



Performance of host blood transcriptomic signatures for TB diagnosis and monitoring TB treatment

Vanessa Mwebaza Muwanga

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Pathology

Faculty of Health Sciences

UNIVERSITY OF CAPE TOWN

On the 28th of June 2024

Supervisor: Professor Thomas J. Scriba

Co-supervisor: Dr Simon C. Mendelsohn

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Declaration

I, Vanessa Mwebaza Muwanga, hereby declare that the work on which this thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is submitted for another degree in this or any other university.

I empower the university to reproduce, for the purpose of research, either the whole or any portion of the contents in any manner whatsoever.

Inclusion of publications

I confirm that I have been granted permission by the University of Cape Town's Doctoral Degrees Board to include the following publication in my thesis and my co-authors have agreed that I may include the publication:

Muwanga, V., Mendelsohn, S.C., Leukes, V., Stanley, K., Mbandi, S.K., Erasmus, M., Flinn, M., Fisher, T.-L., Raphela, R., Bilek, N., Malherbe, S.T., Tromp, G., Van der Spuy, G., Walzl, G., Chegou, N.N., Scriba, T.J., 2024. Blood transcriptomic signatures for symptomatic TB in an African multi-cohort study, in press, Eur.Res.J.

Signature:

Signed by candidate

Date: 28th June 2024

Name: Vanessa Mwebaza Muwanga

Student number: MWNVAN001

Acknowledgements section

I want to take this opportunity to officially thank everyone who has walked with me through this PhD in one way or another. One set of individuals would be my fitness crew i.e. all the folks I have hiked with, ran with, walked with, gone on camps with and done any form of exercise with. It is still shocking that despite how busy my academic life got, I even managed to do a half marathon. I never even knew I could be a fitness ninja, let alone even run. I did my first 5 km with the UCT UCanToo group during COVID.

Another set of individuals I want to thank for the push that I needed to work at times when I wasn't feeling motivated or focused enough are my writing accountability circles: the group of black women in South Africa, the Cohort Sistas initiative by the ever amazing Professor Ijeoma Kola, the manifest network headed by a friend Rutendo Nyaku and then the little hustle group I set up for myself at the start of 2024. The Shut Up and Write sessions that Taahira Hoosen set up introduced me to Richmond Ateko and Raymond Moseki. We fought like mad in 2022 and wrote together as our lives depended on it. Being part of the 2023 UCT 'writing for completion' program was another great push for me and really helped me learn and fight how to put aside dedicated work hours for my thesis. I want to thank Donald Simon for being part of my 4am - 6am writing bonanza in 2023. Kennedy Kipkoech Mutai is another beautiful soul I am so glad was in the same flat complex as me.

My family, particularly my parents: Dr Fred T Muwanga and Mrs Edith K Muwanga and my siblings are the MVPs that have dealt with all my drama, demands, emotional phases, rantings and round-the-clock work hours throughout this PhD process. I want to thank them for being able to be there for me in whatever way I needed them to be. I am home once a year but despite that, I feel like WhatsApp, Microsoft Teams and Google meets have been the gifts that have allowed us to keep in touch and still be part of each other's lives in ways we have never been in before. It takes two to tango, so I am glad for their willingness to hop on and be part of whatever madness I proposed.

I want to thank my supervisors, Tom and Simon for taking a chance on me. I did not think I even had what it took to make it, but I was willing to give things a shot and find out all that I could about what it takes to make it in academia, but most importantly contribute to the field of TB biomarkers. The SATVI Fluidigm team, Rodney, Zwandile and Lungi stood with me through COVID and all the drama that comes with trying to do labwork within that period.

Moving to another country that you have not lived in before in your late twenties as a postgrad can sometimes mean a little more effort has to be made in order to find friends and a community. As an introvert, I did not feel up to it but knew that it was something I had to do if I wanted to do more than just survive this PhD. I want to thank God for these amazing people who I have gotten to do life with through the course of this PhD: Seno Nilonga Namwandi, Tebatso Koma, Sibongile, Thato, Rutendo Nyaku and Wanjiru.

Last, but not least, the crew at SATVI who saw me through the starting years of this PhD and did their best to get me acquainted with everything at UCT and in the research group. They also handled all my panic and anxiety-filled questions. Thank you to Dr.Cheleka Mpande, Dr Melissa Murphy, Dr Carly Young-Bailie and Dr Fatoumatta Darboe. The code of conduct and solidarity you believe should exist among SATVI students is something I have strived to make sure has been passed on to the next set of students here. The two human beings who have dealt with all my R and data analysis cries: are Miguel Rodo and Dr.Denis Awany. I want to thank both of them for saving my life and the life of this project more times than I can count. The patience and time they have put into helping me get through my errors and code setup has been award-winning. It takes a village to get data analysis done.

Thesis summary

Performance of host blood transcriptomic signatures for TB diagnosis and monitoring TB treatment

Non-sputum-based diagnostic tests are necessary to enable tuberculosis (TB) diagnosis and monitoring of TB treatment.

This work aimed to identify which published blood transcriptomic TB signatures have the best diagnostic performance to distinguish between TB cases and other respiratory diseases (ORDs) in symptomatic adults. These signatures were also evaluated to monitor treatment in young children, to identify signatures that might indicate when treatment could be stopped.

Diagnostic performance of signatures was evaluated in adults presenting for care with symptoms associated with TB, who were recruited from primary healthcare clinics in six African countries. To identify signatures that can monitor treatment response, the same set of signatures was evaluated in young children randomised to 4- or 6-month TB treatment in the SHINE trial. Twenty transcriptomic signatures were selected and measured in whole blood using multiplex, microfluidic RT-qPCR.

In the adult diagnostic cohort, nine signatures achieved equivalent performance for differentiating patients with ORDs from all TB cases. Factors associated with signature scores included HIV and country. With sensitivity benchmarked at 90%, these nine signatures achieved specificities between 44%-54%. In pooled analyses, none of the signatures met the minimal target product profile criteria for a TB triage test. Country-specific analyses, however, showed that several signatures met the minimal criteria in some countries.

In children from the SHINE trial, baseline scores for all signatures were highest in children with confirmed, relative to unconfirmed and unlikely TB. Baseline scores were also higher among children with disease classified as more severe on chest x-ray. Scores declined progressively during TB treatment in the confirmed and unconfirmed TB groups; no changes in scores were observed for most signatures in the unlikely TB group. Scores were higher at the end of treatment in the 4-month than the 6-month treatment arm in the confirmed TB group; no differences were observed between treatment arms in the unconfirmed and unlikely groups. These results suggest that host blood transcriptomic signatures have potential to monitor TB treatment response in children.

This work supports further development of transcriptomic signatures as TB triage tests and treatment monitoring tools.

List of abbreviations

3PR - 3-protein pair-ratio

ACS - Adolescent Cohort Study

AE-TBC - African European Tuberculosis Consortium

AUROC - Area Under the Receiver Operating Characteristic Curves

BCG - Bacillus Calmette-Guerin

CAD - Computer-Aided Detection

CAD4TBv6 - Computer-Aided Detection for Tuberculosis version 6

CCC - Concordance Correlation Co-efficient

cDNA - complementary Deoxyribonucleic acid

CI - Confidence Interval

CRP - C-reactive protein

CTBC - Cross-sectional TB Cohort

CV - coefficient of variance

CXR - Chest X-ray

DEG - Differentially expressed genes

DR-TB - Drug resistant TB

DS-TB - Drug susceptible TB

DTTC - Desmond Tutu TB Centre

ECDC - European Centre for Disease Prevention and Control

ECG - electrocardiogram

ELISA - Enzyme-linked immunosorbent assay

ELISpot - enzyme-linked immunospot

EpiScript RT - EpiScript™ RNase H- Reverse Transcriptase

FACS - Fluorescence-activated cell sorting

FDG-PET/CT - [18F]-fluoro-2-deoxy-D glucose positron emission/computed tomography

FDR - false discovery rate

FIND - Foundation for Innovative New Diagnostics

GC674 - Grand Challenges 6-74

HC - Healthy control

HHC - household contact

HKG – Housekeeping gene
 HREC - Human Research Ethics Committee
 IFC - integrated fluidic circuit
 IFN - Interferon
 IGRA - Interferon Gamma Release Assay
 IPC - Internal Positive Control
 IRF - IFN regulatory factor
 ISG - Interferon Stimulated Gene
 LAM - Lipoarabinomannan
 LED - light-emitting diode
 LF-LAM - lateral flow lipoarabinomannan
 LJ - Löwenstein-Jensen
 LTBI - Latent TB infection
 MBT - Multiple Biomarker Test
 MDP - molecular degree of perturbation
 MDR - Multi-Drug Resistant
 MGIT - Mycobacteria Growth Indicator Tube
 mRNA - messenger Ribonucleic acid
Mtb - *Mycobacterium tuberculosis*
 MTBC - *Mycobacterium tuberculosis* complex
 MVLs - mercury vapour lamps
 NAATs - Nucleic acid amplification tests
 NGS - Next-generation sequencing
 NoRT - No Reverse Transcriptase
 NTC – No Template Control
 NTM - Nontuberculous mycobacteria
 OD - Other Disease
 ORD - Other Respiratory Disease
 PBMC - Peripheral blood mononuclear cell
 PET-CT - Positron emission tomography/ computed tomography
 PLWHIV - people living with HIV

POC - Point of care

POCUS - point-of-care ultrasound

PPD - purified protein derivative

PPMMs - primer-probe master mixes

QDs - quantum dots

QFT - QuantiFERON-TB

RIF - rifampicin

RNA-seq - RNA sequencing

ROC - Receiver Operating Characteristic

RR-TB - Rifampicin-resistant TB

RT-qPCR - quantitative reverse transcriptase polymerase chain reaction

SAM - Severe Acute Malnutrition

ScreenTB - The Evaluation of Host Biomarker-based POC Tests for Targeted Screening for Active TB

SD - standard of deviation

SHINE - Shorter Treatment for Nonsevere Tuberculosis in African and Indian Children

SSM - sputum smear microscopy

TB-HIRA - TB-Host Immune Response Assay

TBNET - TB Network European Trials group

TPP - Target Product Profile

TPT - TB preventive treatment

TRM5 - TB Risk Model 5

TST - Tuberculin Skin Test

VOCs - volatile organic compounds

WGCNA - Weighted gene co-expression network analysis

WHO - World Health Organization

XDR - Extensively drug-resistant

Xpert-MTB-HR - Xpert MTB Host Response

ZN - Ziehl-Neelsen

List of Figures

Figure 2.1: BioMark™ 96.96 Dynamic Array IFC for gene expression.	70
Figure 2.2: Heatmap of gene expression.	72
Figure 3.3: QC Workflow.	73
Figure 3.1: Primers for PCR reactions. PCR reactions require two complementary oligonucleotide primers.	85
Figure 3.2: Representative primer-probe assays that met the minimal pass criteria for assay qualification.	90
Figure 3.3: Representative primer-probe assays that failed assay qualification.	91
Figure 3.4: Representative primer-probe assays that were selected for the panel despite not meeting the criteria.	92
Figure 3.5: Signature score distribution for Kaforou25.	97
Figure 3.6: Signature score distribution for Rajan5.	98
Figure 3.7: Sample pass rate of transcriptomic signatures during reparameterization.	100
Figure 3.8: Evaluation of primer-probe assay performance for Kaforou25.	103
Figure 3.9: Part of the heatmap for reparameterization run1 on CTBC samples.	109
Figure 3.10: Concordance correlation co-efficients for IPC_LOW.	110
Figure 3.11: Concordance correlation coefficients for IPC_MEDIUM.	111
Figure 3.12: Concordance correlation coefficients for IPC_HIGH.	112
Figure 3.13: Evaluation of RISK6 scores in IPCs.	113
Figure 4.1: Batch correction and normalization of signature scores.	125
Figure 4.2: Flow diagram of participant enrolment, allocation, and analyses.	127
Figure 4.3: Sample pass rate in the Screen-TB and AE-TBC sub-study.	130
Figure 4.4: Signature score distributions and diagnostic performance of transcriptomic signatures in TB cases and controls with other respiratory diseases (ORD).	134
Figure 4.5: Summary of signature diagnostic performance, before and after batch correction, for all TB cases versus controls with other respiratory diseases.	136
Figure 4.6: Comparison of diagnostic performance between candidate transcriptomic signatures.	137
Figure 4.7: Signature scores for most signatures are highly correlated.	138
Figure 4.8: Signature score distributions and diagnostic performance of transcriptomic signatures for all TB cases and controls with other respiratory diseases (ORD), stratified by country.	145
Figure 5.1: Flow diagram of participant enrolment, allocation, and analyses.	164
Figure 5.2: Signature pass rates in the SHINE sub-study.	167
Figure 5.3: Signature score distributions and performance of transcriptomic signatures as treatment monitoring biomarkers at baseline in confirmed, unconfirmed, and unlikely TB cases.	171
Figure 5.4: Signature score distributions at baseline stratified by chest x-ray severity.	173
Figure 5.5: Change in signature scores during treatment in children with confirmed, unconfirmed, and unlikely TB.	176
Figure 5.6: Signature score distributions during treatment stratified by TB classification.	177
Figure 5.7: Signature score distributions at the end of treatment stratified by treatment arm. ...	178
Figure 5.8: Performance of transcriptomic signatures as treatment monitoring biomarkers in children with confirmed TB.	180
Figure 5.9: Performance of transcriptomic signatures as treatment monitoring biomarkers in children with unconfirmed TB.	181
Figure 5.10: Performance of transcriptomic signatures as treatment monitoring biomarkers in children with unlikely TB.	183

Figure 5.11: Summary of performance of transcriptomic signatures as treatment monitoring biomarkers in children with confirmed TB, before and after batch correction.....	185
Figure 6.1: Innovations in microfluidic lab-on-a-chip systems for molecular diagnostics.....	201

List of Tables

Table 1.1: Characteristics of studies on transcriptomic signatures for active TB	27
Table 1.2: Approved drugs for TB treatment	35
Table 1.3: Current methods for TB treatment monitoring	37
Table 2.1: Participant cohorts analysed in this thesis	66
Table 3.1: Panel of transcriptomic signatures	78
Table 3.2: Representation of transcriptomic signatures across different systematic reviews	79
Table 3.3: Panel of HKGs from the CORTIS study	81
Table 3.4: Pass rate of HKGs based on results from CORTIS data.	82
Table 3.5: Rankings of the 10 HKGs from all the CORTIS runs based on results from four different methods for selection of HKGs.	83
Table 3.6: Layout of samples for pre-amplification in preparation for assay qualification	87
Table 3.7: Dilution series for assay qualification	88
Table 3.8: Primer-probe assays retained on the panel.	93
Table 3.9: The 96 primer-probe assays included in the final gene expression panel.	94
Table 3.10: Transcriptomic signatures on the panel that were reparameterized	96
Table 3.11: Performance of signatures that required reparameterization.	99
Table 3.12: Sample pass rate during reparameterization for the selected 96 primer-probe assays	101
Table 3.13: Number of failed reactions for Kaforou25 transcripts	103
Table 3.14: TaqMan PCR primer-probe assays for transcripts evaluated in this study	105
Table 3.15: Primer-probe assays for each IPC with gene expression results that did not pass quality control	108
Table 4.1: Diagnostic classification of participants in the ScreenTB and AE-TBC cohorts	123
Table 4.2: Participant characteristics	128
Table 4.3: Assay pass rate	131
Table 4.4: Diagnostic performance of candidate signatures for definite TB, probable TB, and all TB cases vs ORD controls.	135
Table 4.5: Diagnostic performance of 20 published host response TB transcriptomic signatures, ordered by area under the ROC curve for all TB patients vs ORD controls.	139
Table 4.6: Regression analyses determining clinical and demographic variables associated with signature scores.	141
Table 4.7: Numbers of TB cases and ORD controls, by HIV status.	142
Table 4.8: Diagnostic performance of candidate transcriptomic signatures for all TB cases vs ORD controls, disaggregated by HIV status.	143
Table 4.9: Numbers of TB cases and ORD controls, by country.	145
Table 4.10: Diagnostic performance of candidate transcriptomic signatures for all TB cases vs ORD controls, disaggregated by country.	146
Table 5.1: Classification of TB for children in the SHINE study	160
Table 5.2: Whole blood PAXgene tube samples available from study sites in Pune (India), Cape Town (South Africa) and Kampala (Uganda).	161
Table 5.3: Participant characteristics	166
Table 5.4: Sample pass-rate of individual transcripts in the SHINE sub-study	168
Table 5.5: Performance of candidate transcriptomic signatures in differentiating between the different TB classification groups at baseline.	172
Table 5.6: Regression analyses determining clinical and demographic variables associated with signature scores at baseline	174

Table 5.7: Regression analyses determining clinical and demographic variables associated with signature scores at baseline, including weight for age z scores at baseline	175
Table 5.8: Performance of candidate transcriptomic signatures in confirmed TB cases for discriminating between baseline and week 2, week 8, and end of treatment time points.....	180
Table 5.9: Performance of candidate transcriptomic signatures in unconfirmed TB cases for discriminating between baseline and week 2, week 8, and end of treatment all time points.	182
Table 5.10: Performance of candidate transcriptomic signatures in unlikely TB cases for discriminating between baseline and week 2, week 8, and end of treatment all time points.	184
Table 6.1: WHO performance benchmarks for a triage or referral test for TB.....	197
Table 6.2: WHO performance benchmarks for a rapid biomarker test for TB.....	197
Table 6.3: Desired characteristics of a biomarker for tuberculosis in children.....	205
Table 6.4: Priorities for paediatric TB biomarker development.....	205
Table 6.5: WHO TPPS for tests aimed at TB treatment monitoring and optimization.....	206

Contents

Declaration.....	ii
Inclusion of publications.....	iii
Acknowledgements section	iv
Thesis summary	v
List of abbreviations	vi
List of Figures	ix
List of Tables	xi
Chapter 1: Introduction	1
Section 1: Transcriptomics and its intersection with the study of TB.....	2
General overview of methods used to study transcriptomics	2
Transcriptomics enhancing our knowledge of the immune response to TB.....	4
Transcriptomics and the spectrum of TB disease	8
Section 2: Point of care diagnostic tests for TB	13
Current point-of-care or near-point-of-care tests	13
Smear microscopy.....	13
Sputum culture	15
Nucleic acid amplification tests (NAATs)	15
Chest X-ray (CXR)	16
Lipoarabinomannan (LAM).....	17
Development of point-of-care diagnostic biomarkers for TB.....	19
Host sensitization to <i>Mtb</i> antigens	19
Human antibodies against <i>Mtb</i> antigens	21
Human proteins	22
Transcriptomic signatures	25
Section 3: Current issues surrounding TB treatment.....	33
Overview of TB treatment.....	33
TB treatment shortening	34
Monitoring TB treatment	36
Need for non-sputum based biomarkers of treatment response	38
Transcriptomics for monitoring responses to TB treatment	39
Summary.....	40
Section 4: Aims and Objectives	41
References	44
Chapter 2: Materials and methods	66

2.1 Institutional ethics review	66
2.2 Collection of whole blood samples	66
2.3 RNA extraction, processing and storage	67
2.4 Measurement of gene expression.....	68
2.4.1 cDNA synthesis	68
2.4.2 Pre-amplification PCR.....	69
2.4.3 Confirmatory PCR for quality control of the pre-amplification process	69
2.4.4 High-throughput Gene Expression Measurement with RT-qPCR.....	70
2.5 Quality control (QC) procedures for gene expression data	72
2.6 Contributions	74
References	75
Chapter 3: Optimization of the panel of transcripts.....	77
3.1: Introduction	77
3.1.1 Selection of transcripts for the panel.....	77
3.1.2 Selection of Housekeeper genes	81
3.1.3 Selecting and designing primer probe assays	85
3.2: Assay qualification.....	86
3.2.1 Sample selection	86
3.2.2 Pre-amplification	87
3.2.3 Dilution series on pre-amped cDNA.....	88
3.2.4 Gene expression measurement	88
3.2.5 Criteria used to qualify assays.....	89
3.2.6 Outcomes from assay qualification and further selection of assays for the 96-transcript panel	89
3.2.7 Conclusions from assay qualification.....	93
3.3: Reparameterization.....	96
3.3.1 Laboratory setup for reparameterization	99
3.3.2 Pass rate of signatures during reparameterization	100
3.3.3 Conclusions from reparameterization	104
3.4: Development of an internal positive control (IPC).....	108
3.4.1 Criteria for selecting an IPC	108
3.5: Contributions	114
References	115
Chapter 4.....	119
Blood transcriptomic signatures for symptomatic TB in an African multi-cohort study	119
4.1: Abstract.....	120

Background	120
Methods	120
Results:	120
Conclusion:	120
Keywords	120
4.2: Introduction	121
4.3: Methods	122
Study design	122
Ethics statement	124
Measurement of signatures	124
Blinding	126
Statistical analysis	126
Missing data	126
4.4: Results	127
Selection and matching of study participants	127
Pass rate of transcriptomic signatures	130
Nine signatures had indistinguishable performance in discriminating TB cases and ORD controls	133
No signatures met the minimal WHO TPP benchmarks when cohorts are pooled	138
HIV status and country are associated with signature scores	140
4.5: Discussion	147
4.6: Contributors/Author Contributions	151
Declaration of interests	151
Data sharing	151
Acknowledgements	151
References	152
Chapter 5: Evaluation of transcriptomic signatures as treatment monitoring tools for pulmonary paediatric TB	158
5.1 Introduction	158
5.2 Methods	160
Study design	160
Ethics statement	161
Transcriptomic signatures	161
Signature measurement	162
Blinding	162
Statistical analyses	163

5.3: Results	164
Selection of study participants	164
Study participant characteristics differ based on TB classification	165
Pass-rate of transcriptomic signatures	167
Signature scores at baseline distinguished between confirmed and unlikely TB cases	170
Signature scores increase with CXR severity	173
Participant characteristics associated with signature scores at baseline	174
Changes in scores during treatment are greater in confirmed and unconfirmed TB groups	176
Children with confirmed TB showed differences in scores at the end of treatment	178
Transcriptomic signatures distinguish baseline from all other treatment time points	179
5.4: Discussion	186
5.5 Contributions	190
References	191
Chapter 6: Conclusions	196
6.1 Summary of major findings from this thesis	196
6.2 Rethinking WHO TPPs for non-sputum based biomarker tests for TB	197
6.3 Technology in place for translation of transcriptomic signatures as point-of-care tests	199
6.4 Addressing the presence of ISGs as part of transcriptomic signatures	201
6.5 Single biomarkers may not be enough	202
6.6 Bottlenecks with studies on non-sputum based biomarkers for TB diagnosis and treatment monitoring in children	203
6.7 How to move biomarker studies for paediatric TB forward	204
6.8 Future work in the field of TB transcriptomic signatures	206
References	209

Chapter 1: Introduction

Timelier and more accurate diagnosis of individuals with TB would greatly contribute to fewer cases of TB being missed. Individuals who go undiagnosed and untreated are more likely to continue contributing to transmission, ultimately leading to even more TB cases arising. Without an increase in the number of TB cases identified, TB will continue to be a leading cause of mortality and morbidity worldwide. In 1993, WHO declared TB a global public health emergency and yet three decades later, it continues to be a top cause of death by a single infectious disease agent (Lessells, 2019). Based on global death estimates in 2019, TB was ranked as the 13th leading cause of death (World Health Organization, 2022a). Between 2020 and 2021, TB was ranked as the second leading cause of death from a single infectious agent, after COVID-19. TB has now overtaken COVID as a leading cause of death. In 2021, over 1.4 million deaths, globally, were officially classified as being caused by TB. This number was more than double that caused by HIV/AIDS (0.65 million) (World Health Organization, 2022a). From the 2023 Global TB report, the number of TB cases detected was lowest in low-income countries and yet historically, these regions have had the greatest TB burden (World Health Organization, 2023). This signifies that more still needs to be done to improve TB diagnosis, by widening access to diagnostic tests. More rapid, cheaper and easier-to-use diagnostic tests, that can be deployed in most low-income countries would be one solution to this problem. The field of transcriptomics is an area with great potential to aid the discovery of biomarkers for TB diagnostic development. From the work highlighted in this thesis and this chapter, we propose that detecting transcriptomic signatures measured in whole blood should be another diagnostic development avenue.

Chapter 1 is divided into four main sections, each highlighting literature that supports the research rationale presented in this thesis. Section 1 discusses various methods for conducting transcriptomic studies, concluding with a justification for choosing the RT-qPCR technique for this project. Section 1 also brings to light the contributions that transcriptomic studies have made with regards to our understanding of the host immune response to TB and how this line of work has already advanced with the new classification of TB disease. Sections 2 and 3 explore the potential of transcriptomic signatures as non-sputum-based biomarkers for diagnostics and treatment monitoring in tuberculosis (TB). These sections also address existing gaps and challenges in the TB field, emphasizing the need for these biomarkers. The chapter concludes with an overview of the aims and objectives of the two primary projects detailed in the results chapters of this thesis.

Section 1: Transcriptomics and its intersection with the study of TB

General overview of methods used to study transcriptomics

Transcriptomics involves the analysis of a collection of RNA molecules present within a particular cell, tissue, or organ. By analysing the transcription, expression levels, and locations of a group of genes expressed as mRNA transcripts we are able to understand more about their function within the human body and the role they play in different disease processes (Scanlan and Wu, 2023; Skerrett-Byrne Anthony et al., 2023). Various high-throughput methods have been used to elucidate insights into the transcriptome, including microarray, RNA sequencing (RNA-seq), and quantitative reverse transcriptase polymerase chain reaction (RT-qPCR).

Microarrays detect nucleic acid and enable the quantification of the expression of transcripts through hybridization to an array of complementary probes on microchips (Lowe et al., 2017). With the improved specificity of probes that have been developed, microarrays have enabled more genes to be detected on a single array. The fluorescence intensity at each probe location directly translates into the level of abundance of each transcript of interest (Lowe et al., 2017). Microarrays are particularly useful for analyzing large mammalian transcriptomes. The drawback with microarrays is that they can only detect known sequences and so, with regards to the discovery of novel transcripts, other platforms such as RNA-seq are needed. Furthermore, background hybridization and the saturation of probes limits the accurate measurement of transcripts that have very low and very high expression (Wang et al., 2009; Zhao et al., 2014).

Next-generation sequencing (NGS) techniques have been in development from the early 2000's, representing a significant advancement over Sanger-based sequencing. NGS has enabled sequencing of whole genomes and exomes, and thus paved a way for RNA sequencing (Alfaro et al., 2019; Behjati and Tarpey, 2013). NGS made the automation of sample processing techniques simpler as it did not require the use of gels, thereby enabling the sequences for a larger number of samples to be determined. Unlike microarray or DNA chip-based methods, prior knowledge about genome sequences is not a requirement. The presence and prevalence of transcripts from previously unknown genes can be identified, giving more room to discover and characterize novel transcripts, alternative splice variants, low-abundance RNAs, and rare cell types (Mortazavi et al., 2008).

RNA sequencing involves the conversion of total or fractionated RNA into a library of cDNA fragments (Wang et al., 2009). Adaptors are then added to one or both ends of each cDNA fragment and short sequences obtained using a high-throughput sequencing technology such as Illumina IG, Applied Biosystems, or Roche 454 Life Sciences (Wang et al., 2009). RNA sequencing does not have an upper limit of quantification and due to being highly sensitive for genes expressed at both high and low levels, it has a wider dynamic range than microarray technology. RNA sequencing, however, does come with its own challenges, such as the setup of cDNA libraries (Wang et al., 2009). The way libraries are constructed has direct implications on the level of transcriptomic profiling that can be performed as well as the way RNA sequencing data is analyzed.

To make sense of the high-quality reads that arise from RNA sequencing, bioinformatics analysis pipelines are required to store and process all the genomic data generated. Finally, costs dictate the number of transcripts that can be analyzed, and the level of sequencing depth. For rare transcripts, costs implications may be higher as greater sequencing depth may be required to ably identify them (Wang et al., 2009). With the continuous upgrade of systems used to study transcriptomics, there are current technologies in development beyond RNA sequencing, such as Nanopore sequencing (Kono and Arakawa, 2019), Nanostring systems (Marx, 2021; Williams et al., 2022), and the use of microfluidics to separate cells for single-cell transcriptomics (Lamanna et al., 2020; Zhou et al., 2021).

The importance of knowing the particular location of molecular features, and how the position of cells in relation to others can influence gene expression, has led to the development of spatially resolved transcriptomics (Asp et al., 2020).

Studies where both RNA sequencing and microarrays have been conducted have reported a correlation in gene expression results (Lowe et al., 2017; Zhao et al., 2014). Zhao and colleagues also reported that despite correlations seen, RNA sequencing was particularly better at identifying transcripts with a low abundance. Interpretation and analysis of RNA sequencing results was easier since the platform was not subject to technical discrepancies that arise from the presence of numerous redundant probe sets and cross hybridization. Researchers have also highlighted the potential loss of gene information owing to the mismatch in transcripts identified between both methods (Rachinger et al., 2021). These are the instances in which validation with an independent gene expression method, such as RT-qPCR, is useful (Lowe et al., 2017; Rachinger et al., 2021). RT-qPCR is a PCR-based technique that combines amplification of a DNA sequence of interest with real-time quantification of the concentration of that DNA species (Dymond, 2013). The process of reverse transcription, carried out by the reverse transcriptase enzyme, leads to the generation of cDNA from each RNA template (Adams, 2020).

RT-qPCR, being a targeted gene expression technique, does have restrictions on the number of transcripts that can be analysed. It does, however, come with the advantage of being cheaper and requiring lower amounts of starting material. Real time PCR methods provide an added advantage over conventional PCR, in that amplification of DNA is monitored in real time. The popularity of RT-qPCR, reflected in its widespread use across various disciplines, results from the speed and simplicity with which experiments are conducted (Bustin et al., 2009). For the work present in this thesis, multiplex RT-qPCR was the method of choice for generation of gene expression data on multiple transcripts of interest. Chapter 2 highlights the laboratory pipeline for the RT-qPCR experiments that were conducted, and Chapter 3 gives details of the transcripts selected for analysis and all the optimization that led to the development of the gene expression panel.

Transcriptomics enhancing our knowledge of the immune response to TB

The various gene expression methods that are currently available have led to the measurement and discovery of various genes, giving us a better idea of how abundant mRNA transcripts are in different tissues, cell populations, and individual cells. This, in turn, has led to the identification of various transcriptomic signatures, comprised of small sets of mRNA transcripts.

Transcriptional profiling and resulting host whole blood signatures can represent changes in cell numbers and cell composition within the blood. These changes may be a direct result of the disease state that an individual is in. The different immune pathways and processes that cells are involved in, during different disease states, can also be represented by the different levels of gene expression that are captured by resulting transcriptomic signatures (Blankley et al., 2014). Through the use of bioinformatic analysis tools, transcriptomic signatures that are as specific as possible for a particular disease of interest, can be discovered (Blankley et al., 2014).

Transcriptomics can identify biomarkers with the potential for translation to TB diagnostic tests. The biomarkers can also be indicative of immune mechanisms that hold the key to understanding what processes within the host are protective and can therefore be targeted for therapeutic development such as vaccines (Burel et al., 2019). A fuller picture of the immune responses to *Mycobacterium tuberculosis* (*Mtb*) can further our understanding of what plays a role in people eventually succumbing to active TB disease (Cliff et al., 2015). Burel and colleagues suggest that three key variables can make a difference in whether signatures discovered end up being more diagnostic or mechanistic in nature. These three variables include: the cohort(s) selected during study design, the nature of samples collected, and how data are generated and analysed (Burel et al., 2019).

To identify and develop mechanistic signatures, Burel and colleagues suggest processing blood to get PBMCs or isolating particular cells of interest using FACS sorting or other separation techniques. These methods may provide more room to study and highlight the role that certain cells play in the immune response to TB, as well as determine the extent of their presence (Burel et al., 2019). Obtaining samples that are closer to the primary site of TB disease, such as lung biopsies could also reveal novel pathways and immune responses that occur in the granuloma, for example. To be able to decipher novel mechanisms that drive the immune response to TB disease, the application of NGS approaches is advised. These methods could give an unbiased view of the transcriptome and increase the possibility of identifying rare transcripts since these approaches have a wider range of detection (Burel et al., 2019).

Transcriptomic signatures associated with TB disease progression typically incorporate inflammatory markers that are up- or down-regulated as an individual transitions towards TB disease. The most striking and consistent transcripts that make up majority of TB transcriptomic signatures to date have been interferon stimulated genes (ISGs) (Berry et al., 2010a; Scriba et al., 2017; Singhania et al., 2018a). Microarray experiments on whole blood from patients with TB, those with latent TB, and healthy controls, led to discovery of a signature with 393 transcripts that could identify patients with active TB (Berry et al., 2010a). Transcriptional modular analysis showed decreased expression of genes mapping to B- and T-cell modules and an upregulation in myeloid-related transcripts. Interferon (IFN) signalling was the most significantly represented pathway in the 393 transcriptomic signature. Genes downstream of IFN- γ and IFN- $\alpha\beta$ receptor signalling were significantly expressed in blood from individuals with active TB (Berry et al., 2010a). The increase in type I IFN- $\alpha\beta$ -inducible transcripts in blood from active TB patients correlated with disease severity. In patients with active TB, these transcripts were overexpressed in neutrophils, and monocytes, but not in T cells, when compared to controls. The active TB signature was also present in 10-20% of latent TB patients,

leading the authors to hypothesize that the signature can also identify individuals who eventually progress to active TB disease (Berry et al., 2010a).

Singhania and colleagues aimed to identify an ideal set of transcripts that could distinguish TB cases from those that are infected, as well as individuals with other diseases. Weighted gene co-expression network analysis (WGCNA) was applied to whole blood transcriptomic datasets that included cohorts of individuals who had TB, bacterial infections, viral infections (Berry et al., 2010b; Kaforou et al., 2013; Zak et al., 2016), and other diseases such as lung cancer and sarcoidosis (Kaforou et al., 2013). WGCNA demonstrated that the IFN modules, pre-dominantly comprised of type I interferon-inducible genes, that were overly abundant in the modular signature for active TB were also present in acute influenza infection (Singhania et al., 2018a). These IFN modules, however, were not as striking and almost absent in bacterial infections (Singhania et al., 2018a). IFN- α/β -mediated signalling that mediates type I IFN immunity came up as a highly significant pathway with regards to differential gene regulation in individuals that had active TB (Ottenhoff et al., 2012).

The role of IFN- γ in mounting a protective immune response has been well documented and widely accepted. There is a debate, however, over whether the existence of type I IFNs, such as IFN- α and IFN- β , as shown in DEGs for active TB, is a result of a possible protective immune response or a contributor to pathogenesis (Moreira-Teixeira et al., 2018). Most transcriptomic studies, to date, seem to support prior immunological research work that points more towards type I IFNs having detrimental effects (Singhania et al., 2018).

Studies have shown that an upregulation of type I IFN signalling has been associated with suppression of immune mechanisms that protect against TB disease (de Wetering et al., 2011; Korbel et al., 2008). The upregulation of type I IFNs is regarded as a possible immune evasion strategy by *Mtb* infection (Ottenhoff et al., 2012). This upregulation of type I IFNs is driven by the expression of transcription factors IFN regulatory factor (IRF) 3 and NF- κ B that are present as a result of the *Mtb* ESX-1 secretion system and subsequent cascades that lead to STING activation (Donovan et al., 2017). Type I IFNs are implicated in suppressing IL-1 α/β responses and the recruitment of macrophages. They are also inducers of IL-10, which may drive individuals more towards TB disease, as a result of impaired immune responses (Donovan et al., 2017).

An investigation into the influence of influenza virus infection before and during acute *Mtb* infection in mice demonstrated a negative effect of influenza on mycobacterial control in the lungs (Redford et al., 2014). Mice co-infected with both influenza A subtype X31 or PR8 and *Mtb* via aerosol had enhanced mycobacterial load and increased levels of inflammation. Influenza A virus was hypothesized to induce long-term effects in the lung microenvironment that were driven by type I IFN signalling. This was confirmed when IFN- $\alpha\beta$ R^{-/-} mice coinfecting with *Mtb* and influenza had mycobacterial loads similar to control IFN- $\alpha\beta$ R^{-/-} mice with *Mtb* only (Redford et al., 2014).

In another study, control mice infected with the HN878 *Mtb* strain had higher bacterial loads than the IFN- $\alpha\beta$ receptor gene disrupted knockout mice. These effects were noted during the chronic stage of disease (Ordway et al., 2007). Although the HN878 strain had the highest virulence, a pattern of better control of bacterial load was observed in IFN- $\alpha\beta$ knockout mice infected with H37Rv, CSU 123, Erdman-KO1, and CSU93 *Mtb* strains via aerosol (Ordway et al., 2007).

To determine the virulence levels of the HN878 and NHN5 *Mtb* strains, pulmonary infection was established in mice. Mice infected with the HN878 clinical isolate had higher colony forming units that kept on increasing 28 days after infection. Mortality rates were also significantly higher in HN878-infected mice. The hypervirulence of the HN878 strain was found to be a result of reduced

Th1 activation due to lower mRNA levels of IFN- γ , TNF- α , IL-6, and IL-12 in the lungs of infected mice. This was coupled with increased mRNA levels of IFN- α (Manca et al., 2001). The inhibitory effects of type I IFNs on the protective immune response to *Mtb* were determined when HN878-infected mice administered with intranasal IFN- α/β had higher bacterial loads, reduced survival times, and reduced levels of IL-12 mRNA in the lungs, in comparison to mice treated with IFN- γ . In fact, there were very minimal differences in mice treated with IFN- γ versus control mice that were infected with *Mtb* but received no treatment (Manca et al., 2001). Type I IFNs were also shown to act by suppressing the production of IL-1 in macrophages and dendritic cells (Mayer-Barber et al., 2011). IL-10 was one of the cytokines produced downstream, because of IFN- β . Together, they suppressed the protective immune response of IFN- γ in macrophages (McNab et al., 2014).

The immunological studies highlighted above point toward type I IFNs most likely playing a role in driving TB disease progression, or at least contributing to the onset of disease. There are, however, some studies that point to a potential protective role of type I IFNs in the context of *Mtb* infection (Desvignes et al., 2012). In C57BL/6 mice carrying null mutations for both type I and type II IFN receptors, bacterial loads and pathology in the lungs were greater than that seen in mice with null mutations for only type II IFN. The mice with impaired signalling for both type I and type II IFNs also died faster. Based on the findings the authors hypothesized that the non-redundant protective role of type I IFNs must be more evident in the absence of IFN- γ (Desvignes et al., 2012).

Alongside IFN-inducible genes, transcriptomic profiling has also highlighted the role of several genes involved in inflammation and the myeloid compartment. This is usually coupled with a downregulation in genes related to B and T-cell compartments (Cliff et al., 2015a). The abundance of type I IFNs has also been coupled with an upregulation of genes involved in the complement cascade (Berry et al., 2010a; Cliff et al., 2013; Kaforou et al., 2013; Maertzdorf et al., 2012).

Aside from complement C5 and C7, that are hypothesized to play a role in TB pathology (Schorey et al., 1997), not much is known about the function of other components of the complement system in the context of TB disease. Based on whole blood microarray results from a study that reported the differential expression of C1q between latent and active TB cases, Cai and colleagues sought to investigate the role of C1q in human TB (Cai et al., 2014). Microarray assays performed on total RNA from PBMCs revealed a significant increase in gene expression of C1qA, C1qB, and C1qC in active TB cases compared to household contacts and those with latent TB. Flow cytometry analysis of peripheral blood from TB patients revealed CD14+ and CD68+ monocytes and granulocytes as the main cell types expressing C1q (Cai et al., 2014). The association of C1q with active TB disease was confirmed by increased mRNA levels of C1qA/B/C in smear positive active TB patients in comparison to those that were smear negative (Cai et al., 2014). In a cohort of HIV negative household contacts recruited from South Africa, The Gambia, and Ethiopia, a single gene pair, *C1QC/TRAV27* (complement C1q C-chain / T-cell receptor- α variable gene 27) was identified as being able to consistently predict progression to active TB in all three countries (Suliman et al., 2018). The authors noted that C1q may be a marker of early TB pathology since it is activated during TB disease progression (Scriba et al., 2017), but downregulated after initiation of TB treatment (Cliff et al., 2013). In the study by Cai and colleagues, C1qC was the subcomponent that had the highest diagnostic accuracy in distinguishing active TB patients from those who were latently infected, with an AUC of 0.84.

Transcriptomic profiling in active TB disease has also revealed an underabundance of IFN- γ and *TBX21* (Singhania et al., 2018a). Through the application of an advanced modular approach, this was identified in individuals with TB and those classified as latent TB outliers (individuals that clustered with the active TB group) (Singhania et al., 2018a). *TBX21* is the gene that encodes the transcription

factor T-bet. T-bet directly activates *IFNG*, the gene that encodes IFN- γ (Lazarevic and Glimcher, 2011). This makes it one of the key regulators of inflammatory and immunoregulatory processes that CD4⁺ T cells are involved in. By upregulating *IFNG*, T-bet promotes the differentiation of Th1 cells (Lazarevic and Glimcher, 2011). Mice with T-bet deficient CD4⁺ T-cells are more susceptible to *Mtb* infection, have decreased IFN- γ production and have higher amounts of IL-10. Likewise, a deletion of the *TBX21* gene was found to increase susceptibility to *Mtb*. T-bet has also been associated with differentiation into effector memory cells since a higher frequency of these cells has been identified in CD4⁺T-bet⁺ *Mtb* antigen-specific CD4⁺ T cells in comparison to CD4⁺T-bet⁻ cells (Yang et al., 2015).

The downregulation in *IFNG* expression, as well as signalling by high levels of type I IFNs, have been shown to contribute to TB pathology (McNab et al., 2015). Through their advanced modular approach to analysing gene expression, Singhania and colleagues found the *IFNG* module underabundant in those that progressed to TB. This led the authors to conclude that their findings support the hypothesis that the ratio of type I IFNs versus *IFNG*-inducible genes plays a pivotal role in determining protection or progression to TB disease (Singhania et al., 2018a).

Transcriptomics and the spectrum of TB disease

With an ever-growing understanding of the immune system and how it responds to *Mtb*, we now know that there are several disease outcomes that can arise after exposure to *Mtb*. These outcomes surpass the conventional binary classification of ‘active’ and ‘latent’ TB (Lin and Flynn, 2018). In light of this, many researchers (Barry et al., 2009a), as well as WHO (World Health Organization, 2020, p. 1), have proposed moving away from the use of the term ‘latent’ and now describe this phase as ‘an on-going state of infection’ or simply as ‘infection’ (Coussens et al., 2024; Migliori et al., 2021).

Latent TB refers to a condition where *Mtb* infection is present, but without any symptoms and this is attributed to the fact that the immune system has been able to contain the bacterium. Several individuals with latent TB can progress to active TB disease, and potentially spread *Mtb* to others around them (Lin and Flynn, 2018).

The tuberculin skin test (TST) gives an indication of whether a person may have an *Mtb* infection by detecting the presence of a delayed hypersensitivity reaction to mycobacterial purified protein derivative (PPD) antigens. The PPD antigens are injected intradermally and within 48 to 72 hours a measure of induration, in millimetres, is taken. The firm swelling seen signifies the presence of *Mtb*-specific T-cells, resulting from *Mtb* infection. Another test, with greater specificity for detecting *Mtb* infection, is the IFN- γ release assay (IGRA) (Barry et al., 2009b). After stimulation with ESAT-6 and CFP-10, the presence of IFN- γ is measured in whole blood.

Individuals termed as ‘resisters’ are those who, despite consistent exposure to *Mtb*, do not have any evidence of *Mtb* infection as seen from a negative TST or IGRA for a sustained period of time. This period could be two years or more. Resisters are thought to have cleared *Mtb* infection very early following exposure (Kaipilyawar and Salgame, 2019). This could result from the innate immune system clearing *Mtb* infection without induction of adaptive immune responses (Lu et al., 2019) or a result of IFN- γ independent adaptive responses.

Within the category of latently infected individuals are those termed ‘reverters’. Reverters initially have a positive TST or IGRA but then change to have a negative result (Lin and Flynn, 2018). Reversions possibly reflect individuals clearing *Mtb* infection or may simply be a result of variability in both laboratory and IGRA test procedures (Pai and O’Brien, 2007). Genome-wide association studies have found consistent TST negativity to be linked to chromosomal regions 2q21–2q24 and 5p13–5q22, TST positivity (termed TST1) to chromosome 11p14 and the intensity of TST reactivity (called TST2) impacted by a locus located on chromosome 5p15 (Abel et al., 2014). These studies contribute to the growing body of evidence of host genetic factors influencing the progression to TB disease. Between the states of latent *Mtb* infection and active TB disease, individuals progress through incipient and subclinical TB, as outlined below.

The New Diagnostics Working Group under the Stop TB Partnership defines incipient TB as “the asymptomatic phase of early disease during which pathology evolves, prior to clinical presentation as active TB disease” (“Model to evaluate the impact of a test for incipient TB | Stop TB Partnership,” n.d.). Although individuals with incipient TB are thought to be infected with viable *Mtb* bacilli, they have no evidence from radiological and microbiological tests that point towards active TB disease (Drain et al., 2018). The state of incipient TB is seen to be the most difficult to describe and detect within the spectrum. This is because there are no approved tests to diagnose individuals with incipient TB. This in part, may be a direct result of very limited bacilli present that are able to be detected with current tests for active TB. Incipient TB may therefore only be defined retrospectively after an individual with incipient TB ultimately progresses to clinical TB disease. Subclinical disease, on the other hand, is a state where viable, and therefore actively replicating bacilli are present with

evidence from either radiological or bacteriological tests, but in the absence of any symptoms (Esmail et al., 2022; Lin and Flynn, 2018; Migliori et al., 2021).

One of the greatest drawbacks of tests for latent TB, such as TST and IGRA, is the inability to detect those at highest risk of progressing to active TB disease. New tests that can identify such individuals and help narrow down the number of people required to receive preventive treatment are needed (Esmail et al., 2020). In recognition of this need, WHO has put together a target product profile (TPP) for novel TB tests that can predict progression to active TB over a period of two years (World Health Organization, 2017a). The ability to accurately identify individuals with the highest risk of progressing to TB, so that they can receive treatment, would allow for preventive therapy to be targeted. This could be done by first administering preventive therapy only to those who are positive for a particular biomarker or have a certain score for whatever transcriptomic signature is being measured.

When one is infected with *Mtb* and the mycobacteria in their lungs or in any other tissue lead to them being sick the consequence is active TB disease. This results in signs and symptoms that are evident on clinical presentation and a disease status that can often be confirmed from radiological and microbiological tests. TB disease is typically diagnosed by detection of *Mtb* in sputum. The diagnostic tests for sputum testing have limited sensitivity leading to a proportion of patients who do have TB not accurately being diagnosed. (Migliori et al., 2021).

The reliance on the collection of sputum limits the utility of currently available and widely used TB diagnostics, particularly at earlier stages of disease progression. This is because most individuals at early stages within the spectrum are asymptomatic and cannot produce sputum. Alternative tests suggested to identify people during earlier stages of disease are biomarker-based, with an emphasis on easily accessible sample types other than sputum, such as blood. These biomarkers are indicators of the host immune (Lin and Flynn, 2018) response to *Mtb*, while others can identify parts of the bacterium (Esmail et al., 2022).

In settings that are highly endemic for TB, there are very high proportions of adults who have already been exposed and infected with *Mtb* (Mpande et al., 2021). In these areas, the provision of TB preventive therapy (TPT) would not be affordable nor feasible to carry out (Mpande et al., 2021). Not everyone who is infected with *Mtb* ends up progressing to TB disease and so provision of TPT to everyone may expose many individuals to unnecessary side-effects.

To further streamline who would require TPT, biomarkers that can distinguish recent from remote TB infection have been suggested. This arises from the finding that the highest risk of TB progression occurs within the first two years after primary infection (Machingaidze et al., 2012). Features of *Mtb* specific T-cells such as activation (Musvosvi et al., 2018a) and differentiation (Petrucchioli et al., 2015) have shown diagnostic capability in distinguishing *Mtb* infection from TB. Through the design of a simple flow cytometry assay, Δ HLA-DR, a measure of activated T-cells, was discovered as a blood-based biomarker that could identify individuals with recent QuantiFERON-TB Gold (QFT) conversion, as well as TB disease progressors with higher levels of *Mtb*-specific T-cell activation (Mpande et al., 2021).

T cell activation markers (TAMs) such as CD27, CD38, and HLA-DR, identified in whole blood or PBMCs, have been shown to differentiate individuals with active TB from those who are latently infected (Acharya et al., 2021; Portevin et al., 2014). These findings led Hiza and colleagues to test the diagnostic capability of a simplified TAM-based assay that takes real-time measurements from a millilitre of blood (Hiza et al., 2021). The authors also sought to compare the performance of both

CD27 and CD38, as other literatures suggests that CD38 could perform better than CD27 (Adekambi et al., 2015; Ahmed et al., 2019). The CD38-based TAM assay performed better than CD27 in distinguishing active TB from non-TB cases, however CD38 had decreased specificity for presumptive TB cases with recent exposure (Hiza et al., 2021). No differences in diagnostic accuracy were noted when performance of the CD38-TAM assay was compared between HIV positive and negative individuals. This led (Hiza et al., 2021) to the same conclusions as (Wilkinson et al., 2016) about TAMs being able to differentiate active from latent TB irrespective of HIV status. The CD38-TAM assay achieved a specificity 93.4% and sensitivity of 82.2%, therefore meeting the minimum sensitivity (80%) and approaching the minimal specificity (98%) of the WHO TPP for non-sputum based diagnostic tests (World Health Organization, 2014).

There is growing evidence supporting the role that other non-sputum based biomarkers, such as whole blood transcriptomic signatures, can play in the diagnosis and identification of individuals within early stages of the TB disease spectrum (Migliori et al., 2021). With an aim to identify transcriptomic signatures that measure the risk of one developing TB, a cohort of South African adolescents infected with *Mtb* were followed up for two years, alongside HIV-negative adults from South Africa and The Gambia who had household exposure to an adult with sputum smear positive TB disease. The disease risk signature that was developed provided evidence that whole blood transcriptomic signatures can detect progression to active TB disease up to eighteen months before the presence of any TB disease manifestations (Zak et al., 2016).

A subsequent randomised control trial determined both the diagnostic and prognostic ability of a blood transcriptomic signature, RISK11, to predict incident TB in South African adults (Scriba et al., 2021a). The prognostic performance of RISK11 for incident TB within 6 months of testing exceeded the optimal TPP for an incipient TB test. Within 12 months, the RISK11 signature approached the minimal TPP, but failed to meet this criterion beyond the 15-month mark (Scriba et al., 2021a). In a prospective study of RISK11 carried out in HIV positive adults within South Africa, the signature's performance approached, but did not meet the WHO benchmarks for a screening and prognostic test for TB (Mendelsohn et al., 2021). In one follow-up study, the diagnostic performance of eight signatures in symptomatic and asymptomatic individuals was benchmarked against the RISK11 signature and the WHO benchmarks for a triage test and an incipient TB test (Mendelsohn et al., 2022b). Several of the signatures met or approached the WHO TPP for a triage test. In the work discussed in Chapter 4, we report similar findings from an evaluation of 20 transcriptomic signatures for differentiating TB cases from controls with other respiratory diseases.

In a systematic review and patient-level pooled meta-analysis of the performance of 17 candidate signatures as biomarkers for incipient TB, none of the signatures met the WHO benchmarks for biomarkers targeted at identifying individuals with incipient TB over a two-year period (Gupta et al., 2020). However, the performance of some signatures to identify those with incident TB within 3-6 months after testing, exceeded the TPP criteria. This led the authors to conclude that the prognostic signatures would only have utility if they were measured serially at 9-12 month intervals (Gupta et al., 2020).

Changes in gene expression, represented by transcriptomic signatures, can effectively track the evolution of the host immune response during the course of progression along the TB disease spectrum (Lin and Flynn, 2018; Tabone et al., 2021). Tabone and colleagues were able to decipher the immune responses present in individuals across the spectrum of TB disease by following up study participants with microbiologically-confirmed TB, and household contacts of active TB patients that progressed to TB, for a total of 24 months. A clinical criterion was used to define each category within the spectrum of TB disease (Tabone et al., 2021). The authors carried out modular transcriptional

analysis and identified differentially expressed genes over time in TB contacts who progressed to TB disease (Tabone et al., 2021). These analyses revealed reduced expression of gene modules for T-cells, B-cells, and NK-cells more than 40 days before diagnosis. Modules for complement and type I IFN-inducible genes were increased up to 20 days prior to diagnosis. The levels of these genes were downregulated to the same level as that in healthy controls at the 4-month treatment time point (Tabone et al., 2021).

Tabone and colleagues also developed 30-gene signatures for incipient, subclinical, and active TB and demonstrated that in individuals classified as having incipient TB, no significant rise in gene expression was observed for the incipient, subclinical and active TB signatures that were developed. An increase in gene expression, however, was seen in both subclinical and active TB cases, with maximum expression noted in active TB cases (Tabone et al., 2021).

Individuals with latent TB, and those with sub-clinical disease have shown increased expression of transcripts involved in the complement cascade (Singhania et al., 2018b). A group of asymptomatic HIV-1 infected individuals were able to be classified as having either sub-clinical or active TB disease based on evidence captured on [18F]-fluoro-2-deoxy-D glucose positron emission/computed tomography (FDG-PET/CT) imaging (Esmail et al., 2018). Researchers went a step further to carry out transcriptomic profiling on these same individuals to identify any features that were unique for those with sub-clinical disease. Of the 82 transcripts that distinguished sub-clinical from active TB patients, those involved in the classical complement pathway were overly expressed in peripheral blood of sub-clinical cases (Esmail et al., 2018).

Esmail and colleagues (Esmail et al., 2018) validated their findings by re-analysing RNA-seq data from the ACS study (Zak et al., 2016) discussed earlier. Apart from the type 1 IFN signalling pathway, components of the classical complement pathway such as C5 and C1QA were significantly differentially expressed less than 180 days before adolescents were diagnosed with TB. Even at time points between 181 - 360 days to TB presentation, Ingenuity Pathway Analysis identified the complement pathway as one of the most significantly enriched biological processes. Similar findings were reported from earlier work carried out on by Scriba and colleagues, on the same adolescent cohort (Scriba et al., 2017).

In light of the spectrum of TB disease progression, new screening tools and diagnostic tests are required. It is also not known whether newer, shorter regimens would be sufficient to treat people with incipient and sub-clinical TB, and how to make the changes in treatment algorithms easier to implement (Esmail et al., 2022).

What is yet to be fully understood, is how much individuals with sub-clinical disease contribute to the spread, and burden of TB. Another question is whether individuals with incipient TB are infectious (Migliori et al., 2021). Evidence gathered on these matters would allow TB control efforts to include the early stages of TB (Migliori et al., 2021). By coming up with tools that best identify people in both the incipient and subclinical TB disease categories, we can learn more about these stages to better inform what occurs both clinically and with regards to policies on prevention of TB transmission.

A compartmental mathematical model was constructed to determine the effect that the inclusion of subclinical and “pre-diagnostic” disease (defined as a stage where individuals are identified as being positive for symptoms during screening, but these symptoms do not warrant seeking further healthcare intervention) may have on current strategies to identify active TB cases (Dowdy et al., 2013). The inclusion of a 9-month sub-clinical phase into models assessing the impact of different

diagnostic interventions led to reductions in TB incidence up 2.7 times greater than any enhanced diagnostic strategies. Although we are yet to learn more about the incipient and sub-clinical disease phases, the authors strongly urge for the inclusion of transmission, during the sub-clinical phase, as part of current epidemiological models. This could help prevent an overestimation of the impact that TB diagnostics may have in TB prevention and control programs (Dowdy et al., 2013).

In low-middle income settings, there is limited evidence for treating individuals who have no symptoms, and have negative microbiological tests, even though there is evidence of some pathology. Clinical trials will have to take place to determine the appropriate treatment to give during earlier stages of disease, or whether to give any treatment at all. The CORTIS study is an example of a randomised controlled trial that evaluated the performance of the RISK11 transcriptomic signature in guiding initiation of TPT, for individuals who were RISK11 positive (Scriba et al., 2021a). These individuals were classified as having a high risk of progressing to incident TB disease. RISK11 performed better as a screening test for TB in symptomatic individuals recruited from TB-endemic communities, compared to those who were asymptomatic. To our knowledge, the CORTIS study is the only one to date that has reported the use of a transcriptomic signature to determine giving treatment.

In summary, the work highlighted in this section of Chapter 1 demonstrates the role that transcriptomic signatures have played in identifying individuals across the entire spectrum of TB disease progression, and in enhancing our understanding of immune responses present at each stage.

With the shortcomings already mentioned that are still present regarding the use of sputum-based diagnostics for active TB, one of the areas of focus in this thesis (**Chapter 4**) is on transcriptomic signatures, particularly for active TB. Since many transcriptomic signatures have already been discovered, our main objective was to evaluate multiple transcriptomic signatures for TB screening and triage. The signatures identified to have the highest diagnostic accuracy as triage tools could be considered for point-of-care (POC) diagnostic development. With continuous generation of evidence for the use of transcriptomic signatures as TB biomarkers, the POC test(s) eventually developed could also be used to identify individuals with subclinical and incipient TB. We have already seen potential for transcriptomic signatures in identifying individuals with incipient TB (Gupta et al., 2020). Several of the signatures evaluated by Gupta and colleagues (Gupta et al., 2020) were part of our 20 transcriptomic signature panel (**Chapter 3**). Some of these signatures also allowed identification of individuals who eventually progressed to active TB (Mendelsohn et al., 2024).

The next section of this chapter will highlight work on non-sputum-based biomarkers that are currently in use as well as those with potential for TB diagnosis, use in primary health care settings and at the POC. Transcriptomic signatures and their ability to stand out from other potential non-sputum-based biomarkers has also been demonstrated.

Section 2: Point of care diagnostic tests for TB

Section 2 of this chapter will give an introduction to current diagnostic platforms that exist within primary health care settings, at the POC, as well as those with potential for use in these settings. The studies mentioned provide an overview of the major advantages and disadvantages of these diagnostic platforms and improvements that have been made to some of them. An emphasis has been put on diagnostic tests in primary health care settings, and at the POC since gaps in TB diagnosis, at these levels, are where we believe transcriptomic signatures can make a difference.

The term primary health care describes the most basic health care level where individuals first interface with health care workers (García-Basteiro et al., 2018). POC testing describes testing people admitted to hospital, those seeking care at peripheral clinics, people at home and those within communities (Lessells, 2019). Another feature that defines POC testing is the lack of a requirement for a laboratory or skilled technician to operate.

Tests aimed at being deployed at the POC, particularly primary health care settings, should lead to improvements in patient care and management. This would arise from a combination of more rapid diagnosis and timely initiation of appropriate treatment (García-Basteiro et al., 2018). For that to occur, POC tests should have a fast turnaround time, ideally not require electricity or instead run on a battery, be easy to interpret, and have little reliance on the need for sophisticated lab settings alongside highly trained staff. Such tests may have to be developed and tested on samples other than sputum (García-Basteiro et al., 2018).

POC tests will also need to be more affordable and accessible beyond centralized health care centres, in order for more people to have an opportunity to receive a TB diagnosis (Hong et al., 2022). Mathematical modelling has provided insight into how much impact improved access to TB diagnostics can have (Lessells, 2019). A decision-tree model was used to simulate the introduction of theoretical tests for pulmonary TB, for use in adults presenting with chronic persistent cough (Keeler et al., 2006). Improved diagnostic tests, deployed in the four highest burden WHO regions were noted to have the potential to reduce annual mortality by up to 625,000 lives or 36% of the total. These results, however, were based on 100% access to a rapid diagnostic test that has 100% sensitivity and 100% specificity (Keeler et al., 2006). Although no tests for TB currently exist that meet these criteria, TPPs that represent the most urgent diagnostic needs of end-users have been developed by the WHO (World Health Organization, 2014). They outline the optimal and minimal performance characteristics required of non-sputum based POC diagnostic tests, triage tests, and sputum tests that can replace sputum smear microscopy (SSM).

Great strides, however, have been made when it comes to TB diagnostic development, including tests that have been endorsed by WHO such as Xpert MTB/RIF, Xpert MTB/RIF Ultra, and urine lateral flow lipoarabinomannan (LF-LAM). GeneXpert tests have been applied more at referral, regional, district and subdistrict levels, but are also thought to be useful at the POC despite their limitations (García-Basteiro et al., 2018). These limitations are discussed further in the section under the heading 'nucleic acid amplification tests'.

Current point-of-care or near-point-of-care tests

Smear microscopy

SSM has been a standard test for TB diagnosis across primary health centres in most resource limited settings (Kik et al., 2014b). Although it cannot detect drug-resistant TB and has low sensitivity, SSM continues to be used because it is rapid, simple to perform, affordable, and does not require a

sophisticated lab setting (Desikan, 2013). In areas with a high TB prevalence and where samples have high bacterial loads, SSM tends to be highly specific for detecting *Mtb* (Desikan, 2013).

The Ziehl–Neelsen (ZN) stain for SSM has been the most common staining method in developing countries (Abdelaziz et al., 2016). Gabbett’s cold stain and modified cold staining are both modifications of the ZN technique that do not involve a heating step and have a single decolorisation and counterstaining step. In a study that evaluated all three staining methods versus sputum culture, Gabbett’s stain had the highest specificity and sensitivity for detecting acid-fast bacilli, followed by the conventional ZN stain. This led the authors to suggest Gabbett’s stain as a replacement for ZN (Abdelaziz et al., 2016).

The major disadvantages of SSM are that viable and dead mycobacteria are indistinguishable and that sensitivity is low leading to low bacterial counts in sputum samples being detected as negative. In individuals who cannot expectorate sputum easily, such as children and those with HIV, TB diagnosis based on SSM is not helpful. SSM is most often negative leading to an incorrect diagnosis being made. Another major limitation is the inability to distinguish *Mtb* from non-TB mycobacteria. This is particularly relevant in certain populations, such as those with a high proportion of PLWHIV, and settings with a high burden of environmental mycobacteria.

Owing to the current contributions that microscopy has made to TB control programs worldwide, methods to improve it, as opposed to replacing it completely, have been identified such as auramine fluorescent microscopy. TB diagnosis based on fluorescent microscopy has a faster turnaround time and noted to be up to 10% more sensitive than regular SSM, however specificity has not been found to be any different (Desikan, 2013).

In a study carried out in India comparing both fluorescent and ZN microscopy techniques, fluorescent microscopy was found to be superior to ZN microscopy, detecting 24% more cases of *Mtb* in sputum samples (Hajra Tasneem Masali, Aparna Takpere and Praveen Shahapur, 2021). Similar findings were reported from a study conducted in Ghana where fluorescent microscopy was more sensitive than the ZN technique for diagnosis of pulmonary TB (Dzodanu et al., 2019). The higher sensitivity and specificity of fluorescent microscopy was attributed to its stronger contrast compared to ZN staining and the ability to screen slides at a lower magnification of 200x and 400x (Hajra Tasneem Masali, Aparna Takpere and Praveen Shahapur, 2021). The procedure of staining for fluorescent microscopy is also less complicated compared to the ZN stain. Unlike the ZN procedure, fluorescent microscopy does not require an immersion objective, immersion oil, or xylene.

Fluorescent microscopy does come with its own cost implications, one major one arising from mercury vapour lamps (MVLs) that are expensive and have a short life span. This, coupled with the requirement of stable electricity and replacement bulbs, have hindered the widespread roll-out of fluorescent microscopy for TB diagnosis in primary health care settings (Whitelaw et al., 2011).

An alternative to fluorescent microscopy recommended by WHO is light-emitting diode (LED) based microscopy. LEDs are up to 10 percent cheaper than MVLs, have a longer life span (>50,000 h vs ~250 hours), do not require a dark-room, and can run on batteries (Desikan, 2013; Whitelaw et al., 2011). LED microscopy is more sensitive than the conventional ZN microscopy method, but this is accompanied a loss in specificity (World Health Organization, 2011a). On both direct and concentrated smear samples, LED microscopy was cheaper than performing ZN-based microscopy (US\$1.63 versus US\$2.10) (Whitelaw et al., 2011). Performance, however, was noted to be similar across LED, MVL fluorescent, and ZN microscopy techniques with lowered sensitivity in HIV-positive patients (Whitelaw et al., 2011).

Sputum culture

Detection of *Mtb* through culture is still the gold standard for the diagnosis of active TB. *Mtb* is slow-growing leading to the turnaround time for solid culture results being between four to eight weeks. *Mtb* culture can be done on either solid (Löwenstein-Jensen, LJ) or liquid media such as Middlebrook 7H9. Solid culture is cheaper and less likely to have contamination by other micro-organisms in comparison to liquid media (Huang et al., 2022). Liquid cultures, however, have growth detected automatically, allowing a faster turnaround time for results (Chien et al., 2000).

In a systematic review conducted by Rageade and colleagues, six different culture-based diagnostic tests were assessed. The BACTEC mycobacteria growth indicator tube (MGIT) 960 System (Becton-Dickinson) had the highest sensitivity when used alone or in combination with LJ media. LJ media, however, had the highest specificity for detecting *Mtb* (Rageade et al., 2014).

Liquid medium for *Mtb* culture is a faster and more sensitive method for detection of *Mtb*. One study assessed the diagnostic yield from sputum using SSM, solid (LJ), and liquid media (modified Middlebrook 7H9 broth) (Boldi et al., 2023). Both solid and liquid culture media had higher sensitivities and specificities for the diagnosis of pulmonary TB in comparison to both direct and concentrated SSM methods. Similar findings were reported in a study that assessed SSM, PCR, and culture for TB diagnosis in multiple respiratory samples, including sputum, induced sputum, bronchial aspirate, and broncho-alveolar lavage. Culture had the highest specificity and sensitivity of all three diagnostic techniques for all respiratory samples tested (Boldi et al., 2023).

Although *Mtb* cultures are advised for immunosuppressed patients, the diagnostic accuracy of these tests in people living with HIV (PLWHIV) is not clear. Furthermore, there is no set number of culture tests that should be conducted to ensure definitive results. In clinical settings, however, it is not feasible or cost-effective to perform two specimens for culture per patient.

Nucleic acid amplification tests (NAATs)

NAATs offer better diagnostic accuracy than SSM (Peralta et al., 2016). One of the first NAATs to be rolled out by national TB programmes was the Xpert MTB/RIF assay (Cepheid), a cartridge-based test able to detect *Mycobacterium tuberculosis* complex (MTBC) in sputum samples, as well as determine whether the bacterium is resistant to rifampicin (RIF) in less than 2 hours. The turnaround time for the Xpert MTB/RIF assay gives it an advantage over culture tests that take four to eight weeks, and conventional drug resistance tests that can take up to an additional three weeks (Peralta et al., 2016).

The performance of the Xpert MTB/RIF test was compared to two other NAATs for TB, the LightCycler Mycobacterium Detection assay (Roche) and the MTBDRplus assay (Hain Lifescience) (Scott et al., 2011). Using sputum samples from culture-confirmed and clinically-defined adult cases of TB from an area with a high HIV prevalence, Xpert MTB/RIF had the highest sensitivity of all three assays (86%, 95% confidence interval (CI) 76%–93%). All three NAATs had similar sensitivities among HIV uninfected participants. Among HIV-infected participants, however, sensitivities for the LCTB and MTBDRplus assays dropped. The NAATs had higher specificities for smear-positive TB cases but among smear-negative culture-positive TB cases, the Xpert MTB/RIF assay had the highest specificity (Scott et al., 2011).

The sensitivity of Xpert MTB/RIF has not yet superseded that of culture, as false negatives have been known to arise (Blakemore et al., 2010; Boehme et al., 2011, 2010; Helb et al., 2010). This is a major limitation of Xpert MTB/RIF/Ultra and other NAATs which stems from them failing to differentiate between genomic material from live versus dead mycobacteria. This becomes particularly problematic for individuals who have recently been cured or treated.

Aspects that directly limit the use of Xpert MTB/RIF in resource limited settings include servicing and maintenance of the GeneXpert instrument, the shelf life of the cartridges, the requirement for a constant supply of electricity, as well as particular humidity and temperature requirements (Evans, 2011). In 2011, MTB/RIF machines were estimated to cost between US\$17,000–\$62,000, with costs for each disposable test-cartridge at US\$17–\$120 (Evans, 2011). In 2023, it was announced that the cost of Xpert Ultra cartridges would be reduced to US\$7.97. This was the result of a partnership that Cepheid had with The Global Fund, the Stop TB Partnership, the U.S. Agency for International Development and Danaher technology. While the costs of Xpert MTB/RIF have gone down in the last decade, it is still far more expensive than SSM (Evans, 2011; Lawn et al., 2013). The rollout of Xpert machines in several LMICs still depends largely on donations and bodies such as WHO lobbying for subsidized prices from Cepheid for increased uptake in resource limited settings (Evans, 2011).

There are other factors health workers have noted that have implications on the use of GeneXpert in POC clinics and community hospitals within high TB burden settings. These include high specimen loads that in-turn affect the turnaround time for results, availability of technicians locally in most LMIC settings to repair instruments in a timely manner, as well as limitations on the number of tests that can be performed on Gene Xpert machines that have a maximum of 4 modules (Pongpeeradech et al., 2022). GeneXpert machines, however, have gone through several upgrades to run up to 1,000 tests in 24 hours on devices such as the GeneXpert Infinity machine that has a total of 48 modules. The Xpert MTB/RIF Ultra assay has also been developed that has a lower limit of detection for MTBC, resulting in improved sensitivity compared to the original Xpert machine. Cepheid has also developed the GeneXpert Omni, a wireless, web-enabled machine that can be operated via a smartphone and is battery-operated. This allows it to be used in rural areas or mobile clinics without electricity (Zhang et al., 2020). With the current advances in technology that have supported the development and roll-out of NAATs, such as GeneXpert, there are increasing opportunities to revolutionize TB diagnostic development and with that, do away with the reliance on sputum as the major sample for TB diagnosis (Piatek et al., 2013).

Chest X-ray (CXR)

CXRs have been used for TB screening for over 70 years by detecting pathology consistent with TB. In South Africa, CXR vans performed mass screening between 1950-1970. Mass screening using CXR was also carried out in the Czech Republic as early as 1960 (Schwalb et al., 2022). With the ability to pick up other respiratory conditions, CXRs can also provide differential diagnoses to TB. The low specificity of CXRs (MacPherson et al., 2021; Mungai et al., 2022), however, has been associated with over-diagnosis of TB. This in turn leads to the initiation of treatment in people who would not benefit from treatment (Cudahy et al., 2017; Davis et al., 2010; Jaeger et al., 2013). This is why positive CXR results, at the point of triage, should ideally be followed up with a bacteriological confirmatory test such as Xpert MTB/RIF, that have a higher specificity (MacPherson et al., 2021).

CXRs have performed poorly at diagnosing smear-negative, culture-positive TB in HIV positive individuals (Cudahy et al., 2017; Davis et al., 2010). For example, one study showed that with *Mtb* culture as the reference standard, CXR had a sensitivity of 67.9% (95% CI 56.6–77.8%) and specificity of 53.7% (95% CI 47.7–59.6%) for the diagnosis of smear-negative TB in HIV positive cases (Nakiyingi et al., 2021). When combined with Xpert MTB/RIF, CXR interpretation did increase the sensitivity of Xpert [87.7% (95%CI 78.5–93.9%) vs. 65.4% (95%CI 54.0–75.7%)] but brought down specificity [51.6% (95%CI 45.6–57.5%) vs. 95.8% (95%CI 92.8–97.8%)] (Nakiyingi et al., 2021).

Addressing operational costs, including electricity, X-ray films, and machinery maintenance, is crucial for effective CXR implementation (Nakiyingi et al., 2021). A case study on the inclusion of CXR in

triage algorithms in India highlighted how cost implications, availability of CXR equipment and the presence of trained personnel hampered the implementation of these strategies (World Health Organization, 2016). Although CXRs have been an established screening tool for TB, they require highly trained staff that can operate the machinery and interpret results (Bigio et al., 2021).

Technological advances, such as the development of Computer-Aided Detection (CAD), are a possible step forward to address some of these factors (Bigio et al., 2021). Like SSM, the sensitivity and specificity of CXR is variable depending on how health personnel score results. Automated scoring, based on CAD algorithms, is one strategy with potential to not only reduce inter- and intra-observer bias, but can also be used in areas where human capital is limited (Muyoyeta et al., 2014). For example, a prospective study was conducted to determine the sensitivity and specificity of a CAD program for scoring CXR in comparison to GeneXpert (Muyoyeta et al., 2014). CAD had a sensitivity of 100% (95%CI 96.2–100), specificity of 23.2% (95%CI 18.2–28.9), PPV of 33.0% (95%CI 27.6–38.7), and NPV of 100% (95% 93.9–100). In HIV positive individuals, CAD did have reduced specificity (18.0% vs 29.2%; $p = 0.039$) (Muyoyeta et al., 2014).

To identify the contribution of CAD software to community-based TB screening, its accuracy was assessed in individuals that were part of the 2016 Kenya National TB prevalence survey (Mungai et al., 2022). The Computer-Aided Detection for Tuberculosis version 6 (CAD4TBv6) system met both the minimum and optimum TPP for a community-based referral test (Mungai et al., 2022). Based on this outcome, the authors supported WHO recommendations to have CAD in resource-limited settings for interpreting digital x-rays for TB screening.

A randomised trial was conducted in Malawi to determine the cost implications and the ability for the inclusion of CAD to improve the timeframes in which a full HIV and TB diagnosis was reached (MacPherson et al., 2021). Oral HIV testing coupled with CAD and Xpert MTB/RIF confirmation reduced the time to treatment initiation by eleven days. This was in comparison to TB/HIV screening led by a medical professional. Analysis done over an 8-week period during the trial did not find CAD cost-effective. McPherson and colleagues highlighted Xpert testing as one of the main cost drivers, and that this could be addressed by increasing the CAD4TB score thresholds for identifying individuals with presumptive TB. Another suggestion was to have another rapid screening tool or the use of a risk prediction score to further streamline which individuals should go forward for definitive sputum-based diagnostic testing (MacPherson et al., 2021).

Although CAD has the potential to address the limitations of normal CXRs, aspects such as the need for CXR hardware, access to internet, and equipment required to run the CAD4TB software still limit the widespread deployment of CAD in resource-limited settings (Bigio et al., 2021). With the limitations still present regarding CXR and CAD, point-of-care ultrasound (POCUS) devices have been considered as another possible TB triage tool. POCUS devices are more affordable than CXR, easier to transport and do not come with radiation exposure. They do, however, take longer to perform in comparison to CXR plus CAD and require a skilled ultrasonographer. Although POCUS devices have been utilized for other infectious diseases, more evidence will have to be gathered to assess their diagnostic accuracy in comparison to CXR and CAD as well as their ability to reach the WHO TPPs for a non-sputum based triage test (Bigio et al., 2021).

Lipoarabinomannan (LAM)

LAM is a heat-stable lipoglycan and a major component of the cell-wall of mycobacteria, excreted in urine by some people with TB (Thwaites, 2014). Urine collection is a simple and non-invasive procedure, making the detection of urinary LAM a simple method for non-sputum TB diagnosis. This is easier and more suitable for individuals who cannot expectorate sputum or have paucibacillary TB,

as is the case in most children (Togun et al., 2017) and PLWHIV. The detection of the LAM antigen in urine, however, has varied specificity and a low sensitivity for TB diagnosis. This sensitivity has been reported to be better in people with severe immunocompromise, such as advanced HIV disease (Hanifa et al., 2016). It is because of this that LAM assays, such as Determine TB LAM (AlereLAM) have been approved by WHO for diagnosis of TB in individuals living with HIV (Ketema et al., 2022).

A policy update from WHO, however, does warn against the use of lateral flow urine lipoarabinomannan assay (LF-LAM) as a triage test. Furthermore, in outpatients, they recommend against the use of LF-LAM without an assessment of TB symptoms, when CD4 count is unknown or when CD4 count is greater than 200 cells/mm³ as there is very low certainty of the test's accuracy in such scenarios ("Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV, 2019 Update," n.d.). Among outpatients in established HIV care in South Africa with CD4 counts less than 200x10⁶ cells/l, LF-LAM had a very low sensitivity of 13%. The findings brought Hanifa and colleagues to different conclusions from WHO, recommending against the use of LF-LAM as a screening tool in ambulatory patients (Hanifa et al., 2016).

Since children with severe acute malnutrition (SAM) present with immune-suppression, Schramm and colleagues looked at the ability of LF-LAM to detect TB in HIV-negative, hospitalized children with SAM (Schramm et al., 2021). Children younger than five years were recruited from an in-patient feeding centre in Niger and placed into two categories; group 1 that were probable TB cases and group 2 that were unlikely TB cases. Group 1 had a higher LAM-positivity than group 2 (52.0% versus 37.0%; $p = 0.032$). All children in the study who were negative for Xpert but clinically diagnosed with TB were positive for the Determine-LAM test. Although HIV negative children do not fall under the LF-LAM target group outlined in WHO policy recommendations, Schramm and colleagues suggest the use of this test for fast tracking TB treatment initiation, but only among children with SAM and LAM positivity higher than grade 1 (Schramm et al., 2021).

Development of point-of-care diagnostic biomarkers for TB

Of the 6.2 million people diagnosed with pulmonary TB worldwide in 2021, 37% were not bacteriologically confirmed (World Health Organization, 2023). To add to the burden of active TB, disease states such as incipient and sub-clinical TB have a very low likelihood of being identified by sputum-based tests, since individuals at these points within the TB spectrum usually have paucibacillary disease (Coussens et al., 2024). Individuals with incipient or subclinical disease, however, may still transmit *Mtb* and would therefore also benefit from timely diagnosis and preventive treatment (Drain et al., 2018; Migliori et al., 2021).

Various forms of TB disease, such as extrapulmonary TB, and paucibacillary TB in children and PLWHIV are difficult to bacteriologically confirm and also contribute to the active TB disease burden (MacLean and Nathavitharana, 2021). Current tests such as SSM, Xpert MTB/RIF, and mycobacterial cultures require the collection of sputum samples from children and PLWHIV, however these individuals may not always be able to produce sputum. Furthermore, these tests do not have sufficient sensitivity to ably detect TB in children with low bacterial loads. Extrapulmonary TB requires non-respiratory samples for diagnosis which are challenging to obtain and current diagnostic tests may not be able to sensitively detect *Mtb* in these samples (MacLean and Nathavitharana, 2021).

Further drawbacks with sputum-based tests come from the fact that they may put health workers at a higher risk of *Mtb* infection since the collection of sputum samples can potentially expose health workers to aerosols that are infectious. To address these limitations, non-sputum based biomarkers are needed, particularly at the POC. Recently, there has been prolific discovery of host and pathogen biomarkers with potential for development as TB diagnostic tests, some of which include *Mtb* antigens, volatile organic compounds (VOCs), human antibodies against *Mtb* antigens, human proteins, and human mRNA.

Multiple reviews (Abdulgader et al., 2022; MacLean et al., 2019a; Nathavitharana et al., 2022; Togun et al., 2018; Zimmer et al., 2022) already exist that have discussed the pros and cons of several promising biomarkers for TB diagnosis and treatment monitoring. Other reviews, specific to some of the biomarkers discussed in this section are mentioned and referenced in the text. This section on biomarkers will draw from these reviews and independent studies that showcase the role that biomarkers can play regarding TB diagnosis. Some shortcomings of these biomarkers and areas for improvement have also been highlighted. We believe that these aspects have the potential to be addressed by non-sputum based biomarkers, particularly transcriptomic signatures, that have been identified in less invasive and more easily accessible samples such as whole blood.

Host sensitization to *Mtb* antigens

TST and IGRAs (previously described in section 1 of Chapter 1) are used to determine whether an individual has *Mtb* infection (Dewi et al., 2019). IGRAs were developed as more specific tests for *Mtb* infection since cross reactivity between PPD antigens and several mycobacteria decrease the specificity of TST as a test for *Mtb* infection (Delogu and Goletti, 2014). Two *ex vivo* T-cell tests for the diagnosis of latent TB include the QuantiFERON (QFT)-TB Gold and T-SPOT.TB assays.

The main advantage of TST over IGRA is its applicability for wide-scale screening since it is less costly and does not require a laboratory setup. IGRAs, however have a lower number of false-negative results in immunocompromised individuals and are negative in BCG-vaccinated individuals since they are specific for *Mtb* infection (van Ingen et al., 2009). PPD is not a suitable biomarker for TB diagnosis because of its poor specificity that results from people already being sensitized to it through exposure to NTMs or BCG (van Pinxteren et al., 2000). Ultimately, a test for *Mtb* infection should

allow early preventive treatment to prevent progression to TB disease. The drawback with both TST and IGRA is the poor accuracy in predicting progression to active TB (Rangaka et al., 2012), which defeats the purpose of *Mtb* infection diagnosis, which is to intervene early enough to prevent development of active TB.

A meta-analysis comparing the performance of both IGRAs and TSTs for the diagnosis of latent TB showed similar sensitivities between both tests (Menzies et al., 2007). Individuals in this study were classified based on a clinically defined gradient of exposure. In study participants who received BCG vaccination, both the IGRA and ELISpot tests were more specific than the TST. In studies that carried out both TST and IGRA tests, there were high numbers of discordant results among people who received BCG vaccination. ELISpot had more positive results compared to QFT, which is in line with it having the highest sensitivity in this study, even in individuals who were immunocompromised. An update to the results from the systematic review incorporated 20 more studies and further affirmed that the T-SPOT.TB assay was more sensitive than the QFT and TST for diagnosis of latent TB (Pai et al., 2008). The IGRA tests, particularly QuantiFERON-TB Gold and QuantiFERON-TB Gold plus had superb specificity in BCG vaccinated individuals, whereas TST had a high specificity in populations where individuals were not BCG vaccinated.

Conditions such as HIV can lower the specificity of IGRAs (Pai et al., 2008). HIV and any other drivers of immunosuppression have been demonstrated to cause false negative TST results (Menzies et al., 2007). A systematic review on the diagnostic accuracy of IGRAs for TB infection in PLWHIV reported QFTs having a better diagnostic odds ratio and area under the curve compared to T-SPOT assays (Chen et al., 2022). Across different studies included in the review, IGRAs had similar sensitivity results with an average of 0.64 for detecting TB infection in PLWHIV. IGRAs were concluded to be useful among PLWHIV, especially for determining who should receive preventive therapy. These tests, however, had very low sensitivity for ruling out TB disease in PLWHIV, and a very poor ability to determine individuals with latent TB who would eventually progress to active TB disease (Chen et al., 2022). Since PLWHIV have higher risk of progressing to active TB disease, the ability for IGRAs to predict progression would be very useful.

When the predictive performance of TSTs and IGRAs for incident TB was compared, no differences were found between these the two tests (Hamada et al., 2023). When performance was stratified based on country-specific TB incidence, performance was similar in countries with a high TB incidence. In low incidence settings, however, IGRAs had a higher predictive performance than the TST. The authors note that the difference observed was largely driven by results from only one study (Diel et al., 2011) and thus any advantage in overall predictive performance may be minimal (Hamada et al., 2023). The findings support recommendations by WHO on the interchangeable use of IGRAs or TSTs in settings where testing for *Mtb* infection is encouraged. As new tests are in development, combining IGRA or TST results with study participant data could improve the performance of these tests, particularly in high TB incidence settings.

Another systematic review that assessed whether IGRAs can predict progression to active TB and made a comparison to TSTs, concluded that both tests did not have any significant prognostic value (Rangaka et al., 2012). Overall, the work carried out on tests for latent TB and the measure of IFN- γ responses demonstrates that IFN- γ alone is not sufficient to determine progression to TB disease or assess TB risk. The identification of an array of biomarkers, with greater predictive power, and the combination of these biomarkers with known risk factors, or existing diagnostic algorithms could greatly enhance prediction of TB risk.

To directly detect *Mtb*-derived antigens in blood samples, Liu and colleagues developed a peptide-based approach using antibody-conjugated porous discoidal nanoparticles (NanoDisks) (Liu et al., 2018). The quantification of serum-derived ESAT-6 and CFP-10 was detected using bench-top MALDI-TOF mass spectrometry (MS). The NanoDisk-MS assay diagnosed 88.3% of TB cases, achieving a specificity of 95.8%, and sensitivities of 91.6% and 85.3% respectively for culture-positive and culture-negative TB cases. The assay also had 90% sensitivity for extrapulmonary TB, which exceeded the diagnostic performance of mycobacterial culture for these TB cases in particular (Liu et al., 2018). The authors hypothesize that this methodology could overcome the pitfalls that contribute to the poor diagnostic performance of *Mtb* antigen immunoassays. Although ESAT-6 and CFP-10 are *Mtb* specific, Pittius and colleagues caution on the use of these antigens as potential diagnostic markers since they can be found in the genomes of *M. leprae*, *M. smegmatis*, *M. africanum*, *M. kansasii*, *M. marinum*, *M. szulgai*, and *M. bovis* (Gey van Pittius et al., 2002). This is further backed-up by work from Geluk et al that demonstrated cross-reactivity between ESAT-6 found in *Mtb*, and its homologue in *M. leprae* (Geluk et al., 2002).

To aid development of more specific immunodiagnostic tests that can differentiate responses to *Mtb*-specific ESAT-6 and CFP-10 from responses to homologues from environmental NTM, children in Malawi with IFN- γ responses were followed up eight years post-BCG vaccination (Hur et al., 2014). A protein signature made up of IL-5, IL-9, IL-13, and IL-17 differentiated latent *Mtb* infection from *M. avium* and *M. kansasii* exposure. These results support the possibility of cytokine or chemokine markers providing another avenue for detection of *Mtb* infection. This also supports that reliance on only IFN- γ responses after stimulation with both ESAT-6 and CFP-10, may not be enough to effectively or accurately determine *Mtb* infection.

A systematic review by the TB Network European Trials group (TBNET) and European Centre for Disease Prevention and Control (ECDC) found a very low number of studies conducted in children assessing the diagnostic performance of TST and IGRAs for the diagnosis of active TB (European Centre for Disease Prevention and Control, 2011). From the literature that was identified, both tests had a very low specificity and sensitivity which did not support their use as rule-out tests for both latent and active TB in children. It was advised that determining whether a child should start or receive preventive treatment should also be based on an overall risk assessment and not just IGRA or TST results alone (European Centre for Disease Prevention and Control, 2011).

Children are more likely to have false-negative or indeterminate IGRA results compared to adults. Children, especially those below the age of five years have immature immune systems that may account for this (Meier et al., 2019). Aside from age, factors such as malnutrition and presence of co-morbidities are all presumed to contribute to the higher rate of indeterminate IGRA results in children (Meier et al., 2019).

Human antibodies against *Mtb* antigens

The user-friendly format of antibody-based detection tests for TB provides potential for them to replace SSM (World Health Organization, 2011b). The development of TB diagnostics based on serology, however, faces several challenges. The first one is that not everyone produces antibodies to the same antigens (Khan et al., 2011). Secondly, exposure to mycobacteria from the environment as well as that from BCG vaccination, can confound results (Khan et al., 2011).

