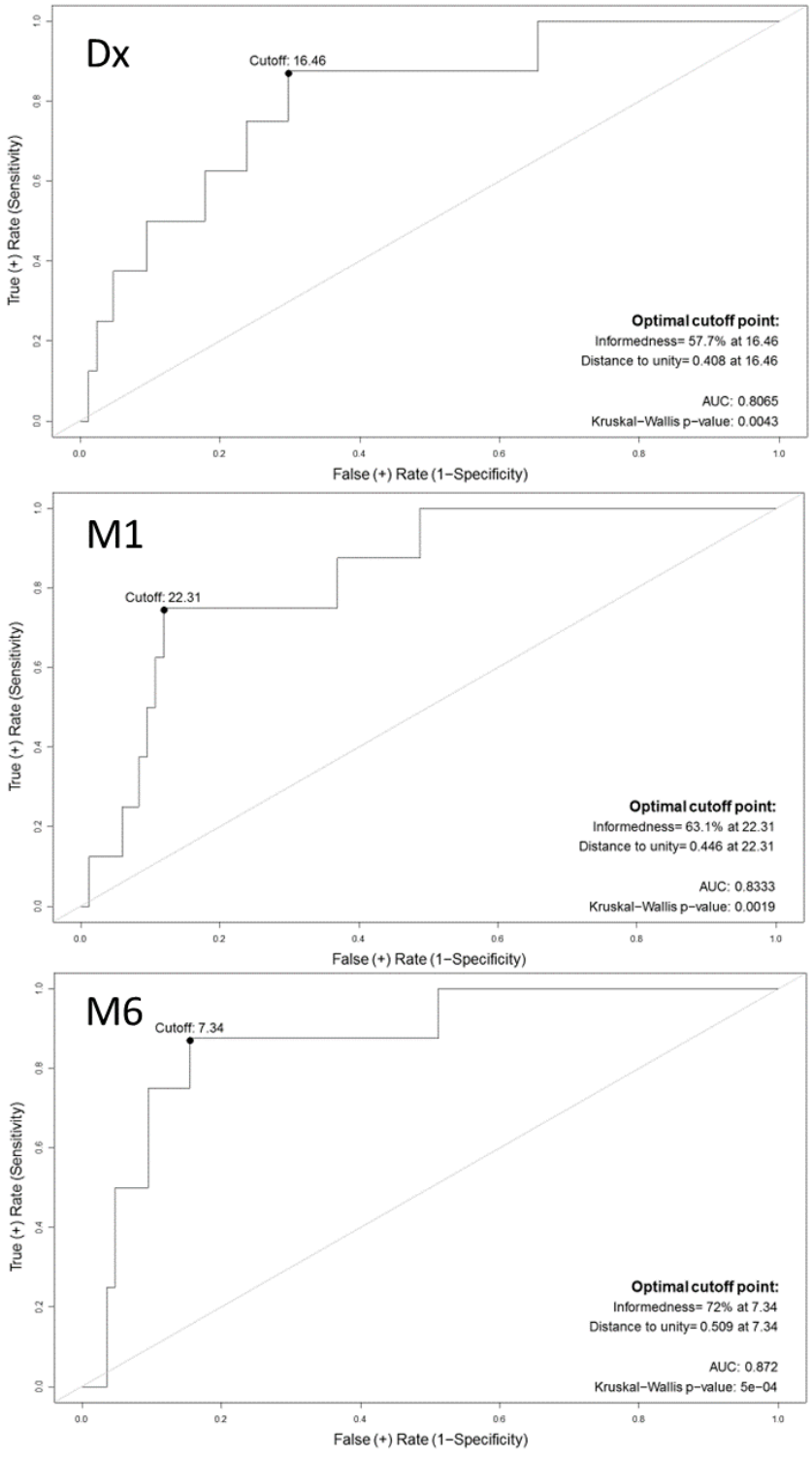


Figure 29 Receiver operating characteristic (ROC) curves of cavity volume predicting failed treatment.

Apart from cavity volume, other parameters reflecting the volumetric lesion extent also performed fairly well at predicting failure at Dx, while parameters reflecting glycolytic activity performed stronger at M6 (Figure *27*).

In addition to cavity volume, M6 cavity wall thickness also correlated to failure. For each M6 scan, we measured the wall of the cavity that appeared the thickest, at the level of the

maximum transverse cavity diameter. In most cured cases, M6 cavity volume ranged from zero (no cavity), to 3mm. Month 6 cavity wall thickness in failed cased were significantly different (Student's T-test for independent samples; $P < 0.001$) and ranged from 2.5mm to 7mm. Recurrent cases were not significantly different and M6 cavity wall thickness ranged from zero to 4mm. The distribution for the groups is shown in Figure 30.

Measurement of cavity walls required a degree of interpretation due to irregular borders, confluence with other lesions and structures, as well as fibrotic changes (examples shown in Figure 31). Therefor it likely to be subject to inter-reader variability.

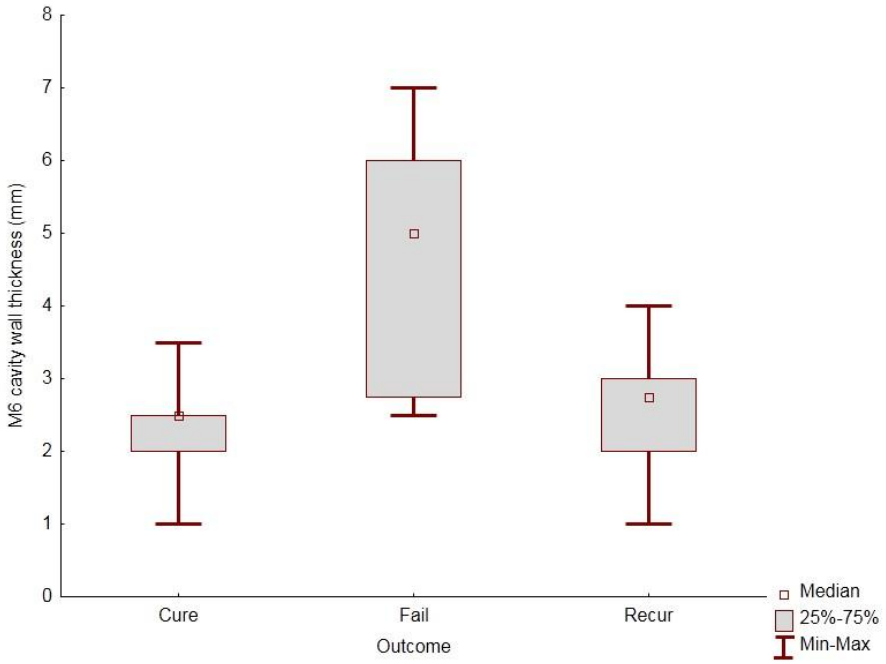


Figure 30 Box and whisker plot of M6 cavity wall thickness, showing the median, 25th & 75th percentile (Box) and range (whiskers).

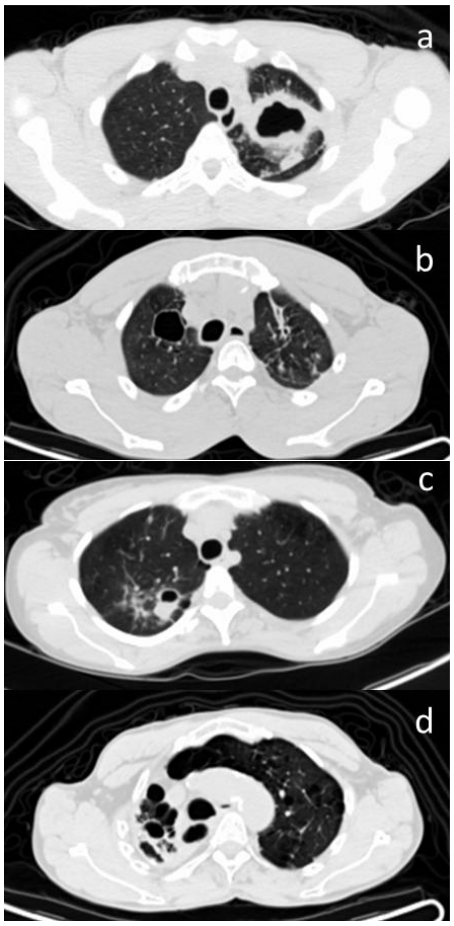


Figure 31 Examples of cavity lesions on M6 CT images.

a) A failed treatment cases with a well-circumscribed cavity with a 6 mm wall. b) Cavity with a 1.5mm wall on the CT images of a cured patient. c) A cured patient with cavity and confluent nodule. d) A cured patient with extensive right upper lobe fibrotic changes and destruction.

Prognostic factors of recurrence

The twelve patients diagnosed with recurrent disease within 2 years after treatment did not show a comparatively large lesion burden at Dx (Figure 19), but did show a relative slow rate of improvement. The mean TGAI for the recurrent group improved at a similar rate to patients with a TTN of 24 weeks from Dx to M1 and M6, even though the median TTN for recurrent cases were 8 weeks (

Table 2).

A similar trend was seen for V_{total} and Z_{mean}, but was less apparent. Mean cavity volume decreased over the first month of treatment, but deteriorated between M1 and M6.

At M1, there was a trend for the recurrent group to have a smaller reduction in TGAI and $TGAI_{com}$ burden. At M6, the difference between the groups were significant (Figure 32 & Figure 33). No other parameters were significantly different between cured and recurrent groups.

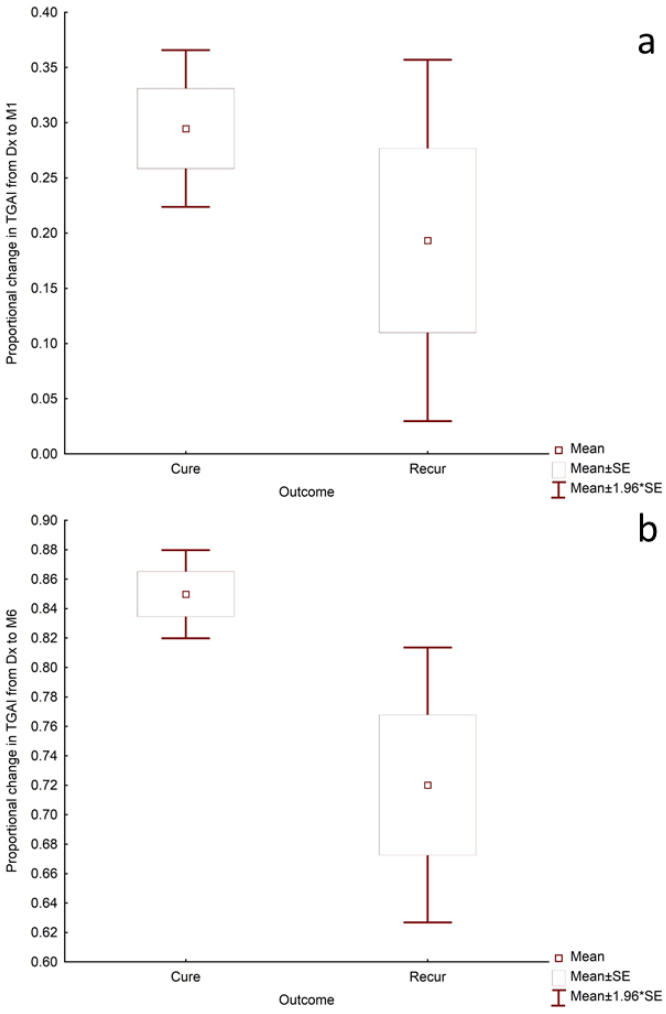


Figure 32 Mean ± SE ± 1.96xSE Change in TGAI for cured and recurrent groups

a) From Dx to M1 (p = 0.3) b) from Dx to M6 (p = 0.003).

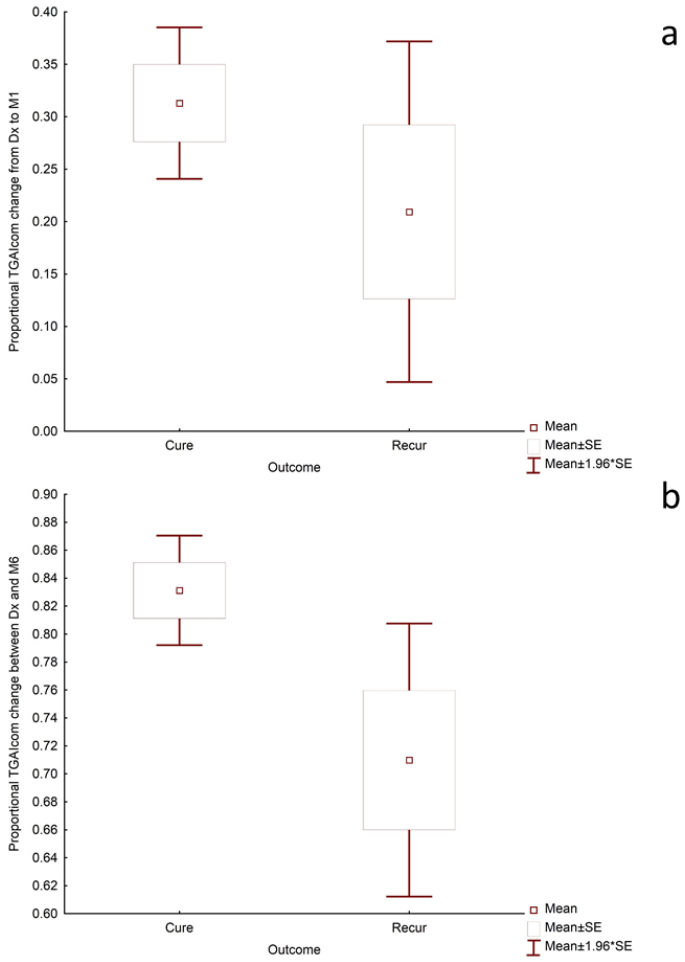


Figure 33 Mean ± SE ± 1.96 x SE Change in $TGAI_{com}$ for cured and recurrent groups

a) From Dx to M1 (p = 0.3) and b) from Dx to M6 (p = 0.02).

Criteria for unfavourable outcomes

We then pooled patients with unfavourable outcomes (failed and recurrent treatment) and performed further analysis to find the best criteria to predict recurrent disease. We analysed the most promising scan parameters' distribution per cured and recurrent groups, combined with ROC curve analysis to determine the most informed thresholds.

Table 23 Summary of Contingency table statistics for scan parameters.

Ranked according to relative risk of unfavourable outcome. Fisher exact test was performed to determine significance. PPV (positive predictive value), NPV (Negative predictive value), Change (change from baseline), intensified (at least one intensified or new lesion).

Parameter	P - value	Relative risk	Sensitivity	Specificity	PPV	NPV	Criteria	No. that met criteria (n = 96) Cured	Fail	Recur
TGAI Change M6	**<0.0001**	6.97	0.80	0.75	0.46	0.94	< 80%	19	7	9
TGAIcom Change M6	**<0.0001**	6.67	0.80	0.74	0.44	0.93	< 80%	20	7	9
Cavity M6	<0.001	4.36	0.55	0.86	0.52	0.88	> 7mm^3	10	7	4
Cav Change M6	**<0.01**	4.30	0.65	0.79	0.45	0.90	< 60%	16	6	7
TGAIcom3	**<0.001**	4.05	0.50	0.88	0.52	0.87	> 1000	9	6	4
TGAI M6	<0.001	4.05	0.50	0.88	0.53	0.87	> 600	9	6	4
Mixed response M6	**0.02**	2.86	0.60	0.72	0.36	0.87	Intensified	21	8	4
Cav Change M1	**0.01**	2.80	0.50	0.80	0.40	0.86	< 33%	15	5	5
Cavity M1	**0.04**	2.50	0.35	0.87	0.41	0.84	>20mm^3	10	6	1
TGAI Change M1	0.07	2.24	0.40	0.82	0.36	0.84	<5%	14	4	4
TGAIcom Change M1	0.16	1.80	0.40	0.76	0.31	0.83	15%	18	3	5
Cavity Dx	0.18	1.72	0.45	0.72	0.29	0.83	> 16.5mm^3	22	6	3

Less than 80% reduction in total glycolytic activity from Dx to M6 was the best predictor of unfavourable outcomes with an almost 7-fold risk if these criteria was met. Adding the assigned cavity activity ($TGAI_{cav}$) to generate $TGAI_{com}$, did not improve accuracy, despite cavity volume improvement of less than 60% from baseline also performing well as an independent predictor of unfavourable outcomes.

A total M6 Cavity volume greater than 7mm^3 and a M6 TGAI of greater than 1000 also carried an increased risk to have an unfavourable outcome of more than 4 times. Cavity volume was slightly more sensitive and TGAI more specific. Combining the variables did not improve accuracy. $TGAI_{com}$ performed identical to TGAI. We also applied criteria in which *either* M6 cavity volume > 7mm^3, *or* TGAI > 1000 put a patient in the high-risk group, but this generated the same sensitivity and lower specificity.

All these M6 quantified parameters performed better than the lesion based qualitative measure. A mixed response pattern at M6 was associated with a 2.86 times increased risk, which was comparable to the best quantified parameters at M1.

At month 1, both a cavity volume greater than 20mm^3 and a lack of at least 33% cavity volume reduction from Dx were significantly associated with a poor outcome – showing a 2.5 and 2.8 times increased risk of unfavourable outcome respectively. A trend was found for high cavity volume at Dx to be associated with unfavourable outcomes – largely driven by a strong association to failed treatment cases.

A trend was found that less than 5% reduction of TGAI at from Dx to M1 was associated with an unfavourable outcome. Due to recurrent cases presenting with a lower mean TGAI burden at Dx than cured patients, total TGAI values at Dx and M1 did not perform well as a predictor of pooled unfavourable outcomes.

5.3 Quantified scan results: EOT + 1y.

Residual lesions

As discussed in section 3.4 (page 38) 50 patients were scanned again 1 year after the end of treatment. Forty-two of these patients maintained cure, while eight were diagnosed with recurrent disease (Figure 11). At EOT + 1y, only 32% of EOT + 1y scans were completely resolved. Significant residual lesion(s) were found in 68%, of which half had improvement of all lesions and the other 34% a mixed lesion response compared to the M6 scan (Page 38). However, there was improvement in size and intensity of most residual lesions when compared M6.