A systematic review on the diagnostic accuracy of commercial serological tests for both pulmonary and extra-pulmonary TB showed no capability for serological tests to improve patient outcomes, as several studies reported high rates of false-positives and negatives (Steingart et al., 2011). Based on these findings, WHO released a policy statement recommending against the use of these tests for

diagnosis of both pulmonary and extra-pulmonary TB (World Health Organization, 2011b). Serological tests are still appealing due to the potential simplicity of such applications in clinical laboratory settings. This has led some investigators to try and develop antibody-based tests that have much better sensitivity and specificity.

On the basis that detection of antibodies against multiple *Mtb* antigens can be fruitful for the detection of *Mtb* infection, Khan and colleagues measured the serological response to 28 recombinant *Mtb* antigens in patients with TB versus healthy controls (Khan et al., 2011). Antibody profiles were generated from a multiplex microbead immunoassay based on the xMAP technology platform. The assay achieved a sensitivity of 90% for TB patients and a specificity of 80%. Only a subset of *Mtb* antigens, however, were found to have some usefulness in TB serodetection. Based on the antibody results, 16 antigens were selected and found to correlate with results from SSM and culture tests.

In another study a custom, portable MBio multiplexed immunoassay system was used to measure antibody responses to 57 selected antigens in blood from more than 750 prospectively enrolled adults with presumptive pulmonary TB in Vietnam and Peru (Broger et al., 2017). IgG antibody responses to a larger set of 132 antigens were also measured in frozen sera, using a multiplex bead based Luminex assay at a reference research laboratory. In line with reports of heterogeneity in antibody responses from different individuals (Khan et al., 2011; Kunnath-Velayudhan et al., 2012; Lyashchenko et al., 1998), Broger and colleagues noted that performance of 7 out of the 22 high ranked antigens were significantly associated with study location (Broger et al., 2017). The authors proceeded to assess several potential antigen models (Broger et al., 2017). Although 3-antigen and multiantigen models performed better than single antigen ones, no combination of antigens, on both multiplex assay platforms, met the performance criterion set for a TB triage test, or for tests aimed at diagnosing TB. When specificity was set at 90%, the sensitivity of the best performing 3-antigen model was 35% (95% CI 31–40).

Human proteins

Acute phase proteins (APPs), such as C-reactive Protein (CRP) and Alpha-1-acid glycoprotein (AGP) have been investigated as possible biomarkers for TB diagnosis (Bishara, 2012). APPs refer to a set of proteins, whose secretion by hepatocytes in the liver, is induced in response to the onset of inflammation. As a result of their elevated levels during inflammation in various disease states, APPs have a low specificity (Santos et al., 2019).

With a selection of studies that used either solid or liquid cultures as a reference standard, a systematic review was carried out to determine the utility of CRP as a POC test. In studies that enrolled out-patients, CRP had a pooled sensitivity of 93% and a pooled specificity of 60% (C. Yoon et al., 2017). These results were similar in individuals with and without HIV. In studies carried out among in-patients, however, pooled sensitivity was 78% and pooled specificity 21%. A follow-up prospective diagnostic accuracy study, by Yoon and colleagues was carried out in PLWHIV in Uganda. With Xpert MTB/RIF as the reference standard, CRP was found to have similar sensitivity to the WHO symptom screen (94% vs. 99%), but higher specificity (ranges 62-80% versus 4-19%) (Christina Yoon et al., 2017).

Their sensitivity, however, remains high and various CRP tests are available in a portable format, making them a very viable option for TB screening (García-Basteiro et al., 2018). One POC test for CRP is the POC LumiraDx CRP test, that has also already been recommended for diagnosis of lower respiratory tract infections across multiple infectious disease and respiratory treatment guidelines in Europe. The test is based on microfluidic technology and makes use of fingerstick blood, venous

blood, or plasma to quantify CRP levels based on the amount of fluorescence detected by the immunoassay.

When performance of the LumiraDx CRP test was compared to the Dimension® Xpand® Plus Integrated Chemistry System (Siemens; Munich, Germany), an established laboratory-based reference method, the former demonstrated equivalent performance in patients with symptoms of infection and any inflammatory conditions (Ellis et al., 2024, 2022).

In 2021, WHO guidelines added CRP as a screening test for TB among PLWHIV, in settings where there is also a high TB burden (“WHO consolidated guidelines on tuberculosis,” n.d.). WHO also recommended that assessments still need to be conducted for use of CRP alone and with other potential screening tools for TB, across different populations. A prospective, cross-sectional study conducted in both Zambia and South Africa assessed the performance of a CRP test in high TB burden settings (Ruperez et al., 2023). Alone, CRP did not meet the sensitivity benchmarks set by WHO for a screening test. This led the authors to suggest that combining the CRP test with symptom screening, for example, may improve its ability to identify TB cases within communities.

Host proteins such as cytokines and chemokines are another category of TB diagnostic biomarkers reported quite frequently in literature (MacLean et al., 2019b). Out of 29 promising biomarkers measured in QFT supernatants using the Luminex cytokine assay, eight distinguished active TB from latent TB (Chegou et al., 2009). These results potentially open up another diagnostic use for the QFT test. EGF and sCD40L were the two single markers with the greatest discriminatory power between both disease states while EGF_{Nil}, EGF_{Ag-Nil}, and MIP-1 β _{Ag-Nil} emerged as the top three-marker combination. The authors suggest the use of a modified QFT test that measures a combination of either of these 3 biomarkers EGF, sCD40L, MIP-1 β , vascular endothelial growth factor (VEGF), TGF- α , or IL-1 α) holds promise for development of a TB diagnostic test (Chegou et al., 2009).

Using a Luminex multiplex immunoassay, the baseline concentrations of 74 host markers in plasma samples from TB patients were compared to levels in patients with other respiratory diseases (Jacobs et al., 2016). Eighteen of these markers individually classified TB with AUROCs of 0.80 or above. When patient data was fitted into general discriminant analysis models, irrespective of HIV status, combinations of up to 6 different markers performed best. The most promising set of six markers were NCAM, SAP, IL- 1 β , sCD40L, IL-13, and Apo A-1, that diagnosed TB disease with a sensitivity 100% (95% CI, 86.3-100%) and specificity of 89.3% (95%CI, 67.6-97.3%). The positive and negative predictive values of the six-marker biosignature were 87.5% (95% CI, 66.5-96.7%) and 100% (95%CI, 83.4-100%), respectively.

In another study based on AE-TBC cohort that was also studied in this thesis (Chapter 4), twenty-two host protein biomarkers in serum were assessed, using a multiplex Luminex assay, to identify which ones could differentiate adults with TB from those with other respiratory diseases (Chegou et al., 2016). These were all individuals, with symptoms suggestive of TB, seeking care from primary health care clinics in five different African countries: South Africa, The Gambia, Uganda, Malawi, and Namibia. A seven-marker serum protein signature, for TB screening, was discovered that had a sensitivity of 94%, specificity of 73.3%, positive predictive value of 61% and a negative predictive value of 96%, irrespective of HIV status and study site.

A follow-up by Chegou and colleagues evaluated 12 previously identified biomarkers, from QFT supernatants (Chegou et al., 2009) in the same AE-TBC cohort of individuals presenting for care at primary health care clinics (Chegou et al., 2016). This follow-up study also included individuals from Ethiopia. A four-marker biosignature, comprised of unstimulated levels of IFN- γ and TGF- α and the

antigen specific levels of IL-1Ra and MIP-1 β had the highest accuracy for TB diagnosis with a sensitivity of 81.3% and a specificity of 56.0%. This four-marker protein signature attained the highest diagnostic accuracy irrespective of HIV status and study site. The authors note that no single biomarker may be sufficient for TB diagnosis, since even unstimulated levels of IFN- γ were not sufficient in this study, even though they have shown promise before (Chegou et al., 2018a).

Oliveira de Souza and colleagues applied the molecular degree of perturbation (MDP) scoring method to plasma concentrations of 15 protein and lipid mediator biomarkers measured by ELISA in active TB cases and healthy uninfected donors from China and India (Oliveira-de-Souza et al., 2019). The MDP method gives an indication of differences in immune and inflammatory responses against *Mtb* and has been an approach applied to gene expression data for transcriptomic analyses (Berry et al., 2010b; Prada-Medina et al., 2017). Discriminant analyses based on canonical correlations highlighted several biomarkers that were different between active TB cases and healthy uninfected controls in both countries (Oliveira-de-Souza et al., 2019). In Indian study participants, sIL-1RI contributed most to differences seen, with significantly higher values in active TB patients compared to uninfected controls. In study participants from China, the marker with the highest discriminatory power was LXA4. In study participants from both countries, LXA4, IFN- β and TNF were the markers that drove a distinction in inflammatory responses. All these three markers have been identified as playing significant roles in TB pathogenesis (Mayer-Barber and Sher, 2015).

With a renewed interest in TB blood-based triage tests versus sputum-based tests for TB diagnosis, Ahmad and colleagues sought to evaluate a blood protein-based immunoassay that can be deployed at the POC (Ahmad et al., 2019). A four-protein signature, comprised of IL-6, IL-8, IL-18, and VEGF distinguished active TB from other TB-like disease in adult patients from Tanzania and The Philippines. The authors note, however, that the limited analytical sensitivity of the Luminex assay resulted in many missing values for these 4 markers, from the plasma samples in the discovery cohort. This may be a similar issue that could easily be faced when attempts are made to downscale and adapt any technology for immunoassays that is aimed at measuring proteins within primary health care settings, and at the POC. If these proteins were to be measured in smaller volumes, such as from a fingerstick of blood, they would require methods that are more sensitive and have a wider analytical range. Mass spectrometry would provide room to identify more novel proteins, however Simoa assays are far more sensitive. The Simoa HD-1 platform can provide results in under an hour, but the current costs of a fully automated platform go beyond the optimal target of <US\$ 4.00 and minimum target of < US\$ 6.00 for a POC test (World Health Organization, 2014).

Proteins in plasma have also demonstrated the capability to predict who will eventually progress to active TB. With the SOMAscan multiplexed assay, 361 proteins had differential abundance between South African adolescents from the ACS study who developed TB disease and those who remained healthy (Penn-Nicholson et al., 2019a). TB Risk Model 5 (TRM5), a five-protein signature for risk of progression to TB, was discovered that had an AUC of 96% as a prognostic marker within six months of TB diagnosis. To promote ease into POC development, the authors discovered an even more parsimonious 3-protein pair-ratio (3PR) signature, that could predict progression to TB within a year of diagnosis. Even though neither TRM5 nor 3PR met the minimal WHO benchmark criteria for an incipient TB test, the study would be the first to report on validated prognostic protein biomarkers, proof-of-concept for identification of protein biomarkers for TB risk.

Through consolidated research efforts from studies such as ScreenTB, AE-TBC, and PredictTB, serum biomarkers have also reached the field-testing stage. An example would be the multi-biomarker test (MBT) test that is one of the POC tests under evaluation through the TriageTB study (Richardson et

al., 2023). The MBT test is a prototype of a lateral flow-based device that measures the individual concentrations of three TB-associated biomarkers: CRP, SAA, and IP-10 in fingerstick blood. Results from the test are accessible through a standalone portable and battery-operated reader (Richardson et al., 2023).

The presence of readily available commercial platforms makes it relatively easy to measure cytokines, chemokines, and other metabolites in serum (MacLean et al., 2019). The downside of this, however, has been repetition of work within the literature due to various researchers carrying out studies on the same limited range of commercially available panels. Unbiased approaches such as 'omics' techniques have now been used to enhance the likelihood of discovering novel biomarkers (MacLean et al., 2019). The focus of this thesis has been the evaluation of multiple host blood mRNA signatures, identified through transcriptomic studies. The overarching aims of the research done has been to determine the potential of these transcriptomic signatures as non-sputum based biomarkers for development of a POC triage test for TB, and as non-sputum based tools for tracking responses during TB treatment.

Transcriptomic signatures

The application of 'omics' techniques for the identification of biomarkers has greatly expanded over the past decade, with transcriptomics being at the forefront. This has ultimately led to the discovery of more than 70 transcriptomic signatures (Nogueira et al., 2022). A search in Medline was conducted for studies published between 1 January 2005 and 31 October 2023 using search terms for "blood", "transcriptomic", "signature", "TB", and "diagnosis". Forty-five studies (**Table 1.1**) describing up to 60 signatures, from discovery cohorts, were identified. This demonstrates that many researchers worldwide have used gene expression profiling to identify mRNA transcripts that are differentially expressed between TB patients and different control groups. Researchers have gone a step further to use different statistical and machine learning methods to train transcriptomic signatures comprising multiple transcripts for TB diagnosis and validate their findings in external cohorts that are entirely separate/independent of their discovery cohorts. The performance of most of these signatures has been previously evaluated in systematic reviews conducted by members of the SATVI research group (Mendelsohn et al., 2023; Mulenga et al., 2020) as well as other researchers (Warsinske et al., 2019). Currently, there are no blood-based tests for the diagnosis of active TB that have been endorsed by WHO or are in clinical use globally (Hamada et al., 2022). Emerging evidence, however, demonstrates the potential for transcriptomic profiling, carried out on whole blood samples, to address the need for POC non-sputum based triage tests.

Gene expression in blood leukocytes changes in response to conditions experienced by the host, including infections, inflammation, and stress. Host blood transcriptomic signatures mirror the host immune response and can detect active TB (Berry et al., 2010b; Mistry et al., 2007), predict risk of progression to TB (Darboe et al., 2018; Scriba et al., 2021a; Zak et al., 2016), and monitor treatment responses (Heyckendorf et al., 2021; Thompson et al., 2017a). A shift to the discovery of signatures comprised of six transcripts or less may facilitate easier detection and translation of these signatures into POC triage tests (Hong et al., 2022). This is evidenced by the selection of parsimonious signatures such as Sweeney3 (Sweeney et al., 2016) by Cepheid, and RISK6 (Bayaa et al., 2021; Penn-Nicholson et al., 2020a) by QuantuMDx for potential commercial development of diagnostics based on transcriptomic signatures (Branigan et al., 2020; Hong et al., 2022; Nathavitharana et al., 2022).

Considerable efforts have been made to measure and detect the 3-gene Sweeney3 signature (Sweeney et al., 2016) using the Xpert MTB Host Response (Xpert-MTB-HR) cartridge system. In its first accuracy study, the Xpert-MTB-HR-Prototype performed similarly to CRP, when tested in people

living with HIV (PLWHIV), reporting with symptoms suggestive of TB (Södersten et al., 2021). The Xpert-MTB-HR-Prototype cartridge was evaluated further in fingerpick blood samples obtained from individuals presenting for care at primary health centres in South Africa, The Gambia, Uganda, and Vietnam (Sutherland et al., 2022). The diagnostic accuracy of this prototype in distinguishing microbiologically-confirmed TB cases from individuals with other respiratory diseases met the WHO minimal TPP for a triage test. The Xpert-MTB-HR-Prototype cartridge was also able to detect changes in gene expression of the 3-gene signature in whole blood samples of TB patients taking anti-tubercular medication, and classified as cured at the end of treatment (Zimmer et al., 2021).

The future application of molecular diagnostic tests, especially at the POC, is said to lie with the development of technology that will allow successful translation of this research work into clinical settings (Gliddon et al., 2016). DNA nanotechnology systems such as the 'toehold-mediated strand displacement' method, and the use of nanoparticle-based sensors, have emerged as methodologies that can make it easier and much cheaper to detect potential genetic-based biomarkers (Gliddon et al., 2016). One promising example of these nanoparticle-based biological sensors are 'quantum dots' (QDs; also known as fluorescent inorganic semiconductor nanoparticles) that have already been incorporated into assays for nucleic acid detection (Silvi and Credi, 2015; Zhou, 2012).

Based on the above, Gliddon et al developed a multiplexed assay that makes use of fluorescent QDs as signal transducers and toehold-mediated strand displacement as the core switching mechanism (Gliddon et al., 2016). The two transcripts selected for detection by this assay were GBP6 and TMMC1, derived from a 44 gene signature that could distinguish active TB from individuals with ORDs (Kaforou et al., 2013). With GBP6 being upregulated and TMMC1 being downregulated in individuals with active TB, the application of a disease risk score (Gliddon et al., 2021; Kaforou et al., 2013) was employed to ease interpretation of gene expression readouts from this multiplexed assay (Gliddon et al., 2016).

To advance POC diagnostic development for TB, we need both enhanced technology for transcriptomic signatures and more evidence confirming their ability to distinguish active TB patients from those with similar clinical symptoms. (Darboe et al., 2019; Hong et al., 2022; MacLean et al., 2019b). This will help streamline who should be a priority for further definitive tests and can therefore start TB treatment promptly.

Table 1.1: Characteristics of studies on transcriptomic signatures for active TB

Publication	Study design	Control group	Age group	Sample type	Geography
(Anderson et al., 2014)	Case control	OD	Children	Whole blood	Malawi, South Africa
	Case control	OD	Children	Whole blood	Kenya
(Berry et al., 2010b)	Cross-sectional	LTBI	Adults	Whole blood	South Africa, Malawi
	Cross-sectional	OD	Adults	Whole blood	
(Bloom et al., 2013)	Case control	OD+HC	Adults	Whole Blood	South Africa
(Cai et al., 2014)	Case control	HC	Adults	PBMC	China
	Case control	LTBI	Adults	PBMC	China
(Chendi et al., 2023)	Longitudinal, observational prospective cohort	OD	Adults	Whole blood	Norway
(da Costa et al., 2015)	Case control	OD	Adults	Whole Blood	South Africa
	Case control	OD	Adults	Whole Blood	
(Darboe et al., 2018)	Case control	LTBI	Adults	PBMC	South Africa
(Darboe et al., 2019)	Case control	No recurrent TB	Adults	PBMC	South Africa
(Dawany et al., 2014)	Case control	HC	Adults	Whole Blood	UK
	Case control	LTBI	Adults	Whole Blood	South Africa
	Case control	LTBI+HC	Adults	Whole Blood	UK
	Case control	LTBI+HC	Adults	Whole Blood	UK
	Case control	LTBI	Adults	Whole Blood	South Africa
	Case control	HC	Adults	Whole Blood	UK
(de Araujo et al., 2021)	Case control	HC	Adults	Whole Blood	UK
	Case control	LTBI	Adults	Whole Blood	South Africa
	Case control	HC	Adults	Whole Blood	Germany
	Case control	HC	Adults	Whole Blood	UK, France
(Francisco et al., 2017)	Case control	OD	Adults	Whole Blood	China
	Case control	OD	Adults	Whole Blood	China
	Case control	OD	Adults	PBMC	China
	Case control	OD	Adults	PBMC	China

	Case control	HC	Adults	Whole Blood	China
	Case control	HC	Adults	Whole Blood	China
(Gupta et al., 2020)	Prospective cohort	Non-progressors (no TB)	Adults	Whole blood	UK
	Case-control	Non-progressors (no TB)	Adolescents	Whole blood	South Africa
	Case-control	Non-progressors (no TB)	Adults	Whole blood	South Africa, Ethiopia, The Gambia
	Prospective cohort	Non-progressors (no TB)	Adults	Whole blood	UK
(Ho et al., 2020)	Case-control	LTBI + HC	Adults	Whole blood	Vietnam
(Hoang et al., 2021)	Prospective cohort	OD	Adults	Whole blood	UK
(Huang et al., 2022)	Case control	HC	Adults	PBMC	China
	Case control	OD+HC	Adults	PBMC	China
(Kaforou et al., 2013)	Case control	LTBI	Adults	Whole blood	South Africa, Malawi
	Case control	OD	Adults	Whole blood	South Africa, Malawi
	Case control	LTBI	Adults	Whole blood	South Africa, Malawi
	Case control	OD	Adults	Whole blood	South Africa, Malawi
(Kaipilyawar et al., 2022)	Case control	LTBI	Adults	Whole blood	India
(Kulkarni et al., 2021)	Case control	HC	Adults	Whole blood	India
	Case control	HC	Adults	Whole blood	Uganda
(Lee et al., 2016)	Case control	LTBI+HC	Adults	PBMC	Taiwan
	Case control	LTBI+HC	Adults	PBMC	Taiwan
(Leong et al., 2018)	Case control	LTBI	Adults	Whole blood	India
	Case control	LTBI	Adults	Whole blood	India
	Case control	LTBI	Adults	Whole blood	India
	Case control	LTBI	Adults	Whole blood	India

	Case control	LTBI	Adults	Whole blood	India
	Case control	LTBI	Adults	Whole blood	India
	Case control	LTBI	Adults	Whole blood	India
	Case control	LTBI	Adults	Whole blood	India
(Lesho et al., 2011)	Case control	LTBI+HC	Adults	PBMC	USA
(Liu et al., 2018)	Case control	LTBI	Adolescents and adults	PBMC	China
(Lyu et al., 2023)	Case-control	HC	Adults	Whole blood	China
	Case-control	HC	Adults	Whole blood	China
(Maertzdorf et al., 2012)	Case control	LTBI	Adolescents and adults	Whole blood	South Africa
	Case control	LTBI	Adolescents and adults	Whole blood	South Africa
(Mendelsohn et al., 2021)	Prospective cohort	LTBI	Adults	Whole blood	South Africa
	Prospective cohort	HC	Adults	Whole blood	South Africa
	Prospective cohort	HC	Adults	Whole blood	South Africa
	Prospective cohort	OD	Adults	Whole blood	South Africa
(Natarajan et al., 2022)	Case control	LTBI	Adults	Whole blood	Malawi
	Case control	LTBI	Adults	Whole blood	South Africa
(Pan et al., 2017)	Case control	HC	Adults	PBMC	China
	Case control	LTBI	Adults	PBMC	China
(Penn-Nicholson et al., 2020b)	Prospective cohort	HC	Adolescents	Whole blood	South Africa
	Case control	LTBI + HC	Adults	Whole blood	South Africa
	Prospective cohort	OD	Adults	Whole blood	South Africa
(Perumal et al., 2021)	Case control	LTBI	Adults	Whole blood	India, UK
(Petrilli et al., 2020)	Case control	LTBI + OD + HC	Adults	Whole blood	Brazil
(Qiu et al., 2022)	Multicohort analysis (case-control)	HC	Adults	Whole blood	South Africa
	Multicohort analysis (case-control)	HC	Adults	Whole blood	UK

	Multicohort analysis (case-control)	HC	Adults	Whole blood	Indonesia
	Multicohort analysis (case-control)	HC	Adults	Whole blood	South Africa, Malawi
	Multicohort analysis (case-control)	HC	Adults	Whole blood	India
	Multicohort analysis (case-control)	HC	Adults	Whole blood	India
	Multicohort analysis (case-control)	HC	Adults	Whole blood	China
	Multicohort analysis (case-control)	HC	Adults	Whole blood	Taiwan
	Multicohort analysis (case-control)	HC	Adults	Whole blood	Mali
	Multicohort analysis (case-control)	HC	Adults	Whole blood	Germany
	Multicohort analysis (case-control)	HC	Adults	Whole blood	South Africa, Malawi
(Rajan et al., 2019)	Case control study	LTBI + HC	Adults	Whole blood	Uganda
(Sambarey et al., 2017)	Case control	LTBI	Adults	Whole blood	Malawi
	Case control	OD	Adults	Whole blood	Malawi
	Case control	LTBI	Adults	Whole blood	South Africa
	Case control	OD	Adults	Whole blood	South Africa
(Satproedprai et al., 2015)	Case control	HC	Adults	Whole blood	Thailand
(Serrano et al., 2016)	Case control	LTBI+HC	Adults	PBMC	Mexico
(Södersten et al., 2021)	Case control	OD	Adults	Whole blood	Peru, South Africa
(Sutherland et al., 2022)	Prospective cohort	OD	Adults	Whole blood	The Gambia, South Africa, Uganda, Vietnam
(Sweeney et al., 2016)	Multi-cohort analysis (Case-control)	HC	Adults	Whole blood	Gambia
	Multi-cohort analysis (Case-control)	HC	Adolescents and adults	Whole blood	Germany

	Multi-cohort analysis (Case-control)	HC	Adults	Whole blood	Indonesia
	Multi-cohort analysis (Case-control)	LTBI	Adults	Whole blood	Gambia
	Multi-cohort analysis (Case-control)	OD	Adults	Whole blood	Germany
(Tornheim et al., 2020)	Case control	HC	Children	Whole blood	India
(Turner et al., 2020)	Prospective cohort	OD	Adults	Whole blood	South Africa
(Walter et al., 2016)	Case control	OD	Adults	Whole blood	USA
	Case control	LTBI	Adults	Whole blood	USA
	Case control	OD+LTBI	Adults	Whole blood	USA
	Case control	LTBI	Adults	Whole blood	USA
	Case control	OD	Adults	Whole blood	USA
	Case control	OD	Adults	Whole blood	USA
	Case control	OD+LTBI	Adults	Whole blood	USA
	Case control	LTBI	Adults	Whole blood	USA
	Case control	OD	Adults	Whole blood	USA
(Warsinske et al., 2018)	Case control nested in a prospective study	OD+HC+LTBI	Adolescents and adults	Whole blood	China
(Wu et al., 2023)	Prospective	LTBI + OD + HC	Adults	Whole blood	China
(Wu et al., 2007)	Case control	LTBI	Adults	PBMC	USA
(Zak et al., 2016)	Prospective cohort	HC	Adolescents	Whole blood	South Africa
	Case control nested in a prognostic study	LTBI	Adults	Whole blood	UK
	Case control nested in a prognostic study	LTBI	Adults	Whole blood	South Africa
	Case control nested in a prognostic study	HC	Adults	Whole blood	UK
	Case control nested in a prognostic study	LTBI	Adults	Whole blood	Malawi

	Case control nested in a prognostic study	LTBI	Adults	Whole blood	South Africa
	Case control nested in a prognostic study	HC	Adults	Whole blood	UK, France
	Case control nested in a prognostic study	OD	Adults	Whole blood	UK, France
	Case control nested in a prognostic study	OD	Adults	Whole blood	UK, France
	Case control nested in a prognostic study	OD	Adults	Whole blood	UK, France
	Case control nested in a prognostic study	OD	Adults	Whole blood	Malawi
	Case control nested in a prognostic study	OD	Adults	Whole blood	South Africa
	Case control nested in a prognostic study	LTBI	Adults	Whole blood	UK, France

LTBI, latent TB infection. HC, household control. OD, other diseases.

Section 3: Current issues surrounding TB treatment

The third section of this chapter introduces recommended treatment guidelines for children and outlines the way treatment monitoring is done. The emphasis on children is because Chapter 5 of this thesis evaluated transcriptomic signatures for treatment monitoring in children. In this section, the topic of TB treatment shortening and the role that non-sputum biomarkers can play in determining this has also been introduced. Section 3 ends with evidence as reported in the literature that supports transcriptomic signatures as non-sputum tools for tracking host immune responses during TB treatment. Altogether, Section 3 aims to provide sufficient rationale for the evaluation of transcriptomic signatures in children as this is an area that has not been fully addressed yet.

Overview of TB treatment

WHO guidelines on the management of TB in children and adolescents recommend a 4-month regimen comprised of two months of isoniazid, rifampicin and pyrazinamide, with or without ethambutol followed by two months of isoniazid and rifampicin (2HRZ(E)/2HR) for children between three months and 16 years with non-severe TB and no evidence of drug-resistance (World Health Organization, 2022b, p. 4). Children who are HIV negative and are in a setting with a low prevalence of isoniazid resistance, can be treated with HRZ for two months while those from high HIV prevalence settings should be treated with HRZE during the first two months. Irrespective of the setting that children are from, it is recommended that they receive HR in the subsequent two months (World Health Organization, 2022c, p. 5). Children identified as having a severe form of pulmonary TB should receive the standard six-month regimen that also includes ethambutol (2HRZE/4HR) during the intensive phase.

Multiple treatment options have been suggested for adolescents aged 12 years and above (World Health Organization, 2022b, p. 4). One of these includes a 4-month regimen of isoniazid, rifapentine, moxifloxacin and pyrazinamide (HPMZ) (World Health Organization, 2022b, p. 4). For children 16 years and above, either a 4-month HPMZ regimen or the standard 6-month regimen of 2HRZE/4HR is recommended. It is advised that the selection of an adequate treatment regimen should take into consideration clinical factors such as disease severity, the presence of co-morbidities as well as the availability of particular drug regimens within a country (World Health Organization, 2022b, p. 4). Children below six years of age with MDR or RR-TB are recommended to take an all-oral treatment regimen containing bedaquiline (World Health Organization, 2022b, p. 4).

TB treatment shortening

For adults, WHO recommends that individuals with DS-TB be treated using a standard 6-month drug regimen, irrespective of how severe TB disease is at the start (World Health Organization, 2017b). Due to differences in the way the host interacts with *Mtb*, the extent of lung disease, bacterial burden, and immune responses, a minority of individuals are not able to achieve cure from 6 months of treatment. TB disease can be complex due to differences in pathology that people present with. These variations, however, also mean that we can personalise treatment durations and come up with better and more effective treatment plans (Dartois and Rubin, 2022).

Although the WHO Guidelines Development Group (GDG) recommends against the use of FLQ-containing regimens for less than 6 months due to higher rates of relapse, they do acknowledge that being able to shorten the duration of TB treatment is a goal (World Health Organization, 2017). The development of shorter treatment regimens that can maximize treatment completion and cure rates (Churchyard, 2018). Shorter treatment regimens also result in fewer issues around adherence, lower the costs incurred by families and health care systems and reduce risks related to treatment such as drug-related toxicity as well as drug-drug interactions in children and adolescents living with HIV (CALHIV) (Dartois and Rubin, 2022; World Health Organization, 2022c).

The REMoxTB (Gillespie et al., 2014), RIFAQUIN (Jindani et al., 2014), and OFLOTUB (Merle et al., 2014) trials evaluated the later-generation FLQs (gatifloxacin and moxifloxacin) as a substitute for drug regimens containing either ETM or INH. Although the 4-month regimens did not meet the non-inferiority criteria in comparison to the standard 6-month treatment, up to four-fifths of study participants were cured (Dartois and Rubin, 2022; Imperial et al., 2018).

A phase 3 trial, Tuberculosis Trials Consortium Study 31/AIDS Clinical Trials Group A5349 (Study 31/A5349) (Dorman Susan E. et al., 2021) was conducted in 13 countries comparing two 4-month rifapentine regimens to a standard 6-month regimen comprised of rifampin, isoniazid, pyrazinamide, and ethambutol. The latter was administered as the control. Rifapentine with moxifloxacin was non-inferior to the control, whereas rifapentine without moxifloxacin was inferior to the control (Dorman Susan E. et al., 2021).

In a pooled analysis of individual participant datasets from these three aforementioned trials, Imperial and colleagues were able to determine the characteristics of those who were cured following the 4-month regimen (Imperial et al., 2018). Based on baseline participant characteristics and features of disease severity, the study population was classified into easy and hard-to-treat cases (Imperial et al., 2018). These findings support a stratified approach to determining treatment duration (Dartois and Rubin, 2022).

The subject of individualized, as opposed to standard treatment duration is increasingly prominent and appears to be operationalized to a greater extent in low-burden and well-resourced areas (Stadler, 2022). The heightened interest in individualized treatment durations stems from substantial heterogeneity in individuals' responses to *Mtb*, leading to varying durations required for achieving TB cure. Having treatment durations that are specific for individuals could be particularly helpful for those that have drug-resistant TB (Heyckendorf et al., 2021), especially with the side effects associated with these drugs (**Table 1.2**).

Table 1.2: Approved drugs for TB treatment

Disease indication	Drug or drug class	Route of administration	Notable side-effects
Drug susceptible TB (DS-TB)	Isoniazid (INH)	Oral	Hepatotoxicity, peripheral neuropathy [#] , optic neuritis, CYP450 interactions
	Pyrazinamide (PZA)	Oral	GI disturbance, hepatotoxicity, gout
	Rifampicin (RIF)	Oral	Hepatotoxicity, nausea, dyspepsia, abdominal pain, rash, CYP450 interactions
	Ethambutol (EMB)	Oral	Optic neuropathy
	Rifapentine (RPT)	Oral	Flu-like reactions, hepatotoxicity
Multidrug resistant (MDR) TB (resistant to INH and RIF)	Fluoroquinolones (levofloxacin, moxifloxacin)	Oral	QT prolongation, tendonitis, hypoglycaemia, psychiatric disturbance
	Aminoglycosides (streptomycin, amikacin)	Injectable (i.m or i.v.)	Nephrotoxicity, ototoxicity
	Rifabutin (RBT)	Oral	hepatic injury
	Capreomycin	Injectable (i.m or i.v.)	liver injury
	Pyrazinamide	Oral	GI disturbance, hepatotoxicity, gout
	Ethambutol	Oral	Optic neuropathy
	Thioamides (ethionamide, prothionamide)	Oral	Psychiatric disturbance, peripheral neuropathy, hepatotoxicity, hypothyroidism, dysglycaemia
	Para-amino-salicylate	Oral	Hepatotoxicity, dysglycaemia, hypothyroidism
	Cycloserine	Oral	Psychiatric disturbance, peripheral neuropathy
	Bedaquiline	Oral	QT prolongation, CYP450 interactions
	Nitroimidazole delamanid	Oral	QT prolongation, CYP450 interactions
Extensively drug resistant (XDR) TB (resistant to INH, RIF, fluoroquinolones and one injectable)	Nitroimidazole pretomanid	Oral	gastrointestinal symptoms, hepatic disorders, skin or subcutaneous tissue disorders, headache
	Clofazimine	Oral	Hepatotoxicity, GI disturbance, neurological disturbance, QT prolongation, altered skin pigmentation
	Linezolid	Oral	Myelosuppression, dysglycaemia, peripheral neuropathy [#] , optic neuropathy
	MDR-TB drugs as appropriate based on drug susceptibility profile		

Adapted from ("Capreomycin," 2012; Dartois and Rubin, 2022; Gill et al., 2022; Horne et al., 2011; Occhinieri et al., 2022; Pease et al., 2018). #: given with pyridoxine prophylaxis to ameliorate risk

Different stratification approaches to determine TB treatment durations should be tested in clinical trials, especially for paediatric TB, extra-pulmonary TB, and drug-resistant TB. Stratification based on molecular diagnostics and non-sputum based biomarkers could also be tested alongside their feasibility to be included as part of TB drug monitoring algorithms (Churchyard, 2018). Predict TB (using biomarkers to predict TB treatment duration), another treatment-shortening clinical trial, used PET-CT scans to stratify people into severity groups and then at month two randomised those with less severe TB into a 4-month vs a 6-month arm.t (R. Y. Chen et al., 2017). One aim of the Predict TB trial is to identify immunological and transcriptomic signatures of treatment outcomes. Chen and colleagues believe that these markers have more capability to be scaled up worldwide and could be cheaper as well as more readily accessible than PET-CT scans. Biomarkers could help individualize treatment durations, based on how individuals respond to a drug regimen (Migliori et al., 2021). This could be especially helpful for earlier stages of TB, such as sub-clinical disease, since putting these individuals on an entire 6-month regimen may not be necessary, could be refused or not adhered to.

Trials that have addressed shorter treatment regimens and different drug combinations have shed light on suitable treatment guidelines (Migliori et al., 2021). These studies, however, have majorly been conducted in adult populations. Studies in children remain scarce due to limitations in collecting respiratory samples and identifying many cases of microbiologically confirmed disease (Turkova et al., 2022). Any recommendations made regarding treatment shortening in children continue to result from trials carried out in adults (Turkova et al., 2022). The success seen from trials on shorter treatment regimens in adults have shown that there are similarities between smear-negative, paucibacillary TB in adults and minimal TB disease in children. The paucibacillary nature of TB disease in children provides the possibility for treatment regimens of less than 6 months. As an example, the SHINE trial (Chabala et al., 2018; Turkova et al., 2022), which is the subject of my studies in chapter 5, demonstrated that four months of treatment for children with drug-sensitive, non-severe TB was non-inferior to six months treatment. This led to WHO making changes in March 2022 to its consolidated guidelines on the management of TB in children and adolescents (“WHO consolidated guidelines on tuberculosis,” n.d.).

The potential of transcriptomic signatures to aid decisions on treatment shortening for paediatric TB is an area that is yet to be addressed. We aim to contribute to this, in Chapter 5 of this thesis, through a sub-study involving children from the SHINE clinical trial (Turkova et al., 2022). The SHINE trial is the only treatment shortening trial to date that has been conducted solely in a paediatric cohort and provides a unique opportunity to assess whether transcriptomic signatures can track responses to treatment in children.

Monitoring TB treatment

Timely initiation of anti-tubercular treatment forms one of the backbones of TB control programmes. Monitoring patient responses to treatment aids in detecting potential drug resistance, facilitating a prompt transition to more effective treatment regimens. This approach also provides insights into the likely causes of treatment failure, such as poor adherence or drug resistance, leading to improved patient management and enhanced treatment outcomes (Gaikwad and Gaikwad, 2018; Parwati et al., 2022).

From WHO recommendations on monitoring and evaluation of children on treatment, it is advised that evidence of culture conversions be shown for children with microbiologically confirmed TB (World Health Organization, 2022c). WHO guidelines recommend treatment monitoring by culture and SSM occur on a monthly basis (World Health Organization, 2023). Results on smear and culture conversion may be challenging for children without culture-confirmed TB. This represents majority of

paediatric TB cases since it is usually paucibacillary. For these children and those who have converted to negative culture results, monitoring the resolution of clinical symptoms and determining weight gain may be more useful as indicators of improvement, in comparison to attempts at repeated respiratory sampling (World Health Organization, 2022c). Repeat radiological assessments are not considered to be mandatory since the abnormalities seen can persist even after treatment completion. The presence of radiological abnormalities is therefore not taken as a definitive indicator of treatment failure or of poor treatment responses (World Health Organization, 2023). For children with MDR-TB on regimens with two or more drugs that prolong the QT interval, it is recommended that electrocardiography be done, and particular biochemical tests made present (World Health Organization, 2022c).

In adults, monitoring of sputum with direct microscopy is suggested to occur at two, five and six months (World Health Organization, 2023). Culture tests are advised to be done alongside smear microscopy to determine possible resistance. Depending on the medication that a patient is on, liver and kidney function tests alongside additional blood tests should be done (World Health Organization, 2023). One example would be an electrocardiogram (ECG) for patients on 6(H)RZE-Lfx with the risk of a prolonged QT interval prolongation.

Although SSM has a faster turnaround time, sputum cultures are more sensitive at detecting persisting mycobacteria. Even with this, initial culture conversion is not always maintained, with some patients reverting to positive (Gammino et al., 2011). It is therefore noted that culture conversion should not be equated to cure even though it is regarded as the most important marker of improvement.

When a patient is unable to expectorate sputum, which occurs in most paediatric TB cases, there is no possibility for carrying out culture (Parwati et al., 2022). Another downside of the reliance on SSM for treatment monitoring is its inability to distinguish between viable and non-viable bacilli. Also, in patients who are immunosuppressed or have extra-pulmonary involvement, SSM tends to have very poor sensitivity (Gaikwad and Gaikwad, 2018).

Although there is more access to SSM at various levels of healthcare, sputum culture, on either solid or liquid media is the gold standard (Gaikwad and Gaikwad, 2018). Cultures, however, can have a turnaround time of 4-6 weeks, require trained staff and biosafety level 3 laboratory facilities to be carried out. Sputum cultures are also expensive and those performed on solid media can be prone to contamination (Gaikwad and Gaikwad, 2018).

An ultimate shift away from sputum-based methods (**Table 1.3**) for treatment monitoring to the use of markers that can help assess treatment efficacy is now under consideration. The need for biomarker-based non-sputum based assays for TB is highlighted in the recent WHO TPPs for TB treatment monitoring and optimization (Gupta-Wright et al., 2023; “Target product profiles for tests for tuberculosis treatment monitoring and optimization,” n.d.). Aside from sputum, radiologic evaluation through chest X-rays is used, but is noted to have a low specificity in individuals with previous lung damage. At present, there is also no standardised interpretation of radiographic changes over the course of TB treatment (Parwati et al., 2022).

Table 1.3: Current methods for TB treatment monitoring

Method	Advantages	Disadvantages
Clinical Assessment	Useful in extrapulmonary and smear-negative pulmonary tuberculosis	Subjective; influenced by concurrent illness

Smear microscopy	Simple, inexpensive, good operational indicator for TB control programmes	Poor sensitivity and specificity
<i>Mtb</i> viability stains	Easy to perform, rapid, Good negative predictive value	Moderate sensitivity and specificity, requirement for fluorescent microscope and reagents
Culture conversion	Highly specific hence considered as 'The gold standard'	Long turn around time, expensive, requires containment facilities

Adapted from; Gaikwad & Gaikwad, 2018

Other promising modalities for the assessment of TB treatment response include immunological biomarkers (Adekambi et al., 2015; Clifford et al., 2015), the Xpert MTB/RIF assay (Friedrich et al., 2013), detection of mycobacterial mRNA in sputum (Desjardin et al., 1999; Li et al., 2010; Montenegro et al., 2014; Moore et al., 1996), and changes in whole blood transcriptomic signatures (Berry et al., 2010b; Blankley et al., 2014; Bloom et al., 2012a; Cliff et al., 2015b). What is even more important is to expand assessment of these aforementioned markers from adults to groups of individuals that need them the most. These include patients with extra-pulmonary TB, paediatric TB cases (Kaforou et al., 2022), people with co-morbidities such as diabetes (Doorn et al., 2022), and those with DR-TB (Parwati et al., 2022).

Need for non-sputum based biomarkers of treatment response

The WHO definitions for treatment response and outcomes in both DS-TB and DR-TB are heavily reliant on sputum conversion from either smear or culture (Rockwood et al., 2016). The drawback with this is that during treatment as patients get better, sputum samples may not be readily available at latter time points. This makes the dependence on sputum an unreliable way of fully evaluating how patients respond to medication throughout the entire course of their treatment (Heyckendorf et al., 2021). Another crucial disadvantage of using culture to predict relapse is that it has very poor predictive value.

The Tuberculosis Network European Trials (TBNET) definition for cure is determined by negative-culture status at 6 months and no subsequent positive culture for up to a year post-treatment (Stadler, 2022). This approach was viewed as being 'simpler' since it reduced the requirements for culture to determine cure or failure, but calls for follow-up for a period of one year after treatment (Stadler, 2022).

Non-sputum biomarkers might be better for showing how well treatment is working and helping decide how long treatment should last (Günther et al., 2021). Biomarkers associated with risk of relapse and TB treatment failure have been identified through various omics studies (Pitaloka et al., 2022). Successful integration into clinical practice, however, requires evidence from longitudinal studies, ensuring broader applicability in diverse populations, including children and those with co-morbidities.

At present, TB diagnosis algorithms, treatment recommendations, and therapeutic drug monitoring guidelines, such as those proposed by units like TBNET, may not universally apply across many African countries. In spite of this, the current research and technological advances we see with molecular-based testing, alongside the push towards non-sputum based biomarkers demonstrates how transcriptomic signatures may make it as part of TB management algorithms. This ideally requires the right support and adequate evidence to show their full implementation in clinical

settings (Stadler, 2022). With the work highlighted in Chapter 5, we aim to generate evidence that supports the use of transcriptomic signatures as non-sputum tools for treatment monitoring in children. The results of this work could encourage other researchers to conduct studies that determine the effectiveness of transcriptomic signatures once integrated into treatment monitoring algorithms.

Transcriptomics for monitoring responses to TB treatment

The lack of sputum production by many patients during TB treatment means that sputum samples cannot be used to determine treatment success or treatment failure for every patient. Current TB treatment monitoring guidelines state that there should be confirmation of microbiological conversion for serial smear and culture tests to negative (Günther et al., 2021). As a result of the reliance on sputum-based treatment monitoring tests, current treatment guidelines miss many patients with treatment failure and those at risk for relapse (Günther et al., 2021). With reduced sensitivity of sputum-based tests, especially at the latter stages of TB treatment, they may not be the most effective or consistent markers for treatment monitoring (Günther et al., 2021). Biomarkers reflecting the host immune response to TB and changes in inflammation can enhance treatment monitoring, predict treatment outcomes, and potentially guide decisions on personalising TB treatment (Günther et al., 2021). Transcriptomic signatures offer a promising alternative, as demonstrated in the current literature.

Cliff and colleagues report the reduced expression of genes in a T- and B-cell signature for active TB two months after some individuals started TB treatment. This reduction in gene expression was even most evident twelve months after treatment initiation (Cliff et al., 2015a). Two successive longitudinal studies (Bloom et al., 2012b; Cliff et al., 2013) reported similar findings, with reductions in gene expression after two weeks of TB treatment. Scores of transcriptomic signatures specific for active and TB treatment progressively reduced up to 6 months post-treatment initiation (Bloom et al., 2012a).

A study carried out on whole blood from the South African Catalysis treatment response cohort led to the discovery of a 5-gene signature, Thompson5 (referred to as RESPONSE5 in the publication), that could differentiate changes in gene expression at baseline from those at week 24 between individuals who were cured and those who had treatment failure (Thompson et al., 2017a). They went on to demonstrate that Thompson5, in combination with results from the Xpert MTB/RIF, was able to predict treatment failures and cures at all time points with 100% accuracy, including week 24 of treatment (Thompson et al., 2017b). Thompson and colleagues also reported a correlation between Thompson5 scores and total glycolytic activity index measurements from PET-CT scans before and during TB treatment (Thompson et al., 2017a). Warsinske and colleagues reported similar findings when they evaluated a 3-gene TB score for treatment monitoring in the same Catalysis cohort (Warsinske et al., 2018). Scores at diagnosis and at the end of treatment (EoT) also significantly correlated with total glycolytic activity index measured using PET-CT scans. Scores were also found to be significantly lower in patients who were radiologically clear in comparison to patients with continuous lung inflammation at the end of 6 months of treatment.

In study participants from a Brazilian cohort and the South African Catalysis cohort, RISK6 scores decreased during treatment and differentiated between pre-treatment and post-treatment blood samples from patients who were cured irrespective of their HIV status (Penn-Nicholson et al., 2020a). In the Catalysis cohort, RISK6 differentiated between individuals with treatment failure and those who achieved clinical cure and were able to predict failure prior to the start of treatment. The RISK6 host-blood transcriptomic signature (Penn-Nicholson et al., 2020a) was also evaluated as a treatment

monitoring tool for adult non-immunocompromised patients with bacteriologically-confirmed active pulmonary TB (Bayaa et al., 2021). These individuals were enrolled as part of a multicentre prospective nested case-control study in Bangladesh, Georgia, Lebanon and Madagascar. In patients with treatment success, RISK6 dropped significantly after just 2 months of treatment and continued to do so up until the point of treatment completion. RISK6 scores for individuals in the DS-TB category decreased throughout treatment.

In one study aimed at discovering a transcriptomic signature for TB treatment monitoring (Long et al., 2021), the authors demonstrated changes in gene expression that occur from the point of diagnosis through end of treatment. A 10-gene signature had lowest scores at the end of treatment and when externally validated, changes in gene expression occurred as early as a week after treatment initiation. The 10-gene signature could predict which patients were cured and which ones would relapse after 24 weeks of treatment.

A 22-gene signature, TB22, was developed from transcriptomic analysis conducted on whole blood samples from two patient cohorts in Germany (Heyckendorf et al., 2021). One cohort had patients with drug-susceptible TB while the other comprised patients with drug-resistant TB. The authors describe a “therapy-end-model”, based on TB22 scores, that was used to determine the individual end-of-therapy time points for TB patients in both their discovery and validation cohorts. The therapy durations indicated by the model were associated with relapse-free cure up to a year after treatment was completed. The association of signature scores with clinical end points, such as smear conversion, culture status, and radiological findings, were used to train the signature and develop the “therapy-end-model”. The therapy-end-model was attributed as a contributing factor to the superior performance reported in the paper, in comparison to other published signatures.

To address improved treatment outcomes for individuals with both TB and diabetes, two blood transcriptomic signatures were identified that can distinguish between people with good and poor treatment outcomes, at the point of TB diagnosis and two weeks after treatment initiation (van Doorn et al., 2022). This showcases the potential for transcriptomic signatures to aid early changes to treatment regimens for patients predicted to have poor outcomes. The application of such transcriptomic signatures has the potential to also improve TB treatment monitoring and care of patients with a diabetes co-morbidity that are predicted to have poor outcomes (van Doorn et al., 2022).

Novel non-sputum-based biomarkers, such as blood transcriptomic signatures, as part of treatment monitoring efforts, could lead to improved patient outcomes. The use of transcriptomic signatures to track treatment responses in children is an aspect that is explored within this thesis (**Chapter 5**). The sub-study involves a cohort of children with paediatric TB who were recruited under the SHINE clinical trial. For transcriptomic signatures to be fully used in clinical settings, researchers need to focus on validating their findings. This means they must show how they discovered the biomarkers and how well these biomarkers perform in independent clinical cohorts.

Summary

Chapter 1 has demonstrated the gaps and issues in TB diagnostics and treatment monitoring highlighting the need for non-sputum biomarkers. An emphasis has been placed on the role that transcriptomic signatures can play in these two areas since this thesis focuses on them as a promising type of non-sputum biomarker. The final section of this chapter outlines the aims and objectives of this project considering the existing literature.

Section 4: Aims and Objectives

This section introduces the aims and objectives of the two major studies that comprise the results chapters of this thesis.

Aim 1 and its objectives focus on a diagnostic study conducted in adults. This project addresses several gaps highlighted earlier in this chapter. Conducting this study in adults from multiple African countries helps evaluate the broader applicability of transcriptomic signatures across the continent. This is crucial because, as shown in **Table 1.1**, most transcriptomic signatures have been discovered, evaluated, and validated primarily in South African cohorts. To develop a POC test effective for the majority of LMICs in Africa, it is essential to identify biomarkers, specifically transcriptomic signatures in whole blood, that exhibit reliable diagnostic performance across multiple countries within different regions of the continent. Rather than discovering new signatures, this study aims to further evaluate existing signatures in a different context and cohort.

Aim 1: To evaluate and compare the diagnostic performance of multiple published transcriptomic gene signatures in cryopreserved samples collected from symptomatic individuals in six African countries who had TB or other respiratory diseases.

Objective 1.1: To select a set of published gene signatures that have the potential to distinguish TB disease from other common respiratory infections.

Objective 1.2: To design an optimal microfluidic qRT-PCR assay and analysis plan that allows measurement of the selected published transcriptomic signatures in whole blood RNA.

Objective 1.3: To compare the diagnostic performance of multiple blood transcriptomic TB biosignatures for differentiating TB or ORDs.

Hypothesis: At least half of the signatures will differentiate between individuals with TB and those with ORDs with sensitivities and specificities that meet the minimum criteria of the WHO TPP for a TB triage test.

Rationale: Many transcriptomic signatures (**Table 1.1**) have been identified in the literature that demonstrate the potential to differentiate individuals with TB from healthy controls (Cai et al., 2014; Huang et al., 2022; Tornheim et al., 2020;) or, in some instances, from those with ORDs (Chendi et al., 2023; Kaforou et al., 2013; Turner et al., 2020) with promising diagnostic performance.

Objective 1.4: To identify the transcriptomic signatures that most accurately differentiate between symptomatic patients with TB and patients who present with ORDs such that the most optimal one can be developed into a POC test.

Hypothesis: At least one transcriptomic signature with the best diagnostic performance to differentiate between TB cases and individuals with ORDs will meet the WHO TPP for a triage test.

Rationale: Multiple transcriptomic signatures in the literature that have been tested on cohorts of TB cases and ORD controls have met and even surpassed the criteria set by WHO for a potential triage test for TB (Mendelsohn et al., 2022b; Penn-Nicholson et al., 2020b; Södersten et al., 2021; Sutherland et al., 2022; Sweeney et al., 2016; Warsinske et al., 2018).

The results from Aim 1 are presented in Chapters 3 and 4.

Aim 2 focuses on the second project of this PhD, which explores the potential of transcriptomic signatures to track treatment responses in children undergoing TB therapy. As discussed in earlier sections of this chapter, especially Section 3, extensive work has been done in adults to evaluate transcriptomic signatures as treatment monitoring tools and to investigate treatment shortening. Although this project did not specifically address treatment shortening, we anticipate that our findings will support the use of transcriptomic signatures to shorten and personalize treatment durations in children. This is based on evidence from adult studies and the paucibacillary nature of TB in children. The SHINE clinical trial demonstrated the feasibility of treatment shortening in cases of non-severe TB, leading to changes in WHO treatment guidelines. We are fortunate to assess transcriptomic signatures as potential treatment monitoring tools in a subset of children from the SHINE trial. The transcriptomic signatures evaluated in Aim 2 were the same ones identified and assessed under Objective 1.1 of Aim 1 of this thesis.

Aim 2: To describe the change in gene expression over time in South African children with non-severe drug-sensitive pulmonary TB on anti-TB chemotherapy in the Shorter Treatment for Non-severe Tuberculosis in African and Indian Children (SHINE) trial.

Objective 2.1: To compare changes in transcriptomic signature scores during TB treatment between children who have TB disease (confirmed and unconfirmed TB disease) and children judged to have unlikely TB disease.

Hypothesis: Children with confirmed and unconfirmed TB are more like to have significantly higher scores that decrease during treatment in comparison to children with unlikely TB.

Rationale: Evidence to support this hypothesis comes from a diagnostic study in children that noted a higher disease risk in children with culture-positive TB in comparison to children with probable and possible TB (Anderson et al., 2014).

Objective 2.2: To compare transcriptomic signature scores at enrolment and on treatment for children with treatment failure/relapse compared to children with successful outcomes.

Hypothesis: Children with treatment failure/relapse will have higher signature scores at enrolment that take longer to resolve (decrease) or do not resolve at all during treatment in comparison to children with successful outcomes.

Rationale: Data to support this comes from a study in individuals with TB and diabetes co-morbidity where individuals with poor treatment outcomes had higher scores before treatment initiation, that did not change much during TB treatment, compared to individuals who had good treatment outcomes (van Doorn et al., 2022).

Objective 2.3: To compare transcriptomic signature scores at the end of treatment in children randomized to the shorter treatment arm (4 months), compared to children randomised to the standard (6 month) treatment arm in order to determine if the additional 2 months of treatment has an impact on transcriptomic signature scores.

Hypothesis: Signature scores at the end of treatment will not be different between children randomized to the shorter (4 month) treatment arm and those in the standard (6 month) treatment arm.

Rationale: No differences in scores are expected between children under both arms since all TB cases were children with non-severe TB and the trial results showed that 4 months of treatment was non-inferior to 6-months with no difference in adverse effects seen for children under both treatment arms. We also trust the randomization process eliminated any bias that would influence outcomes between children in both treatment arms (Chabala et al., 2018; Turkova et al., 2022).

Objective 2.4: To compare transcriptomic signature scores in children at 4 months and at 6 months with healthy controls, recruited from the same geographical area.

Hypothesis: Signature scores in children treated for 4 months or 6 months will not be different to those in healthy controls.

Rationale: Evidence to support this hypothesis comes from treatment monitoring studies in adults that show no difference between end-of-treatment scores from TB patients and scores measured in healthy controls (Bayaa et al., 2021; Penn-Nicholson et al., 2020a).

Objective 2.5: To evaluate the impact of various demographic factors as well as clinical parameters at enrolment and during the course of treatment on change in gene expression or transcriptomic signature scores.

Hypothesis: Some clinical and demographic factors such as age, sex, TB category and HIV status will have a significant impact on signature scores and predict variation in signature scores.

Rationale: A treatment monitoring study in children from India with confirmed TB showed that signature scores increased with an increase in age and that both age and sex were associated with signature scores (Tornheim et al., 2019). Other studies have shown that signature scores are higher in PLWHIV (Mendelsohn et al., 2022). RISK6 scores were higher in male progressors compared to female progressors who were part of an adolescent cohort study (Davies et al., 2024). RISK6 scores were also found to correlate better with risk of TB progression in males compared to females (Davies et al., 2024). In Chapter 4 of this thesis, we demonstrate that HIV status and TB category have a significant impact on scores (Muwanga et al., 2023).

Objective 2.6: To compare the performance of these transcriptomic signatures across different time points for their potential as possible TB treatment monitoring biomarkers.

The results from Aim 2 are presented in Chapter 5.

References

- Abdelaziz, M.M., Bakr, W.M.K., Hussien, S.M., Amine, A.E.K., 2016. Diagnosis of pulmonary tuberculosis using Ziehl-Neelsen stain or cold staining techniques? *J. Egypt. Public Health Assoc.* 91, 39–43. <https://doi.org/10.1097/01.EPX.0000481358.12903.af>
- Abdulgader, S.M., Okunola, A.O., Ndlangalavu, G., Reeve, B.W.P., Allwood, B.W., Koegelenberg, C.F.N., Warren, R.M., Theron, G., 2022. Diagnosing Tuberculosis: What Do New Technologies Allow Us to (Not) Do? *Respiration* 101, 797–813. <https://doi.org/10.1159/000525142>
- Acharya, M.P., Pradeep, S.P., Murthy, V.S., Chikkannaiah, P., Kambar, V., Narayanashetty, S., Burugina Nagaraja, S., Gangadhar, N., Yoganand, R., Satchidanandam, V., 2021. CD38+CD27–TNF- α on Mtb-specific CD4+ T Cells Is a Robust Biomarker for Tuberculosis Diagnosis. *Clin. Infect. Dis.* 73, 793–801. <https://doi.org/10.1093/cid/ciab144>
- Adams, G., 2020. A beginner's guide to RT-PCR, qPCR and RT-qPCR. *The Biochemist* 42, 48–53. <https://doi.org/10.1042/BIO20200034>
- Adekambi, T., Ibegbu, C.C., Cagle, S., Kalokhe, A.S., Wang, Y.F., Hu, Y., Day, C.L., Ray, S.M., Rengarajan, J., 2015. Biomarkers on patient T cells diagnose active tuberculosis and monitor treatment response. *J. Clin. Invest.* 125, 1827–1838. <https://doi.org/10.1172/JCI77990>
- Ahmad, R., Xie, L., Pyle, M., Suarez, M.F., Broger, T., Steinberg, D., Ame, S.M., Lucero, M.G., Szucs, M.J., MacMullan, M., Berven, F.S., Dutta, A., Sanvictores, D.M., Tallo, V.L., Bencher, R., Eisinger, D.P., Dhingra, U., Deb, S., Ali, S.M., Mehta, S., Fawzi, W.W., Riley, I.D., Sazawal, S., Premji, Z., Black, R., Murray, C.J.L., Rodriguez, B., Carr, S.A., Walt, D.R., Gillette, M.A., 2019. A rapid triage test for active pulmonary tuberculosis in adult patients with persistent cough. *Sci. Transl. Med.* 11, eaaw8287. <https://doi.org/10.1126/scitranslmed.aaw8287>
- Ahmed, M.I.M., Ziegler, C., Held, K., Dubinski, I., Ley-Zaporozhan, J., Geldmacher, C., von Both, U., 2019. The TAM-TB Assay—A Promising TB Immune-Diagnostic Test With a Potential for Treatment Monitoring. *Front. Pediatr.* 7.
- Anderson, S.T., Kaforou, M., Brent, A.J., Wright, V.J., Banwell, C.M., Chagaluka, G., Crampin, A.C., Dockrell, H.M., French, N., Hamilton, M.S., Hibberd, M.L., Kern, F., Langford, P.R., Ling, L., Mlotha, R., Ottenhoff, T.H.M., Pienaar, S., Pillay, V., Scott, J.A.G., Twahir, H., Wilkinson, R.J., Coin, L.J., Heyderman, R.S., Levin, M., Eley, B., 2014. Diagnosis of Childhood Tuberculosis and Host RNA Expression in Africa. *N. Engl. J. Med.*
- Asp, M., Bergenstr hle, J., Lundberg, J., 2020. Spatially Resolved Transcriptomes—Next Generation Tools for Tissue Exploration. *BioEssays* 42, 1900221. <https://doi.org/10.1002/bies.201900221>
- Barry, C.E., Boshoff, H.I., Dartois, V., Dick, T., Ehr t, S., Flynn, J., Schnappinger, D., Wilkinson, R.J., Young, D., 2009a. The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat. Rev. Microbiol.* 7, 845–855. <https://doi.org/10.1038/nrmicro2236>
- Barry, C.E., Boshoff, H.I., Dartois, V., Dick, T., Ehr t, S., Flynn, J., Schnappinger, D., Wilkinson, R.J., Young, D., 2009b. The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat. Rev. Microbiol.* 7, 845–855. <https://doi.org/10.1038/nrmicro2236>
- Bayaa, R., Ndiaye, M.D.B., Chedid, C., Kokhraidze, E., Tukvadze, N., Banu, S., Uddin, M.K.M., Biswas, S., Nasrin, R., Ranaivomanana, P., Raherinandrasana, A.H., Rakotonirina, J., Rasolofo, V., Delogu, G., De Maio, F., Goletti, D., Endtz, H., Ader, F., Hamze, M., Ismail, M.B., Pouzol, S., Rakotosamimanana, N., Hoffmann, J., The HINTT working group within the GABRIEL network, Russomando, G., Acosta, C.C.D., Arenas, R., 2021. Multi-country evaluation of RISK6, a 6-gene blood transcriptomic signature, for tuberculosis diagnosis and treatment monitoring. *Sci. Rep.* 11, 13646. <https://doi.org/10.1038/s41598-021-93059-1>
- Berry, M.P.R., Graham, C.M., McNab, F.W., Xu, Z., Bloch, S.A.A., Oni, T., Wilkinson, K.A., Banchereau, R., Skinner, J., Wilkinson, R.J., Quinn, C., Blankenship, D., Dhawan, R., Cush, J.J., Mejias, A., Ramilo, O., Kon, O.M., Pascual, V., Banchereau, J., Chaussabel, D., O'Garra, A., 2010a. An

- interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature* 466, 973–977. <https://doi.org/10.1038/nature09247>
- Berry, M.P.R., Graham, C.M., McNab, F.W., Xu, Z., Bloch, S.A.A., Oni, T., Wilkinson, K.A., Banchereau, R., Skinner, J., Wilkinson, R.J., Quinn, C., Blankenship, D., Dhawan, R., Cush, J.J., Mejias, A., Ramilo, O., Kon, O.M., Pascual, V., Banchereau, J., Chaussabel, D., O’Garra, A., 2010b. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature* 466, 973–977. <https://doi.org/10.1038/nature09247>
- Bigio, J., Kohli, M., Klinton, J.S., MacLean, E., Gore, G., Small, P.M., Ruhwald, M., Weber, S.F., Jha, S., Pai, M., 2021. Diagnostic accuracy of point-of-care ultrasound for pulmonary tuberculosis: A systematic review. *PLOS ONE* 16, e0251236. <https://doi.org/10.1371/journal.pone.0251236>
- Bishara, N., 2012. Chapter 18 - The Use of Biomarkers for Detection of Early- and Late-Onset Neonatal Sepsis, in: Ohls, R.K., Maheshwari, A. (Eds.), *Hematology, Immunology and Infectious Disease: Neonatology Questions and Controversies (Second Edition)*. W.B. Saunders, Philadelphia, pp. 303–315. <https://doi.org/10.1016/B978-1-4377-2662-6.00018-3>
- Blakemore, R., Story, E., Helb, D., Kop, J., Banada, P., Owens, M.R., Chakravorty, S., Jones, M., Alland, D., 2010. Evaluation of the Analytical Performance of the Xpert MTB/RIF Assay. *J. Clin. Microbiol.* 48, 2495–2501. <https://doi.org/10.1128/jcm.00128-10>
- Blankley, S., Berry, M.P.R., Graham, C.M., Bloom, C.I., Lipman, M., O’Garra, A., 2014. The application of transcriptional blood signatures to enhance our understanding of the host response to infection: the example of tuberculosis. *Philos. Trans. R. Soc. B Biol. Sci.* 369, 20130427. <https://doi.org/10.1098/rstb.2013.0427>
- Bloom, C.I., Graham, C.M., Berry, M.P.R., Rozakeas, F., Redford, P.S., Wang, Y., Xu, Z., Wilkinson, K.A., Wilkinson, R.J., Kendrick, Y., Devouassoux, G., Ferry, T., Miyara, M., Bouvry, D., Dominique, V., Gorocho, G., Blankenship, D., Saadatian, M., Vanhems, P., Beynon, H., Vancheeswaran, R., Wickremasinghe, M., Chaussabel, D., Banchereau, J., Pascual, V., Ho, L., Lipman, M., O’Garra, A., 2013. Transcriptional Blood Signatures Distinguish Pulmonary Tuberculosis, Pulmonary Sarcoidosis, Pneumonias and Lung Cancers. *PLOS ONE* 8, e70630. <https://doi.org/10.1371/journal.pone.0070630>
- Bloom, C.I., Graham, C.M., Berry, M.P.R., Wilkinson, K.A., Oni, T., Rozakeas, F., Xu, Z., Rossello-Urgell, J., Chaussabel, D., Banchereau, J., Pascual, V., Lipman, M., Wilkinson, R.J., O’Garra, A., 2012a. Detectable Changes in The Blood Transcriptome Are Present after Two Weeks of Antituberculosis Therapy. *PLOS ONE* 7, e46191. <https://doi.org/10.1371/journal.pone.0046191>
- Bloom, C.I., Graham, C.M., Berry, M.P.R., Wilkinson, K.A., Oni, T., Rozakeas, F., Xu, Z., Rossello-Urgell, J., Chaussabel, D., Banchereau, J., Pascual, V., Lipman, M., Wilkinson, R.J., O’Garra, A., 2012b. Detectable Changes in The Blood Transcriptome Are Present after Two Weeks of Antituberculosis Therapy. *PLOS ONE* 7, e46191. <https://doi.org/10.1371/journal.pone.0046191>
- Boehme, C.C., Nabeta, P., Hillemann, D., Nicol, M.P., Shenai, S., Krapp, F., Allen, J., Tahirli, R., Blakemore, R., Rustomjee, R., Milovic, A., Jones, M., O’Brien, S.M., Persing, D.H., Ruesch-Gerdes, S., Gotuzzo, E., Rodrigues, C., Alland, D., Perkins, M.D., 2010. Rapid Molecular Detection of Tuberculosis and Rifampin Resistance. *N. Engl. J. Med.* 363, 1005–1015. <https://doi.org/10.1056/NEJMoa0907847>
- Boehme, C.C., Nicol, M.P., Nabeta, P., Michael, J.S., Gotuzzo, E., Tahirli, R., Gler, M.T., Blakemore, R., Worodria, W., Gray, C., Huang, L., Caceres, T., Mehdiyev, R., Raymond, L., Whitelaw, A., Sagadevan, K., Alexander, H., Albert, H., Cobelens, F., Cox, H., Alland, D., Perkins, M.D., 2011. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *The Lancet* 377, 1495–1505. [https://doi.org/10.1016/S0140-6736\(11\)60438-8](https://doi.org/10.1016/S0140-6736(11)60438-8)
- Boldi, M.-O., Denis-Lessard, J., Neziri, R., Brouillet, R., von-Garnier, C., Chavez, V., Mazza-Stalder, J., Jatton, K., Greub, G., Opota, O., 2023. Performance of microbiological tests for tuberculosis

- diagnostic according to the type of respiratory specimen: A 10-year retrospective study. *Front. Cell. Infect. Microbiol.* 13. <https://doi.org/10.3389/fcimb.2023.1131241>
- Branigan, D., Deborggraeve, S., Kohli, M., MacLean, E., McKenna, L., Ruhwald, M., 2020. The Tuberculosis Diagnostics Pipeline Report: Advancing the Next Generation of Tools 24.
- Broger, T., Basu Roy, R., Filomena, A., Greef, C.H., Rimmele, S., Havumaki, J., Danks, D., Schneiderhan-Marra, N., Gray, C.M., Singh, M., Rosenkrands, I., Andersen, P., Husar, G.M., Joos, T.O., Gennaro, M.L., Lochhead, M.J., Denking, C.M., Perkins, M.D., 2017. Diagnostic Performance of Tuberculosis-Specific IgG Antibody Profiles in Patients with Presumptive Tuberculosis from Two Continents. *Clin. Infect. Dis.* 64, 947–955. <https://doi.org/10.1093/cid/cix023>
- Burel, J.G., Babor, M., Pomaznoy, M., Lindestam Arlehamn, C.S., Khan, N., Sette, A., Peters, B., 2019. Host Transcriptomics as a Tool to Identify Diagnostic and Mechanistic Immune Signatures of Tuberculosis. *Front. Immunol.* 10. <https://doi.org/10.3389/fimmu.2019.00221>
- Cai, Y., Yang, Q., Tang, Y., Zhang, M., Liu, H., Zhang, G., Deng, Q., Huang, J., Gao, Z., Zhou, B., Feng, C.G., Chen, X., 2014. Increased Complement C1q Level Marks Active Disease in Human Tuberculosis. *PLOS ONE* 9, e92340. <https://doi.org/10.1371/journal.pone.0092340>
- Capreomycin, 2012. , in: *LiverTox: Clinical and Research Information on Drug-Induced Liver Injury*. National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda (MD).
- Chabala, C., Turkova, A., Thomason, M.J., Wobudeya, E., Hissar, S., Mave, V., van der Zalm, M., Palmer, M., Kapasa, M., Bhavani, P.K., Balaji, S., Raichur, P.A., Demers, A.-M., Hoddinott, G., Owen-Powell, E., Kinikar, A., Musoke, P., Mulenga, V., Aarnoutse, R., McIlleron, H., Hesselning, A., Crook, A.M., Cotton, M., Gibb, D.M., SHINE trial team, 2018. Shorter treatment for minimal tuberculosis (TB) in children (SHINE): a study protocol for a randomised controlled trial. *Trials* 19, 237. <https://doi.org/10.1186/s13063-018-2608-5>
- Chegou, N.N., Black, G.F., Kidd, M., van Helden, P.D., Walzl, G., 2009. Host markers in Quantiferon supernatants differentiate active TB from latent TB infection: preliminary report. *BMC Pulm. Med.* 9, 21. <https://doi.org/10.1186/1471-2466-9-21>
- Chegou, N.N., Sutherland, J.S., Malherbe, S., Crampin, A.C., Corstjens, P.L.A.M., Geluk, A., Mayanja-Kizza, H., Loxton, A.G., Spuy, G. van der, Stanley, K., Kotzé, L.A., Vyver, M. van der, Rosenkrands, I., Kidd, M., Helden, P.D. van, Dockrell, H.M., Ottenhoff, T.H.M., Kaufmann, S.H.E., Walzl, G., 2016. Diagnostic performance of a seven-marker serum protein biosignature for the diagnosis of active TB disease in African primary healthcare clinic attendees with signs and symptoms suggestive of TB. *Thorax* 71, 785–794. <https://doi.org/10.1136/thoraxjnl-2015-207999>
- Chegou, N.N., Sutherland, J.S., Namuganga, A.-R., Corstjens, P.L., Geluk, A., Gebremichael, G., Mendy, J., Malherbe, S., Stanley, K., van der Spuy, G.D., Kriel, M., Loxton, A.G., Kriel, B., Simukonda, F., Bekele, Y., Sheehama, J.A., Nelongo, J., van der Vyver, M., Gebrexabher, A., Hailu, H., Esterhuyse, M.M., Rosenkrands, I., Aagard, C., Kidd, M., Kassa, D., Mihret, A., Howe, R., Cliff, J.M., Crampin, A.C., Mayanja-Kizza, H., Kaufmann, S.H.E., Dockrell, H.M., Ottenhoff, T.H.M., Walzl, G., 2018. Africa-wide evaluation of host biomarkers in QuantiFERON supernatants for the diagnosis of pulmonary tuberculosis. *Sci. Rep.* 8, 2675. <https://doi.org/10.1038/s41598-018-20855-7>
- Chen, H., Nakagawa, A., Takamori, M., Abe, S., Ueno, D., Horita, N., Kato, S., Seki, N., 2022. Diagnostic accuracy of the interferon-gamma release assay in acquired immunodeficiency syndrome patients with suspected tuberculosis infection: a meta-analysis. *Infection* 50, 597–606. <https://doi.org/10.1007/s15010-022-01789-9>
- Chen, R.Y., Via, L.E., Dodd, L.E., Walzl, G., Malherbe, S.T., Loxton, A.G., Dawson, R., Wilkinson, R.J., Thienemann, F., Tameris, M., Hatherill, M., Diacon, A.H., Liu, X., Xing, J., Jin, X., Ma, Z., Pan, S., Zhang, G., Gao, Q., Jiang, Q., Zhu, H., Liang, L., Duan, H., Song, T., Alland, D., Tartakovsky, M., Rosenthal, A., Whalen, C., Duvenhage, M., Cai, Y., Goldfeder, L.C., Arora, K., Smith, B., Winter, J., Iii, C.E.B., Group, P.T.S., 2017. Using biomarkers to predict TB treatment duration

- (Predict TB): a prospective, randomized, noninferiority, treatment shortening clinical trial. <https://doi.org/10.12688/gatesopenres.12750.1>
- Chendi, B.H., Jooste, T., Scriba, T.J., Kidd, M., Mendelsohn, S., Tonby, K., Walzl, G., Dyrhol-Riise, A.M., Chegou, N.N., 2023. Utility of a three-gene transcriptomic signature in the diagnosis of tuberculosis in a low-endemic hospital setting. *Infect. Dis. Lond. Engl.* 55, 44–54. <https://doi.org/10.1080/23744235.2022.2129779>
- Chien, H.P., Yu, M.C., Wu, M.H., Lin, T.P., Luh, K.T., 2000. Comparison of the BACTEC MGIT 960 with Löwenstein-Jensen medium for recovery of mycobacteria from clinical specimens. *Int. J. Tuberc. Lung Dis.* 4, 866–870.
- Churchyard, G.J., 2018. A stratified approach to tuberculosis treatment. *Nat. Med.* 24, 1639–1641. <https://doi.org/10.1038/s41591-018-0244-y>
- Cliff, J.M., Kaufmann, S.H.E., McShane, H., Helden, P. van, O'Garra, A., 2015a. The human immune response to tuberculosis and its treatment: a view from the blood. *Immunol. Rev.* 264, 88–102. <https://doi.org/10.1111/imr.12269>
- Cliff, J.M., Kaufmann, S.H.E., McShane, H., van Helden, P., O'Garra, A., 2015b. The human immune response to tuberculosis and its treatment: a view from the blood. *Immunol. Rev.* 264, 88–102. <https://doi.org/10.1111/imr.12269>
- Cliff, J.M., Lee, J.-S., Constantinou, N., Cho, J.-E., Clark, T.G., Ronacher, K., King, E.C., Lukey, P.T., Duncan, K., Van Helden, P.D., Walzl, G., Dockrell, H.M., 2013. Distinct Phases of Blood Gene Expression Pattern Through Tuberculosis Treatment Reflect Modulation of the Humoral Immune Response. *J. Infect. Dis.* 207, 18–29. <https://doi.org/10.1093/infdis/jis499>
- Clifford, V., Zufferey, C., Street, A., Denholm, J., Tebruegge, M., Curtis, N., 2015. Cytokines for monitoring anti-tuberculous therapy: A systematic review. *Tuberculosis* 95, 217–228. <https://doi.org/10.1016/j.tube.2015.01.003>
- Coussens, A.K., Zaidi, S.M.A., Allwood, B.W., Dewan, P.K., Gray, G., Kohli, M., Kredo, T., Marais, B.J., Marks, G.B., Martinez, L., Ruhwald, M., Scriba, T.J., Seddon, J.A., Tisile, P., Warner, D.F., Wilkinson, R.J., Esmail, H., Houben, R.M.G.J., Alland, D., Behr, M.A., Beko, B.B., Burhan, E., Churchyard, G., Cobelens, F., Denholm, J.T., Dinkel, R., Ellner, J.J., Fatima, R., Haigh, K.A., Hatherill, M., Horton, K.C., Kendall, E.A., Khan, P.Y., MacPherson, P., Malherbe, S.T., Mave, V., Mendelsohn, S.C., Musvosvi, M., Nemes, E., Penn-Nicholson, A., Ramamurthy, D., Rangaka, M.X., Sahu, S., Schwalb, A., Shah, D.K., Sheerin, D., Simon, D., Steyn, A.J.C., Anh, N.T., Walzl, G., Weller, C.L., Williams, C.M., Wong, E.B., Wood, R., Xie, Y.L., Yi, S., 2024. Classification of early tuberculosis states to guide research for improved care and prevention: an international Delphi consensus exercise. *Lancet Respir. Med.* 0. [https://doi.org/10.1016/S2213-2600\(24\)00028-6](https://doi.org/10.1016/S2213-2600(24)00028-6)
- Cudahy, P.G.T., Dawson, R., Allwood, B.W., Maartens, G., Wilson, D., 2017. Diagnostic Outcomes After Chest Radiograph Interpretation in Patients With Suspected Tuberculosis and Negative Sputum Smears in a High-Burden Human Immunodeficiency Virus and Tuberculosis Setting. *Open Forum Infect. Dis.* 4, ofx123. <https://doi.org/10.1093/ofid/ofx123>
- da Costa, L.L., Delcroix, M., Dalla Costa, E.R., Prestes, I.V., Milano, M., Francis, S.S., Unis, G., Silva, D.R., Riley, L.W., Rossetti, M.L.R., 2015. A real-time PCR signature to discriminate between tuberculosis and other pulmonary diseases. *Tuberc. Edinb. Scotl.* 95, 421–425. <https://doi.org/10.1016/j.tube.2015.04.008>
- Darboe, F., Mbandi, S.K., Naidoo, K., Yende-Zuma, N., Lewis, L., Thompson, E.G., Duffy, F.J., Fisher, M., Filander, E., van Rooyen, M., Bilek, N., Mabwe, S., McKinnon, L.R., Chegou, N., Loxton, A., Walzl, G., Tromp, G., Padayatchi, N., Govender, D., Hatherill, M., Karim, S.A., Zak, D.E., Penn-Nicholson, A., Scriba, T.J., Team, T.S.C.I., 2019. Detection of Tuberculosis Recurrence, Diagnosis and Treatment Response by a Blood Transcriptomic Risk Signature in HIV-Infected Persons on Antiretroviral Therapy. *Front. Microbiol.* 10. <https://doi.org/10.3389/fmicb.2019.01441>

- Darboe, F., Mbandi, S.K., Thompson, E.G., Fisher, M., Rodo, M., van Rooyen, M., Filander, E., Bilek, N., Mabwe, S., Hatherill, M., Zak, D.E., Penn-Nicholson, A., Scriba, T.J., Matiwane, S., Jaxa, L., Xoyana, N., Schreuder, C., Botes, J., Africa, H., Makhethhe, L., Steyn, M., 2018. Diagnostic performance of an optimized transcriptomic signature of risk of tuberculosis in cryopreserved peripheral blood mononuclear cells. *Tuberculosis* 108, 124–126. <https://doi.org/10.1016/j.tube.2017.11.001>
- Dartois, V.A., Rubin, E.J., 2022. Anti-tuberculosis treatment strategies and drug development: challenges and priorities. *Nat. Rev. Microbiol.* 20, 685–701. <https://doi.org/10.1038/s41579-022-00731-y>
- Davies, L.R.L., Wang, C., Steigler, P., Bowman, K.A., Fischinger, S., Hatherill, M., Fisher, M., Mbandi, S.K., Rodo, M., Ottenhoff, T.H.M., Dockrell, H.M., Sutherland, J.S., Mayanja-Kizza, H., Boom, W.H., Walzl, G., Kaufmann, S.H.E., Nemes, E., Scriba, T.J., Lauffenburger, D., Alter, G., Fortune, S.M., 2024. Age and sex influence antibody profiles associated with tuberculosis progression. *Nat. Microbiol.* 1–13. <https://doi.org/10.1038/s41564-024-01678-x>
- Davis, J.L., Worodria, W., Kitembo, H., Metcalfe, J.Z., Cattamanchi, A., Kawooya, M., Kyeyune, R., Boon, S. den, Powell, K., Okello, R., Yoo, S., Huang, L., 2010. Clinical and Radiographic Factors Do Not Accurately Diagnose Smear-Negative Tuberculosis in HIV-infected Inpatients in Uganda: A Cross-Sectional Study. *PLOS ONE* 5, e9859. <https://doi.org/10.1371/journal.pone.0009859>
- Dawany, N., Showe, L.C., Kossenkova, A.V., Chang, C., Ive, P., Conradie, F., Stevens, W., Sanne, I., Azzoni, L., Montaner, L.J., 2014. Identification of a 251 Gene Expression Signature That Can Accurately Detect M. tuberculosis in Patients with and without HIV Co-Infection. *PLOS ONE* 9, e89925. <https://doi.org/10.1371/journal.pone.0089925>
- de Araujo, L.S., Ribeiro-Alves, M., Wipperfurth, M.F., Vorkas, C.K., Pessler, F., Saad, M.H.F., 2021. Transcriptomic Biomarkers for Tuberculosis: Validation of NPC2 as a Single mRNA Biomarker to Diagnose TB, Predict Disease Progression, and Monitor Treatment Response. *Cells* 10, 2704. <https://doi.org/10.3390/cells10102704>
- de Wetering, D. van, van Wengen, A., Savage, N.D.L., van de Vosse, E., van Dissel, J.T., 2011. IFN- α cannot substitute lack of IFN- γ responsiveness in cells of an IFN- γ R1 deficient patient. *Clin. Immunol.* 138, 282–290. <https://doi.org/10.1016/j.clim.2010.12.005>
- Delogu, G., Goletti, D., 2014. The Spectrum of Tuberculosis Infection: New Perspectives in the Era of Biologics. *J. Rheumatol. Suppl.* 91, 11–16. <https://doi.org/10.3899/jrheum.140097>
- Desikan, P., 2013. Sputum smear microscopy in tuberculosis: Is it still relevant? *Indian J. Med. Res.* 137, 442–444.
- Desjardin, L.E., Perkins, M.D., Wolski, K., Haun, S., Teixeira, L., Chen, Y., Johnson, J.L., Ellner, J.J., Dietze, R., Bates, J., Donald Cave, M., Eisenach, K.D., 1999. Measurement of Sputum Mycobacterium tuberculosis Messenger RNA as a Surrogate for Response to Chemotherapy. *Am. J. Respir. Crit. Care Med.* 160, 203–210. <https://doi.org/10.1164/ajrccm.160.1.9811006>
- Desvignes, L., Wolf, A.J., Ernst, J.D., 2012. Dynamic Roles of Type I and Type II IFNs in Early Infection with *Mycobacterium tuberculosis*. *J. Immunol.* 188, 6205–6215. <https://doi.org/10.4049/jimmunol.1200255>
- Dewi, D.N.S.S., Mertaniasih, N.M., Soedarsono, Ozeki, Y., Artama, W.T., Fihiruddin, Niki, M., Tateishi, Y., Ato, M., Matsumoto, S., 2019. Characteristic profile of antibody responses to PPD, ESAT-6, and CFP-10 of Mycobacterium tuberculosis in pulmonary tuberculosis suspected cases in Surabaya, Indonesia. *Braz. J. Infect. Dis.* 23, 246–253. <https://doi.org/10.1016/j.bjid.2019.07.001>
- Diel, R., Loddenkemper, R., Niemann, S., Meywald-Walter, K., Nienhaus, A., 2011. Negative and Positive Predictive Value of a Whole-Blood Interferon- γ Release Assay for Developing Active Tuberculosis. *Am. J. Respir. Crit. Care Med.* 183, 88–95. <https://doi.org/10.1164/rccm.201006-0974OC>

- Donovan, M.L., Schultz, T.E., Duke, T.J., Blumenthal, A., 2017. Type I Interferons in the Pathogenesis of Tuberculosis: Molecular Drivers and Immunological Consequences. *Front. Immunol.* 8, 1633. <https://doi.org/10.3389/fimmu.2017.01633>
- Doorn, C.L.R. van, Eckold, C., Ronacher, K., Ruslami, R., Veen, S. van, Lee, J.-S., Kumar, V., Kerry-Barnard, S., Malherbe, S.T., Kleynhans, L., Stanley, K., Hill, P.C., Joosten, S.A., Crevel, R. van, Wijmenga, C., Critchley, J.A., Walzl, G., Alisjahbana, B., Haks, M.C., Dockrell, H.M., Ottenhoff, T.H.M., Vianello, E., Cliff, J.M., 2022. Transcriptional profiles predict treatment outcome in patients with tuberculosis and diabetes at diagnosis and at two weeks after initiation of anti-tuberculosis treatment. *eBioMedicine* 82. <https://doi.org/10.1016/j.ebiom.2022.104173>
- Dorman Susan E., Nahid Payam, Kurbatova Ekaterina V., Phillips Patrick P.J., Bryant Kia, Dooley Kelly E., Engle Melissa, Goldberg Stefan V., Phan Ha T.T., Hakim James, Johnson John L., Lourens Madeleine, Martinson Neil A., Muzanyi Grace, Narunsky Kim, Nerette Sandy, Nguyen Nhung V., Pham Thuong H., Pierre Samuel, Purfield Anne E., Samaneka Wadzanai, Savic Radojka M., Sanne Ian, Scott Nigel A., Shenje Justin, Sizemore Erin, Vernon Andrew, Waja Ziyaad, Weiner Marc, Swindells Susan, Chaisson Richard E., 2021. Four-Month Rifapentine Regimens with or without Moxifloxacin for Tuberculosis. *N. Engl. J. Med.* 384, 1705–1718. <https://doi.org/10.1056/NEJMoa2033400>
- Dowdy, D.W., Basu, S., Andrews, J.R., 2013. Is passive diagnosis enough? The impact of subclinical disease on diagnostic strategies for tuberculosis. *Am. J. Respir. Crit. Care Med.* 187, 543–551. <https://doi.org/10.1164/rccm.201207-1217OC>
- Drain, P.K., Bajema, K.L., Dowdy, D., Dheda, K., Naidoo, K., Schumacher, S.G., Ma, S., Meermeier, E., Lewinsohn, D.M., Sherman, D.R., 2018. Incipient and Subclinical Tuberculosis: a Clinical Review of Early Stages and Progression of Infection. *Clin. Microbiol. Rev.* 31. <https://doi.org/10.1128/CMR.00021-18>
- Dymond, J.S., 2013. Chapter Twenty Three - Explanatory Chapter: Quantitative PCR, in: Lorsch, J. (Ed.), *Methods in Enzymology, Laboratory Methods in Enzymology: DNA*. Academic Press, pp. 279–289. <https://doi.org/10.1016/B978-0-12-418687-3.00023-9>
- Dzodanu, E.G., Afrifa, J., Acheampong, D.O., Dadzie, I., 2019. Diagnostic Yield of Fluorescence and Ziehl-Neelsen Staining Techniques in the Diagnosis of Pulmonary Tuberculosis: A Comparative Study in a District Health Facility. *Tuberc. Res. Treat.* 2019, 4091937. <https://doi.org/10.1155/2019/4091937>
- Ellis, J., Harnett, J., Cameron, G., Moss, P., Gray, A., 2024. Performance evaluation of the LumiraDx quantitative microfluidic point-of-care CRP test. *Pract. Lab. Med.* 38, e00349. <https://doi.org/10.1016/j.plabm.2023.e00349>
- Ellis, J., MacLuskie, S., Craig, D., Lehane, L., McInnes, G., Harnett, J., Cameron, G., Moss, P., Gray, A., 2022. Evaluation of the performance of a quantitative point-of-care CRP test (preprint). *Respiratory Medicine*. <https://doi.org/10.1101/2022.05.20.22275259>
- Esmail, H., Lai, R.P., Lesosky, M., Wilkinson, K.A., Graham, C.M., Horswell, S., Coussens, A.K., Barry, C.E., O'Garra, A., Wilkinson, R.J., 2018. Complement pathway gene activation and rising circulating immune complexes characterize early disease in HIV-associated tuberculosis. *Proc. Natl. Acad. Sci.* 115, E964–E973. <https://doi.org/10.1073/pnas.1711853115>
- Esmail, H., Macpherson, L., Coussens, A.K., Houben, R.M.G.J., 2022. Mind the gap – Managing tuberculosis across the disease spectrum. *eBioMedicine* 78. <https://doi.org/10.1016/j.ebiom.2022.103928>
- European Centre for Disease Prevention and Control, 2011. Use of interferon-gamma release assays in support of TB diagnosis : ad hoc scientific panel opinion. Publications Office, LU.
- Evans, C.A., 2011. GeneXpert—A Game-Changer for Tuberculosis Control? *PLOS Med.* 8, e1001064. <https://doi.org/10.1371/journal.pmed.1001064>
- Francisco, N.M., Fang, Y.-M., Ding, L., Feng, S., Yang, Y., Wu, M., Jacobs, M., Ryffel, B., Huang, X., 2017. Diagnostic accuracy of a selected signature gene set that discriminates active