When compared to Dx, only nine cases (five of which were recurrences) showed less than 80% improvement of TGAI by EOT + 1 and only three cases (2 of which recurrences) had more than 7mm^3 total cavity volume. The distribution of TGAI and cavity volume across time-points are shown in Figure 34 and Figure 35.

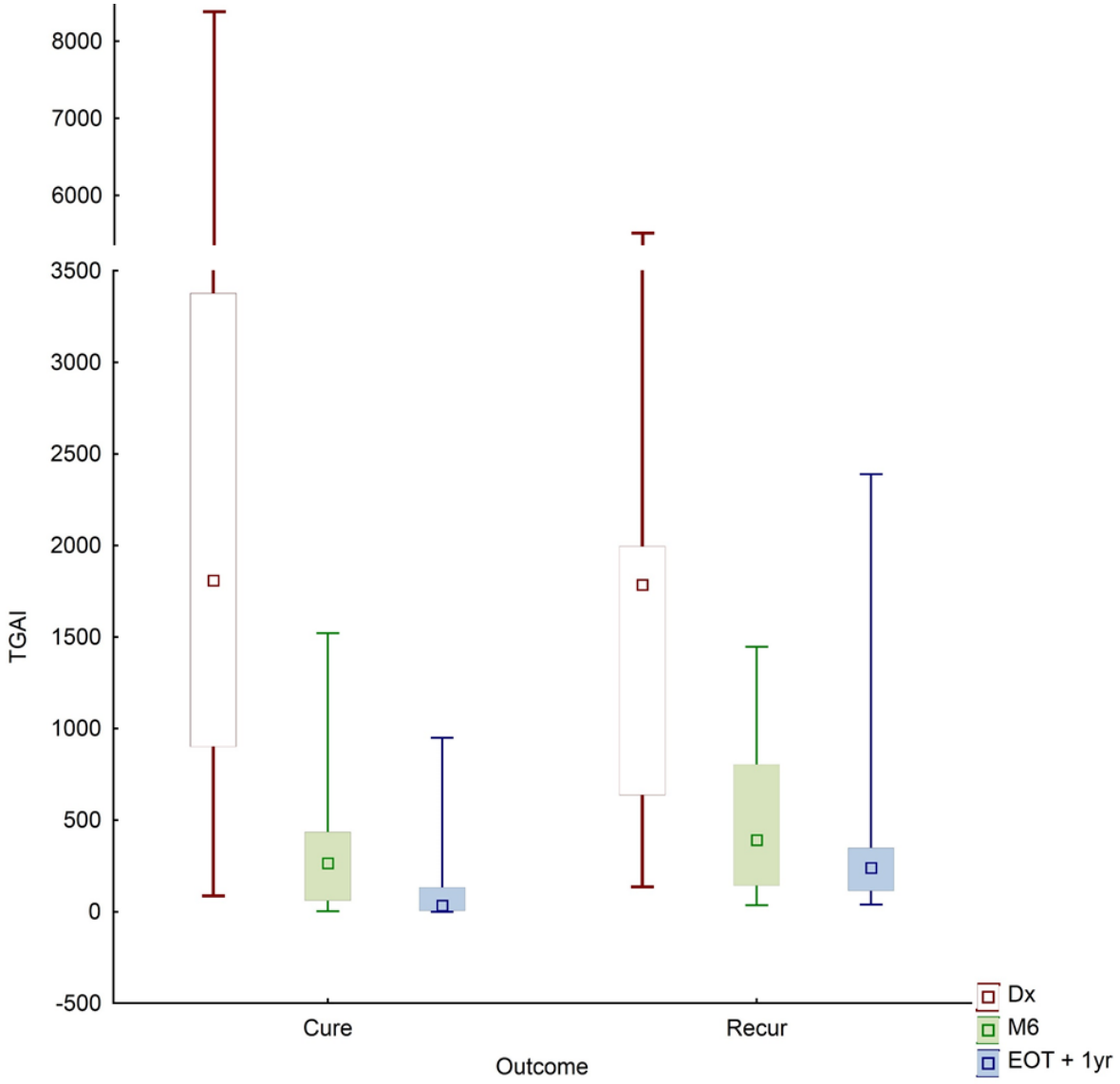


Figure 34 Box and whisker plot showing median, quartiles and range for TGAI at Dx, M6 and EOT + 1y.

Grouped in Cured and recurrent patients. Y-axis truncated.

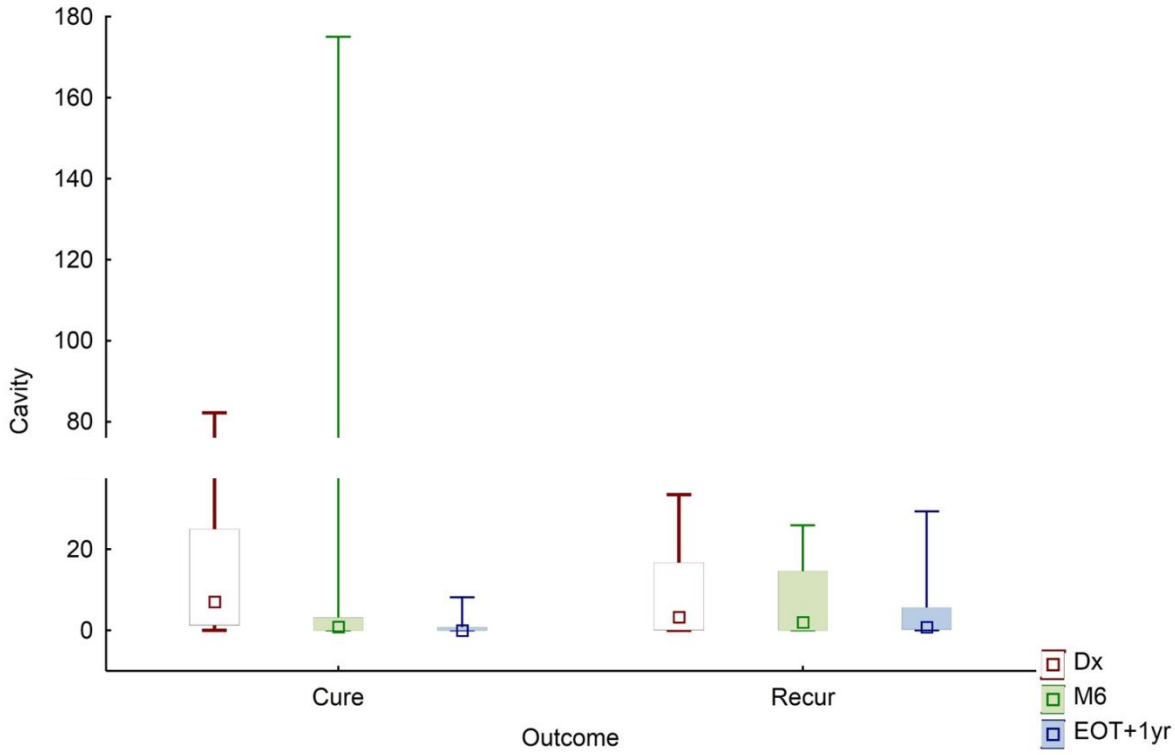


Figure 35 Box and whisker plot for Cavity volume at Dx, M6 and EOT + 1y.

Grouped in Cured and recurrent patients, showing median, quartiles and range. *Y-axis truncated.*

MLV decreased to a median of 0.69% of lung volume at EOT + 1y. V_{total} showed more residual lesion volume, with a median of 3.58% of total lung volume at EOT + 1y.

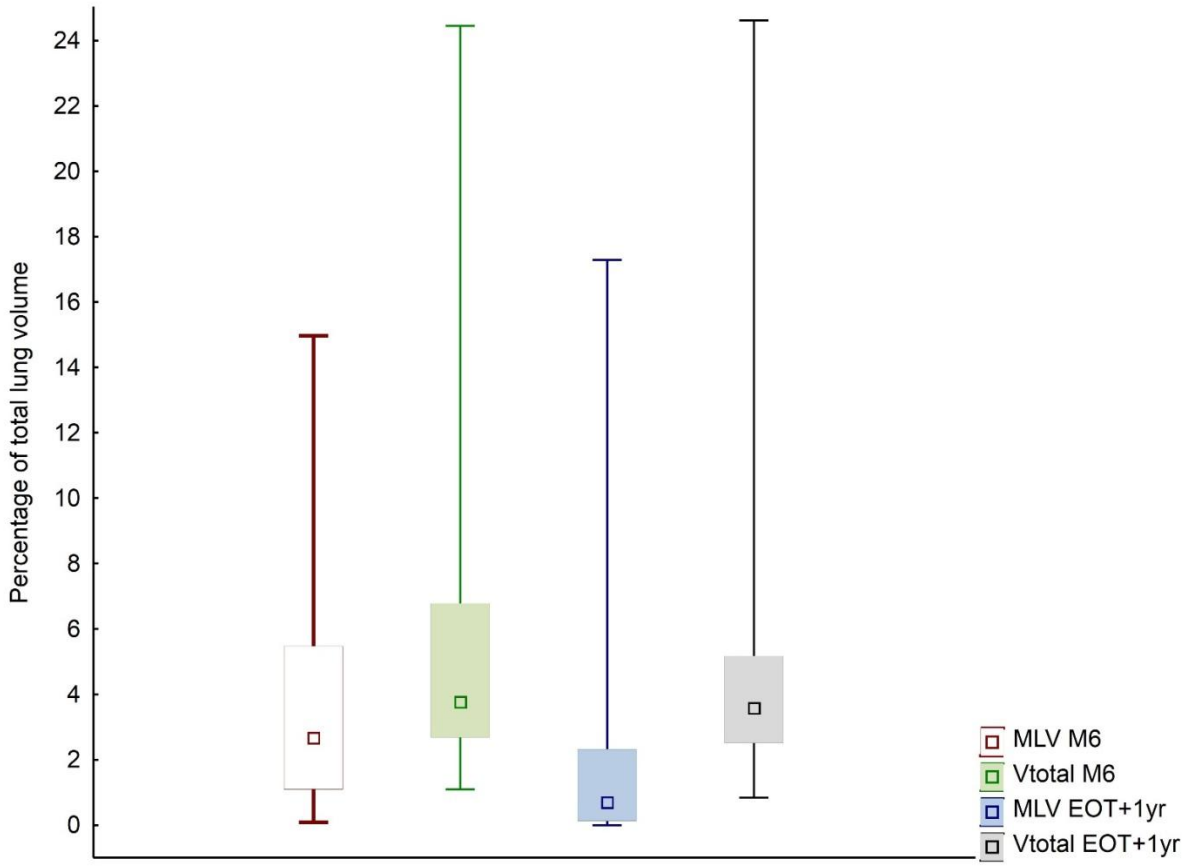


Figure 36 Box and whisker plot for MLV and V_{total} at Dx, M6 and EOT + 1y.

Showing median, quartiles and range.

For recurrent cases, the range of glycolytic activity and cavity volume, were not as high at EOT + 1y as we found at original diagnosis. Two factors would likely influence these findings. Firstly, three of the patients already received partial re-treatment. Secondly, the fact that the patients were actively followed up, may have decreased the time from reactivation or reinfection until diagnosis.

Correlations between M6 and EOT + 1 year findings.

There were no significant association between a mixed response at M6 and a mixed response at EOT + 1 year (*Table 24*). Figure 37 and Figure 10 shows examples of dynamic lesion progression and resolution during and after treatment.

Table 24 Association between mixed scan response patterns at M6 and EOT + 1y.

Fisher exact test p = 0.18.

	M6 Mixed	M6 Resolved/Improved	Totals
EOT + 1y Mixed	7	10	17
EOT + 1y Resolved/Improved	8	25	33
Totals	15	35	50

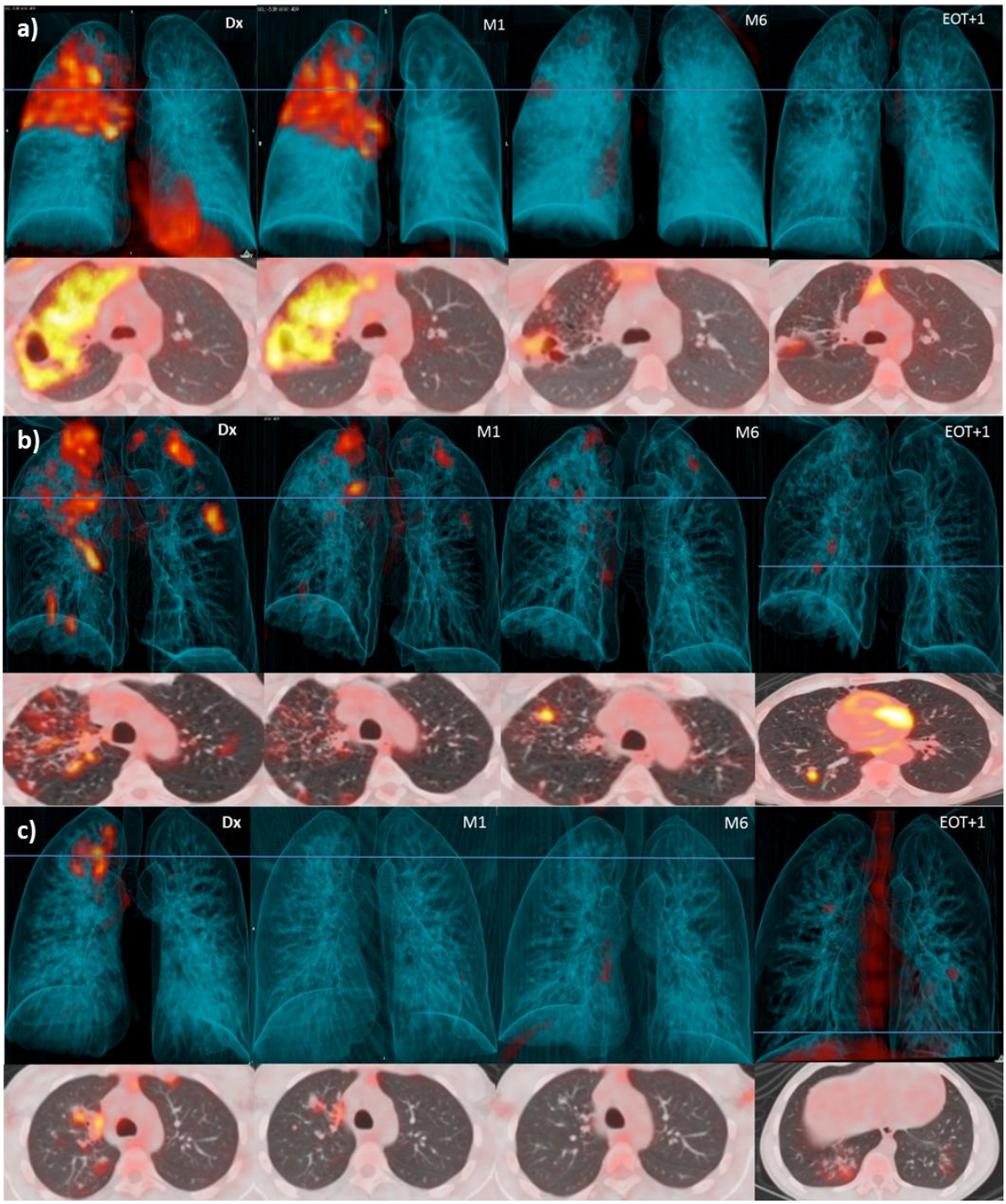


Figure 37 Dx, M1, M6 and EOT + 1y PET/CTs for 3 representative cases.

3-dimensional anterior and transverse slices at the level of horizontal blue line. a) Residual cavitation with moderate FDG-avidity at M6, improves over the next year, leaving nodular infiltrate with mild activity. b) New nodule with high intensity seen at M6 resolves, but 2 new nodules are seen at EOT + 1y. c) All lesions resolved at M6, but 3 new areas with small nodular and tree-in-bud infiltrates seen. All 3 patients received 6 months of standard treatment and maintained cure.

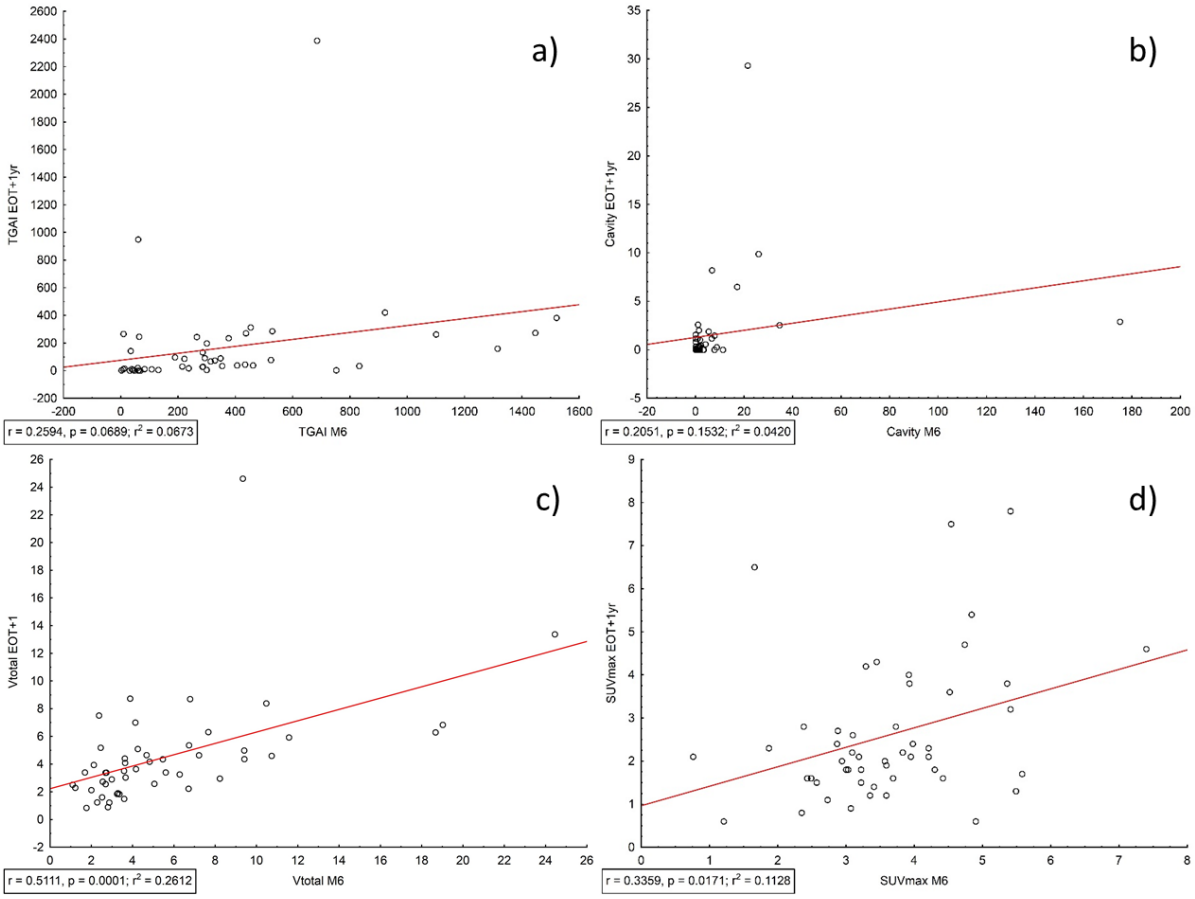


Figure 38 Scatterplots to show correlation of PET/CT scan parameters at M6 and EOT + 1y.

a) TGAI, weak positive correlation, non-significant. b) Cavitation volume, no significant correlation. c) V_{total}, significant positive correlation. d) SUVmax, moderate, but significant positive correlation.