- pulmonary tuberculosis and other pulmonary diseases. *J. Infect.* 75, 499–510. <https://doi.org/10.1016/j.jinf.2017.09.012>
- Friedrich, S.O., Rachow, A., Saathoff, E., Singh, K., Mangu, C.D., Dawson, R., Phillips, P.P., Venter, A., Bateson, A., Boehme, C.C., Heinrich, N., Hunt, R.D., Boeree, M.J., Zumla, A., McHugh, T.D., Gillespie, S.H., Diacon, A.H., Hoelscher, M., 2013. Assessment of the sensitivity and specificity of Xpert MTB/RIF assay as an early sputum biomarker of response to tuberculosis treatment. *Lancet Respir. Med.* 1, 462–470. [https://doi.org/10.1016/S2213-2600\(13\)70119-X](https://doi.org/10.1016/S2213-2600(13)70119-X)
- Gaikwad, U.N., Gaikwad, N.R., 2018. Modalities to monitor the treatment response in tuberculosis. *Indian J. Tuberc.* 65, 109–117. <https://doi.org/10.1016/j.ijtb.2017.12.014>
- Gammino, V.M., Taylor, A.B., Rich, M.L., Bayona, J., Becerra, M.C., Bonilla, C., Gelmanova, I., Hollo, V., Jaramillo, E., Keshavjee, S., Leimane, V., Mitnick, C.D., Quelapio, M.I.D., Rieksina, V., Tupasi, T.E., Wells, C.D., Zignol, M., Cegielski, P.J., 2011. Bacteriologic monitoring of multidrug-resistant tuberculosis patients in five DOTS-Plus pilot projects. *Int. J. Tuberc. Lung Dis. Off. J. Int. Union Tuberc. Lung Dis.* 15, 1315–1322. <https://doi.org/10.5588/ijtld.10.0221>
- García-Basteiro, A.L., DiNardo, A., Saavedra, B., Silva, D.R., Palmero, D., Gegia, M., Migliori, G.B., Duarte, R., Mambuque, E., Centis, R., Cuevas, L.E., Izco, S., Theron, G., 2018. Point of care diagnostics for tuberculosis. *Pulmonology, World TB day 2018: targeting tuberculosis* 24, 73–85. <https://doi.org/10.1016/j.rppnen.2017.12.002>
- Geluk, A., van Meijgaarden, K.E., Franken, K.L.M.C., Subronto, Y.W., Wieles, B., Arend, S.M., Sampaio, E.P., de Boer, T., Faber, W.R., Naafs, B., Ottenhoff, T.H.M., 2002. Identification and Characterization of the ESAT-6 Homologue of *Mycobacterium leprae* and T-Cell Cross-Reactivity with *Mycobacterium tuberculosis*. *Infect. Immun.* 70, 2544–2548. <https://doi.org/10.1128/iai.70.5.2544-2548.2002>
- Gey van Pittius, N.C., Warren, R.M., van Helden, P.D., 2002. ESAT-6 and CFP-10: What Is the Diagnosis? *Infect. Immun.* 70, 6509–6511. <https://doi.org/10.1128/iai.70.11.6509-6511.2002>
- Gill, C.M., Dolan, L., Piggott, L.M., McLaughlin, A.M., 2022. New developments in tuberculosis diagnosis and treatment. *Breathe* 18, 210149. <https://doi.org/10.1183/20734735.0149-2021>
- Gillespie, S.H., Crook, A.M., McHugh, T.D., Mendel, C.M., Meredith, S.K., Murray, S.R., Pappas, F., Phillips, P.P.J., Nunn, A.J., 2014. Four-Month Moxifloxacin-Based Regimens for Drug-Sensitive Tuberculosis. *N. Engl. J. Med.* 371, 1577–1587. <https://doi.org/10.1056/NEJMoa1407426>
- Gliddon, H.D., Howes, P.D., Kaforou, M., Levin, M., Stevens, M.M., 2016. A nucleic acid strand displacement system for the multiplexed detection of tuberculosis-specific mRNA using quantum dots. *Nanoscale* 8, 10087–10095. <https://doi.org/10.1039/C6NR00484A>
- Gliddon, H.D., Kaforou, M., Alikian, M., Habgood-Coote, D., Zhou, C., Oni, T., Anderson, S.T., Brent, A.J., Crampin, A.C., Eley, B., Heyderman, R., Kern, F., Langford, P.R., Ottenhoff, T.H.M., Hibberd, M.L., French, N., Wright, V.J., Dockrell, H.M., Coin, L.J., Wilkinson, R.J., Levin, M., 2021. Identification of Reduced Host Transcriptomic Signatures for Tuberculosis Disease and Digital PCR-Based Validation and Quantification. *Front. Immunol.* 12.
- Günther, G., Heyckendorf, J., Zellweger, J.P., Reimann, M., Claassens, M., Chesov, D., van Leth, F., 2021. Defining Outcomes of Tuberculosis (Treatment): From the Past to the Future. *Respiration* 100, 843–852. <https://doi.org/10.1159/000516392>
- Gupta, R.K., Turner, C.T., Venturini, C., Esmail, H., Rangaka, M.X., Copas, A., Lipman, M., Abubakar, I., Noursadeghi, M., 2020. Concise whole blood transcriptional signatures for incipient tuberculosis: a systematic review and patient-level pooled meta-analysis. *Lancet Respir. Med.* 0. [https://doi.org/10.1016/S2213-2600\(19\)30282-6](https://doi.org/10.1016/S2213-2600(19)30282-6)
- Gupta-Wright, A., den Boon, S., MacLean, E.L., Cirillo, D., Cobelens, F., Gillespie, S.H., Kohli, M., Ruhwald, M., Savic, R., Brigden, G., Gidado, M., Goletti, D., Hanna, D., Hasan, R., Hewison, C., Koura, K.G., Lienhardt, C., Lungu, P., McHugh, T.D., McKenna, L., Scott, C., Scriba, T., Sekaggya-Wiltshire, C., Kasaeva, T., Zignol, M., Denking, C.M., Falzon, D., 2023. Target

- product profiles: tests for tuberculosis treatment monitoring and optimization. *Bull. World Health Organ.* 101, 730–737. <https://doi.org/10.2471/BLT.23.290901>
- Hajra Tasneem Masali, Aparna Takpere and Praveen Shahapur, 2021. A Comparative Study of Ziehl-Neelsen Stain and Fluorescent Stain Microscopy in the Diagnosis of Pulmonary Tuberculosis. *J. Pure Appl. Microbiol.* URL <https://microbiologyjournal.org/a-comparative-study-of-ziehl-neelsen-stain-and-fluorescent-stain-microscopy-in-the-diagnosis-of-pulmonary-tuberculosis/> (accessed 5.8.24).
- Hamada, Y., Gupta, R.K., Quartagno, M., Izzard, A., Acuna-Villaorduna, C., Altet, N., Diel, R., Dominguez, J., Floyd, S., Gupta, A., Huerga, H., Jones-López, E.C., Kinikar, A., Lange, C., Van Leth, F., Liu, Q., Lu, W., Lu, P., Rueda, I.L., Martinez, L., Mbandi, S.K., Muñoz, L., Padilla, E.S., Paradkar, M., Scriba, T., Sester, M., Shanaube, K., Sharma, S.K., Sloot, R., Sotgiu, G., Thiruvengadam, K., Vashishtha, R., Abubakar, I., Rangaka, M.X., 2023. Predictive performance of interferon-gamma release assays and the tuberculin skin test for incident tuberculosis: an individual participant data meta-analysis. *eClinicalMedicine* 56, 101815. <https://doi.org/10.1016/j.eclinm.2022.101815>
- Hanifa, Y., Fielding, K.L., Chihota, V.N., Adonis, L., Charalambous, S., Karstaedt, A., McCarthy, K., Nicol, M.P., Ndlovu, N.T., Sahid, F., Churchyard, G.J., Grant, A.D., 2016. Diagnostic Accuracy of Lateral Flow Urine LAM Assay for TB Screening of Adults with Advanced Immunosuppression Attending Routine HIV Care in South Africa. *PLOS ONE* 11, e0156866. <https://doi.org/10.1371/journal.pone.0156866>
- Helb, D., Jones, M., Story, E., Boehme, C., Wallace, E., Ho, K., Kop, J., Owens, M.R., Rodgers, R., Banada, P., Safi, H., Blakemore, R., Lan, N.T.N., Jones-López, E.C., Levi, M., Burday, M., Ayakaka, I., Mugerwa, R.D., McMillan, B., Winn-Deen, E., Christel, L., Dailey, P., Perkins, M.D., Persing, D.H., Alland, D., 2010. Rapid Detection of *Mycobacterium tuberculosis* and Rifampin Resistance by Use of On-Demand, Near-Patient Technology. *J. Clin. Microbiol.* 48, 229–237. <https://doi.org/10.1128/jcm.01463-09>
- Heyckendorf, J., Marwitz, S., Reimann, M., Avsar, K., DiNardo, A.R., Günther, G., Hoelscher, M., Ibraim, E., Kalsdorf, B., Kaufmann, S.H.E., Kontsevaya, I., van Leth, F., Mandalakas, A.M., Maurer, F.P., Müller, M., Nitschkowski, D., Olaru, I.D., Popa, C., Rachow, A., Rolling, T., Rybníček, J., Salzer, H.J.F., Sanchez-Carballo, P., Schuhmann, M., Schaub, D., Spinu, V., Suárez, I., Terhalle, E., Unnewehr, M., Weiner, J., Goldmann, T., Lange, C., 2021. Prediction of anti-tuberculosis treatment duration based on a 22-gene transcriptomic model. *Eur. Respir. J.* 58, 2003492. <https://doi.org/10.1183/13993003.03492-2020>
- Hiza, H., Hella, J., Arbués, A., Magani, B., Sasamalo, M., Gagneux, S., Reither, K., Portevin, D., 2021. Case–control diagnostic accuracy study of a non-sputum CD38-based TAM-TB test from a single milliliter of blood. *Sci. Rep.* 11, 13190. <https://doi.org/10.1038/s41598-021-92596-z>
- Ho, J., Bokil, N.J., Nguyen, P.T.B., Nguyen, T.A., Liu, M.Y., Hare, N., Fox, G.J., Saunders, B.M., Marks, G.B., Britton, W.J., 2020. A transcriptional blood signature distinguishes early tuberculosis disease from latent tuberculosis infection and uninfected individuals in a Vietnamese cohort. *J. Infect.* 81, 72–80. <https://doi.org/10.1016/j.jinf.2020.03.066>
- Hoang, L.T., Jain, P., Pillay, T.D., Tolosa-Wright, M., Niazi, U., Takwoingi, Y., Halliday, A., Berrocal-Almanza, L.C., Deeks, J.J., Beverley, P., Kon, O.M., Lalvani, A., 2021. Transcriptomic signatures for diagnosing tuberculosis in clinical practice: a prospective, multicentre cohort study. *Lancet Infect. Dis.* 21, 366–375. [https://doi.org/10.1016/S1473-3099\(20\)30928-2](https://doi.org/10.1016/S1473-3099(20)30928-2)
- Hong, J.M., Lee, H., Menon, N.V., Lim, C.T., Lee, L.P., Ong, C.W.M., 2022. Point-of-care diagnostic tests for tuberculosis disease. *Sci. Transl. Med.* 14, eabj4124. <https://doi.org/10.1126/scitranslmed.abj4124>
- Horne, D.J., Spitters, C., Narita, M., 2011. Experience with Rifabutin Replacing Rifampin in the Treatment of Tuberculosis. *Int. J. Tuberc. Lung Dis.* 15, 1485–i. <https://doi.org/10.5588/ijtld.11.0068>

- Huang, H.-L., Lee, J.-Y., Lo, Y.-S., Liu, I.-H., Huang, S.-H., Huang, Y.-W., Lee, M.-R., Lee, C.-H., Cheng, M.-H., Lu, P.-L., Wang, J.-Y., Yang, J.-M., Chong, I.-W., 2022. Whole-Blood 3-Gene Signature as a Decision Aid for Rifapentine-based Tuberculosis Preventive Therapy. *Clin. Infect. Dis.* ciac003. <https://doi.org/10.1093/cid/ciac003>
- Hur, Y.-G., Crampin, A.C., Chisambo, C., Kanyika, J., Houben, R., Ndhlovu, R., Mzembe, T., Lalor, M.K., Saul, J., Branson, K., Stanley, C., Ngwira, B., French, N., Ottenhoff, T.H., Dockrell, H.M., Gorak-Stolinska, P., 2014. Identification of Immunological Biomarkers Which May Differentiate Latent Tuberculosis from Exposure to Environmental Nontuberculous Mycobacteria in Children. *Clin. Vaccine Immunol.* 21, 133–142. <https://doi.org/10.1128/CVI.00620-13>
- Imperial, M.Z., Nahid, P., Phillips, P.P.J., Davies, G.R., Fielding, K., Hanna, D., Hermann, D., Wallis, R.S., Johnson, J.L., Lienhardt, C., Savic, R.M., 2018. A patient-level pooled analysis of treatment-shortening regimens for drug-susceptible pulmonary tuberculosis. *Nat. Med.* 24, 1708–1715. <https://doi.org/10.1038/s41591-018-0224-2>
- Jacobs, R., Malherbe, S., Loxton, A.G., Stanley, K., Spuy, G. van der, Walzl, G., Chegou, N.N., 2016. Identification of novel host biomarkers in plasma as candidates for the immunodiagnosis of tuberculosis disease and monitoring of tuberculosis treatment response. *Oncotarget* 7, 57581–57592. <https://doi.org/10.18632/oncotarget.11420>
- Jindani, A., Harrison, T.S., Nunn, A.J., Phillips, P.P.J., Churchyard, G.J., Charalambous, S., Hatherill, M., Geldenhuys, H., McIlleron, H.M., Zvada, S.P., Mungofa, S., Shah, N.A., Zizhou, S., Magweta, L., Shepherd, J., Nyirenda, S., van Dijk, J.H., Clouting, H.E., Coleman, D., Bateson, A.L.E., McHugh, T.D., Butcher, P.D., Mitchison, D.A., 2014. High-Dose Rifapentine with Moxifloxacin for Pulmonary Tuberculosis. *N. Engl. J. Med.* 371, 1599–1608. <https://doi.org/10.1056/NEJMoa1314210>
- Kaforou, M., Broderick, C., Vito, O., Levin, M., Scriba, T.J., Seddon, J.A., 2022. Transcriptomics for child and adolescent tuberculosis*. *Immunol. Rev.* 309, 97–122. <https://doi.org/10.1111/imr.13116>
- Kaforou, M., Wright, V.J., Oni, T., French, N., Anderson, S.T., Bangani, N., Banwell, C.M., Brent, A.J., Crampin, A.C., Dockrell, H.M., Eley, B., Heyderman, R.S., Hibberd, M.L., Kern, F., Langford, P.R., Ling, L., Mendelson, M., Ottenhoff, T.H., Zgambo, F., Wilkinson, R.J., Coin, L.J., Levin, M., 2013. Detection of Tuberculosis in HIV-Infected and -Uninfected African Adults Using Whole Blood RNA Expression Signatures: A Case-Control Study. *PLOS Med.* 10, e1001538. <https://doi.org/10.1371/journal.pmed.1001538>
- Kaipilyawar, V., Zhao, Y., Wang, X., Joseph, N.M., Knudsen, S., Prakash Babu, S., Muthaiah, M., Hochberg, N.S., Sarkar, S., Horsburgh, C.R., Ellner, J.J., Johnson, W.E., Salgame, P., 2022. Development and Validation of a Parsimonious Tuberculosis Gene Signature Using the digital NanoString nCounter Platform. *Clin. Infect. Dis.* 75, 1022–1030. <https://doi.org/10.1093/cid/ciac010>
- Keeler, E., Perkins, M.D., Small, P., Hanson, C., Reed, S., Cunningham, J., Aledort, J.E., Hillborne, L., Rafael, M.E., Giroi, F., Dye, C., 2006. Reducing the global burden of tuberculosis: the contribution of improved diagnostics. *Nature* 444, 49–57. <https://doi.org/10.1038/nature05446>
- Ketema, W., Woubishet, K., Tesfaye, S., Gutema, S., Taye, K., Shibeshi, M.S., Tadesse, N., 2022. A Breakthrough in the Challenges of Tuberculosis Diagnosis: Lateral Flow Urine Lipoarabinomannan (LAM) Assay for the Diagnosis of Active Tuberculosis in a Subset of Human Immuno Deficiency Virus (HIV) Patients at Hawassa University Comprehensive Specialized Hospital, Hawassa, Ethiopia. *Int. Med. Case Rep. J.* 15, 393–397. <https://doi.org/10.2147/IMCRJ.S373197>
- Khan, I.H., Ravindran, R., Krishnan, V.V., Awan, I.N., Rizvi, S.K., Saqib, M.A., Shahzad, M.I., Tahseen, S., Ireton, G., Goulding, C.W., Felgner, P., DeRiemer, K., Khanum, A., Luciw, P.A., 2011.

- Plasma Antibody Profiles as Diagnostic Biomarkers for Tuberculosis. *Clin. Vaccine Immunol.* 18, 2148–2153. <https://doi.org/10.1128/CVI.05304-11>
- Kik, S.V., Denking, C.M., Chedore, P., Pai, M., 2014. Replacing smear microscopy for the diagnosis of tuberculosis: what is the market potential? *Eur. Respir. J.* 43, 1793–1796. <https://doi.org/10.1183/09031936.00217313>
- Kono, N., Arakawa, K., 2019. Nanopore sequencing: Review of potential applications in functional genomics. *Dev. Growth Differ.* 61, 316–326. <https://doi.org/10.1111/dgd.12608>
- Korbel, D.S., Schneider, B.E., Schaible, U.E., 2008. Innate immunity in tuberculosis: myths and truth. *Microbes Infect.*, 10th anniversary of *Microbes and Infection* 10, 995–1004. <https://doi.org/10.1016/j.micinf.2008.07.039>
- Kulkarni, V., Queiroz, A.T.L., Sangle, S., Kagal, A., Salvi, S., Gupta, A., Ellner, J., Kadam, D., Rolla, V.C., Andrade, B.B., Salgame, P., Mave, V., 2021. A Two-Gene Signature for Tuberculosis Diagnosis in Persons With Advanced HIV. *Front. Immunol.* 12.
- Kunnath-Velayudhan, S., Davidow, A.L., Wang, H.-Y., Molina, D.M., Huynh, V.T., Salamon, H., Pine, R., Michel, G., Perkins, M.D., Xiaowu, L., Felgner, P.L., Flynn, J.L., Catanzaro, A., Gennaro, M.L., 2012. Proteome-Scale Antibody Responses and Outcome of *Mycobacterium tuberculosis* Infection in Nonhuman Primates and in Tuberculosis Patients. *J. Infect. Dis.* 206, 697–705. <https://doi.org/10.1093/infdis/jis421>
- Lamanna, J., Scott, E.Y., Edwards, H.S., Chamberlain, M.D., Dryden, M.D.M., Peng, J., Mair, B., Lee, A., Chan, C., Sklavounos, A.A., Heffernan, A., Abbas, F., Lam, C., Olson, M.E., Moffat, J., Wheeler, A.R., 2020. Digital microfluidic isolation of single cells for -Omics. *Nat. Commun.* 11, 5632. <https://doi.org/10.1038/s41467-020-19394-5>
- Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV, 2019 Update [WWW Document], n.d. URL <https://www.who.int/publications-detail-redirect/9789241550604> (accessed 5.9.23).
- Lawn, S.D., Mwaba, P., Bates, M., Piatek, A., Alexander, H., Marais, B.J., Cuevas, L.E., McHugh, T.D., Zijenah, L., Kapata, N., Abubakar, I., McNerney, R., Hoelscher, M., Memish, Z.A., Migliori, G.B., Kim, P., Maeurer, M., Schito, M., Zumla, A., 2013. Advances in tuberculosis diagnostics: the Xpert MTB/RIF assay and future prospects for a point-of-care test. *Lancet Infect. Dis.* 13, 349–361. [https://doi.org/10.1016/S1473-3099\(13\)70008-2](https://doi.org/10.1016/S1473-3099(13)70008-2)
- Lazarevic, V., Glimcher, L.H., 2011. T-bet in disease. *Nat. Immunol.* 12, 597–606. <https://doi.org/10.1038/ni.2059>
- Lee, S.-W., Wu, L.S.-H., Huang, G.-M., Huang, K.-Y., Lee, T.-Y., Weng, J.T.-Y., 2016. Gene expression profiling identifies candidate biomarkers for active and latent tuberculosis. *BMC Bioinformatics* 17, S3. <https://doi.org/10.1186/s12859-015-0848-x>
- Leong, S., Zhao, Y., Joseph, N.M., Hochberg, N.S., Sarkar, S., Pleskunas, J., Hom, D., Lakshminarayanan, S., Horsburgh, C.R., Roy, G., Ellner, J.J., Johnson, W.E., Salgame, P., 2018. Existing blood transcriptional classifiers accurately discriminate active tuberculosis from latent infection in individuals from south India. *Tuberculosis* 109, 41–51. <https://doi.org/10.1016/j.tube.2018.01.002>
- Lesho, E., Forestiero, F.J., Hirata, M.H., Hirata, R.D., Cecon, L., Melo, F.F., Paik, S.H., Murata, Y., Ferguson, E.W., Wang, Z., Ooi, G.T., 2011. Transcriptional responses of host peripheral blood cells to tuberculosis infection. *Tuberculosis, Special Issue: TB in India* 91, 390–399. <https://doi.org/10.1016/j.tube.2011.07.002>
- Lessells, R., 2019. Rapid Point-of-Care Diagnostic Tests for Tuberculosis, in: *Revolutionizing Tropical Medicine*. John Wiley & Sons, Ltd, pp. 105–125. <https://doi.org/10.1002/9781119282686.ch5>
- Li, L., Mahan, C.S., Palaci, M., Horter, L., Loeffelholz, L., Johnson, J.L., Dietze, R., Debanne, S.M., Joloba, M.L., Okwera, A., Boom, W.H., Eisenach, K.D., 2010. Sputum *Mycobacterium tuberculosis* mRNA as a Marker of Bacteriologic Clearance in Response to Antituberculosis Therapy. *J. Clin. Microbiol.* 48, 46–51. <https://doi.org/10.1128/jcm.01526-09>

- Lin, P.L., Flynn, J.L., 2018. The End of the Binary Era: Revisiting the Spectrum of Tuberculosis. *J. Immunol.* 201, 2541–2548. <https://doi.org/10.4049/jimmunol.1800993>
- Liu, C., Lyon, C.J., Bu, Y., Deng, Z., Walters, E., Li, Y., Zhang, L., Hesselning, A.C., Graviss, E.A., Hu, Y., 2018. Clinical Evaluation of a Blood Assay to Diagnose Paucibacillary Tuberculosis via Bacterial Antigens. *Clin. Chem.* 64, 791–800. <https://doi.org/10.1373/clinchem.2017.273698>
- Long, N.P., Phat, N.K., Yen, N.T.H., Park, S., Park, Y., Cho, Y.-S., Shin, J.-G., 2021. A 10-gene biosignature of tuberculosis treatment monitoring and treatment outcome prediction. *Tuberculosis* 131, 102138. <https://doi.org/10.1016/j.tube.2021.102138>
- Lowe, R., Shirley, N., Bleackley, M., Dolan, S., Shafee, T., 2017. Transcriptomics technologies. *PLOS Comput. Biol.* 13, e1005457. <https://doi.org/10.1371/journal.pcbi.1005457>
- Lu, L.L., Smith, M.T., Yu, K.K.Q., Luedemann, C., Suscovich, T.J., Grace, P.S., Cain, A., Yu, W.H., McKittrick, T.R., Lauffenburger, D., Cummings, R.D., Mayanja-Kizza, H., Hawn, T.R., Boom, W.H., Stein, C.M., Fortune, S.M., Seshadri, C., Alter, G., 2019. IFN- γ -independent immune markers of Mycobacterium tuberculosis exposure. *Nat. Med.* 25, 977–987. <https://doi.org/10.1038/s41591-019-0441-3>
- Lyashchenko, K., Colangeli, R., Houde, M., Al Jahdali, H., Menzies, D., Gennaro, M.L., 1998. Heterogeneous Antibody Responses in Tuberculosis. *Infect. Immun.* 66, 3936–3940. <https://doi.org/10.1128/iai.66.8.3936-3940.1998>
- Lyu, M., Zhou, J., Zhou, Y., Chong, W., Xu, W., Lai, H., Niu, L., Hai, Y., Yao, X., Gong, S., Wang, Q., Chen, Y., Wang, Y., Chen, L., Zengwanggema, Zeng, J., Wang, C., Ying, B., 2023. From tuberculosis bedside to bench: UBE2B splicing as a potential biomarker and its regulatory mechanism. *Signal Transduct. Target. Ther.* 8, 1–14. <https://doi.org/10.1038/s41392-023-01346-2>
- Machingaidze, S., Verver, S., Mulenga, H., Abrahams, D.-A., Hatherill, M., Hanekom, W., Hussey, G.D., Mahomed, H., 2012. Predictive Value of Recent QuantiFERON Conversion for Tuberculosis Disease in Adolescents. *Am. J. Respir. Crit. Care Med.* 186, 1051–1056. <https://doi.org/10.1164/rccm.201206-1134OC>
- MacLean, E., Broger, T., Yerlikaya, S., Fernandez-Carballo, B.L., Pai, M., Denking, C.M., 2019a. A systematic review of biomarkers to detect active tuberculosis. *Nat. Microbiol.* 4, 748–758. <https://doi.org/10.1038/s41564-019-0380-2>
- MacLean, E., Broger, T., Yerlikaya, S., Fernandez-Carballo, B.L., Pai, M., Denking, C.M., 2019b. A systematic review of biomarkers to detect active tuberculosis. *Nat. Microbiol.* 4, 748–758. <https://doi.org/10.1038/s41564-019-0380-2>
- MacLean, E., Nathavitharana, R.R., 2021. Progress toward Developing Sensitive Non-Sputum-Based Tuberculosis Diagnostic Tests: the Promise of Urine Cell-Free DNA. *J. Clin. Microbiol.* 59, e0070621. <https://doi.org/10.1128/JCM.00706-21>
- MacPherson, P., Webb, E.L., Kamchedzera, W., Joekes, E., Mjoli, G., Lalloo, D.G., Divala, T.H., Choko, A.T., Burke, R.M., Maheswaran, H., Pai, M., Squire, S.B., Nliwasa, M., Corbett, E.L., 2021. Computer-aided X-ray screening for tuberculosis and HIV testing among adults with cough in Malawi (the PROSPECT study): A randomised trial and cost-effectiveness analysis. *PLOS Med.* 18, e1003752. <https://doi.org/10.1371/journal.pmed.1003752>
- Maertzdorf, J., Weiner, J., Mollenkopf, H.-J., Network, Tb., Bauer, T., Prasse, A., Müller-Quernheim, J., Kaufmann, S.H.E., Bellinger, O., Diel, R., Ehlers, S., Heykes-Uden, H., Hölscher, C., Kaufmann, S.H.E., Lange, C., Meyer, C., Niemann, S., Nöthlings, U., Reiling, N., Schaberg, T., Stenger, S., 2012. Common patterns and disease-related signatures in tuberculosis and sarcoidosis. *Proc. Natl. Acad. Sci. U. S. A.* 109, 7853–7858. <https://doi.org/10.1073/pnas.1121072109>
- Marx, V., 2021. Method of the Year: spatially resolved transcriptomics. *Nat. Methods* 18, 9–14. <https://doi.org/10.1038/s41592-020-01033-y>
- Mayer-Barber, K.D., Sher, A., 2015. Cytokine and lipid mediator networks in tuberculosis. *Immunol. Rev.* 264, 264–275. <https://doi.org/10.1111/imr.12249>

- McNab, F., Mayer-Barber, K., Sher, A., Wack, A., O'Garra, A., 2015. Type I interferons in infectious disease. *Nat. Rev. Immunol.* 15, 87–103. <https://doi.org/10.1038/nri3787>
- Meier, N.R., Volken, T., Geiger, M., Heininger, U., Tebruegge, M., Ritz, N., 2019. Risk Factors for Indeterminate Interferon-Gamma Release Assay for the Diagnosis of Tuberculosis in Children—A Systematic Review and Meta-Analysis. *Front. Pediatr.* 7. <https://doi.org/10.3389/fped.2019.00208>
- Mendelsohn, S.C., Andrade, B.B., Mbandi, S.K., Andrade, A.M.S., Muwanga, V.M., Figueiredo, M.C., Erasmus, M., Rolla, V.C., Thami, P.K., Cordeiro-Santos, M., Penn-Nicholson, A., Kritski, A.L., Hatherill, M., Sterling, T.R., Scriba, T.J., the RePORT-South Africa and RePORT-Brazil Consortia, 2024. Transcriptomic signatures of progression to TB disease among close contacts in Brazil. *J. Infect. Dis.* jiae237. <https://doi.org/10.1093/infdis/jiae237>
- Mendelsohn, S.C., Fiore-Gartland, A., Penn-Nicholson, A., Mulenga, H., Mbandi, S.K., Borate, B., Hadley, K., Hikuam, C., Musvosvi, M., Bilek, N., Erasmus, M., Jaxa, L., Raphela, R., Nombida, O., Kaskar, M., Sumner, T., White, R.G., Innes, C., Brumskine, W., Hiemstra, A., Malherbe, S.T., Hassan-Moosa, R., Tameris, M., Walzl, G., Naidoo, K., Churchyard, G., Scriba, T.J., Hatherill, M., Abrahams, C., Africa, H., Ahlers, P., Arendsen, D., Badimo, T., Baepanye, K., Baepanye, K.E., Bande, B., Batyi, N.C., Beukes, R., Bontsi, L.T., Boo, O.P., Botha, M.C., Braaf, S., Buhlungu, S., Carstens, A., Chauke, K.V., Chinappa, T., Chung, E., Chung, M., Clarke, K., Cloete, Y., Coetzee, L., Collignon, M., Companie, A., Corris, C., Cwaile, M.T., Cwele, T., Davids, I., Davies, I.J., Klerk, E.D., Kock, M. de, Dhlamini, A.L., Diamond, B., Didloff, M., Dlamini, C., Dolo, P., Eyre, C., Feni, T., Ferreira, J., Ferus, C., Fisher, M., Flinn, M., Fransman, B., Galane, W.P., Geldenhuys, H., Gempies, D., Goliath, T., Govender, D., Gregg, Y., Gumede, G., Gwamada, Z., Halti, S., Hassiem, R., Herling, R., Herselman, Y., Hughes, E., Issel, H., Iyemosolo, B.M., Jali, Z., Rensburg, B.J.V., Jansen, R., Jeleni, J.M., Jonkane, O., Julies, F., Kafaar, F., Kasongo, C.M., Keffers, S., Kekana, B.S., Kekana, S.J., Kelepu, X., Khanyile, L., Khobedi, G.V., Khomba, G., Khoza, L.S., King, M., Kolobe, G.K., Kruger, S., Kruger, J., Kunene, N.I., Lakay, S., Lakhi, A., Langa, N., Ledwaba, H., Lekagane, L.J., Lekotloane, S.C., Leopeng, T., Louw, I.J., Luabeya, A.K.K., Lusale, S.T., Maatjie, P.T., Mabasa, I., Mabe, T.D., Mabena, K.F., Mabuza, N.C., Mabwe, S., Madikwe, J.T., Madikwe, O.M., Maebana, R.L., Magwasha, M.S., Majola, M., Makhetha, M., Makhetha, L., Malay, V., Manzini, V.-I.-V.F., Maphanga, J., Maphanga, N., Market, J., Maroele, I.S., Masibi, O.P., Mathabanzini, J.R., Mathode, T.I., Matsane, E.D., Mbata, L., Meyer, F., Mhandire, N.K., Miga, T., Mkhize, N.C.T., Mkhokho, C., Mkwilase, N.H., Mnqonywa, N., Moche, K., Modisaotsile, B.M., Mokgetsengoane, P.P., Mokone, S.M.C., Molatlhegi, K.M., Molefe, T.A., Moloko, J.P., Molosi, K., Molotsi, M.E., Montwedi, T.E., Monyemangene, B.D., Mooketsi, H.M., Moses, M., Mosito, B., Mosito, T.M., Mosweu, I.L., Mothaga, P., Motlagomang, B.O., Mouton, A., Mpofu, O., Mthembu, F.N., Mtlali, M., Ndlovu, N., Ngcobo, N., Noble, J., Ntamo, B.B., Ntanjana, G., Ntshauba, T., Opperman, F., Padayatchi, N., Papalagae, T., Petersen, C., Phakathi, T., Phatshwane, M.O., Plaatjie, P., Pretorius, A., Rameetse, V.K., Ramjit, D., Ratangee, F., Ratangee, M., Sanyaka, P.N., Sato, A., Schoeman, E., Schreuder, C., Seabela, L., Segatsho, K.M., Sein, N.N., Selepe, R.A.P., Senne, M.N., September, A., September, C., Serake, M., Shenje, J., Shezi, T., Shezi, S.C., Sing, P., Singh, C., Sithetho, Z., Solomons, D., Stanley, K., Steyn, M., Stofile, B., Stryers, S., Swanepoel, L., Swarts, A., Thaba, M.M., Theko, L.C., Thembela, P., Thompo, M., Toefy, A., Toto, K., Tsagae, D.S., Tsamane, A., Tshikovi, V., Tswaile, L.I., Tyambetyu, P., Tönsing, S., Valley, H., Merwe, L. van der, Rooyen, E. van, Veldsman, A., Veldtsman, H., Vollenhoven, K., Zaca, L., Zimri, E., Zulu, M., 2021. Validation of a host blood transcriptomic biomarker for pulmonary tuberculosis in people living with HIV: a prospective diagnostic and prognostic accuracy study. *Lancet Glob. Health* 9, e841–e853. [https://doi.org/10.1016/S2214-109X\(21\)00045-0](https://doi.org/10.1016/S2214-109X(21)00045-0)
- Mendelsohn, S.C., Mbandi, S.K., Fiore-Gartland, A., Penn-Nicholson, A., Musvosvi, M., Mulenga, H., Fisher, M., Hadley, K., Erasmus, M., Nombida, O., Tameris, M., Walzl, G., Naidoo, K.,

- Churchyard, G., Hatherill, M., Scriba, T.J., 2022. Prospective multicentre head-to-head validation of host blood transcriptomic biomarkers for pulmonary tuberculosis by real-time PCR. *Commun. Med.* 2, 26. <https://doi.org/10.1038/s43856-022-00086-8>
- Menzies, D., Pai, M., Comstock, G., 2007. Meta-analysis: New Tests for the Diagnosis of Latent Tuberculosis Infection: Areas of Uncertainty and Recommendations for Research. *Ann. Intern. Med.* 146, 340–354. <https://doi.org/10.7326/0003-4819-146-5-200703060-00006>
- Merle, C.S., Fielding, K., Sow, O.B., Gninafon, M., Lo, M.B., Mthiyane, T., Odhiambo, J., Amukoye, E., Bah, B., Kassa, F., N'Diaye, A., Rustumjee, R., de Jong, B.C., Horton, J., Perronne, C., Sismanidis, C., Lapujade, O., Olliaro, P.L., Lienhardt, C., 2014. A Four-Month Gatifloxacin-Containing Regimen for Treating Tuberculosis. *N. Engl. J. Med.* 371, 1588–1598. <https://doi.org/10.1056/NEJMoa1315817>
- Migliori, G.B., Ong, C.W.M., Petrone, L., D'Ambrosio, L., Centis, R., Goletti, D., 2021. The definition of tuberculosis infection based on the spectrum of tuberculosis disease. *Breathe* 17, 210079. <https://doi.org/10.1183/20734735.0079-2021>
- Mistry, R., Cliff, J.M., Clayton, C.L., Beyers, N., Mohamed, Y.S., Wilson, P.A., Dockrell, H.M., Wallace, D.M., Helden, P.D. van, Duncan, K., Lukey, P.T., 2007. Gene-Expression Patterns in Whole Blood Identify Subjects at Risk for Recurrent Tuberculosis. *J. Infect. Dis.* 195, 357–365. <https://doi.org/10.1086/510397>
- Model to evaluate the impact of a test for incipient TB | Stop TB Partnership [WWW Document], n.d. URL <https://www.stoptb.org/task-force-tbi-and-test-of-progression/model-to-evaluate-impact-of-test-incipient-tb> (accessed 5.29.23).
- Montenegro, R.A., Guarines, K.M., Montenegro, L.M.L., Lira, L.A.S., Falcão, J., Melo, F.L., Santos, F.C.F., Nascimento, A.L.A., Zuzarte, M.S., Leite, R.C., Schindler, H.C., 2014. Assessment of messenger RNA (mRNA) of *Mycobacterium tuberculosis* as a marker of cure in patients with pulmonary tuberculosis. *J. Appl. Microbiol.* 117, 266–272. <https://doi.org/10.1111/jam.12508>
- Moore, D.F., Curry, J.I., Knott, C.A., Jonas, V., 1996. Amplification of rRNA for assessment of treatment response of pulmonary tuberculosis patients during antimicrobial therapy. *J. Clin. Microbiol.* 34, 1745–1749. <https://doi.org/10.1128/jcm.34.7.1745-1749.1996>
- Moreira-Teixeira, L., Mayer-Barber, K., Sher, A., O'Garra, A., 2018. Type I interferons in tuberculosis: Foe and occasionally friend. *J. Exp. Med.* 215, 1273–1285. <https://doi.org/10.1084/jem.20180325>
- Mpande, C.A.M., Musvosvi, M., Rozot, V., Mosito, B., Reid, T.D., Schreuder, C., Lloyd, T., Bilek, N., Huang, H., Obermoser, G., Davis, M.M., Ruhwald, M., Hatherill, M., Scriba, T.J., Nemes, E., Mahomed, H., Hanekom, W.A., Kafaar, F., Workman, L., Mulenga, H., Ehrlich, R., Erasmus, M., Abrahams, D., Hawkridge, A., Hughes, E.J., Moyo, S., Gelderbloem, S., Tameris, M., Geldenhuys, H., Hussey, G., 2021. Antigen-Specific T-Cell Activation Distinguishes between Recent and Remote Tuberculosis Infection. *Am. J. Respir. Crit. Care Med.* 203, 1556–1565. <https://doi.org/10.1164/rccm.202007-2686OC>
- Mungai, B., Ong'angò, J., Ku, C.C., Henrion, M.Y.R., Morton, B., Joekes, E., Onyango, E., Kiplimo, R., Kirathe, D., Masini, E., Sitienei, J., Manduku, V., Mugi, B., Squire, S.B., MacPherson, P., Consortium, T.I., 2022. Accuracy of computer-aided chest X-ray in community-based tuberculosis screening: Lessons from the 2016 Kenya National Tuberculosis Prevalence Survey. *PLOS Glob. Public Health* 2, e0001272. <https://doi.org/10.1371/journal.pgph.0001272>
- Musvosvi, M., Duffy, D., Filander, E., Africa, H., Mabwe, S., Jaxa, L., Bilek, N., Llibre, A., Rouilly, V., Hatherill, M., Albert, M., Scriba, T.J., Nemes, E., 2018. T-cell biomarkers for diagnosis of tuberculosis: candidate evaluation by a simple whole blood assay for clinical translation. *Eur. Respir. J.* 51. <https://doi.org/10.1183/13993003.00153-2018>
- Muwanga, V., Mendelsohn, S.C., Leukes, V., Stanley, K., Mbandi, S.K., Erasmus, M., Flinn, M., Fisher, T.-L., Raphela, R., Bilek, N., Malherbe, S.T., Tromp, G., Van der Spuy, G., Walzl, G., Chegou,

- N.N., Scriba, T.J., 2023. Diagnostic Performance of Host Blood Transcriptomic Signatures for Pulmonary Tuberculosis in a Multi-Cohort Study of Symptomatic Patients in Africa. <https://doi.org/10.2139/ssrn.4645194>
- Muyoyeta, M., Maduskar, P., Moyo, M., Kasese, N., Milimo, D., Spooner, R., Kapata, N., Hogeweg, L., Ginneken, B. van, Ayles, H., 2014. The Sensitivity and Specificity of Using a Computer Aided Diagnosis Program for Automatically Scoring Chest X-Rays of Presumptive TB Patients Compared with Xpert MTB/RIF in Lusaka Zambia. *PLOS ONE* 9, e93757. <https://doi.org/10.1371/journal.pone.0093757>
- Nakiyingi, L., Bwanika, J.M., Ssengooba, W., Mubiru, F., Nakanjako, D., Joloba, M.L., Mayanja-Kizza, H., Manabe, Y.C., 2021. Chest X-ray interpretation does not complement Xpert MTB/RIF in diagnosis of smear-negative pulmonary tuberculosis among TB-HIV co-infected adults in a resource-limited setting. *BMC Infect. Dis.* 21, 63. <https://doi.org/10.1186/s12879-020-05752-7>
- Natarajan, S., Ranganathan, M., Hanna, L.E., Tripathy, S., 2022. Transcriptional Profiling and Deriving a Seven-Gene Signature That Discriminates Active and Latent Tuberculosis: An Integrative Bioinformatics Approach. *Genes* 13, 616. <https://doi.org/10.3390/genes13040616>
- Nathavitharana, R.R., Garcia-Basteiro, A.L., Ruhwald, M., Cobelens, F., Theron, G., 2022. Reimagining the status quo: How close are we to rapid sputum-free tuberculosis diagnostics for all? *eBioMedicine* 78, 103939. <https://doi.org/10.1016/j.ebiom.2022.103939>
- Nogueira, B.M.F., Krishnan, S., Barreto-Duarte, B., Araújo-Pereira, M., Queiroz, A.T.L., Ellner, J.J., Salgame, P., Scriba, T.J., Sterling, T.R., Gupta, A., Andrade, B.B., 2022. Diagnostic biomarkers for active tuberculosis: progress and challenges. *EMBO Mol. Med.* e14088. <https://doi.org/10.15252/emmm.202114088>
- Occhineri, S., Matucci, T., Rindi, L., Tiseo, G., Falcone, M., Riccardi, N., Besozzi, G., 2022. Pretomanid for tuberculosis treatment: an update for clinical purposes. *Curr. Res. Pharmacol. Drug Discov.* 3, 100128. <https://doi.org/10.1016/j.crphar.2022.100128>
- Oliveira-de-Souza, D., Vinhaes, C.L., Arriaga, M.B., Kumar, N.P., Cubillos-Angulo, J.M., Shi, R., Wei, W., Yuan, X., Zhang, G., Cai, Y., Barry, C.E., Via, L.E., Sher, A., Babu, S., Mayer-Barber, K.D., Nakaya, H.I., Fukutani, K.F., Andrade, B.B., 2019. Molecular degree of perturbation of plasma inflammatory markers associated with tuberculosis reveals distinct disease profiles between Indian and Chinese populations. *Sci. Rep.* 9, 8002. <https://doi.org/10.1038/s41598-019-44513-8>
- Ottenhoff, T.H.M., Dass, R.H., Yang, N., Zhang, M.M., Wong, H.E.E., Sahiratmadja, E., Khor, C.C., Alisjahbana, B., van Crevel, R., Marzuki, S., Seielstad, M., van de Vosse, E., Hibberd, M.L., 2012. Genome-Wide Expression Profiling Identifies Type 1 Interferon Response Pathways in Active Tuberculosis. *PLoS ONE* 7, e45839. <https://doi.org/10.1371/journal.pone.0045839>
- Pai, M., Zwerling, A., Menzies, D., 2008. Systematic Review: T-Cell-based Assays for the Diagnosis of Latent Tuberculosis Infection: An Update. *Ann. Intern. Med.* 149, 177–184. <https://doi.org/10.7326/0003-4819-149-3-200808050-00241>
- Pan, L., Wei, N., Jia, H., Gao, M., Chen, X., Wei, R., Sun, Q., Gu, S., Du, B., Xing, A., Zhang, Z., 2017. Genome-wide transcriptional profiling identifies potential signatures in discriminating active tuberculosis from latent infection. *Oncotarget* 8, 112907–112916. <https://doi.org/10.18632/oncotarget.22889>
- Parwati, I., Pitaloka, D.A.E., Chaidir, L., 2022. Transcriptional Biomarkers for Treatment Monitoring of Pulmonary Drug-Resistant Tuberculosis: Protocol for a Prospective Observational Study in Indonesia. *Trop. Med. Infect. Dis.* 7, 326. <https://doi.org/10.3390/tropicalmed7110326>
- Pease, C., Hutton, B., Yazdi, F., Wolfe, D., Hamel, C., Barbeau, P., Skidmore, B., Alvarez, G.G., 2018. A systematic review of adverse events of rifapentine and isoniazid compared to other treatments for latent tuberculosis infection. *Pharmacoepidemiol. Drug Saf.* 27, 557–566. <https://doi.org/10.1002/pds.4423>

- Penn-Nicholson, A., Hraha, T., Thompson, E.G., Sterling, D., Mbandi, S.K., Wall, K.M., Fisher, M., Suliman, S., Shankar, S., Hanekom, W.A., Janjic, N., Hatherill, M., Kaufmann, S.H.E., Sutherland, J., Walzl, G., Groote, M.A.D., Ochsner, U., Zak, D.E., Scriba, T.J., Groups, A. and G.-74 cohort study, 2019. Discovery and validation of a prognostic proteomic signature for tuberculosis progression: A prospective cohort study. *PLOS Med.* 16, e1002781. <https://doi.org/10.1371/journal.pmed.1002781>
- Penn-Nicholson, A., Mbandi, S.K., Thompson, E., Mendelsohn, S.C., Suliman, S., Chegou, N.N., Malherbe, S.T., Darboe, F., Erasmus, M., Hanekom, W.A., Bilek, N., Fisher, M., Kaufmann, S.H.E., Winter, J., Murphy, M., Wood, R., Morrow, C., Van Rhijn, I., Moody, B., Murray, M., Andrade, B.B., Sterling, T.R., Sutherland, J., Naidoo, K., Padayatchi, N., Walzl, G., Hatherill, M., Zak, D., Scriba, T.J., 2020a. RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response. *Sci. Rep.* 10, 8629. <https://doi.org/10.1038/s41598-020-65043-8>
- Penn-Nicholson, A., Mbandi, S.K., Thompson, E., Mendelsohn, S.C., Suliman, S., Chegou, N.N., Malherbe, S.T., Darboe, F., Erasmus, M., Hanekom, W.A., Bilek, N., Fisher, M., Kaufmann, S.H.E., Winter, J., Murphy, M., Wood, R., Morrow, C., Van Rhijn, I., Moody, B., Murray, M., Andrade, B.B., Sterling, T.R., Sutherland, J., Naidoo, K., Padayatchi, N., Walzl, G., Hatherill, M., Zak, D., Scriba, T.J., 2020b. RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response. *Sci. Rep.* 10, 8629. <https://doi.org/10.1038/s41598-020-65043-8>
- Peralta, G., Barry, P., Pascopella, L., 2016. Use of Nucleic Acid Amplification Tests in Tuberculosis Patients in California, 2010–2013. *Open Forum Infect. Dis.* 3, ofw230. <https://doi.org/10.1093/ofid/ofw230>
- Perumal, P., Abdullatif, M.B., Garland, H.N., Honeyborne, I., Lipman, M., McHugh, T.D., Southern, J., Breen, R., Santis, G., Ellappan, K., Kumar, S.V., Belgode, H., Abubakar, I., Sinha, S., Vasan, S.S., Joseph, N., Kempell, K.E., 2021. Validation of Differentially Expressed Immune Biomarkers in Latent and Active Tuberculosis by Real-Time PCR. *Front. Immunol.* 11. <https://doi.org/10.3389/fimmu.2020.612564>
- Petrilli, J.D., Araújo, L.E., da Silva, L.S., Laus, A.C., Müller, I., Reis, R.M., Netto, E.M., Riley, L.W., Arruda, S., Queiroz, A., 2020. Whole blood mRNA expression-based targets to discriminate active tuberculosis from latent infection and other pulmonary diseases. *Sci. Rep.* 10, 22072. <https://doi.org/10.1038/s41598-020-78793-2>
- Petrucchioli, E., Petrone, L., Vanini, V., Cuzzi, G., Navarra, A., Gualano, G., Palmieri, F., Girardi, E., Goletti, D., 2015. Assessment of CD27 expression as a tool for active and latent tuberculosis diagnosis. *J. Infect.* 71, 526–533. <https://doi.org/10.1016/j.jinf.2015.07.009>
- Piatek, A.S., Van Cleeff, M., Alexander, H., Coggin, W.L., Rehr, M., Van Kampen, S., Shinnick, T.M., Mukadi, Y., 2013. GeneXpert for TB diagnosis: planned and purposeful implementation. *Glob. Health Sci. Pract.* 1, 18–23. <https://doi.org/10.9745/GHSP-D-12-00004>
- Pitaloka, D.A.E., Syamsunarno, M.R.A.A., Abdulah, R., Chaidir, L., 2022. Omics Biomarkers for Monitoring Tuberculosis Treatment: A Mini-Review of Recent Insights and Future Approaches. *Infect. Drug Resist.* 15, 2703–2711. <https://doi.org/10.2147/IDR.S366580>
- Portevin, D., Moukambi, F., Clowes, P., Bauer, A., Chachage, M., Ntinginya, N.E., Mfinanga, E., Said, K., Haraka, F., Rachow, A., Saathoff, E., Mpina, M., Jugheli, L., Lwilla, F., Marais, B.J., Hoelscher, M., Daubenberger, C., Reither, K., Geldmacher, C., 2014. Assessment of the novel T-cell activation marker–tuberculosis assay for diagnosis of active tuberculosis in children: a prospective proof-of-concept study. *Lancet Infect. Dis.* 14, 931–938. [https://doi.org/10.1016/S1473-3099\(14\)70884-9](https://doi.org/10.1016/S1473-3099(14)70884-9)
- Prada-Medina, C.A., Fukutani, K.F., Pavan Kumar, N., Gil-Santana, L., Babu, S., Lichtenstein, F., West, K., Sivakumar, S., Menon, P.A., Viswanathan, V., Andrade, B.B., Nakaya, H.I., Kornfeld, H., 2017. Systems Immunology of Diabetes-Tuberculosis Comorbidity Reveals Signatures of Disease Complications. *Sci. Rep.* 7, 1999. <https://doi.org/10.1038/s41598-017-01767-4>

- Qiu, Q., Peng, A., Zhao, Y., Liu, D., Liu, C., Qiu, S., Xu, J., Cheng, H., Xiong, W., Chen, Y., 2022. Diagnosis of pulmonary tuberculosis via identification of core genes and pathways utilizing blood transcriptional signatures: a multicohort analysis. *Respir. Res.* 23, 125. <https://doi.org/10.1186/s12931-022-02035-4>
- Rachinger, N., Fischer, S., Böhme, I., Linck-Paulus, L., Kuphal, S., Kappelmann-Fenzl, M., Bosserhoff, A.K., 2021. Loss of Gene Information: Discrepancies between RNA Sequencing, cDNA Microarray, and qRT-PCR. *Int. J. Mol. Sci.* 22, 9349. <https://doi.org/10.3390/ijms22179349>
- Rageade, F., Picot, N., Blanc-Michaud, A., Chatellier, S., Mirande, C., Fortin, E., van Belkum, A., 2014. Performance of solid and liquid culture media for the detection of *Mycobacterium tuberculosis* in clinical materials: meta-analysis of recent studies. *Eur. J. Clin. Microbiol. Infect. Dis.* 33, 867–870. <https://doi.org/10.1007/s10096-014-2105-z>
- Rajan, J.V., Semitala, F.C., Mehta, T., Seielstad, M., Montalvo, L., Andama, A., Asege, L., Nakaye, M., Katende, J., Mwebe, S., Kamya, M.R., Yoon, C., Cattamanchi, A., 2019. A Novel, 5-Transcript, Whole-blood Gene-expression Signature for Tuberculosis Screening Among People Living With Human Immunodeficiency Virus. *Clin. Infect. Dis.* 69, 77–83. <https://doi.org/10.1093/cid/ciy835>
- Rangaka, M.X., Wilkinson, K.A., Glynn, J.R., Ling, D., Menzies, D., Mwansa-Kambafwile, J., Fielding, K., Wilkinson, R.J., Pai, M., 2012. Predictive value of interferon- γ release assays for incident active tuberculosis: a systematic review and meta-analysis. *Lancet Infect. Dis.* 12, 45–55. [https://doi.org/10.1016/S1473-3099\(11\)70210-9](https://doi.org/10.1016/S1473-3099(11)70210-9)
- Richardson, T.R., Smith, B., Malherbe, S.T., Shaw, J.A., Noor, F., MacDonald, C., van der Spuy, G.D., Stanley, K., Carstens, A., Fisher, T.-L., van Rensburg, I., Flinn, M., Snyders, C., Johnson, I., Fransman, B., Dockrell, H., Thwaites, G., Thuong, N.T.T., Schacht, C., Mayanja-Kizza, H., Nsereko, M., Tjon Kon Fat, E.M., Corstjens, P.L.A.M., Geluk, A., Ruhwald, Morton, Penn-Nicholson, A., Chegou, N.N., Sutherland, J., Walzl, G., Hiemstra, A., Tonsing, S., Tromp, G., Owolabi, M., Mendy, J., Gindeh, A., Barry, A., Mbayo, G., Buech, J., Streitz, M., Nalukwago, S., Namuganga, A.R., Lamunu, D., Odie, M., Pierneef, L., van Hooij, A., Ruhwald, Morten, Belisle, J., Dobos, K., Hatherill, M., Scriba, T., Winter, J., TrENDx consortium, 2023. Field evaluation of a point-of-care triage test for active tuberculosis (TriageTB). *BMC Infect. Dis.* 23, 447. <https://doi.org/10.1186/s12879-023-08342-5>
- Ruperez, M., Shanaube, K., Mureithi, L., Wapamesa, C., Burnett, M.J., Kosloff, B., Haas, P. de, Hayes, R., Fidler, S., Gachie, T., Schaap, A., Floyd, S., Klinkenberg, E., Ayles, H., Africa, A., Amofa-Skeyi, M., Bond, V., Cheeba, M., Dodd, P., Kalisvaart, N., Kangololo, B., Kasese, N., Mainga, T., Mwinga, A., Nikolayevskyy, V., Nyondo, B., Paulsen, R., Simwanga, M., Sisam, C., Telisinghe, L., Thomas, R., Vermaak, R., Vijn, F., 2023. Use of point-of-care C-reactive protein testing for screening of tuberculosis in the community in high-burden settings: a prospective, cross-sectional study in Zambia and South Africa. *Lancet Glob. Health* 11, e704–e714. [https://doi.org/10.1016/S2214-109X\(23\)00113-4](https://doi.org/10.1016/S2214-109X(23)00113-4)
- Sambarey, A., Devaprasad, A., Mohan, A., Ahmed, A., Nayak, S., Swaminathan, S., D'Souza, G., Jesuraj, A., Dhar, C., Babu, S., Vyakarnam, A., Chandra, N., 2017. Unbiased Identification of Blood-based Biomarkers for Pulmonary Tuberculosis by Modeling and Mining Molecular Interaction Networks. *eBioMedicine* 15, 112–126. <https://doi.org/10.1016/j.ebiom.2016.12.009>
- Santos, V.S., Goletti, D., Kontogianni, K., Adams, E.R., Molina-Moya, B., Dominguez, J., Crudu, V., Martins-Filho, P.R.S., Ruhwald, M., Lawson, L., Bimba, J.S., Garcia-Basteiro, A.L., Petrone, L., Kabeer, B.S., Reither, K., Cuevas, L.E., 2019. Acute phase proteins and IP-10 as triage tests for the diagnosis of tuberculosis: systematic review and meta-analysis. *Clin. Microbiol. Infect.* 25, 169–177. <https://doi.org/10.1016/j.cmi.2018.07.017>
- Satproedprai, N., Wichukchinda, N., Suphankong, S., Inunchot, W., Kuntima, T., Kumpeerasart, S., Wattanapokayakit, S., Nedsuwan, S., Yanai, H., Higuchi, K., Harada, N., Mahasirimongkol, S.,

2015. Diagnostic value of blood gene expression signatures in active tuberculosis in Thais: a pilot study. *Genes Immun.* 16, 253–260. <https://doi.org/10.1038/gene.2015.4>
- Schorey, J.S., Carroll, M.C., Brown, E.J., 1997. A Macrophage Invasion Mechanism of Pathogenic *Mycobacteria*. *Science* 277, 1091–1093. <https://doi.org/10.1126/science.277.5329.1091>
- Schramm, B., Nganaboy, R.C., Uwiragiye, P., Mukenba, D., Abdoubara, A., Abdou, I., Nshimiymana, J.-C., Sounna, S., Hiffler, L., Flevaud, L., Huerga, H., 2021. Potential value of urine lateral-flow lipoarabinomannan (LAM) test for diagnosing tuberculosis among severely acute malnourished children. *PLOS ONE* 16, e0250933. <https://doi.org/10.1371/journal.pone.0250933>
- Schwalb, A., Emery, J.C., Houben, R.M.G.J., 2022. Use of chest radiography screening for TB: a re-evaluation of the Kolín study. *Int. J. Tuberc. Lung Dis.* 26, 983–985. <https://doi.org/10.5588/ijtld.22.0216>
- Scott, L.E., McCarthy, K., Gous, N., Nduna, M., Rie, A.V., Sanne, I., Venter, W.F., Duse, A., Stevens, W., 2011. Comparison of Xpert MTB/RIF with Other Nucleic Acid Technologies for Diagnosing Pulmonary Tuberculosis in a High HIV Prevalence Setting: A Prospective Study. *PLOS Med.* 8, e1001061. <https://doi.org/10.1371/journal.pmed.1001061>
- Scriba, T.J., Fiore-Gartland, A., Penn-Nicholson, A., Mulenga, H., Mbandi, S.K., Borate, B., Mendelsohn, S.C., Hadley, K., Hikuam, C., Kaskar, M., Musvosvi, M., Bilek, N., Self, S., Sumner, T., White, R.G., Erasmus, Mzwandile, Jaxa, L., Raphela, R., Innes, C., Brumskine, W., Hiemstra, Andriette, Malherbe, S.T., Hassan-Moosa, R., Tameris, M., Walzl, G., Naidoo, K., Churchyard, G., Hatherill, M., Baepanye, K., Baepanye, T., Clarke, K., Collignon, M., Dlamini, A., Eyre, C., Feni, T., Fikizolo, M., Galane, P., Goliath, T., Gangat, A., Malefo-Grootboom, S., Rensburg, E.J. van, Rensburg, B.J. van, Kekana, S., Zietsman, M., Kock, A., Kunene, I., Lakhi, A., Langa, N., Ledwaba, H., Luphoko, M., Mabasa, I., Mabe, D., Mabuza, N., Majola, M., Makhetha, M., Makoanyane, M., Makhubalo, B., Malay, V., Market, J., Matshego, S., Mbipa, N., Mmotsa, T., Modipa, S., Mopati, S., Moswegu, P., Mothaga, P., Muller, D., Nchwe, G., Nel, M., Nhlangulela, L., Ntamo, B., Ntoahae, L., Ntshauba, T., Sanyaka, N., Seabela, L., Selepe, P., Senne, M., Serake, M.G., Thlapi, M., Tshikovhi, V., Tswaile, L., Aswegen, A. van, Mbata, L., Takavamanya, C., Pinho, P., Mdlulu, J., Taljaard, M., Slabbert, N., Sayed, S., Nielson, T., Senne, M., Sein, N.N., Mbata, L., Govender, D., Chinappa, T., Zulu, M.I., Maphanga, N.B., Hlathi, S.R., Gumede, G.K., Shezi, T.Y., Maphanga, J.L., Jali, Z.P., Cwele, T., Gwamanda, N.Z.E., Dlamini, C., Sing, Z.P.P., Ntanjana, N.G., Nzimande, S.S., Mbatha, S., Maharaj, B., Moosa, A., Corris, C.-M., Kafaar, F., Geldenhuys, H., Luabeya, A.K.K., Shenje, J., Botes, N., Rossouw, S., Africa, H., Diamond, B., Braaf, S., Stryers, S., Carstens, A., Jansen, R., Mabwe, S., Mulenga, H., Herling, R., Veldsman, A., Makhete, L., Steyn, M., Buhlungu, S., Erasmus, Margareth, Davids, I., Plaatjie, P., Companie, A., Ratangee, F., Veldtsman, H., Petersen, C., Abrahams, C., Moses, M., Kelepu, X., Gregg, Y., Swanepoel, L., Magawu, N., Cetywayo, N., Mactavie, L., Valley, H., Filander, E., Nqakala, N., Maasdorp, E., Khoury, J., Kriel, B., Smith, B., Muller, L., Tonsing, S., Loxton, A., Hiemstra, Andriette, Ahlers, P., Flinn, M., Chung, E., Chung, M., Sato, A., 2021. Biomarker-guided tuberculosis preventive therapy (CORTIS): a randomised controlled trial. *Lancet Infect. Dis.* 21, 354–365. [https://doi.org/10.1016/S1473-3099\(20\)30914-2](https://doi.org/10.1016/S1473-3099(20)30914-2)
- Scriba, T.J., Penn-Nicholson, A., Shankar, S., Hraha, T., Thompson, E.G., Sterling, D., Nemes, E., Darboe, F., Suliman, S., Amon, L.M., Mahomed, H., Erasmus, M., Whatney, W., Johnson, J.L., Boom, W.H., Hatherill, M., Valvo, J., Groote, M.A.D., Ochsner, U.A., Aderem, A., Hanekom, W.A., Zak, D.E., Team, other members of the A. cohort study, 2017. Sequential inflammatory processes define human progression from *M. tuberculosis* infection to tuberculosis disease. *PLOS Pathog.* 13, e1006687. <https://doi.org/10.1371/journal.ppat.1006687>
- Serrano, C.J., Cuevas-Córdoba, B., Macías-Segura, N., González-Curiel, R.A., Martínez-Balderas, V.Y., Enciso-Moreno, L., Small, P., Hernández-Pando, R., Enciso-Moreno, J.A., 2016. Transcriptional profiles discriminate patients with pulmonary tuberculosis from non-

- tuberculous individuals depending on the presence of non-insulin diabetes mellitus. *Clin. Immunol.* 162, 107–117. <https://doi.org/10.1016/j.clim.2015.11.008>
- Silvi, S., Credi, A., 2015. Luminescent sensors based on quantum dot–molecule conjugates. *Chem. Soc. Rev.* 44, 4275–4289. <https://doi.org/10.1039/C4CS00400K>
- Singhania, A., Verma, R., Graham, C.M., Lee, J., Tran, T., Richardson, M., Lecine, P., Leissner, P., Berry, M.P.R., Wilkinson, R.J., Kaiser, K., Rodrigue, M., Woltmann, G., Haldar, P., O’Garra, A., 2018a. A modular transcriptional signature identifies phenotypic heterogeneity of human tuberculosis infection. *Nat. Commun.* 9, 1–17. <https://doi.org/10.1038/s41467-018-04579-w>
- Singhania, A., Wilkinson, R.J., Rodrigue, M., Haldar, P., O’Garra, A., 2018b. The value of transcriptomics in advancing knowledge of the immune response and diagnosis in tuberculosis. *Nat. Immunol.* 19, 1159–1168. <https://doi.org/10.1038/s41590-018-0225-9>
- Södersten, E., Ongarello, S., Mantsoki, A., Wyss, R., Persing, D.H., Banderby, S., Strömquist Meuzelaar, L., Prieto, J., Gnanashanmugam, D., Khatri, P., Schumacher, S.G., Denking, C.M., 2021. Diagnostic Accuracy Study of a Novel Blood-Based Assay for Identification of Tuberculosis in People Living with HIV. *J. Clin. Microbiol.* 59, e01643-20. <https://doi.org/10.1128/JCM.01643-20>
- Stadler, J.A.M., 2022. Updated WHO definitions for tuberculosis outcomes: Simplified, unified and future-proofed. *Afr. J. Thorac. Crit. Care Med.* 28, 10.7196/AJTCCM.2022.v28i2.224. <https://doi.org/10.7196/AJTCCM.2022.v28i2.224>
- Steingart, K.R., Flores, L.L., Dendukuri, N., Schiller, I., Laal, S., Ramsay, A., Hopewell, P.C., Pai, M., 2011. Commercial serological tests for the diagnosis of active pulmonary and extrapulmonary tuberculosis: an updated systematic review and meta-analysis. *PLoS Med.* 8, e1001062. <https://doi.org/10.1371/journal.pmed.1001062>
- Suliman, S., Thompson, E.G., Sutherland, J., Weiner, J., Ota, M.O.C., Shankar, S., Penn-Nicholson, A., Thiel, B., Erasmus, M., Maertzdorf, J., Duffy, F.J., Hill, P.C., Hughes, E.J., Stanley, K., Downing, K., Fisher, M.L., Valvo, J., Parida, S.K., van der Spuy, G., Tromp, Gerard, Adetifa, I.M.O., Donkor, S., Howe, R., Mayanja-Kizza, H., Boom, W.H., Dockrell, H.M., Ottenhoff, T.H.M., Hatherill, M., Aderem, A., Hanekom, W.A., Scriba, T.J., Kaufmann, S.H.E., Zak, D.E., Walzl, G., Walzl, G., Black, G.F., van der Spuy, G., Stanley, K., Kriel, M., Du Plessis, N., Nene, N., Roberts, T., Kleyhans, L., Gutschmidt, A., Smith, B., Nene, N., Loxton, A.G., Chegou, N.N., Tromp, Gerhardus, Tabb, D., Ottenhoff, T.H.M., Klein, M.R., Haks, M.C., Franken, K.L.M.C., Geluk, A., van Meijgaarden, K.E., Joosten, S.A., Boom, W.H., Thiel, B., Mayanja-Kizza, H., Joloba, M., Zalwango, S., Nsereko, M., Okwera, B., Kisingo, H., Kaufmann, S.H.E., Parida, S.K., Golinski, R., Maertzdorf, J., Weiner, J., Jacobson, M., Dockrell, H., Smith, S., Gorak-Stolinska, P., Hur, Y.-G., Lalor, M., Lee, J.-S., Crampin, A.C., French, N., Ngwira, B., Ben-Smith, A., Watkins, K., Ambrose, L., Simukonda, F., Mvula, H., Chilongo, F., Saul, J., Branson, K., Suliman, S., Scriba, T.J., Mahomed, H., Hughes, E.J., Bilek, N., Erasmus, M., Xasa, O., Veldsman, A., Downing, K., Fisher, M., Penn-Nicholson, A., Mulenga, H., Abel, B., Bowmaker, M., Kagina, B., Chung, W.K., Hanekom, W.A., Sadoff, J., Sizemore, D., Ramachandran, S., Barker, L., Brennan, M., Weichold, F., Muller, S., Geiter, L., Kassa, D., Abebe, A., Mesele, T., Tegbaru, B., Baarle, D. van, Miedema, F., Howe, R., Mihret, A., Aseffa, A., Bekele, Y., Iwnetu, R., Tafesse, M., Yamuah, L., Ota, M., Sutherland, J., Hill, P., Adegbola, R., Corrah, T., Antonio, M., Togun, T., Adetifa, I., Donkor, S., Andersen, P., Rosenkrands, I., Doherty, M., Weldingh, K., Schoolnik, G., Dolganov, G., Van, T., Kafaar, F., Workman, L., Mulenga, H., Scriba, T.J., Hughes, E.J., Bilek, N., Erasmus, M., Xasa, O., Veldsman, A., Cloete, Y., Abrahams, D., Moyo, S., Gelderbloem, S., Tameris, M., Geldenhuys, H., Hanekom, W., Hussey, G., Ehrlich, R., Verver, S., Geiter, L., 2018. Four-Gene Pan-African Blood Signature Predicts Progression to Tuberculosis. *Am. J. Respir. Crit. Care Med.* 197, 1198–1208. <https://doi.org/10.1164/rccm.201711-2340OC>
- Sutherland, J.S., van der Spuy, G., Gindeh, A., Thuong, N.T.T., Namuganga, A., Owolabi, O., Mayanja-Kizza, H., Nsereko, M., Thwaites, G., Winter, J., Dockrell, H.M., Scriba, T.J., Geluk, A.,

- Corstjens, P., Stanley, K., Richardson, T., Shaw, J.A., Smith, B., Malherbe, S.T., Walzl, G., TrENDx-TB Consortium, Sutherland, J., Owolabi, O., Secka, A., Daffeh, G., Gindeh, A., Mendy, J., Sarr, B., Riley, A.-J., Jobe, A., Davies, M., Kanyi, K., Jallow, M., Barry, S., Cham, O., Nkereuwem, E., Walzl, G., Malherbe, S., Smith, B., van der Spuy, G., Stanley, K., Shaw, J., Chetram, A., Richardson, T., Finn, M., Hiemstra, A., Chegou, N., Kuivaniemi, H., Tromp, G., Tonsing, S., Smit, E., Carstens, B., Mayanja-Kizza, H., Nsereko, M., Namuganga, A., Nalukwago, S., Akol, J., Lamunu, D., Ordie, M., Thwaites, G., Nguyen, T., Le, V., Vo Thanh, S., Thi, H.N., Thi Ngoc, H.V., Le Hong, N., Belisle, J., Dobos, K., Dockrell, H., Scriba, T., Hatherill, M., Hadley, K., Shenje, J., Kimbung, S., Mulenga, H., Oelofse, R., Bilek, N., van Rooyen, E., Mabwe, S., Corstjens, P., Geluk, A., Kon Fat, E.T., Pierneef, L., van Hooij, A., Winter, J., Ruhwald, M., Moreau, E., Penn-Nicholson, A., Schacht, C., Büech, J., Streitz, M., 2022. Diagnostic Accuracy of the Cepheid 3-gene Host Response Fingertick Blood Test in a Prospective, Multi-site Study: Interim Results. *Clin. Infect. Dis.* 74, 2136–2141. <https://doi.org/10.1093/cid/ciab839>
- Sweeney, T.E., Braviak, L., Tato, C.M., Khatri, P., 2016. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. *Lancet Respir. Med.* 4, 213–224. [https://doi.org/10.1016/S2213-2600\(16\)00048-5](https://doi.org/10.1016/S2213-2600(16)00048-5)
- Tabone, O., Verma, R., Singhanian, A., Chakravarty, P., Branchett, W.J., Graham, C.M., Lee, J., Trang, T., Reynier, F., Leissner, P., Kaiser, K., Rodrigue, M., Woltmann, G., Haldar, P., O’Garra, A., 2021. Blood transcriptomics reveal the evolution and resolution of the immune response in tuberculosis. *J. Exp. Med.* 218, e20210915. <https://doi.org/10.1084/jem.20210915>
- Target product profiles for tests for tuberculosis treatment monitoring and optimization [WWW Document], n.d. URL <https://www.who.int/publications-detail-redirect/9789240081178> (accessed 12.22.23).
- Thompson, E.G., Du, Y., Malherbe, S.T., Shankar, S., Braun, J., Valvo, J., Ronacher, K., Tromp, G., Tabb, D.L., Alland, D., Shenai, S., Via, L.E., Warwick, J., Aderem, A., Scriba, T.J., Winter, J., Walzl, G., Zak, D.E., Du Plessis, N., Loxton, A.G., Chegou, N.N., Lee, M., 2017a. Host blood RNA signatures predict the outcome of tuberculosis treatment. *Tuberculosis* 107, 48–58. <https://doi.org/10.1016/j.tube.2017.08.004>
- Thompson, E.G., Du, Y., Malherbe, S.T., Shankar, S., Braun, J., Valvo, J., Ronacher, K., Tromp, G., Tabb, D.L., Alland, D., Shenai, S., Via, L.E., Warwick, J., Aderem, A., Scriba, T.J., Winter, J., Walzl, G., Zak, D.E., Du Plessis, N., Loxton, A.G., Chegou, N.N., Lee, M., 2017b. Host blood RNA signatures predict the outcome of tuberculosis treatment. *Tuberculosis* 107, 48–58. <https://doi.org/10.1016/j.tube.2017.08.004>
- Thwaites, G., 2014. 40 - Tuberculosis, in: Farrar, J., Hotez, P.J., Junghanss, T., Kang, G., Lalloo, D., White, N.J. (Eds.), *Manson’s Tropical Infectious Diseases (Twenty-Third Edition)*. W.B. Saunders, London, pp. 468-505.e3. <https://doi.org/10.1016/B978-0-7020-5101-2.00041-8>
- Togun, T., Kampmann, B., Pai, M., 2017. Diagnosis of Childhood Tuberculosis, in: *Reference Module in Biomedical Sciences*. Elsevier. <https://doi.org/10.1016/B978-0-12-801238-3.64157-0>
- Togun, T.O., MacLean, E., Kampmann, B., Pai, M., 2018. Biomarkers for diagnosis of childhood tuberculosis: A systematic review. *PLoS ONE* 13, e0204029. <https://doi.org/10.1371/journal.pone.0204029>
- Tornheim, J.A., Madugundu, A.K., Paradkar, M., Fukutani, K.F., Queiroz, A.T.L., Gupte, N., Gupte, A.N., Kinikar, A., Kulkarni, V., Balasubramanian, U., Sreenivasamurthy, S., Raja, R., Pradhan, N., Shivakumar, S.V.B.Y., Valvi, C., Hanna, L.E., Andrade, B.B., Mave, V., Pandey, A., Gupta, A., CTRIUMPh RePORT India Study Team, 2020. Transcriptomic Profiles of Confirmed Pediatric Tuberculosis Patients and Household Contacts Identifies Active Tuberculosis, Infection, and Treatment Response Among Indian Children. *J. Infect. Dis.* 221, 1647–1658. <https://doi.org/10.1093/infdis/jiz639>
- Tornheim, J.A., Madugundu, A.K., Paradkar, M., Fukutani, K.F., Queiroz, A.T.L., Gupte, N., Gupte, A.N., Kinikar, A., Kulkarni, V., Balasubramanian, U., Sreenivasamurthy, S., Raja, R., Pradhan,

- N., Shivakumar, S.V.B.Y., Valvi, C., Hanna, L.E., Andrade, B.B., Mave, V., Pandey, A., Gupta, A., CTRIUMPh RePORT India Study Team, 2019. Transcriptomic Profiles of Confirmed Pediatric Tuberculosis Patients and Household Contacts Identifies Active Tuberculosis, Infection, and Treatment Response Among Indian Children. *J. Infect. Dis.* jiz639. <https://doi.org/10.1093/infdis/jiz639>
- Turkova, A., Wills, G.H., Wobudeya, E., Chabala, C., Palmer, M., Kinikar, A., Hissar, S., Choo, L., Musoke, P., Mulenga, V., Mave, V., Joseph, B., LeBeau, K., Thomason, M.J., Mboizi, R.B., Kapasa, M., van der Zalm, M.M., Raichur, P., Bhavani, P.K., McIlhleron, H., Demers, A.-M., Aarnoutse, R., Love-Koh, J., Seddon, J.A., Welch, S.B., Graham, S.M., Hesselning, A.C., Gibb, D.M., Crook, A.M., 2022. Shorter Treatment for Nonsevere Tuberculosis in African and Indian Children. *N. Engl. J. Med.* 386, 911–922. <https://doi.org/10.1056/NEJMoa2104535>
- Turner, C.T., Gupta, R.K., Tsaliki, E., Roe, J.K., Mondal, P., Nyawo, G.R., Palmer, Z., Miller, R.F., Reeve, B.W., Theron, G., Noursadeghi, M., 2020. Blood transcriptional biomarkers for active pulmonary tuberculosis in a high-burden setting: a prospective, observational, diagnostic accuracy study. *Lancet Respir. Med.* 8, 407–419. [https://doi.org/10.1016/S2213-2600\(19\)30469-2](https://doi.org/10.1016/S2213-2600(19)30469-2)
- van Doorn, C.L.R., Eckold, C., Ronacher, K., Ruslami, R., van Veen, S., Lee, J.-S., Kumar, V., Kerry-Barnard, S., Malherbe, S.T., Kleynhans, L., Stanley, K., Hill, P.C., Joosten, S.A., van Crevel, R., Wijmenga, C., Critchley, J.A., Walzl, G., Alisjahbana, B., Haks, M.C., Dockrell, H.M., Ottenhoff, T.H.M., Vianello, E., Cliff, J.M., 2022. Transcriptional profiles predict treatment outcome in patients with tuberculosis and diabetes at diagnosis and at two weeks after initiation of anti-tuberculosis treatment. *eBioMedicine* 82, 104173. <https://doi.org/10.1016/j.ebiom.2022.104173>
- Walter, N.D., Miller, M.A., Vasquez, J., Weiner, M., Chapman, A., Engle, M., Higgins, M., Quinones, A.M., Rosselli, V., Canono, E., Yoon, C., Cattamanchi, A., Davis, J.L., Phang, T., Stearman, R.S., Datta, G., Garcia, B.J., Daley, C.L., Strong, M., Kechris, K., Fingerlin, T.E., Reves, R., Geraci, M.W., 2016. Blood Transcriptional Biomarkers for Active Tuberculosis among Patients in the United States: a Case-Control Study with Systematic Cross-Classifer Evaluation. *J. Clin. Microbiol.* 54, 274–282. <https://doi.org/10.1128/JCM.01990-15>
- Wang, Z., Gerstein, M., Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10, 57–63. <https://doi.org/10.1038/nrg2484>
- Warsinske, H.C., Rao, A.M., Moreira, F.M.F., Santos, P.C.P., Liu, A.B., Scott, M., Malherbe, S.T., Ronacher, K., Walzl, G., Winter, J., Sweeney, T.E., Croda, J., Andrews, J.R., Khatri, P., 2018. Assessment of Validity of a Blood-Based 3-Gene Signature Score for Progression and Diagnosis of Tuberculosis, Disease Severity, and Treatment Response. *JAMA Netw. Open* 1, e183779. <https://doi.org/10.1001/jamanetworkopen.2018.3779>
- Whitelaw, A., Peter, J., Sohn, H., Viljoen, D., Theron, G., Badri, M., Davids, V., Pai, M., Dheda, K., 2011. Comparative cost and performance of light-emitting diode microscopy in HIV–tuberculosis-co-infected patients. *Eur. Respir. J.* 38, 1393–1397. <https://doi.org/10.1183/09031936.00023211>
- WHO consolidated guidelines on tuberculosis: module 2: screening: systematic screening for tuberculosis disease [WWW Document], n.d. URL <https://www.who.int/publications-detail-redirect/9789240022676> (accessed 3.21.24).
- Wilkinson, K.A., Oni, T., Gideon, H.P., Goliath, R., Wilkinson, R.J., Riou, C., 2016. Activation Profile of Mycobacterium tuberculosis–Specific CD4+ T Cells Reflects Disease Activity Irrespective of HIV Status. *Am. J. Respir. Crit. Care Med.* 193, 1307–1310. <https://doi.org/10.1164/rccm.201601-0116LE>
- Williams, C.G., Lee, H.J., Asatsuma, T., Vento-Tormo, R., Haque, A., 2022. An introduction to spatial transcriptomics for biomedical research. *Genome Med.* 14, 68. <https://doi.org/10.1186/s13073-022-01075-1>

- World Health Organization, 2023. 5.5.1 Monitoring treatment response and outcome assignment | TB Knowledge Sharing [WWW Document]. URL <https://tbksp.org/en/node/2516> (accessed 5.18.24).
- World Health Organization, 2022a. Global Tuberculosis Report 2022 [WWW Document]. WHO Glob. Tuberc. Rep. 2022. URL <https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2022> (accessed 11.16.22).
- World Health Organization, 2022b. WHO consolidated guidelines on tuberculosis: module 4: treatment: drug-susceptible tuberculosis treatment [WWW Document]. URL <https://www.who.int/publications-detail-redirect/9789240048126> (accessed 5.20.24).
- World Health Organization, 2022c. WHO consolidated guidelines on tuberculosis: module 5: management of tuberculosis in children and adolescents [WWW Document]. URL <https://www.who.int/publications-detail-redirect/9789240046764> (accessed 5.1.24).
- World Health Organization, 2020. WHO consolidated guidelines on tuberculosis: module 1: prevention: tuberculosis preventive treatment [WWW Document]. URL <https://www.who.int/publications-detail-redirect/9789240001503> (accessed 5.25.24).
- World Health Organization, 2017a. Consensus meeting report: development of a target product profile (TPP) and a framework for evaluation for a test for predicting progression from tuberculosis infection to active disease.
- World Health Organization, 2017b. Guidelines for treatment of drug-susceptible tuberculosis and patient care, 2017 update. ed. World Health Organization, Geneva.
- World Health Organization, 2016. Chest radiography in tuberculosis detection: summary of current WHO recommendations and guidance on programmatic approaches. World Health Organization, Geneva.
- World Health Organization, 2014. High priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting, 28-29 April 2014, Geneva, Switzerland. World Health Organization, Geneva.
- World Health Organization, 2011a. Fluorescent light-emitting diode (LED) microscopy for diagnosis of tuberculosis: policy statement 16.
- World Health Organization, 2011b. Commercial serodiagnostic tests for diagnosis of tuberculosis: policy statement (No. WHO/HTM/TB/2011.5). World Health Organization.
- World Health Organization, W., 2023. Global Tuberculosis Report 2023 [WWW Document]. WHO Glob. Tuberc. Rep. 2023. URL <https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2023> (accessed 3.26.24).
- Wu, B., Huang, C., Kato-Maeda, M., Hopewell, P.C., Daley, C.L., Krensky, A.M., Clayberger, C., 2007. Messenger RNA Expression of IL-8, FOXP3, and IL-12 β Differentiates Latent Tuberculosis Infection from Disease. *J. Immunol.* 178, 3688–3694. <https://doi.org/10.4049/jimmunol.178.6.3688>
- Wu, X., Tan, G., Ma, J., Yang, J., Guo, Y., Lu, H., Ke, H., Li, M., Tang, Y.-W., Sha, W., Yu, F., 2023. Assessment of the Cepheid 3-gene Host Response Fingerstick Blood Test (MTB-HR) on rapid diagnosis of tuberculosis. *Emerg. Microbes Infect.* 12, 2261561. <https://doi.org/10.1080/22221751.2023.2261561>
- Yang, B., Zhai, F., Jiang, J., Wang, X., Cao, Z., Cheng, X., 2015. Elevated expression of T-bet in mycobacterial antigen-specific CD4⁺ T cells from patients with tuberculosis. *Cell. Immunol.* 298, 1–8. <https://doi.org/10.1016/j.cellimm.2015.08.002>
- Yoon, C., Chaisson, L.H., Patel, S.M., Allen, I.E., Drain, P.K., Wilson, D., Cattamanchi, A., 2017. Diagnostic accuracy of C-reactive protein for active pulmonary tuberculosis: a meta-analysis. *Int. J. Tuberc. Lung Dis.* 21, 1013–1019. <https://doi.org/10.5588/ijtld.17.0078>
- Yoon, Christina, Semitala, F.C., Atuhumuza, E., Katende, J., Mwebe, S., Asege, L., Armstrong, D.T., Andama, A.O., Dowdy, D.W., Davis, J.L., Huang, L., Kamya, M., Cattamanchi, A., 2017. Point-of-care C-reactive protein-based tuberculosis screening for people living with HIV: a

- diagnostic accuracy study. *Lancet Infect. Dis.* 17, 1285–1292. [https://doi.org/10.1016/S1473-3099\(17\)30488-7](https://doi.org/10.1016/S1473-3099(17)30488-7)
- Zak, D.E., Penn-Nicholson, A., Scriba, T.J., Thompson, E., Suliman, S., Amon, L.M., Mahomed, H., Erasmus, M., Whatney, W., Hussey, G.D., Abrahams, D., Kafaar, F., Hawkridge, T., Verver, S., Hughes, E.J., Ota, M., Sutherland, J., Howe, R., Dockrell, H.M., Boom, W.H., Thiel, B., Ottenhoff, T.H.M., Mayanja-Kizza, H., Crampin, A.C., Downing, K., Hatherill, M., Valvo, J., Shankar, S., Parida, S.K., Kaufmann, S.H.E., Walzl, G., Aderem, A., Hanekom, W.A., 2016. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. *The Lancet* 387, 2312–2322. [https://doi.org/10.1016/S0140-6736\(15\)01316-1](https://doi.org/10.1016/S0140-6736(15)01316-1)
- Zhang, M., Xue, M., He, J., 2020. Diagnostic accuracy of the new Xpert MTB/RIF Ultra for tuberculosis disease: A preliminary systematic review and meta-analysis. *Int. J. Infect. Dis.* 90, 35–45. <https://doi.org/10.1016/j.ijid.2019.09.016>
- Zhao, S., Fung-Leung, W.-P., Bittner, A., Ngo, K., Liu, X., 2014. Comparison of RNA-Seq and Microarray in Transcriptome Profiling of Activated T Cells. *PLOS ONE* 9, e78644. <https://doi.org/10.1371/journal.pone.0078644>
- Zhou, D., 2012. Quantum dot–nucleic acid/aptamer bioconjugate-based fluorimetric biosensors. *Biochem. Soc. Trans.* 40, 635–639. <https://doi.org/10.1042/BST20120059>
- Zhou, W., Yan, Y., Guo, Q., Ji, H., Wang, H., Xu, T., Makabel, B., Pilarsky, C., He, G., Yu, X., Zhang, J., 2021. Microfluidics applications for high-throughput single cell sequencing. *J. Nanobiotechnology* 19, 312. <https://doi.org/10.1186/s12951-021-01045-6>
- Zimmer, A.J., Lainati, F., Aguilera Vasquez, N., Chedid, C., McGrath, S., Benedetti, A., MacLean, E., Ruhwald, M., Denking, C.M., Kohli, M., 2022. Biomarkers That Correlate with Active Pulmonary Tuberculosis Treatment Response: a Systematic Review and Meta-analysis. *J. Clin. Microbiol.* 60, e01859-21. <https://doi.org/10.1128/jcm.01859-21>
- Zimmer, A.J., Schumacher, S.G., Södersten, E., Mantsoki, A., Wyss, R., Persing, D.H., Banderby, S., Strömqvist Meuzelaar, L., Prieto, J., Gnanashanmugam, D., Khatri, P., Ongarello, S., Ruhwald, M., Denking, C.M., 2021. A novel blood-based assay for treatment monitoring of tuberculosis. *BMC Res. Notes* 14, 247. <https://doi.org/10.1186/s13104-021-05663-z>

Chapter 2: Materials and methods

This chapter aims to describe the methods used across the different aims in this thesis. Details of each human participant cohort, statistical analyses and methods specific for each chapter are included in the results chapters.

2.1 Institutional ethics review

The transcriptomic work detailed in Chapters 4 and 5 of this thesis is covered by the PhD study protocol reviewed and approved by the Human Research Ethics Committee (HREC) of the University of Cape Town (UCT) Faculty of Health Sciences (HREC reference number 522/2020). The respective clinical studies (**Table 2.1**) under which study participants were enrolled were reviewed and approved by ethics committees at other institutions: Stellenbosch University, Makerere University in Uganda, Medical Research Council Unit The Gambia, Karonga Prevention Study in Malawi, the University of Namibia, Armauer Hansen Research Institute in Ethiopia, the Ethiopian Public Health Institute, Desmond Tutu TB Centre in South Africa, Imperial College London, and University College London.

The study protocol for the transcriptomic work carried out on the CTBC diagnostic study in Chapter 3 was reviewed and approved by the UCT HREC (HREC reference number 761/2015).

Further details concerning the design and associated ethical considerations and review committees for each individual cohort (**Table 2.1**) are described in the respective results chapters.

Table 2.1: Participant cohorts analysed in this thesis

Cohort	Cohort design	Sample type analysed	Specific results chapter
Cross-sectional TB Cohort (CTBC) (Darboe et al., 2019)	Active TB cases vs asymptomatic/ <i>Mtb</i> sensitised controls	Whole blood	Chapter 3
Evaluation of host biomarker-based point-of-care tests for targeted screening for active TB (ScreenTB) (Manngo et al., 2019)	Symptomatic TB vs other respiratory diseases (ORDs)	Whole blood	Chapter 4
African European Tuberculosis Consortium (AE-TBC)(Chegou et al., 2018, 2016)	Symptomatic TB vs other respiratory diseases (ORDs)	Whole blood	Chapter 4
Shorter Treatment for Nonsevere Tuberculosis in African and Indian Children (SHINE) (Chabala et al., 2018; Turkova et al., 2022)	Non-severe paediatric TB cases treated for either 4 or 6 months	Whole blood	Chapter 5

2.2 Collection of whole blood samples

PAXgene and Tempus blood tubes are currently the two main collection methods used in studies aimed at gene expression analysis based on *ex vivo* whole blood. Both PAXgene and Tempus tubes contain a stabilizing solution that prevents RNA degradation. RNA preservation in PAXgene tubes is

noted to be relatively consistent resulting in RNA expression signatures obtained from PAXgene tubes to be more consistent and replicable across various samples and conditions (Skogholt et al., 2017).

An assessment of gene expression profiles performed on peripheral blood collected from both blood collection tubes concluded that results were highly comparable in terms of RNA quality (Menke et al., 2012). This was in line with 3 other studies (Asare et al., 2008; Matheson et al., 2008; Skogholt et al., 2017), with one of them involving samples being shipped at different temperatures (Matheson et al., 2008). Menke et al. also reports no difference in the detection rates of transcripts using both blood collection methods (Menke et al., 2012). Matheson et al. and Asare et al. do report on an effect of these blood collection methods on gene expression results with Matheson and colleagues further emphasizing the need for the use of only one blood collection method throughout a study to avoid inconsistencies during RNA extraction (Matheson et al., 2008). Donahue and colleagues also made a similar recommendation from a study comparing RNAgard and PAXgene collection methods for gene expression profiling of whole blood (Donahue et al., 2019).

Dried blood spots (DBS) have emerged as another alternative for blood collection as they are cheaper compared to PAXgene and Tempus tubes, do not need trained personnel and are far less invasive since blood can be obtained from a finger prick. The number of analytes that can be extracted from DBS is small leading to this approach being debated for its use in generating data for genomic analyses (Agrawal et al., 2021). Data however is emerging to determine how suitable DBS are for next-generation and whole genome sequencing. Reust and colleagues, for example, found a strong correlation between transcriptomes from next-generation RNA sequencing, derived from RNA isolated from DBS and that from whole blood (Reust et al., 2018).

Ultimately, the choice of the blood collection method depends on factors such as the setting in which research is being conducted, the intended use of the samples, downstream analysis and the length of long-time storage. For all the projects highlighted in this thesis, all whole blood samples were collected using PAXgene tubes. This system was maintained since all optimization and validation procedures, as well as downstream analyses were all developed and tested on blood collected in these tubes. To ensure consistency and reproducibility of results, we therefore decided to proceed with PAXgene tubes for blood collection.

2.3 RNA extraction, processing and storage

The aim of RNA extraction was to obtain high quality purified RNA from whole blood. RNA extraction was therefore a very crucial step for all RNA-based techniques described in this chapter. This is because it formed the basis for the successful real-time quantitative PCR (RT-qPCR) experiments and downstream gene expression analysis described and carried out in Chapters 3, 4, and 5 (Samadani et al., 2015; Shi and Bressan, 2006).

RNA extraction from 2.5ml whole blood collected in PAXgene blood RNA tubes (Qiagen, Hilden, Germany), and subsequent cDNA synthesis, was carried out in a semi-automated manner using the Tecan Freedom EVO 150 liquid handling and robotics workstation (Tecan Group Ltd, Männedorf, Switzerland). All work downstream, up to the point of performing multiplex RT-qPCR, involved processing 93 samples on each 96.96 gene expression integrated fluidic circuits (IFC) nanofluidic chip (Standard BioTools Inc, South San Francisco, California, United States) (**Section 2.3.4 of Chapter 2**).

Frozen PAXgene tubes were left to thaw overnight at room temperature before RNA extraction for higher RNA yields. PAXgene tubes were centrifuged twice in the Allegra X-12R (Beckman Coulter Inc, Brea, California, United States) each at 3,000g for 10 minutes at room temperature. At the end of each centrifugation, the supernatant was removed and discarded and the PAXgene tubes patted dry, at least

three times on a paper towel to remove any residual supernatant. After the first centrifugation, the tubes were vortexed and whole blood resuspended in the remaining liquid and an additional 5ml 1X PBS (R & D Systems by Bio-Techne Corporation, Minneapolis, Minnesota, United States). After the second centrifugation, the pellet in each PAXgene tube was completely resuspended for 30 seconds in 200µl of a pre-mixed homogenization solution containing 1-Thioglycerol (Promega Corporation, Madison, Wisconsin, United States).

All steps from this point on were carried out on the Tecan Freedom EVO liquid handler instrument and were therefore fully automated by executing an RNA extraction programme designed and set up on the Freedom EVOware software (Tecan Group Ltd, Männedorf, Switzerland). RNA extraction was done with the Maxwell® HT simplyRNA Custom Kit (Promega Corporation, Madison, Wisconsin, United States).

The Tecan instrument aliquoted RNA into three daughter plates with 24µl of RNA per FluidX tube (Azenta, Massachusetts, United States) in 96-tube rack for long term storage, and 6µl per well aliquoted into a 96-well barcoded PCR plate (Applied Biosystems by Thermo Fisher Scientific, Waltham, Massachusetts, United States) for immediate cDNA synthesis. The well for the No Template Control (NTC) contained 6µl of molecular-grade water.

2.4 Measurement of gene expression

2.4.1 cDNA synthesis

Complementary DNA (cDNA) synthesis involves the use of an RNA template to generate cDNA through the process of reverse transcription. To detect mRNA during PCR experiments, cDNA synthesis has to occur. One of the advantages that DNA has over RNA is that it is a more stable molecule making it better suited for long term storage of genetic information. The combination of cDNA synthesis and RT-qPCR also aids in detection of low abundance mRNAs in a sample. Ultimately, cDNA led to the availability of more template, enabling more copies of all sequences of interest to be available for analysis in this thesis.

With the use of EpiScript RNase H– Reverse Transcriptase (EpiScript RT, Lucigen, Hoddesdon, United Kingdom), cDNA was synthesised using the extracted RNA as template. A master mix containing 143µl of Oligo(dT) (0.5 µg/µl stock) (EpiScript RT, Lucigen, Hoddesdon, United Kingdom) and 430µl of molecular grade water (Invitrogen by Thermo Fisher Scientific, Waltham, Massachusetts, United States) was prepared and added automatically by running a “RNA Aliquoting and cDNA synthesis” script on the Tecan Freedom EVO system. Four microlitres of master mix were transferred into to each well of the 96-well barcoded plate that contained 6µl of RNA.

Plates were sealed with Polyolefin PCR plate seal (Starlab Ltd, Milton Keynes, United Kingdom), briefly centrifuged, and moved to a T100 PCR thermocycler (Bio-Rad Laboratories, Hercules, California, United States) with a heating step set at 65°C for 2 minutes followed by a 4°C hold. Apart from the well for the No Reverse Transcriptase (NoRT) control that did contain any RT reagents, 10µl containing a mix of dNTP (EpiScript RT, Lucigen, Hoddesdon, United Kingdom), EpiScript RT, EpiScript RT 10x buffer, 100 mM DTT, RiboGuard (EpiScript RT, Lucigen, Hoddesdon, United Kingdom), and molecular grade water was added to every other well on the plate.

Plates were sealed with domed PCR caps (Thermo Fisher Scientific, Waltham, Massachusetts, United States), centrifuged briefly, and moved to the thermocycler where samples were incubated at 37°C for 60 minutes, heat inactivated at 85°C for 5 minutes, and thereafter held at 4°C. The plates were centrifuged briefly and 80µl of molecular grade water added to each well. Each plate was then labelled

as '1st strand cDNA' along with the processing date, plate number, and operator initials. The plates were used immediately for pre-amplification of genes of interest or stored for later use in a -40°C freezer.

2.4.2 Pre-amplification PCR

High-throughput qPCR instruments, such as the BioMark HD System from Standard BioTools (Standard BioTools Inc, South San Francisco, California, United States), require samples that have a high concentration of genetic material (Korenková et al., 2015). With up to 96 microfluidic PCR reaction chambers present on the 96.96 dynamic array gene expression chips (Standard BioTools Inc, South San Francisco, California, United States), 4096 amplicons, resulting from at least 12 pre-amplification cycles, are required to be loaded on the chips. Pre-amplification is a multiplexed PCR reaction carried out for a specified number of cycles. This ensures that enough cDNA is evenly distributed across all 96 microfluidic PCR reaction chambers (Dominguez et al., 2013).

Primer-probe master mixes (PPMMs) were comprised of a pool of all 96 primer-probe assays (Thermo Fisher Scientific, Waltham, Massachusetts, United States) on the panel. Each primer-probe assay was purchased at 20x concentration and diluted in water to a concentration of 0.2x in the PPMM; 2.3µM. In a Falcon tube, 40µl of molecular grade water was added alongside 10µl of each primer-probe assay to achieve a total volume of 1000µl. PPMM aliquots of 275µl each were made.

The format of 3 controls per plate (i.e. NTC, NoRT and internal positive control [IPC]), from was maintained from cDNA synthesis, pre-amplification, and multiplex RT-qPCR carried out on the Standard BioTools BioMark HD platform.

On an MicroAmp optical 96-well PCR plates (Applied Biosystems by Thermo Fisher Scientific, Waltham, Massachusetts, United States), 5µl of DreamTaq PCR Master Mix (2x) (Thermo Fisher Scientific, Waltham, Massachusetts, United States) and 2.5µl of PPMM (0.2x) were added to each well, followed by 2.5µl of cDNA. This brought the total volume for each pre-amplification reaction, per well, to 10µl. The 96-well PCR plate was then sealed with domed PCR caps, centrifuged briefly, and placed in a T100 PCR thermocycler where genes of interest were pre-amplified for 16 cycles. The conditions on the thermocycler for pre-amplification were set as follows: 1) 95°C for 10 minutes, 2) 95°C for 15 seconds, and 3) 60°C for 4 minutes. Steps two and three were repeated sixteen times before the reaction was put on hold at 4°C. The samples were then diluted 1:25 by adding 240µl of molecular grade water to each well. From this point on, these samples are referred to as pre-amplified cDNA.

2.4.3 Confirmatory PCR for quality control of the pre-amplification process

To ensure that the preamplification process was successful and that no contamination was present, a confirmatory RT-qPCR assay was carried out on a QuantStudio 3 Real-Time PCR System (Applied Biosystems by Thermo Fisher Scientific, Waltham, Massachusetts, United States) using one of the primer-probe assays in the panel. The assay selected would be one known to be highly expressed and therefore easily detected within most samples under investigation.

The confirmatory PCR assay was done on 3 samples, each in duplicate. There were two pre-amplified cDNA samples, each from different columns of the 96-well plate to test that pre-amplification occurred evenly across the plate. To check for contamination, the NTC was the third sample tested. For all 6 reactions, a master mix comprised of 6µl of one primer-probe assay at its original 20x concentration, 30µl of molecular grade water and 60µl of Taqman Universal PCR Master Mix (Applied Biosystems by Thermo Fisher Scientific, Waltham, Massachusetts, United States) was made. In each PCR reaction tube, 16µl of the master mix was added followed by 4µl of pre-amplified cDNA.

The setup of the RT-qPCR on the QuantStudio 3 machine was as follows: 50°C for 2 minutes, 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute.

Observation of no amplification in the NTC replicates confirmed that no contamination was present. If any fluorescence was detected, the pre-amplification step would have to be repeated and another confirmatory PCR done. If fluorescence was detected within the NTC, it meant that contamination most likely occurred during cDNA synthesis. In cases where contamination was detected, the entire process from cDNA synthesis was repeated before multiplex RT-qPCR was performed on the BioMark instrument.

2.4.4 High-throughput Gene Expression Measurement with RT-qPCR

The Biomark HD system was used to detect, measure, and quantify the expression of all 96 transcripts of interest. It is an automated system for high throughput RT-qPCR that uses microfluidic technology to determine gene expression. Using 96.96 gene expression IFC nanofluidic chips (**Figure 2.1**), 93 unique samples and 3 different controls, present in singlets, were evaluated per processing run to quantify 96 mRNA transcripts using different primer-probe assays.

Individual 96-well plates, containing pre-amplified cDNA and primer-probe assays, were thawed at room temperature. For the gene expression IFCs to be prepared for sample transfer, they had to be primed first. This was done by injecting control fluid into each accumulator (**Figure 2.1**). The IFC was then placed into the IFC controller HX instrument (Standard BioTools Inc, South San Francisco, California, United States) with a chip prime program set for 20 minutes.

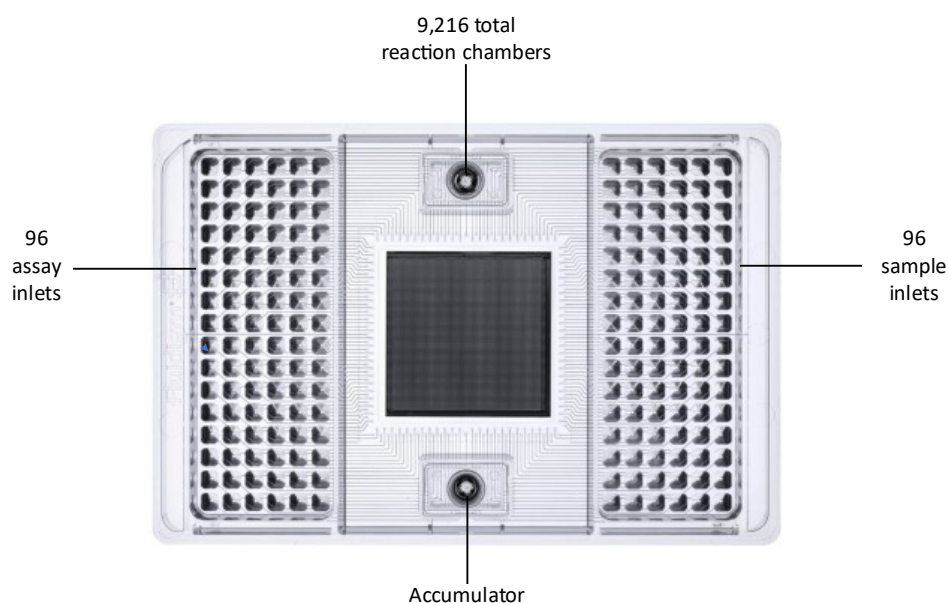


Figure 2.1: BioMark™ 96.96 Dynamic Array IFC for gene expression.

The IFC allows a total of 9,216 reactions to take place by making use of all 96 primer probe assays and samples.

After the 96-well skirted PCR plate (Azenta Life Sciences, Dorking, United Kingdom) containing all 96 assays thawed, it was centrifuged and 3.5µl of assay loading reagent (Standard BioTools Inc, South San Francisco, California, United States) pipetted into each well on the plate. The plate was sealed with

Polyolefin PCR plate seal, centrifuged briefly, and stored at 4°C until the assays were pipetted into the IFC inlets on the chip.

A sample pre-mix was made for the entire 96-well pre-amped cDNA plate. This contained 360µl of 2x TaqMan Universal PCR Master Mix (Applied Biosystems by Thermo Fisher Scientific, Waltham, Massachusetts, United States) and 36µl of 20x gene expression sample loading reagent (Standard BioTools Inc, South San Francisco, California, United States). A volume of 3.3µl of sample pre-mix was added to each well of a fresh, empty MicroAmp optical 96-well PCR plate. 2.7µl of the IPC was pipetted into the well designated for it. Pre-amplified cDNA, including the NTC and NoRT, at a volume of 2.7µl, were added to the remainder of the wells on the PCR plate.

Samples and assays were pipetted into the designated inlets on the IFC (**Figure 2.1**), each at a volume of 5µl with the assays loaded on the left side inlets and samples on the right side of the IFC. To prepare and combine all 9,216 reactions, the chip was loaded into an IFC controller HX instrument with a pre-set program that runs for 90 minutes. Once completed, the chip was removed from the IFC controller and the chip surface cleaned with scotch tape to get rid of any dust particles.

A new run was set up per chip using the BioMark Data Collection software. The chip was placed into the Biomark HD instrument. The application settings for the chip runs were entered as follows: Application type was set to “gene expression”, ROX (carboxyrhodamine) was chosen as the passive reference dye for signal normalisation during qPCR, and FAM-TAMRA selected as the fluorescent primer probe. In this instance we used single probes per assay. One of the pre-determined thermal-cycling protocols provided by the manufacturer, GE 96x96 Standard v1.pcl, was selected for real-time PCR on the BioMark HD instrument.

After each run was completed, heat maps (**Figure 2.2**) with gene expression levels represented by different colours, were viewed using Real-Time PCR Analysis Software (Standard BioTools Inc, South San Francisco, California, United States) to visually confirm a successful run. Analysis settings were set up as follows: quality threshold of 0.3 for each amplification curve, baseline correction done using the “linear (derivative)” method, and cycle thresholds (Ct) for the entire chip calculated using the “auto (global)” method.

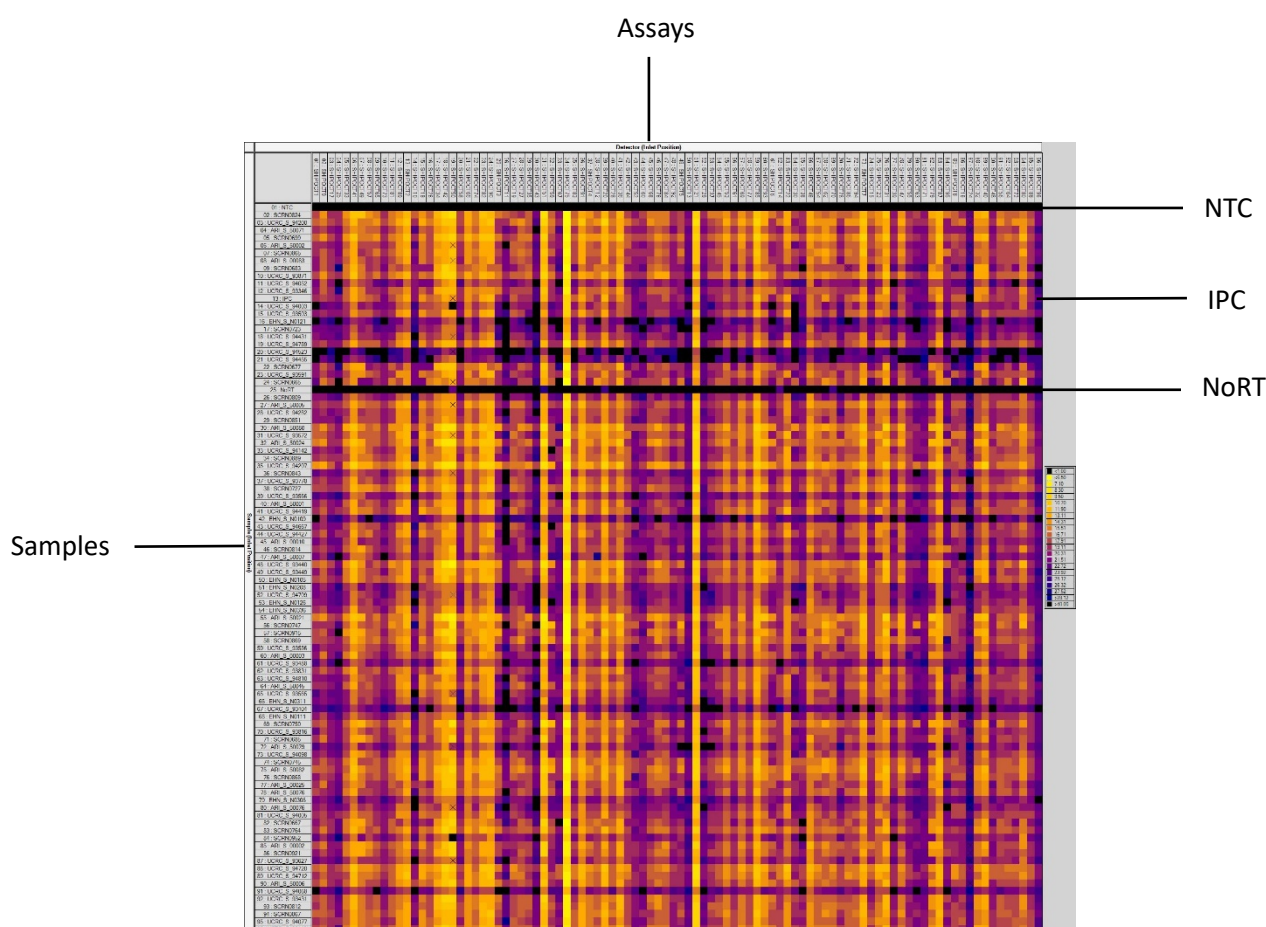


Figure 2.2: Heatmap of gene expression.

non-template control (NTC), no reverse transcriptase (NoRT) and internal positive control (IPC) shown.

2.5 Quality control (QC) procedures for gene expression data

QC criteria (**Figure 2.3**), implemented using a script written in R, was applied to each set of results after every chip run. This was done to ensure that each run adhered to a number of criteria before signature scores were calculated and further data analyses performed.

The QC script evaluated different parameters that are described below. Deviations from the thresholds set for different parameters (**Figure 2.3**) on the QC script determined whether any chips needed be repeated to attain results that were suitable for analysis. The script verified whether the gene expression settings for RT-qPCR on the BioMark instrument were the same for every chip, to ensure consistency in the output generated.

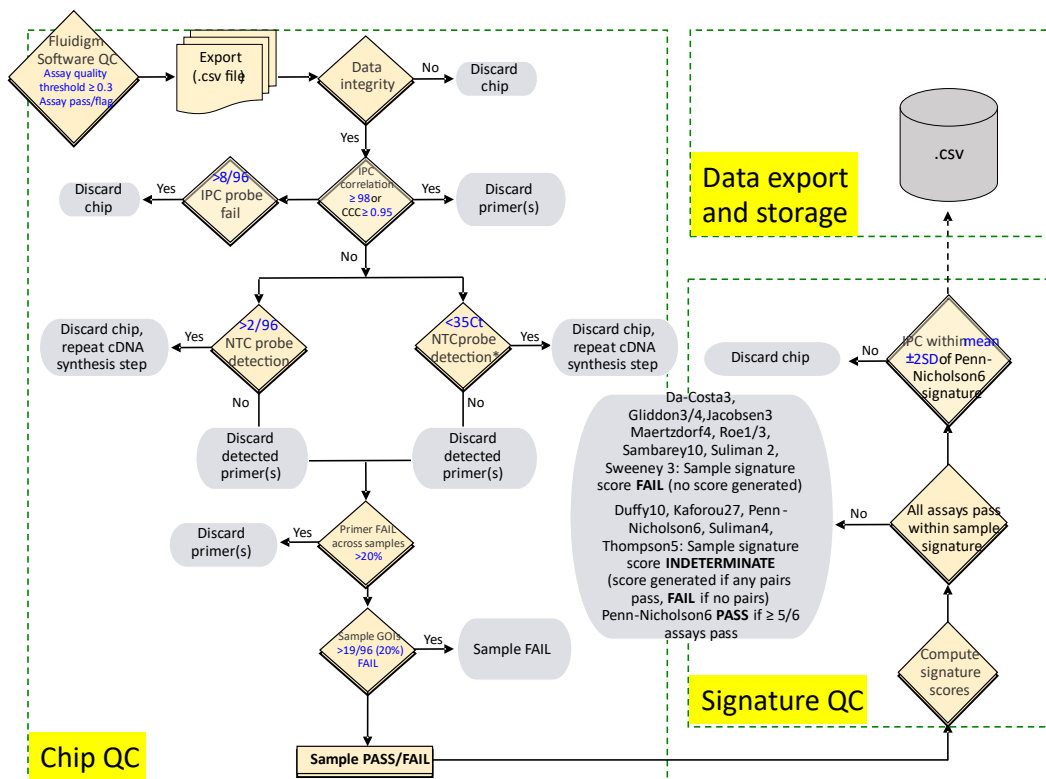


Figure 2.3: QC Workflow.

The graphical layout represents the QC steps applied using code written in the R programming language.

The QC script assessed the presence of the NTC on each chip and if it was labelled correctly. Amplification within the NTC was checked to ensure that no contamination was present. The NTC was deemed as failed if amplification, for more than 2 assays within it, was detected at Ct values below 35. A chip with a failed NTC was excluded from analysis. If amplification for less than 2 of the 96 assays within the NTC was detected, but at a Ct value less than 35, only that primer-probe assay would be excluded from further analysis.

Within the QC script, thresholds were also set for the number of failed reactions (no amplification) for samples and assays on the chip. This was used to determine the quality of each reaction and, in cases of multiple failures, whether the entire chip failed, or if there was an issue with certain assays or samples on that chip. Reactions failed if quality Ct values could not be generated. This usually resulted from irregular amplification detected within the samples and assays on a chip during the RT-qPCR process.

If more than 20% of assays failed within a sample, that particular sample failed QC. If more than 20% of the samples on a chip failed, then the entire chip failed. If more than 30% of samples failed for an assay, that particular assay failed and assay-specific results for all samples were then excluded from any signature score calculations for that chip. If more than 10% of the primer-probe assays failed on a chip, then the entire chip failed, and all signature scores invalid for further analysis.

Criteria for the IPC to have fully passed QC included (1) that Ct values for the primer probe assays within the IPC sample were within ± 4 standard deviations of the range of Ct values for IPCs from historical chip runs. Not more than 8 assays were allowed to be outside of this range; (2) Ct values for the primer probes in the IPC sample had to have a correlation of no less than 0.98 or have a concordance correlation coefficient no less than 0.95 from the range of Ct values for IPCs from previous runs.

The IPC was prepared based on samples with RISK6 scores (Penn-Nicholson et al., 2020a) in a range of 40 to 60 (**Section 3.4 of Chapter 3**). RISK6 scores were therefore also evaluated during QC to determine whether the IPC passed. The IPC on the chip passed if it had a RISK6 score within $\pm 2SD$ of the mean RISK6 scores from historical chip runs.

The NoRT control was not consistently present on all chips used to process samples from the ScreenTB and AE-TBC cohorts (**Chapter 4**). The NoRT control was not evaluated as part of the QC criterion. Evaluation of the NTC, that was consistently present on all chips, for all study cohorts, was included in the QC script pipeline. The NTC was deemed sufficient enough to determine the presence of contamination resulting from the processing of whole blood samples.

2.6 Contributions

I wrote the PhD study proposal, designed the experiments as well as wrote this chapter under the supervision of Professor Thomas J. Scriba and Dr Simon C. Mendelsohn. All RNA extraction and cDNA synthesis were performed by Mzwandile Erasmus, Rodney Raphela and Lungisa Jaxa. All further gene expression work, up to the point of QC was performed by me.

References

- Asare, A.L., Kolchinsky, S.A., Gao, Z., Wang, R., Raddassi, K., Bourcier, K., Seyfert-Margolis, V., 2008. Differential gene expression profiles are dependent upon method of peripheral blood collection and RNA isolation. *BMC Genomics* 9, 474.
<https://doi.org/10.1186/1471-2164-9-474>
- Chabala, C., Turkova, A., Thomason, M.J., Wobudeya, E., Hissar, S., Mave, V., van der Zalm, M., Palmer, M., Kapasa, M., Bhavani, P.K., Balaji, S., Raichur, P.A., Demers, A.-M., Hoddinott, G., Owen-Powell, E., Kinikar, A., Musoke, P., Mulenga, V., Aarnoutse, R., McIlleron, H., Hesselting, A., Crook, A.M., Cotton, M., Gibb, D.M., SHINE trial team, 2018. Shorter treatment for minimal tuberculosis (TB) in children (SHINE): a study protocol for a randomised controlled trial. *Trials* 19, 237. <https://doi.org/10.1186/s13063-018-2608-5>
- Chegou, N.N., Sutherland, J.S., Malherbe, S., Crampin, A.C., Corstjens, P.L.A.M., Geluk, A., Mayanja-Kizza, H., Loxton, A.G., Spuy, G. van der, Stanley, K., Kotzé, L.A., Vyver, M. van der, Rosenkrands, I., Kidd, M., Helden, P.D. van, Dockrell, H.M., Ottenhoff, T.H.M., Kaufmann, S.H.E., Walzl, G., 2016. Diagnostic performance of a seven-marker serum protein biosignature for the diagnosis of active TB disease in African primary healthcare clinic attendees with signs and symptoms suggestive of TB. *Thorax* 71, 785–794.
<https://doi.org/10.1136/thoraxjnl-2015-207999>
- Chegou, N.N., Sutherland, J.S., Namuganga, A.-R., Corstjens, P.L., Geluk, A., Gebremichael, G., Mendy, J., Malherbe, S., Stanley, K., van der Spuy, G.D., Kriel, M., Loxton, A.G., Kriel, B., Simukonda, F., Bekele, Y., Sheehama, J.A., Nelongo, J., van der Vyver, M., Gebrexabher, A., Hailu, H., Esterhuyse, M.M., Rosenkrands, I., Aagard, C., Kidd, M., Kassa, D., Mihret, A., Howe, R., Cliff, J.M., Crampin, A.C., Mayanja-Kizza, H., Kaufmann, S.H.E., Dockrell, H.M., Ottenhoff, T.H.M., Walzl, G., 2018. Africa-wide evaluation of host biomarkers in QuantiFERON supernatants for the diagnosis of pulmonary tuberculosis. *Sci. Rep.* 8, 2675.
<https://doi.org/10.1038/s41598-018-20855-7>
- Darboe, F., Mbandi, S.K., Naidoo, K., Yende-Zuma, N., Lewis, L., Thompson, E.G., Duffy, F.J., Fisher, M., Filander, E., van Rooyen, M., Bilek, N., Mabwe, S., McKinnon, L.R., Chegou, N., Loxton, A., Walzl, G., Tromp, G., Padayatchi, N., Govender, D., Hatherill, M., Karim, S.A., Zak, D.E., Penn-Nicholson, A., Scriba, T.J., Team, T.S.C.I., 2019. Detection of Tuberculosis Recurrence, Diagnosis and Treatment Response by a Blood Transcriptomic Risk Signature in HIV-Infected Persons on Antiretroviral Therapy. *Front. Microbiol.* 10.
<https://doi.org/10.3389/fmicb.2019.01441>
- Dewandel, I., van Niekerk, M., Ghimenton-Walters, E., Palmer, M., Anthony, M.G., McKenzie, C., Croucamp, R., Alter, G., Demers, A.-M., van Zyl, G., Claassen, M., Goussard, P., Swanepoel, R., Hoddinott, G., Bosch, C., Dunbar, R., Allwood, B., McCollum, E.D., Schaaf, H.S., Hesselting, A.C., van der Zalm, M.M., 2023. UMOYA: a prospective longitudinal cohort study to evaluate novel diagnostic tools and to assess long-term impact on lung health in South African children with presumptive pulmonary TB—a study protocol. *BMC Pulm. Med.* 23, 97.
<https://doi.org/10.1186/s12890-023-02329-3>
- Dominguez, M.H., Chattopadhyay, P.K., Ma, S., Lamoreaux, L., McDavid, A., Finak, G., Gottardo, R., Koup, R.A., Roederer, M., 2013. Highly multiplexed quantitation of gene expression on single cells. *J. Immunol. Methods* 391, 10.1016/j.jim.2013.03.002.
<https://doi.org/10.1016/j.jim.2013.03.002>
- Donohue, D.E., Gautam, A., Miller, S.-A., Srinivasan, S., Abu-Amara, D., Campbell, R., Marmar, C.R., Hammamieh, R., Jett, M., 2019. Gene expression profiling of whole blood: A comparative assessment of RNA-stabilizing collection methods. *PLoS ONE* 14, e0223065.
<https://doi.org/10.1371/journal.pone.0223065>

- Korenková, V., Scott, J., Novosadová, V., Jindřichová, M., Langerová, L., Švec, D., Šídová, M., Sjöback, R., 2015. Pre-amplification in the context of high-throughput qPCR gene expression experiment. *BMC Mol. Biol.* 16, 5. <https://doi.org/10.1186/s12867-015-0033-9>
- Matheson, L.A., Duong, T.T., Rosenberg, A.M., Yeung, R.S.M., 2008. Assessment of sample collection and storage methods for multicenter immunologic research in children. *J. Immunol. Methods* 339, 82–89. <https://doi.org/10.1016/j.jim.2008.08.003>
- Menke, A., Rex-Haffner, M., Klengel, T., Binder, E.B., Mehta, D., 2012. Peripheral blood gene expression: it all boils down to the RNA collection tubes. *BMC Res. Notes* 5, 1. <https://doi.org/10.1186/1756-0500-5-1>
- Manngo, P.M., Gutschmidt, A., Snyders, C.I., Mutavhatsindi, H., Manyelo, C.M., Makhoba, N.S., Ahlers, P., Hiemstra, A., Stanley, K., McAnda, S., Kidd, M., Malherbe, S.T., Walzl, G., Chegou, N.N., 2019. Prospective evaluation of host biomarkers other than interferon gamma in QuantiFERON Plus supernatants as candidates for the diagnosis of tuberculosis in symptomatic individuals. *J. Infect.* 79, 228–235. <https://doi.org/10.1016/j.jinf.2019.07.007>
- Penn-Nicholson, A., Mbandi, S.K., Thompson, E., Mendelsohn, S.C., Suliman, S., Chegou, N.N., Malherbe, S.T., Darboe, F., Erasmus, M., Hanekom, W.A., Bilek, N., Fisher, M., Kaufmann, S.H.E., Winter, J., Murphy, M., Wood, R., Morrow, C., Van Rhijn, I., Moody, B., Murray, M., Andrade, B.B., Sterling, T.R., Sutherland, J., Naidoo, K., Padayatchi, N., Walzl, G., Hatherill, M., Zak, D., Scriba, T.J., 2020. RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response. *Sci. Rep.* 10, 8629. <https://doi.org/10.1038/s41598-020-65043-8>
- Samadani, A.A., Nikbakhsh, N., Fattahi, S., Pourbagher, R., Aghajanpour Mir, S.M., Mousavi Kani, N., Abedian, Z., Akhavan-Niaki, H., 2015. RNA Extraction from Animal and Human's Cancerous Tissues: Does Tissue Matter? *Int. J. Mol. Cell. Med.* 4, 54–59.
- Shi, H., Bressan, R., 2006. RNA Extraction, in: Salinas, J., Sanchez-Serrano, J.J. (Eds.), *Arabidopsis Protocols, Methods in Molecular Biology™*. Humana Press, Totowa, NJ, pp. 345–348. <https://doi.org/10.1385/1-59745-003-0:345>
- Skogholt, A.H., Ryeng, E., Erlandsen, S.E., Skorpen, F., Schønberg, S.A., Sætrom, P., 2017. Gene expression differences between PAXgene and Tempus blood RNA tubes are highly reproducible between independent samples and biobanks. *BMC Res. Notes* 10, 136. <https://doi.org/10.1186/s13104-017-2455-6>
- Turkova, A., Wills, G.H., Wobudeya, E., Chabala, C., Palmer, M., Kinikar, A., Hissar, S., Choo, L., Musoke, P., Mulenga, V., Mave, V., Joseph, B., LeBeau, K., Thomason, M.J., Mboizi, R.B., Kapasa, M., van der Zalm, M.M., Raichur, P., Bhavani, P.K., McIlleron, H., Demers, A.-M., Aarnoutse, R., Love-Koh, J., Seddon, J.A., Welch, S.B., Graham, S.M., Hesselning, A.C., Gibb, D.M., Crook, A.M., 2022. Shorter Treatment for Nonsevere Tuberculosis in African and Indian Children. *N. Engl. J. Med.* 386, 911–922. <https://doi.org/10.1056/NEJMoa2104535>

Chapter 3: Optimization of the panel of transcripts

3.1: Introduction

This chapter describes the steps taken to set up and optimise a panel of transcriptomic signatures, identify additional transcripts of interest for the panel, selection of primer-probe assays for each transcript, selection of housekeeper genes, and application of the statistical models used to develop each signature.

The optimization steps mentioned above were necessary to ascertain the validity and reproducibility of the results that would be generated from the multiplex RT-qPCR experiments carried out for the subsequent study cohorts described in chapters 4 and 5. The optimization steps were also performed to ensure that the precision of the primer-probe assays and normalization of gene expression using very stable reference (“housekeeper”) genes, all effectively contributed to the accuracy of RT-qPCR analysis (Zhao et al., 2021).

3.1.1 Selection of transcripts for the panel

Criteria to identify transcriptomic signatures were defined to fulfil the objectives and overall aims of the project. Altogether 20 published transcriptomic signatures (**Table 3.1**) were identified, with each one derived from discovery cohorts. The criteria used to select these signatures were as follows:

- 1) Study design used in the publication: Priority was given to signatures developed from studies that had symptomatic individuals with other respiratory diseases (ORDs) as the control group. This was because the primary study (**Chapter 4**), for which this panel was developed, aimed to compare active TB cases to people that presented with ORDs.
- 2) Purpose for development of the signature: Signatures were chosen if they had been developed to be diagnostic, prognostic, or to monitor patients’ responses to TB treatment.
- 3) Diagnostic performance in published literature: Transcriptomic signatures were selected based on their performance in systematic reviews (**Table 3.2**) where head-to-head comparisons of signatures had been made (Gupta et al., 2020; Mendelsohn et al., 2023; Mulenga et al., 2020; Turner et al., 2020; Warsinske et al., 2019), as well as in the original publications where these signatures had been discovered. Signatures underwent additional refinement guided by their performance in external validation cohort results.
- 4) Number of transcripts that comprised each signature: Signatures comprised of 10 transcripts or less were prioritized for inclusion on the panel. This enabled us to evaluate more signatures within the project. Since gene expression data was generated on 96.96 microfluidic chips (**Chapter 2, Section 2.3.4**), a maximum number 93 transcripts could be included on the panel. The three remaining spaces on the chip were taken up by the housekeeper genes (Section 3.1.3). The Kaforou signature would be the exception to this as it was decided from the inception of the project, to be included on the panel and was retained as a result of ongoing collaborative work with Dr. Mysrini Kaforou and her research group at Imperial College London.
- 5) Ability to reconstruct the signature: Each publication needed to provide sufficient information on gene names and IDs as well the primer probe assays used for each transcript. In some instances, authors were contacted for additional information.
- 6) Ability for the signature be adapted to RT-qPCR: The laboratory technique used for gene expression as well as the statistical and machine learning methods used to develop the signature were taken note of. This information helped ascertain that a signature was adaptable to the Fluidigm microfluidic RT-qPCR platform.

Table 3.1: Panel of transcriptomic signatures

Signature name	Signature model	Discovery population	Discovery country	Intended application
da Costa3 (da Costa et al., 2015)	Random forest	Adults	Brazil	Diagnostic; TB vs. ORD
de Araujo1 (de Araujo et al., 2016)	Standardised expression of NPC2	Adults	Brazil	Diagnostic; TB vs. LTBI and HC
Duffy10 (Duffy et al., 2019)	Random forest	Adults	Publicly available datasets from South Africa, and Malawi (Anderson et al., 2014; Kaforou et al., 2013)	Diagnostic; TB vs. LTBI and ORD
Francisco2 (Francisco et al., 2017)	Random forest	Adults	China	Diagnostic; TB vs. ORD and HC
GjØen7 (GjØen et al., 2017a)	LASSO regression	Children	India	Diagnostic; TB vs. ORD
Gliddon3 (Gliddon et al., 2021)	Disease risk score	Adults	South Africa, Malawi	Diagnostic; TB vs. LTBI
Gliddon4 (Gliddon et al., 2021)	Disease risk score	Adults	South Africa, Malawi	Diagnostic; TB vs. ORDs
Jacobsen3 (Jacobsen et al., 2007)	Linear discriminant analysis	Adults	Germany	Diagnostic; TB vs. HC
Kaforou25 [27]† (Kaforou et al., 2013)	Disease risk score	Adults	South Africa, Malawi	Diagnostic; TB vs. LTBI and ORD
Maertzdorf4 (Maertzdorf et al., 2016)	Pair ratio algorithm adopted from Suliman et al (Suliman et al., 2018)	Adults	India	Diagnostic; TB vs. LTBI and HC
Penn-Nicholson6 (Penn-Nicholson et al., 2020b)	Pair ratio algorithm	Adolescents	South Africa	Prognostic; TB progressors vs. non-progressors
Rajan5 (Rajan et al., 2019)	Unsigned sums	Adults	Uganda	Diagnostic; HIV+ TB vs. HIV+ HC
Roe1 (Roe et al., 2016)	Standardised expression of BATF2	Adults	United Kingdom	Diagnostic; TB vs HC and TB
Roe3 (Roe et al., 2019)	SVM (linear kernel)	Adults	United Kingdom	Prognostic; TB progressors vs. non-progressors
Sambarey10 (Sambarey et al., 2017)	Linear discriminant analysis	Adults	India	Diagnostic; TB vs. LTBI and ORD
Satproedprai7 (Satproedprai et al., 2015)	LASSO regression	Adults	Thailand	Diagnostic; TB vs. HC and previous TB patients
Suliman2 (Suliman et al., 2018)	Pair ratio algorithm	Adults	The Gambia, South Africa, Ethiopia	Prognostic; TB progressors vs. non-progressors
Suliman4 (Suliman et al., 2018)	Pair ratio algorithm	Adults	The Gambia, South Africa	Prognostic; TB progressors vs. non-progressors
Sweeney3 (Sweeney et al., 2016)	TB score (GBP5 + DUSP3) / 2 – KLF2	Adults	Meta-analysis (France, Malawi, South Africa, United Kingdom, USA)	Diagnostic; TB vs. HC, LTBI and ORD
Thompson5 (Thompson et al., 2017b)	Pair ratio approach	Adults	South Africa	Monitoring response to TB treatment

LASSO = least absolute shrinkage and selection operator. SVM = support vector machines. HC = healthy controls. ORD = other respiratory diseases. LTBI = latent tuberculosis infection. TB = tuberculosis.

Signature names combine the surname of the first author of the original publication, and the number of constituent transcripts within the signature.

†Kaforou25 had 27 transcripts in the original signature. Primer-probe assays for two FCGR1B transcripts in the original signature could not be identified since the microarray probe was a duplicate of the same FCGR1B transcript. LOC728744 was discontinued and had all information withdrawn from both the NCBI and RefSeq databases.

Table 3.2: Representation of transcriptomic signatures across different systematic reviews

	(Gupta et al., 2020)	(Mendelsohn et al., 2023)	(Mulenga et al., 2020)	(Turner et al., 2020)	(Warsinske et al., 2019)
DA COSTA3			X		X
DE ARAUJO1		X	X	X	
DUFFY9		X		X	
FRANCISCO2			X		
GJØEN7	X	X		X	
GLIDDON3	X	X		X	
GLIDDON4		X		X	
JACOBSEN3					X
KAFOROU25	X	X	X	X	X
MAERTZDORF4	X	X		X	X
PENN-NICHOLSON6		X		X	
RAJAN5	X	X		X	
ROE1	X	X		X	
ROE3	X	X	X	X	
SAMBAREY10		X	X		X
SATPROEDPRAI7					
SULIMAN2	X	X		X	
SULIMAN4	X	X	X	X	
SWEENEY3	X	X	X	X	X
THOMPSON5					

3.1.2 Selection of Housekeeper genes

The use of internal control genes, reference genes or housekeeper genes (HKGs), is used to control for variation in template concentration that strongly influences gene expression results (Eisenberg and Levanon, 2013). Some of the factors corrected for include variations in starting material used i.e., cDNA or RNA, differences in samples collection, variations in reverse transcription efficiency, and the presence of inhibitors.

Using multiple HKGs, in comparison to just one, has been demonstrated as a better and more accurate approach for normalization of gene expression results. From an evaluation done by Vandesompele and colleagues, the use of a single HKG for normalization, in comparison to multiple ones, led to significant errors in normalization for majority of samples tested (Vandesompele et al., 2002). One HKG could not accurately capture variations in gene expression due to differences in experimental conditions. Normalization bias occurs once the expression of HKGs is not constant across samples, also leading to inaccurate quantification of gene expression levels. In another study where up to 10 HKGs were evaluated in different tissues and cells under multiple experimental conditions, the authors recommended the use of a minimum of three HKGs for normalization (Vandesompele et al., 2002). In a third study, where up to 10 HKGs were also evaluated, Pfaffl and colleagues came to a similar conclusion of three being the minimal essential number required for accurate normalization (Pfaffl et al., 2004).

The aim of this selection process was therefore to identify the top three HKGs out of the ten (**Table 3.3**) that were part of the of the panel used in the CORTIS study (Scriba et al., 2021). The main characteristic used to select these top 3 HKGs was their stability in expression across all cases and controls. HKGs required for normalization need to demonstrate little to no fluctuation in gene expression, when measured in different cells under varying experimental conditions (Vandesompele et al., 2002). Gene expression stability was evaluated in R using the geNorm (Vandesompele et al., 2002), NormFinder (Anderson et al., 2014), BestKeeper (Pfaffl et al., 2004), and comparative delta Ct methods (Silver et al., 2006a). More than one method was used to minimize bias and make the selection process for HKGs more robust.

Table 3.3: Panel of HKGs from the CORTIS study

House Keeping Gene	Description
ACTR3.Hs01029159_g1	Encodes the protein Actin-related protein 3
ADRBK1.Hs01056345_g1	Encodes the enzyme G-protein-coupled receptor kinase 2
CDC42.Hs03044122_g1	It is a protein-coding gene that regulates a variety of cellular functions
CSDE1.Hs00918650_m1	A protein coding gene that regulates gene expression and plays a role in human diseases
CYTIP.Hs00188734_m1	Encodes the cytohesin-interacting protein that regulates immune responses
TMBIM6.Hs01012081_m1	Transmembrane Bax Inhibitor Motif-containing 6 (TMBIM6) is a protein that regulates cell death pathways
TMBIM6.Hs00162661_m1	
TMBIM6.Hs01012082_g1	
TPM3.Hs01900726_g1	Encodes the slow muscle alpha (α)-tropomyosin protein that regulates muscle contraction
USF2.Hs01100994_g1	Encodes a member of the basic helix-loop-helix leucine zipper family of transcription factors

Each HKG is represented by its gene name and the ThermoFisher assay ID used to identify it on the ThermoFisher website. The description for each gene was obtained from the NCBI website under NIH.

Vandesompele and colleagues developed a gene-stability metric, M, which was the average pairwise variation between one HKG and all other HKGs under investigation (Vandesompele et al., 2002). M was automatically calculated using the geNorm program (Vandesompele et al., 2002), which also gives a score of stability for each HKG, ranked by the lowest M value.

The NormFinder program aims to identify the single best gene for normalization using a model-based approach to determine the expression stability of a potential HKG (Andersen et al., 2004).

Bestkeeper software determines the most optimal HKGs by carrying out a pair-wise correlation analysis of all pairs of genes (Pfaffl et al., 2004). Gene expression data for all HKGs under investigation are compiled into an index and compared to target genes being analysed under the same experimental setup. This helps determine whether HKGs are differentially expressed. The software also analyses the expression stability of HKGs by determining geometric mean, arithmetic mean, standard deviation (SD), and coefficient of variance (CV) (Pfaffl et al., 2004).

The comparative delta Ct approach integrates the geNorm software technique of assessing the relative expression of pairs of genes (Silver et al., 2006b). This, however, is done using a delta Ct (ΔCt) method. The stability of candidate HKGs was ranked based on the delta Ct value between each pair of HKGs remaining constant when analysed in different samples (Silver et al., 2006b).

For HKGs to be used effectively to normalize gene expression results, they need to be detected in all samples that are part of a study. With that in mind, one aspect that was looked at was the pass rate (i.e. the number of samples in which a Ct value was obtained). In a total of 5084 RNA samples from the CORTIS study, **Table 3.4** shows the total number of samples in which Ct values were obtained from each HKG. The minimum pass rate was 93%.

Table 3.4: Pass rate of HKGs based on results from CORTIS data.

House Keeping Gene	Pass rate (%)
ACTR3.Hs01029159_g1	94.59
ADRBK1.Hs01056345_g1	94.16
CDC42.Hs03044122_g1	94.02
CSDE1.Hs00918650_m1	93.9
CYTIP.Hs00188734_m1	94.34
TMBIM6.Hs01012081_m1	94.59
TMBIM6.Hs00162661_m1	94.59
TMBIM6.Hs01012082_g1	94.59
TPM3.Hs01900726_g1	94.59
USF2.Hs01100994_g1	93.9

Each HKG is represented by its gene name and the ThermoFisher assay ID used to identify it on the ThermoFisher website.

The rankings (**Table 3.5**) were based on the metrics calculated from each of the four methods used to evaluate HKGs. Each ranking gave the order of HKGs identified as having the most stable gene expression. Based on the rankings present, three HKGs were selected for all subsequent RT-qPCR work: TMBIM6.Hs00162661_m1, USF2.Hs01100994_g1, and ACTR3.Hs01029159_g1.

Table 3.5: Rankings of the 10 HKGs from all the CORTIS runs based on results from four different methods for selection of HKGs.

geNorm_rank	geNorm_meanM	NormFinder_rank	NormFinder_Stability	BestKeeper_correlation_rank	BestKeeper_correlation	BestKeeper_SD_rank	BestKeeper_SD	DeltaCt_rank	DeltaCt_MeanStdDev
CDC42.Hs03044122_g1 TMBIM6.Hs00162661_m1	0.048	CDC42.Hs03044122_g1	0.31	TMBIM6.Hs00162661_m1	0.994	USF2.Hs01100994_g1	1.99	TMBIM6.Hs00162661_m1	0.69
NA [†]	NA [†]	TMBIM6.Hs00162661_m1	0.31	CDC42.Hs03044122_g1	0.993	TPM3.Hs01900726_g1	2.14	CDC42.Hs03044122_g1	0.7
ACTR3.Hs01029159_g1	0.052	USF2.Hs01100994_g1	0.41	ACTR3.Hs01029159_g1	0.99	CDC42.Hs03044122_g1	2.16	TMBIM6.Hs01012082_g1	0.73
CSDE1.Hs00918650_m1	0.056	TMBIM6.Hs01012082_g1	0.42	USF2.Hs01100994_g1	0.99	TMBIM6.Hs01012082_g1	2.2	USF2.Hs01100994_g1	0.74
ADRBK1.Hs01056345_g1	0.059	CSDE1.Hs00918650_m1	0.46	TMBIM6.Hs01012082_g1	0.988	ADRBK1.Hs01056345_g1	2.23	CSDE1.Hs00918650_m1	0.75
USF2.Hs01100994_g1	0.06	ACTR3.Hs01029159_g1	0.46	CSDE1.Hs00918650_m1	0.987	CSDE1.Hs00918650_m1	2.23	ACTR3.Hs01029159_g1	0.76
TMBIM6.Hs01012082_g1	0.062	ADRBK1.Hs01056345_g1	0.54	ADRBK1.Hs01056345_g1	0.984	TMBIM6.Hs01012081_m1	2.23	ADRBK1.Hs01056345_g1	0.81
TMBIM6.Hs01012081_m1	0.065	TMBIM6.Hs01012081_m1	0.61	TMBIM6.Hs01012081_m1	0.979	TMBIM6.Hs00162661_m1	2.26	TMBIM6.Hs01012081_m1	0.85
CYTIP.Hs00188734_m1	0.071	CYTIP.Hs00188734_m1	0.86	CYTIP.Hs00188734_m1	0.961	CYTIP.Hs00188734_m1	2.27	CYTIP.Hs00188734_m1	1.02
TPM3.Hs01900726_g1	0.083	TPM3.Hs01900726_g1	1.31	TPM3.Hs01900726_g1	0.904	ACTR3.Hs01029159_g1	2.32	TPM3.Hs01900726_g1	1.41

Each method aims to assess the stability of each housekeeping gene and ranks them, from highest to lowest, according to the metrics calculated.

geNorm generates a stability score (“M”). HKGs are ranked by the lowest stability score (“M”) value. † After geNorm calculates the pairwise SD of Ct values of all genes, it excludes the one with the lowest stability, and goes on to repeat the process.

NormFinder identifies the single best gene for normalization and ranks HKGs according to their expression stability.

BestKeeper analyses the expression stability of HKGs through a pair-wise correlation analysis of all pairs of genes and calculation of the SD of Ct values. BestKeeper ranks HKGs based on a SD <1.

The comparative delta Ct approach assesses the stability of HKGs by calculating the delta Ct value between each pair of genes.

3.1.3 Selecting and designing primer probe assays

All transcripts of interest were identified and detected using either pre-designed or custom-made 5' nuclease TaqMan gene expression assays (Applied Biosystems for ThermoFisher)(**Figure 3.1**). Each assay came in a single tube containing a mix of a forward primer, a reverse primer and a TaqMan minor groove binder (MGB) probe labelled with the fluorophore, 6-carboxyfluorescein (FAM).

Pre-designed assays were identified using the gene name and unique assay IDs provided by authors within published literature for the respective transcriptomic signatures under evaluation. When an assay for a transcript of interest was not available among the pre-designed assays within the ThermoFisher library, gene expression assays were designed, using the Custom TaqMan Assay Design Tool on the ThermoFisher website.

To design custom assays, target sequences, in the 5' to 3' direction, that were between 61 to 5000 nucleotides long, were entered and target positions for the probes specified. Custom-plus assays were designed by further selecting the option for ThermoFisher to carry out additional bioinformatic analysis on the target sequences provided to ensure the best assay design results.

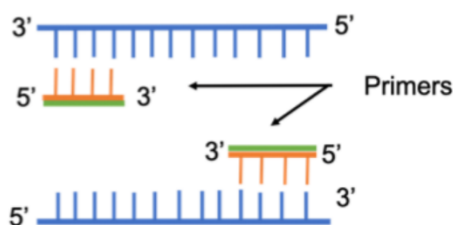


Figure 3.1: Primers for PCR reactions. PCR reactions require two complementary oligonucleotide primers.

Primers anneal to the regions upstream and downstream of the DNA segment to be amplified.

3.2: Assay qualification

All Taqman gene expression assays were validated on the Fluidigm BioMark microfluidic RT-qPCR platform through a process termed assay qualification. The main aim of assay qualification was to choose the most optimal set of primer-probe assays for quantification of gene expression for each transcript on the panel. This was especially important for transcripts where more than one gene expression assay had been identified during the assay selection and design process. A total of 124 assays were ordered and tested during assay qualification, so the top performing 96 assays could be chosen for the panel. Ordering and testing 124 primer-probe assays ensured that alternatives would already be present from the start for several transcripts of interest. These alternative primer-probe assays ensured that certain transcripts would still be represented on the panel or replaced any transcripts that lacked an alternative assay so that a total of 96 transcripts would still be evaluated.

During assay qualification, each Taqman primer-probe set was tested for efficient amplification (i.e., the signal doubles with each PCR cycle) and linearity (i.e., the final signal is proportional to input RNA)(Dominguez et al., 2013).

3.2.1 Sample selection

cDNA from a total of four samples with known RNA concentrations were chosen for assay qualification. There were two samples each with high RNA concentrations and low RNA concentrations. This was factored into the experimental design to ensure that the primer probes that were selected were able to work effectively for different samples, within subsequent research projects, that had varying RNA concentrations.

The concentrations of the 4 samples used were as follows:

- Sample 1: 104.5 ng/ μ l
- Sample 2: 86.5ng/ μ l
- Sample 3: 28.6ng/ μ l
- Sample 4: 9.5ng/ μ l

3.2.2 Pre-amplification

93 transcripts of interest and 3 housekeeping transcripts were pre-amplified (**described in section 2.3.2 of Chapter 2**) using a pool of Taqman primer-probe assays. Samples for pre-amplification (**Table 3.6**) were pipetted into the first two columns of a 96-well PCR plate (Applied Biosystems by Thermo Fisher Scientific, Waltham, MA, USA).

Table 3.6: Layout of samples for pre-amplification in preparation for assay qualification

	1	2	3	4	5	6	7	8	9	10	11	12
A	Water	Sample 1										
B	Water	Sample 2										
C	NTC1	Sample 3										
D	NTC1	Sample 4										
E	NoRT1	Additional sample 1 for PCR confirmation										
F	NoRT2	Additional sample 2 for PCR confirmation										
G	Empty	NTC for PCR confirmation										
H	Empty	Empty										

NTC - No Template Control, NoRT- No Reverse Transcriptase control

Samples 1 to 4 in wells 2A to 2D, (**Table 3.6**) above, were used to prepare the 1 in 2 dilution series (**Table 3.7**) for assay qualification, while samples in wells 2E-2G were used to perform confirmatory PCR test. An internal positive control (IPC) was included and run in duplicate on all microfluidic RT-qPCR chips for assay qualification (see section 3.4 for details). The IPC comprised samples positive for all primer-probe assays that were evaluated during assay qualification.

3.2.3 Dilution series on pre-amped cDNA

Table 3.7: Dilution series for assay qualification

	Dilution	1:2	1:4	1:8	1:16	1:32	1:64	1:128	1:256	1:512	1:1024	1:2048
	1	2	3	4	5	6	7	8	9	10	11	12
A	Water	Sample 1A	Sample 1A	Sample 1A	Sample 1A	Sample 1A	Sample 1A	Sample 1A	Sample 1A	Sample 1A	Sample 1A	Sample 1A
B	Water	Sample 1B	Sample 1B	Sample 1B	Sample 1B	Sample 1B	Sample 1B	Sample 1B	Sample 1B	Sample 1B	Sample 1B	Sample 1B
C	NTC1	Sample 2A	Sample 2A	Sample 2A	Sample 2A	Sample 2A	Sample 2A	Sample 2A	Sample 2A	Sample 2A	Sample 2A	Sample 2A
D	NTC1	Sample 2B	Sample 2B	Sample 2B	Sample 2B	Sample 2B	Sample 2B	Sample 2B	Sample 2B	Sample 2B	Sample 2B	Sample 2B
E	noRT1	Sample 3A	Sample 3A	Sample 3A	Sample 3A	Sample 3A	Sample 3A	Sample 3A	Sample 3A	Sample 3A	Sample 3A	Sample 3A
F	noRT2	Sample 3B	Sample 3B	Sample 3B	Sample 3B	Sample 3B	Sample 3B	Sample 3B	Sample 3B	Sample 3B	Sample 3B	Sample 3B
G	#	Sample 4A	Sample 4A	Sample 4A	Sample 4A	Sample 4A	Sample 4A	Sample 4A	Sample 4A	Sample 4A	Sample 4A	Sample 4A
H	#	Sample 4B	Sample 4B	Sample 4B	Sample 4B	Sample 4B	Sample 4B	Sample 4B	Sample 4B	Sample 4B	Sample 4B	Sample 4B

NB: The wells G1 and H1, designated with a hashtag, #, remained empty and served as place holders for the IPC. The IPC comprised samples positive for all primer-probe assays that were evaluated during assay qualification. The IPC was only plated when samples were pipetted into the wells of the microfluidic chips.

Using pre-amped cDNA in duplicate, a 1 in 2 dilution series was carried out across 96-well plates. (**Table 3.7**). Each plate contained one column of controls (a water control, NTC, noRT, and IPC) present in duplicate.

On a fresh and empty 96-well plate, 4µl of water was added to each well within columns 2 to 12, while 8µl of the respective controls were pipetted into the designated wells under column 2. Four microlitres was progressively transferred from each well under column 2 all the way to column 12 to achieve a final dilution of 1 in 2048 for each sample on the plate.

Each 96-well plate was sealed with plate film, centrifuged briefly, and stored in a -40°C freezer until it was time to perform the subsequent multiplex RT-qPCR experiments.

3.2.4 Gene expression measurement

A confirmatory PCR, to ascertain that pre-amplification was successful, was carried out on the QuantStudio 3 Real-Time PCR System (Applied Biosystems by Thermo Fisher Scientific, Waltham, MA, US) (**Chapter 2, Section 2.3.3**). Gene expression levels of the transcripts were measured by RT-qPCR using the Fluidigm Biomark HD instrument (**Chapter 2, Section 2.3.4**).

3.2.5 Criteria used to qualify assays

The 12-dilution series (**Table 3.7**) was divided into seven segments, with each segment made up of five consecutive dilutions. Each segment was assessed gradually, starting with the lowest amount of RNA. This is what is referred to as a piecewise method for assay qualification (Dominguez et al., 2013). As opposed to assessing the whole series of dilutions at once, this piecewise approach has demonstrated higher pass rates for primer-probe assays (Dominguez et al., 2013). For each segment, the linear least-squares correlation between the logarithm of RNA concentration and Et (40-Ct) values was calculated. Et values increased when a transcript was highly expressed and were therefore taken to be easier to interpret (Dominguez et al., 2013).

A segment within a dilution series passed when:

1. A minimum of 5 consecutive dilution points demonstrated linearity, indicated by R^2 values greater or equal to 0.98 and slope values between 3.1 and 3.6.
2. Efficiency (E) values, calculated as $10^{1/\text{slope}} - 1$, were between 90 and 110%.

For an assay to have passed, these parameters needed to be seen in at least 2 replicates, either from the same samples or different samples per assay. i.e.,

- i. 2 samples: both replicates
- ii. 1 sample: both replicates
- iii. 2 replicates: Even if they were from different samples.

3.2.6 Outcomes from assay qualification and further selection of assays for the 96-transcript panel

The majority (116 out of 124) of the primer-probe assays assessed during the assay qualification met the minimal pass criteria and were thus validated successfully. The primer-probe assay for the CD3E transcript (**Figure 3.2**) is an example of an assay that passed in all four samples and replicates, exceeding the minimal pass criteria. The primer-probe assay for the FCGR1A (Hs00174081_m1) transcript (**Figure 3.2**) is an example of a primer-probe assay that passed in three out of the 4 samples tested and showed consistently uniform amplification across samples. The primer-probe assays for LOC400759 (GBP1P1) (**Figure 3.3**), one of the suggested paediatric gene targets and CERKL.1 (**Figure 3.3**), that was part of the Duffy10 signature (Duffy et al., 2019) consistently failed to meet the minimal pass criteria on any of the 96.96 gene expression integrated fluidic circuit (IFC) nanofluidic chips. These two assays were removed from the panel as there were no alternative primer-probes for them. From this point on, the Duffy signature was evaluated based on the measurement of nine transcripts.

To still have a total of 96 primer-probe assays on our panel, CERKL.1 was replaced with a second FCGR1A assay (AssayID: Hs00174081_m1), and LOC400759 replaced with a custom KLF2 primer-probe assay. These two assays passed assay qualification and had a pass rate of 100% when evaluated in the Grand Challenges 6-74 (GC6-74, GC6) household contact (HHC) study for training predictive transcriptomic biomarker for TB progression (Suliman et al., 2018).

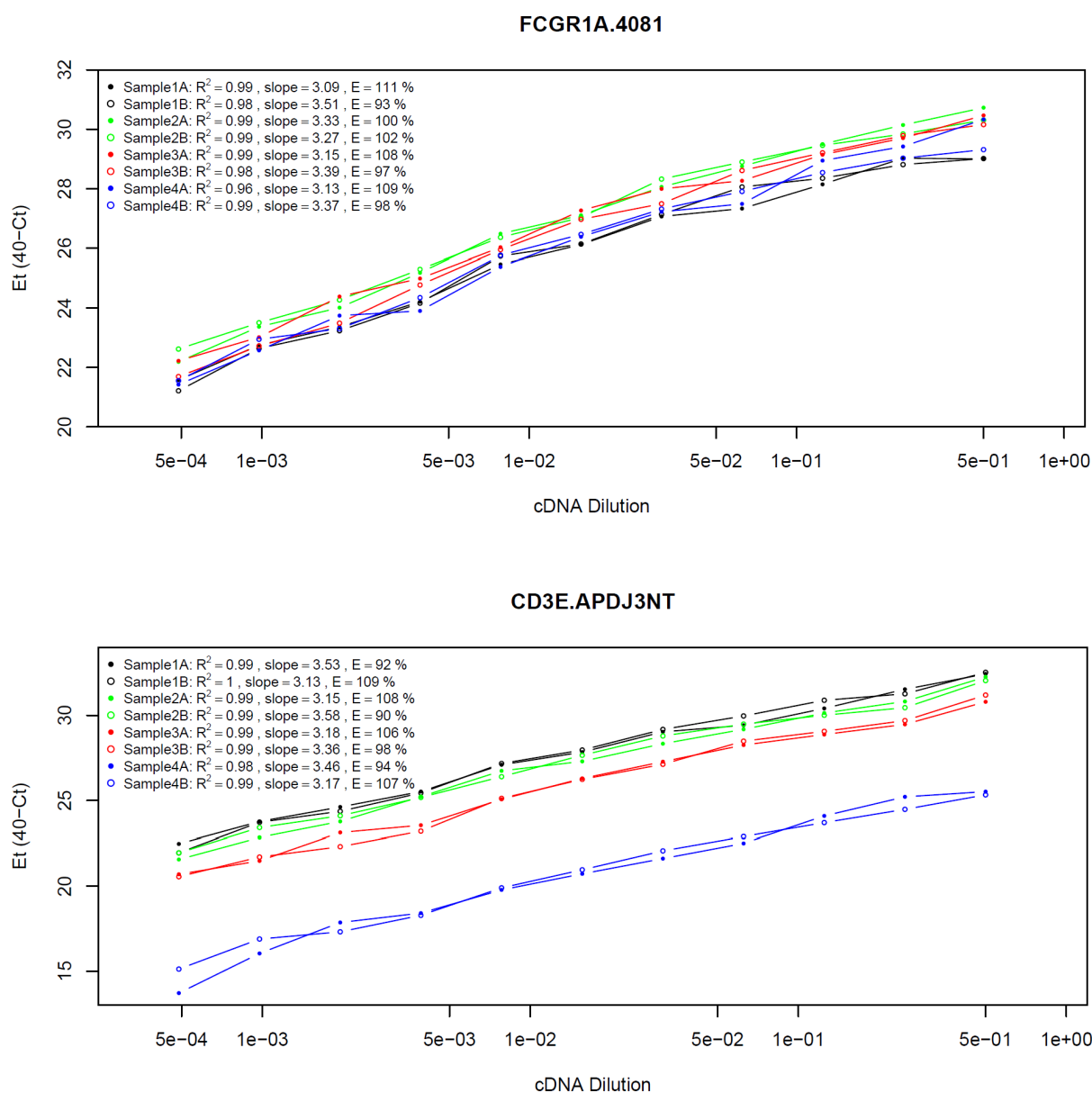


Figure 3.2: Representative primer-probe assays that met the minimal pass criteria for assay qualification.

The threshold expression values (Et values), calculated as 40-Ct, were plotted against RNA concentrations represented on a log scale. Sample 1 is shown in black; Sample 2 is shown in green; Sample 3 is shown in red, and Sample 4 is shown in blue. (a) FCGR1.4081 represents the primer probe for the FGCR1A transcript and has the assay ID Hs00174081_m1. (b) CD3E.APDJ3NT is a custom assay that represents the primer probe for the CD3E transcript and has the assay ID CD3E_APDJ3NT. FCGR1.4081 passed in three out of the four samples evaluated during assay qualification and CD3E.APDJ3NT passed in all samples and replicates with R2 values greater or equal to 0.98 and slope values between 3.1 and 3.6.

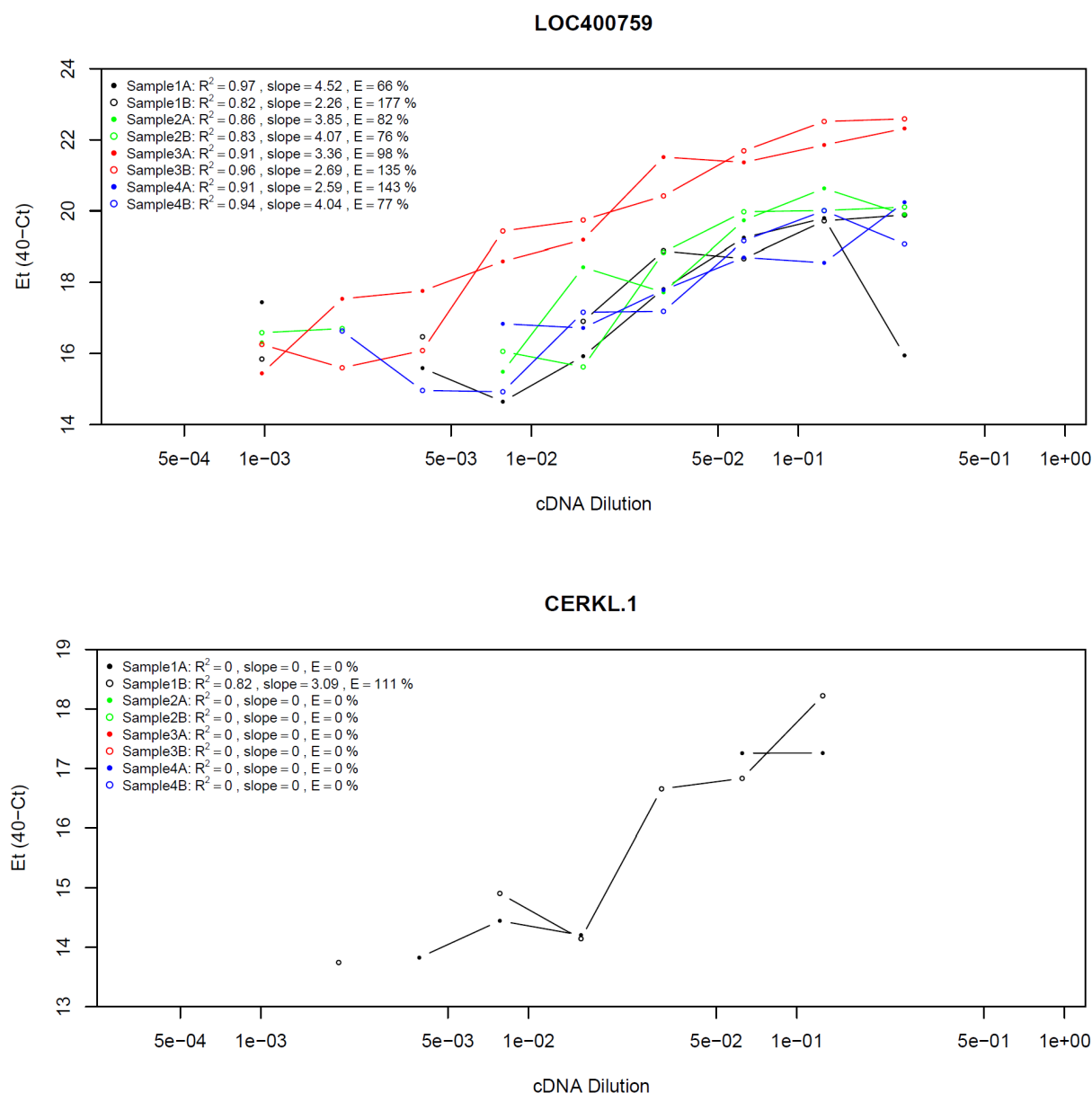


Figure 3.3: Representative primer-probe assays that failed assay qualification.

The threshold expression values (Et values), calculated as 40-Ct, were plotted against RNA concentrations represented on a log scale. Sample 1 is shown in black; Sample 2 is shown in green; Sample 3 is shown in red, and Sample 4 is shown in blue. (a) LOC400759 represents the primer probe for the LOC400759 (GBP1P1) transcript and has the assay ID Hs03299031_m1. (b) CERK.1 represents the primer probe for the CERKL.1 transcript and has the assay ID Hs00923424_m1. Both LOC400759 and CERK.1 failed in all samples and replicates with R2 values less than 0.98 and slope values outside the range of 3.1 - 3.6.

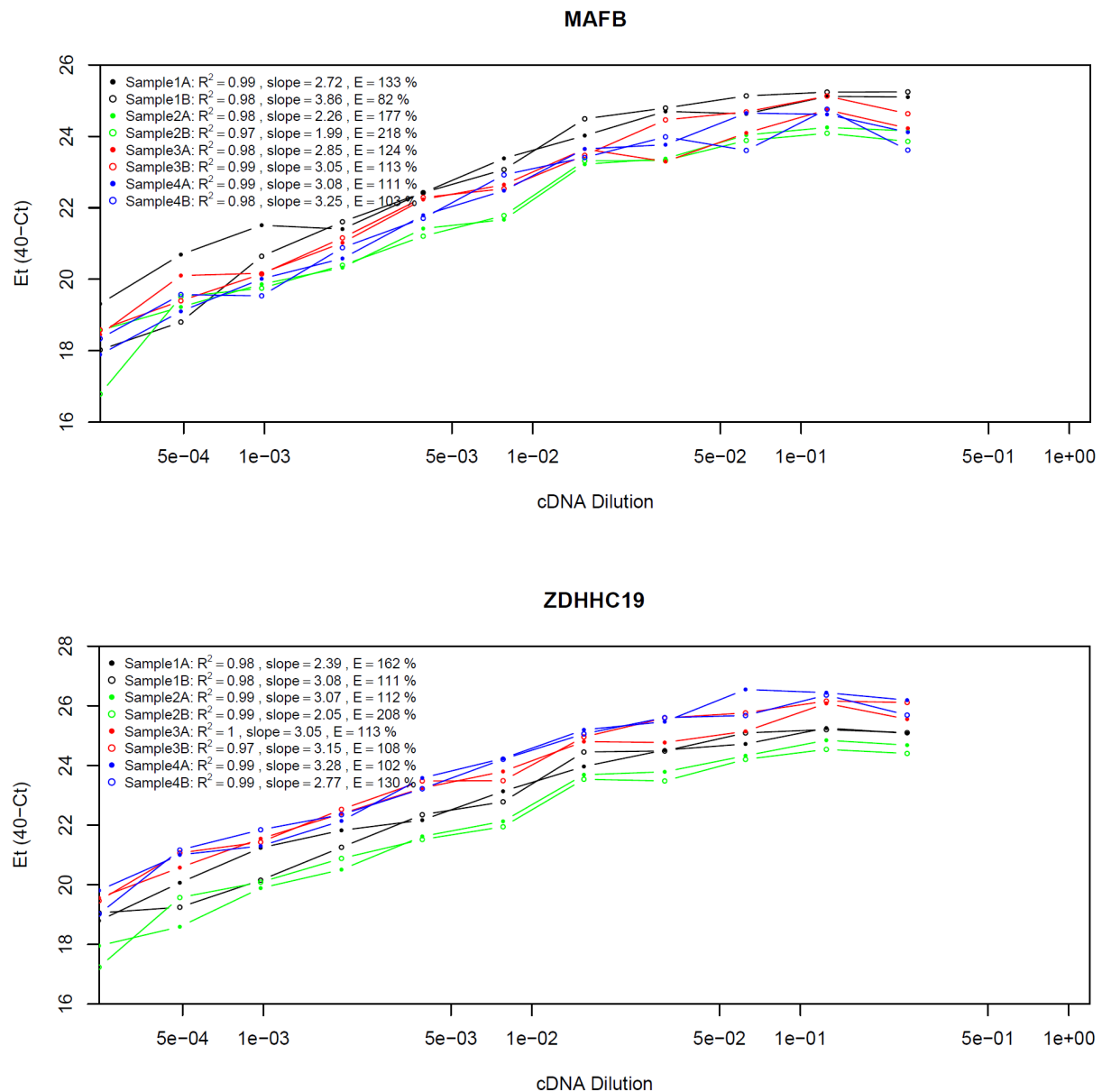


Figure 3.4: Representative primer-probe assays that were selected for the panel despite not meeting the criteria.

The threshold expression values (Et values), calculated as 40-Ct, were plotted against RNA concentrations represented on a log scale. Sample 1 is shown in black; Sample 2 is shown in green; Sample 3 is shown in red, and Sample 4 is shown in blue. (a) MAFB represents the primer probe for the MAFB transcript and has the assay ID MAFB_SAT7_APZTH2N (b) ZDHHC19 represents the primer probe for the ZDHHC19 transcript and has the assay ID Hs00376118_m1.

Six assays (**Table 3.8**) were retained on the panel even though they did not fully meet the assay qualification criteria. This is because they still demonstrated consistent amplification across chips (**Figure 3.4**) and were part of published signatures included in this thesis. TRMT2A was one transcript that was tested on another panel, where it passed assay qualification (Mendelsohn et al., 2022b).

Transcript	AssayID	Signature(s)	Rationale for selecting the assay
GAS6	AR323W3_GAS6_NM_000820_10_11	Kaforou25	The Kaforou25 signature originally had 2 GAS6 transcripts.
GBP6	Hs01584201_m1	Duffy9, Gliddon4, Kaforou25, Rajan5	GBP6 was present in 4 different signatures. It demonstrated consistent amplification across chips.
MAFB	MAFB_SAT7_APZTH2N	Satproedprai7	Demonstrated consistent amplification across chips
ZDHHC19	Hs00376118_m1	Duffy9	Demonstrated consistent amplification across chips
TMCC1	Hs01037666_s1	Gliddon4	Demonstrated consistent amplification across chips
TRMT2A	Hs01000041_g1	Penn-Nicholson6	Demonstrated consistent amplification across chips. This assay passed assay qualification and was assessed in previous study cohorts at SATVI that evaluated the Penn-Nicholson6 signature (Mendelsohn et al., 2022b; Penn-Nicholson et al., 2020b)

Table 3.8: Primer-probe assays retained on the panel.

3.2.7 Conclusions from assay qualification

The most optimal set of 96 primer-probe assays (**Table 3.9**) were assessed and chosen to represent all 20 transcriptomic signatures selected at the start of the optimization process.

Table 3.9: The 96 primer-probe assays included in the final gene expression panel.

Assay	Assay ID	Number of passed samples	Assay qualification result
ACTA2	Hs00426835_g1	4	pass
ACTR3	Hs01029159_g1	5	pass
ANKRD22	Hs00944015_m1	6	pass
APOL1	Hs01066280_m1	5	pass
ARG1	Hs00968979_m1	4	pass
BATF2	Hs00912737_m1	3	pass
BCL6	BCL6_CUSTOM_AREPURD	4	pass
BLK	Hs01017452_m1	4	pass
C1QA	Hs00706358_s1	3	pass
C1QB	Hs00608019_m1	3	pass
C1QC	Hs00757779_m1	2	pass
C4ORF18(FAM198B)	Hs00259260_s1	4	pass
C5	Hs00156197_m1	3	pass
CCR6	Hs01890706_s1	3	pass
CD160	Hs00199894_m1	2	pass
CD1C	Hs00957534_g1	5	pass
CD36	Hs01567186_m1	3	pass
CD3E	CD3E_APDJ3NT	8	pass
CD40L	Hs00163934_m1	3	pass
CD74-j1	Hs04983808_s1	6	pass
CD74j2	CD74-J2_AP9HN47	2	pass
CD79A	Hs00998120_g1	5	pass
CD79B	Hs01058826_g1	7	pass
CDCA7	Hs00230589_m1	2	pass
CDKN1C	Hs00175938_m1	3	pass
CXCR5	Hs00540548_s1	3	pass
CYP4F3	Hs01587860_m1	4	pass
DUSP3	Hs01115776_m1	5	pass
F2RL1	Hs00608346_m1	2	pass
FAM20A	Hs01034066_m1	3	pass
FCGR1A	Hs00174081_m1	8	pass
FCGR1A	FCGR1A_SAT7_AP2XDMK	6	pass
FCGR1B	Hs02341825_m1	4	pass
FCGR1B VARIANT1	FCGR1B-VARIANT1_AP3267H	5	pass
FCGR1B VARIANT2	FCGR1B-VARIANT2_AP47ZTF	4	pass
FCGR1C	Hs00417598_m1	5	pass
FLVCR2	Hs00900390_m1	3	pass
GAS6	GAS6_NM_000820_5_6_AR47XGZ	2	pass
GAS6*	GAS6_NM_000820_10_11_AR323W3	0	fail
GBP1	Hs00977005_m1	5	pass
GBP2	Hs00894846_g1	5	pass
GBP5	Hs00369472_m1	4	pass
GBP6*	Hs01584201_m1	1	fail
GNG7	Hs00192999_m1	5	pass
GYG1	Hs00907542_g1	3	pass
GZMA	Hs00989184_m1	3	pass
HK3	Hs01092850_m1	8	pass
ID3	Hs00954037_g1	4	pass
IFI44L	Hs00915292_m1	6	pass
IFITM3	IFITM3_G_APMF24C	5	pass

Assay	Assay ID	Number of passed samples	Assay qualification result
IFITM3_G	Hs03057129_s1	3	pass
KAZN	KAZN-TOTAL_SAT7_AP7DVDD	4	pass
KIF1B	Hs01114512_g1	5	pass
KLF2	Hs00360439_g1	5	pass
KLF2	KLF2_CUSTOM_ARH6CF4	2	pass
KLHDC8B	Hs00293902_m1	4	pass
LAG3	Hs00958444_g1	2	pass
LHFPL2	Hs00299613_m1	4	pass
LTF	Hs00158924_m1	6	pass
MAFB*	MAFB_SAT7_APZTH2N	0	fail
MAP7D3	Hs00226257_m1	4	pass
MMP9	MMP9_APEPW9P	4	pass
MPO	Hs00165162_m1	2	pass
MTRF1L	Hs01097882_g1	4	pass
NOD2	Hs01550759_g1	3	pass
NPC2	Hs01119244_m1	4	pass
OSBPL10	Hs00215016_m1	5	pass
P2RY14	Hs01848195_s1	4	pass
PRDM1	Hs00153357_m1	6	pass
RAB13	Hs04400188_g1	5	pass
Rab33A	Hs00191243_m1	2	pass
RABL2A	Hs00255244_m1	2	pass
RBBP8	Hs01090329_m1	3	pass
RP11.2	Hs01373568_m1	3	pass
S100A8	Hs00374264_g1	3	pass
SCARF1	Hs01092483_m1	2	pass
SDR39U1	Hs01016970_g1	2	pass
SEPT4	Hs00910208_g1	3	pass
SERPING1	Hs00934329_m1	3	pass
SLPI	Hs00268206_m1	3	pass
SMARCD3	Hs01088251_g1	4	pass
STAT1	Hs01013996_m1	3	pass
STT3A	Hs00967491_m1	2	pass
TIMM10	TIMM10_CUSTOM_APT2DEC	2	pass
TMBIM6	Hs00162661_m1	4	pass
TMCC1*	Hs01037666_s1	1	fail
TNIP1.GEO	TNIP1_GEO_AP47ZYZ	2	pass
TRMT2A*	Hs01000041_g1	1	fail
TUBGCP6	Hs00363509_g1	3	pass
UCP2	Hs01075224_g1	3	pass
USF2	Hs01100994_g1	5	pass
VAMP5	Hs01105383_g1	2	pass
WARS-j1	Hs00998737_m1	2	pass
WARS-j2	WARS-J2_APAAFHX	2	pass
ZDHHC19*	Hs00376118_m1	1	fail
ZNF296	Hs00377132_m1	2	pass

GAS6*, GBP6*, MAFB*, TMCC1*, TRMT2A*, and ZDHHC19* were included on the panel even though they did not fully meet the assay qualification criteria due to inconsistent amplification. TRMT2A met the assay qualification in previous study cohorts(Mendelsohn et al., 2022b; Scriba et al., 2021a)

3.3: Reparameterization

A signature required reparameterization (**Table 3.10**) if it was originally developed using a gene expression analysis technique other than the Standard BioTools BioMark multiplex RT-qPCR platform. This is because different platforms each have their own dynamic ranges and scales for quantifying gene expression levels.

Table 3.10: Transcriptomic signatures on the panel that were reparameterized

Signature	Gene expression technique	Model used
DA-COSTA3	RT-qPCR	Random forest (da Costa et al., 2015)
DUFFY9	Meta-analysis of microarray data	Six-class multinomial random forest (Duffy et al., 2019)
FRANCISCO2	RT-qPCR	Random forest (Francisco et al., 2017)
GJØEN7	Dual-color-Reverse-Transcriptase-Multiplex-Ligation-dependent-Probe-Amplification (dc-RT MLPA)	LASSO regression (Gjøen et al., 2017a)
JACOBSEN3	RNA microarray	Linear discriminant analysis(Jacobsen et al., 2007)
ROE3	RNA sequencing	Support vector machine (linear kernel)
SAMBAREY10	RNA sequencing	Linear discriminant analysis
SATPROEDPRAI7	RT-qPCR	LASSO regression (Satproedprai et al., 2015)

Signature names represent the first author's name of the corresponding publication, followed by the number of constituent transcripts genes. The models used during reparameterization are explained in further detail in the individual publications where each transcriptomic signature was discovered.

Reparameterization was carried out on whole blood samples from a cohort of TB cases and controls. This ensured that machine learning models and statistical methods used for computing signature scores were reconstructed (reparameterized) to optimally discriminate between symptomatic TB cases and healthy controls. All gene expression measurements for TB cases and controls were generated from RT-qPCR runs on the BioMark HD platform.

Two signatures had not been evaluated previously in Fluidigm Biomark RT-qPCR experiments at SATVI, but did not require any reparameterization, even though they were discovered using genomic approaches other than RT-qPCR. These include Kaforou25 (Kaforou et al., 2013) and Rajan5 (Rajan et al., 2019).

Kaforou25 (**Figure 3.5**) was discovered based on microarray data and a Disease Risk Score (DRS) was developed to easily generate results on gene expression. DRS sums up normalized Ct values from upregulated transcripts in the signature and subtracts the normalized Ct values of all downregulated transcripts, making the signature essentially "self-normalizing". Rajan5 was also discovered based on microarray gene expression data, but developed using the signed and unsigned sums methods, which also allowed the signature to be "self-normalizing".