No significant association was found between M6 and EOT + 1y for TGAI or cavity volume [Figure 38 a) & b)]. However, we found a moderate correlation between the time-points for V_{total} and SUVmax.

<u>Differences between recurrence and maintained cure.</u>

Applying two-tailed Student's T-tests for independent samples, we found significant differences at EOT + 1y between cured and recurrent patients for TGAI ($P < 0.01$), Cavity volume ($P < 0.01$), MLV ($P = 0.01$) and SUVmax ($p = 0.04$), but not for Z_{mean}, V_{total} or CT density subgroups (V_{soft}, V_{medium}, V_{hard}).

Since there was no consistent spacing between the timing of the scan and re-treatment initiation, analysis to define an EOT + 1y criteria to differentiate maintained cure from recurrence was not deemed appropriate.

5.4 Statistical associations with quantified PET/CT parameters.

Association with Microbiological markers

We performed student's T-tests, with analysis of variance and homogeneity to test for significant associations between Xpert positivity at M6 and scan parameters.

Table 25 Student's T-test results for difference in M6 PET/CT parameters and M6 Xpert positivity.

	M6 Xpert Neg mean	M6 Xpert pos mean	T-Test P-value	Levene P-value	Mann-Whitney U-test P	Xpert pos (n=)	Xpert neg (n=)
TGAI M6	278.41	964.56	**< 0.001**	**< 0.001**	**< 0.001**	68	30
TGAI Change (Dx-M6)	0.84	0.69	**0.001**	**0.001**	**< 0.001**	68	30
Vtotal3	1.70	2.72	**0.039**	**0.011**	**0.04**	68	30
Cavity3	10.31	11.85	0.846	0.515	**0.001**	68	30
Cav Change (Dx-M6)	0.60	0.60	0.969	**0.082**	**0.01**	68	30
TGAIcom M6	434.87	1190.12	**0.003**	**0.004**	**< 0.001**	68	30
TGAIcom Change (Dx-M6)	0.82	0.69	**0.004**	**0.033**	**0.002**	68	30

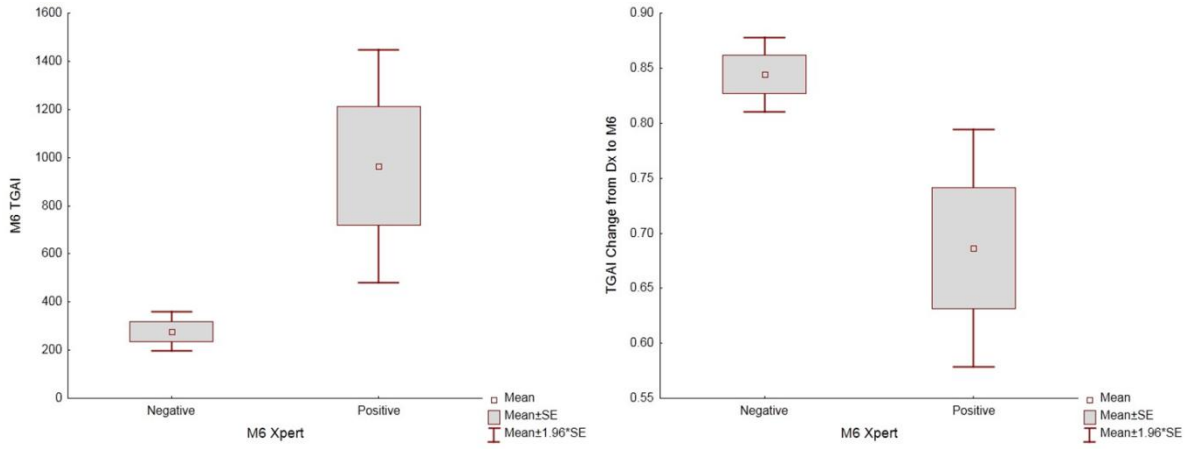


Figure 39 Association between M6 Xpert positivity, M6 TGAI and TGAI change from Dx.

Box and whisker plot showing Mean, ±SE, ±1.96×SE.

At M6 we found the clearest difference between Xpert negative and positive groups for M6 TGAI and proportional change in TGAI from Dx to M6. $TGAI_{com}$ and V_{total} also showed significant differences. Xpert positivity at M6 was associated with less improvement during treatment and a higher lesion burden at the end of treatment. We also found a linear correlation between Xpert Cycle threshold values at M6 and M6 TGAI ($r = -0.43$; $p < 0.0001$) and TGAI change from baseline ($r = 0.43$, $p < 0.0001$). Using the non-parametric Mann-Whitney U-test (likely more accurate due to the positive skew in cavity volume distribution), we also found a significant difference for cavity volume or cavity change from Dx–M6.

We performed the same analysis to test for an association between the main scan parameters and m6 sputum MTB mRNA positivity and found no significant associations. We also found no linear correlation between the number of positive MTB transcript targets and the quantified scan values.

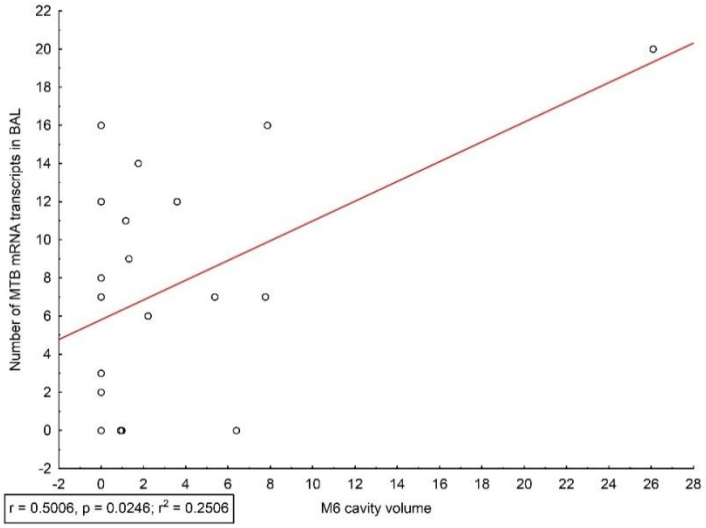


Figure 40 Correlation between the number of MTB mRNA transcripts detected in EOT BAL and M6 cavity volume.

We found a linear correlation between cavitation volume and the number of positive MTB transcript detected in EOT bronchoalveolar lavage fluid (Figure 40), but no other scan parameters.

At Dx, we found a non-significant negative correlation between days to liquid culture positivity (TTP) and cavity volume ($r = -0.2$ $p = 0.05$), and a significant negative correlation with TGAI ($r = -0.27$, $p = 0.008$), which was slightly strengthened by adding the influence of cavitation to create $TGAI_{com}$ (Figure 41)

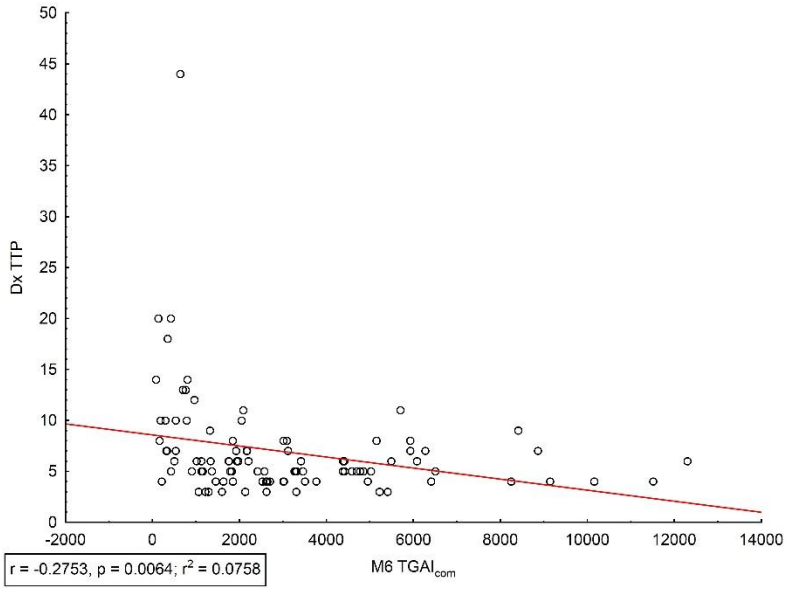


Figure 41 Scatterplot of correlation between TTP and $TGAI_{com}$.

To look for different response in different strains, we collapsed the strain clades into those with the Beijing strain and other strains and performed two-tailed, T-tests for independent samples. The Beijing strain was grouped since this strain was the most prevalent in our cohort and have previously been linked to an increased risk for treatment failure.[41,42]

Table 26 T-test results testing for difference in M6 PET/CT parameters and grouped MTB strains.

	Mean Beijing	Mean Other	T-test P - value	Levene p - Value	Mann-Whitney U-test P	Beijing (n =)	Other (n =)
TGAI Dx	2409.53	2856.10	0.32	0.01	0.82	51	38
TGAI M6	513.42	537.09	0.90	0.19	**0.04**	51	38
TGAI Change (Dx– M1)	0.24	0.28	0.64	0.67	0.03	51	38
TGAI Change (Dx–M6)	0.77	0.82	0.32	0.62	**0.03**	51	38
Cavity1	18.84	23.11	0.55	0.44	0.87	51	38
Cavity3	14.33	7.97	0.43	0.23	**0.05**	51	38
Cav Change (Dx–M1)	0.38	0.68	**0.0004**	0.0005	**0.001**	51	38
Cav Change (Dx–M6)	0.48	0.82	**0.02**	0.001	**0.01**	51	38
Vtotal1	5.53	5.85	0.63	0.03	0.99	51	38
Vtotal3	1.93	2.35	0.42	0.04	0.73	51	38
Vtotal Change (Dx–M6)	0.66	0.64	0.68	0.01	0.48	51	38

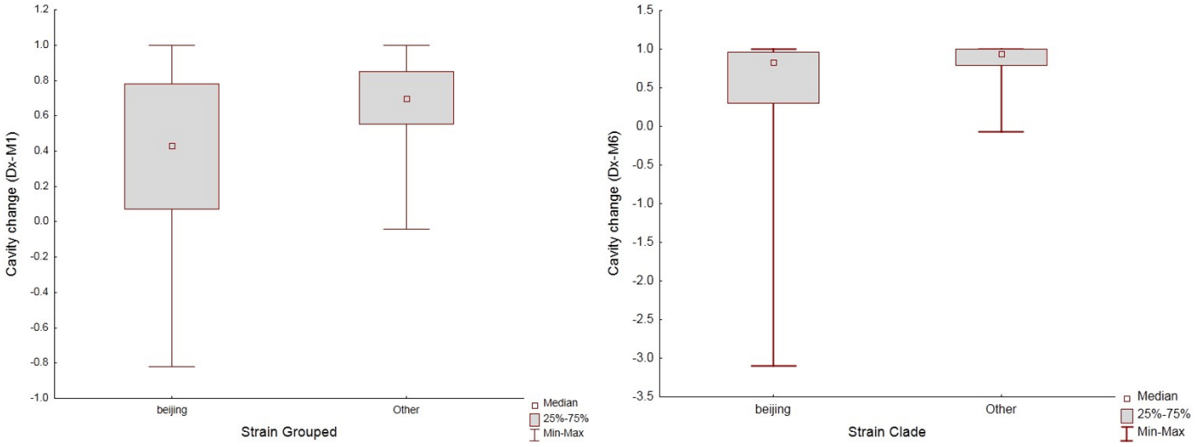


Figure 42 Association between grouped strain clades and change in cavity volume from Dx to M1&M6.

Box and whisker plot showing median, quartiles and range .

We found no difference between the groups at Dx. The group infected with Beijing strains, however, did have a significantly less improvement in proportional cavity volume from baseline to M1 and M6 (Table 26; Figure 42). This resulted in a trend for higher cavity volume at M6 in the Beijing group.

Association with clinical factors

We performed student's T-tests, with analysis of variance and homogeneity to test for significant associations between a history of previous PTB episode(s) and scan parameters.

Table 27 T-test results testing for difference in M6 PET/CT parameters and history of previous PTB.

	1st episode Mean	Previous PTB mean	T-test P-value	Levene P-Value	Mann-Whitney U-test P	1st episode (n =)	Previous PTB (n =)
TGAI Dx	2279.98	3054.43	0.083	0.033	0.83	68	31
TGAI M6	302.07	891.25	**0.001**	**<0.001**	**0.047**	68	31
TGAI Change (Dx–M6)	0.84	0.70	**0.003**	**0.003**	**0.029**	68	31
Cavity1	18.04	22.41	0.526	0.798	0.87	68	31
Cavity3	11.01	9.96	0.893	0.479	0.057	68	31
Cav Change (Dx–M6)	0.65	0.45	0.211	0.209	**0.01**	64	31
TGAIcom (Dx–M6)	472.43	1079.44	**0.016**	**0.005**	0.058	68	31
TGAIcom Change (Dx–M6)	0.82	0.71	**0.025**	0.058	**0.02**	68	31
Vtotal Dx	5.41	5.95	0.428	0.413	0.99	68	31
Vtotal M6	1.83	2.38	0.266	0.725	0.48	68	31

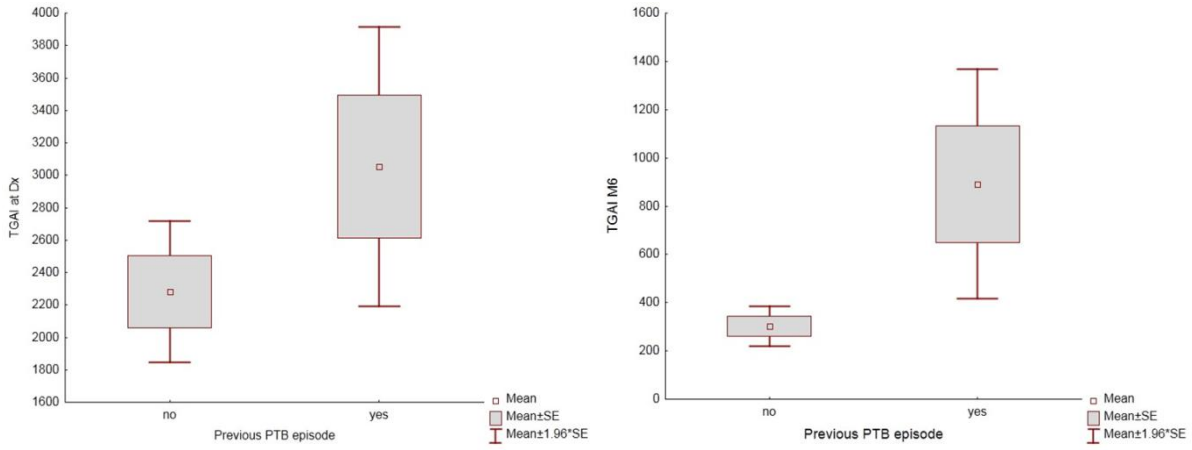


Figure 43 Association with between a history of previous PTB episode(s) and TGAI at Dx and M6. Box and whisker plot showing Mean, ±SE, ±1.96×SE.

There was a trend for patients with previous PTB to have a higher TGAI burden at Dx and a significant association between higher TGAI at M6 and less improvement from Dx to M6 and previous PTB episodes. We found no significant difference for cavity volume or CT density areas. Using non-parametric testing (Mann-Whitney U-test), however, we found significantly less reduction in cavity volume in patients with previous PTB.

We tested for a significant association between documented treatment dosages missed during follow-up and M6 scan parameters, by doing Student's T-Tests, checking for linear correlations and found no significant association. We also performed a Fisher Exact on the cross-tabulation of patients that missed more than 80% of treatment and those with less than 80% TGAI reduction from Dx to M6 and found no significant association.

We also did not find a significant difference between quantified PET/CT parameters and a history of smoking or the PTB symptom score.

6. Discussion: Operator independent technique to quantify widespread complex tuberculous lung lesions on PET/CT.

6.1 The use of the quantification technique

Summary

We devised a quantification technique to reproducably quantify pulmonary tuberculosis lesions on PET and CT. This method is described in section 2.6, page 19 and details of the application provided in section 5.1, page 59. We used patient specific reference volumes to reduce intra- and interscan variability on PET and auto-delineation to reduce inter-reader and inter-operator variability when assessing multiple widely distributed lesions with complex morphology. The technique quantifies lesions throughout the lungs to measure the central trend of disease progression or resolution, thus accounting for the variable response of individual lesions during PTB treatment response.

Optimal adaptive thresholds for auto-segmentation of PET scans were determined by using a small number of control lung scans and scans containing lesions with minimal to mildly increased uptake. Density thresholds for CT scans were determined using values previously reported in literature. We tested the technique in a pilot set of control cases and in PTB cases during treatment. The quantified variables correlated well with manual visual scan interpretation and clinical outcomes.

Novel aspects of quantification technique

The technique introduced some novel concepts in PET/CT analysis, including the use of a patient specific reference volume to standardardise the lung uptake; whole lung automated segmentation of CT lesions; and bivariate quantification of PET and CT images. It also re-applied concepts previously used in other settings, such as co-registration of corresponding scans at different time-points, whole organ automated segmentation of PET uptake and joint-segmentation of PET and CT.[86]

Strengths and limitations of technique

Some challenges and potential drawbacks were encountered during the application.

Using the 3D region grow tool in MRICro was user friendly and easily reproducable, however, some manual imput was required when drawing the lung mask, to include dense lesions extending into the chest wall, and areas affected by misregistration. This may be more time-consuming and could lead to some inter-user variability, however it should lower than manual delineation of each individual lesion. Although the segmented volumes were

visually compared to the source scans, we were not able to compare directly to an objective ground truth, due to lack of available histology or gold standard segmentation techniques for complex lesions. Testing the technique using a phantom scan may be limited, since the results may not be accurate when derived from a scan with an artificially homogenous background. Comparing to a ground truth in future analysis may however be useful, but given that the main goal of this technique is to reduce interscan variability in longitudinal follow-up and not planning surgical interventions or radiation therapy, this is likely to be less crucial.

Co-registering baseline and follow-up scans allowed the user to generate a single lung mask and reference volume to use across all time points. This should reduce interscan variability between follow-up scans and save time. It did not, however, allow the measurement of change in total lung volume over time, which may occur during lesion resolution with associated fibrosis formation. Re-slicing allows direct comparison of PET and CT components using a single VOI. A disadvantage is that there is a smoothing effect on the CT images.

Using the mean and standard deviation of 2 reference volumes to standardise lesion to background activity should notably decrease inter-scan and inter-patient variability compared to techniques that normalise the uptake to reference volumes from other organs or a theoretical whole body concentration.

The technique was useful to report trends, but a possible drawback is that Z-scores are a statistical measure not routinely used in reporting, which may require additional conversions when applying research findings to a clinical setting.

Auto-delineation of the whole lung allowed for the segmentation of multiple lesions with widespread distribution and variable intensity, size and morphology. This is expected to greatly reduce inter-reader variability, especially in a self-controlled study measuring changes after interventions. The technique was not specific enough to delineate cavity volume, and an additional manual step was required to perform this function.

Conclusion

This technique to auto-segment and quantify multi-focal and complex lung lesions of PET and CT showed great promise in a pilot set of subjects, and required limited operator input. Further application to a larger number of subjects were of value to further validate the measures against clinical and research outcomes. Ultimately this methodology, or some of the features, could easily be incorporated into standard clinical analysis programs and reporting protocols.

6.2 Quantified PET/CT results.

Summary of main findings

We successfully implemented the automated quantification technique on all PET/CT scans of patients included in the analysis.

We found a good correlation between quantified PET/CT scan characteristics and micro-biological measurements, such as time to culture negativity, days to liquid culture positivity at diagnosis and Xpert positivity at M6. We also found a good correlation between the scan parameters and treatment outcomes, as well as a history of previous PTB episode(s). While even most cured cases did not manage complete resolution of lesion burden, quantification resulted in much improved prognostic indicators of treatment outcome.

A summary of suggested thresholds for most prominent scan parameters at different time-points and the relative risk, sensitivity, specificity, negative predictive value and positive predictive value are shown in Table 23.

To summarise findings relating to cavity volume: 1) Cavity volume had the strongest correlation with sputum culture positivity and TTN. 2) It showed the best predictive ability at early time-points, with an AUC of > 80% at Dx, M1 and M6 to predict treatment failure. A M6 cavity volume greater than 7mm^3, or a reduction less than 60% was associated with an unfavourable outcome. 3) Cavity volume was the only parameter that showed a significant correlation with mRNA detection in BAL fluid and a non-significant association with mRNA in sputum detection. 4) A lack in cavity volume reduction from diagnosis was a significant risk factor for recurrent disease. 5) Infection with the Beijing strain was associated with a lack of reduction in cavity volume from Dx.

To summarise findings related to total glycolytic activity index: 1) A suboptimal reduction in TGAI was the greatest risk factor for unfavourable outcomes. At least 5% reduction in TGAI during the first month of treatment and at least 80% reduction by the end of treatment should be expected during a favourable response. 2) The change in TGAI from Dx to M6 also had the strongest association with Xpert positivity at M6. 3) TGAI showed the strongest correlation to sputum culture TTP at Dx. 4) In patients with previous episodes of PTB, we found a suboptimal reduction in TGAI from baseline, which resulting in a higher TGAI load at M6.

The volume of high-density lesions on CT (V_{total}) showed a relatively good correlation with TTN and failed treatment, even in early treatment, with Dx to M1 change in V_{medium} showing good correlation. Later in treatment, however, density changes were less dynamic when

compared to TGAI and a weaker association with recurrent disease and subsequently, pooled unfavourable outcomes were found.

Values indicating FDG uptake intensity (SUVmax and Z_{mean}), did not show any correlations early in treatment. However the proportional intensity change over time did correlate to TTN and outcome, although not as strongly as TGAI, which combines intensity and volume.

We also combined promising parameters from PET and CT in different ways: using combined segmentation, statistically adding a weighted cavity volume, assigning an intensity value to cavity volume and creating a combined criteria of both TGAI and cavity volume thresholds. While these combined parameters performed similar to their underlying components, we did not find them to be superior to these components.

We did not observe any correlation between documented poor adherence and scan parameters or pooled unfavourable outcome and found this a noteworthy negative finding. As noted earlier (Table 1), the mean amount of treatment dosages missed were significantly increased for patients that failed treatment.

Quantification of EOT + 1y scans confirmed our previous observations that there was a tendency for all parameters to reduce after treatment, but that a lack of complete resolution was still common and dynamic changes often seen. Dynamic changes were common for PET parameters and resulted in a poor correlation between M6 and EOT + 1y measurements, compared to CT lesions which appeared more persistent after treatment. We did not find a trend for mixed response patterns at M6 to be associated with mixed patterns at EOT + 1y (note that failed treatment cases were not included in EOT + 1y analysis).

As expected, recurrent disease was associated with significantly increased TGAI, Cavity volume, MLV and SUVmax, but not $Z_{mean,}$ or V_{total}.

Comparison with previous literature.

Our quantified findings correspond well with previous reports on animal models. In rabbit and primate models, SUV intensity of individual lesions only loosely correlated with bacterial load prior to treatment, [61,62] while on treatment, both the change in lesion size and the change in intensity, independently showed a better correlation to bacterial lesion load and a negative correlation to the potency of the chemotherapy used. In our study, lesion intensity did not correlate with microbiological markers or outcome, but the proportional change in intensity did. Both the total lesion volume and the change volume correlated with outcome and microbiological markers. On treatment, the main focus in these studies were correlation

of lesion specific scan parameters with the MTB bacterial load in the lesions at necropsy and did not comment on total lung burden and outcome, compared to our study, were histology was unavailable and longitudinal clinical outcome was the main independent variable. Before treatment, however, Via et a.l[61] and Coleman et al.[106] did find the central trends to correlate with progression to active disease and in a mouse model, a rise in mean PET activity (SUVmean) preceded culture positive relapse after initial treatment.[66] Our findings also correspond with the rabbit model,[61] that showed FDG uptake intensity change prior to lesion density changes.

An important difference between animal models and human PTB must be noted: The time of initial infection in laboratory animals is always known and interventions would likely follow at a comparatively early stage and may lead to less complex or less established lesions, such as thick walled cavities.

This is the largest study conducted on the use of ^{18}F-FDG PET/CT in human patients with PTB. The prospective follow up and in depth additional investigations and analysis, sets it apart from most other reports in literature on this subject, which often included small sample sizes, varying levels of follow-up and a lack of objective results for comparison. Our quantification of scan parameters allowed the detailed description of scan changes over time and the comparison of these changes to objective microbiological and clinical outcomes. It was also the first study to report on the fate of residual PET/CT lesions after PTB treatment.

Our findings were in keeping with the two most comparable human studies in regards to study design and the depth of analysis (both conducted in drug resistant PTB cases). Coleman et al. found that adding Linezolid to a failing XDR treatment regimen significantly decreased the total glycolytic activity in lung lesions (segmented by using fixed thresholds) within 6 months of follow-up. The sharpest decline in lesion burden was seen in the first month. Chen et al. [75] addressed the prognostic ability of PET and CT changes, by showing a correlation between the change in total glycolytic activity after 2 months, as well as a change in CT reader scores in the lungs after 6 months and treatment outcome in MDR patients.

Our results also confirm the previous reports that cavitatory disease is associated with a poor outcome.[8,19,111] Unlike some previous reports, we did not find a correlation between month 2 culture conversion and pooled unfavourable outcomes. We also found a no association between adherence and recurrent disease.

We did not find an association between infection with the Beijing strain and failed treatment.[41,42] The Beijing strain, however, was associated with a suboptimal reduction in

cavity volume during treatment. This is an interesting finding and if this could be validated in future studies, might in part help to explain the strains rapid increase in incidence.

The persistence of density changes in the lungs are also in keeping with reports of the high incidence of post-tuberculosis lung impairment.

Study Limitations

Limitation to the study design were discussed in chapter 4, page 57. Limitation to the quantification techniques is discussed in section 6.1, page 93.

An additional limitation in the design, not previously mentioned, was the in-availability of pharmacokinetic testing, due to insufficient resources.

While the strong association found between our quantified results and microbiological and clinical variables would suggest that the segmentation and automated quantification technique is of value, the absence of a ground truth, we cannot exclude the possibility that other analysis methods could improve on the prognostic ability of scan characteristics.

At month 6, the best predictor of unfavourable outcomes showed 80% sensitivity and 75% specificity, which is modest for a diagnostic test but far superior to currently used predictive biomarkers for poor treatment outcomes. Including month 2 sputum culture conversion, sputum smear conversion at month 2 and month 5, as well as Chest X-rays and CT scans. Nevertheless, we noted that some patient outcomes were not predicted correctly, including some poor outcomes in patients with excellent improvement on PET/CT scans and some patients who maintained clinical cure in spite of a large and dynamic FDG-avid lesion load detected on PET/CT M6. For this reason relative risk was reported, which allows scan characteristics to be used a part of risk stratification and risk management.

Quantification of central trends in lesion burden provided continuous variables that allowed multiple options for statistical analysis. These analyses, as was the case for Xpert and a history of previous PTB, showed up correlations more clearly. For both MTB mRNA detection in sputum and BAL, however, we still could not find a strong correlation. We did find a linear correlation between cavity volume and the number of MTB mRNA transcripts detected in BAL, but that may have been influenced by the single failed treatment case, for whom all transcripts were detected and who had an extensive lesion burden.

In regards to EOT + 1y scans, the variation in timing between the scan and the recurrent disease diagnosis limited the conclusions we can draw from the data.

Implications

We also found the heterogenic response pattern at a lesion level in the quantified intensity parameters. This was shown by both changes in SUVmax during early treatment and infrequent complete resolution at the end of treatment. Importantly, however, the more predictable changes in lesion burden throughout the lungs; and the strong correlation to serial microbiological investigations and clinical outcomes would suggest that PET/CT correlates with the extent of TB lung involvement and the MTB bacterial load in the lung.

Therefor it suggests that PET/CT scans are useful for risk stratification and treatment response monitoring in therapeutic trials and assisting treatment decisions in complicated clinical cases, such as treatment of drug resistant TB and evaluation of adverse reactions to medication. It also shows the importance of using quantification tools when reporting PET/CT scans in PTB patients.

We found notable differences between failed and recurrent treatment cases. Failed treatment was associated with extensive lung lesions at baseline and large cavities with thick walls at M6, as well as poor adherence. Recurrent cases displayed a comparatively low lesion burden at baseline, average adherence and time to sputum culture negativity, but insufficient reduction in lesion burden during treatment. We also found insufficient reduction in lesion burden in patients with a history of previous PTB episodes. This raises important questions regarding the mechanisms underlying unfavourable outcomes. It may indicate that patients who develop recurrent PTB disease have an insufficient immune response. However, the role that pharmacokinetics and MTB virulence factors play in this group also need further investigation.

We found less reduction in cavity volume patients infected with the Beijing strain. It is not clear if this was related to virulence factors regarding tolerance to drugs or host immunity.

It was also surprising that the change in total glycolytic activity index showed a stronger correlation to sputum culture TTP at baseline and sputum Xpert positivity at M6, compared to cavity volume, which showed a stronger correlation to TTN, mRNA in BAL. This supports the concept that the type of lesion plays an important part in the culturability of different MTB populations, and not only the bacterial burden, since lesions outside cavities may exist in an altered metabolic state.[37]

The technique produced continuous variables representing the inflammatory lesion burden in the lungs and we proposed optimal thresholds, which could be used to form grouping variables. This could be used as a tool to aid the discovery of biomarkers that could replace

more expensive and labour intensive PET/CT scans in risk stratification and individualised management decisions.

Future research

Various follow-up research projects are already underway or in advanced planning stages.

Examples of using the quantified variables to aid biomarker discovery on the same cohort are shown in the next chapter.

In a pilot project aimed at investigating possible end of treatment MTB persistence and improving its detection, a post-doctoral fellow, Dr Caroline Beltran, from our group is currently collaborating with Prof Bhavesh Kana (University of Witwatersrand). We have recruited 26 additional PTB patients at the end of treatment. We performed PET/CT scans on them and collected bronchoalveolar lavage fluid, post-bronchoscopy sputum, induced sputum and spot sputum samples. We will attempt to replicate the scan analysis and mRNA assays. In addition, we will conduct further gene expression analysis, and attempt to resuscitate viable, not conventionally culturable MTB, by applying and optimising resuscitation techniques adapted from methods developed by Prof Kana.

We hope that more sensitive culture techniques would pave the way for better understanding of antibiotic and immune activity against MTB persistence. In turn, this may lead to better immune modulation therapy to improve treatment outcome and perhaps decrease post tuberculosis lung sequelae.

In an ongoing trial, we are collaborating with the Tuberculosis Research Section of the NIH and TASK applied science. In this trial, PET/CT quantification is used in addition to sputum MTB CFU counts in determining the early bactericidal activity of different combinations of TB medication and the effect on different lesion types. This study includes pharmacokinetic analysis.

We are also involved in the advanced planning of a trial aimed at individualised treatment shortening for patients with a low baseline lesion extent, adequate reduction of total glycolytic activity on PET/CT at M1 and a high Xpert cycle threshold at month 4. End of treatment scans would also be used to optimise and validate quantified criteria. Other host biomarkers discovered in the current project would also optimised and validated with the goal to be used in place of more costly and labour intensive PET/CT scans in the future. This study will also include pharmacokinetic testing.

Conclusions

Quantification of the PET/CT images significantly improved the prognostic accuracy of scans compared to qualitative scan patterns. Risk stratification is already possible during early treatment and improves by the end of treatment. A high cavity volume showed the strongest association with a risk of treatment failure, while a suboptimal reduction of the total glycolytic activity throughout the lung had the strongest association with recurrent disease. Both these variables also correlated to microbiological markers. This suggests a correlation between the quantified lesion burden and the MTB load.

7. Using quantified PET/CT variables to aid biomarker discovery

7.1 Lesion burden and host inflammatory cytokines

We evaluated 32 host protein biomarkers on serum samples from the study cohort, using a multiplex cytokine platform. The laboratory work was led by Dr Novel Chegou and in this section, I present preliminary results from in-depth bio-informatics analysis performed by Prof Gerard Tromp at Stellenbosch University.

Serum samples, collected at Dx, M1, M3 and M6 were analysed on the Luminex platform. We log transformed or scaled the cytokine data to be approximately normally distributed. For the Dx time-point, TGAI lesion burden of all participants was divided into tertiles. Linear support vector machine (LSVM) models were developed from the cytokine data to predict which patients fell into the lowest tertile of TGAI measures (lowest risk for poor outcome). Four models were calculated and all showed AUCs of greater than 86% (Figure 44). In addition, LSVM models were developed to indicate which patients fell within the highest tertile of TGAI burden. The models all showed AUCs of greater that 90% (Figure 46).

This could potentially serve as early low-risk prognostic indicator, which would be based on serum cytokine measurements rather than PET/CT imaging.

Evaluation M6 samples to calculate predictive cytokine models, which accurately indicate patients with a high lesion load at the end of treatment that could serve as a high-risk prognostic indicator would be the next step.

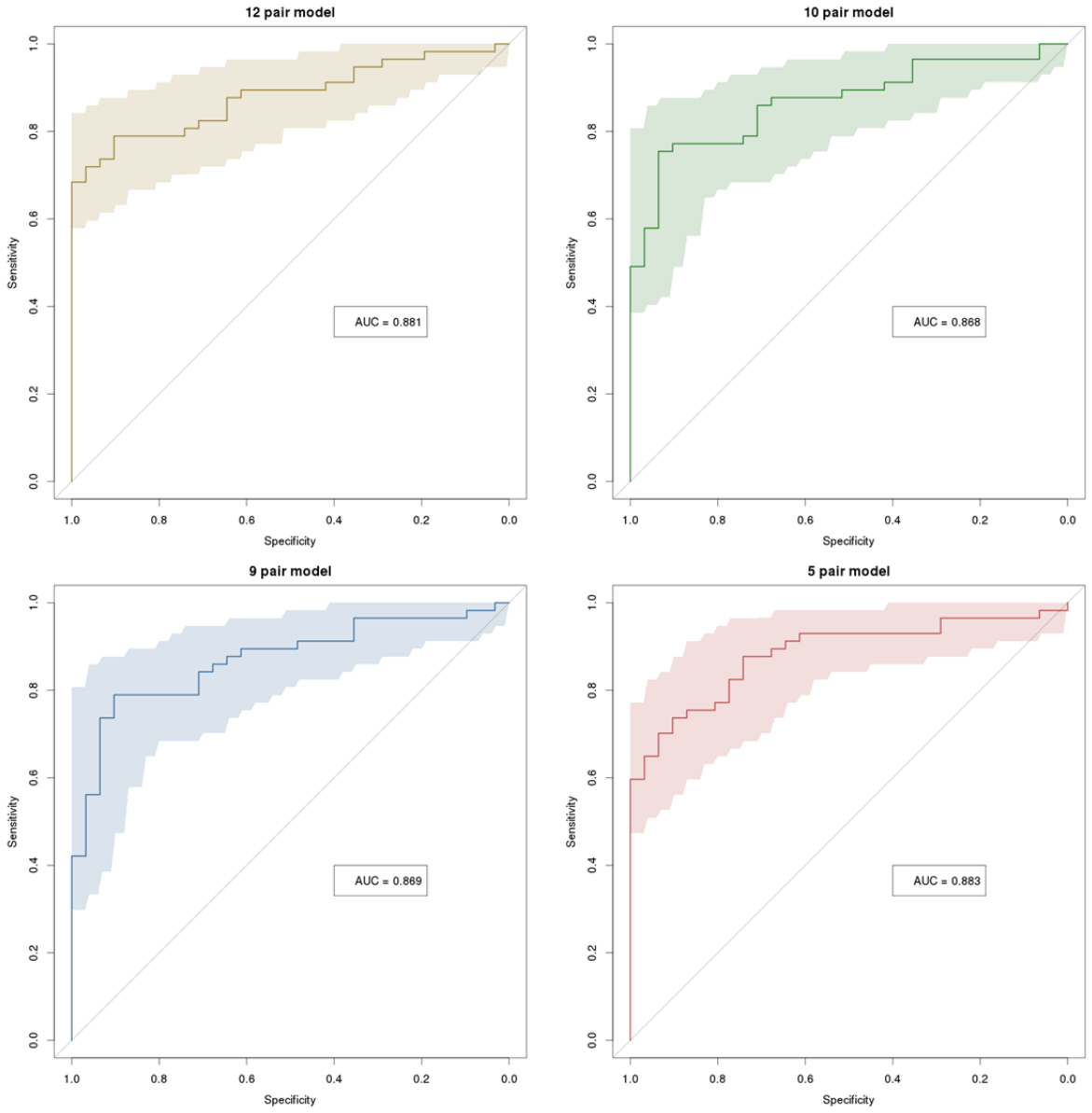


Figure 44 ROC curves to show accuracy of host-protein markers to indicate low TGAI at Dx.