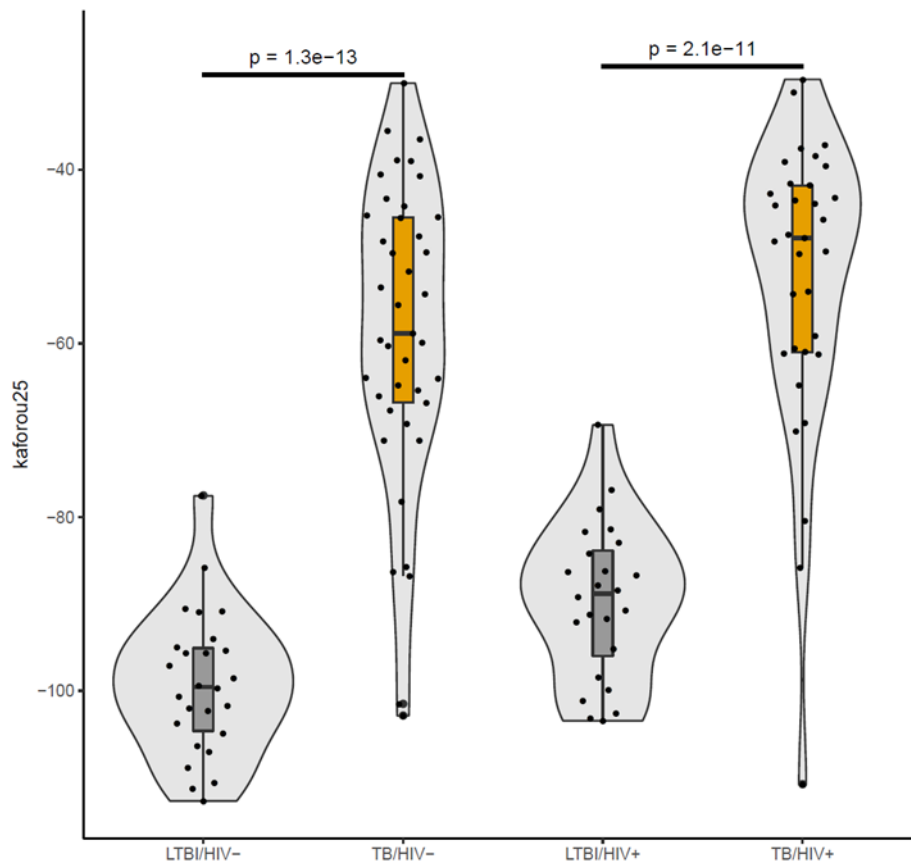


Figure 3.5: Signature score distribution for Kaforou25.

LTBI = Latent TB, TB= active TB, HIV- = HIV negative, HIV+ = HIV positive. The Violin plots show the distribution of scores across different categories of CTBC study participants. Each dot represents a study participant. P-values for comparison of signature scores across groups were calculated using the Mann–Whitney U test. The width of each inner box plot shows the IQR, the midline represents the median, and the whiskers are $1.5 \times \text{IQR}$.

Rajan5 (**Figure 3.6**) was also discovered based on microarray gene expression data, but developed using the signed and unsigned sums methods, which also allowed the signature to be “self-normalizing”.

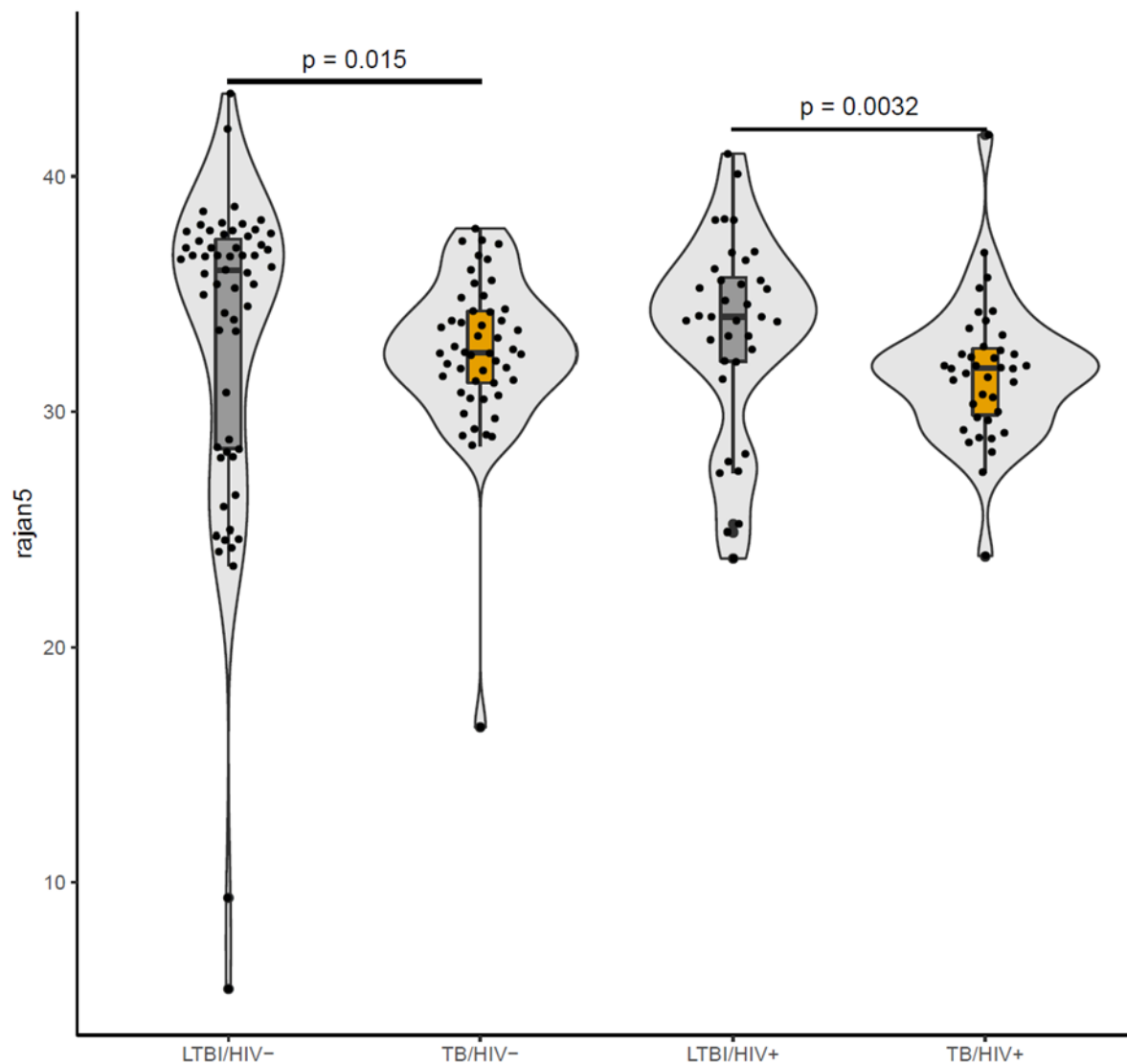


Figure 3.6: Signature score distribution for Rajan5.

LTBI = Latent TB, TB= active TB, HIV- = HIV negative, HIV+ = HIV positive. The violin plots show the distribution of scores across different categories of CTBC study participants. Each dot represents a study participant. P-values for comparison of signature scores across groups were calculated using the Mann–Whitney U test. The width of each inner box plot shows the IQR, the midline represents the median, and the whiskers are $1.5 \times \text{IQR}$.

Apart from Kaforou25 and Rajan5, the other set of signatures (Suliman 2, Suliman4, Thompson5, Roe3, Maertzdorf4, Penn-Nicholson6) that did not require reparameterization, were those developed by the SATVI group, on the Standard BioTools BioMark multiplex RT-qPCR platform, and previously evaluated in other research projects (Mendelsohn et al., 2022b; Penn-Nicholson et al., 2020b; Suliman et al., 2018).

3.3.1 Laboratory setup for reparameterization

Reparameterization was carried using gene expression data on whole blood samples collected from the study participants that were part of the Cross-sectional TB cohort (CTBC) (Darboe et al., 2019; Penn-Nicholson et al., 2020a). CTBC was selected for reparameterization because the cohort had samples from HIV positive and negative TB cases as well as asymptomatic community controls. The inclusion of both HIV positive and negative individuals, from this cohort, allowed for the models tested to be more generalisable. The use of an independent dataset for reparameterization also helped avoid bias.

The train and test set technique (Genç and Tunç, 2019) was used to evaluate the performance of the machine learning models for each signature that required reparameterization. Model training was done on individuals that were HIV negative (59 LTBI; 50 TB). 70% of HIV negative individuals (42 LTBI; 34 TB) made up the train set and the remaining 30% comprised the test set (17 LTBI; 16 TB). Once a model constructed on the training set demonstrated that it worked on the test set, then both sets were combined to reparametrize on the entire dataset (**Table 3.11**). ROC curves were generated to assess the performance of each model's ability to discriminate HIV positive cases (n=39) from HIV negative controls (n=36) (**Table 3.11**).

Table 3.11: Performance of signatures that required reparameterization

Signature	HIV-negative training and test set (combined) AUC (95% CI)	HIV-positive validation set AUC (95% CI)
DA-COSTA3	0.98 (0.93–1)	0.96 (0.93–1)
DUFFY9	TB vs LTBI: 1 (1–1)*	
FRANCISCO2	0.99 (0.98–1)	0.83 (0.74–0.93)
GJØEN7	0.98 (0.94–1)	0.78 (0.67–0.88)
JACOBSEN3	0.99 (0.99–1)	0.95 (0.89–1)
ROE3	0.96 (0.89–1)	0.89 (0.80–0.97)
SAMBAREY10	0.99 (0.96–1)	0.95 (0.90–1)
SATPROEDPRAI7	0.99 (0.98–1)	0.95 (0.90–1)

* HIV-positive samples required to parameterise the multinomial model. No independent validation set available.

3.3.2 Pass rate of signatures during reparameterization

Nineteen of the twenty signatures tested during reparameterization had a pass-rate of 80% and above in CTBC samples (**Figure 3.7**).

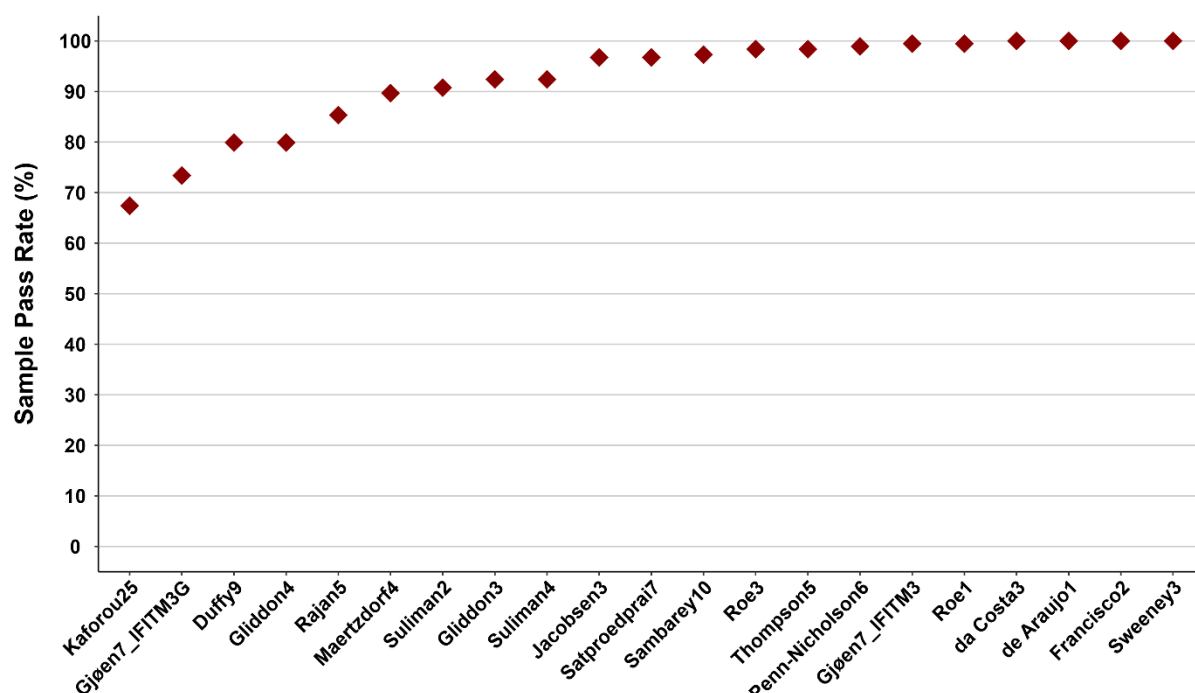


Figure 3.7: Sample pass rate of transcriptomic signatures during reparameterization.

All transcriptomic signatures on the panel, alongside the variations of different signatures generated during reparameterization are shown on the x-axis. On the y-axis, the pass-rate of each signature in samples from CTBC was expressed as a percentage.

The primer-probe assay, IFITM3_G that was used to quantify expression of the IFITM3 transcript in the Gjøen signature (named Gjøen7_IFITMG in **Figure 3.7**) was found to have a very high failure rate (**Figure 3.7** and **Table 3.12**). The second variation of the Gjøen7 signature resulted from substituting the original IFITM3 transcript (IFITM3_G), for another IFITM3 transcript (Hs03057129_s1) on the panel. Substituting IFITM3_G for IFITM3.Hs03057129_s1 greatly improved the pass-rate of the Gjøen signature [99.46% for Gjøen7IFITM3 vs. 73.37% for Gjøen7] (**Figure 3.7**).

Kaforou25, Duffy9, Gliddon4, and Rajan5 all had lower sample pass rates (**Figure 3.7**). This was a result of a high reaction failure rate for the GBP6 primer-probe assay. Kaforou25 having a lower pass rate than the Gjøen7 was because of the failure rate of other assays other than GBP6 and GAS6 (AR323W3_GAS6_NM_000820_10_11) that did not fully meet the assay qualification criteria.

Table 3.12: Sample pass rate during reparameterization for the selected 96 primer-probe assays

Primer-probe assay	Pass rate (%)
ACTA2.Hs00426835_g1	97.83
ACTR3.Hs01029159_g1	100
ANKRD22.Hs00944015_m1	97.28
APOL1.Hs01066280_m1	99.46
ARG1.Hs00968979_m1	98.91
BATF2.Hs00912737_m1	99.46
BCL6.BCL6_CUSTOM_AREPURD	100
BLK.Hs01017452_m1	98.91
C1QA.Hs00706358_s1	100
C1QB.Hs00608019_m1	96.2
C1QC.Hs00757779_m1	80.98
C4ORF18.FAM198B.Hs00259260_s1	100
C5.Hs00156197_m1	90.76
CCR6.Hs01890706_s1	96.2
CD160.Hs00199894_m1	98.37
CD1C.Hs00957534_g1	98.91
CD36.2.Hs01567186_m1	91.3
CD3E.CD3E_APDJ3NT	100
CD40L.Hs00163934_m1	98.91
CD74_j1.Hs04983808_s1	100
CD74_j2.CD74_J2_AP9HN47	100
CD79A.Hs00998120_g1	100
CD79B.Hs01058826_g1	100
CDCA7.Hs00230589_m1	17.39
CDKN1C.Hs00175938_m1	92.39
CXCR5.Hs00540548_s1	100
CYP4F3.Hs01587860_mH	99.46
DUSP3.Hs01115776_m1	100
F2RL1.Hs00608346_m1	88.04
FAM20A.Hs01034066_m1	88.04
FCGR1A.FCGR1A_SAT7_AP2XDMK	100
FCGR1A.Hs00174081_m1	100
FCGR1B.Hs02341825_m1	100
FCGR1B_VARIANT_1.FCGR1B_VARIANT1_AP3267H	100
FCGR1B_VARIANT_2.FCGR1B_VARIANT2_AP47ZTF	98.91
FCGR1C.Hs00417598_m1	100
FLVCR2.Hs00900390_m1	97.28
GAS6.GAS6_NM_000820_10_11_AR323W3	95.11
GAS6.GAS6_NM_000820_5_6_AR47XGZ	94.02
GBP1.Hs00977005_m1	97.83
GBP2.Hs00894846_g1	100
GBP5.Hs00369472_m1	100
GBP6.Hs01584201_m1	85.87
GNG7.Hs00192999_m1	99.46
GYG1.Hs00907542_g1	98.37
GZMA.Hs00989184_m1	100
HK3.Hs01092850_m1	100
ID3.Hs00954037_g1	98.91
IFI44L.Hs00915292_m1	99.46
IFITM3.Hs03057129_s1	100
IFITM3_G.IFITM3_G_APMF24C	73.91
KAZN.KAZN_TOTAL_SAT7_AP7DVDD	97.83

KIF1B.Hs01114512_g1	100
KLF2.Hs00360439_g1	100
KLF2.KLF2_CUSTOM_ARH6CF4	98.37
KLHDC8B.Hs00293902_m1	98.91
LAG3.Hs00958444_g1	100
LHFPL2.Hs00299613_m1	95.11
LTF.Hs00158924_m1	97.83
MAFB.MAFB_SAT7_APZTH2N	98.37
MAP7D3.Hs00226257_m1	98.37
MMP9.MMP9_APEPW9P	100
MPO.Hs00165162_m1	95.11
MTRF1L.Hs01097882_g1	99.46
NOD2.Hs01550759_g1	99.46
NPC2.Hs01119244_m1	100
OSBPL10.Hs00215016_m1	92.39
P2RY14.Hs01848195_s1	90.76
PRDM1.Hs00153357_m1	100
RAB13.Hs04400188_g1	100
Rab33A.Hs00191243_m1	98.37
RABL2A.Hs00255244_m1	99.46
RBBP8.Hs01090329_m1	98.37
RP11.295G20.2.Hs01373568_m1	100
S100A8.Hs00374264_g1	96.74
SCARF1.Hs01092483_m1	98.91
SDR39U1.Hs01016970_g1	100
SEPT4.Hs00910208_g1	95.65
SERPING1.Hs00934329_m1	99.46
SLPI.Hs00268206_m1	100
SMARCD3.Hs01088251_g1	100
STAT1.Hs01013996_m1	100
STT3A.Hs00967491_m1	99.46
TIMM10.TIMM10_CUSTOM_APT2DEC	100
TMBIM6.Hs00162661_m1	100
TMCC1.Hs01037666_s1	89.67
TNIP1.TNIP1_GEO_AP47ZYZ	100
TRMT2A.Hs01000041_g1	99.46
TUBGCP6.Hs00363509_g1	100
UCP2.Hs01075224_g1	100
USF2.Hs01100994_g1	100
VAMP5.Hs01105383_g1	100
WARS_j1.Hs00998737_m1	100
WARS_j2.WARS.J2_APAAFHX	99.46
ZDHHC19.Hs00376118_m1	94.02
ZNF296.Hs00377132_m1	95.11

Four assays within Kaforou25 failed in 9 out of the 184 samples from CTBC (**Table 3.13**) and 5 assays failed in more than 9 samples (**Figure 3.8**). Aside from GBP6, the biggest contributors to the low pass rate of Kaforou25 were C1QC.Hs00757779_m1 (22 failed reactions) and FAM20A.Hs01034066_m1 (35 failed reactions) (**Figure 3.8** and **Table 3.13**).

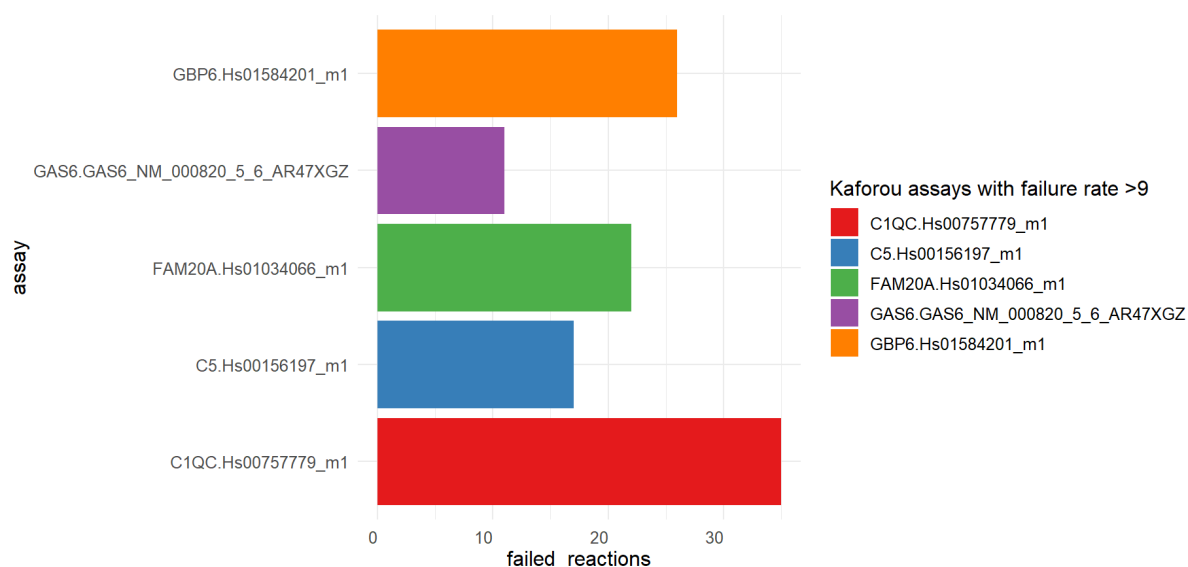


Figure 3.8: Evaluation of primer-probe assay performance for Kaforou25.

The bar plots show the 5 primer-probe assays that failed in more than 9 samples evaluated during reparameterization. The number of failed reactions for all 25 primer-probe assays are shown in **Table 3.13**.

Table 3.13: Number of failed reactions for Kaforou25 transcripts

ANKRD22.Hs00944015_m1	5
C1QB.Hs00608019_m1	7
C1QC.Hs00757779_m1	35
C4ORF18.FAM198B.Hs00259260_s1	0
C5.Hs00156197_m1	17
DUSP3.Hs01115776_m1	0
FAM20A.Hs01034066_m1	22
FCGR1A.Hs00174081_m1	0
FCGR1B.Hs02341825_m1	0
FCGR1C.Hs00417598_m1	0
FLVCR2.Hs00900390_m1	5
GAS6.GAS6_NM_000820_5_6_AR47XGZ	11
GAS6.GAS6_NM_000820_10_11_AR323W3	9
GBP6.Hs01584201_m1	26
LHFPL2.Hs00299613_m1	9
MPO.Hs00165162_m1	9
S100A8.Hs00374264_g1	6
SMARCD3.Hs01088251_g1	0
VAMP5.Hs01105383_g1	0
CCR6.Hs01890706_s1	7
CD79A.Hs00998120_g1	0
CD79B.Hs01058826_g1	0
CXCR5.Hs00540548_s1	0
GNG7.Hs00192999_m1	1
ZNF296.Hs00377132_m1	9

The sample pass rate of Kaforou25 can be greatly improved by excluding assays that frequently failed. Further modifications were therefore made to the Kaforou25 signature after assessing the sample pass rates of both the signature and each of primer-probe assays for its constituent transcripts in another study cohort (**Chapter 4**). This led to the signature eventually being assessed as Kaforou22 in the subsequent chapters (**Chapter 4 and Chapter 5**).

3.3.3 Conclusions from reparameterization

The modelling approach for generation of scores for each signature was the same as that applied in the publications where the original discovery of each signature was reported. The successful validation and identification of 96 primer-probe assays allowed us to evaluate all 20 transcriptomic signatures (**Table 3.14**). These transcriptomic signatures were evaluated in the study cohorts described in subsequent results chapters (**Chapter 4 and Chapter 5**).

Table 3.14: TaqMan PCR primer-probe assays for transcripts evaluated in this study

Gene Symbol	TaqMan Assay ID	Assay excluded**	Reference probe	# of signatures	da Costa3	de Araujo1	Duffy9 (10) [†]	Francisco2	GjØen7	Gliddon3	Gliddon4	Jacobsen3	Kaforou22 (27) [‡]	Maertzdorf4	Penn-Nicholson6	Rajan5	Roe1	Roe3	Sambarey10	Satproedprai7	Suliman2	Suliman4	Sweeney3	Thompson5
ACTR3	Hs01029159_g1		X	14	X	X	X	X	X	X	X	X	X			X	X	X	X	X				
TMBIM6	Hs00162661_m1		X	14	X	X	X	X	X	X	X	X	X			X	X	X	X	X				
USF2	Hs01100994_g1		X	14	X	X	X	X	X	X	X	X	X			X	X	X	X	X				
ACTA2	Hs00426835_g1			1												X								
ANKRD22	Hs00944015_m1			2									X								X			
APOL1	Hs01066280_m1			1																X				
ARG1	Hs00968979_m1			1							X													
BATF2	Hs00912737_m1			2													X	X						
BCL6	AREPURD*			1															X					
BLK	Hs01017452_m1			1																	X			
C1QA	Hs00706358_s1	X		0																				
C1QB	Hs00608019_m1			2			X			X			‡											
C1QC	Hs00757779_m1	X		0									‡											
C4ORF18 / FAM198B	Hs00259260_s1			1									X											
C5	Hs00156197_m1			1									X											
CCR6	Hs01890706_s1			1									X											
CD160	Hs00199894_m1			1			X																	
CD1C	Hs00957534_g1			1																	X			
CD36.2	Hs01567186_m1			1			X																	
CD3E	APDJ3NT*			1					X															
CD40L	Hs00163934_m1			1			X																	
CD74-j1	Hs04983808_s1	X		0																				
CD74-j2	AP9HN47*	X		0																				
CD79A	Hs00998120_g1			1									X											
CD79B	Hs01058826_g1			1									X											
CDCA7	Hs00230589_m1	X		0																				
CDKN1C	Hs00175938_m1	X		0																				
CXCR5	Hs00540548_s1			1									X											
CYP4F3	Hs01587860_mH			1															X					
DUSP3	Hs01115776_m1			2									X										X	
F2RL1	Hs00608346_m1	X		0																				
FAM20A	Hs01034066_m1			1									X											
FCGR1A	Hs00174081_m1			5	X					X		X	X						X					
FCGR1A	AP2XDMK*			1																X				
FCGR1B	Hs02341825_m1			3			X						X		X									
FCGR1B_VARIANT1	AP3267H*			1																X				
FCGR1B_VARIANT2	AP47ZTF*			1																X				
FCGR1C	Hs00417598_m1			1									X											
FLVCR2	Hs00900390_m1			1									X											
GAS6	AR47XGZ*			2									X									X		
GAS6	AR323W3*			1									X											
GBP1	Hs00977005_m1			1										X										
GBP2	Hs00894846_g1			1											X									
GBP5	Hs00369472_m1			5	X			X	X									X					X	
GBP6	Hs01584201_m1			3			X				X		‡			X								
GNG7	Hs00192999_m1			1									X											
GYG1	Hs00907542_g1			1												X								

Gene Symbol	TaqMan Assay ID	Assay excluded**	Reference probe	# of signatures	da Costa3	de Araujo1	Duffy9 (10)†	Francisco2	GjØen7	Gliddon3	Gliddon4	Jacobsen3	Kaforou22 (27)‡	Maertzdorf4	Penn-Nicholson6	Rajan5	Roe1	Roe3	Sambarey10	Satproedprai7	Suliman2	Suliman4	Sweeney3	Thompson5
GZMA	Hs00989184_m1			1	X																			
HK3	Hs01092850_m1			1															X					
ID3	Hs00954037_g1			2			X							X										
IFI44L	Hs00915292_m1			1															X					
IFITM3_G	APMF24C*	X		0																				
IFITM3	Hs03057129_s1			2					X					X										
KAZN	AP7DVDD*			1																X				
KIF1B	Hs01114512_g1			1					X															
KLF2	Hs00360439_g1			2				X															X	
KLF2	ARH6CF4*	X		0																				
KLHDC8B	Hs00293902_m1	X		0																				
LAG3	Hs00958444_g1			1			X																	
LHFPL2	Hs00299613_m1			1									X											
LTF	Hs00158924_m1			1								X												
MAFB	APZTH2N*			1																X				
MAP7D3	Hs00226257_m1			1																				X
MMP9	APEPW9P*			1					X															
MPO	Hs00165162_m1			1									X											
MTRF1L	Hs01097882_g1			1												X								
NOD2	Hs01550759_g1			1					X															
NPC2	Hs01119244_m1			1		X																		
OSBPL10	Hs00215016_m1			1																	X			
P2RY14	Hs01848195_s1			1										X										
PRDM1	Hs00153357_m1			1							X													
RAB13	Hs04400188_g1			1															X					
Rab33A	Hs00191243_m1			1								X												
RABL2A	Hs00255244_m1			1												X								
RBBP8	Hs01090329_m1			1															X					
RP11-295G20.2	Hs01373568_m1			1																				X
S100A8	Hs00374264_g1			1									X											
SCARF1	Hs01092483_m1			1														X						
SDR39U1	Hs01016970_g1			1											X									
SEPT4	Hs00910208_g1			1																	X			
SERPING1	Hs00934329_m1			1											X									
SLPI	Hs00268206_m1			1															X					
SMARCD3	Hs01088251_g1			3									X						X					X
STAT1	Hs01013996_m1			1																X				
STT3A	Hs00967491_m1			1																				X
TIMM10	APT2DEC*			1															X					
TMCC1	Hs01037666_s1			1							X													
TNIP1	AP47ZY*			1					X															
TRMT2A	Hs01000041_g1			1											X									
TUBGCP6	Hs00363509_g1			1											X									
UCP2	Hs01075224_g1			1																				X
VAMP5	Hs01105383_g1			1									X											
WARS-j1	Hs00998737_m1	X		0																				
WARS-j2	APAAFHX*	X		0																				
ZDHC19	Hs00376118_m1			1			X																	
ZNF296	Hs00377132_m1			2						X			X											

Signatures are named by first author and number of transcripts included in the model. Numbers in brackets indicate the original number of transcripts in the published model. Some signatures have a reduced number of transcripts when translated to RT-qPCR due to duplicate transcript symbols (IDs), high-primer probe failure rate (**poor amplification

efficiency), or where transcript sequences from an original discovery cohorts could not be mapped to a more recent reference transcriptome.

* Custom designed primer-probe assays; available on request to corresponding author.

† The CERKL1 primer-probe assay failed during panel optimisation and was removed from the Duffy10 signature.

‡ Three transcripts (C1QB, C1QC, and GBP6) were excluded from the Kaforou27 signature due to a high primer-probe assay failure rate (**). Unique primer-probe assays could not be designed for both of the FCGR1B transcript variants; only one FCGR1B primer-probe was included in the Kaforou22 model. We were unable to design a primer-probe for LOC728744; all information had been withdrawn from both the NCBI and RefSeq databases.

3.4: Development of an internal positive control (IPC)

Positive controls form an integral part of RT-qPCR experiments as they are useful in detecting any false positives, confirming that DNA amplification has occurred successfully (Kavlick, 2018), identifying the presence of any PCR inhibitors and determining that all reagents used during every microfluidic RT-qPCR chip run yielded an expected, reproducible result (Moldovan and Moldovan, 2020). Since the IPC formed part of the main parameters evaluated in the QC script (**Section 2.4 of Chapter 2**), careful consideration was put in when developing one.

More than one IPC was developed, evaluated, and selected, alongside the reparameterization exercise, using samples from CTBC. Since RISK6 (Penn-Nicholson et al., 2020a) scores had been generated already for CTBC study participants (Penn-Nicholson et al., 2020a), these scores became the basis on which samples were selected to generate different IPCs.

RISK6 scores were put into three categories that is: low (0-40), medium (40-60) and high scores (60-100). Three IPCs were generated that were classified as IPC_LOW, IPC_MEDIUM, and IPC_HIGH, based on the RISK6 scores for the samples used to make up each of them.

Stored cDNA from CTBC samples were pre-amplified with all 96 transcripts of interest (**as described in Chapter 2: Section 2.3.2**). Following successful pre-amplification and PCR confirmation steps, five samples each with either low, medium, or high RISK6 scores were identified and combined to generate the 3 individual IPCs, namely IPC_LOW, IPC_MEDIUM, and IPC_HIGH. All three IPCs were evaluated on the same microfluidic chips.

3.4.1 Criteria for selecting an IPC

- All primer-probe assays passed, meaning that the IPC had to have Ct values for all 96 transcripts of interest.
- Consistency in RISK6 scores, assessed using the concordance correlation co-efficient (CCC) of each IPC across chips. The cut-off for CCC was set at 0.95.

None of the three IPCs had 100% pass rate for all primer-probe assays. This was a result of primer-probe assays for the CDCA7 and IFITM3_G transcripts, that had irregular gene expression results for all three microfluidic chip runs where the performance of IPCs was assessed (**Table 3.15**). This demonstrated that the lack of a 100% pass-rate was not particularly an issue of the IPCs, but rather the primer-probe assays themselves. Due to the lack of an IPC with 100% pass rate for all primer-probe assays, the choice of an IPC was based more on the evaluation of RISK6 scores and how concordant the Ct values for transcripts were for the IPCs across chips.

Table 3.15: Primer-probe assays for each IPC with gene expression results that did not pass quality control

	IPC_LOW	IPC_MEDIUM	IPC_HIGH
No expression	C1QC & IFITM3_G	CDCA7	CDCA7
Flag	CDCA7	CDCA7 & IFITM3_G	CDCA7

No expression means there was no amplification and thus no Ct value measured (**Chapter 2 Section 2.3.4**). Flag means that a Ct value was generated, but the amplification curve was outside the quality threshold range set on the BioMark microfluidic RT-qPCR system (**Figure 3.9**).

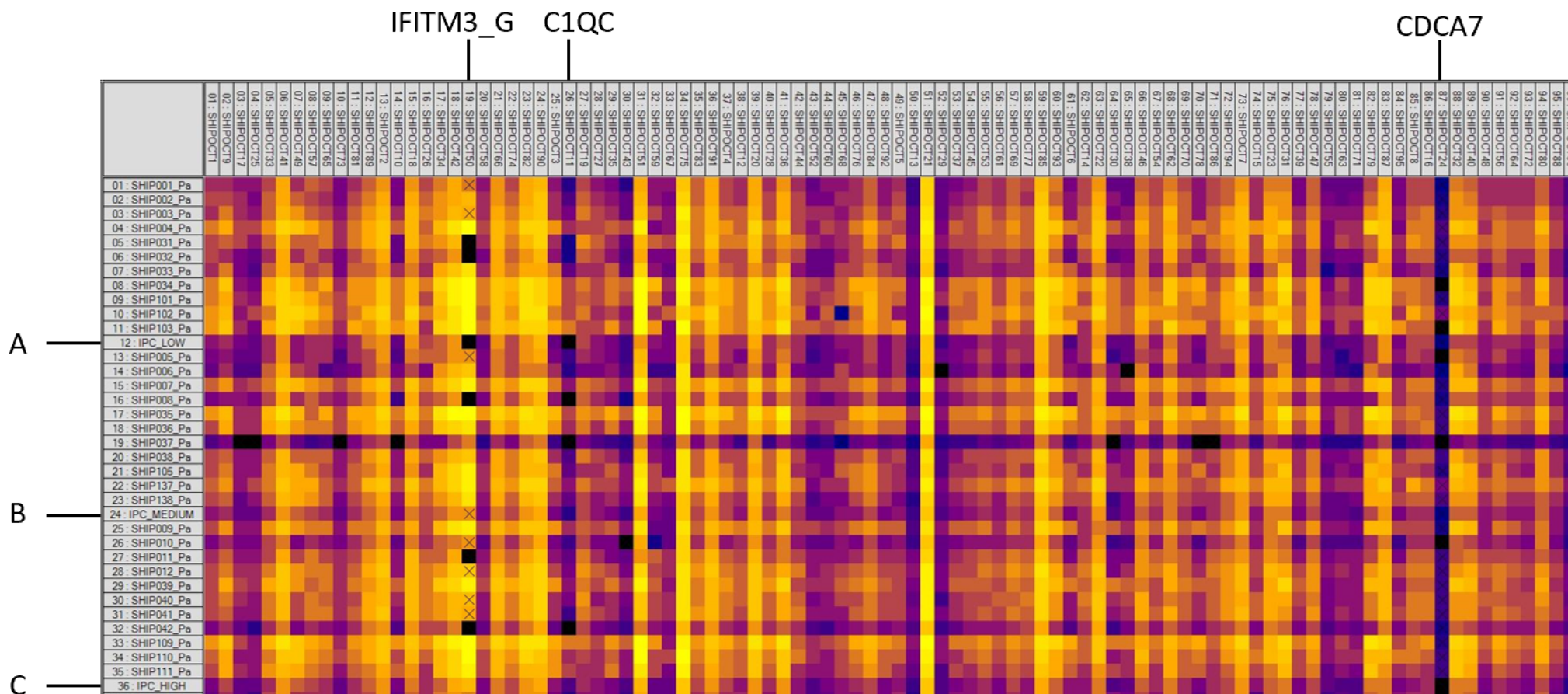


Figure 3.9: Part of the heatmap for reparameterization run1 on CTBC samples.

This snippet of a heatmap demonstrates gene expression results for all 96 primer/probe assays tested within the IPC_LOW (A), IPC_MEDIUM(B) and IPC_LOW(C). The samples are indicated on the left panel and each primer/probe assay shown above. An X marks a flagged gene expression result and black signified no gene expression present. Black would be a Ct value <1.00 or >40.00 that the BioMark microfluidic platform could not verify.

Both IPC_LOW (**Figure 3.10**) and IPC_MEDIUM (**Figure 3.11**) had higher CCCs for Ct values across chips, in comparison to IPC_HIGH (**Figure 3.12**). Regarding RISK6 scores, all IPCs showed variations across chip runs (**Figure 3.13**). Previous studies conducted at SATVI have demonstrated that IPCs generated from whole blood samples of study participants with high RISK6 scores do not show any variation in results even when there may be a technical issue. IPCs comprised of whole blood samples from study participants with RISK6 scores of 50 and below were therefore preferred as these are more likely to pick up on technical variations between runs. IPC_MEDIUM was therefore selected for use in all research projects reported in this thesis (**Chapters 4 and 5**). IPC_MEDIUM also had two of CCCs higher than IPC_LOW.

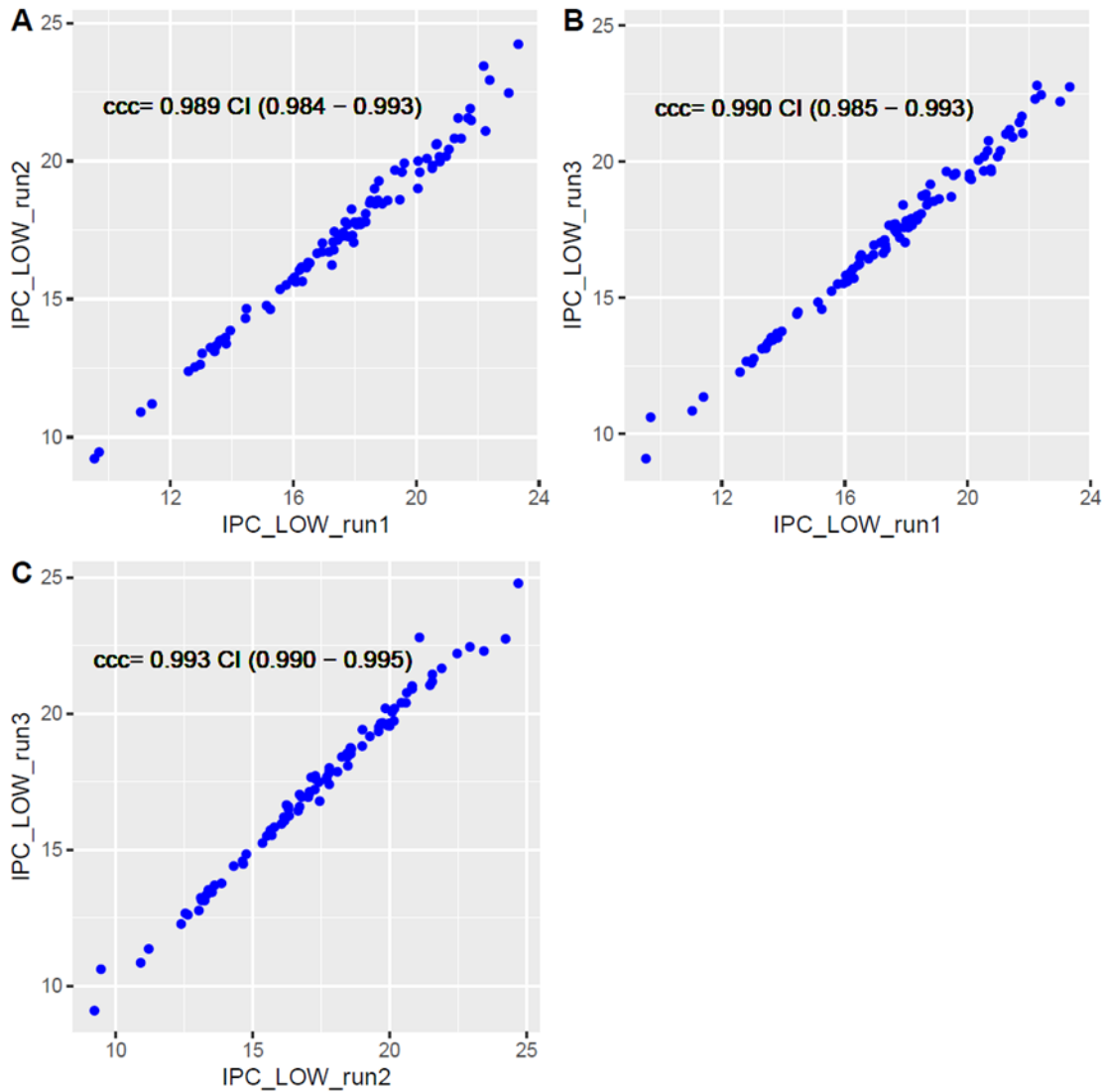


Figure 3.10: Concordance correlation co-efficients for IPC_LOW.

The scatter plots show Lin's Concordance correlation coefficient for Ct values for all 96 transcripts per chip run during reparameterization. (a) shows CCC for Ct values between IPC_LOW in run1 and run2, (b) shows CCC for Ct values between IPC_LOW in run3 and run1, and (c) shows CCC for Ct values between IPC_LOW in run3 and run2.

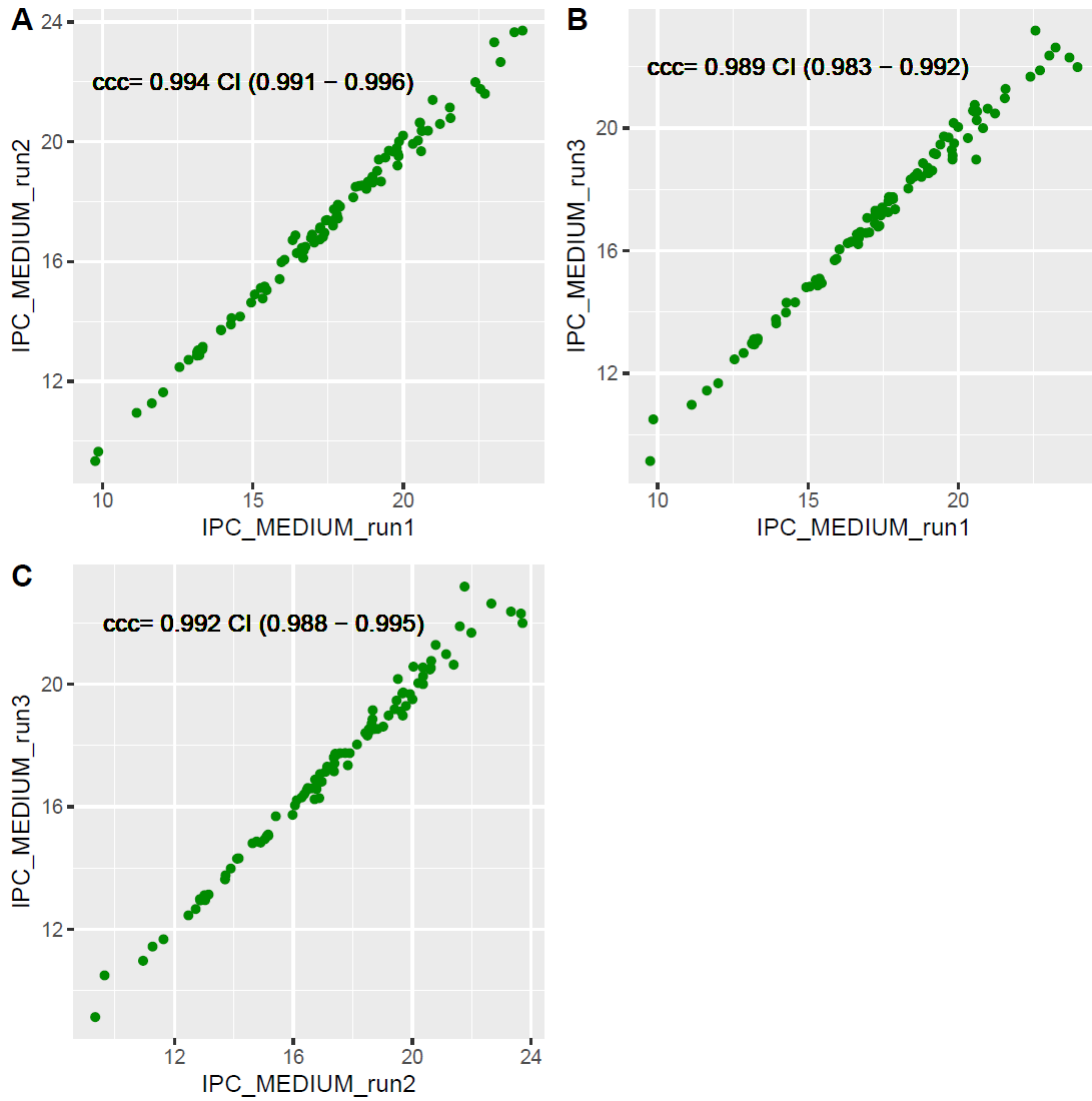


Figure 3.11: Concordance correlation coefficients for IPC_MEDIUM.

The scatter plots show Lin's Concordance correlation coefficient for Ct values for all 96 transcripts per microfluidic chip run during reparameterization. (a) shows CCC for Ct values between IPC_MEDIUM in run1 and run2, (b) shows CCC for Ct values between IPC_MEDIUM in run3 and run1, and (c) shows CCC for Ct values between IPC_MEDIUM in run3 and run2.

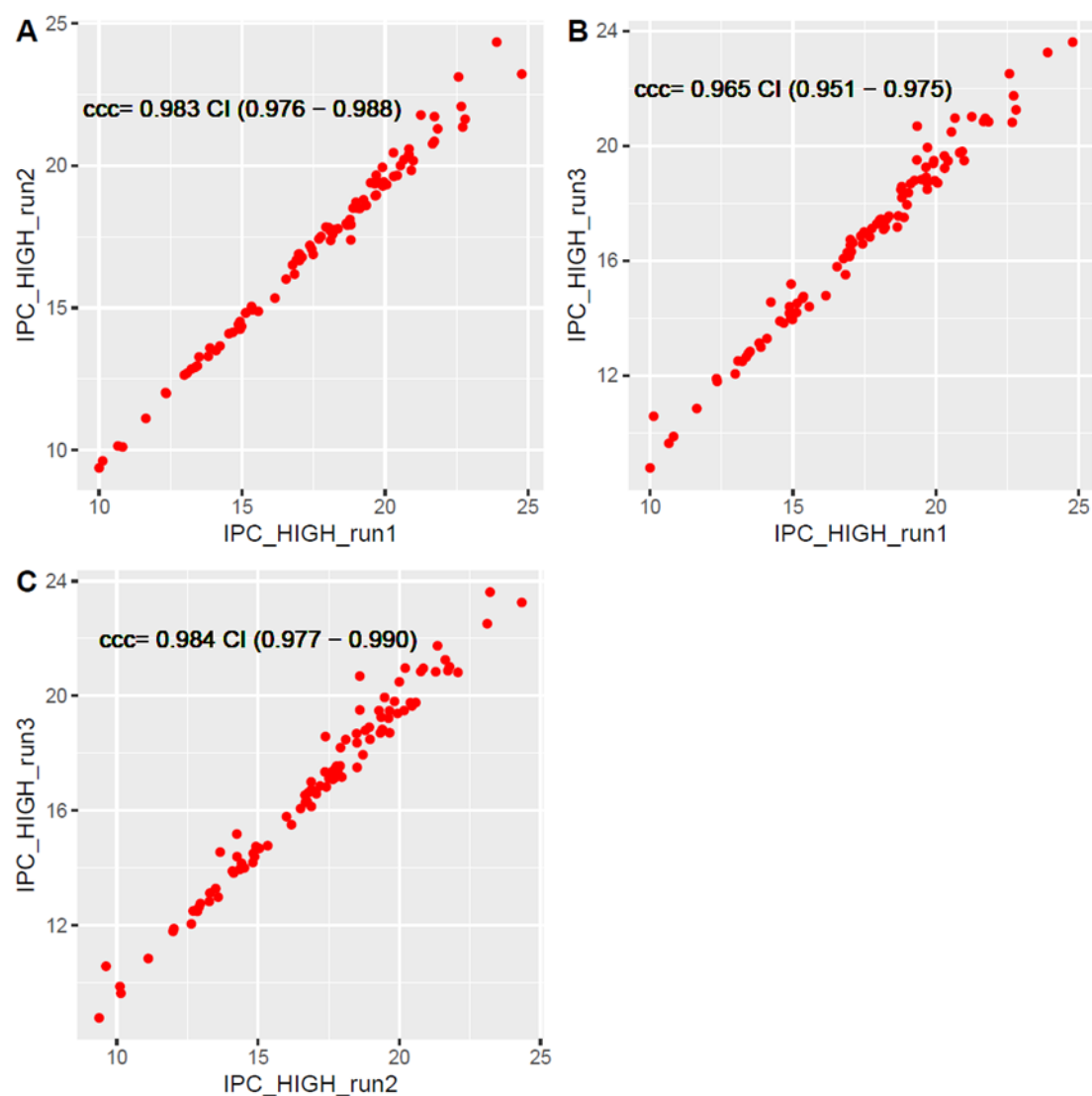


Figure 3.12: Concordance correlation coefficients for IPC_HIGH.

The scatter plots show Lin's Concordance correlation coefficient for Ct values for all 96 transcripts per chip run during reparameterization. (a) shows CCC for Ct values between IPC_HIGH in run1 and run2, (b) shows CCC for Ct values between IPC_HIGH in run3 and run1, and (c) shows CCC for Ct values between IPC_HIGH in run3 and run2.

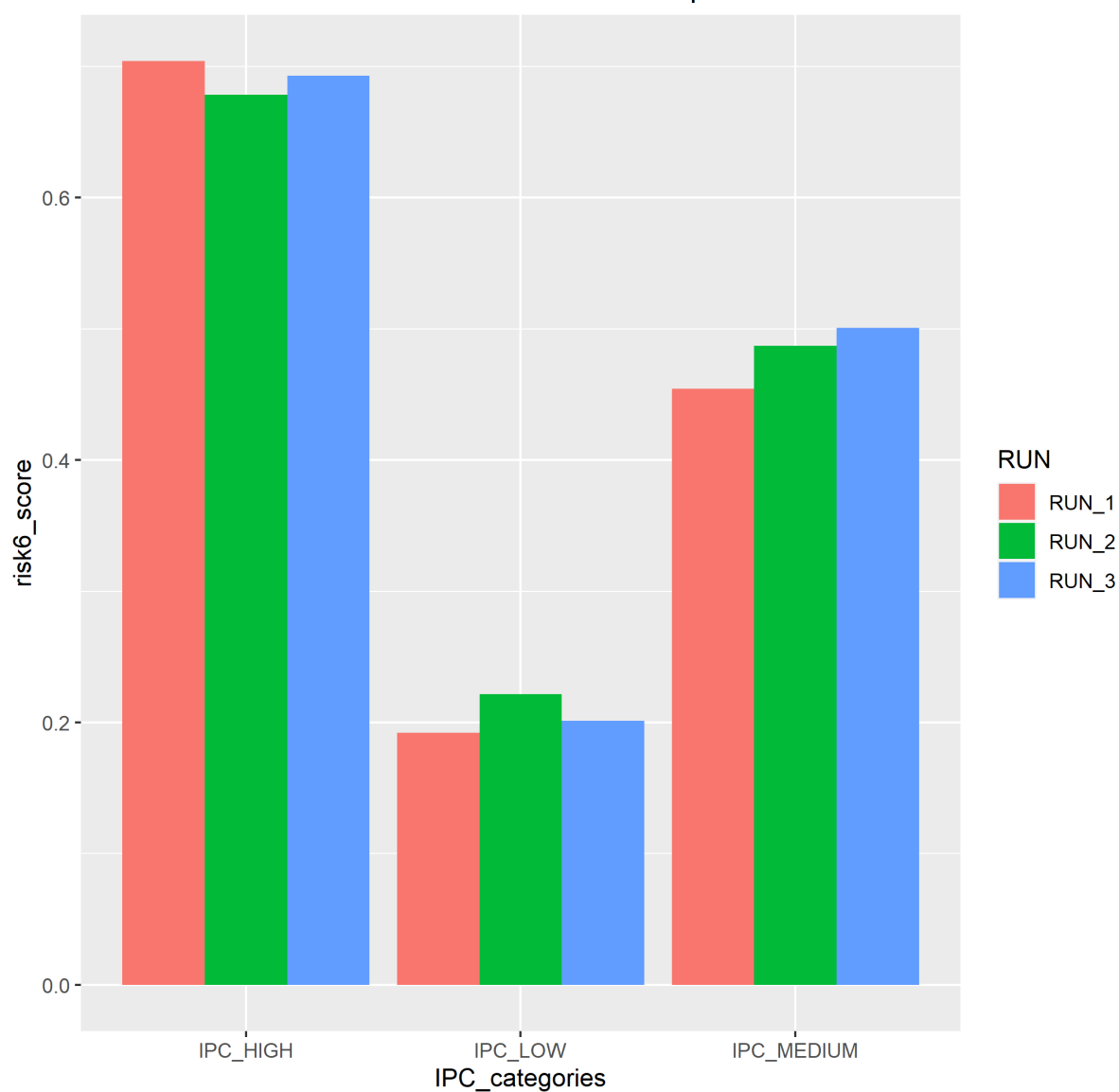


Figure 3.13: Evaluation of RISK6 scores in IPCs.

The bar plots show the variation in RISK6 scores across different chips for reparametization. Each reparameterization run represents RISK6 results from a different Fluidigm chip. The same 96 primer probe assays were assessed on every run. RUN1 and RUN2 were repeat runs carried out on the same set of CTBC samples, while RUN3 was carried out on a different set of samples.

3.5: Contributions

I identified transcriptomic signatures to complete the panel of 96 primer-probe assays, set up the experiments, and performed the laboratory work for assay qualification and reparameterization. I generated, verified, and interpreted the gene expression data. I generated the code in R adapted from analysis pipelines set up from previous projects at SATVI. I carried out the data analysis for all steps involved in the optimization process.

The initial selection of signatures for the panel was done by Prof. Thomas J. Scriba and Dr. Simon C. Mendelsohn. The identification of possible primer-probe assays and the design of custom primer-probe assays was carried out by me, Dr. Simon C. Mendelsohn, and Dr. Stanley K. Mbandi. Dr. Simon C. Mendelsohn designed the approach for validation and selection of HKGs. The code for assay qualification was originally developed by Dr. Simon C. Mendelsohn and then modified by me and Dr. Stanley K. Mbandi to meet the objectives of this PhD project. The statistical methodology and code for reparameterization of transcriptomic signatures was designed by Dr. Stanley K. Mbandi.

References

- Andersen, C.L., Jensen, J.L., Ørntoft, T.F., 2004. Normalization of Real-Time Quantitative Reverse Transcription-PCR Data: A Model-Based Variance Estimation Approach to Identify Genes Suited for Normalization, Applied to Bladder and Colon Cancer Data Sets. *Cancer Res.* 64, 5245–5250. <https://doi.org/10.1158/0008-5472.CAN-04-0496>
- Anderson, S.T., Kaforou, M., Brent, A.J., Wright, V.J., Banwell, C.M., Chagaluka, G., Crampin, A.C., Dockrell, H.M., French, N., Hamilton, M.S., Hibberd, M.L., Kern, F., Langford, P.R., Ling, L., Mlotha, R., Ottenhoff, T.H.M., Pienaar, S., Pillay, V., Scott, J.A.G., Twahir, H., Wilkinson, R.J., Coin, L.J., Heyderman, R.S., Levin, M., Eley, B., 2014. Diagnosis of Childhood Tuberculosis and Host RNA Expression in Africa. *N. Engl. J. Med.*
- da Costa, L.L., Delcroix, M., Dalla Costa, E.R., Prestes, I.V., Milano, M., Francis, S.S., Unis, G., Silva, D.R., Riley, L.W., Rossetti, M.L.R., 2015. A real-time PCR signature to discriminate between tuberculosis and other pulmonary diseases. *Tuberc. Edinb. Scotl.* 95, 421–425. <https://doi.org/10.1016/j.tube.2015.04.008>
- Darboe, F., Mbandi, S.K., Naidoo, K., Yende-Zuma, N., Lewis, L., Thompson, E.G., Duffy, F.J., Fisher, M., Filander, E., van Rooyen, M., Bilek, N., Mabwe, S., McKinnon, L.R., Chegou, N., Loxton, A., Walzl, G., Tromp, G., Padayatchi, N., Govender, D., Hatherill, M., Karim, S.A., Zak, D.E., Penn-Nicholson, A., Scriba, T.J., Team, T.S.C.I., 2019. Detection of Tuberculosis Recurrence, Diagnosis and Treatment Response by a Blood Transcriptomic Risk Signature in HIV-Infected Persons on Antiretroviral Therapy. *Front. Microbiol.* 10. <https://doi.org/10.3389/fmicb.2019.01441>
- de Araujo, L.S., Vaas, L.A.I., Ribeiro-Alves, M., Geffers, R., Mello, F.C.Q., de Almeida, A.S., Moreira, A. da S.R., Kritski, A.L., Lapa e Silva, J.R., Moraes, M.O., Pessler, F., Saad, M.H.F., 2016. Transcriptomic Biomarkers for Tuberculosis: Evaluation of DOCK9, EPHA4, and NPC2 mRNA Expression in Peripheral Blood. *Front. Microbiol.* 7.
- Dominguez, M.H., Chattopadhyay, P.K., Ma, S., Lamoreaux, L., McDavid, A., Finak, G., Gottardo, R., Koup, R.A., Roederer, M., 2013. Highly multiplexed quantitation of gene expression on single cells. *J. Immunol. Methods* 391, 10.1016/j.jim.2013.03.002. <https://doi.org/10.1016/j.jim.2013.03.002>
- Duffy, F.J., Thompson, E.G., Scriba, T.J., Zak, D.E., 2019. Multinomial modelling of TB/HIV co-infection yields a robust predictive signature and generates hypotheses about the HIV+TB+ disease state. *PLOS ONE* 14, e0219322. <https://doi.org/10.1371/journal.pone.0219322>
- Eisenberg, E., Levanon, E.Y., 2013. Human housekeeping genes, revisited. *Trends Genet.* 29, 569–574. <https://doi.org/10.1016/j.tig.2013.05.010>
- Francisco, N.M., Fang, Y.-M., Ding, L., Feng, S., Yang, Y., Wu, M., Jacobs, M., Ryffel, B., Huang, X., 2017. Diagnostic accuracy of a selected signature gene set that discriminates active pulmonary tuberculosis and other pulmonary diseases. *J. Infect.* 75, 499–510. <https://doi.org/10.1016/j.jinf.2017.09.012>
- Genç, B., Tunç, H., 2019. Optimal training and test sets design for machine learning. *Turk. J. Electr. Eng. Comput. Sci.* 1534–1545. <https://doi.org/10.3906/elk-1807-212>
- Gjøn, J.E., Jenum, S., Sivakumaran, D., Mukherjee, A., Macaden, R., Kabra, S.K., Lodha, R., Ottenhoff, T.H.M., Haks, M.C., Doherty, T.M., Ritz, C., Grewal, H.M.S., 2017. Novel transcriptional signatures for sputum-independent diagnostics of tuberculosis in children. *Sci. Rep.* 7, 1–9. <https://doi.org/10.1038/s41598-017-05057-x>
- Gliddon, H.D., Kaforou, M., Alikian, M., Habgood-Coote, D., Zhou, C., Oni, T., Anderson, S.T., Brent, A.J., Crampin, A.C., Eley, B., Heyderman, R., Kern, F., Langford, P.R., Ottenhoff, T.H.M., Hibberd, M.L., French, N., Wright, V.J., Dockrell, H.M., Coin, L.J., Wilkinson, R.J., Levin, M., 2021. Identification of Reduced Host Transcriptomic Signatures for Tuberculosis Disease and Digital PCR-Based Validation and Quantification. *Front. Immunol.* 12.
- Gupta, R.K., Turner, C.T., Venturini, C., Esmail, H., Rangaka, M.X., Copas, A., Lipman, M., Abubakar, I., Noursadeghi, M., 2020. Concise whole blood transcriptional signatures for incipient

- tuberculosis: a systematic review and patient-level pooled meta-analysis. *Lancet Respir. Med.* 0. [https://doi.org/10.1016/S2213-2600\(19\)30282-6](https://doi.org/10.1016/S2213-2600(19)30282-6)
- Jacobsen, M., Repsilber, D., Gutschmidt, A., Neher, A., Feldmann, K., Mollenkopf, H.J., Ziegler, A., Kaufmann, S.H.E., 2007. Candidate biomarkers for discrimination between infection and disease caused by *Mycobacterium tuberculosis*. *J. Mol. Med.* 85, 613–621. <https://doi.org/10.1007/s00109-007-0157-6>
- Kaforou, M., Wright, V.J., Oni, T., French, N., Anderson, S.T., Bangani, N., Banwell, C.M., Brent, A.J., Crampin, A.C., Dockrell, H.M., Eley, B., Heyderman, R.S., Hibberd, M.L., Kern, F., Langford, P.R., Ling, L., Mendelson, M., Ottenhoff, T.H., Zgambo, F., Wilkinson, R.J., Coin, L.J., Levin, M., 2013. Detection of Tuberculosis in HIV-Infected and -Uninfected African Adults Using Whole Blood RNA Expression Signatures: A Case-Control Study. *PLOS Med.* 10, e1001538. <https://doi.org/10.1371/journal.pmed.1001538>
- Kavlick, M.F., 2018. Development of a universal internal positive control. *BioTechniques* 65, 275–280. <https://doi.org/10.2144/btn-2018-0034>
- Maertzdorf, J., McEwen, G., Weiner, J., Tian, S., Lader, E., Schriek, U., Mayanja-Kizza, H., Ota, M., Kenneth, J., Kaufmann, S.H., 2016. Concise gene signature for point-of-care classification of tuberculosis. *EMBO Mol. Med.* 8, 86–95. <https://doi.org/10.15252/emmm.201505790>
- Mendelsohn, S.C., Mbandi, S.K., Fiore-Gartland, A., Penn-Nicholson, A., Musvosvi, M., Mulenga, H., Fisher, M., Hadley, K., Erasmus, M., Nombida, O., Tameris, M., Walzl, G., Naidoo, K., Churchyard, G., Hatherill, M., Scriba, T.J., 2022. Prospective multicentre head-to-head validation of host blood transcriptomic biomarkers for pulmonary tuberculosis by real-time PCR. *Commun. Med.* 2, 26. <https://doi.org/10.1038/s43856-022-00086-8>
- Mendelsohn, S.C., Verhage, S., Mulenga, H., Scriba, T.J., Hatherill, M., 2023. Systematic review of diagnostic and prognostic host blood transcriptomic signatures of tuberculosis disease in people living with HIV. *Gates Open Res.* 7, 27. <https://doi.org/10.12688/gatesopenres.14327.1>
- Moldovan, E., Moldovan, V., 2020. Controls in Real-Time Polymerase Chain Reaction Based Techniques. *Acta Marisiensis - Ser. Medica* 66, 79–82. <https://doi.org/10.2478/amma-2020-0025>
- Mulenga, H., Zauchenberger, C.-Z., Bunyasi, E.W., Mbandi, S.K., Mendelsohn, S.C., Kagina, B., Penn-Nicholson, A., Scriba, T.J., Hatherill, M., 2020. Performance of diagnostic and predictive host blood transcriptomic signatures for Tuberculosis disease: A systematic review and meta-analysis. *PLOS ONE* 15, e0237574. <https://doi.org/10.1371/journal.pone.0237574>
- Penn-Nicholson, A., Mbandi, S.K., Thompson, E., Mendelsohn, S.C., Suliman, S., Chegou, N.N., Malherbe, S.T., Darboe, F., Erasmus, M., Hanekom, W.A., Bilek, N., Fisher, M., Kaufmann, S.H.E., Winter, J., Murphy, M., Wood, R., Morrow, C., Van Rhijn, I., Moody, B., Murray, M., Andrade, B.B., Sterling, T.R., Sutherland, J., Naidoo, K., Padayatchi, N., Walzl, G., Hatherill, M., Zak, D., Scriba, T.J., 2020a. RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response. *Sci. Rep.* 10, 8629. <https://doi.org/10.1038/s41598-020-65043-8>
- Penn-Nicholson, A., Mbandi, S.K., Thompson, E., Mendelsohn, S.C., Suliman, S., Chegou, N.N., Malherbe, S.T., Darboe, F., Erasmus, M., Hanekom, W.A., Bilek, N., Fisher, M., Kaufmann, S.H.E., Winter, J., Murphy, M., Wood, R., Morrow, C., Van Rhijn, I., Moody, B., Murray, M., Andrade, B.B., Sterling, T.R., Sutherland, J., Naidoo, K., Padayatchi, N., Walzl, G., Hatherill, M., Zak, D., Scriba, T.J., 2020b. RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response. *Sci. Rep.* 10, 8629. <https://doi.org/10.1038/s41598-020-65043-8>
- Pfaffl, M.W., Tichopad, A., Prgomet, C., Neuvians, T.P., 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pair-wise correlations. *Biotechnol. Lett.* 26, 509–515. <https://doi.org/10.1023/B:BILE.0000019559.84305.47>

- Rajan, J.V., Semitala, F.C., Mehta, T., Seielstad, M., Montalvo, L., Andama, A., Asege, L., Nakaye, M., Katende, J., Mwebe, S., Kanya, M.R., Yoon, C., Cattamanchi, A., 2019. A Novel, 5-Transcript, Whole-blood Gene-expression Signature for Tuberculosis Screening Among People Living With Human Immunodeficiency Virus. *Clin. Infect. Dis.* 69, 77–83. <https://doi.org/10.1093/cid/ciy835>
- Roe, J., Venturini, C., Gupta, R.K., Gurry, C., Chain, B.M., Sun, Y., Southern, J., Jackson, C., Lipman, M.C., Miller, R.F., Martineau, A.R., Abubakar, I., Noursadeghi, M., 2019. Blood Transcriptomic Stratification of Short-term Risk in Contacts of Tuberculosis. *Clin. Infect. Dis.* ciz252. <https://doi.org/10.1093/cid/ciz252>
- Roe, J.K., Thomas, N., Gil, E., Best, K., Tsaliki, E., Morris-Jones, S., Stafford, S., Simpson, N., Witt, K.D., Chain, B., Miller, R.F., Martineau, A., Noursadeghi, M., 2016. Blood transcriptomic diagnosis of pulmonary and extrapulmonary tuberculosis. *JCI Insight* 1. <https://doi.org/10.1172/jci.insight.87238>
- Sambarey, A., Devaprasad, A., Mohan, A., Ahmed, A., Nayak, S., Swaminathan, S., D'Souza, G., Jesuraj, A., Dhar, C., Babu, S., Vyakarnam, A., Chandra, N., 2017. Unbiased Identification of Blood-based Biomarkers for Pulmonary Tuberculosis by Modeling and Mining Molecular Interaction Networks. *eBioMedicine* 15, 112–126. <https://doi.org/10.1016/j.ebiom.2016.12.009>
- Satproedprai, N., Wichukchinda, N., Suphankong, S., Inunchot, W., Kuntima, T., Kumpeerasart, S., Wattanapokayakit, S., Nedsuwan, S., Yanai, H., Higuchi, K., Harada, N., Mahasirimongkol, S., 2015. Diagnostic value of blood gene expression signatures in active tuberculosis in Thais: a pilot study. *Genes Immun.* 16, 253–260. <https://doi.org/10.1038/gene.2015.4>
- Scriba, T.J., Fiore-Gartland, A., Penn-Nicholson, A., Mulenga, H., Mbandi, S.K., Borate, B., Mendelsohn, S.C., Hadley, K., Hikvam, C., Kaskar, M., Musvosvi, M., Bilek, N., Self, S., Sumner, T., White, R.G., Erasmus, Mzwandile, Jaxa, L., Raphela, R., Innes, C., Brumskine, W., Hiemstra, Andriette, Malherbe, S.T., Hassan-Moosa, R., Tameris, M., Walzl, G., Naidoo, K., Churchyard, G., Hatherill, M., Baepanye, K., Baepanye, T., Clarke, K., Collignon, M., Dlamini, A., Eyre, C., Feni, T., Fikizolo, M., Galane, P., Goliath, T., Gangat, A., Malefo-Grootboom, S., Rensburg, E.J. van, Rensburg, B.J. van, Kekana, S., Zietsman, M., Kock, A., Kunene, I., Lakhi, A., Langa, N., Ledwaba, H., Lophoko, M., Mabasa, I., Mabe, D., Mabuza, N., Majola, M., Makhetha, M., Makoanyane, M., Makhubalo, B., Malay, V., Market, J., Matshego, S., Mbipa, N., Mmotsa, T., Modipa, S., Mopati, S., Moswegu, P., Mothaga, P., Muller, D., Nchwe, G., Nel, M., Nhlangulela, L., Ntamo, B., Ntoahae, L., Ntshauba, T., Sanyaka, N., Seabela, L., Selepe, P., Senne, M., Serake, M.G., Thlapi, M., Tshikovhi, V., Tswaile, L., Aswegen, A. van, Mbata, L., Takavamanya, C., Pinho, P., Mdlulu, J., Taljaard, M., Slabbert, N., Sayed, S., Nielson, T., Senne, M., Sein, N.N., Mbata, L., Govender, D., Chinappa, T., Zulu, M.I., Maphanga, N.B., Hlathi, S.R., Gumede, G.K., Shezi, T.Y., Maphanga, J.L., Jali, Z.P., Cwele, T., Gwamanda, N.Z.E., Dlamini, C., Sing, Z.P.P., Ntanjana, N.G., Nzimande, S.S., Mbatha, S., Maharaj, B., Moosa, A., Corris, C.-M., Kafaar, F., Geldenhuys, H., Luabeya, A.K.K., Shenje, J., Botes, N., Rossouw, S., Africa, H., Diamond, B., Braaf, S., Stryers, S., Carstens, A., Jansen, R., Mabwe, S., Mulenga, H., Herling, R., Veldsman, A., Makhete, L., Steyn, M., Buhlungu, S., Erasmus, Margareth, Davids, I., Plaatjie, P., Companie, A., Ratangee, F., Veldtsman, H., Petersen, C., Abrahams, C., Moses, M., Kelepu, X., Gregg, Y., Swanepoel, L., Magawu, N., Cetywayo, N., Mactavie, L., Valley, H., Filander, E., Nqakala, N., Maasdorp, E., Khoury, J., Kriel, B., Smith, B., Muller, L., Tonsing, S., Loxton, A., Hiemstra, Andriette, Ahlers, P., Flinn, M., Chung, E., Chung, M., Sato, A., 2021. Biomarker-guided tuberculosis preventive therapy (CORTIS): a randomised controlled trial. *Lancet Infect. Dis.* 21, 354–365. [https://doi.org/10.1016/S1473-3099\(20\)30914-2](https://doi.org/10.1016/S1473-3099(20)30914-2)
- Silver, N., Best, S., Jiang, J., Thein, S.L., 2006a. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol. Biol.* 7, 33. <https://doi.org/10.1186/1471-2199-7-33>

- Silver, N., Best, S., Jiang, J., Thein, S.L., 2006b. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol. Biol.* 7, 33.
<https://doi.org/10.1186/1471-2199-7-33>
- Suliman, S., Thompson, E.G., Sutherland, J., Weiner, J., Ota, M.O.C., Shankar, S., Penn-Nicholson, A., Thiel, B., Erasmus, M., Maertzdorf, J., Duffy, F.J., Hill, P.C., Hughes, E.J., Stanley, K., Downing, K., Fisher, M.L., Valvo, J., Parida, S.K., van der Spuy, G., Tromp, Gerard, Adetifa, I.M.O., Donkor, S., Howe, R., Mayanja-Kizza, H., Boom, W.H., Dockrell, H.M., Ottenhoff, T.H.M., Hatherill, M., Aderem, A., Hanekom, W.A., Scriba, T.J., Kaufmann, S.H.E., Zak, D.E., Walzl, G., Walzl, G., Black, G.F., van der Spuy, G., Stanley, K., Kriel, M., Du Plessis, N., Nene, N., Roberts, T., Kleyhans, L., Gutschmidt, A., Smith, B., Nene, N., Loxton, A.G., Chegou, N.N., Tromp, Gerhardus, Tabb, D., Ottenhoff, T.H.M., Klein, M.R., Haks, M.C., Franken, K.L.M.C., Geluk, A., van Meijgaarden, K.E., Joosten, S.A., Boom, W.H., Thiel, B., Mayanja-Kizza, H., Joloba, M., Zalwango, S., Nsereko, M., Okwera, B., Kisingo, H., Kaufmann, S.H.E., Parida, S.K., Golinski, R., Maertzdorf, J., Weiner, J., Jacobson, M., Dockrell, H., Smith, S., Gorak-Stolinska, P., Hur, Y.-G., Lalor, M., Lee, J.-S., Crampin, A.C., French, N., Ngwira, B., Ben-Smith, A., Watkins, K., Ambrose, L., Simukonda, F., Mvula, H., Chilongo, F., Saul, J., Branson, K., Suliman, S., Scriba, T.J., Mahomed, H., Hughes, E.J., Bilek, N., Erasmus, M., Xasa, O., Veldsman, A., Downing, K., Fisher, M., Penn-Nicholson, A., Mulenga, H., Abel, B., Bowmaker, M., Kagina, B., Chung, W.K., Hanekom, W.A., Sadoff, J., Sizemore, D., Ramachandran, S., Barker, L., Brennan, M., Weichold, F., Muller, S., Geiter, L., Kassa, D., Abebe, A., Mesele, T., Tegbaru, B., Baarle, D. van, Miedema, F., Howe, R., Mihret, A., Aseffa, A., Bekele, Y., Iwnetu, R., Tafesse, M., Yamuah, L., Ota, M., Sutherland, J., Hill, P., Adegbola, R., Corrah, T., Antonio, M., Togun, T., Adetifa, I., Donkor, S., Andersen, P., Rosenkrands, I., Doherty, M., Weldingh, K., Schoolnik, G., Dolganov, G., Van, T., Kafaar, F., Workman, L., Mulenga, H., Scriba, T.J., Hughes, E.J., Bilek, N., Erasmus, M., Xasa, O., Veldsman, A., Cloete, Y., Abrahams, D., Moyo, S., Gelderbloem, S., Tameris, M., Geldenhuys, H., Hanekom, W., Hussey, G., Ehrlich, R., Verver, S., Geiter, L., 2018. Four-Gene Pan-African Blood Signature Predicts Progression to Tuberculosis. *Am. J. Respir. Crit. Care Med.* 197, 1198–1208.
<https://doi.org/10.1164/rccm.201711-2340OC>
- Sweeney, T.E., Braviak, L., Tato, C.M., Khatri, P., 2016. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. *Lancet Respir. Med.* 4, 213–224.
[https://doi.org/10.1016/S2213-2600\(16\)00048-5](https://doi.org/10.1016/S2213-2600(16)00048-5)
- Thompson, E.G., Du, Y., Malherbe, S.T., Shankar, S., Braun, J., Valvo, J., Ronacher, K., Tromp, G., Tabb, D.L., Alland, D., Shenai, S., Via, L.E., Warwick, J., Aderem, A., Scriba, T.J., Winter, J., Walzl, G., Zak, D.E., Du Plessis, N., Loxton, A.G., Chegou, N.N., Lee, M., 2017. Host blood RNA signatures predict the outcome of tuberculosis treatment. *Tuberculosis* 107, 48–58.
<https://doi.org/10.1016/j.tube.2017.08.004>
- Turner, C.T., Gupta, R.K., Tsaliki, E., Roe, J.K., Mondal, P., Nyawo, G.R., Palmer, Z., Miller, R.F., Reeve, B.W., Theron, G., Noursadeghi, M., 2020. Blood transcriptional biomarkers for active pulmonary tuberculosis in a high-burden setting: a prospective, observational, diagnostic accuracy study. *Lancet Respir. Med.* 8, 407–419. [https://doi.org/10.1016/S2213-2600\(19\)30469-2](https://doi.org/10.1016/S2213-2600(19)30469-2)
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., Speleman, F., 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3, research0034.1.
<https://doi.org/10.1186/gb-2002-3-7-research0034>
- Warsinske, H., Vashisht, R., Khatri, P., 2019. Host-response-based gene signatures for tuberculosis diagnosis: A systematic comparison of 16 signatures. *PLOS Med.* 16, e1002786.
<https://doi.org/10.1371/journal.pmed.1002786>

Chapter 4

The work in this chapter has been accepted for publication in the European Respiratory Journal:

Blood transcriptomic signatures for symptomatic TB in an African multi-cohort study

Vanessa Mwebaza Muwanga¹, Simon C Mendelsohn¹, Vinzeigh Leukes², Kim Stanley², Stanley Kimbung Mbandi¹, Mzwandile Erasmus¹, Marika Flinn², Tarryn-Lee Fisher², Rodney Raphela¹, Nicole Bilek¹, Stephanus T Malherbe², Gerard Tromp², Gian Van Der Spuy³, Gerhard Walzl², Novel N Chegou², Thomas J Scriba¹, for the AE-TBC and ScreenTB Consortium.

Author affiliations:

¹South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular Medicine, Division of Immunology, Department of Pathology, University of Cape Town, Cape Town, South Africa.

²DSI-NRF Centre of Excellence for Biomedical Tuberculosis Research; South African Medical Research Council Centre for Tuberculosis Research; Division of Immunology, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa.

³DSI-NRF Centre of Excellence for Biomedical Tuberculosis Research; South African Medical Research Council Centre for Tuberculosis Research; Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa.

Corresponding author: Thomas J Scriba

Email: Thomas.Scriba@uct.ac.za

Telephone: +27 (0)21 406 6791

Postal address: South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular Medicine and Division of Immunology, Department of Pathology, University of Cape Town, Anzio Road, Observatory, 7935, Cape Town, South Africa.

Summary

In a nested case-control study of blood transcriptomic TB signatures in symptomatic African patients, no signatures met benchmark criteria for a non-sputum TB triage test when all countries were pooled, but several met the minimum criteria in some countries.

4.1: Abstract

Background: Multiple host blood transcriptional signatures have been developed as non-sputum triage tests for tuberculosis (TB). We aimed to compare the diagnostic performance of 20 blood transcriptomic TB signatures for differentiating between symptomatic patients who have TB vs other respiratory diseases (ORD).

Methods: As part of a nested case-control study, individuals presenting with respiratory symptoms at primary health care clinics in Ethiopia, Malawi, Namibia, Uganda, South Africa, and The Gambia were enrolled. TB was diagnosed based on clinical, microbiological, and radiological findings. Transcriptomic signatures were measured in whole blood using microfluidic RT-qPCR. Diagnostic performance was benchmarked against the WHO Target Product Profile (TPP) for a non-sputum TB triage test.

Results: Among 541 participants, 158 had definite, microbiologically-confirmed TB, 32 had probable TB, while 389 participants had ORD. Nine signatures differentiated between ORD and TB with equivalent performance (Satproedprai7: area under the curve 0.83 [95% CI 0.79-0.87], Jacobsen3: 0.83 [0.79-0.86]; Suliman2: 0.82 [0.78-0.86]; Roe1: 0.82 [0.78-0.86]; Kaforou22: 0.82 [0.78-0.86]; Sambarey10: 0.81 [0.77-0.85]; Duffy9: 0.81 [0.76-0.86]; Gliddon3: 0.8 [0.75-0.85]; and Suliman4 0.79 [0.75-0.84]. Benchmarked against a 90% sensitivity, these signatures achieved specificities between 44% (95% CI 38-49) and 54% (49-59), not meeting the TPP criteria. Signature scores significantly varied by HIV status and country. In country-specific analyses several signatures, such as Satproedprai7 and Penn-Nicholson6, met the minimal TPP criteria for a triage test in Ethiopia, Malawi, and South Africa.

Conclusion: No signatures met the TPP criteria in a pooled analysis of all countries, but several signatures met the minimum criteria for a non-sputum TB triage test in some countries.

Keywords: *Mycobacterium tuberculosis*; blood; transcriptomic; mRNA, biomarker; performance; Africa; multi-cohort; Point-of-care; Triage; Non-sputum

4.2: Introduction

Current strategies to control the global TB epidemic largely rely on individuals with symptoms to self-present to clinics or other facilities for investigation and care. This approach, however, continues to contribute to missed TB cases, thereby calling for improvements in identification and diagnosis of these individuals.

Host biomarkers hold promise as non-sputum based diagnostics with fast turn-around times that could be deployed as POC triage tests to either rule in those with evidence of TB or rule out those who do not require further investigation (MacLean et al., 2019a). A frequent strategy to discover such host biomarkers is to profile blood transcripts in a case-control cohort of patients with microbiologically-confirmed TB and healthy control participants to identify transcripts that best discriminate the two groups. This strategy has led to the development of dozens of transcriptomic signatures that have good diagnostic performance in independent validation cohorts (Duffy et al., 2019; Maertzdorf et al., 2016; Penn-Nicholson et al., 2020b; Satproedprai et al., 2015). Most signatures include common transcripts, which largely map to ISG pathways, illustrating that these signatures capture common inflammatory processes characteristic of TB. This underpins confidence that transcriptomic TB signatures developed by different investigators in diverse settings reproducibly detect evidence of TB disease in peripheral blood.

A key limitation in many studies that have developed or validated diagnostic or triage test performance of transcriptomic signatures, and can influence diagnostic accuracy, is that case-control studies have been small, based on a single geographic setting and/or not represented a clinically relevant population (MacLean et al., 2019a). In most routine health-care settings, TB management requires a triage test that can be used to determine who, among those presenting for healthcare with symptoms suggestive of TB, should be referred for intensive investigation. Such individuals may have TB or other infectious or non-infectious respiratory diseases, the appropriate management and treatment of which is dependent on a definitive diagnosis. Respiratory diseases other than TB, such as viral infections, may also induce inflammatory responses that modulate blood transcriptomic signatures. It thus stands to reason that analysis of the diagnostic performance of transcriptomic signatures should be assessed in clinically relevant cohorts of symptomatic individuals seeking care, ideally from diverse geographic and epidemiological settings (MacLean et al., 2019a; Nogueira et al., 2022).

Here, we sought to determine the performance of 20 previously published, parsimonious blood transcriptomic signatures for diagnosing TB in patients who presented for care at clinics in six African countries with presumptive TB. Gene expression was measured by optimised, microfluidic RT-qPCR. For each signature the performance for differentiating between TB and ORD was benchmarked against the TPP criteria for a non-sputum TB triage test, developed by the WHO (World Health Organization, 2014). The use-case for symptomatic patients who seek care would be a test that allows triage of test-positive individuals for TB confirmation and, if confirmed, chemotherapy. Such a triage test must therefore have excellent sensitivity to ensure that those with TB have the best opportunity to receive the appropriate diagnosis and treatment. Under this scenario, a triage test with a sensitivity of at least 90% should achieve a specificity of at least 70% to meet the TPP criteria.

4.3: Methods

Study design

We conducted a nested case-control study that included participants from two previously completed, multi-site phase III diagnostic cohorts, The Evaluation of Host Biomarker-based POC Tests for Targeted Screening for Active TB (ScreenTB) (Manngo et al., 2019) and African European Tuberculosis Consortium (AE-TBC) (Chegou et al., 2018a, 2016) studies. AE-TBC study participants, at all sites, were recruited between November 2010 and November 2012. All 6-month follow-ups were completed in 2013. Study participants for ScreenTB from the South African site at Stellenbosch University were recruited between November 2016 and July 2019. Common eligibility criteria were used to enrol adults (18-70 years) with or without HIV, seeking care at primary health care centres in Ethiopia, Malawi (AE-TBC only), Namibia, South Africa, Uganda, and The Gambia. Individuals presenting with cough for more than two weeks, and at least one other feature suggestive of TB disease (close contact with a TB patient, fever, malaise, weight loss, night sweats, haemoptysis, chest pain, or loss of appetite), were prospectively recruited. Patients were excluded if they were on TB treatment or received isoniazid prophylaxis in the last 90 days, reported quinolone or aminoglycoside antibiotic use in the last 60 days, were pregnant or breastfeeding, had a haemoglobin less than 9g/dL, were without permanent residence, or had not been residing in the study area for more than 3 months.

At enrolment, participants provided sputum samples for *Mycobacterium tuberculosis* (*Mtb*) culture, smear microscopy, and Xpert MTB/RIF (ScreenTB only), underwent chest radiography, and provided venous blood in PAXgene blood RNA tubes (Qiagen, Hilden, Germany). Harmonized protocols were implemented across study sites for collection and cryopreservation of whole blood samples in the PAXgene tubes. Participants were required to attend follow-up visits from week 2 up to week 24 for further sample collection and to assess their response to treatment.

An outcome classification committee, comprised of clinician investigators from the participating sites, classified each participant as having either definite TB (microbiologically confirmed), probable TB, possible TB, or no TB (ORD controls). Case definitions were based on a composite reference standard of microbiological, clinical, and radiological findings (**Table 4.1**). Microbiological confirmation was by TB sputum culture, Xpert MTB/RIF (ScreenTB only), or smear microscopy, while clinical diagnosis was based on chest radiography results, and clinical response to antibiotics or TB treatment, if prescribed. Participants with possible TB were not analysed due to insufficient evidence of TB disease and uncertainty about which group to allocate them to.

Table 4.1: Diagnostic classification of participants in the ScreenTB and AE-TBC cohorts

CLASSIFICATION	DEFINITION
Definite TB	1 Sputum culture positive for MTB OR Gene Xpert MTB/Rif positive for MTB and no previous TB treatment in the past year OR 1 Positive smear plus CXR suggestive of PTB
Probable TB	1 positive smear and symptoms responding to TB treatment OR CXR compatible with PTB and symptoms responding to TB treatment OR Treatment started on clinical suspicion after failed antibiotic trial and shows good response to TB treatment
Questionable/Possible TB*	Positive smear(s), but no other supporting evidence OR CXR compatible with PTB, but no other supporting evidence OR Treatment initiated by healthcare providers on clinical suspicion only. No other supporting evidence OR No sputum Xpert or MTB culture successfully performed
Other Respiratory Diseases (ORDs)	Negative cultures/Xpert, negative smears, negative CXR not compatible with PTB, and TB treatment never initiated by healthcare providers OR CXR compatible with PTB, but negative sputum culture/Xpert, negative smears, and resolution of symptoms without TB treatment or alternative diagnosis confirmed

TB = tuberculosis. PTB = pulmonary tuberculosis. ORD = other respiratory diseases. MTB = *Mycobacterium tuberculosis*. CXR = chest x-ray.