Shaded area shows 95% confidence interval. 5, 9, 10 and 12 marker models were developed, all showed AUC's of greater than 86%.

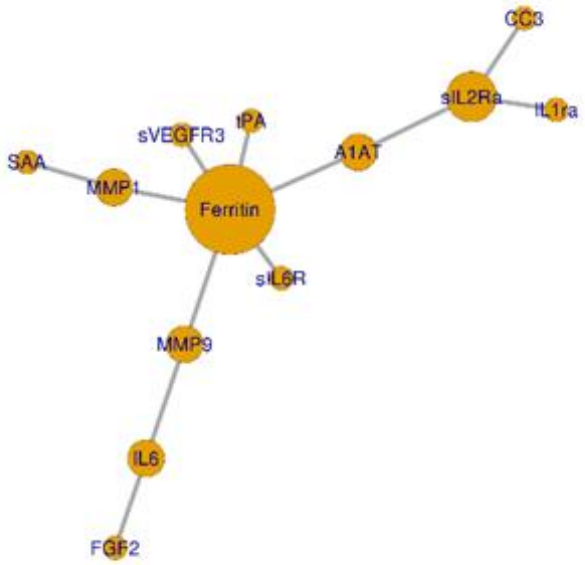


Figure 45 The host proteins used in the 12-marker mathematical model.

Spheres indicate specific host proteins, the size of the spheres the importance of the specific markers in the predictive model and the lines indicate a relationships between the specific markers in the model.

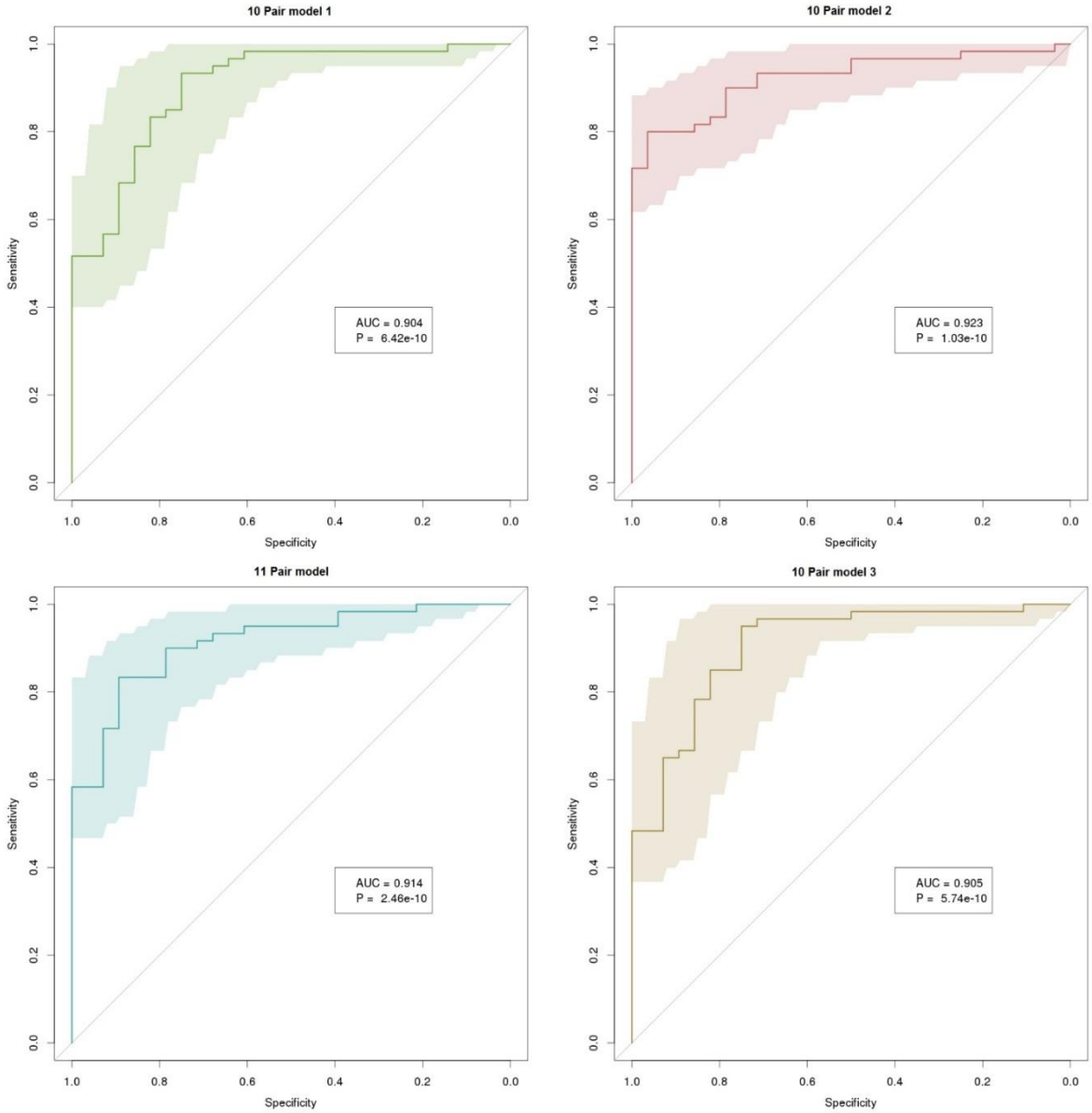


Figure 46 ROC curves to show accuracy of host-protein markers to indicate high TGAI at Dx. Shaded area shows 95% confidence interval. 10 and 11 marker models were developed and all showed AUC's of greater than 90%.

7.2 Lesion burden and host gene expression.

Blood samples collected from the PTB cohort at Dx, M1 and M6 were sent to the Center for Infectious Disease Research in Seattle, Washington State, USA, for host-gene expression analysis (RNA Sequencing) to develop well-defined mathematical models for predicting outcome and discovering genes and pathways associated with different clinical outcomes.

The analysis was led by Daniel Zak and Ethan Thompson. The expression of thousands of host-genes were measured and compared to microbiological test results, clinical outcomes and quantified PET/CT data.

TGAI and $TGAI_{com}$ showed a very strong correlation to thousands of genes – at all time-points and on all patients. Examples of specific genes with a strong correlation include *Gbp5* and *Ankrd22* (shown in Figure 47). More than 8000 genes had a false discovery probability of $P < 0.01$ and more than 2000 genes had a false discovery probability of $P < 0.00000001$ (shown in Figure 48). Host gene expression's correlation with TGAI was stronger than with MGIT culture time to positivity and Xpert cycle time values.

Further analysis are underway to use these correlations in developing a host-gene expression signature that could be used for risk stratification and treatment monitoring.

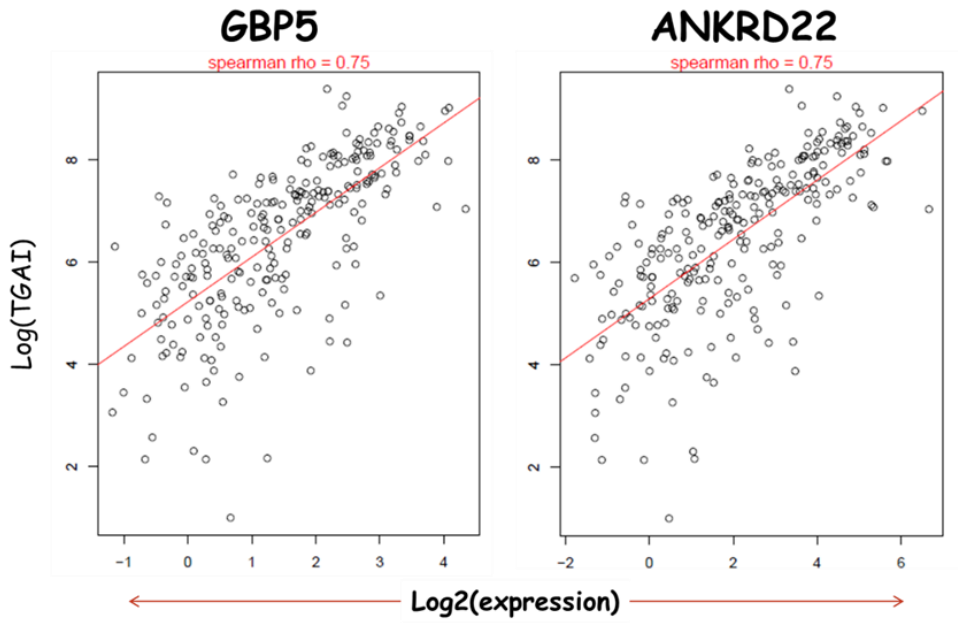


Figure 47 Scatterplot of correlation between the expression of 2 host-genes and log (TGAI).

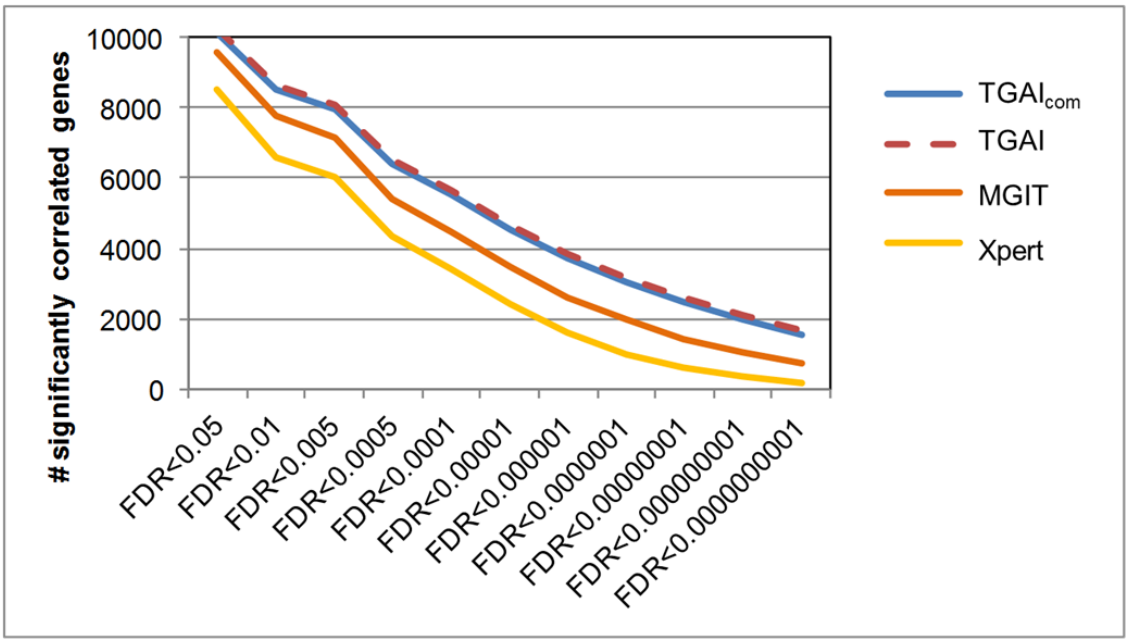


Figure 48 Gene correlation TGAI, TGAIcom, Culture TTP and Xpert Cycle threshold value.

The y-axis represent the number of genes and the x-axis the level of significance.

7.3 Discussion: Biomarker discovery

We found a strong agreement between circulating host proteins and host-gene expression measured in peripheral blood, and the lung lesion burden as measured by automated segmentation techniques. This is reassuring and promising.

It serves as further substantiation for:

1) The usefulness of PET/CT in TB research.

2) The robustness of the user-independent segmentation method.

3) The importance of TGAI as a biomarker.

4) The feasibility of finding more easily measurable biomarkers in blood that correlate with the most important PET/CT variables.

Peripheral blood is easily collected compared to more invasive sampling techniques, like bronchoalveolar or lung specimens and requires less resources than PET/CT scans. It shows great promise that biomarkers measured from blood could be used to provide information regarding the lesion burden in lungs. With further research these measurements could enriched as a prognostic indicator of morbidity (PTB recurrence and post-TB lung impairment) and mortality – leading to improved individualised management. The markers in peripheral blood could also be used to build mechanistic hypotheses.

8. Concluding remarks

Realisation of study objectives.

In a major collaborative team effort, we successfully recruited and followed-up the participant targets set out. We also successfully conducted in depth microbiological testing alongside the FDG-PET CT scans and collected a variety of specimens for biomarker discovery. Very importantly, patients were well characterised and data captured in a study-specific database.

Some additional goals were set (and accomplished) during the study, as new research questions came up. These include: 1) Follow-up after treatment to determine outcome after treatment. 2) Performing bronchoscopies on patients at the end of treatment. 3) Performing an additional PET/CT scan 1 year after the end of treatment.

Defining outcome after treatment proved especially challenging. Factors adding to the challenge included: 1) GeneXpert was introduced in Cape Town clinics during the recruitment and follow-up and some healthcare providers were not yet very well informed about the complexity of a positive Xpert after treatment and the importance of sending sputum for a culture if a patient has recently completed treatment. 2) It is not uncommon for doctors in primary care to initiate TB treatment on clinical suspicion, since the diagnosis of TB can be very difficult to prove, especially in patients with a previous history of TB and delaying treatment could be harmful.

Scans were qualitatively evaluated and we reported a heterogenic response and a surprising discovery that most patients still have residual lesions showing a significant inflammatory response at end of treatment, in spite of clinical cure. We also reported the presence of MTB mRNA in a significant proportion of patients' sputum at the end of treatment and in BAL fluid for all patients that underwent this procedure. This raised up important questions regarding the nature of host-pathogen interaction and the role an adequate immune response plays in maintaining a disease-free state after cure.

We developed a technique requiring limited user input, that allowed that standardisation of PET uptake, based on a reference volume specific to the scan and the organ in question (in this case lungs), the auto-segmentation of lesions and the quantification of a variety of PET/CT characteristics at a voxel level. We successfully implemented the method in all scans of included participants.

The results correlated very well with microbiological and clinical data. Therefore, we were able to define prognostic scan parameters. Analysis are currently underway to use these quantified parameters as a tool for new biomarker discovery. Early results appear very

promising. Prognostic parameters might also be used to assist in the management of complicated clinical cases, such as MDR TB.

In conclusion our results strongly suggest that PET/CT could be a useful tool in the monitoring the response to the treatment of pulmonary tuberculosis. However, PET/CT is an expensive resource and only available in the large imaging facilities of countries with access to radioactive tracers. This would imply that its use would most likely be limited to research and complicated clinical cases, where it could possibly guide treatment in a cost-effective manner.

Dissemination

A manuscript, entitled: '*Persisting PET/CT lesion activity and Mycobacterium tuberculosis mRNA transcripts are common after curative treatment of pulmonary tuberculosis'* (Malherbe et al.) has been published in Nature Medicine.[112]

Another manuscript, entitled '*Bacterial loads measured by the Xpert MTB/RIF assay as markers of culture conversion and bacteriological cure in pulmonary TB'*, (Shenai et al.), has been published in PLoS ONE.[113] This manuscript describes the relationship between Xpert cycle time values and culture positivity and outcome.

We have prepared an advanced draft of a manuscript entitled: 'An operator independent technique to quantify widespread complex tuberculous lung lesions using PET/CT', which we will shortly submit to a technical medical physics journal.

A manuscript describing the quantified scan parameters and prognostic indicators is in advanced planning stages and will be submitted to a nuclear medicine journal.

Further manuscripts describing novel biomarkers and their correlation to scan parameters are being planned.

Personal role in project

As a study clinician, I managed the nursing staff, supervised participant recruitment and assisted in procedures, like PET/CT scans and bronchoscopies. I reviewed test results, including vitals, sputum microbiology, CXR's and PET/CT scans. Firstly, from a clinical perspective, to ensure due diligence and inform the patient and appropriate healthcare providers of any information which might influence case management. Secondly, I helped to clean data and consolidate clinical results. I was also part of the team that determined and applied outcome definitions.

I took part in the discussions regarding the addition of the EOT bronchoscopies and the EOT + 1y scans, made the protocol amendments and applied for ethical approval of the changes.

I designed and applied the method to evaluate the PET/CT scans qualitatively. Since only a minority of scans showed complete resolution, I devised a way to subdivide scan response patterns and grade intensity. I adjusted the intensity grading system (the Deauville score) to allow monitoring of specific lesion progression or resolution. To subdivide the response patterns, I used MIM Fusion software, to track the relative intensity of individual lesions from Dx to M6 and later from M6 to EOT + 1y to determine whether scans were resolved, improved or mixed. I used the combined radiologist and nuclear physician scan reports to corroborate my findings and in case of any uncertainty, I asked Prof Walzl (pulmonologist) and Prof Warwick (nuclear physician) for a second opinion.

The user independent method to quantify PET/CT images was conceptualised by Prof Warwick and Prof Dupont. Prof Dupont and Ilse Kant (medical technology Masters student) wrote the first code for the MATLAB program and Prof Warwick evaluated and optimised it. I assessed the application of the technique and made further changes to the method and script. These included: defining the intensity and density thresholds, improving the way to create a lung mask accurately (by drawing the lung VOI and removing areas affected by motion misregistration). After the technique was optimized, I applied it on all the scans included in analysis.

Statistical analysis were performed on consolidated datasets I provided, by Prof Gerard Tromp, Tom Peppard, Lori Dodd, Yookwan Noh and myself. More details regarding statistical analysis provided above, in Chapter 2.8.

I also advised which qualitative and quantitative scan parameters seem the most important to aid further biomarker discovery.

Personal statement

It was a great privilege to be part of this project. Firstly, I learned the scientific thinking process. I also gained a lot of knowledge in the fields of immunology and microbiology, statistics, pulmonology, radiology and nuclear medicine. My previous experience in multi-disciplinary healthcare helped to give me insight in the important research questions.

I was also introduced to an international collaboration. It was an honour to network with great scientific minds. I look forward to build on these partnerships in the future.

I sincerely believe I can take the lessons I learnt while working on this project with me into the future and make an impact as a clinical researcher.

References:

1. WHO. Global Tuberculosis Report. **20,** (2015).

2. Fox, W. Whither Short-course chemotherapy? *Br. J. Dis. Chest* **75,** 331–337 (1981).

3. Cox, H. S., Morrow, M. & Deutschmann, P. W. Long term efficacy of DOTS regimens for tuberculosis: systematic review. *BMJ* **336,** 484–487 (2008).

4. Gillespie, S. H. *et al.* Four-Month Moxifloxacin-Based Regimens for Drug-Sensitive Tuberculosis. *N. Engl. J. Med.* **371,** 1577–1587 (2014).

5. Jindani, A. *et al.* High-Dose Rifapentine with Moxifloxacin for Pulmonary Tuberculosis. *N. Engl. J. Med.* **371,** 1599–1608 (2014).

6. Merle, C. S. *et al.* A Four-Month Gatifloxacin-Containing Regimen for Treating Tuberculosis. *N. Engl. J. Med.* **371,** 1588–1598 (2014).

7. Luzze, H. *et al.* Relapse more common than reinfection in recurrent tuberculosis 1-2 years post treatment in urban Uganda. *Int. J. Tuberc. Lung Dis.* **17,** 361–367 (2013).

8. Sonnenberg, P. *et al.* HIV-1 and recurrence, relapse, and reinfection of tuberculosis after cure: a cohort study in South African mineworkers. *Lancet* **358,** 1687–1693 (2001).

9. Marx, F. M. *et al.* The temporal dynamics of relapse and reinfection tuberculosis after successful treatment: A retrospective cohort study. *Clin. Infect. Dis.* **58,** 1676–1683 (2014).

10. Middelkoop, K., Bekker, L.-G., Shashkina, E., Kreiswirth, B. & Wood, R. Retreatment tuberculosis in a South African community: the role of re-infection, HIV and antiretroviral treatment. *Int. J. Tuberc. Lung Dis.* **16,** 1510–6 (2012).

11. Shen, G. *et al.* Recurrent Tuberculosis and Exogenous Reinfection, Shanghai. **12,** 11–13 (2006).

12. Wood, R. *et al.* Burden of new and recurrent tuberculosis in a major South African city stratified by age and HIV-status. *PLoS One* **6,** 1–9 (2011).

13. Nunn, a J., Phillips, P. P. J. & Mitchison, D. a. Timing of relapse in short-course chemotherapy trials for tuberculosis. *Int. J. Tuberc. Lung Dis.* **14,** 241–242 (2010).

14. WHO. Treatment of tuberculosis guidelines. **4,** (2010).

15. Hepple, P., Ford, N. & McNerney, R. Microscopy compared to culture for the diagnosis of tuberculosis in induced sputum samples: a systematic review [Review article]. *Int. J. Tuberc. Lung Dis.* **16,** 579–588 (2012).

16. Horne, D. J. *et al.* Sputum monitoring during tuberculosis treatment for predicting outcome: systematic review and meta-analysis. *Lancet Infect. Dis.* **10,** 387–394 (2010).

17. Hesseling, a C. *et al.* Baseline sputum time to detection predicts month two culture conversion and relapse in non-HIV-infected patients. *Int. J. Tuberc. Lung Dis.* **14,** 560–570 (2010).

18. Stop TB Partnership, W. The Global Plan to Stop TB 2016 – 2020. **2020,** 1–4 (2016).

19. Johnson, J. L. *et al.* Shortening treatment in adults with noncavitary tuberculosis and 2-month culture conversion. *Am. J. Respir. Crit. Care Med.* **180,** 558–563 (2009).

20. Warner, D. F. & Mizrahi, V. Shortening Treatment for Tuberculosis - Back to Basics. *N. Engl. J. Med.* **371,** 1642–1643 (2014).

21. Seon, H. J., Kim, Y. I., Lim, S. C., Kim, Y. H. & Kwon, Y. S. Clinical significance of residual lesions in chest computed tomography after anti-tuberculosis treatment. *Int J Tuberc Lung Dis* **18,** 341–346 (2014).

22. Ralph, A. P. *et al.* A simple, valid, numerical score for grading chest x-ray severity in adult smear-positive pulmonary tuberculosis. *Thorax* **65,** 863–869 (2010).

23. Kriel, M., Lotz, J. W., Kidd, M. & Walzl, G. Evaluation of a radiological severity score to predict treatment outcome in adults with pulmonary tuberculosis. *Int. J. Tuberc. Lung Dis.* **19,** 1354–1360 (2015).

24. Nihues, S. de S. E. *et al.* Chronic symptoms and pulmonary dysfunction in post-tuberculosis Brazilian patients. *Brazilian J. Infect. Dis.* **19,** 492–497 (2015).

25. Pasipanodya, J. G. *et al.* Using the St. George Respiratory Questionnaire to ascertain health quality in persons with treated pulmonary tuberculosis. *Chest* **132,** 1591–1598 (2007).

26. Wejse, C. *et al.* TBscore: Signs and symptoms from tuberculosis patients in a low-resource setting have predictive value and may be used to assess clinical course. *Scand. J. Infect. Dis.* **40,** 111–120 (2008).

27. Pasipanodya, J. G. *et al.* Pulmonary impairment after tuberculosis. *Chest* **131,** 1817–1824 (2007).

28. Hnizdo, E., Singh, T. & Churchyard, G. Chronic pulmonary function impairment caused by initial and recurrent pulmonary tuberculosis following treatment. *Thorax* **55,** 32–38 (2000).

29. Baig, I. M., Saeed, W. & Khalil, K. F. Post-tuberculous chronic obstructive pulmonary disease. *J. Coll. Physicians Surg. Pak.* **20,** 542–4 (2010).

30. Pasipanodya, J. G. *et al.* Pulmonary impairment after tuberculosis and its contribution to TB burden. *BMC Public Health* **10,** 259 (2010).

31. Friedrich, S. O. *et al.* Assessment of the sensitivity and specificity of Xpert MTB/RIF assay as an early sputum biomarker of response to tuberculosis treatment. *Lancet Respir. Med.* **2600,** 1–9 (2014).

32. Lahtinen, S. J. *et al.* Degradation of 16S rRNA and attributes of viability of viable but nonculturable probiotic bacteria. *Lett. Appl. Microbiol.* **46,** 693–698 (2008).

33. Mitchison, D. A. Basic Mechanisms of Chemotherapy. *Chest* **76,** (1979).

34. Hu, Y. *et al.* Detection of mRNA transcripts and active transcription in persistent Mycobacterium tuberculosis induced by exposure to rifampin or pyrazinamide. *J. Bacteriol.* **182,** 6358–6365 (2000).

35. Rustad, T. R. *et al.* Global analysis of mRNA stability in Mycobacterium tuberculosis. *Nuclecic Acids Res.* **41,** 509–517 (2013).

36. Ignatov, D. V. *et al.* Dormant non-culturable Mycobacterium tuberculosis retains stable low-abundant mRNA. *BMC Genomics* **16,** 954 (2015).

37. Boshoff, H. I. & Barry 3rd, C. E. Tuberculosis - metabolism and respiration in the absence of growth. *Nat Rev Microbiol* **3,** 70–80 (2005).

38. Smith, I. *Mycobacterium tuberculosis* pathogenesis and molecular determinants of virulence. *Clin. Microbiol. Rev.* **16,** 463–496 (2003).

39. Forrellad, M. A. *et al.* Virulence factors of the Mycobacterium tuberculosis complex. *Virulence* **4,** 3–66 (2013).

40. Glynn, J. R., Whiteley, J., Bifani, P. J., Kremer, K. & Van Soolingen, D. Worldwide Occurrence of Beijing/W Strains of Mycobacterium tuberculosis: A Systematic

Review. *Emerg. Infect. Dis.* **8,** 843–849 (2002).

41. Parwati, I. *et al.* Mycobacterium tuberculosis Beijing genotype is an independent risk factor for tuberculosis treatment failure in Indonesia. *J. Infect. Dis.* **201,** 553–557 (2010).

42. Lan, N. T. N. *et al.* Mycobacterium tuberculosis Beijing Genotype and Risk for Treatment Failure and Relapse, Vietnam. *Emerg. Infect. Dis.* **9,** 1633–1635 (2003).

43. Dooley, K. E. *et al.* Risk factors for tuberculosis treatment failure, default, or relapse and outcomes of retreatment in Morocco. *BMC Public Health* **11,** 140 (2011).

44. Winston, C. A., Hopewell, P. C., Wells, C. D. & Kenyon, T. A. Isoniazid, Rifampin, Ethambutol and Pyrazinamide Pharmacokinetics and Treatment Outcomes among a Predominantly HIV-Infected Cohort of Adults with Tuberculosis. *Clin Infect Dis.* **48,** 1685–1694 (2009).

45. Tostmann, A. *et al.* Pharmacokinetics of first-line tuberculosis drugs in Tanzanian patients. *Antimicrob. Agents Chemother.* **57,** 3208–3213 (2013).

46. Kjellsson, M. C. *et al.* Pharmacokinetic evaluation of the penetration of antituberculosis agents in rabbit pulmonary lesions. *Antimicrob. Agents Chemother.* **56,** 446–57 (2012).

47. Ge, Z., Wang, Z. & Wei, M. Measurement of the concentration of three antituberculosis drugs in the focus of spinal tuberculosis. *Eur. Spine J.* **17,** 1482–7 (2008).

48. Maldonado, A., González-Alenda, F. J., Alonso, M. & Sierra, J. M. PET-CT in clinical oncology. *Clin. Transl. Oncol.* **9,** 494–505 (2007).

49. Hess, S., Hansson, S. H., Pedersen, K. T., Basu, S. & Høilund-Carlsen, P. F. FDG-PET/CT in Infectious and Inflammatory Diseases. *PET Clin.* **9,** 497–519 (2014).

50. Cheebsumon, P. *et al.* Effects of image characteristics on performance of tumor delineation methods: a test-retest assessment. *J. Nucl. Med.* **52,** 1550–8 (2011).

51. Huang, W. *et al.* Standard uptake value and metabolic tumor volume of ^{18}F-FDG PET/CT predict short-term outcome early in the course of chemoradiotherapy in advanced non-small cell lung cancer. *Eur. J. Nucl. Med. Mol. Imaging* **38,** 1628–35 (2011).

52. Dibble, E. H. *et al.* 18F-FDG metabolic tumor volume and total glycolytic activity of oral cavity and oropharyngeal squamous cell cancer: adding value to clinical staging. *J. Nucl. Med.* **53,** 709–15 (2012).

53. Foster, B., Bagci, U., Mansoor, A., Xu, Z. & Mollura, D. J. A review on segmentation of positron emission tomography images. *Comput. Biol. Med.* **50,** 76–96 (2014).

54. Fletcher, J. W. PET/CT Standardized Uptake Values (SUVs) n Clinical Practice Assessing Response to Therapy. *Semin Ultrasouns CT MRI* **31,** 496–505 (2010).

55. Lowe, V. J., Hoffman, J. M., DeLong, D. M., Patz, E. F. & Coleman, R. E. Semiquantitative and visual analysis of FDG-PET images in pulmonary abnormalities. *J. Nucl. Med.* **35,** 1771–6 (1994).

56. Hatt, M. *et al.* Reproducibility of 18F-FDG and 3'-deoxy-3'-18F-fluorothymidine PET tumor volume measurements. *J. Nucl. Med.* **51,** 1368–76 (2010).

57. Boellaard, R., Krak, N. C., Hoekstra, O. S. & Lammertsma, A. a. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values: a simulation study. *J. Nucl. Med.* **45,** 1519–1527 (2004).

58. Boellaard, R. Standards for PET image acquisition and quantitative data analysis. *J. Nucl. Med.* **50 Suppl 1,** 11S–20S (2009).

59. Lin, P. L. *et al.* Radiologic responses in cynomolgous macaques for assessing tuberculosis chemotherapy regimens. *Antimicrob. Agents Chemother.* **57,** 4237–4244 (2013).

60. Coleman, M., Chen, R. & Lee, M. PET/CT imaging reveals a therapeutic response to oxazolidinones in macaques and humans with tuberculosis. *Sci. Transl. Med.* **6,** 1–10 (2014).

61. Via, L. E. *et al.* Infection dynamics and response to chemotherapy in a rabbit model of tuberculosis using [^{18}F]2-fluoro-deoxy-D-glucose positron emission tomography and computed tomography. *Antimicrob. Agents Chemother.* **56,** 4391–402 (2012).

62. Lin, P. L. *et al.* Sterilization of granulomas is common in both active and latent tuberculosis despite extensive within-host variability in bacterial killing. *Nat Med.* **20,** 75–79 (2014).

63. Lin, P. L. *et al.* Early events in Mycobacterium tuberculosis infection in cynomolgus macaques. *Infect. Immun.* **74,** 3790–803 (2006).

64. Coleman, M. T. *et al.* Early Changes by (18)Fluorodeoxyglucose positron emission tomography coregistered with computed tomography predict outcome after Mycobacterium tuberculosis infection in cynomolgus macaques. *Infect. Immun.* **82,** 2400–4 (2014).

65. Via, L. E. *et al.* Differential virulence and disease progression following mycobacterium tuberculosis complex infection of the common marmoset (callithrix jacchus). *Infect. Immun.* **81,** 2909–2919 (2013).

66. Davis, S. L. *et al.* Noninvasive pulmonary [18F]-2-fluoro-deoxy-D-glucose positron emission tomography correlates with bactericidal activity of tuberculosis drug treatment. *Antimicrob. Agents Chemother.* **53,** 4879–4884 (2009).

67. du Toit, R. *et al.* The diagnostic accuracy of integrated positron emission tomography/computed tomography in the evaluation of pulmonary mass lesions in a tuberculosis-endemic area. *South African Med. J.* **105,** 1049–1052 (2015).

68. Shaw, J. A. *et al.* Integrated positron emission tomography/computed tomography for evaluation of mediastinal lymph node staging of non-small-cell lung cancer in a tuberculosisendemic area: A 5-year prospective observational study. *South African Med. J.* **105,** 145–150 (2015).

69. Sathekge, M., Maes, A. & Wiele, C. Van De. FDG-PET Imaging in HIV Infection and Tuberculosis. *Semin. Nucl. Med.* **43,** 349–66 (2013).

70. Martinez, V. *et al.* (18)F-FDG PET/CT in tuberculosis: an early non-invasive marker of therapeutic response. *Int. J. Tuberc. Lung Dis.* **16,** 1180–5 (2012).

71. Dureja, S., Sen, I. & Acharya, S. Potential role of F18 FDG PET-CT as an imaging biomarker for the noninvasive evaluation in uncomplicated skeletal tuberculosis: a prospective clinical observational. *Eur. Spine J.* **23,** 2449–2454 (2014).

72. Demura, Y. *et al.* Usefulness of 18F-fluorodeoxyglucose positron emission tomography for diagnosing disease activity and monitoring therapeutic response in patients with pulmonary mycobacteriosis. *Eur. J. Nucl. Med. Mol. Imaging* **36,** 632–639 (2009).

73. Stelzmueller, I. *et al.* 18F-FDG PET / CT in the Initial Assessment and for Follow-up in Patients With Tuberculosis. *Clin. Nucl. Med.* **41,** 187–194 (2016).

74. Sathekge, M., Maes, A., Kgomo, M., Stoltz, A. & Van de Wiele, C. Use of 18F-FDG

PET to predict response to first-line tuberculostatics in HIV-associated tuberculosis. *J. Nucl. Med.* **52,** 880–885 (2011).

75. Chen, R. Y. *et al.* PET/CT imaging correlates with treatment outcome in patients with multidrug-resistant tuberculosis. *Sci. Transl. Med.* **6,** 265ra166 (2014).

76. Boellaard, R. *et al.* FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. *Eur. J. Nucl. Med. Mol. Imaging* **37,** 181–200 (2010).

77. Fogh, S. E. *et al.* Pathological correlation of PET/CT based auto contouring for radiation planning of lung cancer. *Bodine J.* **78,** 202–203 (2010).

78. Dann, E. J. *et al.* A functional dynamic scoring model to elucidate the significanceof post-induction interim fluorine-18-fluorodeoxyglucose positron emission tomography findings in patients with hodgkin's lymphoma. *Haematologica* **95,** 1198–1206 (2010).

79. Delbeke, D. *et al.* Expert opinions on positron emission tomography and computed tomography imaging in lymphoma. *Oncologist* **14 Suppl 2,** 30–40 (2009).

80. Firouzian, A., Kelly, M. D. & Declerck, J. M. Insight on automated lesion delineation methods for PET data. *EJNMMI Res.* **4,** 1–12 (2014).

81. David Chien, Martin Lodge, R. W. Reproducibility of liver and mediastinal blood pool F-18 activity as normal reference tissues. *J. Nucl. Med.* **52,** (2011).

82. Higashi, K. *et al.* 18F-FDG uptake by primary tumor as a predictor of intratumoral lymphatic vessel invasion and lymph node involvement in non-small cell lung cancer: analysis of a multicenter study. *J. Nucl. Med.* **46,** 267–73 (2005).

83. Barrington, S. F. *et al.* Concordance between four European centres of PET reporting criteria designed for use in multicentre trials in Hodgkin lymphoma. *Eur. J. Nucl. Med. Mol. Imaging* **37,** 1824–1833 (2010).

84. Foster*, B. *et al.* Segmentation of PET Images for Computer Aided Functional Quantification of Tuberculosis in Small Animal Models. *IEEE Trans. Biomed. Eng.* **61,** 711–724 (2013).