* Excluded from this study

We evaluated factors that influenced outcome membership for each participant into “definite”, “probable”, and “negative (ORD controls)”. Site, age and sex were important drivers (significant explanatory variables), with HIV status relevant to probable case status using multinomial regression. Redefining cases as either definite or probable yielded site, sex and age as drivers of outcome. We therefore decided to match on site, sex, and age in the propensity score model, but retaining definite, probable, and negative as outcomes.

There were 447 cases (355 definite and 92 probable) and 978 controls in the ScreenTB and AE-TBC cohorts with PAXgene samples available. We randomly selected 300 cases from the 447 and then used the R package MatchIt to select 600 controls by propensity scoring using the *nearest neighbor without replacement* algorithm. To reduce the risk of overmatching, the set of 300 cases and 600 controls (matched on site, sex and age) were then divided, using a stratified random sampling on site, sex, and classification (to ensure proportional representation of classes with low counts), but without regard to keeping matched controls with the case to which they were matched, into equal sets of 300 subjects for the train, test and validate sets (note that the validate set was not used in this study – it was retained for future validation studies of a newly trained signature).

Thus, the propensity score was used to ensure matching in the aggregate on factors that were significant for outcome membership, but the strict matching was decoupled in the sets used for training, testing and validation.

The analysts who generated the model were blinded to the process of propensity scoring and randomization.

Ethics statement

Study protocols for the ScreenTB and AE-TBC studies were approved by the Health Research Ethics Committee of Stellenbosch University, and at all participating sites. All participants provided written, informed consent in accordance with the Declaration of Helsinki. The transcriptomic signatures sub-study protocol was reviewed and approved by the University of Cape Town HREC (589/2019).

Measurement of signatures

Blood signatures were measured as previously described (Mendelsohn et al., 2022b; Scriba et al., 2021a). Whole blood samples from all participants were thawed once to perform RNA extraction and processed in the same manner using standardised protocols. RNA was extracted from cryopreserved PAXgene blood RNA tubes using the Maxwell HT simplyRNA Custom Kit (Promega, Madison, Wisconsin, United States), and cDNA synthesis performed using EpiScript reverse transcriptase (Lucigen, Hoddesdon, United Kingdom), in a semi-automated manner in 96-well plate format on the Freedom EVO 150 robotics platform (Tecan, Männedorf, Switzerland). Non-template (water) and no reverse transcriptase controls were included in all runs. Following cDNA synthesis, 93 transcripts of interest and 3 reference (“housekeeping”) genes were pre-amplified using a pool of Taqman primer-probe assays (**Chapter 3 Table 3.9**). Measurement of transcript expression was performed by RT-qPCR using 96.96 microfluidic gene expression chips (Standard BioTools, California, United States) on the Biomark HD system (Standard BioTools, California, United States), which multiplexes 96 samples with the 96 TaqMan primer-probe assays. An internal positive control (IPC) sample was included on all chips to monitor batch variability. Gene expression data quality control, including evaluation of control samples, and calculation of signature scores, was performed using a locked-down R script (**Chapter 2 Section 2.4 and Figure 3.3**). Signature score calculations were based on the machine learning models and statistical methods described in the original publications, unless otherwise stated (**Chapter 3**). Batch correction (**Figure 4.1**) was performed to correct for differences in signature score distribution between Fluidigm chip runs using a quantile regression method with the `adjust_batch` function in the `batchtma` R package (Stopsack et al., 2021). This method unifies lower quantiles (25th percentile) and ranges between the lower and upper quantile (75th percentile) between batches, allowing for differences in both parameters due to confounders. Thereafter, signature scores were winsorized (clipped) to account for outliers (Winsorize function, DescTools R package (Signorell et al., 2023)) log transformed to normalise score distribution (log function, base R package), converted to z-scores (scale function, base R package), and compared using the cumulative distribution function (pnorm function, stats R package).

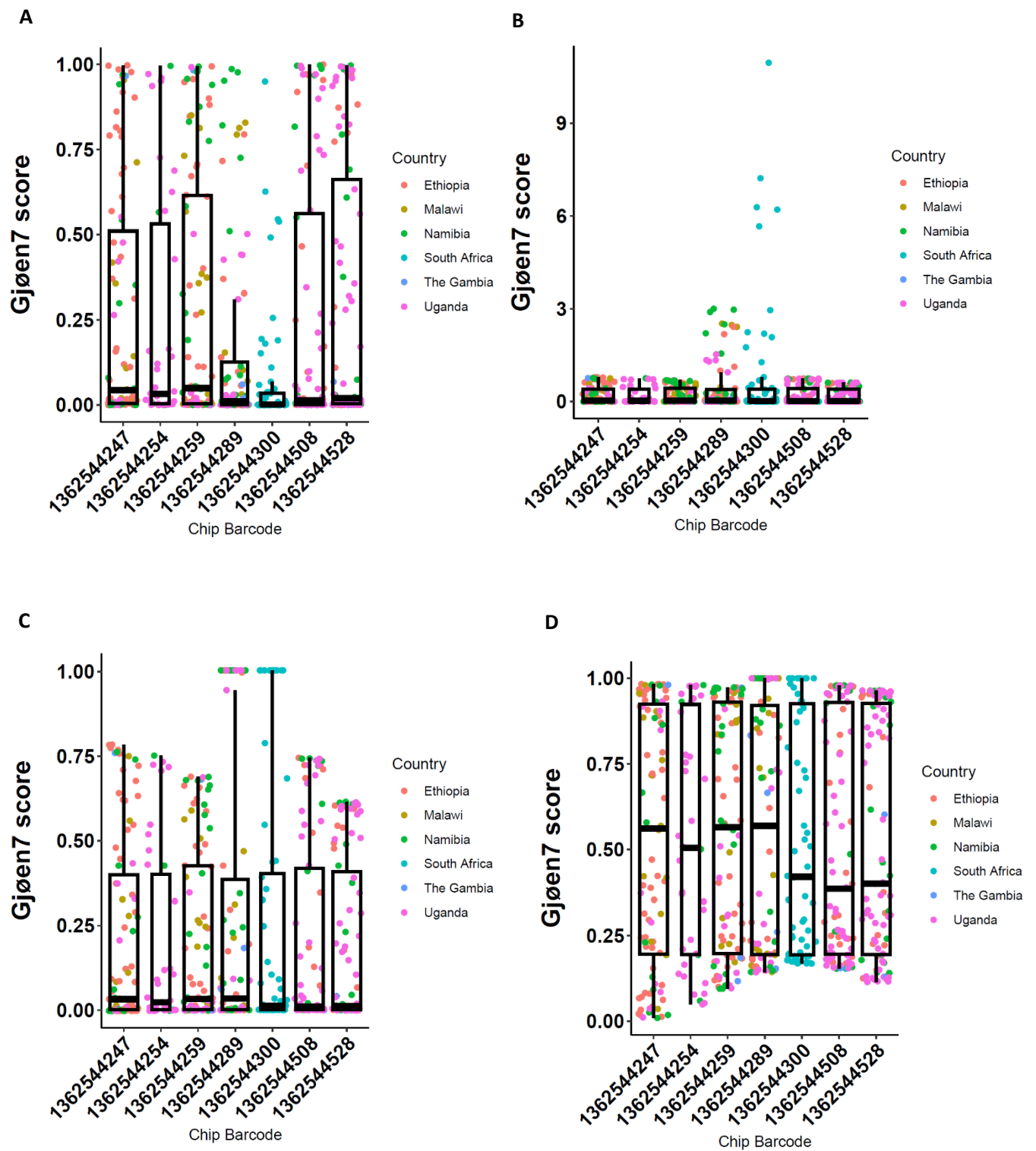


Figure 4.1: Batch correction and normalization of signature scores.

(A) Signature scores for Gjøn7 across chips before batch correction, (B) after batch correction, (C) after winsorization, and (D) following normalization.

Blinding

The ScreenTB and AE-TBC clinical teams determined TB endpoint status prior to measurement of signature scores. The laboratory team were blinded to the TB status of participants during sample processing and signature score generation. Clinical and demographic data (including TB disease status) were shared by collaborators only after all laboratory work was completed and signature scores had been calculated. Statistical analyses were agreed upon and signed off before any data analysis officially began.

Statistical analysis

Sample size was determined based on the availability of PAXgene RNA samples from TB cases and controls enrolled in the ScreenTB and AE-TBC cohorts. All statistical analyses were carried out in R version 4.1.3 (R Core Team 2022). A p-value or, to adjust for multiple testing, adjusted q-value (Benjamini-Hochberg q-value) of <0.05 was considered significant.

The primary diagnostic analyses were performed on individuals classified as having either definite or probable TB, compared to ORD controls. The pROC package (Robin et al., 2011) was used to calculate area under the receiver operating characteristic (ROC) curve (AUC). The DeLong method (Elizabeth R. DeLong et al., 1988) was used to calculate 95% confidence intervals (CI) and compare AUC of each candidate signature in a pairwise approach. The false discovery rate was controlled using the Benjamini-Hochberg (Benjamini and Hochberg, 1995) method to calculate q-values.

Positive and negative predictive values, sensitivity, and specificity for each candidate signature were calculated using routine formulae and benchmarked against the minimal (sensitivity 90% and specificity 70%) and optimal (sensitivity 80% and specificity 95%) WHO TPP criteria for a non-sputum TB triage test (World Health Organization, 2014). The binomial Wilson method was used to calculate 95% CIs for all performance metrics.

The Wilcoxon rank-sum test was used to test for differences in signature scores between groups. Multivariable linear regression was performed to determine effect of participant clinical and demographic characteristics at enrolment on signature scores. Missing data points, including signature scores due to failed primer-probe reactions, were assumed missing at random and excluded from the multivariable regression (i.e. complete case analysis only) and pairwise signature comparisons. Spearman's rank correlation coefficient was used to test association between signatures.

Missing data

Signature pass and fail rate were looked at first to determine an appropriate threshold at which to exclude certain signatures from analysis completely if a signature score could not be computed for too many samples. For exploratory analysis, participant characteristic variables (e.g. BMI, CD4 count, smoking status, haemoglobin) were assumed to be missing at random. Participants with missing data points were excluded from exploratory analyses (i.e. complete case analysis only).

4.4: Results

Selection and matching of study participants

Among 2130 participants enrolled into the ScreenTB and AE-TBC studies, 425 participants with definite TB, 94 with probable TB, and 1128 with ORD were eligible for inclusion in this sub-study, of which 158 with definite TB, 32 with probable TB, and 389 ORD controls were randomised to inclusion (**Figure 4.2**). Gene expression results from 37 study participants were excluded from analysis because of failed housekeeping genes (n=4) or failure rate of primer-probe assays greater than 20% (n=33). Transcriptomic signature scores were obtained from 150 definite TB, 30 probable TB cases, and 361 ORD controls. We compared demographic and clinical variables among TB cases and ORD controls (**Table 4.2**). No differences in sex (p=0.17), age (p=0.28), CD4 count (p=0.28), proportion on ART (p=0.65), history of prior TB (p=0.37), or smoking (p=0.30) were observed. The distribution of cases and controls across countries was different (p=0.007). Definite TB cases had a lower BMI (median 18.5% vs 19.4% vs 20.8%, p<0.001), were more likely to be interferon- γ release assay (IGRA) positive (84.3% vs 45.5 vs 42.7%, p<0.001), present with haemoptysis (37.5% vs 20% vs 22.1%, p=0.042), and report subjective weight loss (76.8% vs 70.8% vs 50%, p<0.001) than probable TB cases or ORD controls, respectively.

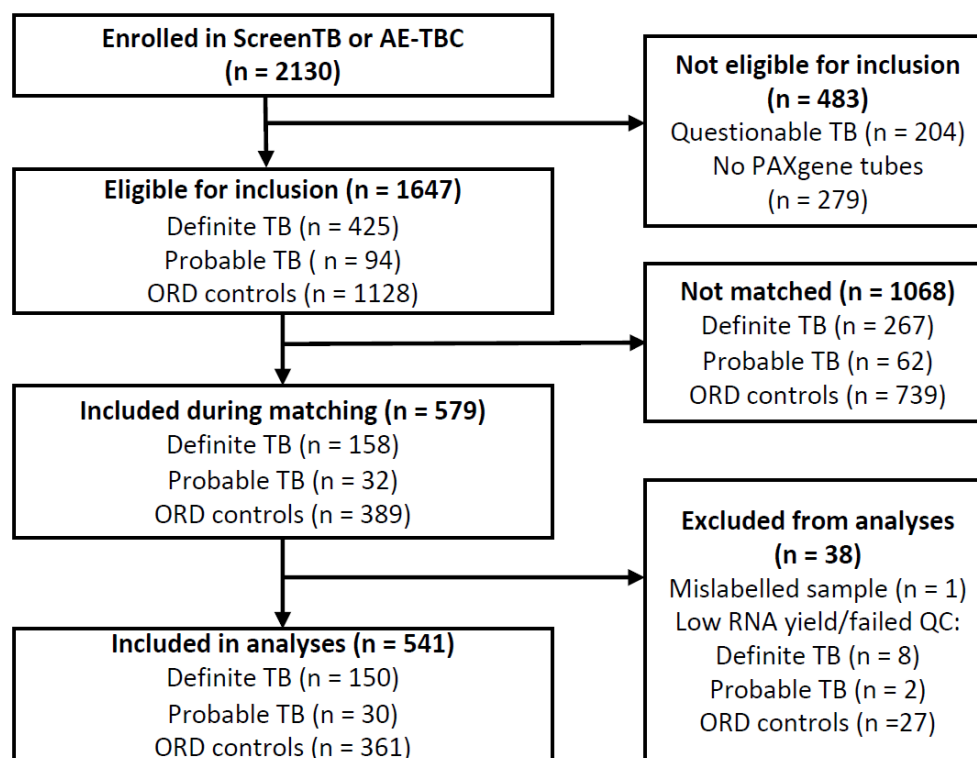


Figure 4.2: Flow diagram of participant enrolment, allocation, and analyses.

Participants in the research study that were previously enrolled into either the ScreenTB (Manngo et al., 2019) or AE-TBC (Chegou et al., 2018a, 2016) studies.

Table 4.2: Participant characteristics

Characteristic	Definite TB n = 150	Probable TB n = 30	ORD n = 361	p-value [†]
Male, n (%)	99 (66.0)	15 (50.0)	214 (59.3)	0.17
Median age, years (IQR)	32.0 (25.0-41.8)	34.5 (29.0-44.0)	32.0 (26.0-41.0)	0.28
Previous TB, n (%)	32 (21.3)	5 (16.7)	58 (16.2)	0.37
Missing, n	0	0	2	
Smoking history, n (%)				0.30
Current	38 (27.0)	5 (20.8)	63 (19.0)	
Past	14 (9.9)	1 (4.2)	28 (8.4)	
Never	89 (63.1)	18 (75.0)	241 (72.6)	
Missing	9	6	29	
Median BMI, kg/m ² (IQR)	18.5 (17.2-20.2)	19.4 (18.1-21.6)	20.8 (18.8-23.1)	<0.001
Missing	10	6	31	
IGRA result, n (%)				<0.001
Positive	86 (84.3)	5 (45.5)	94 (42.7)	
Negative	14 (13.7)	6 (54.5)	109 (49.5)	
Indeterminate	2 (2.0)	0 (0.0)	17 (7.7)	
Missing	48	19	141	
HIV positive, n (%)	33 (22.0)	14 (46.7)	96 (26.6)	0.020
Antiretroviral therapy, n (%)	16/33 (48.5)	5/14 (35.7)	29/96 (30.2)	0.654
Median CD4 count in HIV-infected, cells/mm ³ blood (IQR)	176.0 (95.0-304.0)	220.0 (219.5-233.5)	355.0 (127.0-634.0)	0.28
Missing/NA	135	27	320	
Country, n (%)				0.007
Ethiopia	41 (27.3)	11 (36.7)	105 (29.1)	
Malawi	8 (5.3)	6 (20.0)	29 (8.0)	
Namibia	26 (17.3)	5 (16.7)	55 (15.2)	
South Africa	30 (20.0)	0 (0.0)	35 (9.7)	
The Gambia	4 (2.7)	0 (0.0)	7 (1.9)	
Uganda	41 (27.3)	8 (26.7)	130 (36.0)	
Study, n (%)				<0.001
AE-TBC	57 (38.0)	24 (80.0)	195 (54.0)	
ScreenTB	93 (62.0)	6 (20.0)	166 (46.0)	
Haemoptysis, n (%)	24 (37.5)	4 (20.0)	45 (22.1)	0.042
Missing	86	10	157	
Shortness of breath, n (%)	118 (83.1)	15 (62.5)	233 (70.2)	0.006
Missing	8	6	29	
Weight loss, n (%)				<0.001
Yes	109 (76.8)	17 (70.8)	166 (50.0)	
No	29 (20.4)	7 (29.2)	155 (46.7)	
Unsure	4 (2.8)	0 (0.0)	11 (3.3)	
Missing	8	6	29	
Night sweats, n (%)				0.090
Yes	82 (57.7)	18 (75.0)	165 (49.7)	
No	56 (39.4)	6 (25.0)	160 (48.2)	
Unsure	4 (2.8)	0 (0.0)	7 (2.1)	
Missing	8	6	29	
Median cough duration, months (IQR)	0.7 (0.6-1.0)	1.0 (0.7-2.3)	0.7 (0.7-1.0)	0.091
Missing	16	6	44	

TB = tuberculosis. ORD = other respiratory diseases. BMI = body-mass index. IGRA = interferon- γ release assay. IQR = interquartile range. † Pearson's Chi-squared test; Kruskal-Wallis rank sum test; Fisher's Exact Test for Count Data with simulated p-value (based on 2000 replicates).

Pass rate of transcriptomic signatures

Of the 20 signatures on the panel, five (Francisco2, de Araujo1, Sweeney3, Gjøen7, and da Costa3) had a pass-rate of 100%, seven (Suliman4, Satproedprai7, Jacobsen3, Maertzdorf4, Roe1, Roe3, Thompson5, Penn-Nicholson6, and Sambarey7) had a sample pass-rate between 90% and 100%, while six (Duffy9, Gliddon 4, Rajan5, Gliddon3, Kaforou22, and Suliman2) had a sample pass-rate between 60% and 80% (**Figure 4.3**). These six signatures with a pass rate below 80% all contained the GBP6 transcript whose primer-probe assay had a pass rate of 70.98% in whole blood samples from this cohort (**Table 4.3**). The pass rate of all 96 individual transcripts was also assessed (**Table 4.3**).

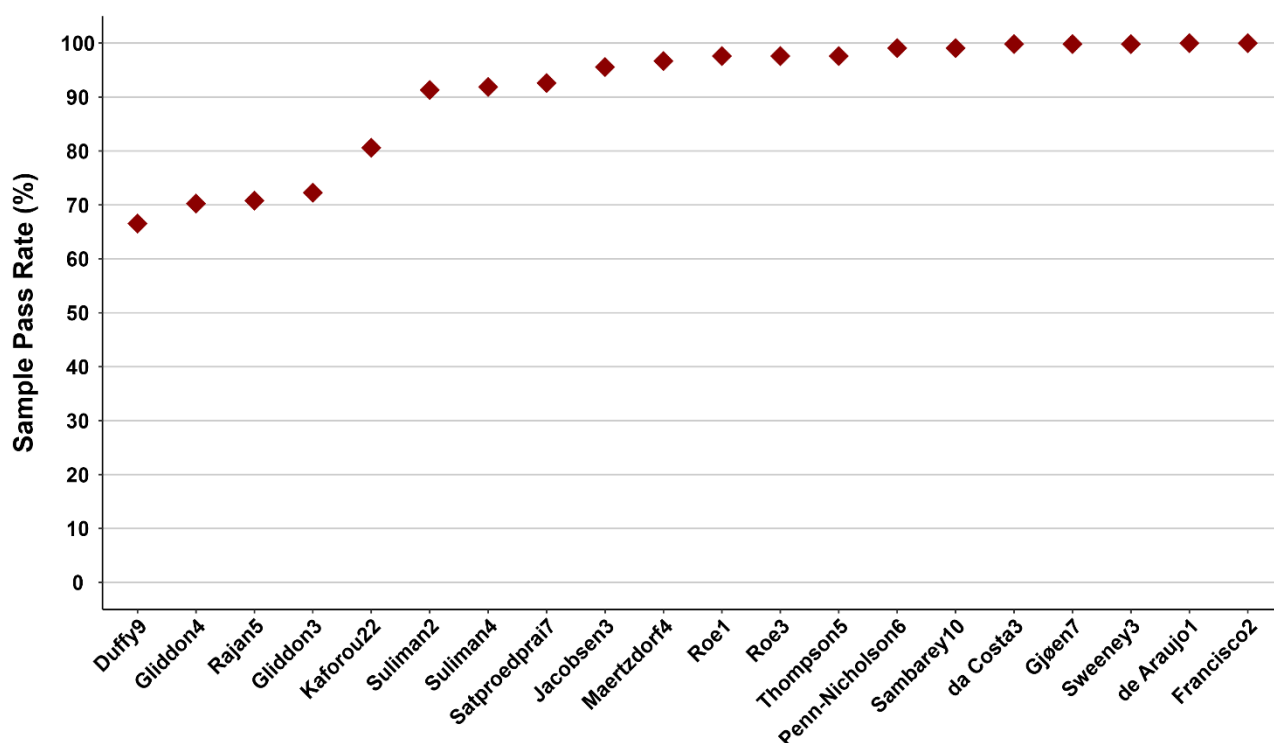


Figure 4.3: Sample pass rate in the Screen-TB and AE-TBC sub-study.

All 20 transcriptomic signatures evaluated are shown on the x-axis in order of sample pass rate from the highest to the lowest. On the y-axis, the pass-rate of each signature in samples from both the ScreenTB and AE-TBC cohorts was expressed as a percentage.

Table 4.3: Assay pass rate

Primer probe assay	Pass rate (%)
ACTA2.Hs00426835_g1	98.15
ACTR3.Hs01029159_g1	100
ANKRD22.Hs00944015_m1	97.78
APOL1.Hs01066280_m1	99.82
ARG1.Hs00968979_m1	98.52
BATF2.Hs00912737_m1	97.6
BCL6.BCL6_CUSTOM_AREPURD	100
BLK.Hs01017452_m1	98.71
C1QA.Hs00706358_s1	99.26
C1QB.Hs00608019_m1	78.37
C1QC.Hs00757779_m1	27.17
C4ORF18.FAM198B.Hs00259260_s1	99.82
C5.Hs00156197_m1	91.13
CCR6.Hs01890706_s1	98.34
CD160.Hs00199894_m1	98.52
CD1C.Hs00957534_g1	97.78
CD36.2.Hs01567186_m1	99.45
CD3E.CD3E_APDJ3NT	100
CD40L.Hs00163934_m1	99.82
CD74_j1.Hs04983808_s1	100
CD74_j2.CD74_J2_AP9HN47	85.4
CD79A.Hs00998120_g1	100
CD79B.Hs01058826_g1	100
CDCA7.Hs00230589_m1	0
CDKN1C.Hs00175938_m1	91.5
CXCR5.Hs00540548_s1	99.82
CYP4F3.Hs01587860_mH	100
DUSP3.Hs01115776_m1	99.82
F2RL1.Hs00608346_m1	75.42
FAM20A.Hs01034066_m1	94.82
FCGR1A.FCGR1A_SAT7_AP2XDMK	100
FCGR1A.Hs00174081_m1	99.82
FCGR1B.Hs02341825_m1	100
FCGR1B_VARIANT_1.FCGR1B_VARIANT1_AP3267H	100
FCGR1B_VARIANT_2.FCGR1B_VARIANT2_AP47ZTF	99.63
FCGR1C.Hs00417598_m1	100
FLVCR2.Hs00900390_m1	98.15
GAS6.GAS6_NM_000820_10_11_AR323W3	95.01
GAS6.GAS6_NM_000820_5_6_AR47XGZ	96.3
GBP1.Hs00977005_m1	100
GBP2.Hs00894846_g1	100
GBP5.Hs00369472_m1	100
GBP6.Hs01584201_m1	70.98
gjeong7IFITM3	99.82
GNG7.Hs00192999_m1	99.82
GYG1.Hs00907542_g1	99.82
GZMA.Hs00989184_m1	100
HK3.Hs01092850_m1	100
ID3.Hs00954037_g1	99.26
IFI44L.Hs00915292_m1	100
IFITM3.Hs03057129_s1	100

IFITM3_G.IFITM3_G_APMF24C	0
KAZN.KAZN_TOTAL_SAT7_AP7DVDD	96.49
KIF1B.Hs01114512_g1	100
KLF2.Hs00360439_g1	100
KLF2.KLF2_CUSTOM_ARH6CF4	85.21
KLHDC8B.Hs00293902_m1	99.26
LAG3.Hs00958444_g1	100
LHFPL2.Hs00299613_m1	97.6
LTF.Hs00158924_m1	96.3
MAFB.MAFB_SAT7_APZTH2N	95.19
MAP7D3.Hs00226257_m1	98.52
MMP9.MMP9_APEPW9P	100
MPO.Hs00165162_m1	95.38
MTRF1L.Hs01097882_g1	100
NOD2.Hs01550759_g1	99.82
NPC2.Hs01119244_m1	100
OSBPL10.Hs00215016_m1	93.35
P2RY14.Hs01848195_s1	97.41
PRDM1.Hs00153357_m1	100
RAB13.Hs04400188_g1	100
Rab33A.Hs00191243_m1	99.26
RABL2A.Hs00255244_m1	99.82
RBBP8.Hs01090329_m1	99.08
RP11.295G20.2.Hs01373568_m1	100
S100A8.Hs00374264_g1	100
SCARF1.Hs01092483_m1	100
SDR39U1.Hs01016970_g1	100
SEPT4.Hs00910208_g1	97.78
SERPING1.Hs00934329_m1	99.82
SLPI.Hs00268206_m1	100
SMARCD3.Hs01088251_g1	100
STAT1.Hs01013996_m1	100
STT3A.Hs00967491_m1	99.08
TIMM10.TIMM10_CUSTOM_APT2DEC	100
TMBIM6.Hs00162661_m1	100
TMCC1.Hs01037666_s1	99.45
TNIP1.TNIP1_GEO_AP47YZ	100
TRMT2A.Hs01000041_g1	99.26
TUBGCP6.Hs00363509_g1	100
UCP2.Hs01075224_g1	100
USF2.Hs01100994_g1	100
VAMP5.Hs01105383_g1	99.45
WARS_j1.Hs00998737_m1	100
WARS_j2.WARS.J2_APAAFHX	99.63
ZDHC19.Hs00376118_m1	93.35
ZNF296.Hs00377132_m1	89.83

Nine signatures had indistinguishable performance in discriminating TB cases and ORD controls

All 20 transcriptomic signatures significantly differentiated between all TB cases and ORD controls with areas under the ROC curve of 0.68 (95% CI 0.63-0.73) or above (**Figure 4.4** and **Table 4.4**). Notably, all signatures had higher diagnostic performance for definite TB than for probable TB (**Table 4.4**). We report signature performance for definite and probable TB cases combined, rather than separately, in the following analyses to reflect more realistic signature performance in real world settings, where people with unconfirmed TB are not excluded. Satpoedprai7 had the highest AUC (0.83, 95% CI 0.79-0.87) for distinguishing between all TB cases and controls (**Figure 4.4**). The diagnostic performance of eight other signatures, Jacobsen3, Kaforou22, Roe1, Suliman2, Sambarey10, Duffy9, Gliddon3, and Suliman4, was not statistically different to Satproedprai7 (**Figure 4.6** and **Table 4.5**). It should be noted that analysis of inter-signature performance required pair-wise comparisons between Satproedprai7 and each other signature and thus the sample size for each comparison was limited to the samples that had scores for both signatures. For example, the comparison of Satproedprai7 (0.83, 95% CI 0.79-0.87) with Duffy9 (0.81, 95% CI 0.76-0.86) was based on 127 paired TB cases and 230 paired controls, whereas the comparison of Satproedprai7 with Penn-Nicholson6 (0.81, 95% CI 0.77-0.84), with a near-identical AUC to Duffy9, was based on a much larger sample size of 167 paired TB cases and 333 paired controls. In this case the AUC for Duffy9 was found to be statistically not different to Satproedprai7, whereas the AUC for Penn-Nicholson6 was lower than that for Satproedprai7 (**Figure 4.6** and **Table 4.5**). Scores for most signatures were highly correlated, with spearman rank correlation coefficients ranging from 0.28 to 0.96 (**Figure 4.7**).

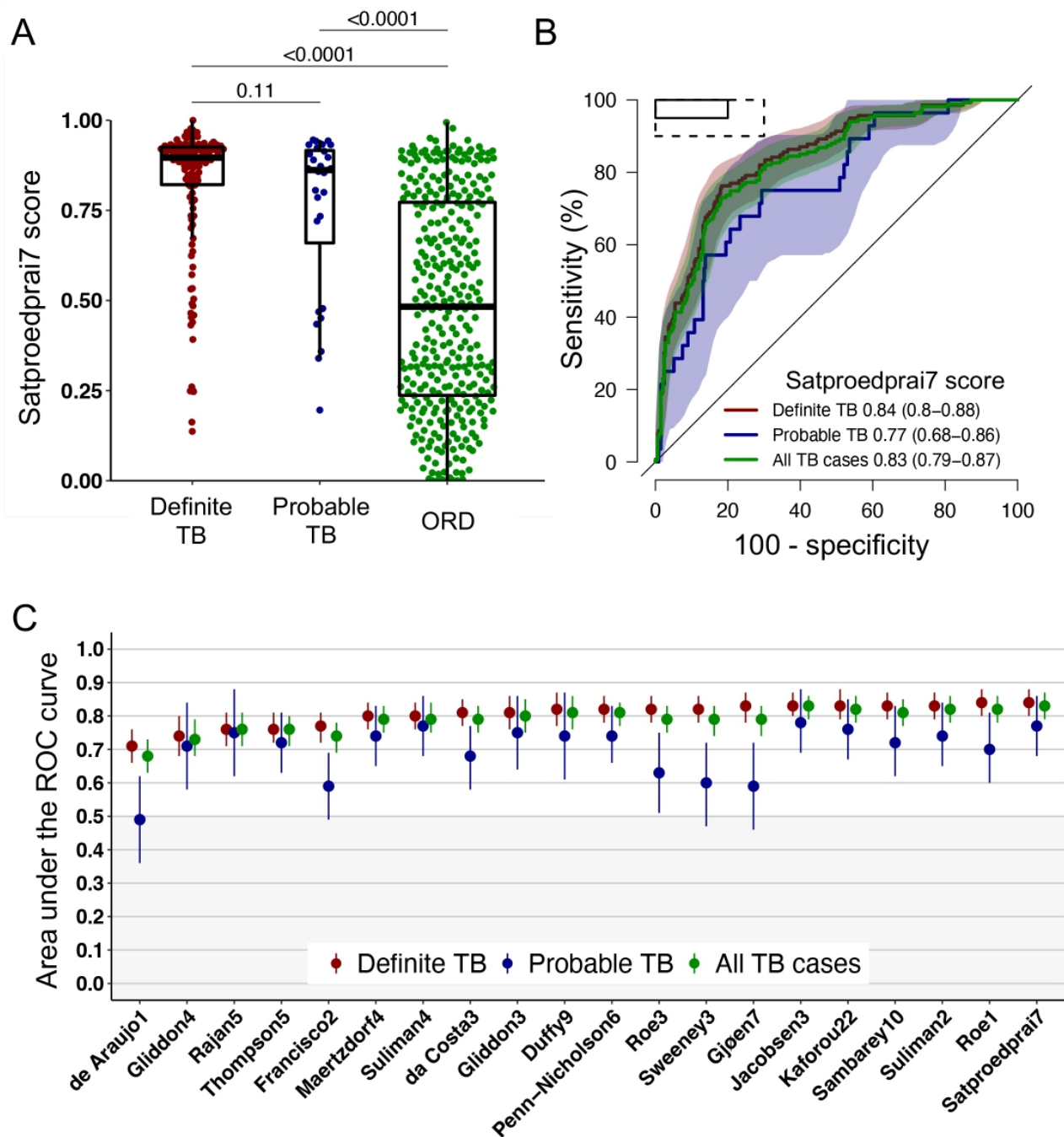


Figure 4.4: Signature score distributions and diagnostic performance of transcriptomic signatures in TB cases and controls with other respiratory diseases (ORD).

(A) Box-and-whisker plot of the signature with the best diagnostic performance, based on area under the receiver operating characteristic (ROC) curve (Satproedprai7). The plot shows signature score distribution (each dot represents a participant) in definite TB cases ($n = 139$), probable TB cases ($n = 28$) and controls with ORD ($n = 334$). P-values for comparison of median signature scores between groups were calculated using the Mann–Whitney U test. Boxes depict the IQR, the midline represent the median, and whiskers represent $1.5 \times \text{IQR}$.

(B) ROC curve depicting diagnostic performance (area under the curve, AUC, with 95% CI) of the Satproedprai7 signature in definite and probable TB cases separately, and in all TB cases combined, vs ORD controls. The shaded areas represent 95% CIs. The solid box represents the optimal criteria (95% sensitivity and 80% specificity) and the dashed box represents the minimal criteria (90% sensitivity and 70% specificity) set out in the WHO Target Product Profile for a TB triage test.

(C) Summary of signature diagnostic performance, expressed as AUC for all TB cases vs ORD (green). Error bars depict the 95% CIs. The diagnostic AUCs for definite TB cases vs ORD (red) and probable TB vs ORD (blue) are also shown.

Signatures were ranked according to AUC for definite TB cases vs ORD. A table of AUCs for all signatures are shown in **Table 4.4** below.

Table 4.4: Diagnostic performance of candidate signatures for definite TB, probable TB, and all TB cases vs ORD controls.

Biomarker	Definite TB Cases	Probable TB Cases	ORD Controls	Definite TB vs ORDs AUC (95% CI)	Probable TB vs ORDs AUC (95% CI)	All TB cases* vs ORDs AUC (95% CI)
da Costa3	150	30	360	0.81 (0.77-0.85)	0.68 (0.58-0.77)	0.79 (0.75-0.83)
de Araujo1	150	30	361	0.71 (0.66-0.76)	0.49 (0.36-0.62)	0.68 (0.63-0.73)
Duffy9	110	19	231	0.82 (0.77-0.87)	0.74 (0.61-0.87)	0.81 (0.76-0.86)
Francisco2	150	30	361	0.77 (0.72-0.81)	0.59 (0.49-0.69)	0.74 (0.69-0.78)
Gj��en7	150	30	360	0.83 (0.78-0.87)	0.59 (0.46-0.72)	0.79 (0.74-0.83)
Gliddon3	108	22	261	0.81 (0.76-0.86)	0.75 (0.64-0.86)	0.80 (0.75-0.85)
Gliddon4	113	20	247	0.74 (0.68-0.80)	0.71 (0.58-0.84)	0.73 (0.68-0.79)
Jacobsen3	142	28	347	0.83 (0.80-0.87)	0.78 (0.69-0.88)	0.83 (0.79-0.86)
Kaforou22	122	25	289	0.83 (0.79-0.88)	0.76 (0.67-0.85)	0.82 (0.78-0.86)
Maertzdorf4	142	29	352	0.80 (0.76-0.84)	0.74 (0.65-0.83)	0.79 (0.75-0.83)
Penn-Nicholson6	148	30	358	0.82 (0.78-0.86)	0.74 (0.66-0.83)	0.81 (0.77-0.84)
Rajan5	113	20	250	0.76 (0.71-0.81)	0.75 (0.62-0.88)	0.76 (0.71-0.81)
Roe1	150	29	349	0.84 (0.80-0.88)	0.70 (0.60-0.81)	0.82 (0.78-0.86)
Roe3	150	29	349	0.82 (0.78-0.86)	0.63 (0.51-0.75)	0.79 (0.75-0.83)
Sambarey10	148	30	358	0.83 (0.79-0.87)	0.72 (0.62-0.83)	0.81 (0.77-0.85)
Satproedprai7	139	28	334	0.84 (0.80-0.88)	0.77 (0.68-0.86)	0.83 (0.79-0.87)
Suliman2	137	28	329	0.83 (0.79-0.87)	0.74 (0.65-0.84)	0.82 (0.78-0.86)
Suliman4	137	28	332	0.80 (0.76-0.84)	0.77 (0.68-0.86)	0.79 (0.75-0.84)
Sweeney3	150	30	360	0.82 (0.78-0.86)	0.60 (0.47-0.72)	0.79 (0.74-0.83)
Thompson5	144	28	356	0.76 (0.72-0.81)	0.72 (0.63-0.81)	0.76 (0.71-0.80)

TB = tuberculosis. ORD = other respiratory diseases. AUC = area under the curve. CI = confidence interval.

* Definite and probable TB combined.

Data batch correction and normalisation processing steps had a negligible effect on the diagnostic performance of transcriptomic signatures (**Figure 4.5**).

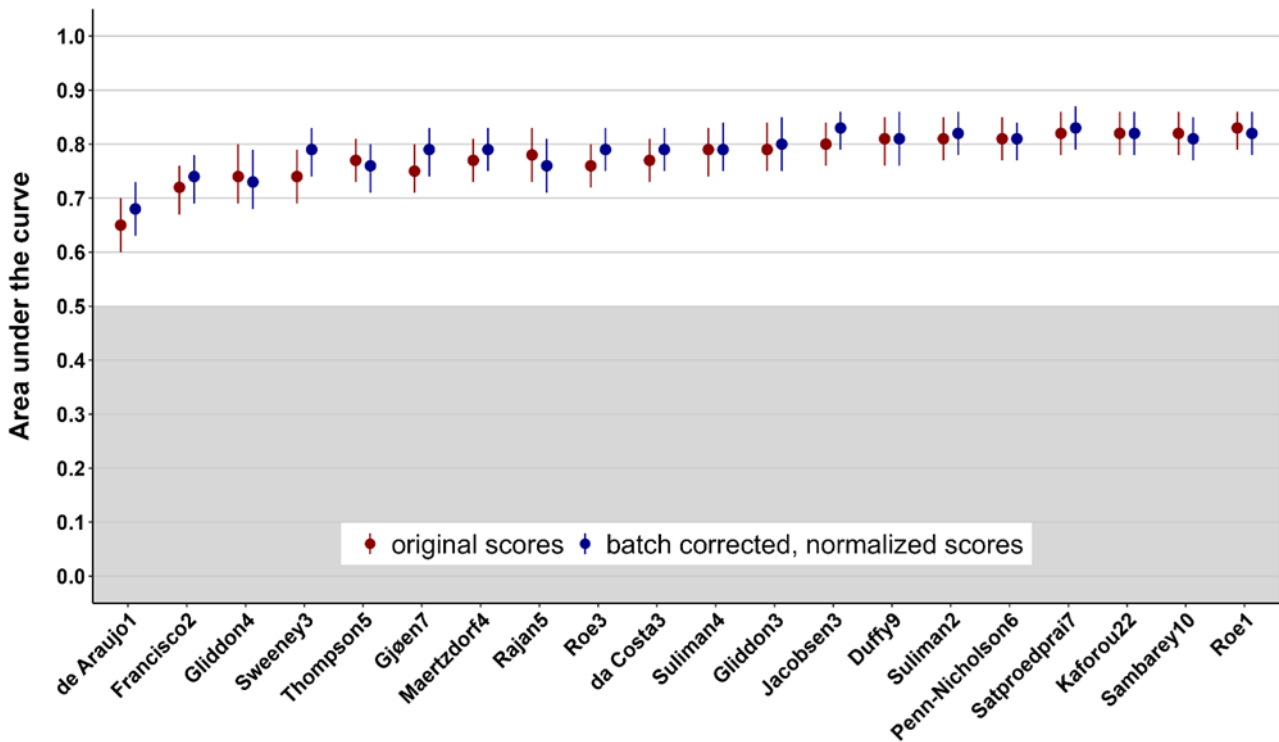


Figure 4.5: Summary of signature diagnostic performance, before and after batch correction, for all TB cases versus controls with other respiratory diseases.

The diagnostic area under curve (AUC) for the original signature scores (red) and batch corrected, normalised signature scores (blue) for all TB cases versus ORD controls, are shown. Error bars depict the 95% CIs.

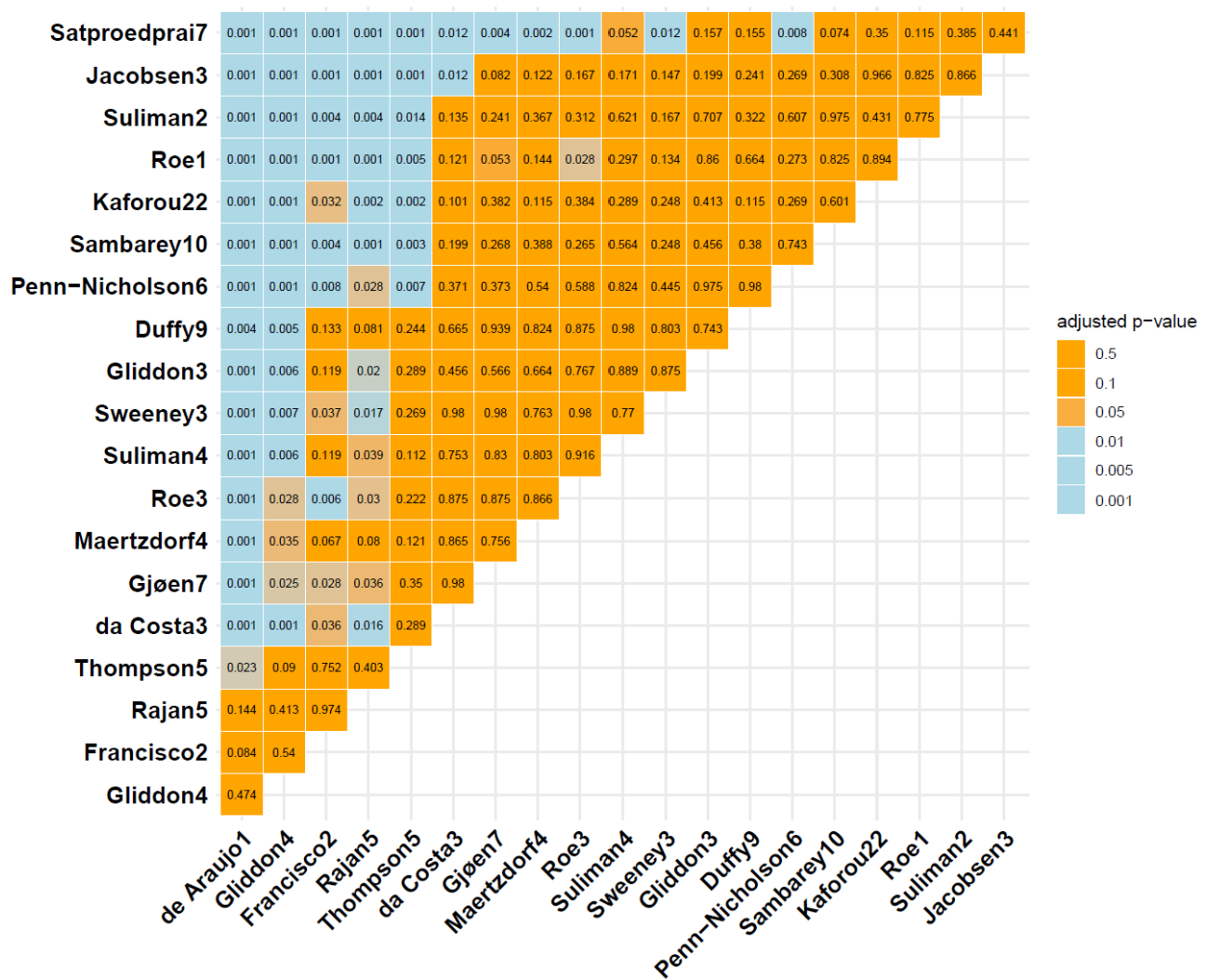


Figure 4.6: Comparison of diagnostic performance between candidate transcriptomic signatures.

Numbers in each square depict adjusted p-values (q-values), generated with the DeLong test for comparing receiver operating characteristic curves computed for each pair of transcriptomic signatures, corrected using the Benjamini–Hochberg method. Comparisons with adjusted p-values below 0.05, i.e. with different diagnostic performance, are shown in blue shading. Signatures on the y-axis and x-axis are listed in order of diagnostic performance, with Satproedprai7 having the highest AUC.

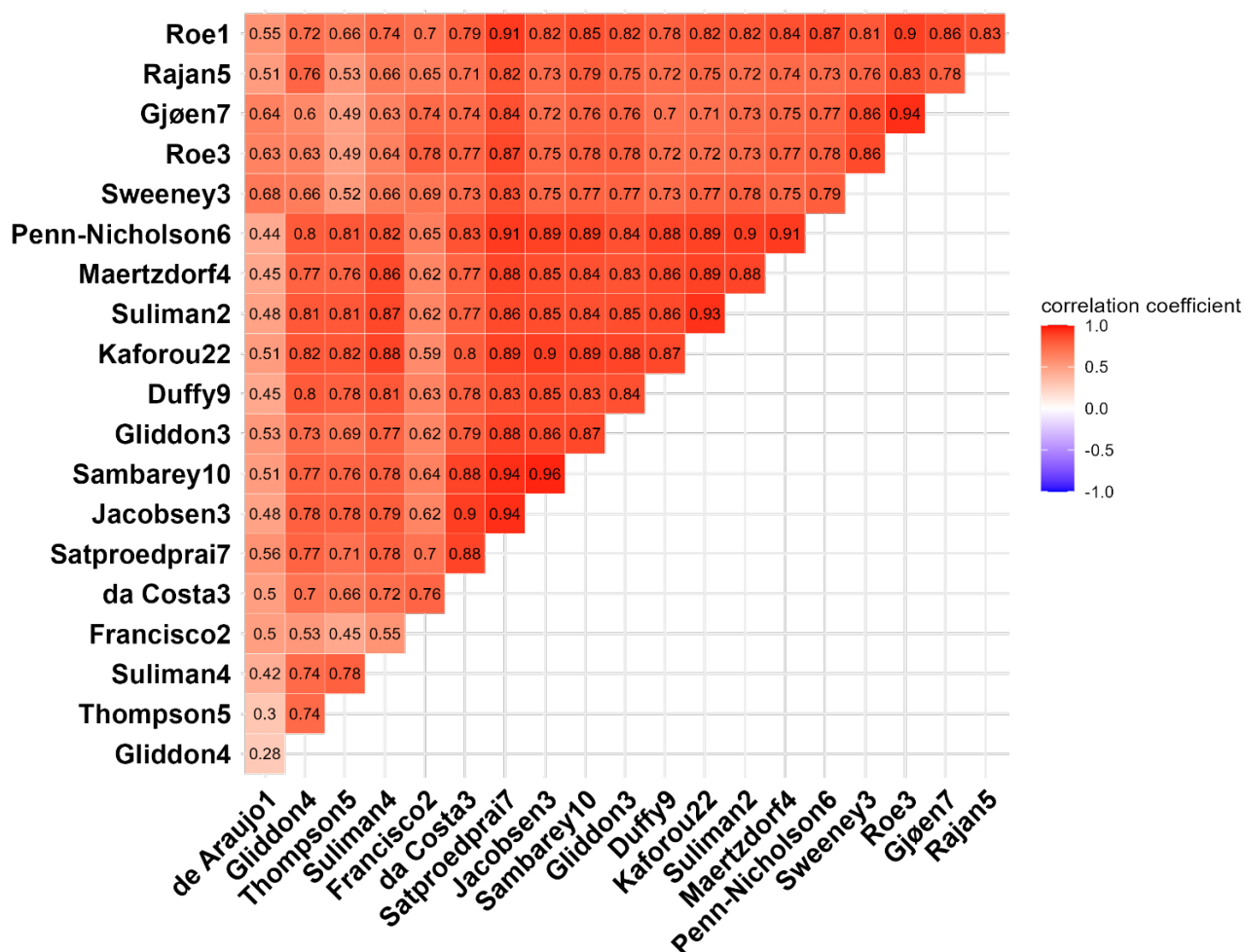


Figure 4.7: Signature scores for most signatures are highly correlated.

Numbers in each square depict Spearman rank correlation coefficients computed for each pair of transcriptomic signatures on samples from TB cases and ORD controls combined. Signature scores for de Araujo1, Rajan5, and Roe1 were negatively correlated with the other signatures, but absolute correlation coefficients were computed to make interpretation easier.

No signatures met the minimal WHO TPP benchmarks when cohorts are pooled

We benchmarked signature performance against the minimal WHO TPP criteria specified for a non-sputum-based triage test. At a fixed specificity of 70%, sensitivities ranged from 0.61 (95% CI 0.54-0.68) for the de Araujo1 signature to 0.81 (95% CI 0.74-0.86) for Satproedprai7; none met the minimum criterion threshold of 90% sensitivity (Table 4.5). At a fixed sensitivity of 90%, specificities ranged from 0.21 (95% CI 0.17-0.26) for the de Araujo1 signature to 0.54 (95% CI 0.49-0.59) for Sambarey10; none met the minimum criterion threshold of 70% specificity (Table 4.5). In this study population with a 30% TB prevalence and a sensitivity of 90%, PPVs ranged from 0.36 (95% CI 0.32-0.41) to 0.51 (95% CI 0.45-0.58) and NPVs ranged from 0.81 (95% CI 0.72-0.88) to 0.91 (95% CI 0.87-0.95).

Table 4.5: Diagnostic performance of 20 published host response TB transcriptomic signatures, ordered by area under the ROC curve for all TB patients vs ORD controls.

Signature	TB¹ vs ORD AUC (95% CI)	Sensitivity (95% CI) [at specificity of 70%²]	Specificity (95% CI) [at sensitivity of 90%³]	Positive predictive value (95% CI)⁴	Negative predictive value (95% CI)⁴	p-value vs signature with highest AUC	q-value (Benjamini Hochberg method) vs signature with highest AUC
Satproedprai7	0.83 (0.79-0.87)	0.81 (0.74-0.86)	0.48 (0.43-0.53)	0.46 (0.41-0.52)	0.91 (0.85-0.94)	NA	NA
Jacobsen3	0.83 (0.79-0.86)	0.80 (0.73-0.85)	0.52 (0.46-0.57)	0.48 (0.42-0.53)	0.91 (0.87-0.95)	0.316	0.441
Kaforou22	0.82 (0.78-0.86)	0.77 (0.69-0.83)	0.51 (0.45-0.57)	0.48 (0.42-0.54)	0.91 (0.85-0.94)	0.226	0.350
Roe1	0.82 (0.78-0.86)	0.80 (0.73-0.85)	0.49 (0.44-0.54)	0.48 (0.42-0.53)	0.91 (0.86-0.94)	0.051	0.115
Suliman2	0.82 (0.78-0.86)	0.78 (0.71-0.83)	0.49 (0.44-0.55)	0.47 (0.42-0.53)	0.91 (0.86-0.94)	0.263	0.385
Sambarey10	0.81 (0.77-0.85)	0.81 (0.74-0.86)	0.54 (0.49-0.59)	0.49 (0.44-0.55)	0.92 (0.87-0.95)	0.029	0.074
Penn-Nicholson6	0.81 (0.77-0.84)	0.80 (0.73-0.85)	0.45 (0.40-0.50)	0.45 (0.40-0.50)	0.90 (0.85-0.94)	0.002	0.008
Duffy9	0.81 (0.76-0.86)	0.81 (0.73-0.87)	0.52 (0.46-0.59)	0.51 (0.45-0.58)	0.90 (0.84-0.94)	0.079	0.155
Gliddon3	0.80 (0.75-0.85)	0.78 (0.71-0.85)	0.51 (0.45-0.57)	0.48 (0.41-0.54)	0.91 (0.85-0.95)	0.080	0.157
Suliman4	0.79 (0.75-0.84)	0.77 (0.70-0.83)	0.44 (0.38-0.49)	0.44 (0.39-0.50)	0.90 (0.84-0.94)	0.020	0.052
da Costa3	0.79 (0.75-0.83)	0.75 (0.68-0.81)	0.39 (0.34-0.44)	0.43 (0.38-0.48)	0.89 (0.83-0.93)	0.003	0.012
Maertzdorf4	0.79 (0.75-0.83)	0.78 (0.71-0.83)	0.49 (0.43-0.54)	0.46 (0.41-0.51)	0.91 (0.86-0.94)	0.001	0.002
Roe3	0.79 (0.75-0.83)	0.77 (0.70-0.83)	0.37 (0.32-0.42)	0.42 (0.37-0.47)	0.88 (0.82-0.92)	0.001	0.001
GjØen7	0.79 (0.74-0.83)	0.76 (0.69-0.82)	0.32 (0.28-0.37)	0.40 (0.35-0.45)	0.87 (0.80-0.91)	0.001	0.004
Sweeney3	0.79 (0.74-0.83)	0.75 (0.68-0.81)	0.40 (0.35-0.45)	0.43 (0.38-0.48)	0.89 (0.83-0.93)	0.003	0.012
Rajan5	0.76 (0.71-0.81)	0.68 (0.59-0.75)	0.41 (0.35-0.47)	0.45 (0.39-0.51)	0.88 (0.81-0.93)	0.001	0.001
Thompson5	0.76 (0.71-0.80)	0.69 (0.62-0.76)	0.40 (0.35-0.45)	0.42 (0.37-0.47)	0.89 (0.83-0.93)	0.003	0.001
Francisco2	0.74 (0.69-0.78)	0.71 (0.64-0.77)	0.26 (0.22-0.31)	0.38 (0.33-0.43)	0.84 (0.76-0.90)	0.001	0.001
Gliddon4	0.73 (0.68-0.79)	0.71 (0.62-0.78)	0.26 (0.21-0.32)	0.40 (0.34-0.45)	0.83 (0.73-0.90)	0.001	0.001
de Araujo1	0.68 (0.63-0.73)	0.61 (0.54-0.68)	0.21 (0.17-0.26)	0.36 (0.32-0.41)	0.81 (0.72-0.88)	0.001	0.001

TB = tuberculosis. ORD = other respiratory diseases. AUC = area under the curve. CI = confidence interval.

¹Definite and probable TB combined

²Minimum specificity criterion of the WHO TPP for a triage test (70%)

³Minimum sensitivity criterion of the WHO TPP for a triage test (90%)

⁴Based on a prevalence of 30% and the specificity at a sensitivity of 90%

HIV status and country are associated with signature scores

Multivariable linear regression was used to identify clinical and demographic features associated with signature scores. HIV positivity and definite or probable TB disease were independently associated with higher signature scores for most signatures; among the top 9 performing signatures, only Jacobsen3 scores were not significantly associated with HIV status (**Table 4.6**). When disaggregated by HIV status, AUC values for all signatures were higher in HIV-negative than HIV-positive participants (**Tables 4.7 - 4.8**).

Country was also a strong predictor of signature scores. We therefore sought to investigate signature scores and performance when disaggregated by country. Due to the low numbers of TB cases (n=4) and ORD controls (n=7), data from The Gambia were excluded from these analyses (**Table 4.2**). It was notable that signature scores differed considerably by country, in particular ORD controls from Namibia had significantly higher scores for most signatures than those for other countries (**Figure 4.8**). This influenced diagnostic performance and most signatures had higher AUCs in Ethiopia, Malawi, and South Africa than Uganda and Namibia (**Figure 4.8** and **Tables 4.9 - 4.10**). Consistent with this, several signatures met the minimum TPP criteria for a triage test in Ethiopia, Malawi, or South Africa. For example, Gliddon3 and Penn-Nicholson6 met the TPP criteria in Ethiopia, Malawi, and South Africa; Satproedprai7 met these criteria in Ethiopia and Malawi; and Sambarey10 met these criteria in Ethiopia (**Figure 4.8**).

Table 4.6: Regression analyses determining clinical and demographic variables associated with signature scores.

Variable	Duffy9 β-coefficient (95% CI); adjusted p- value	Gliddon3 β-coefficient (95% CI); adjusted p- value	Jacobsen3 β-coefficient (95% CI); adjusted p- value	Kaforou22 β-coefficient (95% CI); adjusted p- value	Penn-Nicholson6 β-coefficient (95% CI); adjusted p-value	Roe1 β-coefficient (95% CI); adjusted p- value	Sambarey10 β-coefficient (95% CI); adjusted p- value	Satproedprai7 β-coefficient (95% CI); adjusted p- value	Suliman2 β-coefficient (95% CI); adjusted p- value	Suliman4 β-coefficient (95% CI); adjusted p- value
Age (per 10 years)	0.00 (-0.02, 0.03); 0.92	0.00 (-0.02, 0.02); 0.90	0.00 (-0.03, 0.02); 0.94	0.01 (-0.01, 0.03); 0.79	-0.01 (-0.03, 0.01); 0.42	-0.01 (-0.03, 0.01); 0.65	0.00 (-0.02, 0.03); 0.90	-0.01 (-0.03, 0.02); 0.85	0.00 (-0.02, 0.02); 0.99	0.01 (-0.02, 0.03); 0.79
Sex										
Female	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Male	0.02 (-0.04, 0.08); 0.74	-0.02 (-0.07, 0.03); 0.67	-0.07 (-0.13, -0.01); 0.05	-0.01 (-0.06, 0.03); 0.79	-0.03 (-0.08, 0.01); 0.37	-0.04 (-0.08, 0.01); 0.31	-0.06 (-0.12, 0.00); 0.13	-0.06 (-0.12, -0.01); 0.06	0.00 (-0.05, 0.04); 0.93	-0.02 (-0.07, 0.03); 0.74
HIV status										
Negative	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Positive	0.12 (0.05, 0.19); 0.01	0.10 (0.05, 0.16); <0.001	0.07 (0.01, 0.14); 0.11	0.09 (0.03, 0.14); 0.01	0.10 (0.04, 0.15); <0.001	0.08 (0.03, 0.13); 0.01	0.09 (0.02, 0.16); 0.04	0.13 (0.07, 0.19); <0.001	0.09 (0.04, 0.14); 0.01	0.08 (0.01, 0.14); 0.06
TB status										
ORD controls	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Definite TB	0.30 (0.23, 0.36); <0.001	0.25 (0.20, 0.30); <0.001	0.30 (0.24, 0.36); <0.001	0.27 (0.22, 0.32); <0.001	0.27 (0.22, 0.31); <0.001	0.33 (0.28, 0.38); <0.001	0.36 (0.30, 0.42); <0.001	0.34 (0.29, 0.40); <0.001	0.26 (0.22, 0.31); <0.001	0.27 (0.22, 0.32); <0.001
Probable TB	0.21 (0.08, 0.34); 0.01	0.15 (0.06, 0.25); 0.01	0.20 (0.08, 0.32); 0.01	0.18 (0.09, 0.28); <0.001	0.16 (0.06, 0.25); 0.01	0.17 (0.07, 0.27); <0.001	0.20 (0.08, 0.32); 0.01	0.21 (0.10, 0.32); <0.001	0.18 (0.09, 0.27); <0.001	0.17 (0.07, 0.28); 0.01
Smoking history										
Never smoked	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Current smoker	-0.04 (-0.14, 0.05); 0.67	0.02 (-0.06, 0.09); 0.86	0.00 (-0.08, 0.08); 1.00	0.03 (-0.04, 0.10); 0.68	0.00 (-0.07, 0.07); 0.99	-0.02 (-0.08, 0.05); 0.86	-0.04 (-0.12, 0.05); 0.67	0.01 (-0.07, 0.09); 0.93	-0.01 (-0.08, 0.05); 0.86	0.01 (-0.07, 0.09); 0.90
Past smoker	-0.04 (-0.16, 0.08); 0.77	-0.06 (-0.16, 0.03); 0.41	0.01 (-0.09, 0.10); 0.95	-0.01 (-0.09, 0.06); 0.90	-0.02 (-0.10, 0.06); 0.84	-0.05 (-0.13, 0.02); 0.41	-0.01 (-0.10, 0.09); 0.95	0.01 (-0.08, 0.09); 0.94	-0.03 (-0.11, 0.04); 0.67	0.01 (-0.08, 0.10); 0.90
Previous TB										
No	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Yes	0.01 (-0.07, 0.09); 0.93	0.00 (-0.07, 0.06); 0.99	0.00 (-0.08, 0.07); 0.99	-0.02 (-0.08, 0.04); 0.85	-0.01 (-0.07, 0.05); 0.89	0.02 (-0.04, 0.08); 0.77	0.01 (-0.06, 0.08); 0.90	0.01 (-0.05, 0.08); 0.86	0.00 (-0.06, 0.06); 0.99	0.00 (-0.07, 0.07); 0.99
Country										
Uganda*	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Ethiopia	0.02 (-0.05, 0.08); 0.85	0.00 (-0.05, 0.05); 0.98	0.04 (-0.02, 0.10); 0.46	0.02 (-0.02, 0.07); 0.65	0.03 (-0.02, 0.08); 0.47	0.00 (-0.06, 0.05); 0.93	0.00 (-0.06, 0.06); 0.99	-0.01 (-0.07, 0.05); 0.90	0.06 (0.01, 0.10); 0.07	0.06 (0.01, 0.12); 0.07
Namibia	0.17 (0.07, 0.26); <0.001	0.10 (0.03, 0.18); 0.03	0.18 (0.10, 0.27); <0.001	0.18 (0.11, 0.25); <0.001	0.19 (0.12, 0.25); <0.001	0.14 (0.07, 0.20); <0.001	0.16 (0.07, 0.24); <0.001	0.13 (0.06, 0.21); <0.001	0.14 (0.08, 0.21); <0.001	0.19 (0.11, 0.27); <0.001
South Africa	0.03 (-0.09, 0.15); 0.85	-0.01 (-0.10, 0.08); 0.93	0.03 (-0.08, 0.14); 0.85	-0.01 (-0.10, 0.09); 0.95	0.00 (-0.09, 0.09); 0.98	-0.02 (-0.11, 0.06); 0.83	0.04 (-0.07, 0.15); 0.74	-0.05 (-0.15, 0.05); 0.65	0.03 (-0.05, 0.12); 0.70	0.01 (-0.09, 0.11); 0.91
The Gambia	-0.13 (-0.33, 0.08); 0.50	-0.10 (-0.27, 0.07); 0.54	-0.12 (-0.30, 0.06); 0.42	-0.08 (-0.22, 0.06); 0.57	-0.04 (-0.18, 0.09); 0.78	-0.05 (-0.19, 0.10); 0.77	-0.10 (-0.27, 0.07); 0.53	-0.04 (-0.21, 0.13); 0.86	-0.03 (-0.16, 0.11); 0.88	-0.11 (-0.26, 0.04); 0.35

Signatures are listed in alphabetical order. CI = confidence interval. *Uganda was chosen as the reference because it has the highest number of both cases and controls.

Table 4.7: Numbers of TB cases and ORD controls, by HIV status.

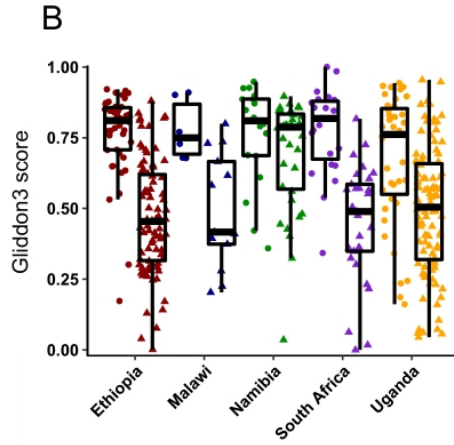
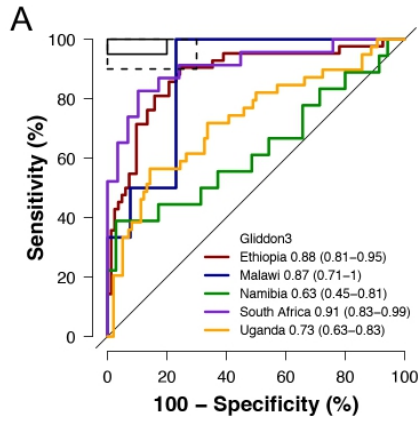
Signature	HIV positive		HIV negative	
	TB	ORD	TB	ORD
da Costa3	47	96	133	264
de Araujo1	47	96	133	265
Duffy9	35	55	94	176
Francisco2	47	96	133	265
GjØen7	47	96	133	264
Gliddon3	33	66	97	195
Gliddon4	35	57	98	190
Jacobsen3	44	91	126	256
Kaforou22	39	78	108	211
Maertzdorf4	45	96	126	256
Penn-Nicholson6	47	95	131	263
Rajan5	36	58	97	192
Roe1	46	96	133	253
Roe3	46	96	133	253
Sambarey10	46	95	132	263
Satproedprai7	42	91	125	243
Suliman2	43	86	122	243
Suliman4	42	86	123	246
Sweeney3	47	96	133	264
Thompson5	46	94	126	262

TB = tuberculosis. ORD = other respiratory diseases.

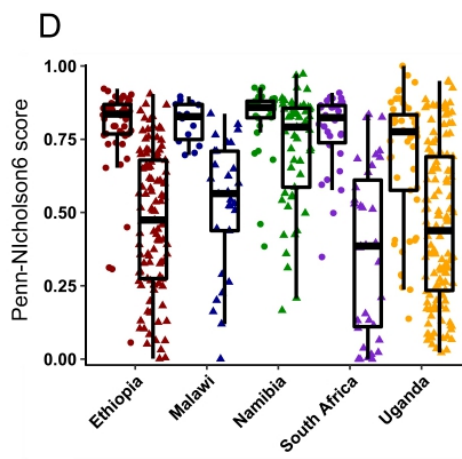
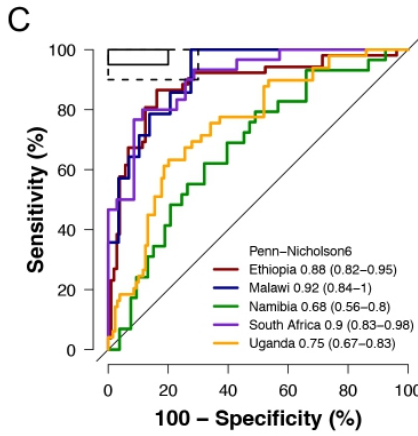
Table 4.8: Diagnostic performance of candidate transcriptomic signatures for all TB cases vs ORD controls, disaggregated by HIV status.

Signature	HIV positive AUC (95% CI)	HIV negative AUC (95% CI)
da Costa3	0.72 (0.63-0.81)	0.81 (0.76-0.85)
de Araujo1	0.56 (0.46-0.67)	0.72 (0.66-0.77)
Duffy9	0.74 (0.63-0.85)	0.83 (0.78-0.88)
Francisco2	0.65 (0.56-0.75)	0.76 (0.71-0.81)
GjØen7	0.72 (0.63-0.81)	0.81 (0.76-0.86)
Gliddon3	0.72 (0.61-0.83)	0.83 (0.78-0.88)
Gliddon4	0.72 (0.60-0.83)	0.75 (0.68-0.81)
Jacobsen3	0.78 (0.70-0.87)	0.84 (0.80-0.89)
Kaforou22	0.75 (0.65-0.85)	0.85 (0.80-0.89)
Maertzdorf4	0.76 (0.68-0.85)	0.82 (0.77-0.87)
Penn-Nicholson6	0.74 (0.66-0.82)	0.83 (0.79-0.88)
Rajan5	0.71 (0.60-0.82)	0.78 (0.73-0.84)
Roe1	0.76 (0.67-0.84)	0.84 (0.80-0.89)
Roe3	0.73 (0.64-0.82)	0.81 (0.76-0.86)
Sambarey10	0.75 (0.67-0.84)	0.84 (0.79-0.88)
Satproedprai7	0.79 (0.70-0.87)	0.86 (0.82-0.90)
Suliman2	0.73 (0.64-0.82)	0.85 (0.81-0.89)
Suliman4	0.72 (0.63-0.81)	0.82 (0.77-0.86)
Sweeney3	0.74 (0.64-0.83)	0.81 (0.76-0.85)
Thompson5	0.74 (0.66-0.83)	0.77 (0.72-0.82)

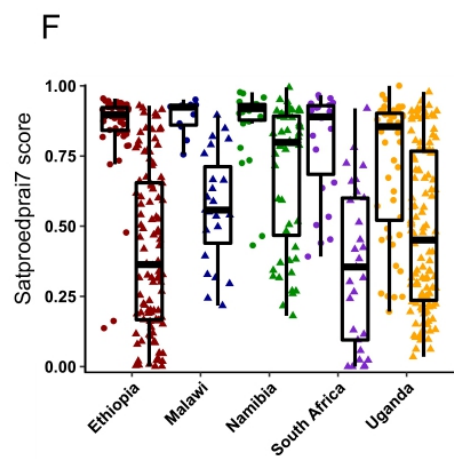
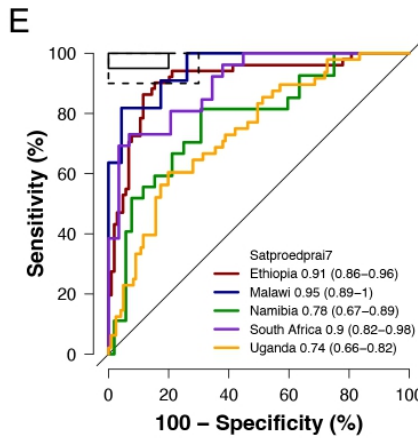
TB = tuberculosis. ORD = other respiratory diseases. AUC = area under the curve. CI = confidence interval



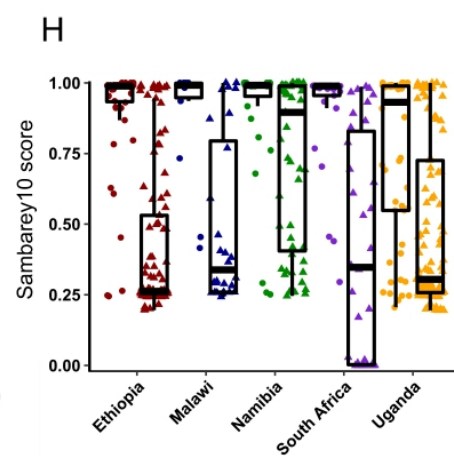
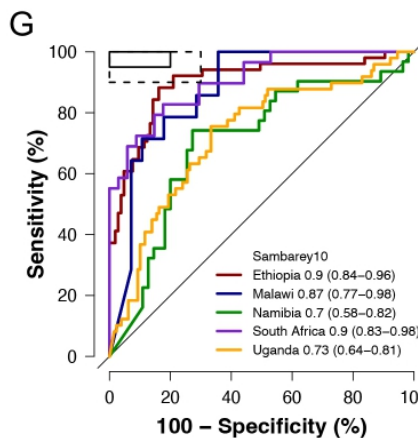
	Ethi.	Mala.	Nam.	S.A.
Mala.	1.00			
Nam.	<0.0001	0.02		
S.A.	1.00	1.00	<0.0001	
Ugan	1.00	1.00	<0.0001	1.00



	Ethi.	Mala.	Nam.	S.A.
Mala.	1.00			
Nam.	<0.0001	<0.0001		
S.A.	0.56	0.24	<0.0001	
Ugan	1.00	1.00	<0.0001	1.00



	Ethi.	Mala.	Nam.	S.A.
Mala.	0.14			
Nam.	<0.0001	0.15		
S.A.	1.00	0.09	<0.0001	
Ugan	0.57	1.00	<0.0001	0.40



	Ethi.	Mala.	Nam.	S.A.
Mala.	0.26			
Nam.	<0.0001	<0.0001		
S.A.	1.00	1.00	<0.0001	
Ugan	0.44	1.00	<0.0001	0.93

Figure 4.8: Signature score distributions and diagnostic performance of transcriptomic signatures for all TB cases and controls with other respiratory diseases (ORD), stratified by country.

Receiver operating characteristic (ROC) curves (A, C, E, and G) and box-and-whisker plots (B, D, F, and H) for (A-B) Gliddon3, (C-D) Penn-Nicholson6, (E-F) Satproedprai7, and (G-H) Sambarey10 for distinguishing between all TB cases and ORD controls from Ethiopia, Malawi, Namibia, South Africa, and Uganda. The Gambia was excluded from individual country analyses due to very small sample size ($n = 11$; **Table 4.2**). Diagnostic performance (area under the curve, 95% CI) disaggregated by country is shown in the ROC plots and for all signatures in **Table 4.10**. Box-and-whisker plots show signature score distribution (each dot represents a participant) between TB cases and controls in Ethiopia, Malawi, Namibia, South Africa, and Uganda. The tables in B, D, F, and H show q-values (Benjamini-Hochberg adjusted p-values) for comparison of signature scores in control groups between the countries, calculated using Mann-Whitney U test. Boxes depict the IQR, the midline represents the medians, and the whiskers represent $1.5 \times$ IQR.

Ethi = Ethiopia. Mala = Malawi. Nam = Namibia. SA = South Africa. Ugan = Uganda.

Table 4.9: Numbers of TB cases and ORD controls, by country.

Signature	Ethiopia		Malawi		Namibia		South Africa		Uganda	
	TB	ORD	TB	ORD	TB	ORD	TB	ORD	TB	ORD
da Costa3	52	105	14	29	31	55	30	35	49	129
de Araujo1	52	105	14	29	31	55	30	35	49	130
Duffy9	42	77	8	9	18	31	24	25	34	86
Francisco2	52	105	14	29	31	55	30	35	49	130
GjØen7	52	105	14	29	31	55	30	35	49	129
Gliddon3	42	82	6	13	18	35	23	29	39	98
Gliddon4	42	79	8	9	20	32	25	28	35	96
Jacobsen3	52	104	9	26	29	52	29	34	48	124
Kaforou22	49	95	10	15	21	41	22	23	42	109
Maertzdorf4	52	105	14	29	28	54	24	28	49	129
Penn-Nicholson6	52	105	14	29	29	53	30	35	49	129
Rajan5	42	80	8	10	21	32	24	28	35	97
Roe1	52	103	14	26	30	55	30	33	49	126
Roe3	52	103	14	26	30	55	30	33	49	126
Sambarey10	51	105	14	28	31	55	29	34	49	129
Satproedprai7	51	104	11	23	27	52	26	29	48	121
Suliman2	50	101	11	23	26	45	26	32	48	122
Suliman4	51	103	12	23	24	47	25	30	49	122
Sweeney3	52	105	14	29	31	55	30	34	49	130
Thompson5	52	105	12	27	28	52	28	35	48	130

TB = tuberculosis. ORD = other respiratory diseases. Only 4 Definite TB cases and 7 ORC controls were available from the Gambia, precluding per-country diagnostic performance analyses.

Table 4.10: Diagnostic performance of candidate transcriptomic signatures for all TB cases vs ORD controls, disaggregated by country.

Signature	Ethiopia AUC (95% CI)	Malawi AUC (95% CI)	Namibia AUC (95% CI)	South Africa AUC (95% CI)	Uganda AUC (95% CI)
da Costa3	0.85 (0.78-0.92)	0.80 (0.64-0.97)	0.73 (0.62-0.84)	0.89 (0.82-0.97)	0.70 (0.61-0.79)
de Araujo1	0.69 (0.59-0.78)	0.74 (0.57-0.90)	0.69 (0.57-0.81)	0.72 (0.58-0.85)	0.65 (0.55-0.74)
Duffy9	0.88 (0.82-0.95)	0.83 (0.63-1)	0.71 (0.55-0.87)	0.89 (0.79-0.98)	0.74 (0.64-0.85)
Francisco2	0.82 (0.75-0.88)	0.70 (0.53-0.87)	0.65 (0.53-0.77)	0.76 (0.62-0.89)	0.68 (0.59-0.78)
GjØen7	0.85 (0.79-0.92)	0.87 (0.77-0.98)	0.78 (0.67-0.89)	0.89 (0.80-0.97)	0.66 (0.56-0.76)
Gliddon3	0.88 (0.81-0.95)	0.87 (0.71-1)	0.63 (0.45-0.81)	0.91 (0.83-0.99)	0.73 (0.63-0.83)
Gliddon4	0.83 (0.75-0.91)	0.82 (0.58-1)	0.60 (0.43-0.78)	0.68 (0.53-0.83)	0.70 (0.59-0.81)
Jacobsen3	0.91 (0.85-0.96)	0.91 (0.77-1)	0.74 (0.62-0.85)	0.90 (0.83-0.97)	0.77 (0.70-0.85)
Kaforou22	0.91 (0.86-0.96)	0.77 (0.57-0.96)	0.75 (0.60-0.90)	0.84 (0.71-0.97)	0.79 (0.71-0.87)
Maertzdorf4	0.87 (0.80-0.93)	0.85 (0.73-0.96)	0.72 (0.60-0.84)	0.87 (0.77-0.98)	0.74 (0.66-0.82)
Penn-Nicholson6	0.88 (0.82-0.95)	0.92 (0.84-1)	0.68 (0.56-0.80)	0.90 (0.83-0.98)	0.75 (0.67-0.83)
Rajan5	0.79 (0.71-0.87)	0.75 (0.50-1)	0.77 (0.63-0.90)	0.74 (0.61-0.88)	0.71 (0.61-0.81)
Roe1	0.88 (0.83-0.94)	0.85 (0.72-0.99)	0.81 (0.71-0.91)	0.92 (0.86-0.99)	0.73 (0.64-0.81)
Roe3	0.87 (0.80-0.93)	0.82 (0.67-0.96)	0.77 (0.66-0.88)	0.89 (0.81-0.97)	0.67 (0.57-0.76)
Sambarey10	0.90 (0.84-0.96)	0.87 (0.77-0.98)	0.70 (0.58-0.82)	0.90 (0.83-0.98)	0.73 (0.64-0.81)
Satproedprai7	0.91 (0.86-0.96)	0.95 (0.89-1)	0.78 (0.67-0.89)	0.90 (0.82-0.98)	0.74 (0.66-0.82)
Suliman2	0.89 (0.84-0.95)	0.76 (0.58-0.94)	0.77 (0.65-0.89)	0.86 (0.76-0.95)	0.77 (0.70-0.85)
Suliman4	0.90 (0.84-0.95)	0.83 (0.69-0.97)	0.58 (0.44-0.72)	0.78 (0.65-0.91)	0.77 (0.69-0.85)
Sweeney3	0.85 (0.78-0.91)	0.76 (0.61-0.92)	0.75 (0.62-0.87)	0.85 (0.75-0.95)	0.73 (0.64-0.81)
Thompson5	0.88 (0.82-0.93)	0.78 (0.62-0.94)	0.65 (0.52-0.77)	0.79 (0.67-0.90)	0.71 (0.63-0.80)

TB = tuberculosis. ORD = other respiratory diseases. AUC = area under the curve. CI = confidence interval.

4.5: Discussion

We sought to determine the diagnostic accuracy of 20 candidate blood transcriptomic TB signatures in symptomatic individuals presenting at primary care facilities. Our results show that 9 transcriptomic signatures had equivalent diagnostic performance in differentiating between patients with TB from those who presented with ORD. All signatures had reduced performance for clinically defined, probable TB, than for microbiologically-confirmed definite TB. In pooled analysis of all countries, none of the signatures met the minimum TPP criteria for a triage test for TB. However, multiple signatures met these criteria in certain countries. Among the clinical and demographic factors explored, HIV positivity, independent of TB classification, was associated with higher signature scores, leading to reduced diagnostic performance of transcriptomic signatures.

We found that, compared to Satproedprai7, the signature with the highest AUC, eight other signatures had statistically indistinguishable diagnostic performance. These included signatures trained on diverse cohorts from different settings, derived using different training strategies and distinct scoring algorithms and signatures that range from a single mRNA transcript to 22 transcripts. This finding suggests that it is unlikely that there is a single “best-performing” signature or strategy. It also calls into question the value of developing yet more signatures for pulmonary TB, especially if such signature discovery approaches follow the approach that has been taken in most studies, where the most differentially expressed genes are selected. More innovative statistical models, such as multinomial signatures, may also perform better. It therefore remains to be established exactly where existing signatures may add most value.

Diagnostic biomarkers that can identify those with TB who do not have a definitive diagnosis based on traditional diagnostic tests, who cannot expectorate sputum, including young children, who have extra pulmonary TB, or who have microbiologically negative subclinical TB, would allow more accurate diagnosis. Many such individuals traditionally do not receive a diagnosis, yet would potentially benefit from TB treatment and may transmit *Mtb* to others (Esmail et al., 2022). Reduced diagnostic performance of TB transcriptomic signatures in those with probable TB has been reported previously (Hoang et al., 2021). Since transcriptomic signatures reflect the host inflammatory response consistent with TB, the lower diagnostic performance of transcriptomic signatures in unconfirmed TB likely reflects an earlier or less pronounced pathological or inflammatory state, as has also been reported for asymptomatic TB (Mendelsohn et al., 2022b, 2022a; Scriba et al., 2021a). It is possible that diagnostic pathogen markers may offer better diagnostic performance than host response markers (Hoang et al., 2021), although it is likely that *Mtb* load in such difficult-to-identify cases is lower than in sputum test-positive, symptomatic cases of pulmonary TB.

A prospective, observational study that compared diagnostic accuracy of 27 transcriptomic signatures in symptomatic adults who presented to a TB clinic found that four signatures met the minimum TPP criteria, but not the optimum criteria for a triage test; these signatures also did not meet the minimum criteria for a confirmatory test (Turner et al., 2020). When assessed as an add-on confirmatory test for patients that had positive Xpert MTB/RIF results, transcriptomic signatures improved specificity (Turner et al., 2020), illustrating one use-case for signatures for TB diagnosis under certain conditions.

Genetic background is known to influence candidate genes identified in transcriptomic studies as well as performance of transcriptomic signatures. When genetic backgrounds are similar, a greater level of overlap is expected between traits (Sarup et al., 2011). The transcriptomic signatures evaluated have been discovered and validated in different populations and yet certain transcripts have been found to comprise more than one signature. The fact that the same genes are the most differentially expressed in TB cases in different studies is a simple reflection of conserved innate responses and this occurs in

very diverse populations. SNPs and eQTLs that regulate inflammation and disease severity are genetic factors that need to be studied further as these can ultimately influence gene expression and the performance of transcriptomic signatures.

The level of correlation between signature scores was not surprising, given the extent of common transcripts, predominated by ISGs (Berry et al., 2010b; Singhania et al., 2018c), that constitute these signatures. For example, FCGR1A is present in six signatures (da Costa et al., 2015; Gliddon et al., 2021; Jacobsen et al., 2007; Kaforou et al., 2013; Sambarey et al., 2017; Satproedprai et al., 2015) and GBP5 in five of the 20 signatures (da Costa et al., 2015; Francisco et al., 2017; Gjøen et al., 2017a; Roe et al., 2016; Sweeney et al., 2016). Our findings further underpin the consistent involvement of type I interferons in TB pathogenesis (Qiu et al., 2022). Several other infections and conditions, most notably viral infections (Mulenga et al., 2022; Singhania et al., 2018a), such as SARS-CoV-2 (Ortego-Prieto & Jimenez-Guardeno, 2024) and HIV (Darboe et al., 2019; Esmail et al., 2018), induce expression of ISGs, leading to erosion of the specificity of ISG-based TB signatures. Capturing a wider scope of immune responses and specifically TB-associated genes that are not IFN-induced, may lead to better specificity of signatures (Esmail et al., 2018; Scagnolari and Antonelli, 2018).

As most signatures on our panel predominantly included ISGs, the effect of HIV infection on signature scores and diagnostic performance was not surprising. Although plasma viral load was not included in our regression analysis, due to a high number of missing values, others have shown an association between detectable HIV viral loads and elevated scores for ISG-based signatures (Mendelsohn et al., 2022b)(Darboe et al., 2019; Penn-Nicholson et al., 2020a).

The identification of other biological pathways and transcripts less affected by HIV may have better utility in HIV-infected individuals and predominantly HIV-infected populations. Genes involved in the complement pathway are promising biomarkers for TB (Cai et al., 2014; Esmail et al., 2018; Lubbers et al., 2020, 2018). It was notable that Thompson5 (Thompson et al., 2017b), a treatment response signature that contains no ISGs, showed only a minimal reduction in AUC in HIV-positive vs HIV-negative participants. This is consistent with findings of our previous study that Thompson5 scores were not affected by HIV viral load (Mendelsohn et al., 2022b). One transcript within the Thompson signature belongs to the category of long noncoding RNAs. These transcripts are known to modulate immune responses (Y. G. Chen et al., 2017; Wei et al., 2021), and have now also been highlighted as potential biomarkers for TB (Zhang et al., 2022).

Satproedprai7 (Satproedprai et al., 2015), the signature with the best diagnostic performance in these African cohorts, was discovered in an adult population from India. Many of the signatures evaluated here, alongside many others not included, have either been developed or validated predominantly in South African populations (Gliddon et al., 2021; Kaforou et al., 2013; Penn-Nicholson et al., 2020a; Sambarey et al., 2017; Suliman et al., 2018; Thompson et al., 2017b) or using publicly available datasets (Anderson et al., 2014; Berry et al., 2010b; Bloom et al., 2013; Kaforou et al., 2013) generated from studies conducted within South Africa (Duffy et al., 2019; Maertzdorf et al., 2016; Satproedprai et al., 2015; Sweeney et al., 2016). This likely accounts for the good performance of most transcriptomic signatures, especially those that met the minimal WHO TPP, in South Africa. It was puzzling, however, that signature performance in neighbouring Namibia, was consistently poorer than in South Africa. We do not know the reason for this but speculate that respiratory infections in Namibian participants with ORD may be markedly different to those in South Africa, Ethiopia, or Malawi. An alternative explanation may be that those who present to clinics in some countries have more severe disease than those in other settings. The ScreenTB and AE-TBC studies were designed to minimise this by using highly standardised protocols and predefined TB classification algorithms across countries, but implementation and program settings may have been different. We also cannot rule out the possibility

that more controls in some settings could have had undiagnosed TB and may have been misclassified or missed as TB cases during follow-up. Season is unlikely to have affected performance since study participants, across all sites, were recruited within the same period across several seasons. Overall, what is clear is that differences present across African countries may be attributed to more than just variations in TB burden and host genetics. Enroth and colleagues for example report that up to about 56% of the variation in biomarker levels between individuals are a result of genetic, clinical and lifestyle factors. (Enroth et al., 2014). Well-designed multi-site biomarker studies that address these underlying factors are needed to understand this better.

Alongside co-morbidities such as HIV, diabetes mellitus and alcohol consumption, malnutrition is another factor that puts people at higher risk of progressing to TB. Human studies have shown the effect that malnutrition has on the immune system such as a decrease in Th1 and pro-inflammatory cytokines (Sinha et al., 2019). There is still limited information regarding how immune responses during a state of malnutrition may affect the expression of transcriptomic signatures that predict risk of progression to TB. VanValkenburg and colleagues sought to address this through analysis of transcriptomic data from household contacts of TB cases with and without malnutrition (VanValkenburg et al., 2022). Household contacts with malnutrition had higher levels of inflammation and higher scores for all 4 transcriptomic signatures analysed in this study.