85. Hatt, M., Cheze le Rest, C., Turzo, A., Roux, C. & Visvikis, D. A Fuzzy Locally Adaptive Bayesian Segmentation Approach for Volume Determination in PET. *IEEE Trans. Med. Imaging* **28,** 881–893 (2009).

86. Ulas Bagci, Jayaram K. Udupa, Neil Mendhiratta, Brent Foster, Ziyue Xu, J. & Yao,

Xinjian Chen, and D. J. M. Joint Segmentation of Anatomical and Functional Images: Applications in Quantification of Lesions from PET, PET-CT, MRI-PET, and MRI-PET-CT Images. *Med. Image Anal.* **17,** 929–945 (2013).

87. Chong, D. *et al.* Reproducibility of volume and densitometric measures of emphysema on repeat computed tomography with an interval of 1 week. *Eur. Radiol.* **22,** 287–94 (2012).

88. Colombi, D. *et al.* Visual vs Fully Automatic Histogram-Based Assessment of Idiopathic Pulmonary Fibrosis (IPF) Progression Using Sequential Multidetector Computed Tomography (MDCT). *PLoS One* **10,** e0130653 (2015).

89. Karimi, R., Tornling, G. & Forsslund, H. Lung density on high resolution computer tomography (HRCT) reflects degree of inflammation in smokers. *Respir. …* **15,** 1–10 (2014).

90. Shin, K. E., Chung, M. J., Jung, M. P., Choe, B. K. & Lee, K. S. Quantitative computed tomographic indexes in diffuse interstitial lung disease: correlation with physiologic tests and computed tomography visual scores. *J. Comput. Assist. Tomogr.* **35,** 266–271 (2011).

91. Lewis Center for Neuroimaging. MRIConvert and mcverter.

92. Welcome Trust Centre for Neuroimaging. Statistical Parametric Mapping. doi:10.1007/978-1-4615-1079-6_16

93. Chris Rorden. MRIcro. http://people.cas.sc.edu/rorden/mricro/index.html

94. Rasband, W. ImageJ. http://imagej.net

95. Walter, N. D. *et al.* Transcriptional adaptation of drug-tolerant Mycobacterium tuberculosis during treatment of human tuberculosis. *J. Infect. Dis.* **212,** 1–9 (2015).

96. Lew JM, Kapopoulou A, Jones LM, C. S. TubercuList - TB Gene Database. Available at: http://tuberculist.epfl.ch/quicksearch.php?gene+name=Rv2031c. (Accessed: 30th June 2015)

97. Jacomelli, M. *et al.* Bronchoscopy for the diagnosis of pulmonary tuberculosis in patients with negative sputum smear results. *J Bras Pneum* **38,** 167–173 (2012).

98. Reingold, A. L., Daley, C. L. & Kritski, A. L. Comparison of Sputum Induction with Fiberoptic Bronchoscopy in the Diagnosis of Tuberculosis Experience at an Acquired

Immune Deficiency Syndrome Reference Center in Rio de Janeiro , Brazil. **162,** 2238–2240 (2000).

99. George, P. M. *et al.* Post-bronchoscopy sputum: Improving the diagnostic yield in smear negative pulmonary TB. *Respir. Med.* **105,** 1726–1731 (2011).

100. Galagan, J. E. *et al.* The Mycobacterium tuberculosis regulatory network and hypoxia. *Nature* **499,** 178–83 (2013).

101. Nicholas D Walter, Gregory M Dolganov, et al. Transcriptional adaptation of drug-tolerant Mycobacterium tuberculosis during treatment of human tuberculosis supplement.

102. Commandeur, S. *et al.* An Unbiased Genome-Wide Mycobacterium tuberculosis Gene Expression Approach To Discover Antigens Targeted by Human T Cells Expressed during Pulmonary Infection. *J. Immunol.* **190,** 1659–1671 (2013).

103. Warren, R. M. *et al.* Patients with active tuberculosis often have different strains in the same sputum specimen. *Am. J. Respir. Crit. Care Med.* **169,** 610–614 (2004).

104. Ramos, L. M. M., Sulmonett, N., Ferreira, C. S., Henriques, J. F. & de Miranda, S. S. Functional profile of patients with tuberculosis sequelae in a university hospital. *J. Bras. Pneumol. publicação Of. da Soc. Bras. Pneumol. e Tisilogia* **32,** 43–7 (2006).

105. Mittl, R. L. *et al.* Radiographic Resolution of Community-acquired Pneumonia. *Am J Respir Crit Care Med.* **149,** 630–635 (1994).

106. Coleman, M. T. *et al.* Early changes by 18F-PET-CT predict outcome after M. tuberculosis infection in cynomolgus macaques. *Infect. Immun.* **82,** 2400–2404 (2014).

107. Yuan, Y., Crane, D. D. & Barry, C. E. Stationary phase-associated protein expression in Mycobacterium tuberculosis: Function of the mycobacterial alpha-crystallin homolog. *J. Bacteriol.* **178,** 4484–4492 (1996).

108. Breen, R. A. M. *et al.* How good are systemic symptoms and blood inflammatory markers at detecting individuals with tuberculosis? *Int J Tuberc Lung Dis* **12,** 44–49 (2008).

109. Marais BJ, Schaaf HS, Hesseling AC, G. R. Tuberculosis case definition : time for critical re-assessment? *Int J Tuberc Lung Dis* **12,** 1217–1218 (2008).

110. Comstock, G. W., Golub, J. E. & Panjabi, R. Recurrent tuberculosis and its risk factors: Adequately treated patients are still at high risk. *Int. J. Tuberc. Lung Dis.* **11,** 828–837 (2007).

111. Nettles, R. E. *et al.* Risk factors for relapse and acquired rifamycin resistance after directly observed tuberculosis treatment: a comparison by HIV serostatus and rifamycin use. *Clin. Infect. Dis.* **38,** 731–736 (2004).

112. Malherbe, S. T. *et al.* Persisting positron emission tomography lesion activity and Mycobacterium tuberculosis mRNA after tuberculosis cure. *Nat. Med.* ePub (2016). doi:10.1038/nm.4177

113. Shenai, S. *et al.* Bacterial Loads Measured by the Xpert MTB/RIF Assay as Markers of Culture Conversion and Bacteriological Cure in Pulmonary TB. *PLoS One* **11,** e0160062 (2016).

Appendices

Appendix A Patient demographics, outcome and scan response pattern.

Px No.	Age	Sex	Missed Rx	M6 Sputum Xpert	Smoker	Outcome	M6 Response Pattern	Max M6 lesion intensity
1	30	F	9	Negative	Yes	Cure	Improved	Mild
2	21	M	33	Positive	Yes	UE	Mixed	High
3	66	M	0	Positive	Yes	Fail	Mixed	Very High
4	40	M	0	Negative	No	Cure	Resolved	Minimal
5	30	F	1	Negative	No	Cure	Improved	High
7	58	M	0	Positive	Yes	Cure	Mixed	High
9	19	M	12	Negative	Yes	Cure	Improved	Mild
13	48	M	0	Negative	Yes	Cure	Improved	Moderate
19	33	F	53	Positive	Yes	Fail	Mixed	High
20	53	F	7	Negative	Yes	Cure	Mixed	Moderate
21	46	M	21	Positive	Yes	Fail	Mixed	High
22	29	F	1	Negative	Yes	Recur	Improved	Very High
24	43	M	4	Negative	Yes	Cure	Mixed	High
27	28	M	27	Positive	Yes	Fail	Mixed	Very High
29	42	M	6	Negative	Yes	Cure	Resolved	Minimal
30	25	M	4	Positive	No	Cure	Improved	Mild
34	33	M	3	Positive	Yes	Recur	Resolved	Minimal
35	28	M	32	Negative	Yes	Cure	Mixed	Moderate
41	35	F	0	Negative	Yes	Cure	Resolved	Minimal
42	32	F	6	Positive	Yes	Cure	Improved	Moderate
43	18	M	0	Negative	No	Fail	Mixed	Moderate
49	21	F	0	Negative	No	Cure	Improved	moderate
50	38	M	0	Negative	Yes	Cure	Improved	High
51	52	M	7	Negative	Yes	Cure	Improved	High
52	44	F	0	Negative	Yes	Cure	Improved	High
53	31	M	17	Negative	Yes	Cure	Mixed	Moderate
54	55	M	0	Positive	No	Cure	Resolved	None
70	27	M	6	Negative	Yes	Cure	Mixed	Moderate
71	33	M	0	Negative	Yes	Cure	Improved	Mild
72	25	M	3	Negative	Yes	Cure	Improved	High
73	26	F	41	Negative	Yes	UE	Resolved	None
76	36	M	0	Negative	Yes	Cure	Resolved	Minimal
77	31	M	15	Negative	Yes	Cure	Improved	Mild
78	39	M	0	Negative	Yes	Cure	Improved	Mild
80	47	F	0	Negative	Yes	UE	Improved	Moderate
81	22	M	2	Negative	Yes	Cure	Mixed	High
83	22	M	5	Negative	Yes	Cure	Improved	High
84	57	M	0	Positive	No	Cure	Mixed	Very High
85	26	M	0	Negative	Yes	Cure	Improved	High
86	19	F	0	Negative	No	Cure	Improved	High
87	21	M	5	Negative	Yes	Cure	Improved	Mild
88	46	M	0	Negative	Yes	Recur	Mixed	High
89	25	F	82	Negative	Yes	Cure	Mixed	Moderate
90	23	F	5	Negative	Yes	Cure	Improved	High
91	28	M	7	Negative	No	Cure	Improved	Moderate
92	42	M	0	Negative	No	Cure	Mixed	High
93	30	M	0	Negative	Yes	Recur	Improved	High

95	46	M	4	Positive	Yes	Recur	Mixed	Moderate
98	46	F	0	Negative	Yes	Cure	Mixed	High
99	37	F	0	Negative	Yes	Cure	Mixed	High
100	55	F	9	Negative	Yes	Cure	mixed	Very High
101	28	F	12	Positive	Yes	Recur	Improved	High
102	39	F	0	Negative	Yes	Cure	Mixed	Moderate
103	43	F	0	Positive	Yes	Cure	improved	Moderate
104	42	M	0	Negative	No	Cure	Mixed	High
105	21	M	0	Negative	Yes	Cure	Resolved	Minimal
108	29	M	0	Negative	Yes	Cure	Improved	Moderate
111	22	M	0	Negative	Yes	Cure	Improved	Moderate
112	52	M	2	Negative	Yes	Recur	Mixed	High
123	40	M	0	Positive	Yes	Cure	Mixed	High
124	24	M	32	Negative	Yes	Cure	Improved	Mild
125	32	M	0	Negative	Yes	Cure	Improved	Mild
126	39	M	1	Negative	No	Cure	Improved	Moderate
128	22	M	15	Positive	Yes	Cure	Improved	High
129	39	M	9	Negative	Yes	Cure	Improved	Moderate
130	29	M	1	Negative	Yes	Recur	Improved	Mild
131	18	F	0	Negative	No	Cure	Improved	Very High
132	17	F	15	Negative	Yes	Cure	Mixed	High
133	33	F	10	Positive	Yes	Cure	Improved	High
134	24	F	0	Negative	Yes	Cure	Resolved	Minimal
135	17	M	2	Negative	No	Cure	Improved	Very High
136	41	M	0	Negative	Yes	Cure	Improved	High
137	44	F	0	Positive	Yes	Recur	Improved	High
138	23	M	11	Negative	Yes	Cure	Improved	High
139	38	M	31	Negative	Yes	Cure	improved	Moderate
140	22	M	25	Positive	Yes	Recur	improved	Mild
141	17	M	5	Negative	Yes	Cure	Resolved	None
142	43	F	0	Positive	Yes	Recur	Mixed	Very High
144	31	M	21	Negative	Yes	Cure	Mixed	High
145	19	F	11	Negative	Yes	Cure	Improved	High
146	20	M	0	Positive	Yes	Cure	Improved	High
147	23	M	0	Negative	No	Cure	Mixed	Very high
149	19	F	4	Negative	No	Cure	Improved	Moderate
150	21	F	56	Positive	No	Cure	Mixed	High
152	23	F	0	Negative	Yes	Recur	Improved	High
153	25	M	53	Positive	Yes	Cure	Resolved	Minimal
154	42	F	0	Positive	Yes	Cure	Improved	High
155	36	F	2	Negative	No	Cure	Resolved	Minimal
156	64	M	0	?	Yes	Cure	Mixed	High
159	19	F	0	Negative	No	Cure	Resolved	Minimal
160	40	F	0	Negative	No	Cure	Improved	Moderate
161	25	F	7	Positive	No	Cure	Improved	High
162	49	M	1	Positive	Yes	Cure	Improved	Moderate
163	25	M	3	Positive	Yes	Fail	Mixed	High
167	41	M	2	Negative	Yes	Cure	Improved	High
168	23	F	86	Positive	No	Fail	Mixed	Very High
169	30	F	91	Positive	Yes	Fail	Mixed	Moderate
170	30	F	33	Negative	Yes	Cure	Resolved	None
171	45	F	0	Positive	Yes	Cure	Improved	High

Appendix B EOT + 1y scan response patterns per patient.

Px ID	Outcome	EOT+1y response pattern	Max EOT+1y intensity	EOT+1y SUVmax	M6 Response Pattern	Max M6 intensity
76	Cure	Mixed	very high	6.5	Resolved	Minimal
77	Cure	Resolved	Minimal	1.6	Improved	Mild
78	Cure	Resolved	Minimal	1.5	Improved	Mild
80	UE	Improved	Mild	2.8	Improved	Moderate
81	Cure	Mixed	Moderate	3.2	Mixed	High
84	Cure	Improved	High	4.3	Mixed	Very high
85	Cure	Improved	Mild	2	Improved	High
86	Cure	Improved	Mild	2.3	Improved	High
88	Recur	Mixed	High	2.7	Mixed	High
89	Cure	Resolved	None	0.9	Mixed	Moderate
90	Cure	Resolved	Minimal	2	Improved	High
91	Cure	Mixed	Moderate	2.3	Improved	Moderate
92	Cure	Improved	Moderate	2.2	Mixed	High
93	Recur	Mixed	High	4.2	Improved	High
95	Recur	Improved	Moderate	2.4	Mixed	Moderate
98	Cure	Mixed	Very high	7.5	Mixed	High
99	Cure	Resolved	Minimal	1.9	Mixed	High
100	Cure	Improved	Mild	2.2	mixed	Very High
102	Cure	Resolved	Minimal	1.2	Mixed	Moderate
103	Cure	Resolved	Minimal	1.6	improved	Moderate
104	Cure	Resolved	Minimal	1.3	Mixed	High
105	Cure	Resolved	None	0.8	Resolved	Minimal
111	Cure	Resolved	Minimal	1.5	Improved	Moderate
112	Recur	Mixed	Very High	5.4	Mixed	High
124	Cure	Resolved	Minimal	1.1	Improved	Mild
125	Cure	Improved	Moderate	2.6	Improved	Mild
126	Cure	Resolved	Minimal	1.6	Improved	moderate
128	Cure	Mixed	High	4.7	Improved	High
129	Cure	Resolved	Minimal	1.2	Improved	Moderate
131	Cure	Resolved	None	0.6	Improved	Very high
132	Cure	Mixed	Moderate	2.8	Mixed	High
133	Cure	Mixed	High	4	Improved	High
135	Cure	Improved	High	4.6	Improved	Very high
137	Recur	Improved	Moderate	3.8	Improved	High
138	Cure	Improved	Mild	1.8	Improved	High
139	Cure	Mixed	High	2.4	improved	Moderate
140	Recur	Improved	Mild	1.6	improved	Mild
141	Cure	Mixed	High	2.1	Resolved	None
142	Recur	Mixed	Very high	7.8	Mixed	Very high
145	Cure	Improved	Moderate	2.1	Improved	High
146	Cure	Mixed	high	3.8	Improved	high
149	Cure	Improved	Mild	1.8	Improved	Moderate
152	Recur	Mixed	Moderate	1.4	Improved	high
154	Cure	Improved	Mild	1.8	Improved	high
155	Cure	Resolved	Minimal	0.6	Resolved	Minimal
156	Cure	Mixed	Moderate	2.1	Mixed	High
161	Cure	Improved	Mild	1.7	Improved	high
162	Cure	Resolved	Minimal	1.8	Improved	Moderate
167	Cure	Mixed	High	3.6	Improved	High
171	Cure	Improved	mild	2.1	Improved	High

Appendix C South Korea patient demographics, outcome and scan response.

Px ID	Age	Gender	Smoking	Previous PTB?	DS/MDR	Outcome	Adherence	M6 scan response pattern	Lesion intensity rank
A225	65	Male	Yes	No	MDR	Cured	Good	Resolved	Minimal
A226	63	Male	Yes	No	DS	Cured	Good	Improved	Moderate
A227	57	Male	Yes	No	R mono	Cured	Good	Mixed	Very high
A228	65	Male	No	No	DS	Cured	Good	Mixed	Very high
A230	57	Male	No	No	DS	Cured	Good	Mixed	Very High
A238	45	Female	Yes	No	DS	Cured	Poor	Improved	Moderate
A240	27	Female	No	No	DS	UE	Good	Mixed	Very High
B449	56	Male	Yes	Yes	DS	Recurrence	Good	Improved	Mild
B450	50	Male	Yes	Yes	DS	Cured	Good	Mixed	Very High
B452	66	Male	Yes	Yes	R mono	Cured	Good	Mixed	Very High
B453	24	Male	No	Yes	MDR	Cured	Good	Improved	Moderate
B459	61	Male	Yes	Yes	MDR	Cured	Good	Improved	Mild
B462	26	Male	No	Yes	DS	Cured	Good	Mixed	Moderate
B463	51	Male	No	Yes	DS	Cured	Good	Improved	High

New lesions: 4

2 Nodular infiltrates

1 Consolidation

1 Tree in bud

Old More intense

3 Cavity, 2 now consolidation

1 Nodule, now consolidation

1 Nodule, more intense

Appendix D Sputum MTB mRNA copy numbers and clinical status.

Pulmonary Tuberculosis patients

Patient	Outcome	85B	DOSR	PRCA	ICL	PSTS1	NUOB	Rv2623	ACR	CARD	TGS1	TB 8.4	ACPM	RRN	Total
41	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
42	Cure	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1
43	Fail	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	2
49	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
50	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
51	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
52	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
53	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
54	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
70	Cure	0.0	1.6	0.0	0.0	0.0	2.4	0.0	72.3	0.0	5.2	0.0	6.6	1.8	6
71	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
72	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
73	UE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
76	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
77	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	1.1	0.0	2
78	Cure	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	2
80	UE	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1
81	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.1	0.0	0.0	0.0	0.0	0.0	1
83	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
84	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
85	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
86	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0	0.0	0.0	0.0	4.9	2
87	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.3	0.0	0.0	0.0	0.0	0.0	1
88	Recur	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
89	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
90	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
91	Cure	0.8	0.0	0.0	0.0	0.0	0.0	0.0	6.0	0.0	0.0	0.0	0.0	0.0	2
92	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
93	Recur	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
95	Recur	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
98	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
99	Cure	0.0	0.0	0.0	0.0	7.7	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0	2
100	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
101	Recur	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		0.0	0.0	0
102	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	20.0	0.0	0.0	0.0	0.0	0.0	1
103	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
104	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
105	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	38.3	0.0	1.3	0.0	1.5	2.6	4
108	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.5	0.0	0.0	0.0	0.0	0.0	1
111	Cure	5.8	0.0	0.0	0.0	0.0	0.0	0.0	9.4	0.0	0.0	0.0	0.0	3.8	3
112	Recur	0.0	0.0	0.0	0.0	0.0	0.0	0.0	32.3	0.0	1.6	0.0	3.6	1.3	4
123	Cure	0.0	0.9	2.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.9	4
124	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.9	0.0	0.0	0.0	1.5	0.8	3
125	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
126	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
128	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	39.2	0.0	0.0	0.0	7.6	3.0	3
129	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
130	Recur	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
131	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.8	0.0	0.0	0.0	0.0	0.0	1
132	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
134	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
135	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.5	0.0	0.0	0.0	0.0	1.3	2
136	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
137	Recur	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
138	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
139	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
142	Recur	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	1
144	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.8	0.0	1
145	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
147	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.0	1
149	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.0	1
150	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
152	Recur	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
154	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
155	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
156	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
160	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
161	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
162	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
163	Fail	0.0	5.9	0.0	0.0	0.0	1.1	0.0	18.6	0.0	1.7	0.0	2.6	4.8	6
167	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
168	Fail	0.0	122.5	0.0	112.4	0.0	0.0	145.2	1584.5	1526.7	0.0	197.1	835.7	224.7	8
169	Fail	4.6	32.6	35.1	14.7	21.7	125.5	45.5	658.0	594.3	118.1	81.0	23.0	71.2	13
170	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
171	Cure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.3	0.0	1
	Total pos	4	6	3	2	2	4	2	20	2	5	2	16	12	29

Controls

ID	Sex	Age	Outcome	Quantiferon®	85B	DOSR	PRCA	ICL	PSTS1	NUOB	TB37.1	HspX	CARD	TGS1	TB 8.4	ACPM	RRN	Total
S6	Male	33	Community	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S10	Female	21	Community	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S11	Female	45	Community	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S18	Female	26	Community	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S36	Female	40	Community	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S37	Female	60	Community	negative	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S38	Male	21	Community	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S39	Female	24	Community	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S45	Female	34	Community	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S48	Male	25	Community	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1
S60	Male	26	Community	Negative	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S61	Male	23	Community	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S62	Male	22	Community	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S63	Male	20	Community	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S65	Female	48	Community	negative	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S66	Female	34	Community	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S67	Female	50	Community	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S68	Female	23	Community	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S181	Male	27	Community	negative	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S182	Male	27	Community	negative	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.7	0.0	1
S172	Female	60	OLD	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S173	Female	62	OLD	negative	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S175	Male	49	OLD	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S177	Female	28	OLD	Positive	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
S178	Male	26	OLD	Negative	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0
Total					0	0	0	0	0	0	0	0	0	1	0	1	0	2

Appendix E MTB mRNA copy numbers in Bronchoalveolar lavage and clinical status

End of treatment PTB patients

Patient	Outcome	recF	eccD3	menA	lprB	Rv1421	gabD2	Rv1910c	hspX	fadD9	tgs1	fadE34	Rv3675	sodA	trxB2	rpsJ	rplV	sigI	sigH	rpsK	Total
139	Cure	0	8	0	36	0	0	0	24	0	0	0	84	0	21	169	0	0	0	110	7
141	Cure	0	0	9	18	0	0	0	0	0	0	0	11	0	26	4	3	0	0	7	7
142	Recur	0	0	0	8	0	0	0	8	0	35	0	0	0	0	0	2	0	6	4	6
144	Cure	0	25	5	49	3	0	0	42	0	98	0	24	123	4	13	1	0	0	0	11
145	Cure	0	6	0	1	13	0	5	56	0	0	0	6	8	0	206	2	33	19	251	12
146	Cure	1	8	11	32	0	9	0	9	8	12	0	0	0	18	26	3	0	2	18	13
150	Cure	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	1
152	Recur	1	18	57	97	14	10	3	72	0	36	9	10	14	1	23	0	0	4	25	16
153	Cure	2	0	0	16	0	0	9	9	0	0	0	45	0	0	20	0	0	0	0	6
154	Cure	0	0	0	0	0	0	0	31	9	0	0	19	0	0	8	0	0	0	8	6
155	Cure	0	0	0	14	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0	2
159	Cure	0	2	0	3	0	0	0	19	0	16	0	1	64	9	21	0	0	0	0	8
161	Cure	0	2	0	0	0	0	0	43	0	0	0	23	0	0	26	0	0	0	7	5
162	Cure	2	0	13	29	24	20	0	32	7	37	0	43	0	10	15	5	40	0	31	14
163	Fail	136	64086	780	79173	3463	3724	2813	886470	6373	64980	453	19725	158895	22423	91261	9794	1358	7047	132232	19
		5	8	6	13	5	4	4	14	4	7	2	11	5	8	12	7	3	5	10	15

Controls

OLD	Sex	Age	Diagnosis	Quantiferon®	recF	eccD3	menA	lprB	Rv 1421	gabD2	Rv1910c	hspX	fadD9	tgs1	fadE34	Rv3675	sodA	trxB2	rpsJ	rplV	sigI	sigH	rpsK	Total mRNA's
CA1	F	78	Infection	Pos	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CA2	F	80	New PTB	Pos	0	75	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	5
CA3	F	56	AdenoCA	Pos	0	0	0	0	61	0	0	217	0	89	0	0	90	0	0	0	0	0	0	1
CA4	M	51	Cancer	Pos	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	134	1
CA5	F	45	Cancer	Pos	0	0	0	0	0	0	0	0	0	0	0	0	0	0	62	0	0	0	0	1
CA6	F	53	Pneumonia	Pos	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CA7	M	68	Cancer	Neg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CA8	F	45	Interstitial fibrosis	Neg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CA9	M	63	Pneumonia	Neg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CA10	F	75	Cancer	Neg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
					0	1	0	0	1	0	0	1	0	1	0	0	1	0	1	0	0	0	2	

Appendix F Primer sequences for MTB mRNA

List of Primer sequences for transcription of MTB RNA targets used for RT-PCR in M6 sputum. (Shubhada and Alland)

No	Targets	Primers and Molecular Beacon Sequences
1.	DOSR	DOS R RT - CCC TTG ATG TCT TTG ACG ACA DOSR F - GAT CCT CAC GTC CTA CAC CTC DOSR R - GAC GAC ATA TCC GCT GGC ACC DOSR MB - AGCCG CTGACGAGGCCATGCTAGATG CG TCGGC
2.	85B	85 B RT - GGG TGC CGT TCC CGC AAT AAA 85B F - CAA CGA CCC TAC GCA GCA GAT C 85B R - TTC CCG CAA TAA ACC CAT AGC 85B MB - CGCTG CAAGCTGGTCGCAAACAACACCCG CAGCG
3.	ICL	ICL RT - GCC GCG AGC CGA GCA GAC GT ICL F - CTG AGA AGA AGT GCG GCC AC ICL R - AGA CGT CAA AGT GCG GAT GTG ICL MB - CTC GGGCGGCAAGGTGTTGATCCC GAG
4.	NUOB	NUOB RT - AGC GCG TTC CCG GTT GAT ACC NUOB F - GCT GCA CGC AAT CCT GAA GCT NUOB R - CCG GTT GAT ACC TAA TGG CAT C NUOB MB - CGC GCACGAAAAGATTCAGCAGATGC GCG
5.	ACPM	ACPM RT - GTC GAT GTC CAG GTC GTC GA ACPM F - CCG GTA TCG AGC CGT CCG AGA ACPM R- GGT CGT CGA CGA ACG ACT TC ACPM MB - ACGC GAGATCACCCCGGAGAAGTC GCGT
6.	PSTS1	PSTS1 RT - TCG AAA TCG CCT GGT TCG CCG PSTS1 F - GCC CGA CGC GCA AAG CAT TCA PSTS1 R - GCC TGG TTC GCC GGG GTT TTC PSTS1 MB - CCACC GCGGCTGGCTTCGCATCGAAAAC GGTGG
7.	PRCA	PRCA RT - GCA GCG ACC GCG ATA CGC AG PRCA F - GAT CGC CAA CGC GCT CAA AGA PRCA R - CGC AGG GCG TCG GTC AGG CT PRCAMB - CCAGC GAGTCGTATGCCGAGAACGCC GCTGG
8.	HspX (ACR)	ACR RT - GGT GGC CTT AAT GTC GTC CTC GTC ACR F - GAC GGT CGC TCG GAA TTC GCG TAC ACR R - CGT CCT CGT CAG CAC CTA CCG GC ACR MB - CGCTCC CCTTCGTTCGCACGGTGTCGCTG GAGCG
9.	CARD	CARD RT - CGG GGC ACG CAA CAC CTG GAA CARD F - CGA CCT GAC AGT ACG AGT TCC C CARD R - GCC TTC CTG CCC GAC GAC ATC CARD MB - AG CCCGCTGAAAACGCCGAATACG GGCT
10.	TGS1	TGS1 RT - CGA GAC CGG CAC TAG CGT ACG TGS1 F - ATG TTG CGC TTG CCG CGA TTA CGG TGS1 R - CGA ATC AAA CCT AGG CCG CTC ACC TGS1 MB - ACGC CTACCGCAACGTCCTCATCCAG GCGT
11.	TB 8.4	TB 8.4 RT - GGT GTT AAT GAC CGC GTC CAC G TB 8.4 F - CGC ATT GAG CGC CGG TGT AGG C TB 8.4 R - CTG CGG AGG CGA CCC CGG CCC TB 8.4 MB - ACG CCGTGGCAATGTCGTTGACCG GCGT
12.	Rv2623	RV2623 RT - CCA CCG CCC ACT TCC GAG ACA AC Rv2623 F - GGC AGC CGT TCC CAC ATT GG Rv2623 R - CCG AGA CAA CCC ACG ACC AT

		Rv2623 MB - CGACC GGTCGACATGTCCAAAGACGC GGTCG
13.	SIGA	SIGA RT - ATC TGG CGG ATG CGT TCC CGG SIGA F - GGC CAG CCG CGC ACC CTT GAC SIGA R - CGG ATG CGT TCC CGG GTC ACG SIGA MB - CGC ACG AGA TCG GCC AGG TCT ACG GCG T GCG
14.	RRN	RRN RT - TTC TCA AAC AAC ACG CTT GCT TG RRN F - CCT ATG GAT ATC TAT GGA TGA C RRN R - GCA ACC CTG CCA GTC TAA TAC AA RRN MB - ACGC CCT GTT CTT GAC TCC ATT GCC GG GCGT

Primer sequences used for qRT-PCR of BAL samples. (Schoolnik, Dolganov, Van)

Gene	**Pre-Amplification Primer_RTF**	**Pre-Amplification Primer_RTR**	**Taqman_TMF**	**Taqman_TMR**	**Taqman_TMP**
recF	GAGCGTGGGGTTTGTCTAGTT	CACCATCAACGCGTAACAGT	CGCACCGTGACGACCTAATAC	CCCCATGGCTAGCAAATCCT	5'-FAM-CGACTAGGCGATCAACCCGCGA-BHQ1
eccD3	GGTGATCTGTTGGTGTTGCAG	ATAGGTCACCGTCAAACCGGTAG	CGACGCCGCGATGATC	TCGTTGGATATGCGCTATGC	5'-FAM-TTTCGACGTCGCGGTTAAAGCCC-BHQ1
menA	GGGGCCTGGCTCTACACC	ATGCCTGCGTGTACTGGGTA	ACACCGGCGGGTCAA	GGCCCGAAGAACACAAACAC	5'-FAM-CCCTACGGCTATGCGGGCTTCG-BHQ1
pabC	CTCCGTGGTATCCAATCCTG	CATGCTCGATACCAACCAAAT	CGAAAGGCTACGACTGCGACTA	CAAATACCTTGGGAATCGAAGAGA	5'-FAM-CGTGCCCTACGCG

					TCGCCG-BHQ1
accD3	CTGATCGGCTTTCTGGGAC	GTCGATGAGCACCGTCAAC	ACGGGTCTATGAGTTGCTCTATGG	CCATGCCGCCGTAGATTCT	5'-FAM-CCCATCCGGCGTCCAAACCG-BHQ1
lprB	TCGAGGACATCAACATCGAC	AAGTCGTCGGAGAACTGGATT	ACGGCCACAGCGGTTTC	TTCCGACTTCACACAGTGAGTCA	5'-FAM-TCGCCATCGGTAACGAGCCCAGT-BHQ1
Rv1421	GGTATCGACGTCGTCTTGGTAAC	CAGCTGGGTGATCCGTGA	GAAGACCTGGGCTGGTATGTG	CCACCATGCGGGTAATCAG	5'-FAM-CCGACAATCTGCCGCCCCA-BHQ1
gabD2	AGAACCGCGAGTTCCTCAT	ACACCGTGGTCTTGCCTATC	GCGCAAGAGGAAATTGTCGAT	GCTTCAGCAGGTCCACACAGA	5'-FAM-CGCGAACGCGAATTATTACGCACGA-BHQ1
Rv1910c	GGAACAGTACACCTGCAAAGG	GAGTTCGGCAGGCTGATTC	GCGGCGCACTCGTTGT	CACGATCCAATGGACGTAAGGT	5'-FAM-TGATCCGGACGCACCTCGCG-BHQ1
hspX	TGAAAGAGGGGCGCTACGAG	CCGCCACCGACACAGTAAGA	ACATTATGGTCCGCGATGGT	CGACCGTCGAAGTCCTTCTG	5'-FAM-TGACCATCAAGGCCGAGCGCA-BHQ1
lipX	CACGATCTTGTGTCCATCGT	ACGAAGTCTTGCTGGTCTGAA	GCACGGCGAGCTGTACAAG	TGGACAAATGACTCGTGGATCA	5'-FAM-ATCGCGCGCCAGGCCCA-BHQ1
fadD9	CATGTATCGCGAGAGCCAG	GCGACGAAGTAGGCGGTA	ATGAGCTTCTGGCGCAAGTC	GTTGGAAAGCGTCCCGTAGAG	5'-FAM-CTGAACTTCATGCCGATGAGCCACG-BHQ1
tgs1	CGAACAGGTGTGCCGGAAAT	GCAGCGGGTTCTCTTGATCC	TGCCGCGATTACGGAAAG	GCGTACGCAGCGAATCAAA	5'-FAM-TACCGCAACGTCCTCATCCAGCG-BHQ1
fadE34	GCCGACCATCCTTGAACAC	GACGTCCACACCTTCTGTCC	GCCGGAACAGATCGAACGTT	CCCGGCTCGGAAAATAGCT	5'-FAM-CCACCATGCGCGGTGAATTCCTTT-BHQ1
Rv3675	GATGTGCATACGCTCGCTC	ATCGAGATCGGTGACAAACAG	TCGCTCGGAATGGAATGC	CAAGCGGCGGTGAATCG	5'-FAM-ACCTGCATCGACGTCAAGCCCG-BHQ1
sodA	GGACTACGGAGCACTGGAAC	TCTTTTCGTTCAGCAAGATCG	TCGGGTCAGATCAACGAGCTT	TCATTGGCGCCCTTTACGT	5'-FAM-CCACAGCAAGCACCACGCCACC-BHQ1
trxB2	AGCTACCTTCCTGACCCGAT	AGCCGGTCACTGTGGTGT	CGGGCTTCCAAAATCATGCT	GTGGTTGGTGAGGAACCGTATC	5'-FAM-TCGCGCCCGCAACAACGAC-BHQ1
rpsJ	AGAACGTGTATTGCGTCATCC	GTTGACGTCGACGCTGG	GGACTCGCGGGAGCACTT	GGCGTGGGATCGATGATG	5'-FAM-TGCGCACACACAAGCGGTTGATC-BHQ1
rplV	GCGGCTACTAAGGCTACCGAGTAT	GTGGCCACCACCAAGGTT	TGGCAGGTCGGTGTCAGA	GCACTGGCGATCACTTTGG	5'-FAM-CGCTCGACATCCTGCGCTGG-BHQ1
sigI	GCTGCCCTACCAGCAAATC	AGCAGGGTGTCCAGGTCTC	CAACGAATCGCGCATTGC	TTGGAGCAGGCTTCGATGAA	5'-FAM-CCAGCCCAGCATCGCGTCGT-BHQ1
TRP	GCAACACTGTATCGGTACTTCGAC	GTAAACCACGCGGCCAAC	TGGTGAAGATCGCCCGATGT	GCAACGTCGTGGTGATGCT	5'-FAM-CGAACCTGCCGAACGGCTGC-BHQ1
sigB	AATCCGGCATTCCAATCGA	GATGAAATCGCCCAAAGGG	TTCCAATCGACAAGATCAACGA	GCCGACCGGCATATCCA	5'-FAM-CTGGAACACAGTCGCGACCCGG-BHQ1
sigH	AATCTCAAGGCCTGGCTCTAC	GCTTCTAACGCTTCGACTTCA	CATCAACAGCTATCGCAAGAAACA	CAGTTGCCAATCGGTGATCTG	5'-FAM-CAACCGGCGGAGT

					ATCCGACCG-BHQ1
rpsK	GGAGAAGAAGAACGTCCCGC	GGGTCGATTTCCGGGAAC	CACGTTCAACAACACGATCGT	TTTCCGGGAACCCTTGAAG	5'-FAM-CAACGTCATTGCCTGGGCATCG-BHQ1