Through the development of a TBSignatureProfiler, Johnson and colleagues sought to determine the effect of malnutrition on the performance of over 45 transcriptomic signatures (Johnson et al., 2021). The data from the TBSignatureProfiler was applied to a cohort of malnourished TB cases and household contacts with LTBI from India. Transcriptomic signatures differentiated TB cases from LTBI controls with AUCs ranging from 0.662 - 0.989. This led the authors to conclude that current signatures for TB can work even in the context of malnutrition. They noted, however, that only a few smaller signatures did not perform as well alongside signatures trained in particular settings (Johnson et al., 2021). Up to 12 signatures evaluated by the authors were included in our work. This would demonstrate that the signatures selected can be robust and perform similarly in different settings. The TBSignatureProfiler has the potential to enable researchers to assess the effects that several other comorbidities and any additional factors could have on the performance of not only existing but also newly developed signatures.

Strengths of our study were the unbiased selection and evaluation of candidate transcriptomic signatures, inclusion of cohorts from multiple African countries representing diverse geographies and populations, and a diagnostic study design that compared TB cases to appropriate controls within the real-world setting where POC tests are needed. Inclusion of probable TB cases was important to ascertain performance of signatures in those who are difficult to diagnose.

Limitations of our study included insufficient information on the cause of respiratory or other infections or diseases in the control group or in the cases, or comorbidities such as diabetes, and to what degree this may have differed by country setting. It is likely that these may have contributed to the differences in signature performance at the different sites. It is likely that these may have contributed to the differences in signature performance at the different sites. A limitation of the analysis done to compare the performance of different signatures is that failed quantification of transcripts in some signatures for some samples led to very different sample sizes for certain signatures, ranging from 360 samples for Duffy9 to 541 samples for de Araujo1. Because inter-signature performance required pair-wise comparisons between Satproedprai7, which had the highest AUC, and each other signature, the sample sizes for each comparison were different. Consequently, it is important to interpret the results of the statistical comparisons of signature AUCs with caution. We also cannot rule out that diagnostic performance of Gliddon4 (Gliddon et al., 2021), Duffy9 (Duffy et

al., 2019), and Rajan5 (Rajan et al., 2019) may have been biased by poor performance of the primer/probe assays for GBP6, which failed in a considerable number of samples. Although the primer-probe assay that represented GBP6 had been published and used in by Gliddon et al. (Gliddon et al., 2021), it did not meet our assay qualification pass criteria.

The pass rate of samples and thus signatures evaluated in this study cohort is a result of both the quality of primer-probe assays selected to represent each transcript as well as the amount of RNA extracted from whole blood samples. From previous work done within SATVI, RNA extraction on the semi-automated TECAN platform has been found to generate lower RNA concentrations in comparison to manual extractions. Lower RNA concentrations in some samples could have evidently decreased the sensitivity of our multiplex RT-qPCR platform to accurately detect gene expression of certain transcripts (Adams, 2020; Kuang et al., 2018; Park et al., 2022). This is also reflected in the number of samples that were excluded from analysis due to failure to fully meet our QC criteria.

In conclusion, we report a head-to-head comparison of multiple signatures for TB, demonstrating their utility within a wider African context. It is critical that transcriptomic signatures be assessed in larger studies of unconfirmed TB, and well-designed studies of extrapulmonary and paediatric TB, where there is an urgent need for better diagnostics.

4.6: Contributors/Author Contributions

NNC and TJS conceived the study. NNC, SKM, and TJS raised funds and provided resources. GW, STM, and members the ScreenTB and AE-TBC study teams were responsible for all site-level activities, including recruitment, clinical management, sample and data collection. NB, TF, MF and VL provided operational or laboratory support and project management. VMM, ME, and RR processed samples and performed the experiments. The data was verified by SKM, KS, GT, and GVS modelled and verified data, then analysed by VMM and SCM. VMM, SCM, and TJS interpreted the results. VMM, SCM, and TJS wrote the manuscript. All authors have read and approved the manuscript.

Declaration of interests

GW and TJS report grants from the Bill & Melinda Gates Foundation during the conduct of the study. TJS and NNC report grants from the South African Medical Research Council during the conduct of the study. GW reports grants from the South African National Research Foundation and EDCTP. TJS has patents of the RISK11 (Darboe11), RISK6 (Penn-Nicholson6) and RISK4 (Suliman4) signatures issued. GW and NC have patents “TB diagnostic markers” (PCT/IB2013/ 054377), “Serum host biomarkers for tuberculosis disease” [PCT/IB2017/052142], “Method for diagnosing TB” (PCT/IB2017/052142) granted but receive no royalties on these patents. All other authors had nothing to disclose.

Data sharing

Deidentified signature scores, PCR probe data, clinical metadata, and TB endpoint data have been deposited in Zivahub (doi.org/10.25375/uct.25650933), an open access data repository hosted by the University of Cape Town’s institutional data repository powered by Figshare for Institutions.

Acknowledgements

We thank Mr Miguel Rodo and Dr Denis Awany for providing statistical support. We acknowledge the substantial contributions of the ‘African European Tuberculosis Consortium (AE-TBC)’ and ‘Evaluation of host biomarker-based point-of-care tests for targeted screening for active TB (ScreenTB)’ study teams (lists of consortia members and their affiliations appear in the supplementary information). Both AE-TBC and ScreenTB were funded by the European and Developing Countries Clinical Trials Partnership (EDCTP) under the 1st and 2nd European & Developing Countries Clinical Trials Partnership Programmes respectively (Grant Agreement Numbers IP_2009_32040 and DRIA2014-311). This project was supported by the Strategic Health Innovation Partnerships (SHIP) Unit of the South African Medical Research Council (SAMRC) with funds received from the South African Department of Science and Technology. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. VMM is a recipient of PhD funding from University of Cape Town through the incoming international students’ bursary scheme.

References

- Adams, G., 2020. A beginner's guide to RT-PCR, qPCR and RT-qPCR. *The Biochemist* 42, 48–53. <https://doi.org/10.1042/BIO20200034>
- Anderson, S.T., Kaforou, M., Brent, A.J., Wright, V.J., Banwell, C.M., Chagaluka, G., Crampin, A.C., Dockrell, H.M., French, N., Hamilton, M.S., Hibberd, M.L., Kern, F., Langford, P.R., Ling, L., Mlotha, R., Ottenhoff, T.H.M., Pienaar, S., Pillay, V., Scott, J.A.G., Twahir, H., Wilkinson, R.J., Coin, L.J., Heyderman, R.S., Levin, M., Eley, B., 2014. Diagnosis of Childhood Tuberculosis and Host RNA Expression in Africa. *N. Engl. J. Med.*
- Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* 57, 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Berry, M.P.R., Graham, C.M., McNab, F.W., Xu, Z., Bloch, S.A.A., Oni, T., Wilkinson, K.A., Banchereau, R., Skinner, J., Wilkinson, R.J., Quinn, C., Blankenship, D., Dhawan, R., Cush, J.J., Mejias, A., Ramilo, O., Kon, O.M., Pascual, V., Banchereau, J., Chaussabel, D., O'Garra, A., 2010. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature* 466, 973–977. <https://doi.org/10.1038/nature09247>
- Bloom, C.I., Graham, C.M., Berry, M.P.R., Rozakeas, F., Redford, P.S., Wang, Y., Xu, Z., Wilkinson, K.A., Wilkinson, R.J., Kendrick, Y., Devouassoux, G., Ferry, T., Miyara, M., Bouvry, D., Dominique, V., Gorocho, G., Blankenship, D., Saadatian, M., Vanhems, P., Beynon, H., Vancheeswaran, R., Wickremasinghe, M., Chaussabel, D., Banchereau, J., Pascual, V., Ho, L., Lipman, M., O'Garra, A., 2013. Transcriptional Blood Signatures Distinguish Pulmonary Tuberculosis, Pulmonary Sarcoidosis, Pneumonias and Lung Cancers. *PLOS ONE* 8, e70630. <https://doi.org/10.1371/journal.pone.0070630>
- Cai, Y., Yang, Q., Tang, Y., Zhang, M., Liu, H., Zhang, G., Deng, Q., Huang, J., Gao, Z., Zhou, B., Feng, C.G., Chen, X., 2014. Increased Complement C1q Level Marks Active Disease in Human Tuberculosis. *PLOS ONE* 9, e92340. <https://doi.org/10.1371/journal.pone.0092340>
- Chegou, N.N., Sutherland, J.S., Malherbe, S., Crampin, A.C., Corstjens, P.L.A.M., Geluk, A., Mayanja-Kizza, H., Loxton, A.G., Spuy, G. van der, Stanley, K., Kotzé, L.A., Vyver, M. van der, Rosenkrands, I., Kidd, M., Helden, P.D. van, Dockrell, H.M., Ottenhoff, T.H.M., Kaufmann, S.H.E., Walzl, G., 2016. Diagnostic performance of a seven-marker serum protein biosignature for the diagnosis of active TB disease in African primary healthcare clinic attendees with signs and symptoms suggestive of TB. *Thorax* 71, 785–794. <https://doi.org/10.1136/thoraxjnl-2015-207999>
- Chegou, N.N., Sutherland, J.S., Namuganga, A.-R., Corstjens, P.L., Geluk, A., Gebremichael, G., Mendy, J., Malherbe, S., Stanley, K., van der Spuy, G.D., Kriel, M., Loxton, A.G., Kriel, B., Simukonda, F., Bekele, Y., Sheehama, J.A., Nelongo, J., van der Vyver, M., Gebrexabher, A., Hailu, H., Esterhuyse, M.M., Rosenkrands, I., Aagard, C., Kidd, M., Kassa, D., Mihret, A., Howe, R., Cliff, J.M., Crampin, A.C., Mayanja-Kizza, H., Kaufmann, S.H.E., Dockrell, H.M., Ottenhoff, T.H.M., Walzl, G., 2018. Africa-wide evaluation of host biomarkers in QuantiFERON supernatants for the diagnosis of pulmonary tuberculosis. *Sci. Rep.* 8, 2675. <https://doi.org/10.1038/s41598-018-20855-7>
- Chen, Y.G., Satpathy, A.T., Chang, H.Y., 2017. Gene regulation in the immune system by long noncoding RNAs. *Nat. Immunol.* 18, 962–972. <https://doi.org/10.1038/ni.3771>
- da Costa, L.L., Delcroix, M., Dalla Costa, E.R., Prestes, I.V., Milano, M., Francis, S.S., Unis, G., Silva, D.R., Riley, L.W., Rossetti, M.L.R., 2015. A real-time PCR signature to discriminate between tuberculosis and other pulmonary diseases. *Tuberc. Edinb. Scotl.* 95, 421–425. <https://doi.org/10.1016/j.tube.2015.04.008>
- Darboe, F., Mbandi, S.K., Naidoo, K., Yende-Zuma, N., Lewis, L., Thompson, E.G., Duffy, F.J., Fisher, M., Filander, E., van Rooyen, M., Bilek, N., Mabwe, S., McKinnon, L.R., Chegou, N., Loxton, A.,

- Walzl, G., Tromp, G., Padayatchi, N., Govender, D., Hatherill, M., Karim, S.A., Zak, D.E., Penn-Nicholson, A., Scriba, T.J., Team, T.S.C.I., 2019. Detection of Tuberculosis Recurrence, Diagnosis and Treatment Response by a Blood Transcriptomic Risk Signature in HIV-Infected Persons on Antiretroviral Therapy. *Front. Microbiol.* 10. <https://doi.org/10.3389/fmicb.2019.01441>
- DeLong, E.R., DeLong, D.M., Clarke-Pearson, D.L., 1988. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* 44, 837–845. <https://doi.org/10.2307/2531595>
- Duffy, F.J., Thompson, E.G., Scriba, T.J., Zak, D.E., 2019. Multinomial modelling of TB/HIV co-infection yields a robust predictive signature and generates hypotheses about the HIV+TB+ disease state. *PLOS ONE* 14, e0219322. <https://doi.org/10.1371/journal.pone.0219322>
- Esmail, H., Lai, R.P., Lesosky, M., Wilkinson, K.A., Graham, C.M., Horswell, S., Coussens, A.K., Barry, C.E., O'Garra, A., Wilkinson, R.J., 2018. Complement pathway gene activation and rising circulating immune complexes characterize early disease in HIV-associated tuberculosis. *Proc. Natl. Acad. Sci.* 115, E964–E973. <https://doi.org/10.1073/pnas.1711853115>
- Esmail, H., Macpherson, L., Coussens, A.K., Houben, R.M.G.J., 2022. Mind the gap – Managing tuberculosis across the disease spectrum. *eBioMedicine* 78. <https://doi.org/10.1016/j.ebiom.2022.103928>
- Enroth, S., Johansson, Å., Enroth, S.B., Gyllensten, U., 2014. Strong effects of genetic and lifestyle factors on biomarker variation and use of personalized cutoffs. *Nat. Commun.* 5, 4684. <https://doi.org/10.1038/ncomms5684>
- Francisco, N.M., Fang, Y.-M., Ding, L., Feng, S., Yang, Y., Wu, M., Jacobs, M., Ryffel, B., Huang, X., 2017. Diagnostic accuracy of a selected signature gene set that discriminates active pulmonary tuberculosis and other pulmonary diseases. *J. Infect.* 75, 499–510. <https://doi.org/10.1016/j.jinf.2017.09.012>
- GjØen, J.E., Jenum, S., Sivakumaran, D., Mukherjee, A., Macaden, R., Kabra, S.K., Lodha, R., Ottenhoff, T.H.M., Haks, M.C., Doherty, T.M., Ritz, C., Grewal, H.M.S., 2017. Novel transcriptional signatures for sputum-independent diagnostics of tuberculosis in children. *Sci. Rep.* 7, 1–9. <https://doi.org/10.1038/s41598-017-05057-x>
- Gliddon, H.D., Kaforou, M., Alikian, M., Habgood-Coote, D., Zhou, C., Oni, T., Anderson, S.T., Brent, A.J., Crampin, A.C., Eley, B., Heyderman, R., Kern, F., Langford, P.R., Ottenhoff, T.H.M., Hibberd, M.L., French, N., Wright, V.J., Dockrell, H.M., Coin, L.J., Wilkinson, R.J., Levin, M., 2021. Identification of Reduced Host Transcriptomic Signatures for Tuberculosis Disease and Digital PCR-Based Validation and Quantification. *Front. Immunol.* 12.
- Hoang, L.T., Jain, P., Pillay, T.D., Tolosa-Wright, M., Niazi, U., Takwoingi, Y., Halliday, A., Berrocal-Almanza, L.C., Deeks, J.J., Beverley, P., Kon, O.M., Lalvani, A., 2021. Transcriptomic signatures for diagnosing tuberculosis in clinical practice: a prospective, multicentre cohort study. *Lancet Infect. Dis.* 21, 366–375. [https://doi.org/10.1016/S1473-3099\(20\)30928-2](https://doi.org/10.1016/S1473-3099(20)30928-2)
- Jacobsen, M., Repsilber, D., Gutschmidt, A., Neher, A., Feldmann, K., Mollenkopf, H.J., Ziegler, A., Kaufmann, S.H.E., 2007. Candidate biomarkers for discrimination between infection and disease caused by *Mycobacterium tuberculosis*. *J. Mol. Med.* 85, 613–621. <https://doi.org/10.1007/s00109-007-0157-6>
- Johnson, W.E., Odom, A., Cintron, C., Muthaiah, M., Knudsen, S., Joseph, N., Babu, S., Lakshminarayanan, S., Jenkins, D.F., Zhao, Y., Nankya, E., Horsburgh, C.R., Roy, G., Ellner, J., Sarkar, S., Salgame, P., Hochberg, N.S., 2021. Comparing tuberculosis gene signatures in malnourished individuals using the TBSignatureProfiler. *BMC Infect. Dis.* 21, 106. <https://doi.org/10.1186/s12879-020-05598-z>
- Kaforou, M., Wright, V.J., Oni, T., French, N., Anderson, S.T., Bangani, N., Banwell, C.M., Brent, A.J., Crampin, A.C., Dockrell, H.M., Eley, B., Heyderman, R.S., Hibberd, M.L., Kern, F., Langford, P.R., Ling, L., Mendelson, M., Ottenhoff, T.H., Zgambo, F., Wilkinson, R.J., Coin, L.J., Levin, M., 2013. Detection of Tuberculosis in HIV-Infected and -Uninfected African Adults Using Whole

- Blood RNA Expression Signatures: A Case-Control Study. *PLOS Med.* 10, e1001538. <https://doi.org/10.1371/journal.pmed.1001538>
- Kuang, J., Yan, X., Genders, A.J., Granata, C., Bishop, D.J., 2018. An overview of technical considerations when using quantitative real-time PCR analysis of gene expression in human exercise research. *PLOS ONE* 13, e0196438. <https://doi.org/10.1371/journal.pone.0196438>
- Lubbers, R., Sutherland, J.S., Goletti, D., de Paus, R.A., Dijkstra, D.J., van Moorsel, C.H.M., Veltkamp, M., Vestjens, S.M.T., Bos, W.J.W., Petrone, L., Malherbe, S.T., Walzl, G., Gelderman, K.A., Groeneveld, G.H., Geluk, A., Ottenhoff, T.H.M., Joosten, S.A., Trouw, L.A., 2020. Expression and production of the SERPING1-encoded endogenous complement regulator C1-inhibitor in multiple cohorts of tuberculosis patients. *Mol. Immunol.* 120, 187–195. <https://doi.org/10.1016/j.molimm.2020.02.006>
- Lubbers, R., Sutherland, J.S., Goletti, D., de Paus, R.A., van Moorsel, C.H.M., Veltkamp, M., Vestjens, S.M.T., Bos, W.J.W., Petrone, L., Del Nonno, F., Bajema, I.M., Dijkman, K., Verreck, F.A.W., Walzl, G., Gelderman, K.A., Groeneveld, G.H., Geluk, A., Ottenhoff, T.H.M., Joosten, S.A., Trouw, L.A., 2018. Complement Component C1q as Serum Biomarker to Detect Active Tuberculosis. *Front. Immunol.* 9, 2427. <https://doi.org/10.3389/fimmu.2018.02427>
- MacLean, E., Broger, T., Yerlikaya, S., Fernandez-Carballo, B.L., Pai, M., Denking, C.M., 2019. A systematic review of biomarkers to detect active tuberculosis. *Nat. Microbiol.* 4, 748–758. <https://doi.org/10.1038/s41564-019-0380-2>
- Maertzdorf, J., McEwen, G., Weiner, J., Tian, S., Lader, E., Schriek, U., Mayanja-Kizza, H., Ota, M., Kenneth, J., Kaufmann, S.H., 2016. Concise gene signature for point-of-care classification of tuberculosis. *EMBO Mol. Med.* 8, 86–95. <https://doi.org/10.15252/emmm.201505790>
- Manngo, P.M., Gutschmidt, A., Snyders, C.I., Mutavhatsindi, H., Manyelo, C.M., Makhoba, N.S., Ahlers, P., Hiemstra, A., Stanley, K., McAnda, S., Kidd, M., Malherbe, S.T., Walzl, G., Chegou, N.N., 2019. Prospective evaluation of host biomarkers other than interferon gamma in QuantiFERON Plus supernatants as candidates for the diagnosis of tuberculosis in symptomatic individuals. *J. Infect.* 79, 228–235. <https://doi.org/10.1016/j.jinf.2019.07.007>
- Mendelsohn, S.C., Fiore-Gartland, A., Awany, D., Mulenga, H., Mbandi, S.K., Tameris, M., Walzl, G., Naidoo, K., Churchyard, G., Scriba, T.J., Hatherill, M., 2022a. Clinical predictors of pulmonary tuberculosis among South African adults with HIV. *eClinicalMedicine* 45. <https://doi.org/10.1016/j.eclinm.2022.101328>
- Mendelsohn, S.C., Mbandi, S.K., Fiore-Gartland, A., Penn-Nicholson, A., Musvosvi, M., Mulenga, H., Fisher, M., Hadley, K., Erasmus, M., Nombida, O., Tameris, M., Walzl, G., Naidoo, K., Churchyard, G., Hatherill, M., Scriba, T.J., 2022b. Prospective multicentre head-to-head validation of host blood transcriptomic biomarkers for pulmonary tuberculosis by real-time PCR. *Commun. Med.* 2, 26. <https://doi.org/10.1038/s43856-022-00086-8>
- Mulenga, H., Fiore-Gartland, A., Mendelsohn, S.C., Penn-Nicholson, A., Mbandi, S.K., Borate, B., Musvosvi, M., Tameris, M., Walzl, G., Naidoo, K., Churchyard, G., Scriba, T.J., Hatherill, M., 2022. The effect of host factors on discriminatory performance of a transcriptomic signature of tuberculosis risk. *eBioMedicine* 77, 103886. <https://doi.org/10.1016/j.ebiom.2022.103886>
- Nogueira, B.M.F., Krishnan, S., Barreto-Duarte, B., Araújo-Pereira, M., Queiroz, A.T.L., Ellner, J.J., Salgame, P., Scriba, T.J., Sterling, T.R., Gupta, A., Andrade, B.B., 2022. Diagnostic biomarkers for active tuberculosis: progress and challenges. *EMBO Mol. Med.* e14088. <https://doi.org/10.15252/emmm.202114088>
- Ortega-Prieto AM, Jimenez-Guardeño JM. 2024. Interferon-stimulated genes and their antiviral activity against SARS-CoV-2. *mBio* 15:e02100-24. <https://doi.org/10.1128/mbio.02100-24>
- Park, H., Jung, W., Jang, H., Namkoong, K., Choi, K.-Y., 2022. One-Step RT-qPCR for Viral RNA Detection Using Digital Analysis. *Front. Bioeng. Biotechnol.* 10.
- Penn-Nicholson, A., Mbandi, S.K., Thompson, E., Mendelsohn, S.C., Suliman, S., Chegou, N.N., Malherbe, S.T., Darboe, F., Erasmus, M., Hanekom, W.A., Bilek, N., Fisher, M., Kaufmann, S.H.E., Winter, J., Murphy, M., Wood, R., Morrow, C., Van Rhijn, I., Moody, B., Murray, M.,

- Andrade, B.B., Sterling, T.R., Sutherland, J., Naidoo, K., Padayatchi, N., Walzl, G., Hatherill, M., Zak, D., Scriba, T.J., 2020a. RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response. *Sci. Rep.* 10, 8629. <https://doi.org/10.1038/s41598-020-65043-8>
- Qiu, Q., Peng, A., Zhao, Y., Liu, D., Liu, C., Qiu, S., Xu, J., Cheng, H., Xiong, W., Chen, Y., 2022. Diagnosis of pulmonary tuberculosis via identification of core genes and pathways utilizing blood transcriptional signatures: a multicohort analysis. *Respir. Res.* 23, 125. <https://doi.org/10.1186/s12931-022-02035-4>
- Rajan, J.V., Semitala, F.C., Mehta, T., Seielstad, M., Montalvo, L., Andama, A., Asege, L., Nakaye, M., Katende, J., Mwebe, S., Kanya, M.R., Yoon, C., Cattamanchi, A., 2019. A Novel, 5-Transcript, Whole-blood Gene-expression Signature for Tuberculosis Screening Among People Living With Human Immunodeficiency Virus. *Clin. Infect. Dis.* 69, 77–83. <https://doi.org/10.1093/cid/ciy835>
- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., Müller, M., 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12, 77. <https://doi.org/10.1186/1471-2105-12-77>
- Roe, J.K., Thomas, N., Gil, E., Best, K., Tsaliki, E., Morris-Jones, S., Stafford, S., Simpson, N., Witt, K.D., Chain, B., Miller, R.F., Martineau, A., Noursadeghi, M., 2016. Blood transcriptomic diagnosis of pulmonary and extrapulmonary tuberculosis. *JCI Insight* 1. <https://doi.org/10.1172/jci.insight.87238>
- Sambarey, A., Devaprasad, A., Mohan, A., Ahmed, A., Nayak, S., Swaminathan, S., D'Souza, G., Jesuraj, A., Dhar, C., Babu, S., Vyakarnam, A., Chandra, N., 2017. Unbiased Identification of Blood-based Biomarkers for Pulmonary Tuberculosis by Modeling and Mining Molecular Interaction Networks. *eBioMedicine* 15, 112–126. <https://doi.org/10.1016/j.ebiom.2016.12.009>
- Sarup, P., Sørensen, J.G., Kristensen, T.N., Hoffmann, A.A., Loeschcke, V., Paige, K.N., Sørensen, P., 2011. Candidate Genes Detected in Transcriptome Studies Are Strongly Dependent on Genetic Background. *PLOS ONE* 6, e15644. <https://doi.org/10.1371/journal.pone.0015644>
- Satproedprai, N., Wichukchinda, N., Suphankong, S., Inunchot, W., Kuntima, T., Kumpeerasart, S., Wattanapokayakit, S., Nedsuwan, S., Yanai, H., Higuchi, K., Harada, N., Mahasirimongkol, S., 2015. Diagnostic value of blood gene expression signatures in active tuberculosis in Thais: a pilot study. *Genes Immun.* 16, 253–260. <https://doi.org/10.1038/gene.2015.4>
- Scagnolari, C., Antonelli, G., 2018. Type I interferon and HIV: Subtle balance between antiviral activity, immunopathogenesis and the microbiome. *Cytokine Growth Factor Rev.* 40, 19–31. <https://doi.org/10.1016/j.cytogfr.2018.03.003>
- Scriba, T.J., Fiore-Gartland, A., Penn-Nicholson, A., Mulenga, H., Mbandi, S.K., Borate, B., Mendelsohn, S.C., Hadley, K., Hikuam, C., Kaskar, M., Musvosvi, M., Bilek, N., Self, S., Sumner, T., White, R.G., Erasmus, Mzwandile, Jaxa, L., Raphela, R., Innes, C., Brumskine, W., Hiemstra, Andriëtte, Malherbe, S.T., Hassan-Moosa, R., Tameris, M., Walzl, G., Naidoo, K., Churchyard, G., Hatherill, M., Baepanye, K., Baepanye, T., Clarke, K., Collignon, M., Dlamini, A., Eyre, C., Feni, T., Fikizolo, M., Galane, P., Goliath, T., Gangat, A., Malefo-Grootboom, S., Rensburg, E.J. van, Rensburg, B.J. van, Kekana, S., Zietsman, M., Kock, A., Kunene, I., Lakhi, A., Langa, N., Ledwaba, H., Lophoko, M., Mabasa, I., Mabe, D., Mabuza, N., Majola, M., Makhetha, M., Makoanyane, M., Makhubalo, B., Malay, V., Market, J., Matshego, S., Mbipa, N., Mmotsa, T., Modipa, S., Mopati, S., Moswegu, P., Mothaga, P., Muller, D., Nchwe, G., Nel, M., Nhlangulela, L., Ntamo, B., Ntoahae, L., Ntshauba, T., Sanyaka, N., Seabela, L., Selepe, P., Senne, M., Serake, M.G., Thlapi, M., Tshikovhi, V., Tswaile, L., Aswegen, A. van, Mbata, L., Takavamanya, C., Pinho, P., Mdlulu, J., Taljaard, M., Slabbert, N., Sayed, S., Nielson, T., Senne, M., Sein, N.N., Mbata, L., Govender, D., Chinappa, T., Zulu, M.I., Maphanga, N.B., Hlathi, S.R., Gumede, G.K., Shezi, T.Y., Maphanga, J.L., Jali, Z.P., Cwele, T., Gwamanda, N.Z.E.,

- Dlamini, C., Sing, Z.P.P., Ntanjana, N.G., Nzimande, S.S., Mbatha, S., Maharaj, B., Moosa, A., Corris, C.-M., Kafaar, F., Geldenhuys, H., Luabeya, A.K.K., Shenje, J., Botes, N., Rossouw, S., Africa, H., Diamond, B., Braaf, S., Stryers, S., Carstens, A., Jansen, R., Mabwe, S., Mulenga, H., Herling, R., Veldsman, A., Makhete, L., Steyn, M., Buhlungu, S., Erasmus, Margareth, Davids, I., Plaatjie, P., Companie, A., Ratangee, F., Veldtsman, H., Petersen, C., Abrahams, C., Moses, M., Kelepu, X., Gregg, Y., Swanepoel, L., Magawu, N., Cetywayo, N., Mactavie, L., Valley, H., Filander, E., Nqakala, N., Maasdorp, E., Khoury, J., Kriel, B., Smith, B., Muller, L., Tonsing, S., Loxton, A., Hiemstra, Andriette, Ahlers, P., Flinn, M., Chung, E., Chung, M., Sato, A., 2021. Biomarker-guided tuberculosis preventive therapy (CORTIS): a randomised controlled trial. *Lancet Infect. Dis.* 21, 354–365. [https://doi.org/10.1016/S1473-3099\(20\)30914-2](https://doi.org/10.1016/S1473-3099(20)30914-2)
- Signorell, A., Aho, K., Alfons, A., Anderegg, N., Aragon, T., Arachchige, C., Arppe, A., Baddeley, A., Barton, K., Bolker, B., Borchers, H.W., Caeiro, F., Champely, S., Chessel, D., Chhay, L., Cooper, N., Cummins, C., Dewey, M., Doran, H.C., Dray, S., Dupont, C., Eddelbuettel, D., Ekstrom, C., Elff, M., Enos, J., Farebrother, R.W., Fox, J., Francois, R., Friendly, M., Galili, T., Gamer, M., Gastwirth, J.L., Gegzna, V., Gel, Y.R., Graber, S., Gross, J., Grothendieck, G., Jr, F.E.H., Heiberger, R., Hoehle, M., Hoffmann, C.W., Hojsgaard, S., Hothorn, T., Huerzeler, M., Hui, W.W., Hurd, P., Hyndman, R.J., Jackson, C., Kohl, M., Korpela, M., Kuhn, M., Labes, D., Leisch, F., Lemon, J., Li, D., Maechler, M., Magnusson, A., Mainwaring, B., Malter, D., Marsaglia, G., Marsaglia, J., Matei, A., Meyer, D., Miao, W., Millo, G., Min, Y., Mitchell, D., Mueller, F., Naepflin, M., Navarro, D., Nilsson, H., Nordhausen, K., Ogle, D., Ooi, H., Parsons, N., Pavoiné, S., Plate, T., Prendergast, L., Rapold, R., Revelle, W., Rinker, T., Ripley, B.D., Rodriguez, C., Russell, N., Sabbe, N., Scherer, R., Seshan, V.E., Smithson, M., Snow, G., Soetaert, K., Stahel, W.A., Stephenson, A., Stevenson, M., Stubner, R., Templ, M., Lang, D.T., Therneau, T., Tille, Y., Torgo, L., Trapletti, A., Ulrich, J., Ushey, K., VanDerWal, J., Venables, B., Verzani, J., Iglesias, P.J.V., Warnes, G.R., Wellek, S., Wickham, H., Wilcox, R.R., Wolf, P., Wollschlaeger, D., Wood, J., Wu, Y., Yee, T., Zeileis, A., 2023. DescTools: Tools for Descriptive Statistics.
- Singhania, A., Verma, R., Graham, C.M., Lee, J., Tran, T., Richardson, M., Lecine, P., Leissner, P., Berry, M.P.R., Wilkinson, R.J., Kaiser, K., Rodrigue, M., Woltmann, G., Haldar, P., O'Garra, A., 2018a. A modular transcriptional signature identifies phenotypic heterogeneity of human tuberculosis infection. *Nat. Commun.* 9, 1–17. <https://doi.org/10.1038/s41467-018-04579-w>
- Singhania, A., Wilkinson, R.J., Rodrigue, M., Haldar, P., O'Garra, A., 2018b. Transcriptomics in TB: the immune response and diagnosis. *Nat. Immunol.* 19, 1159–1168. <https://doi.org/10.1038/s41590-018-0225-9>
- Sinha, P., Davis, J., Saag, L., Wanke, C., Salgame, P., Mesick, J., Horsburgh, C.R., Hochberg, N.S., 2019. Undernutrition and Tuberculosis: Public Health Implications. *J. Infect. Dis.* 219, 1356–1363. <https://doi.org/10.1093/infdis/jiy675>
- Stopsack, K.H., Tyekucheva, S., Wang, M., Gerke, T.A., Vasselkiv, J.B., Penney, K.L., Kantoff, P.W., Finn, S.P., Fiorentino, M., Loda, M., Lotan, T.L., Parmigiani, G., Mucci, L.A., 2021. Extent, impact, and mitigation of batch effects in tumor biomarker studies using tissue microarrays. *eLife* 10, e71265. <https://doi.org/10.7554/eLife.71265>
- Suliman, S., Thompson, E.G., Sutherland, J., Weiner, J., Ota, M.O.C., Shankar, S., Penn-Nicholson, A., Thiel, B., Erasmus, M., Maertzdorf, J., Duffy, F.J., Hill, P.C., Hughes, E.J., Stanley, K., Downing, K., Fisher, M.L., Valvo, J., Parida, S.K., van der Spuy, G., Tromp, Gerard, Adetifa, I.M.O., Donkor, S., Howe, R., Mayanja-Kizza, H., Boom, W.H., Dockrell, H.M., Ottenhoff, T.H.M., Hatherill, M., Aderem, A., Hanekom, W.A., Scriba, T.J., Kaufmann, S.H.E., Zak, D.E., Walzl, G., Walzl, G., Black, G.F., van der Spuy, G., Stanley, K., Kriel, M., Du Plessis, N., Nene, N., Roberts, T., Kleynhans, L., Gutschmidt, A., Smith, B., Nene, N., Loxton, A.G., Chegou, N.N., Tromp, Gerhardus, Tabb, D., Ottenhoff, T.H.M., Klein, M.R., Haks, M.C., Franken, K.L.M.C., Geluk, A., van Meijgaarden, K.E., Joosten, S.A., Boom, W.H., Thiel, B., Mayanja-Kizza, H., Joloba, M., Zalwango, S., Nsereko, M., Okwera, B., Kisingo, H., Kaufmann, S.H.E., Parida, S.K., Golinski, R., Maertzdorf, J., Weiner, J., Jacobson, M., Dockrell, H., Smith, S., Gorak-Stolinska, P., Hur,

- Y.-G., Lalor, M., Lee, J.-S., Crampin, A.C., French, N., Ngwira, B., Ben-Smith, A., Watkins, K., Ambrose, L., Simukonda, F., Mvula, H., Chilongo, F., Saul, J., Branson, K., Suliman, S., Scriba, T.J., Mahomed, H., Hughes, E.J., Bilek, N., Erasmus, M., Xasa, O., Veldsman, A., Downing, K., Fisher, M., Penn-Nicholson, A., Mulenga, H., Abel, B., Bowmaker, M., Kagina, B., Chung, W.K., Hanekom, W.A., Sadoff, J., Sizemore, D., Ramachandran, S., Barker, L., Brennan, M., Weichold, F., Muller, S., Geiter, L., Kassa, D., Abebe, A., Mesele, T., Tegbaru, B., Baarle, D. van, Miedema, F., Howe, R., Mihret, A., Aseffa, A., Bekele, Y., Iwnetu, R., Tafesse, M., Yamuah, L., Ota, M., Sutherland, J., Hill, P., Adegbola, R., Corrah, T., Antonio, M., Togun, T., Adetifa, I., Donkor, S., Andersen, P., Rosenkrands, I., Doherty, M., Welding, K., Schoolnik, G., Dolganov, G., Van, T., Kafaar, F., Workman, L., Mulenga, H., Scriba, T.J., Hughes, E.J., Bilek, N., Erasmus, M., Xasa, O., Veldsman, A., Cloete, Y., Abrahams, D., Moyo, S., Gelderbloem, S., Tameris, M., Geldenhuys, H., Hanekom, W., Hussey, G., Ehrlich, R., Verver, S., Geiter, L., 2018. Four-Gene Pan-African Blood Signature Predicts Progression to Tuberculosis. *Am. J. Respir. Crit. Care Med.* 197, 1198–1208. <https://doi.org/10.1164/rccm.201711-2340OC>
- Sweeney, T.E., Braviak, L., Tato, C.M., Khatri, P., 2016. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. *Lancet Respir. Med.* 4, 213–224. [https://doi.org/10.1016/S2213-2600\(16\)00048-5](https://doi.org/10.1016/S2213-2600(16)00048-5)
- Thompson, E.G., Du, Y., Malherbe, S.T., Shankar, S., Braun, J., Valvo, J., Ronacher, K., Tromp, G., Tabb, D.L., Alland, D., Shenai, S., Via, L.E., Warwick, J., Aderem, A., Scriba, T.J., Winter, J., Walzl, G., Zak, D.E., Du Plessis, N., Loxton, A.G., Chegou, N.N., Lee, M., 2017. Host blood RNA signatures predict the outcome of tuberculosis treatment. *Tuberculosis* 107, 48–58. <https://doi.org/10.1016/j.tube.2017.08.004>
- Turner, C.T., Gupta, R.K., Tsaliki, E., Roe, J.K., Mondal, P., Nyawo, G.R., Palmer, Z., Miller, R.F., Reeve, B.W., Theron, G., Noursadeghi, M., 2020. Blood transcriptional biomarkers for active pulmonary tuberculosis in a high-burden setting: a prospective, observational, diagnostic accuracy study. *Lancet Respir. Med.* 8, 407–419. [https://doi.org/10.1016/S2213-2600\(19\)30469-2](https://doi.org/10.1016/S2213-2600(19)30469-2)
- VanValkenburg, A., Kaipilyawar, V., Sarkar, S., Lakshminarayanan, S., Cintron, C., Prakash Babu, S., Knudsen, S., Joseph, N.M., Horsburgh, C.R., Sinha, P., Ellner, J.J., Narasimhan, P.B., Johnson, W.E., Hochberg, N.S., Salgame, P., 2022. Malnutrition leads to increased inflammation and expression of tuberculosis risk signatures in recently exposed household contacts of pulmonary tuberculosis. *Front. Immunol.* 13, 1011166. <https://doi.org/10.3389/fimmu.2022.1011166>
- Wei, L., Liu, K., Jia, Q., Zhang, H., Bie, Q., Zhang, B., 2021. The Roles of Host Noncoding RNAs in *Mycobacterium tuberculosis* Infection. *Front. Immunol.* 12.
- World Health Organization, 2014. High priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting, 28-29 April 2014, Geneva, Switzerland. World Health Organization, Geneva.
- Zhang, X., Chen, C., Xu, Y., 2022. Long Non-coding RNAs in Tuberculosis: From Immunity to Biomarkers. *Front. Microbiol.* 13, 883513. <https://doi.org/10.3389/fmicb.2022.883513>

Chapter 5: Evaluation of transcriptomic signatures as treatment monitoring tools for pulmonary paediatric TB

5.1 Introduction

In settings with a high TB burden and in which there is a very close link to TB and HIV epidemics, children contributed to up to 30% of incident cases (Hesseling et al., 2009; Schaaf et al., 2009, 2007). In 2022, children accounted for 12% of TB cases reported worldwide (World Health Organization, 2023).

The field of childhood and adolescent TB faces several challenges, including the very low number of cases detected through microbiological diagnosis, predicting disease progression in *Mtb*-exposed children, identifying children who do not respond to TB treatment, tailoring therapy to the severity of the disease, and identifying those who will develop disease-related morbidity (Kaforou et al., 2022). Monitoring responses to TB treatment in children includes assessing the resolution of TB-related symptoms, tracking weight gain and routine checks for sputum conversion. Sputum conversion from either culture or smear microscopy is considered the most important evidence of improvement. The drawback with this is that many children with TB do not spontaneously expectorate sputum and present with non-severe, paucibacillary disease that is often negative for TB microbiological tests (Turkova et al., 2022). This presents limitations for strategies to find, diagnose and effectively treat children (Perez-Velez and Marais, 2012). The higher risk for extrapulmonary disease in children also makes diagnosis and microbiological confirmation difficult (Anderson et al., 2014; Kaforou et al., 2022).

The low number of children with positive sputum smear, GeneXpert, and culture results lead to many of them being excluded from care (Perez-Velez and Marais, 2012). In clinical practice, children who do not respond well to treatment are often missed, due to the absence of adequate treatment monitoring tools that are not dependent on sputum (Kaforou et al., 2022). For example, while sputum conversion is often used as a marker of treatment response in adults, this is not possible to implement in children who do not present with microbiologically confirmed TB. Sputum-based tests have also not been able to identify children who require shorter, or individualized treatment durations. This gap, however, may be addressed by non-sputum-based treatment monitoring tools.

With the development of 'omics techniques, high throughput assays have been developed that have facilitated discovery of several promising TB biomarkers. An example of these includes host response biomarkers such as transcriptomic signatures that detect inflammation and reflect immune stimulation or activation. These biomarkers have also correlated with disease severity in the lungs, measured through PET-CT imaging (Penn-Nicholson et al., 2020; Thompson et al., 2017a). Now there are more possibilities to discover biomarkers that can identify treatment failure in children, as well as children most likely to relapse (Kaforou et al., 2022). Biomarkers that can discriminate between severe childhood TB cases and non-severe ones would aid efforts to personalize treatment. Furthermore, treatment response biomarkers that give an indication of when the inflammatory response has returned to pre-TB levels can guide decisions on when to stop treatment (Kaforou et al., 2022).

Published work has already shown the potential of transcriptomic signatures as diagnostic (Anderson et al., 2014; Gjøen et al., 2017; Verhagen et al., 2013), prognostic (Bobak et al., 2023; Zak et al., 2016), and treatment-monitoring biomarkers (Tornheim et al., 2020) in children. The measurement of these signatures in clinical settings can help address challenges faced with diagnosis and treatment-response monitoring of TB in children.

Studies have been carried out, mainly in adults, to evaluate whether transcriptomic signatures can track responses to anti-tubercular treatment (Bayaa et al., 2021; Doorn et al., 2022; Eckold et al., 2023; Penn-Nicholson et al., 2020; Thompson et al., 2017a), however there is limited evidence in children. To our knowledge, only one study has been carried out exclusively in children to assess the application of published signatures. This study, however, only included confirmed TB cases and compared their gene expression profiles to *Mtb* exposed household contacts (Tornheim et al., 2020). The diagnostic and treatment monitoring signatures that were identified outperformed the adult signatures that were included for comparison. Very little overlap in transcripts was present between the paediatric signatures that were discovered and the adult signatures. Age and sex of the children with confirmed TB had a significant impact on signature performance.

Data on treatment shortening in children is limited with recommendations predominantly derived from trials conducted in adults, particularly those with smear positive TB (Turkova et al., 2022). The application of personalized medicine for other conditions such as breast cancer and myeloid leukaemia has brought consideration for this in the field of TB (Kaforou et al., 2022). The paucibacillary nature of TB disease in children also offers the possibility for children to receive treatment for less than 6 months. The recent SHINE treatment-shortening clinical trial provided evidence that 4-months was non-inferior to 6-months of treatment for non-severe TB, demonstrating the potential of treatment shortening for paediatric TB based on current drug regimens (Turkova et al., 2022).

The work outlined in this thesis chapter would be the first conducted in children, across all three categories of TB: confirmed, unconfirmed and unlikely TB (Graham et al., 2015), evaluating changes in transcriptomic signatures during TB treatment.

The overall aim of this work was to determine the trend in gene expression over time in children with non-severe TB, recruited from South Africa under the SHINE trial. We compared changes in signature scores during treatment among children who had TB disease as well as signature scores between children randomized to the 4- and 6-month arms. We evaluated the impact of demographic factors as well as clinical parameters at enrolment on change in signature scores and compared the performance of these transcriptomic signatures across different timepoints, for their potential as possible TB treatment biomarkers. We hypothesized that significant changes in signature scores were more likely to occur in children with confirmed TB in comparison to those with unconfirmed TB and unlikely TB disease. We also hypothesized that there would be no difference in signature scores at the end of treatment, between children randomized to either the 4 month or 6-month treatment arm.

5.2 Methods

Study design

SHINE was an open-label, parallel-group, randomized, controlled trial that assessed whether a 4-month treatment regimen could be effective for children with non-severe TB (Chabala et al., 2018). It was a multi-site trial that took place in India, Uganda, Zambia, and South Africa. Children who were part of this trial were younger than 16 years of age and had non severe, symptomatic, smear-negative TB presumed to also be drug-sensitive (Chabala et al., 2018; Turkova et al., 2022). Non-severe TB comprised respiratory TB restricted to one lobe and had none of the following: cavities, signs of miliary TB, complex pleural effusion, airway obstruction, or peripheral lymph node TB.

Children were randomly allocated to either the 4- or 6-month arm and received standard first line treatment for TB based on WHO-recommended doses for children. At baseline, children were classified as having confirmed, unconfirmed, or unlikely TB by an independent endpoint review committee, using established international consensus criteria (Graham et al., 2015) (**Table 5.1**). During the 8-week intensive phase, all children received a fixed-dose combination of isoniazid, rifampin, and pyrazinamide, with or without ethambutol. In the continuation phase, children received a fixed-dose combination of isoniazid and rifampin for either 8 weeks in the 4-month arm or 16 weeks in the 6-month arm. The results from the trial demonstrated that 4 months of treatment was non-inferior to 6 months of treatment (Turkova et al., 2022).

Table 5.1: Classification of TB for children in the SHINE study

Confirmed TB	Unconfirmed TB	Unlikely TB
Bacteriological confirmation (culture or Xpert MTB/RIF assay) obtained	Bacteriological confirmation not obtained	Bacteriological confirmation not obtained
	Symptoms/signs suggestive of TB	Criteria for unconfirmed TB not met with or without <i>Mtb</i> infection.
	Chest radiograph consistent with TB	
	Close TB exposure or immunologic evidence of <i>Mtb</i> infection (TST and/or IGRA positive)	
	Positive response to TB Rx (based on documentation that gave evidence of a positive clinical response) with or without evidence of <i>Mtb</i> infection	

Adapted from Graham et al, 2015

This sub-study on assessing blood transcriptomic signatures was conducted solely on samples collected from the Desmond Tutu TB Centre (DTTC), Stellenbosch University in South Africa, because a limited number of PAXgene samples were collected from participants at the Ugandan and Indian sites (**Table 5.2**). This sub-study was therefore conducted on whole blood collected in PAXgene tubes at DTTC at weeks 0, 2, 8, and 16 (4-month treatment arm) or 24 (6-month treatment arm). Whole blood samples were also collected at the time of relapse/recurrence.

Table 5.2: Whole blood PAXgene tube samples available from study sites in Pune (India), Cape Town (South Africa) and Kampala (Uganda)

	Pune	Cape Town	Kampala
Week 0	9	252	42
Week 2	12	245	35
Week 4	0	23	1
Week 8	0	241	16
Week 12	0	1	0
Week 16	0	107	0
Week 20	0	0	0
Week 24	0	137	0
Week 28	0	0	0
Week 36	0	1	0
Week 48	0	2	0
Unscheduled	0	11	0

Ethics statement

The study protocol for the SHINE clinical trial was approved by the individual ethics committees at the each of the participating sites in South Africa (Stellenbosch University and University of Cape Town Research Ethics Committees), Uganda (Joint Clinical Research Centre Institutional Review Board), Zambia (University of Zambia Biomedical Research Ethics Committee and National Health Research Ethics Committee), India (BJMC Ethics Committee-Pune, National Institute of Research in Tuberculosis Ethics Committee-Chennai and Health Ministry Screening Committee), and by University College London (UCL Research Ethics Committee). The study also received ethical clearance from regulatory authorities in all the participating countries. The transcriptomic signatures sub-study protocol was reviewed and approved by the University of Cape Town HREC (522/2020).

Recruitment to the SHINE trial begun in 2016 and trial registration details were as follows; International Standard Randomised Controlled Trials Number: ISRCTN63579542 on 14 October 2014, Pan African Clinical Trials Registry Number: PACTR201505001141379 on 14 May 2015 and Clinical Trial Registry-India, registration number: CTRI/2017/07/009119 on 27 July 2017.

Parents or caregivers provided written informed consent while assent was provided by children whose comprehension was assessed as sufficient. In accordance with ICH-GCP, the confidentiality of all study participants was upheld through the entire course of the trial.

Transcriptomic signatures

The same panel of signatures highlighted in **Chapter 4** were evaluated in this cohort. Transcriptomic signatures were selected as described in **Section 3.1 of Chapter 3** and in the methods section of **Chapter 4**.

Signature measurement

Whole blood transcriptomic signatures were measured as previously described in **Sections 2.2 - 2.4 of Chapter 2** and in the literature (Mendelsohn et al., 2022; Scriba et al., 2021).

Briefly, RNA was extracted from cryopreserved PAXgene blood RNA tubes using the Maxwell HT simplyRNA Custom Kit (Promega, Madison, WI, USA), and cDNA synthesis performed using EpiScript reverse transcriptase (Lucigen, Hoddesdon, UK), in a semi-automated manner in 96-well plate format on the Freedom EVO 150 robotics platform (Tecan, Männedorf, Switzerland). Measurement of transcript expression was performed by RT-qPCR using 96.96 microfluidic gene expression chips (Standard BioTools, CA, USA) on the Biomark HD system (Standard BioTools, CA, USA), which multiplexes 96 samples with the 96 TaqMan primer-probe assays (**Table 3.9 of Chapter 3** and **Section 3.2 of Chapter 3**). An internal positive control (IPC) sample was included on all chips to monitor inter-chip batch variability. Gene expression data quality control, including evaluation of control samples, and calculation of signature scores, was performed using a locked-down R script (**Section 2.4 of Chapter 2**). Signature score calculations and batch correction methods were carried out as outlined in **Chapter 4**.

Blinding

The laboratory team were blinded to the TB status and the treatment arm assignment of the participants while samples were being processed. Clinical and demographic data (including TB disease category) were shared by collaborators only after all laboratory work was completed and signature scores had been calculated. A statistical analysis plan was generated, agreed upon, and signed off before any data analysis officially began.

Statistical analyses

Sample size was determined based on the availability of PAXgene whole blood samples from children recruited in the SHINE study. The sample size would be the largest so far in comparison to existing studies on transcriptomic signatures for treatment response monitoring in children. All statistical analyses were carried out in R version 4.1.3 (R Core Team 2022).

A p-value of <0.05 was considered significant for all analyses. Where appropriate, correction for multiple testing was performed using the Benjamini-Hochberg FDR method, or, for a small number of comparisons (below 5), the Bonferroni method.

Signature pass and fail rate was assessed to determine which signatures to exclude from analysis. Participant characteristic variables were assumed to be missing at random and participants with missing data points were excluded from analyses (i.e. complete case analysis only).

Participant demographic, clinical, and laboratory characteristics was summarized based on the TB diagnostic category. Pearson's Chi-squared test (categorical data), with values summarized as percentages, or Kruskal-Wallis /Mann-Whitney U/Wilcoxon Rank Sum (continuous data), with median and interquartile range reported, were used to test for any differences in demographic or baseline data.

An initial assumption was made that scores were not symmetrically distributed and therefore non-parametric tests were used. Where required and deemed necessary, however, an evaluation of whether signature scores and the study participants under different categories were symmetrically distributed was carried out before proceeding to apply either parametric or non-parametric tests. When comparing signature scores between two groups, the Mann-Whitney U-test was used. Kruskal-Wallis test was used to compare differences in signature scores across three or more groups. We used a paired, non-parametric test, the Wilcoxon signed-rank test, to compare scores across different time points in the same participants.

Study participants characteristics included in the multivariate linear regression model needed to be present for no less than 90% of study participants included in this sub-study. BCG vaccination status was excluded since all children in South Africa were immunized against it in childhood. Variables such as weight and age, that would be highly correlated, were not included together in the regression model. To better interpret the effect that age had on signature scores when included as a variable in the linear regression model, age was split into bands based on the median at baseline.

The pROC package (Robin et al., 2011) in R was used to calculate area under the receiver operating characteristic (ROC) curves (AUROC). AUROC were reported with 95% CI calculated using the DeLong method (DeLong et al., 1988) for each candidate signature when comparing paired scores, at different timepoints during treatment.

The zscorer package in R that makes use of WHO Child Growth Standards and Growth References for children and adolescents was used to generate anthropometric z-scores.

5.3: Results

Selection of study participants

Among 1204 study participants enrolled into the SHINE study, 315 were enrolled at the site in South Africa (**Figure 5.1**). Study participants from South Africa who were included in this sub-study ($n = 198$) were selected based on the availability of complete sets of PAXgene tubes at the 4 timepoints (weeks 0, 2, 8, and 16 or 24). Study participants from the Ugandan site and the Pune site in India were excluded due to limited availability of PAXgene tubes. Of the 198 study participants from South Africa included in this sub-study, 91 were in the 4-month treatment arm and 107 in the 6-month treatment arm (**Figure 5.1**). Transcriptomic signature scores from all 198 study participants were included in analyses. Due to very low numbers of children that relapsed or had treatment failure, a comparison of signature scores to children with treatment success was not conducted.

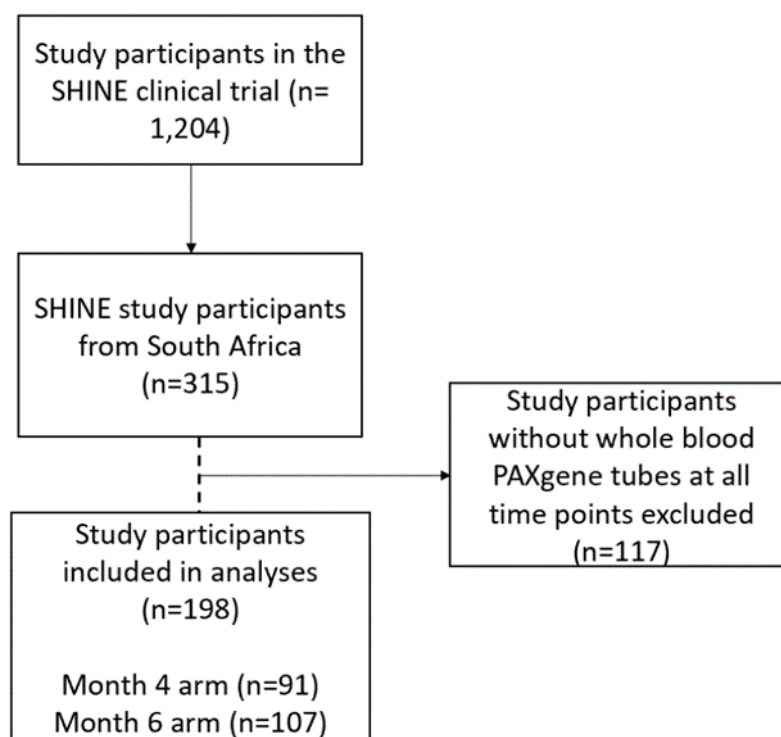


Figure 5.1: Flow diagram of participant enrolment, allocation, and analyses.

SHINE clinical trial participants who were included in the transcriptomic signature research study (Chabala et al., 2018; Turkova et al., 2022).

Study participant characteristics differ based on TB classification

A comparison of demographic and clinical variables based on TB classification showed differences in age ($p=0.002$), weight ($p=0.002$), clinical presentation ($p=0.003$) and chest x-ray (CXR) severity ($p<0.001$) (**Table 5.3**); children with confirmed TB weighed more (median weight of 18kg versus 13kg in unconfirmed TB cases and 12.5kg for children with unlikely TB), and a higher proportion were most were older than 3 years (29% versus 51.2% with unconfirmed TB and 37.5% with unlikely TB). Confirmed TB cases also had the highest proportion of children (11.4%) with a typical TB, severe CXR classification as compared to children with unconfirmed (1%) and unlikely TB (0%). A higher proportion of children with unconfirmed TB had clinical presentations of mixed (41.9%) and respiratory TB (58.1%) in comparison to children with confirmed and unlikely TB. Based on this, analyses of signature scores were performed within TB classification strata.

Table 5.3: Participant characteristics

Characteristic	Confirmed TB, N = 37	Unconfirmed TB, N = 129	Unlikely TB, N = 32	p-value [†]
Treatment arm, n (%)				0.31
4-month arm	14 (37.8)	59 (45.7)	18 (56.3)	
6-month arm	23 (62.2)	70 (54.3)	14 (43.8)	
Age, years, n (%)				0.002
> 3 years	29 (78.4)	66 (51.2)	12 (37.5)	
< 3 years	8 (21.6)	63 (48.8)	20 (62.5)	
Female sex, n (%)	16 (43.2)	66 (51.2)	17 (53.1)	0.65
Median weight at recruitment, kg (IQR)	18.0 (12.2-24.4)	13.0 (10.5-16.9)	12.5 (9.5-17.4)	0.002
BCG vaccination, n (%)	37 (100.0)	125 (96.9)	31 (96.9)	0.66
History of TB treatment, n (%)				0.28
Never treated	34 (91.9)	123 (95.3)	32 (100.0)	
Treated for TB >2years ago, n (%)	3 (8.1)	6 (4.7)	0 (0.0)	
HIV positive, n (%)	1 (2.7)	2 (1.6)	0 (0.0)	0.73
Clinical presentation, n (%)				0.003
Mixed TB	22 (59.5)	54 (41.9)	6 (18.8)	
Respiratory TB	15 (40.5)	75 (58.1)	26 (81.3)	
CXR severity, n (%)				<0.001
Normal	4 (11.4)	35 (33.0)	19 (63.3)	
Non-typical TB	7 (20.0)	23 (21.7)	11 (36.7)	
Typical TB, non- severe	20 (57.1)	47 (44.3)	0 (0.0)	
Typical TB, severe	4 (11.4)	1 (0.9)	0 (0.0)	
Missing	2	23	2	
Adherent to medication, n (%)	36 (97.3)	127 (98.4)	30 (93.8)	0.16
Endpoint status, n (%)				0.78
Favourable	36 (97.3)	122 (94.6)	31 (96.9)	
Unassessed	1 (2.7)	5 (3.9)	0 (0.0)	
Unfavourable	0 (0.0)	2 (1.6)	1 (3.1)	

TB = tuberculosis. CXR = chest x-ray. BCG = Bacillus Calmette-Guérin. IQR = interquartile range. [†]Pearson's Chi-squared test; Kruskal-Wallis rank sum test; Fisher's exact test

Pass-rate of transcriptomic signatures

Of the 20 signatures on the panel, 11 (Francisco2, Sweeney3, de Araujo1, Gj  en7, da Costa3, Roe3, Sambarey10, Suliman2, Penn-Nicholson6, Roe1, Thompson5) had a sample pass-rate of 100% (**Figure 5.2**). The sample pass-rates, for individual transcripts, measured in the SHINE sub-study cohort (**Table 5.4**) are higher, in comparison to the ScreenTB and AE-TBC studies (**Table 4.3, Chapter 4**). Primer-probes for only two transcripts on the panel had a sample pass-rate below 98%. These were C1QC.Hs00757779_m1 (78.91%) and the KLF2.KLF2_CUSTOM_ARH6CF4 (88.40%) (**Table 5.4**)

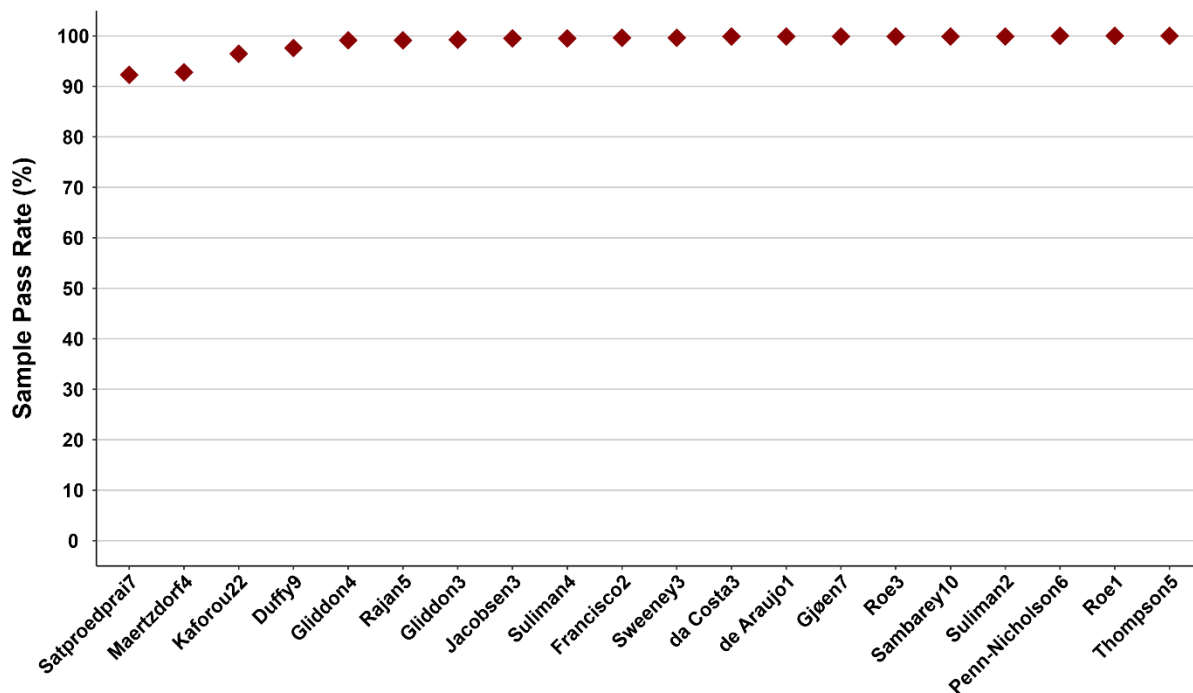


Figure 5.2: Signature pass rates in the SHINE sub-study.

Sample pass rates represent the percentage of whole blood samples from the entire study for which signature scores were calculated. Transcriptomic signatures with sample pass-rate of 90% and above were included for analysis in this sub-study. All 20 transcriptomic signatures evaluated are shown on the x-axis in order of sample pass-rate. On the y-axis, the pass-rate of each signature in samples from the SHINE clinical trial is expressed as a percentage.

Table 5.4: Sample pass-rate of individual transcripts in the SHINE sub-study

Transcript	Pass rate (%)
ACTA2.Hs00426835_g1	99.51
ACTR3.Hs01029159_g1	99.63
ANKRD22.Hs00944015_m1	99.51
APOL1.Hs01066280_m1	99.51
ARG1.Hs00968979_m1	99.63
BATF2.Hs00912737_m1	99.63
BCL6.BCL6_CUSTOM_AREPURD	99.63
BLK.Hs01017452_m1	99.63
C1QA.Hs00706358_s1	99.63
C1QB.Hs00608019_m1	99.01
C1QC.Hs00757779_m1	78.91
C4ORF18.FAM198B.Hs00259260_s1	99.51
C5.Hs00156197_m1	97.66
CCR6.Hs01890706_s1	99.51
CD160.Hs00199894_m1	99.51
CD1C.Hs00957534_g1	99.63
CD36.2.Hs01567186_m1	99.51
CD3E.CD3E_APDJ3NT	99.63
CD40L.Hs00163934_m1	99.63
CD74_j1.Hs04983808_s1	99.63
CD74_j2.CD74_J2_AP9HN47	99.63
CD79A.Hs00998120_g1	99.63
CD79B.Hs01058826_g1	99.63
CDKN1C.Hs00175938_m1	99.63
CXCR5.Hs00540548_s1	99.63
CYP4F3.Hs01587860_mH	99.51
DUSP3.Hs01115776_m1	99.63
F2RL1.Hs00608346_m1	95.81
FAM20A.Hs01034066_m1	98.15
FCGR1A.FCGR1A_SAT7_AP2XDMK	99.63
FCGR1A.Hs00174081_m1	99.63
FCGR1B.Hs02341825_m1	99.63
FCGR1B_VARIANT_1.FCGR1B_VARIANT1_AP3267H	99.63
FCGR1B_VARIANT_2.FCGR1B_VARIANT2_AP47ZTF	99.38
FCGR1C.Hs00417598_m1	99.63
FLVCR2.Hs00900390_m1	99.38
GAS6.GAS6_NM_000820_10_11_AR323W3	99.26
GAS6.GAS6_NM_000820_5_6_AR47XGZ	99.14
GBP1.Hs00977005_m1	99.63
GBP2.Hs00894846_g1	99.63
GBP5.Hs00369472_m1	99.51
GBP6.Hs01584201_m1	98.77
GNG7.Hs00192999_m1	99.63
GYG1.Hs00907542_g1	99.51
GZMA.Hs00989184_m1	99.63
HK3.Hs01092850_m1	99.63
ID3.Hs00954037_g1	99.63

IFI44L.Hs00915292_m1	99.63
IFITM3.Hs03057129_s1	99.63
KAZN.KAZN_TOTAL_SAT7_AP7DVDD	91.12
KIF1B.Hs01114512_g1	99.63
KLF2.Hs00360439_g1	99.38
KLF2.KLF2_CUSTOM_ARH6CF4	88.04
KLHDC8B.Hs00293902_m1	99.63
LAG3.Hs00958444_g1	99.63
LHFPL2.Hs00299613_m1	99.63
LTF.Hs00158924_m1	99.14
MAFB.MAFB_SAT7_APZTH2N	99.51
MAP7D3.Hs00226257_m1	99.63
MMP9.MMP9_APEPW9P	99.63
MPO.Hs00165162_m1	99.14
MTRF1L.Hs01097882_g1	99.63
NOD2.Hs01550759_g1	99.63
NPC2.Hs01119244_m1	99.51
OSBPL10.Hs00215016_m1	99.63
P2RY14.Hs01848195_s1	92.48
PRDM1.Hs00153357_m1	99.63
RAB13.Hs04400188_g1	99.63
Rab33A.Hs00191243_m1	99.63
RABL2A.Hs00255244_m1	99.63
RBBP8.Hs01090329_m1	99.63
RP11.295G20.2.Hs01373568_m1	99.63
S100A8.Hs00374264_g1	99.38
SCARF1.Hs01092483_m1	99.63
SDR39U1.Hs01016970_g1	99.63
SEPT4.Hs00910208_g1	99.63
SERPING1.Hs00934329_m1	99.63
SLPI.Hs00268206_m1	99.63
SMARCD3.Hs01088251_g1	99.63
STAT1.Hs01013996_m1	99.63
STT3A.Hs00967491_m1	99.63
TIMM10.TIMM10_CUSTOM_APT2DEC	99.63
TMBIM6.Hs00162661_m1	99.63
TMCC1.Hs01037666_s1	99.63
TNIP1.TNIP1_GEO_AP47ZYZ	99.63
TRMT2A.Hs01000041_g1	99.63
TUBGCP6.Hs00363509_g1	99.63
UCP2.Hs01075224_g1	99.63
USF2.Hs01100994_g1	99.63
VAMP5.Hs01105383_g1	99.63
WARS_j1.Hs00998737_m1	99.63
WARS_j2.WARS.J2_APAAFHX	99.63
ZDHC19.Hs00376118_m1	98.52
ZNF296.Hs00377132_m1	99.38

Signature scores at baseline distinguished between confirmed and unlikely TB cases

Before assessing signature scores during treatment, we first sought to determine whether transcriptomic signatures could distinguish between confirmed, unconfirmed or unlikely TB at baseline. Confirmed TB cases had higher scores at baseline compared to unconfirmed and unlikely TB cases. All 20 transcriptomic signatures differentiated between confirmed TB and unlikely TB cases at baseline, with AUROC between 0.59 (95% CI 0.45-0.73) and 0.81 (95 % CI 0.7-0.92) (**Figure 5.3** and **Table 5.5**). All transcriptomic signatures had reduced performance in differentiating between unconfirmed and unlikely TB cases, in comparison to confirmed versus unlikely TB. Gliddon4 was the only signature with a significant difference in scores between unconfirmed and unlikely TB cases ($p=0.0047$). Gliddon4 had the highest AUC (0.81, 95% CI 0.7-0.92) for distinguishing confirmed and unconfirmed TB cases combined (represented as conf_unconf) from unlikely TB cases (**Figure 5.3**).

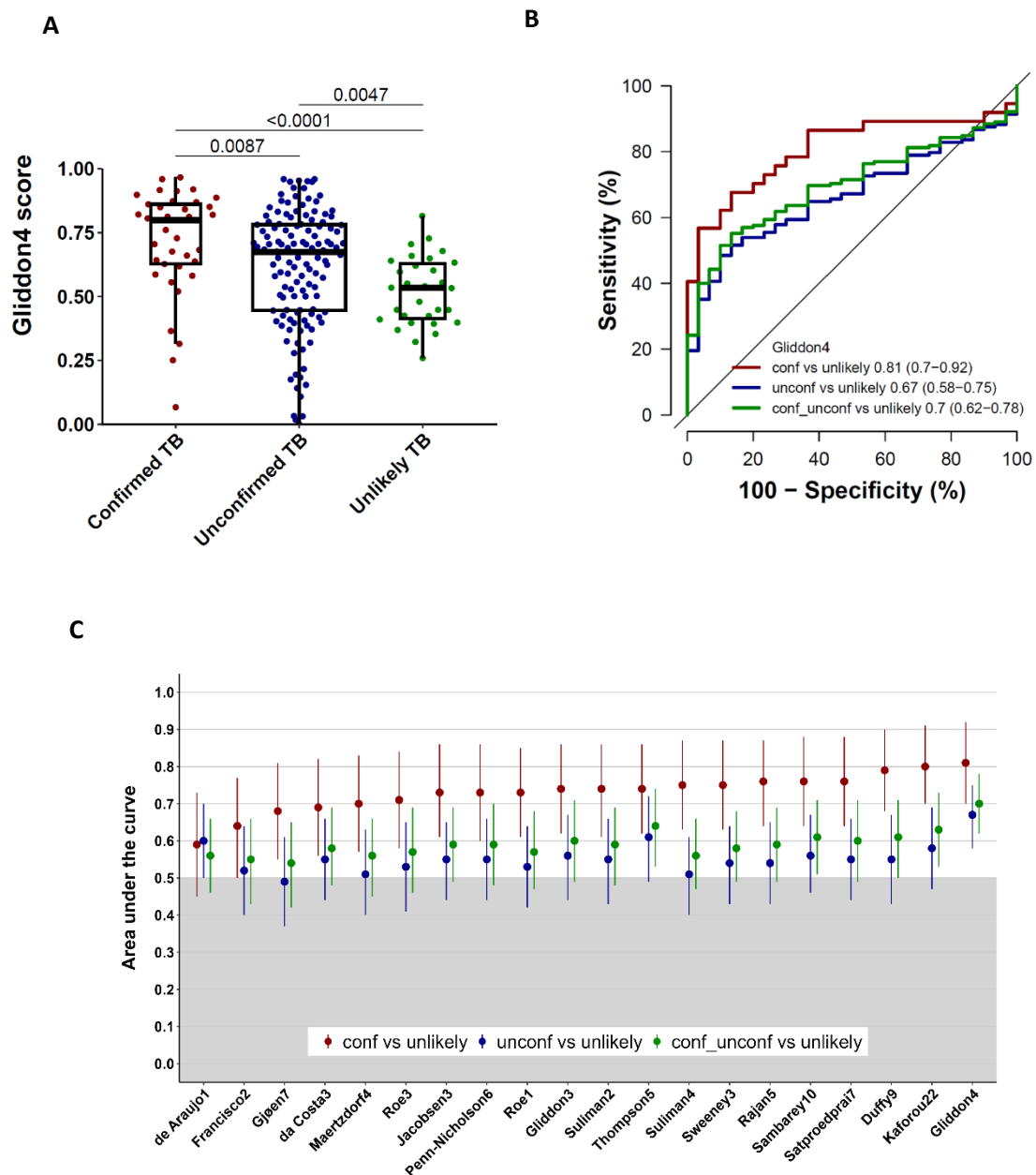


Figure 5.3: Signature score distributions and performance of transcriptomic signatures as treatment monitoring biomarkers at baseline in confirmed, unconfirmed, and unlikely TB cases.

conf = confirmed TB, unconf = unconfirmed TB, unlikely = unlikely TB. conf_unconf = confirmed TB + unconfirmed TB.

(A) Box-and-whisker plot of the signature with the highest AUC for distinguishing confirmed from unlikely TB cases at baseline (Gliddon4). The plot shows signature score distribution (each dot represents a participant) in confirmed TB cases (n = 37), unconfirmed TB cases (n = 128) and unlikely TB cases (n = 30) at baseline. P-values for comparison of median signature scores between groups were calculated using the Mann–Whitney U test. Boxes depict the IQR, the midline represent the median, and whiskers represent $1.5 \times$ IQR.

(B) ROC curve depicting baseline performance (area under the curve, AUC, with 95% CI) of the Gliddon4 signature in differentiating confirmed TB (conf) vs unlikely TB, unconfirmed TB (unconf) vs unlikely TB, and confirmed and unconfirmed TB cases combined (conf_unconf) vs unlikely TB.

(C) Summary of signature performance, ranked by AUC for confirmed TB vs unlikely TB (red). Error bars depict the 95% CIs. The AUCs for confirmed vs unlikely TB cases (red), unconfirmed TB cases vs unlikely TB (blue) and confirmed and unconfirmed TB cases combined vs unlikely TB cases (green) are also shown. AUCs for all signatures are shown in the Table 5.5.

Table 5.5: Performance of candidate transcriptomic signatures in differentiating between the different TB classification groups at baseline.

Biomarker	conf vs unlikely AUC (95% CI)	unconf vs unlikely AUC (95% CI)	conf_unconf vs unlikely AUC (95% CI)
da Costa3	0.69 (0.56-0.82)	0.55 (0.44-0.66)	0.58 (0.48-0.69)
de Araujo1	0.59 (0.45-0.73)	0.6 (0.5-0.7)	0.56 (0.46-0.66)
Duffy9	0.79 (0.68-0.9)	0.55 (0.43-0.67)	0.61 (0.5-0.71)
Francisco2	0.64 (0.5-0.77)	0.52 (0.4-0.64)	0.55 (0.43-0.66)
GjØen7	0.68 (0.55-0.81)	0.49 (0.37-0.61)	0.54 (0.42-0.65)
Gliddon3	0.74 (0.62-0.86)	0.56 (0.44-0.67)	0.6 (0.49-0.71)
Gliddon4	0.81 (0.7-0.92)	0.67 (0.58-0.75)	0.7 (0.62-0.78)
Jacobsen3	0.73 (0.61-0.86)	0.55 (0.44-0.65)	0.59 (0.49-0.69)
Kaforou22	0.8 (0.7-0.91)	0.58 (0.47-0.69)	0.63 (0.53-0.73)
Maertzdorf4	0.7 (0.57-0.83)	0.51 (0.4-0.63)	0.56 (0.45-0.66)
Penn-Nicholson6	0.73 (0.6-0.86)	0.55 (0.44-0.66)	0.59 (0.48-0.7)
Rajan5	0.76 (0.64-0.87)	0.54 (0.43-0.65)	0.59 (0.49-0.69)
Roe1	0.73 (0.61-0.85)	0.53 (0.42-0.64)	0.57 (0.47-0.68)
Roe3	0.71 (0.58-0.84)	0.53 (0.41-0.65)	0.57 (0.46-0.69)
Sambarey10	0.76 (0.64-0.88)	0.56 (0.46-0.67)	0.61 (0.51-0.71)
Satproedprai7	0.76 (0.64-0.88)	0.55 (0.44-0.66)	0.6 (0.49-0.71)
Suliman2	0.74 (0.61-0.86)	0.55 (0.43-0.66)	0.59 (0.48-0.69)
Suliman4	0.75 (0.63-0.87)	0.51 (0.4-0.61)	0.56 (0.47-0.66)
Sweeney3	0.75 (0.63-0.87)	0.54 (0.43-0.64)	0.58 (0.49-0.68)
Thompson5	0.74 (0.62-0.86)	0.61 (0.49-0.72)	0.64 (0.53-0.74)

conf = confirmed TB. unconf = unconfirmed TB. unlikely = unlikely TB. conf_unconf = confirmed TB + unconfirmed TB. AUC = area under the curve. CI = confidence interval.

Signature scores increase with CXR severity

Median scores for all signatures increased with enhanced CXR severity (**Figure 5.4**). CXR presentation and classifications of TB severity were determined by the radiographers and clinicians that were part of the SHINE clinical trial end-point review committee. Children with a typical TB, severe CXR presentation had significantly higher signature scores than children with normal CXRs ($p = 0.015$ for both Gliddon4 and Kaforou22).

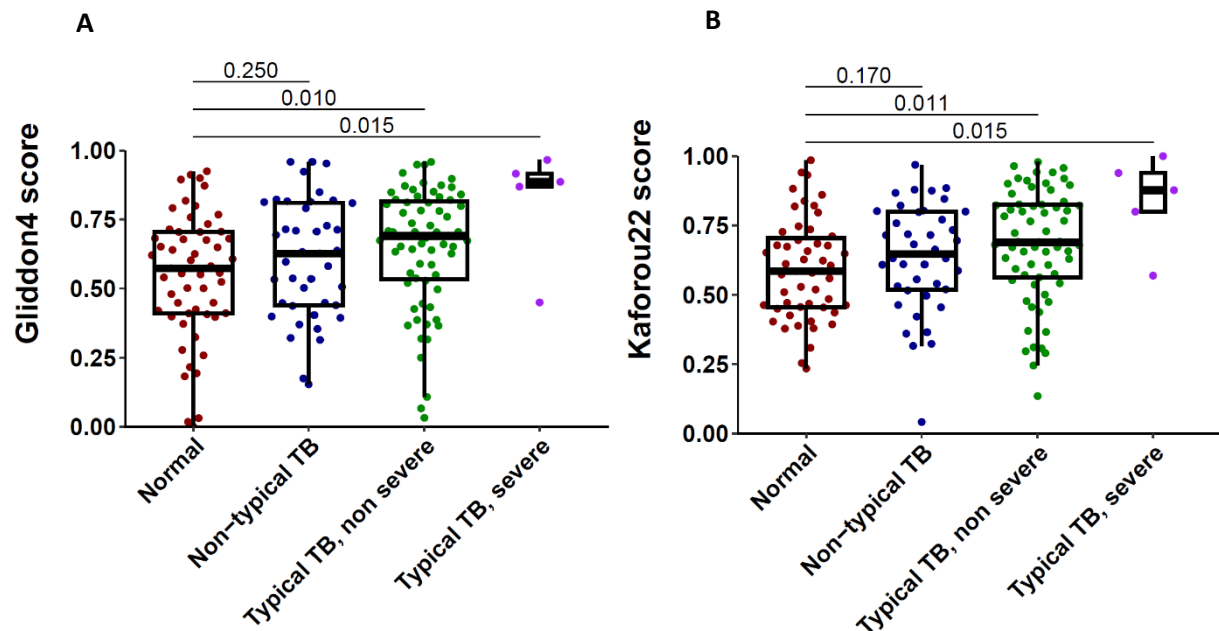


Figure 5.4: Signature score distributions at baseline stratified by chest x-ray severity.

Box-and-whisker plots for (A) Gliddon4 and (B) Kaforou22 showing signature score distribution (each dot represents a participant) at baseline for disease severity based on chest x-ray results. P-values for comparison of signature score distribution between chest x-ray classification groups were calculated using the Mann-Whitney U test. Boxes depict the IQR, the midline represents the median, and whiskers represent $1.5 \times \text{IQR}$.

Participant characteristics associated with signature scores at baseline

Multivariate linear regression was used to identify clinical and demographic features associated with signature scores. This was carried out using signature scores at baseline for all SHINE study participants. Results for the four signatures shown (**Table 5.6**) are representative of the overall trend seen for all signatures evaluated. TB classification was associated with signature scores, at baseline, for 14 out of 20 signatures. Unconfirmed and unlikely TB were associated with lower signature scores as compared to confirmed TB.

Table 5.6: Regression analyses determining clinical and demographic variables associated with signature scores at baseline

Variable	Duffy9 β-coefficient (95% CI); adjusted p-value	Gliddon4 β-coefficient (95% CI); adjusted p- value	Jacobsen3 β-coefficient (95% CI); adjusted p- value	Kaforou22 β-coefficient (95% CI); adjusted p- value
Age				
Age> 3 years	Reference	Reference	Reference	Reference
< 3 years	0.04 (-0.05, 0.13); 0.62	0.03 (-0.04, 0.10); 0.65	0.00 (-0.09, 0.09); 1.00	0.01 (-0.06, 0.07); 0.94
Sex				
Female	Reference	Reference	Reference	Reference
Male	0.02 (-0.06, 0.11); 0.77	0.02 (-0.05, 0.08); 0.77	0.01 (-0.08, 0.10); 0.94	0.02 (-0.03, 0.08); 0.65
Clinical presentation				
Respiratory TB	Reference	Reference	Reference	Reference
Mixed TB	0.00 (-0.09, 0.08); 0.98	-0.06 (-0.13, 0.01); 0.29	-0.01 (-0.10, 0.09); 0.97	-0.01 (-0.07, 0.06); 0.94
TB classification				
Confirmed TB	Reference	Reference	Reference	Reference
Unconfirmed TB	-0.23 (-0.34, -0.11); 0.01	-0.10 (-0.20, -0.01); 0.16	-0.18 (-0.30, -0.06); 0.03	-0.15 (-0.24, -0.07); 0.01
Unlikely TB	-0.23 (-0.39, -0.07); 0.05	-0.18 (-0.31, -0.05); 0.05	-0.18 (-0.35, -0.01); 0.18	-0.16 (-0.28, -0.05); 0.05
CXR severity				
Normal	Reference	Reference	Reference	Reference
Non-typical TB	0.00 (-0.11, 0.11); 1.00	0.04 (-0.05, 0.13); 0.65	0.04 (-0.08, 0.16); 0.70	0.02 (-0.06, 0.10); 0.75
Typical TB, non-severe	0.08 (-0.03, 0.19); 0.35	0.05 (-0.04, 0.14); 0.47	0.09 (-0.02, 0.21); 0.29	0.05 (-0.03, 0.12); 0.47
Typical TB, severe	0.17 (-0.09, 0.44); 0.42	0.18 (-0.04, 0.39); 0.29	0.03 (-0.25, 0.30); 0.94	0.13 (-0.06, 0.31); 0.41

CI: confidence interval. CXR: chest x-ray

Weight-for-age z scores were generated to standardize weight data for comparability across different age groups, and control for age effects in the regression analysis. With inclusion of weight-for-age z scores in the regression model, no clinical and demographic features were significantly associated with signature scores (**Table 5.7**). These results are shown separately to account for possible collinearity between z-scores and TB classification (**Table 5.7**).

Table 5.7: Regression analyses determining clinical and demographic variables associated with signature scores at baseline, including weight for age z scores at baseline

Variable	Duffy9 β -coefficient (95% CI); adjusted p-value	Gliddon4 β -coefficient (95% CI); adjusted p-value	Jacobsen3 β -coefficient (95% CI); adjusted p-value	Kaforou22 β -coefficient (95% CI); adjusted p-value
Age				
Age > 3 years	Reference	Reference	Reference	Reference
< 3 years	0.05 (-0.04, 0.14); 0.64	0.03 (-0.04, 0.10); 0.68	0.01 (-0.09, 0.10); 0.96	0.01 (-0.05, 0.08); 0.88
Weight-for-age z-score	0.01 (-0.03, 0.05); 0.84	0.00 (-0.04, 0.03); 0.94	0.00 (-0.04, 0.05); 0.95	-0.01 (-0.04, 0.03); 0.86
Sex				
Female	Reference	Reference	Reference	Reference
Male	0.02 (-0.07, 0.11); 0.84	0.00 (-0.07, 0.07); 0.96	0.01 (-0.09, 0.11); 0.94	0.02 (-0.05, 0.08); 0.82
Clinical presentation				
Respiratory TB	Reference	Reference	Reference	Reference
Mixed TB	0.00 (-0.09, 0.09); 1.00	-0.06 (-0.14, 0.01); 0.42	-0.01 (-0.11, 0.09); 0.94	-0.01 (-0.07, 0.06); 0.94
TB classification				
Confirmed TB	Reference	Reference	Reference	Reference
Unconfirmed TB	-0.19 (-0.31, -0.06); 0.14	-0.08 (-0.18, 0.02); 0.44	-0.16 (-0.29, -0.02); 0.22	-0.14 (-0.23, -0.04); 0.14
Unlikely TB	-0.19 (-0.37, -0.02); 0.24	-0.17 (-0.31, -0.03); 0.22	-0.15 (-0.34, 0.03); 0.43	-0.14 (-0.27, -0.02); 0.23
CXR severity				
Normal	Reference	Reference	Reference	Reference
Non-typical TB	0.01 (-0.11, 0.14); 0.94	0.04 (-0.05, 0.14); 0.67	0.03 (-0.10, 0.16); 0.84	0.03 (-0.06, 0.12); 0.77
Typical TB, non-severe	0.08 (-0.03, 0.20); 0.46	0.05 (-0.04, 0.14); 0.64	0.08 (-0.05, 0.20); 0.54	0.04 (-0.04, 0.13); 0.65
Typical TB, severe	0.17 (-0.12, 0.47); 0.58	0.19 (-0.04, 0.43); 0.43	-0.01 (-0.33, 0.31); 0.98	0.13 (-0.08, 0.34); 0.54

CI = confidence interval. CXR = chest x-ray. Sixteen children were not part of this analysis since wfaz could not be calculated for them. The zscore package in R used to generate z scores considered children above ten years to be outliers.

Changes in scores during treatment are greater in confirmed and unconfirmed TB groups

Confirmed TB cases had the highest median scores at baseline ($p = <0.001$ for both Gliddon4 and Kaforou22) (**Figure 5.5**). Results for Gliddon4 and Kaforou22 represent the general trend in signature scores, during treatment, for the 20 signatures evaluated. All signatures showed differences in scores at baseline across the three TB categories. By week 8, there were no differences in scores across TB categories for 19 out of the 20 signatures (week 8: $p = 0.087$ for Gliddon4 and $p = 0.33$ for Kaforou22; end of treatment: $p = 0.051$ for Gliddon4 and $p = 0.16$ for Kaforou22).

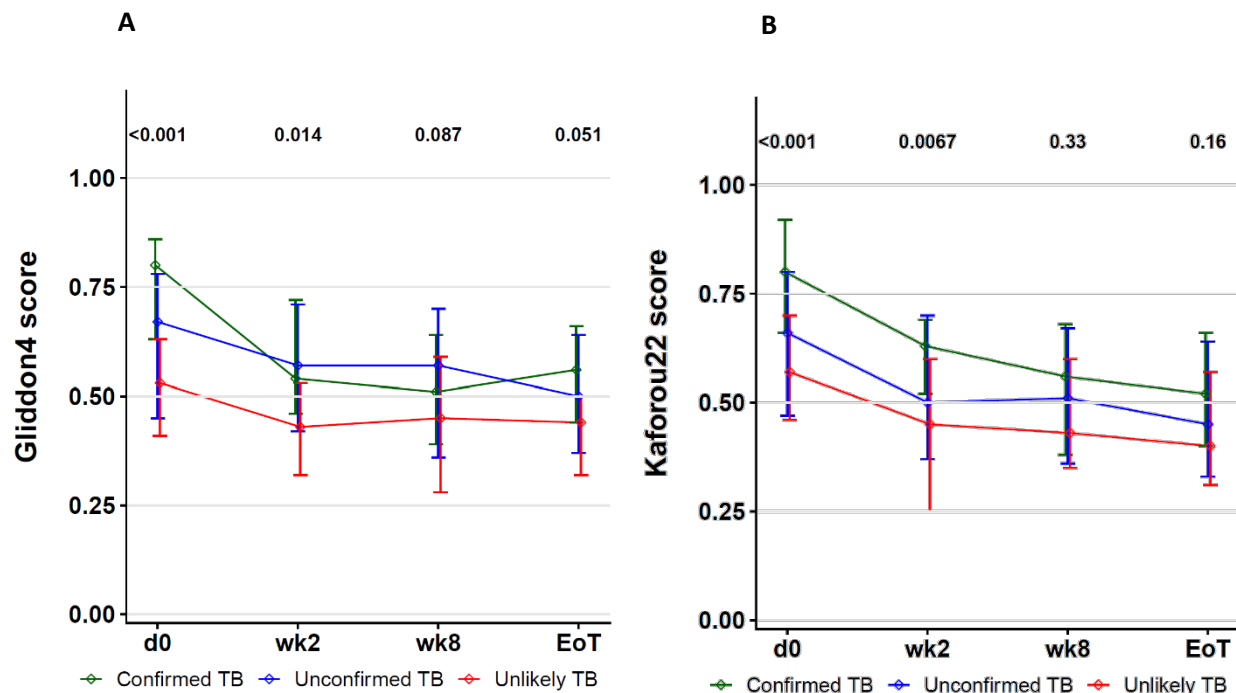


Figure 5.5: Change in signature scores during treatment in children with confirmed, unconfirmed, and unlikely TB.

d0 = day zero, wk2 = week 2, wk8 = week 8, EoT = end of treatment.

Line graphs show differences in signature scores for (A) Gliddon4 and (B) Kaforou22 in confirmed, unconfirmed, and unlikely TB cases during treatment. P-values for comparison of median signature scores among children in the three TB categories at each time point were calculated using the Kruskal Wallis test. Each line depicts the IQR, with the midpoint representing the median. EoT is a combination of results from week 16 for children who in the 4-month arm and week 24 for children in the 6-month arm.

Paired analysis for children under each TB category showed that scores decreased significantly in the first two weeks of treatment, with not much change thereafter (**Figure 5.6**). This was consistent for both the confirmed and unconfirmed TB cases. Gliddon4 and Kaforou22 are shown as a representation of all 20 transcriptomic signatures.

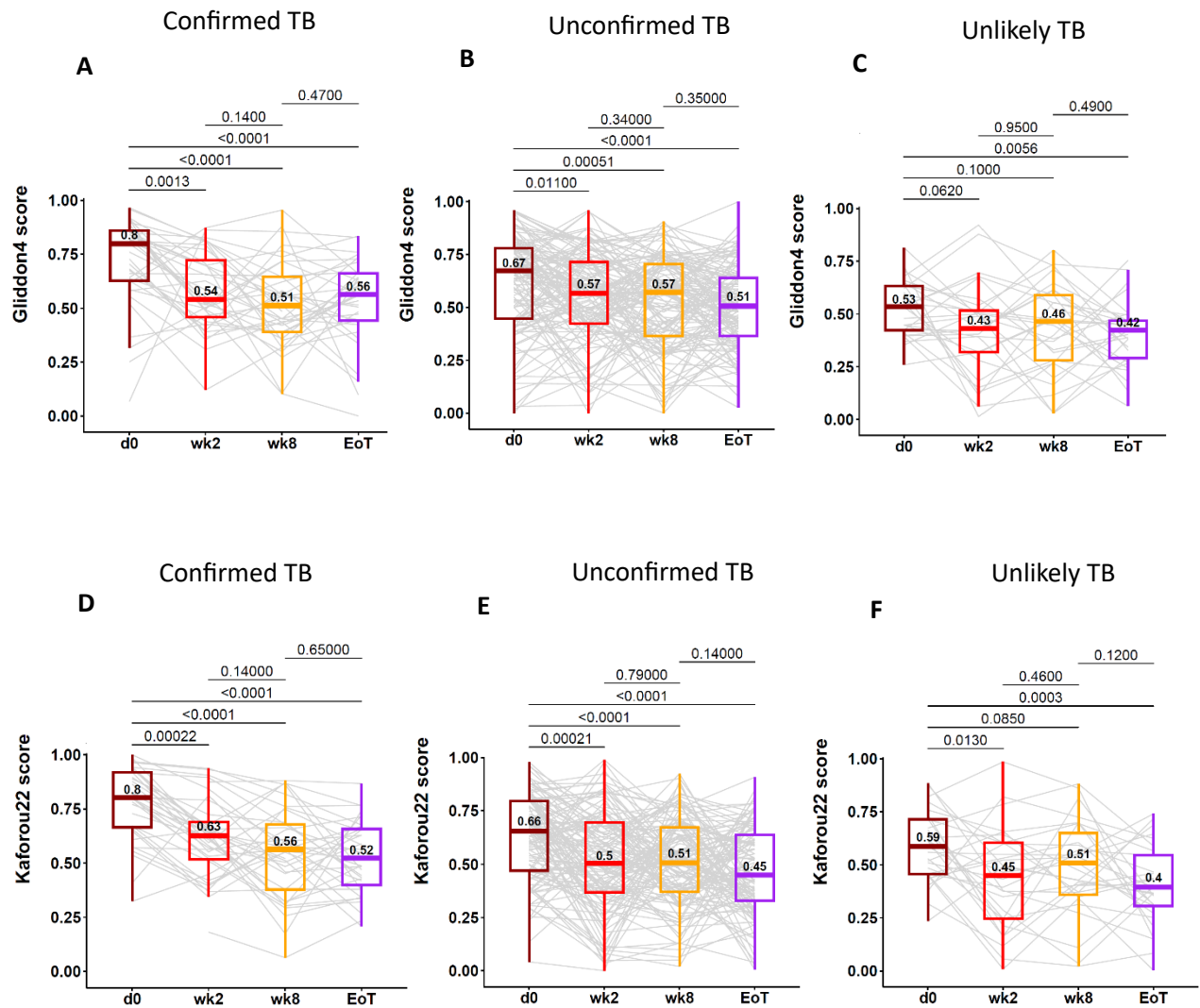


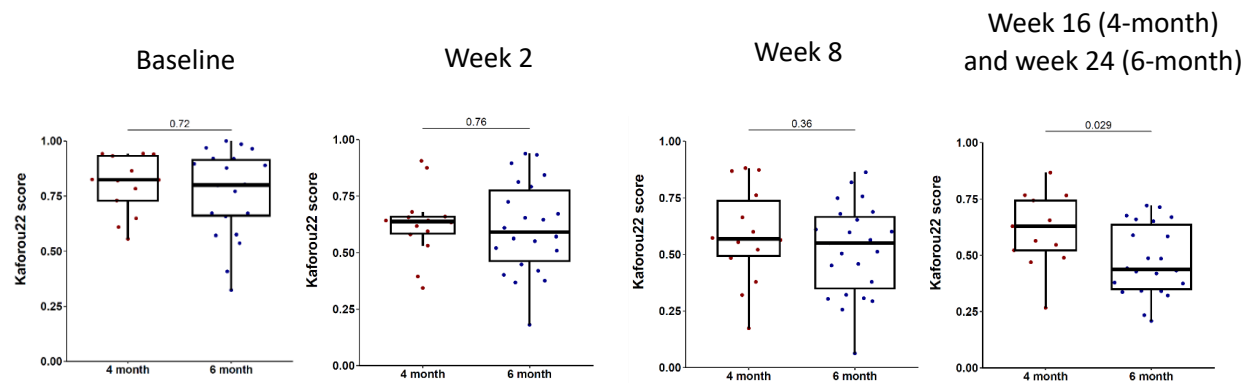
Figure 5.6: Signature score distributions during treatment stratified by TB classification.

Box-and-whisker plots for (A-C) Gliddon, (D-F), and Kaforou22. Box-and-whisker plots show paired signature score distribution (each grey line represents a participant) at each time point in children with confirmed TB (A, D), unconfirmed TB (B, E) and unlikely TB (C, F). P-values for comparison of median signature scores between time points were calculated using the Wilcoxon signed-rank test. Boxes depict the IQR, the midline represents the median, and whiskers represent $1.5 \times$ IQR.

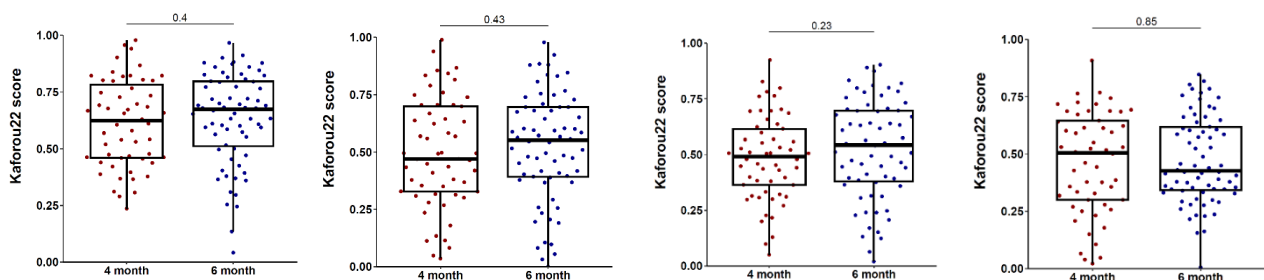
Children with confirmed TB showed differences in scores at the end of treatment

We aimed to determine whether there were any differences in signature scores for children under both treatment arms. Due to differences in signature scores across TB categories (**Figures 5.5 and 5.6**), analysis of scores for children in the 4- and 6-month arms was stratified based on TB classification. For confirmed TB cases, 13 out of 20 signatures had differences in scores at the end of treatment. Kaforou22 (**Figure 5.7**) is shown to represent these signatures. In the unconfirmed and unlikely TB groups, there were no signatures with differences in scores between treatment arms, across all timepoints.

Confirmed TB



Unconfirmed TB



Unlikely TB

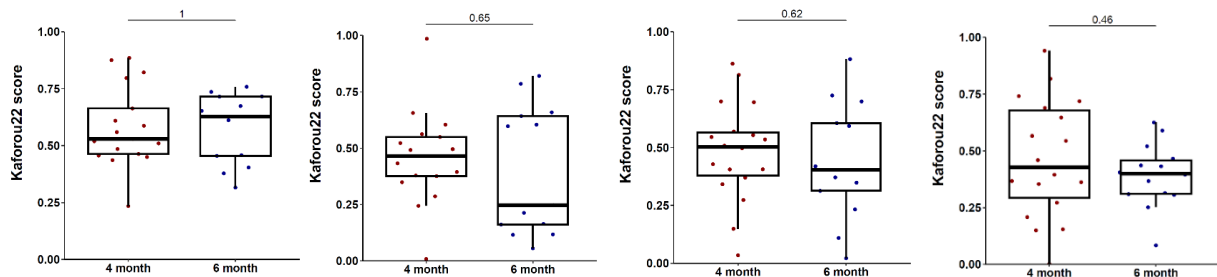


Figure 5.7: Signature score distributions at the end of treatment stratified by treatment arm.

Box-and-whisker plots for Kaforou22 show signature score distribution (each dot represents a participant) at baseline, week 2, week 8, and the end of treatment in children with confirmed, unconfirmed, and unlikely TB. P-values for comparison of median signature scores between treatment arms were calculated using the Mann–Whitney U test. Boxes depict the IQR, the midline represents the median, and whiskers represent $1.5 \times \text{IQR}$.

Transcriptomic signatures distinguish baseline from all other treatment time points

We compared the performance of transcriptomic signatures in differentiating TB cases, across treatment time points, to determine their potential as possible TB treatment monitoring biomarkers. Kaforou22 had the highest AUC for discriminating between baseline and EoT for confirmed TB (AUC 0.85, 95% CI 0.76-0.94) (**Figure 5.8** and **Table 5.8**) and unconfirmed TB (AUC 0.71, 95% 0.65-0.77) (**Figure 5.9** and **Table 5.9**). Jacobsen3 best discriminated between baseline and EoT for unlikely TB (AUC 0.73, 95% 0.59-0.86) (**Figure 5.10** and **Table 5.10**). Across the three TB categories, all signatures had the lowest AUCs for differentiating TB cases at baseline from TB cases after two weeks of treatment. This was in comparison to signatures differentiating between children at baseline and at the end of treatment as well at week eight.

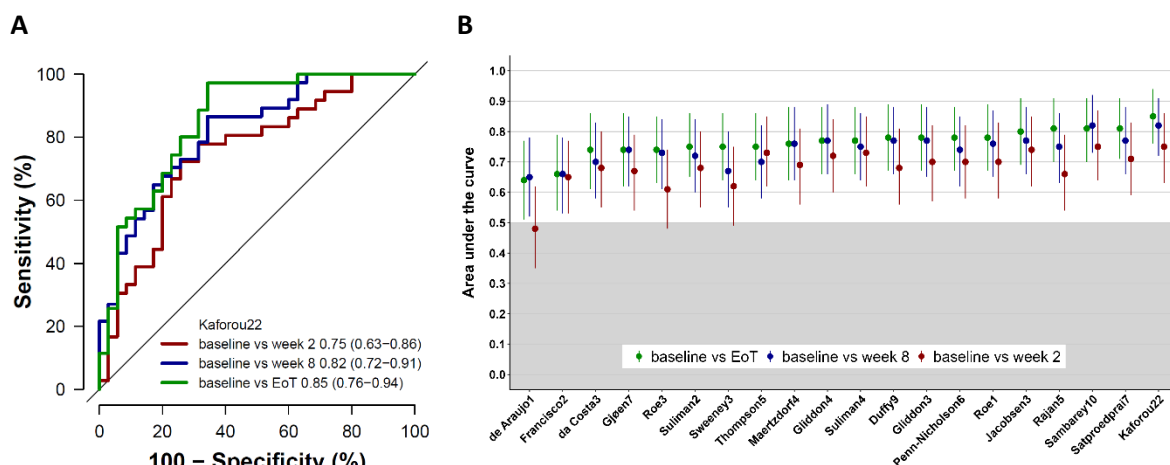


Figure 5.8: Performance of transcriptomic signatures as treatment monitoring biomarkers in children with confirmed TB.

(A) ROC curve depicting performance (area under the curve, AUC, with 95% CI) of the Kaforou22 signature in confirmed TB cases comparing baseline signature scores across all time points during treatment.

(B) Summary of signature performance, ranked by AUC for baseline vs end of treatment (EoT) (green) scores. Error bars depict the 95% CIs. A table of AUCs for all signatures is shown in **Table 5.8**.

Table 5.8: Performance of candidate transcriptomic signatures in confirmed TB cases for discriminating between baseline and week 2, week 8, and end of treatment time points.

Biomarker	Baseline vs week 2 AUC (95% CI)	Baseline vs week 8 AUC (95% CI)	Baseline vs EoT AUC (95% CI)
da Costa3	0.68 (0.55-0.8)	0.70 (0.58-0.83)	0.74 (0.61-0.86)
de Araujo1	0.48 (0.35-0.62)	0.65 (0.52-0.78)	0.64 (0.51-0.77)
Duffy9	0.68 (0.56-0.81)	0.77 (0.66-0.88)	0.78 (0.67-0.89)
Francisco2	0.65 (0.53-0.77)	0.66 (0.53-0.78)	0.66 (0.54-0.79)
Gjøen7	0.67 (0.54-0.79)	0.74 (0.62-0.85)	0.74 (0.62-0.86)
Gliddon3	0.70 (0.57-0.82)	0.77 (0.65-0.88)	0.78 (0.67-0.89)
Gliddon4	0.72 (0.6-0.84)	0.77 (0.66-0.89)	0.77 (0.66-0.88)
Jacobsen3	0.74 (0.62-0.85)	0.77 (0.66-0.88)	0.8 (0.69-0.91)
Kaforou22	0.75 (0.63-0.86)	0.82 (0.72-0.91)	0.85 (0.76-0.94)
Maertzdorf4	0.69 (0.56-0.81)	0.76 (0.64-0.88)	0.76 (0.64-0.88)
Penn-Nicholson6	0.70 (0.58-0.82)	0.74 (0.62-0.85)	0.78 (0.67-0.88)
Rajan5	0.66 (0.54-0.79)	0.75 (0.63-0.86)	0.81 (0.7-0.91)
Roe1	0.70 (0.58-0.83)	0.76 (0.65-0.87)	0.78 (0.67-0.89)
Roe3	0.61 (0.48-0.74)	0.73 (0.61-0.84)	0.74 (0.63-0.85)
Sambarey10	0.75 (0.64-0.87)	0.82 (0.73-0.92)	0.81 (0.7-0.91)
Satproedprai7	0.71 (0.59-0.83)	0.77 (0.66-0.88)	0.81 (0.71-0.91)
Suliman2	0.68 (0.55-0.8)	0.72 (0.6-0.84)	0.75 (0.65-0.86)
Suliman4	0.73 (0.62-0.85)	0.75 (0.64-0.86)	0.77 (0.66-0.88)
Sweeney3	0.62 (0.49-0.75)	0.67 (0.55-0.8)	0.75 (0.64-0.86)
Thompson5	0.73 (0.62-0.85)	0.70 (0.58-0.82)	0.75 (0.64-0.86)

EoT = end of treatment. AUC = area under the curve. CI = confidence interval.

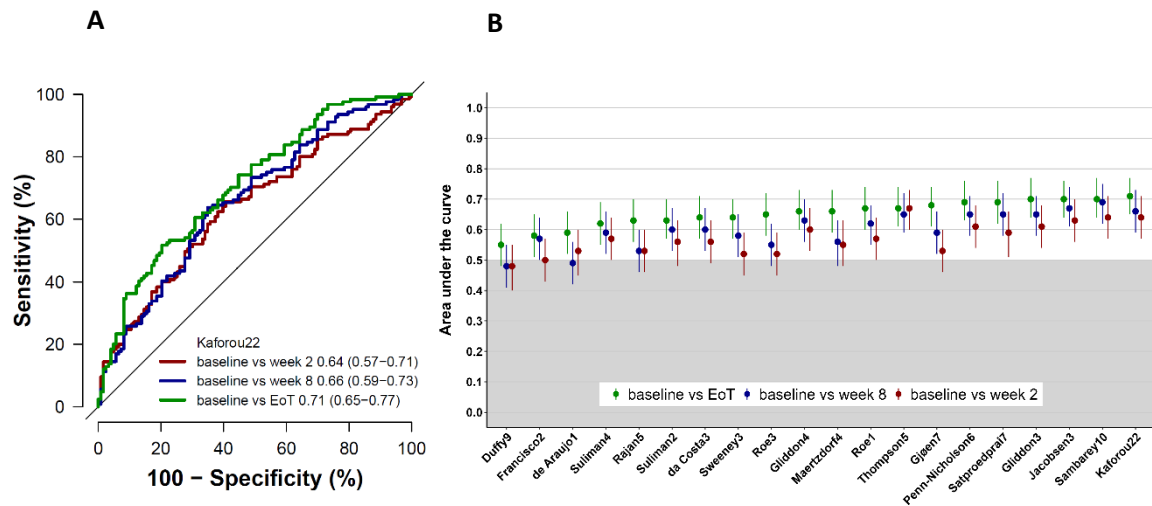


Figure 5.9: Performance of transcriptomic signatures as treatment monitoring biomarkers in children with unconfirmed TB.

(A) ROC curve depicting performance (area under the curve, AUC, with 95% CI) of the Kaforou22 signature in unconfirmed TB cases comparing baseline signature scores across all time points during treatment.

(B) Summary of signature performance, ranked by AUC for baseline vs end of treatment (EoT) (green) scores. Error bars depict the 95% CIs. A table of AUCs for all signatures is shown in **Table 5.9**.

Table 5.4: Performance of candidate transcriptomic signatures in unconfirmed TB cases for discriminating between baseline and week 2, week 8, and end of treatment all time points.

Biomarker	Baseline vs week 2 AUC (95% CI)	Baseline vs week 8 AUC (95% CI)	Baseline vs EoT AUC (95% CI)
da Costa3	0.56 (0.49-0.63)	0.60 (0.53-0.67)	0.64 (0.57-0.71)
de Araujo1	0.53 (0.45-0.6)	0.49 (0.42-0.56)	0.59 (0.52-0.66)
Duffy9	0.48 (0.4-0.55)	0.48 (0.41-0.55)	0.55 (0.48-0.62)
Francisco2	0.50 (0.43-0.57)	0.57 (0.5-0.64)	0.58 (0.51-0.65)
GjØen7	0.53 (0.46-0.6)	0.59 (0.52-0.66)	0.68 (0.61-0.74)
Gliddon3	0.61 (0.54-0.68)	0.65 (0.58-0.71)	0.70 (0.64-0.77)
Gliddon4	0.60 (0.53-0.67)	0.63 (0.56-0.7)	0.66 (0.6-0.73)
Jacobsen3	0.63 (0.56-0.7)	0.67 (0.61-0.74)	0.70 (0.64-0.76)
Kaforou22	0.64 (0.57-0.71)	0.66 (0.59-0.73)	0.71 (0.65-0.77)
Maertzdorf4	0.55 (0.48-0.63)	0.56 (0.48-0.63)	0.66 (0.59-0.73)
Penn-Nicholson6	0.61 (0.54-0.68)	0.65 (0.58-0.71)	0.69 (0.63-0.76)
Rajan5	0.53 (0.46-0.6)	0.53 (0.46-0.6)	0.63 (0.56-0.7)
Roe1	0.57 (0.5-0.64)	0.62 (0.55-0.68)	0.67 (0.6-0.74)
Roe3	0.52 (0.45-0.59)	0.55 (0.48-0.62)	0.65 (0.58-0.72)
Sambarey10	0.64 (0.57-0.71)	0.69 (0.62-0.75)	0.70 (0.64-0.77)
Satproedprai7	0.59 (0.51-0.66)	0.65 (0.58-0.72)	0.69 (0.62-0.76)
Suliman2	0.56 (0.48-0.63)	0.60 (0.53-0.67)	0.63 (0.57-0.7)
Suliman4	0.57 (0.5-0.64)	0.59 (0.52-0.66)	0.62 (0.55-0.69)
Sweeney3	0.52 (0.45-0.59)	0.58 (0.51-0.65)	0.64 (0.57-0.7)
Thompson5	0.67 (0.6-0.73)	0.65 (0.59-0.72)	0.67 (0.61-0.74)

EoT = end of treatment. AUC = area under the curve. CI = confidence interval.

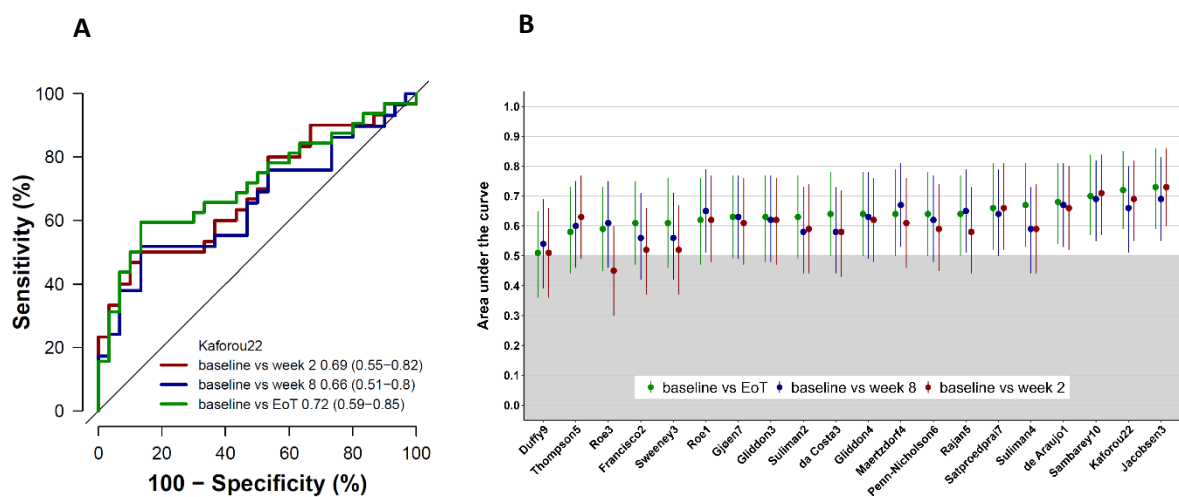


Figure 5.10: Performance of transcriptomic signatures as treatment monitoring biomarkers in children with unlikely TB.

(A) ROC curve depicting performance (area under the curve, AUC, with 95% CI) of the Kaforou22 signature in unlikely TB cases comparing baseline signature scores across all time points during treatment.

(B) Summary of signature performance, ranked by AUC for baseline vs end of treatment (EoT) (green) scores. Error bars depict the 95% CIs. A table of AUCs for all signatures is shown in the **Table 5.10**.

Table 5.10: Performance of candidate transcriptomic signatures in unlikely TB cases for discriminating between baseline and week 2, week 8, and end of treatment all time points.

Biomarker	Baseline vs week 2 AUC (95% CI)	Baseline vs week 8 AUC (95% CI)	Baseline vs EoT AUC (95% CI)
da Costa3	0.58 (0.43-0.72)	0.58 (0.44-0.73)	0.64 (0.5-0.78)
de Araujo1	0.66 (0.52-0.8)	0.67 (0.53-0.81)	0.68 (0.54-0.81)
Duffy9	0.51 (0.36-0.66)	0.54 (0.39-0.69)	0.51 (0.36-0.65)
Francisco2	0.52 (0.37-0.66)	0.56 (0.42-0.71)	0.61 (0.47-0.75)
GjØen7	0.61 (0.47-0.76)	0.63 (0.49-0.77)	0.63 (0.49-0.77)
Gliddon3	0.62 (0.47-0.76)	0.62 (0.48-0.77)	0.63 (0.48-0.77)
Gliddon4	0.62 (0.48-0.76)	0.63 (0.49-0.78)	0.64 (0.5-0.78)
Jacobsen3	0.73 (0.6-0.86)	0.69 (0.55-0.83)	0.73 (0.59-0.86)
Kaforou22	0.69 (0.55-0.82)	0.66 (0.51-0.8)	0.72 (0.59-0.85)
Maertzdorf4	0.61 (0.46-0.76)	0.67 (0.53-0.81)	0.64 (0.5-0.79)
Penn-Nicholson6	0.59 (0.45-0.74)	0.62 (0.48-0.77)	0.64 (0.5-0.78)
Rajan5	0.58 (0.44-0.73)	0.65 (0.51-0.79)	0.64 (0.5-0.77)
Roe1	0.62 (0.48-0.77)	0.65 (0.51-0.79)	0.62 (0.47-0.76)
Roe3	0.45 (0.3-0.6)	0.61 (0.46-0.75)	0.59 (0.45-0.73)
Sambarey10	0.71 (0.57-0.84)	0.69 (0.55-0.82)	0.70 (0.57-0.84)
Satproedprai7	0.66 (0.52-0.81)	0.64 (0.5-0.79)	0.66 (0.52-0.81)
Suliman2	0.59 (0.44-0.74)	0.58 (0.44-0.73)	0.63 (0.49-0.77)
Suliman4	0.59 (0.44-0.74)	0.59 (0.44-0.73)	0.67 (0.53-0.81)
Sweeney3	0.52 (0.37-0.67)	0.56 (0.42-0.71)	0.61 (0.46-0.76)
Thompson5	0.63 (0.49-0.77)	0.60 (0.46-0.75)	0.58 (0.44-0.73)

EoT = end of treatment. AUC = area under the curve. CI = confidence interval.

A sensitivity analysis was performed to determine whether batch correction and normalisation had any effect on results. A comparison of performance for transcriptomic signatures was made before and after batch correction procedures. Data batch correction and normalisation processing steps were found to have a negligible effect on the performance of transcriptomic signatures as treatment monitoring tools (**Figure 5.11**).

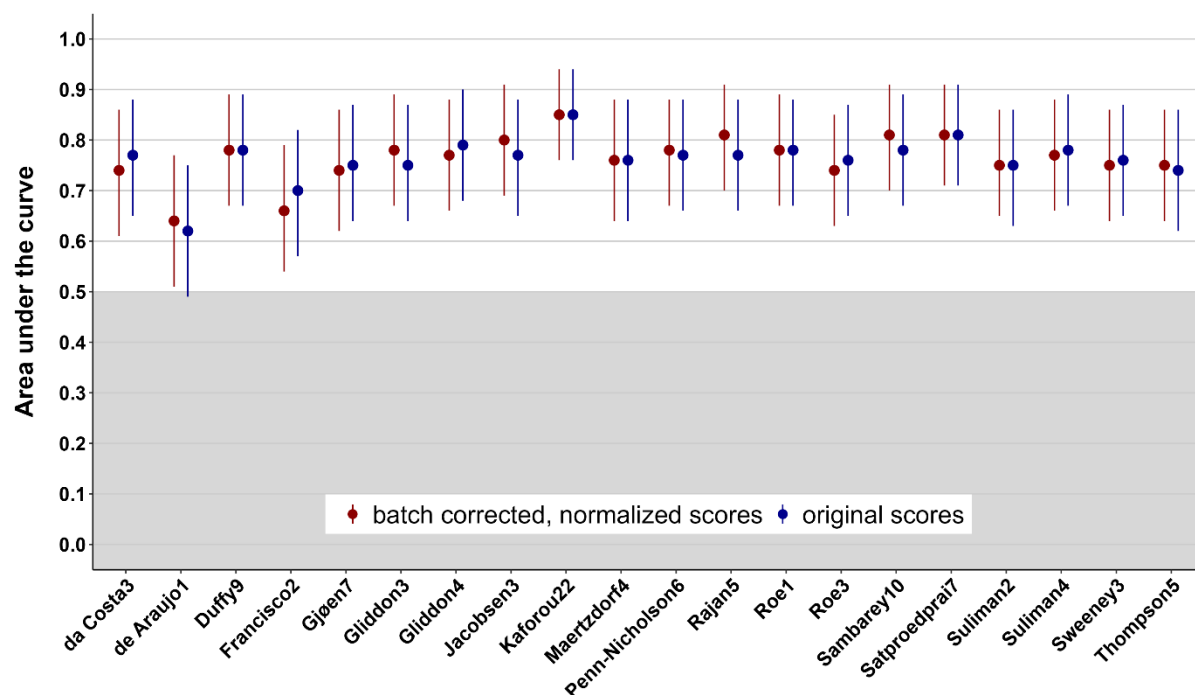


Figure 5.11: Summary of performance of transcriptomic signatures as treatment monitoring biomarkers in children with confirmed TB, before and after batch correction.

Performance is shown for comparison of scores at baseline versus end of treatment. The area under curve (AUC) for the original signature scores (blue) and batch corrected, normalised signature scores (red) for confirmed TB cases at baseline versus confirmed TB cases at the end of treatment are shown. Error bars depict the 95% CIs.

5.4: Discussion

We evaluated published blood transcriptional signatures of TB to determine their utility as non-sputum-based tools for treatment monitoring in young children with TB in the SHINE trial. Our analyses revealed five major points: baseline scores for all signatures were highest in the confirmed, relative to the unconfirmed and unlikely TB groups, baseline scores were also higher among children with disease classified as radiologically more severe on chest X-ray, scores declined progressively during TB treatment in the confirmed and unconfirmed TB groups; no changes in scores were observed for most signatures in the unlikely TB group and lastly scores were higher at the EoT in the 4-month than the 6-month treatment arm in the confirmed TB group; no differences were observed between treatment arms in the unconfirmed and unlikely groups.

We found that majority of children with confirmed TB in this cohort were older. This was consistent with a previous study that incorporated data from transcriptomic and proteomic assays. The authors reported a raised median age in their paediatric cohort, which made it easier for them to identify confirmed TB cases (Dutta et al., 2020).

Signature scores at baseline were highest in children with the most severe form of TB, based on CXR results. Similar findings have been reported in adult studies for two transcriptomic signatures that we evaluated in this sub-study. RISK6 scores significantly correlated with PET-CT scans for a cohort of TB patients from South Africa during treatment (Penn-Nicholson et al., 2020). Thompson and colleagues also report a correlation between Thompson5 scores and total glycolytic activity index measurements from PET-CT scans before and during TB treatment (Thompson et al., 2017a). Signature scores could also predict PET-CT results during treatment (Thompson et al., 2017a). This evidence further supports the hypothesis that transcriptomic signatures measured in whole blood do reflect pulmonary inflammation associated with TB disease. This evidence of consistent correlation with *Mtb*-related pathology indicates the relevance and potential of transcriptomic signatures as effective non-sputum-based biomarkers for TB treatment monitoring.

We found differences in signature scores at the end of treatment between confirmed TB cases in the 4- and 6-month arms. This finding suggests that for some children 4 months may not have been sufficient and that they would have possibly benefitted from a longer treatment duration. Higher signature scores at the end of treatment for children in the 4-month arm could also reflect inflammation that is still resolving at this early timepoint. Three percent of children in the SHINE clinical trial had an unfavourable outcome after month 4 (Turkova et al., 2022). More unfavourable outcomes in the 4-month arm were reported since there was a greater possibility of treatment extension in this group. This, however, was attributed to the open-label design of the SHINE trial and was also highlighted as one of its limitations (Turkova et al., 2022). According to the predefined statistical criteria for the SHINE trial, the 4-month arm was not inferior to the 6-month arm.

As children get older and move into their teenage years, they often present with TB which more closely resembles that in adults. The finding that median scores at baseline were higher in the confirmed TB group than the unconfirmed and unlikely groups is consistent with the premise that the former have more extensive adult-type TB disease, associated with greater systemic inflammation. Since signatures ultimately reflect levels of innate inflammation, this leads to a more prominent expression of transcriptomic signatures in the confirmed TB group. Those with unlikely TB had lowest signature scores at baseline and the lowest decrease in scores when treated. This is consistent with lower levels of inflammation, suggesting less extensive disease and pathology. It is not known how many of these had TB and the number that may have been treated unnecessarily. A

TB-specific biomarker, such as *Mtb*-specific T cell activation (TASA) (Duffy et al., 2020; Mpande et al., 2021; Musvosvi et al., 2023, 2018) may shed more light on who among this group, may have had TB.

Consistent with results from treatment response studies in adults (Bloom et al., 2012; Cliff et al., 2013; Darboe et al., 2019; Heyckendorf et al., 2021; Kaforou et al., 2013; Penn-Nicholson et al., 2019; Thompson et al., 2017b), significant drops in signature scores occurred as early as 2 weeks after treatment initiation. This drop in scores was consistent for most signatures when looked at in children with confirmed and unconfirmed TB, but not unlikely TB. Due to far lower levels of inflammation present in the unlikely TB group, the identification of signatures that represent other biological pathways may be more valuable for children in this TB category. The case definition of unlikely TB included children with or without immunologic sensitization to *Mtb*, indicating *Mtb* infection. This suggests that at least some children in this category may not have had TB at all (Graham et al., 2015), and could account for the lower scores in children with unlikely TB.

Although age and sex were not associated with signature scores in this cohort, these two factors can influence the host immune response to TB in children (Gutiérrez-González et al., 2021). From a principal component analysis, age and sex were the factors most associated with differences in transcriptomic profiles between confirmed TB cases from India and household contacts without TB that were also from India (Tornheim et al., 2020). This finding may be due to the enrolment of a healthy cohort of children without co-infections. Co-infections caused by bacteria, viruses, and fungi also modify the host immune response induced by *Mtb* infection (Gutiérrez-González et al., 2021).

Despite differences in immune responses to *Mtb* between children and adults, there are still shared host responses. We see this in one report (Tornheim et al., 2020) of up to 25.9 - 37.5% overlap between gene sets, for diagnosis and treatment monitoring discovered in a paediatric cohort from India (Tornheim et al., 2020), and those from signatures discovered in adult African populations (Kaforou et al., 2013; Sweeney et al., 2016; Zak et al., 2016). The authors claim that none of the three published treatment response signatures, from studies in adults, outperformed the 25-gene paediatric treatment response signature they discovered in confirmed TB cases. Despite this, we demonstrate that more than one signature, identified in adults, performed better across all paediatric TB categories, in comparison to the only signature on our panel, Gjøn7, that was discovered solely in a paediatric cohort (Gjøn et al., 2017). What this may point to is the importance of the methodology used to discover, train, and validate a signature and how that ultimately influences its performance. Although Kaforou22 was not assessed against their treatment response signature, this was the signature reported to have the most overlap with one of their paediatric gene lists (Tornheim et al., 2020). In our cohort, Kaforou22 had the highest AUC in differentiating confirmed and unconfirmed TB cases across more than one time point during treatment.

In a diagnostic study aimed at identifying transcriptional immune biomarkers that distinguished children with TB from their asymptomatic siblings, the authors also report overlaps in transcripts identified and those reported on in adult diagnostic studies assessing transcripts that best classify individuals with TB (Haks et al., 2015; Mihret et al., 2014; Sutherland et al., 2014). Although it is warranted to identify transcriptomic signatures for children from paediatric cohorts, the shared biological pathways driving the host immune responses in adults and children does suggest that adult signatures could also have some utility in diagnosis and treatment monitoring of paediatric TB.

Olbrich and colleagues carried out a diagnostic assessment of the MTB-HR prototype GeneXpert cartridge in children younger than 15 years through a multi-site study comprised of five low-income and middle-income countries. The MTB-HR TB-score had the highest diagnostic accuracy for culture-confirmed TB but decreased performance for Based on a strict reference standard, the MTB-HR TB-

score distinguished between children with culture-confirmed TB and unlikely TB cases with an AUC of 0.85 (95% CI 0.80–0.89 and a sensitivity of 59.8% (95% CI 50.8–68.4). A drop in performance was noted based on their composite reference standard i.e confirmed and unconfirmed TB cases combined which most likely led the authors to conclude the use of MTB-HR cartridge for culture-confirmed TB. These results are consistent with what is reported in this chapter where transcriptomic signatures worked best for confirmed TB cases. This continues to indicate the need for biomarkers that are specific for unconfirmed and unlikely TB cases in kids. Although the authors also had unconfirmed cases in their cohort at all sites, the results from the paper focused on differentiating confirmed from unlikely TB. While the analysis done is still relevant, a less biased approach that also accounts for the performance in distinguishing confirmed from unconfirmed cases would be helpful in further determining just how sensitive and specific the MTB-HR score is across all categories of TB. The authors were however, quite thorough regarding their use of three different reference standards. The composite reference standard may be the one that would work best across all three categories of TB, however the other two i.e. the strict reference standard (culture confirmation) and the composite microbiological standard may have limitations when comparing results between the unconfirmed and unlikely TB for example (Olbrich et al., 2024).

Within the field of biomarkers for paediatric TB, there is still very limited work specifically on biomarkers for monitoring treatment responses. To our knowledge, most of the work in the literature has focussed on immunological biomarkers. In Polish children with symptoms of TB and in recent TB contacts, plasma IP-10 levels, but not those of IFN- γ significantly decreased in children with active TB. This was noted particularly in children who had the highest plasma concentrations of IP-10 at baseline (Strzelak et al., 2024). Using ELISAs, dos Santos and colleagues measured the levels of serum immunoglobulin against the following antigens present within the *Mtb* cell envelope; cardiolipin, sulfatide, mycolic acid and Mce1A protein (dos Santos et al., 2020). IgM and IgG antibodies against the Mce1A protein significantly reduced during treatment whereas only IgM antibody levels against cardiolipin, sulfatide and mycolic acid decreased during treatment. Only children and adolescents with confirmed and unconfirmed TB were recruited in this study (dos Santos et al., 2020) and so the applicability of these serological biomarkers in children with unlikely TB is still to be determined. Furthermore, analyses of antibody levels were done on confirmed and unconfirmed TB cases combined, and so the extent to which antigen levels really decreased in unconfirmed TB cases is also unknown. The combined use of a QFT-Plus test and a heparin-binding haemagglutinin (HBHA)-based IGRA assay also showed promise for monitoring children's response to treatment. This was evaluated in children with symptoms of TB and those who were asymptomatic, with either radiological or microbiological evidence of disease were all classified under the active TB group (Buonsenso et al., 2023).

A combination of biomarkers across more than one 'omics' platform has also been assessed (Dutta et al., 2020). In a follow-up to the identification and evaluation of transcriptomic signatures in children from India with confirmed TB (Tornheim et al., 2020), Dutta and colleagues identified potential metabolomic markers from the same cohort (Dutta et al., 2020). Four metabolites; gammaglutamylalanine, gamma-glutamylglycine, glutamine, and pyridoxate had the highest diagnostic accuracy in differentiating pre-treatment from post-treatment samples (AUC of 0.86). This performance was not very far from that of the Kaforou22 signature that achieved an AUC of 0.85 in confirmed TB cases (Dutta et al., 2020).

When both metabolomic and transcriptomic data were consolidated, N-acetylneuraminate had the highest correlation between metabolite abundance and gene expression data (Dutta et al., 2020). The integration of transcriptomic and metabolomic data enabled the authors to verify the

involvement of certain biological pathways in TB. N-acetylneuraminate, for example, was associated with mRNA transcription in the lymphoid immunoregulatory and neutrophil degranulation pathways (Dutta et al., 2020).

Overall, biomarkers that monitor how children respond to treatment can also potentially identify children that may benefit from shorter treatment regimens (Buonsenso et al., 2023). This would require more studies to be conducted that assess biomarker-guided treatment strategies, such as that carried out with the RISK11 transcriptomic signature (Scriba et al., 2021). This would help determine the extent to which these biomarkers could be helpful in driving efforts to personalize paediatric TB treatment durations.

Strengths of our study were the comprehensive and largely unbiased selection and evaluation of candidate transcriptomic signatures in a cohort of children from the only clinical trial on treatment shortening for paediatric TB to date. To our knowledge, our study is the only one to carry out a head-to-head comparison of transcriptomic signatures in children across all three clinical classifications of paediatric TB.

Improvements to our semi-automated RNA extraction process on the TECAN led to better and higher RNA yields from whole blood samples for SHINE study participants. This is reflected in higher pass-rates, for individual transcripts, measured in the SHINE sub-study cohort, in comparison to the ScreenTB and AE-TBC studies.

PAXgene samples were not collected at the same timepoints in the 4- and 6-month arms, such that the 4-month arm only had samples at week 16, while the 6-month arm only had samples at week 24, precluding comparisons of signature scores at week 16 or week 24. Also, the lack of matching whole blood PAXgene samples at relevant time points from children recruited at SHINE study sites in India and Uganda limited our ability to evaluate these signatures in children from these geographical settings and limits the generalizability of our findings.

Another limitation of the work done comes from the lack of a healthy control group that we could compare signature scores to at the end of treatment for children under both treatment arms. These comparisons, however, have been carried out in studies that have assessed transcriptomic signatures for treatment monitoring in adults. The studies have demonstrated that signature scores at the end of treatment are not different from scores in healthy individuals (Bayaa et al., 2021; Penn-Nicholson et al., 2020). We also lacked information on any other infections that children under each of the three TB cases could have had.

We were also not able to assess the ability of the transcriptomic signatures evaluated to predict treating outcomes. Only three children that were part of our sub-study had a TB episode after completing the trial and were thus classified as having a poor treatment outcome.

We were not able to assess Satproedprai7 because a defect in one microfluidic chip led to failed PCR reactions for the KAZN primer-probe assay in more than 30% of samples. This resulted in a sample pass-rate just above 90% for Satproedprai7, which was the lowest sample pass rate in this study cohort. We also cannot rule out that the evaluation of Gliddon4 (Gliddon et al., 2021), Duffy9 (Duffy et al., 2019), and Rajan5 (Rajan et al., 2019) may have been affected by the pass-rate of the primer/probe assays for GBP6. 1.23% of samples had failed reactions for the GBP6 primer-probe assay. The pass rate for GBP6 inevitably affected signature score calculations for signatures that contain it, ultimately lowering the sample pass-rate of these signatures in this study. This is why the sample pass-rate of these signatures was at the tail end.

In conclusion, we report a head-to-head comparison of multiple blood transcriptomic signatures for TB, demonstrating their possible utility as treatment monitoring tools in children. Transcriptomic signatures or other biomarkers that are tailored to either diagnosis or treatment monitoring specifically in children may yield more definitive results, especially in those with unconfirmed TB. More focussed multi-cohort studies of children with paediatric TB, across all three clinical classifications are required. These studies should also include appropriate control groups, comprised of healthy children. More of these studies are essential since paediatric TB continues to be an area where there is still a need for an easier and more accurate assessment of responses to treatment. The next step with the evaluation of transcriptomic signatures as treatment monitoring tools for paediatric TB would be assessing whether they can predict treatment outcomes at baseline.

5.5 Contributions

Members of the SHINE clinical trial team were responsible for all site-level activities, including recruitment, clinical management, and sample and data collection. The SHINE sub-study was conceived by Dr. James Seddon. The statistical analysis plan was developed by me with input and final approval from Dr James Seddon, Professor Thomas J. Scriba, Dr Simon C. Mendelsohn, Dr Myrsini Kaforou, Dr Ortensia Vito and Dr Marieke Van der Zalm. Mzwandile Erasmus and Rodney Raphela performed the RNA extractions and cDNA synthesis for all the samples received. I carried out all the subsequent gene expression experiments. All data verification and analysis were performed by me under the supervision of Professor Thomas J Scriba and Dr Simon C. Mendelsohn with input from Dr James Seddon. I interpreted the results and wrote this thesis chapter.

References

- Anderson, S.T., Kaforou, M., Brent, A.J., Wright, V.J., Banwell, C.M., Chagaluka, G., Crampin, A.C., Dockrell, H.M., French, N., Hamilton, M.S., Hibberd, M.L., Kern, F., Langford, P.R., Ling, L., Mlotha, R., Ottenhoff, T.H.M., Pienaar, S., Pillay, V., Scott, J.A.G., Twahir, H., Wilkinson, R.J., Coin, L.J., Heyderman, R.S., Levin, M., Eley, B., 2014. Diagnosis of Childhood Tuberculosis and Host RNA Expression in Africa. *N. Engl. J. Med.*
- Bayaa, R., Ndiaye, M.D.B., Chedid, C., Kokhraidze, E., Tukvadze, N., Banu, S., Uddin, M.K.M., Biswas, S., Nasrin, R., Ranaivomanana, P., Raherinandrasana, A.H., Rakotonirina, J., Rasolofo, V., Delogu, G., De Maio, F., Goletti, D., Endtz, H., Ader, F., Hamze, M., Ismail, M.B., Pouzol, S., Rakotosamimanana, N., Hoffmann, J., The HINTT working group within the GABRIEL network, Russomando, G., Acosta, C.C.D., Arenas, R., 2021. Multi-country evaluation of RISK6, a 6-gene blood transcriptomic signature, for tuberculosis diagnosis and treatment monitoring. *Sci. Rep.* 11, 13646. <https://doi.org/10.1038/s41598-021-93059-1>
- Bloom, C.I., Graham, C.M., Berry, M.P.R., Wilkinson, K.A., Oni, T., Rozakeas, F., Xu, Z., Rossello-Urgell, J., Chaussabel, D., Banchereau, J., Pascual, V., Lipman, M., Wilkinson, R.J., O'Garra, A., 2012. Detectable Changes in The Blood Transcriptome Are Present after Two Weeks of Antituberculosis Therapy. *PLOS ONE* 7, e46191. <https://doi.org/10.1371/journal.pone.0046191>
- Bobak, C.A., Botha, M., Workman, L., Hill, J.E., Nicol, M.P., Holloway, J.W., Stein, D.J., Martinez, L., Zar, H.J., 2023. Gene Expression in Cord Blood and Tuberculosis in Early Childhood: A Nested Case-Control Study in a South African Birth Cohort. *Clin. Infect. Dis.* 77, 438–449. <https://doi.org/10.1093/cid/ciad268>
- Buonsenso, D., Delogu, G., del Carmen Pereyra Boza, M., De Maio, F., Palucci, I., Martino, L., Pata, D., Sanguinetti, M., Valentini, P., Sali, M., 2023. Commercially available CD4 + and CD8 + IFN- γ release assays combined with an HBHA-induced IGRA improve the characterization of the tuberculosis spectrum and monitoring of treatment in children. *Eur. J. Pediatr.* 182, 2155–2167. <https://doi.org/10.1007/s00431-023-04844-1>
- Chabala, C., Turkova, A., Thomason, M.J., Wobudeya, E., Hissar, S., Mave, V., van der Zalm, M., Palmer, M., Kapasa, M., Bhavani, P.K., Balaji, S., Raichur, P.A., Demers, A.-M., Hoddinott, G., Owen-Powell, E., Kinikar, A., Musoke, P., Mulenga, V., Aarnoutse, R., McIlleron, H., Hesselting, A., Crook, A.M., Cotton, M., Gibb, D.M., SHINE trial team, 2018. Shorter treatment for minimal tuberculosis (TB) in children (SHINE): a study protocol for a randomised controlled trial. *Trials* 19, 237. <https://doi.org/10.1186/s13063-018-2608-5>
- Cliff, J.M., Lee, J.-S., Constantinou, N., Cho, J.-E., Clark, T.G., Ronacher, K., King, E.C., Lukey, P.T., Duncan, K., Van Helden, P.D., Walzl, G., Dockrell, H.M., 2013. Distinct Phases of Blood Gene Expression Pattern Through Tuberculosis Treatment Reflect Modulation of the Humoral Immune Response. *J. Infect. Dis.* 207, 18–29. <https://doi.org/10.1093/infdis/jis499>
- Darboe, F., Mbandi, S.K., Naidoo, K., Yende-Zuma, N., Lewis, L., Thompson, E.G., Duffy, F.J., Fisher, M., Filander, E., van Rooyen, M., Bilek, N., Mabwe, S., McKinnon, L.R., Chegou, N., Loxton, A., Walzl, G., Tromp, G., Padayatchi, N., Govender, D., Hatherill, M., Karim, S.A., Zak, D.E., Penn-Nicholson, A., Scriba, T.J., Team, T.S.C.I., 2019. Detection of Tuberculosis Recurrence, Diagnosis and Treatment Response by a Blood Transcriptomic Risk Signature in HIV-Infected Persons on Antiretroviral Therapy. *Front. Microbiol.* 10. <https://doi.org/10.3389/fmicb.2019.01441>
- DeLong, E.R., DeLong, D.M., Clarke-Pearson, D.L., 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44, 837–845.
- Doorn, C.L.R. van, Eckold, C., Ronacher, K., Ruslami, R., Veen, S. van, Lee, J.-S., Kumar, V., Kerry-Barnard, S., Malherbe, S.T., Kleynhans, L., Stanley, K., Hill, P.C., Joosten, S.A., Crevel, R. van, Wijmenga, C., Critchley, J.A., Walzl, G., Alisjahbana, B., Haks, M.C., Dockrell, H.M., Ottenhoff, T.H.M., Vianello, E., Cliff, J.M., 2022. Transcriptional profiles predict treatment outcome in

- patients with tuberculosis and diabetes at diagnosis and at two weeks after initiation of anti-tuberculosis treatment. *eBioMedicine* 82. <https://doi.org/10.1016/j.ebiom.2022.104173>
- dos Santos, D.C.M., Lovero, K.L., Schmidt, C.M., Barros, A.C.M.W., Quintanilha, A.P., Barbosa, A.P., Pone, M.V.S., Pone, S.M., Araujo, J.M., Martins, C. de P., Cosme, E.M., Dourado de Oliveira, T.R., Miceli, A.L., Vieira, M.L., Queiroz, A., Velarde, L.G.C., Kritski, A., March, M. de F.P., Sias, S.M. de A., Sant'Anna, C.C., Riley, L.W., Araújo Cardoso, C.A., 2020. Serological biomarkers for monitoring response to treatment of pulmonary and extrapulmonary tuberculosis in children and adolescents. *Tuberc. Edinb. Scotl.* 123, 101960. <https://doi.org/10.1016/j.tube.2020.101960>
- Duffy, D., Nemes, E., Llibre, A., Rouilly, V., Musvosvi, M., Smith, N., Filander, E., Africa, H., Mabwe, S., Jaxa, L., Charbit, B., Mulenga, H., Tameris, M., Walzl, G., Malherbe, S., Thomas, S., Hatherill, M., Bilek, N., Scriba, T.J., Albert, M.L., 2020. Immune Profiling Enables Stratification of Patients With Active Tuberculosis Disease or Mycobacterium tuberculosis Infection. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* 73, e3398–e3408. <https://doi.org/10.1093/cid/ciaa1562>
- Duffy, F.J., Thompson, E.G., Scriba, T.J., Zak, D.E., 2019. Multinomial modelling of TB/HIV co-infection yields a robust predictive signature and generates hypotheses about the HIV+TB+ disease state. *PLOS ONE* 14, e0219322. <https://doi.org/10.1371/journal.pone.0219322>
- Dutta, N.K., Tornheim, J.A., Fukutani, K.F., Paradkar, M., Tiburcio, R.T., Kinikar, A., Valvi, C., Kulkarni, V., Pradhan, N., Shivakumar, S.V.B.Y., Kagal, A., Gupte, A., Gupte, N., Mave, V., Gupta, A., Andrade, B.B., Karakousis, P.C., 2020. Integration of metabolomics and transcriptomics reveals novel biomarkers in the blood for tuberculosis diagnosis in children. *Sci. Rep.* 10, 19527. <https://doi.org/10.1038/s41598-020-75513-8>
- Eckold, C., van Doorn, C.L.R., Ruslami, R., Ronacher, K., Riza, A.-L., van Veen, S., Lee, J.-S., Kumar, V., Kerry-Barnard, S., Malherbe, S.T., Kleynhans, L., Stanley, K., Joosten, S.A., Critchley, J.A., Hill, P.C., van Crevel, R., Wijmenga, C., Haks, M.C., Ioana, M., Alisjahbana, B., Walzl, G., Ottenhoff, T.H.M., Dockrell, H.M., Vianello, E., Cliff, J.M., Consortium\$, the T., 2023. Impaired resolution of blood transcriptomes through tuberculosis treatment with diabetes comorbidity. *Clin. Transl. Med.* 13, e1375. <https://doi.org/10.1002/ctm2.1375>
- Gjøl, J.E., Jenum, S., Sivakumaran, D., Mukherjee, A., Macaden, R., Kabra, S.K., Lodha, R., Ottenhoff, T.H.M., Haks, M.C., Doherty, T.M., Ritz, C., Grewal, H.M.S., 2017. Novel transcriptional signatures for sputum-independent diagnostics of tuberculosis in children. *Sci. Rep.* 7, 5839. <https://doi.org/10.1038/s41598-017-05057-x>
- Gliddon, H.D., Kaforou, M., Alikian, M., Habgood-Coote, D., Zhou, C., Oni, T., Anderson, S.T., Brent, A.J., Crampin, A.C., Eley, B., Heyderman, R., Kern, F., Langford, P.R., Ottenhoff, T.H.M., Hibberd, M.L., French, N., Wright, V.J., Dockrell, H.M., Coin, L.J., Wilkinson, R.J., Levin, M., 2021. Identification of Reduced Host Transcriptomic Signatures for Tuberculosis Disease and Digital PCR-Based Validation and Quantification. *Front. Immunol.* 12.
- Graham, S.M., Cuevas, L.E., Jean-Philippe, P., Browning, R., Casenghi, M., Detjen, A.K., Gnanashanmugam, D., Hesselink, A.C., Kampmann, B., Mandalakas, A., Marais, B.J., Schito, M., Spiegel, H.M.L., Starke, J.R., Worrell, C., Zar, H.J., 2015. Clinical Case Definitions for Classification of Intrathoracic Tuberculosis in Children: An Update. *Clin. Infect. Dis.* 61, S179–S187. <https://doi.org/10.1093/cid/civ581>
- Gutiérrez-González, L.H., Juárez, E., Carranza, C., Carreto-Binaghi, L.E., Alejandre, A., Cabello-Gutiérrez, C., Gonzalez, Y., 2021. Immunological Aspects of Diagnosis and Management of Childhood Tuberculosis. *Infect. Drug Resist.* 14, 929–946. <https://doi.org/10.2147/IDR.S295798>
- Haks, M.C., Goeman, J.J., Magis-Escurra, C., Ottenhoff, T.H.M., 2015. Focused human gene expression profiling using dual-color reverse transcriptase multiplex ligation-dependent probe amplification. *Vaccine, Systems Vaccinology: How Big Data Can Accelerate Vaccine Development* 33, 5282–5288. <https://doi.org/10.1016/j.vaccine.2015.04.054>

- Heyckendorf, J., Marwitz, S., Reimann, M., Avsar, K., DiNardo, A.R., Günther, G., Hoelscher, M., Ibraim, E., Kalsdorf, B., Kaufmann, S.H.E., Kontsevaya, I., van Leth, F., Mandalakas, A.M., Maurer, F.P., Müller, M., Nitschkowski, D., Olaru, I.D., Popa, C., Rachow, A., Rolling, T., Rybniker, J., Salzer, H.J.F., Sanchez-Carballo, P., Schuhmann, M., Schaub, D., Spinu, V., Suárez, I., Terhalle, E., Unnewehr, M., Weiner, J., Goldmann, T., Lange, C., 2021. Prediction of anti-tuberculosis treatment duration based on a 22-gene transcriptomic model. *Eur. Respir. J.* 58, 2003492. <https://doi.org/10.1183/13993003.03492-2020>
- Kaforou, M., Broderick, C., Vito, O., Levin, M., Scriba, T.J., Seddon, J.A., 2022. Transcriptomics for child and adolescent tuberculosis*. *Immunol. Rev.* 309, 97–122. <https://doi.org/10.1111/imr.13116>
- Kaforou, M., Wright, V.J., Oni, T., French, N., Anderson, S.T., Bangani, N., Banwell, C.M., Brent, A.J., Crampin, A.C., Dockrell, H.M., Eley, B., Heyderman, R.S., Hibberd, M.L., Kern, F., Langford, P.R., Ling, L., Mendelson, M., Ottenhoff, T.H., Zgambo, F., Wilkinson, R.J., Coin, L.J., Levin, M., 2013. Detection of Tuberculosis in HIV-Infected and -Uninfected African Adults Using Whole Blood RNA Expression Signatures: A Case-Control Study. *PLOS Med.* 10, e1001538. <https://doi.org/10.1371/journal.pmed.1001538>
- Mendelsohn, S.C., Mbandi, S.K., Fiore-Gartland, A., Penn-Nicholson, A., Musvosvi, M., Mulenga, H., Fisher, M., Hadley, K., Erasmus, M., Nombida, O., Tameris, M., Walzl, G., Naidoo, K., Churchyard, G., Hatherill, M., Scriba, T.J., 2022. Prospective multicentre head-to-head validation of host blood transcriptomic biomarkers for pulmonary tuberculosis by real-time PCR. *Commun. Med.* 2, 1–13. <https://doi.org/10.1038/s43856-022-00086-8>
- Mihret, A., Loxton, A.G., Bekele, Y., Kaufmann, S.H., Kidd, M., Haks, M.C., Ottenhoff, T.H., Aseffa, A., Howe, R., Walzl, G., 2014. Combination of gene expression patterns in whole blood discriminate between tuberculosis infection states. *BMC Infect. Dis.* 14, 257. <https://doi.org/10.1186/1471-2334-14-257>
- Mpande, C.A.M., Musvosvi, M., Rozot, V., Mosito, B., Reid, T.D., Schreuder, C., Lloyd, T., Bilek, N., Huang, H., Obermoser, G., Davis, M.M., Ruhwald, M., Hatherill, M., Scriba, T.J., Nemes, E., Mahomed, H., Hanekom, W.A., Kafaar, F., Workman, L., Mulenga, H., Ehrlich, R., Erasmus, M., Abrahams, D., Hawkridge, A., Hughes, E.J., Moyo, S., Gelderbloem, S., Tameris, M., Geldenhuys, H., Hussey, G., 2021. Antigen-Specific T-Cell Activation Distinguishes between Recent and Remote Tuberculosis Infection. *Am. J. Respir. Crit. Care Med.* 203, 1556–1565. <https://doi.org/10.1164/rccm.202007-2686OC>
- Musvosvi, M., Duffy, D., Filander, E., Africa, H., Mabwe, S., Jaxa, L., Bilek, N., Llibre, A., Rouilly, V., Hatherill, M., Albert, M., Scriba, T.J., Nemes, E., 2018. T-cell biomarkers for diagnosis of tuberculosis: candidate evaluation by a simple whole blood assay for clinical translation. *Eur. Respir. J.* 51. <https://doi.org/10.1183/13993003.00153-2018>
- Musvosvi, M., Huang, H., Wang, C., Xia, Q., Rozot, V., Krishnan, A., Acs, P., Cheruku, A., Obermoser, G., Leslie, A., Behar, S.M., Hanekom, W.A., Bilek, N., Fisher, M., Kaufmann, S.H.E., Walzl, G., Hatherill, M., Davis, M.M., Scriba, T.J., 2023. T cell receptor repertoires associated with control and disease progression following Mycobacterium tuberculosis infection. *Nat. Med.* 29, 258–269. <https://doi.org/10.1038/s41591-022-02110-9>
- Olbrich, L., Verghese, V.P., Franckling-Smith, Z., Sabi, I., Ntinginya, N.E., Mfinanga, A., Banze, D., Viegas, S., Khosa, C., Semphere, R., Nliwasa, M., McHugh, T.D., Larsson, L., Razid, A., Song, R., Corbett, E.L., Nabeta, P., Trollip, A., Graham, S.M., Hoelscher, M., Geldmacher, C., Zar, H.J., Michael, J.S., Heinrich, N., Baard, C.B., Munro, J.D., Prins, M., Benzi, N., Bateman, L.C., Ryan, A., Booi, K., Paulo, N., Heydenrych, A., Petersen, W., Brookes, R., Mento, M., Centner, C., Dalgarno, C., Rieß, F., Mutuku, S., Saathoff, E., Held, K., Ninan, M.M., Chacko, A., Kumari, R., Dhanabhagyan, R., Muniswamy, N., Nicol, M.P., Mtafya, B., Mwambola, H., Manyama, C., Mahiga, H., Sichone, E., Sudi, L., Maueia, C., Madeira, C., Cambuie, J., Ribeiro, J., Chiume, L., Mnyanga, A., Sikwese, T., Masakasa, H., Kachere, D., Kosaka, M., Niemann, S., Chegou, N., Horn, L., 2024. Diagnostic accuracy of a three-gene Mycobacterium tuberculosis host

- response cartridge using fingerstick blood for childhood tuberculosis: a multicentre prospective study in low-income and middle-income countries. *Lancet Infect. Dis.* 24, 140–149. [https://doi.org/10.1016/S1473-3099\(23\)00491-7](https://doi.org/10.1016/S1473-3099(23)00491-7)
- Penn-Nicholson, A., Mbandi, S.K., Thompson, E., Mendelsohn, S., Suliman, S., Chegou, N.N., Malherbe, S.T., Darboe, F., Erasmus, M., Hanekom, W.A., Bilek, N., Fisher, M., Kaufmann, S.H., Winter, J., Murphy, M., Wood, R., Morrow, C., Rhijn, I.V., Moody, D.B., Murray, M., Andrade, B., Sterling, T., Sutherland, J., Naidoo, K., Padayatchi, N., Walzl, G., Hatherill, M., Zak, D., Scriba, T., 2019. RISK6, a universal 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response. *medRxiv* 19006197. <https://doi.org/10.1101/19006197>
- Penn-Nicholson, A., Mbandi, S.K., Thompson, E., Mendelsohn, S.C., Suliman, S., Chegou, N.N., Malherbe, S.T., Darboe, F., Erasmus, M., Hanekom, W.A., Bilek, N., Fisher, M., Kaufmann, S.H.E., Winter, J., Murphy, M., Wood, R., Morrow, C., Van Rhijn, I., Moody, B., Murray, M., Andrade, B.B., Sterling, T.R., Sutherland, J., Naidoo, K., Padayatchi, N., Walzl, G., Hatherill, M., Zak, D., Scriba, T.J., 2020. RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response. *Sci. Rep.* 10, 8629. <https://doi.org/10.1038/s41598-020-65043-8>
- Perez-Velez, C.M., Marais, B.J., 2012. Tuberculosis in Children. *N. Engl. J. Med.* 367, 348–361. <https://doi.org/10.1056/NEJMra1008049>
- Rajan, J.V., Semitala, F.C., Mehta, T., Seielstad, M., Montalvo, L., Andama, A., Asege, L., Nakaye, M., Katende, J., Mwebe, S., Kanya, M.R., Yoon, C., Cattamanchi, A., 2019. A Novel, 5-Transcript, Whole-blood Gene-expression Signature for Tuberculosis Screening Among People Living With Human Immunodeficiency Virus. *Clin. Infect. Dis.* 69, 77–83. <https://doi.org/10.1093/cid/ciy835>
- Scriba, T.J., Fiore-Gartland, A., Penn-Nicholson, A., Mulenga, H., Mbandi, S.K., Borate, B., Mendelsohn, S.C., Hadley, K., Hikuam, C., Kaskar, M., Musvosvi, M., Bilek, N., Self, S., Sumner, T., White, R.G., Erasmus, Mzwandile, Jaxa, L., Raphela, R., Innes, C., Brumskine, W., Hiemstra, Andriëtte, Malherbe, S.T., Hassan-Moosa, R., Tameris, M., Walzl, G., Naidoo, K., Churchyard, G., Hatherill, M., Baepanye, K., Baepanye, T., Clarke, K., Collignon, M., Dlamini, A., Eyre, C., Feni, T., Fikizolo, M., Galane, P., Goliath, T., Gangat, A., Malefo-Grootboom, S., Rensburg, E.J. van, Rensburg, B.J. van, Kekana, S., Zietsman, M., Kock, A., Kunene, I., Lakhi, A., Langa, N., Ledwaba, H., Lophoko, M., Mabasa, I., Mabe, D., Mabuza, N., Majola, M., Makhetha, M., Makoanyane, M., Makhubalo, B., Malay, V., Market, J., Matshego, S., Mbipa, N., Mmotsa, T., Modipa, S., Mopati, S., Moswegu, P., Mothaga, P., Muller, D., Nchwe, G., Nel, M., Nhlangulela, L., Ntamo, B., Ntoahae, L., Ntshauba, T., Sanyaka, N., Seabela, L., Selepe, P., Senne, M., Serake, M.G., Thlapi, M., Tshikovhi, V., Tswaile, L., Aswegen, A. van, Mbata, L., Takavamanya, C., Pinho, P., Mdlulu, J., Taljaard, M., Slabbert, N., Sayed, S., Nielson, T., Senne, M., Sein, N.N., Mbata, L., Govender, D., Chinappa, T., Zulu, M.I., Maphanga, N.B., Hlathi, S.R., Gumede, G.K., Shezi, T.Y., Maphanga, J.L., Jali, Z.P., Cwele, T., Gwamanda, N.Z.E., Dlamini, C., Sing, Z.P.P., Ntanjana, N.G., Nzimande, S.S., Mbatha, S., Maharaj, B., Moosa, A., Corris, C.-M., Kafaar, F., Geldenhuys, H., Luabeya, A.K.K., Shenje, J., Botes, N., Rossouw, S., Africa, H., Diamond, B., Braaf, S., Stryers, S., Carstens, A., Jansen, R., Mabwe, S., Mulenga, H., Herling, R., Veldsman, A., Makhete, L., Steyn, M., Buhlungu, S., Erasmus, Margareth, Davids, I., Plaatjie, P., Companie, A., Ratangee, F., Veldtsman, H., Petersen, C., Abrahams, C., Moses, M., Kelepu, X., Gregg, Y., Swanepoel, L., Magawu, N., Cetywayo, N., Mactavie, L., Valley, H., Filander, E., Nqakala, N., Maasdorp, E., Khoury, J., Kriel, B., Smith, B., Muller, L., Tonsing, S., Loxton, A., Hiemstra, Andriette, Ahlers, P., Flinn, M., Chung, E., Chung, M., Sato, A., 2021. Biomarker-guided tuberculosis preventive therapy (CORTIS): a randomised controlled trial. *Lancet Infect. Dis.* 21, 354–365. [https://doi.org/10.1016/S1473-3099\(20\)30914-2](https://doi.org/10.1016/S1473-3099(20)30914-2)
- Strzelak, A., Komorowska-Piotrowska, A., Borowa, A., Krasińska, M., Feleszko, W., Kulus, M., 2024. IP-10 for the Diagnosis and Treatment Monitoring of Tuberculosis in Children. *Diagnostics* 14, 177. <https://doi.org/10.3390/diagnostics14020177>

- Sutherland, J.S., Loxton, A.G., Haks, M.C., Kassa, D., Ambrose, L., Lee, J.-S., Ran, L., van Baarle, D., Maertzdorf, J., Howe, R., Mayanja-Kizza, H., Boom, W.H., Thiel, B.A., Crampin, A.C., Hanekom, W., Ota, M.O.C., Dockrell, H., Walzl, G., Kaufmann, S.H.E., Ottenhoff, T.H.M., 2014. Differential gene expression of activating Fcγ receptor classifies active tuberculosis regardless of human immunodeficiency virus status or ethnicity. *Clin. Microbiol. Infect.* 20, O230–O238. <https://doi.org/10.1111/1469-0691.12383>
- Sweeney, T.E., Braviak, L., Tato, C.M., Khatri, P., 2016. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. *Lancet Respir. Med.* 4, 213–224. [https://doi.org/10.1016/S2213-2600\(16\)00048-5](https://doi.org/10.1016/S2213-2600(16)00048-5)
- Thompson, E.G., Du, Y., Malherbe, S.T., Shankar, S., Braun, J., Valvo, J., Ronacher, K., Tromp, G., Tabb, D.L., Alland, D., Shenai, S., Via, L.E., Warwick, J., Aderem, A., Scriba, T.J., Winter, J., Walzl, G., Zak, D.E., Du Plessis, N., Loxton, A.G., Chegou, N.N., Lee, M., 2017a. Host blood RNA signatures predict the outcome of tuberculosis treatment. *Tuberculosis* 107, 48–58. <https://doi.org/10.1016/j.tube.2017.08.004>
- Thompson, E.G., Du, Y., Malherbe, S.T., Shankar, S., Braun, J., Valvo, J., Ronacher, K., Tromp, G., Tabb, D.L., Alland, D., Shenai, S., Via, L.E., Warwick, J., Aderem, A., Scriba, T.J., Winter, J., Walzl, G., Zak, D.E., Du Plessis, N., Loxton, A.G., Chegou, N.N., Lee, M., 2017b. Host blood RNA signatures predict the outcome of tuberculosis treatment. *Tuberculosis* 107, 48–58. <https://doi.org/10.1016/j.tube.2017.08.004>
- Tornheim, J.A., Madugundu, A.K., Paradkar, M., Fukutani, K.F., Queiroz, A.T.L., Gupte, N., Gupte, A.N., Kinikar, A., Kulkarni, V., Balasubramanian, U., Sreenivasamurthy, S., Raja, R., Pradhan, N., Shivakumar, S.V.B.Y., Valvi, C., Hanna, L.E., Andrade, B.B., Mave, V., Pandey, A., Gupta, A., CTRIUMPh RePORT India Study Team, 2020. Transcriptomic Profiles of Confirmed Pediatric Tuberculosis Patients and Household Contacts Identifies Active Tuberculosis, Infection, and Treatment Response Among Indian Children. *J. Infect. Dis.* 221, 1647–1658. <https://doi.org/10.1093/infdis/jiz639>
- Turkova, A., Wills, G.H., Wobudeya, E., Chabala, C., Palmer, M., Kinikar, A., Hissar, S., Choo, L., Musoke, P., Mulenga, V., Mave, V., Joseph, B., LeBeau, K., Thomason, M.J., Mboizi, R.B., Kapasa, M., van der Zalm, M.M., Raichur, P., Bhavani, P.K., McIlleron, H., Demers, A.-M., Aarnoutse, R., Love-Koh, J., Seddon, J.A., Welch, S.B., Graham, S.M., Hesselning, A.C., Gibb, D.M., Crook, A.M., 2022. Shorter Treatment for Nonsevere Tuberculosis in African and Indian Children. *N. Engl. J. Med.* 386, 911–922. <https://doi.org/10.1056/NEJMoa2104535>
- Verhagen, L.M., Zomer, A., Maes, M., Villalba, J.A., del Nogal, B., Eleveld, M., van Hijum, S.A., de Waard, J.H., Hermans, P.W., 2013. A predictive signature gene set for discriminating active from latent tuberculosis in Warao Amerindian children. *BMC Genomics* 14, 74. <https://doi.org/10.1186/1471-2164-14-74>
- World Health Organization, W., 2023. Global Tuberculosis Report 2023 [WWW Document]. WHO Glob. Tuberc. Rep. 2023. URL <https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2023> (accessed 3.26.24).
- Zak, D.E., Penn-Nicholson, A., Scriba, T.J., Thompson, E., Suliman, S., Amon, L.M., Mahomed, H., Erasmus, M., Whatney, W., Hussey, G.D., Abrahams, D., Kafar, F., Hawkrigde, T., Verver, S., Hughes, E.J., Ota, M., Sutherland, J., Howe, R., Dockrell, H.M., Boom, W.H., Thiel, B., Ottenhoff, T.H.M., Mayanja-Kizza, H., Crampin, A.C., Downing, K., Hatherill, M., Valvo, J., Shankar, S., Parida, S.K., Kaufmann, S.H.E., Walzl, G., Aderem, A., Hanekom, W.A., 2016. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. *The Lancet* 387, 2312–2322. [https://doi.org/10.1016/S0140-6736\(15\)01316-1](https://doi.org/10.1016/S0140-6736(15)01316-1)

Chapter 6: Conclusions

This chapter summarises major findings from chapters four and five. Here I also highlight lessons learnt and insights on further work to be done in the field of transcriptomics, to advance the discovery of non-sputum-based TB biomarkers.

6.1 Summary of major findings from this thesis

Chapter 4

Nine of the twenty transcriptomic signatures evaluated had indistinguishable performance for differentiating active TB cases from individuals with ORDs. These signatures were discovered and validated in different populations using various statistical and modelling approaches, and yet none of them met the WHO TPP benchmarks when data from all study sites were analysed altogether. The likelihood that any other signatures that are developed meet all the WHO TPP benchmarks in a similar multi-site study is unlikely. This partly results from the bottleneck still present regarding the influence and presence of ISGs within several signatures. This becomes an issue due to ISGs also being upregulated in response to viral infections. Due to the presence of ISGs, there was also a high level of correlation among signatures and overlap in transcripts. This would be one of the most likely contributing factors to no differences in diagnostic performance for several signatures. Most efforts to discover signatures rely on a feature selection step that seeks to identify the most differentially expressed genes, which consistently end up being mostly ISGs. Studies that seek to replicate this same strategy are unlikely to identify signatures that significantly outperform those we assessed here. Country had an effect on signature scores which we hypothesize to be driven mainly by the control group of individuals classified as having ORDs. This most likely contributed to transcriptomic signatures achieving the minimal WHO benchmarks in some countries (Ethiopia, Malawi, and South Africa) but not all that were part of our study.

Chapter 5

We demonstrated the capacity for transcriptomic signatures, majority of which were discovered in adults, in tracking responses to TB treatment in children with paediatric TB. Although the identification and evaluation of transcriptomic signatures, solely from paediatric cohorts, would be more appropriate, the results from this work show that there are shared biological pathways involved in the host immune response to both adult and paediatric TB, that transcriptomic signatures are able to capture. This is further supported by reports in the literature on the overlap in transcripts between signatures discovered in children and those in adults. The differences in signature scores, in children across TB categories, would support greater utility of adult transcriptomic signatures in confirmed TB cases as opposed to unconfirmed and unlikely TB cases. This arises from most confirmed TB cases being children who were older and therefore had inflammatory responses that were closer to those of adults. These responses were more easily picked up and represented by the transcriptomic signatures evaluated in this sub-study.

In the following sections and sub-headings, I highlight the major themes related to transcriptomic signatures and non-sputum biomarkers that emerge from this thesis and its results. Relevant findings from the literature are also included to further support the ideas presented.

6.2 Rethinking WHO TPPs for non-sputum based biomarker tests for TB

6.2 Rethinking WHO TPPs for non-sputum based biomarker tests for TB

In 2014, WHO released TPPs detailing the most pressing needs for TB diagnostics (World Health Organization, 2014). These TPPs provide a helpful guide for both researchers and product developers by describing the characteristics of tests that need to be brought into the market, as well as the requirements that these tests should meet (Kik et al., 2014; MacLean et al., 2023; World Health Organization, 2014). One example of these would be the TPP for a point-of-care triage test (Table 6.1), which were used to benchmark the performance of signatures in Chapter 4. A triage test was defined as a simple, low-cost test that can be used by first contact health providers to rule-out TB and adequately identify patients who need further testing. The TPP also highlighted the need for a triage test to have potential as a screening tool for active case finding.

Table 6.1: WHO performance benchmarks for a triage or referral test for TB

	Optimal (ideal) requirements	Minimal requirements
Diagnostic sensitivity	95%	90%
Diagnostic specificity	80%	70%

A TPP was also developed for a point-of-care non-sputum based test able to detect all forms of TB in both adults and children (Table 6.2). This test would be based on the measurement of either biomarkers or biosignatures, be able to increase the number of TB cases detected and be able to be used by health workers with minimum training.

Table 6.2: WHO performance benchmarks for a rapid biomarker test for TB

	Optimal (ideal) requirements	Minimal requirements
Diagnostic sensitivity for pulmonary TB in adults	98%	65%
Diagnostic sensitivity for extrapulmonary TB in adults	80%	None defined
Diagnostic sensitivity in children	66%	None defined
Diagnostic specificity	98% when compared against a microbiological reference standard or at least be as specific as Xpert MTB/RIF assay	

As these TPPs are just a starting point, a test that does not fully meet the stipulated requirements, but is in a position to impact healthcare, can still be considered for WHO policy recommendations (MacLean et al., 2023). Several transcriptomic signatures such as Sweeney3, Penn-Nicholson6 (RISK6) and Heyckendorf22, for example, have demonstrated potential for test development, despite some of them not consistently meeting all use-case scenarios and performance criteria stipulated within the TPPs. A joint understanding and appreciation for this is something that WHO and other global

health stakeholders may have to openly communicate so that more transcriptomic signatures, and other potential biomarkers, can still make it as diagnostic tests on the market. Lipoarabinomannan (LAM), for example, is a biomarker that has been endorsed by WHO for the detection of active TB in PLWHIV even though it does not meet the TPPs. Lateral flow assays based on the detection of LAM in urine have however still been recommended and used in clinical practice. This is because they are the best-performing diagnostic for TB in very immunocompromised individuals, where there are no other good diagnostics present. From our diagnostic study in adults (**Chapter 4**), the Sweeney3 signature did not meet TPP thresholds and was not among the top-ranking signatures that distinguished TB cases from ORDs, and yet it has been developed into a host response (HR) cartridge device by Cepheid. What is yet to be seen is whether this host response cartridge will be licensed as a triage test.

Since the inception of the TPPs, the needs of the TB community have evolved and as such TPPs may have to be revised to continue reflecting the current diagnostic gaps in the TB field, as well as lessons learnt with regards to product development during the past decade (Abdulgader et al., 2022). The WHO TPPs may therefore need to be amended to acknowledge the existence of TB disease across a spectrum. Considering this, TPPs that are specific for incipient and sub-clinical disease may be necessary.

Current changes in action have involved the expansion of TPPs to now include those for a test predicting progression from *Mtb* infection to active disease, next-generation drug-susceptibility testing at peripheral centres, and the use of NGS for detection of resistance-associated mutations in *Mtb* complex (FIND, 2023). Amendments have already been made to the TPP for next-generation *Mtb* drug susceptibility testing at peripheral centres (MacLean et al., 2023) and now a TPP for tests aimed at TB treatment monitoring and optimization has recently been published (Gupta-Wright et al., 2023). Aside from TB diagnostic needs, WHO has gone a step further to develop TPPs outlining standards for future TPT regimens (Den Boon et al., 2022).

From the results generated in Chapters 4 and 5, different signatures performed best in distinguishing TB versus ORDs and in tracking responses to treatment in children. Within the paediatric cohort, different signatures performed best, for each of the three TB categories, in distinguishing children at baseline from those at all other time points. While WHO TPPs may call for biomarkers that suit all use cases for TB diagnosis or treatment monitoring across different population groups, the identification of an ideal and universal biomarker is challenging, especially when addressing paediatric TB (Gutiérrez-González et al., 2021). This is attributed to the host immune response and the biomarkers that reflect it being influenced by multiple factors such as age, nutritional status, geography, disease burden, genetics, and co-morbidities (Gutiérrez-González et al., 2021). Based on differences in the magnitude of host immune responses between adults and children, WHO TPPs may even need to consider different thresholds for non-sputum based biomarkers aimed at diagnosis and treatment monitoring in adults versus those for children. In line with this, researchers, product developers and policy makers should keep in mind pre-test probability when it comes to cut-offs for sensitivity and specificity of biomarkers being tested. This is particularly relevant as we push to have more validation studies and multi-site studies carried out in different settings where factors like prevalence of TB disease may differ.

The complexities of identifying biomarkers that are both highly specific and sensitive for *Mtb* or active TB, irrespective of different factors that affect the host immune response, may require us to do away with the notion of a “one size fits all” biomarker (Gutiérrez-González et al., 2021; Hamada et al., 2022). In light of the above, transcriptomic signatures and any other biomarkers for either diagnosis, triage, or treatment monitoring that fall short of the WHO TPP metrics should not be cast

aside, but rather continuously tested and validated in different patient groups and disease settings to identify the most appropriate scenarios in which they can be used (Abdulgader et al., 2022).

None of the signatures we evaluated met any of the WHO TPP benchmarks when data from all study participants were analysed together. When diagnostic performance (**Chapter 4**) was stratified by country, there were transcriptomic signatures that did meet the TPP in some countries. This would support the application of different test thresholds for distinct patient groups, populations, countries and regions. This is further backed by the finding that country had a significant impact on scores in our diagnostic study. Since HIV status also had a significant impact on signature scores, TPP thresholds that take into consideration comorbidities such as HIV and diabetes, as well as common ORDs, are steps that should be taken to make TPPs more achievable. The application of a combination of tests could increase the likelihood of the optimal TPP criteria being met (Abdulgader et al., 2022). Different combinations of tests would have to be modelled and included within diagnostic algorithms to determine acceptable thresholds or cut-offs. Host-based biomarker tests for example can be added to increase the sensitivity of clinical algorithms as sensitive tools for TB screening at the POC (Basile et al., 2022). Other technologies that are coming up for triage tests include CRP, digital chest X-ray, CAD and the Xpert host response test, which can also be used in combination with confirmatory tests such as Truenat and GeneXpert. While the goal is to also have non-sputum based confirmatory tests soon, the absence of one should not delay the use of potential triage tests that are already present (Abdulgader et al., 2022).

6.3 Technology in place for translation of transcriptomic signatures as point-of-care tests

Transcriptomic signatures have applications in the following areas: triage, diagnosis, predicting progression to active TB and tracking responses to treatment. In our diagnostic study, the aim was to identify transcriptomic signatures with the potential for the development of a triage test. Specificity is crucial for signatures and any other biomarkers aimed at translation as triage tests, to minimize false positives when ruling out who has TB. The WHO minimal threshold for specificity for a triage test was only met by some signatures in three out of five countries that were part of our analyses. Several signatures that we evaluated such as da Costa3, Francisco2, Sweeney3, Sambarey10 and Penn-Nicholson6 (RISK6) have met the minimal thresholds for a triage test as demonstrated in the original publications that report on the discovery of these signatures (da Costa et al., 2015; Francisco et al., 2017; Penn-Nicholson et al., 2019; Sambarey et al., 2017; Sweeney et al., 2016). Translation of these and other transcriptomic signatures will however require them to also meet pricing and operational targets.

POC molecular testing is one avenue with potential for expansion and capacity to deliver rapid and accurate test results. We see evidence of this from the development of more compact and easier-to-use devices such as the GeneXpert host response cartridge, the TB-Host Immune Response Assay (TB-HIRA) and the Cepheid Xpert® TB Fingerstick (Xpert® TB FS).

Efforts to improve the accuracy and interpretation of results from potential POC diagnostic platforms have now involved the incorporation of machine learning and artificial intelligence systems. Examples of these include the Keynome® ID and Keynome® gAST machine learning algorithms, developed by Zero Day Diagnostics. These systems interpret whole genome sequencing data on clinical samples and translate them into interpretable diagnostic results for clinicians.

So far, RT-PCR technologies have been the most widely used platforms for POC molecular diagnostic development ("Global POC Molecular Diagnostics Market Driven by Increasing Demand for Rapid and Accurate Testing," 2023). In the TB field, two tests, based on RT-PCR technology, the GeneXpert by Cepheid and Truenat by Molbio have been endorsed by WHO for TB diagnosis. Since majority of

upcoming POC tests for TB have made use of PCR technology, future biomarker studies should most likely consider the application of PCR techniques when validating biomarkers. This could ease the scaling down and adaptation of any potential biomarkers as POC tests. From the panel of signatures evaluated in this thesis, Sweeney³ and Penn-Nicholson⁶ (RISK6) have both been adapted into potential POC tests. This could partly stem from them being validated using PCR platforms. Although platforms such as RNA-seq and microarrays are better suited for initial biomarker discovery, these are not platforms with capacity to be easily scaled down for development of rapid as well as affordable screening, triage and diagnostic tests for TB.

The development of the GeneXpert HR cartridge is a good example of the translation of transcriptomic signatures into actual TB tests. Although the Xpert HR cartridge has been assessed for both TB diagnosis (Sutherland et al., 2022) and treatment monitoring (Zimmer et al., 2021), there may still be hindrances with its applicability as a true POC test for TB. The high costs of the GeneXpert have been a limiting step to its implementation in national TB guidelines for several highly endemic countries. The increased uptake of GeneXpert MTB/RIF cartridges from 2012 was a result of FIND negotiating concessional prices for 145 high-burden and developing countries. The cost of each cartridge was brought down further from \$18.68 to \$9.98 because of funding from bodies such as United States Agency for International Development and the Bill & Melinda Gates Foundation (Lawn et al., 2013).

Regarding application at the level of primary health care, GeneXpert machines require a stable and constant electricity supply which, in primary health care facilities in most low- and middle-income settings, is not always guaranteed unless a form of backup power is available. This, along with the size of most GeneXpert machines would also limit their portability. Availability of cartridges, spare parts, maintenance, regular servicing and technical support are additional aspects that limit the versatility of GeneXpert in primary healthcare settings and at the POC where access to these services can be limited in several countries. Regulatory approval for GeneXpert machines and maintenance of quality assurance standards for POC use, in diverse settings, can be challenging especially in resource-limited areas with varying regulatory frameworks. The Xpert HR cartridge is not commercially available yet as it will require approval from regulatory bodies such the Food and Drug Administration, European Medicines Agency and the South African Health Products Regulatory Authority.

If Cepheid is not able to address these aforementioned factors that have hindered the full implementation of GeneXpert testing in peripheral healthcare settings, these factors will most likely also impede the wide-scale use of the Xpert HR cartridge as a POC test for TB (Lawn et al., 2013; Nalugwa et al., 2020).

QuantuMDx is also developing a multiplex one-step RT-qPCR assay, called the TB-HIRA. This assay quantifies the expression of transcripts that make up the RISK6 signature (Penn-Nicholson et al., 2020). The TB-HIRA test runs on the QuantuMDx Q-POCTM POC platform that can generate results in less than 30 minutes when testing blood from a fingerstick sample (Flanagan, 2024).

Although still in the early development phase, microfluidic platforms also hold potential for POC diagnostic development (Su et al., 2015; Tay et al., 2016). These integrated microfluidics circuits are referred to as lab-on-a-chip systems (**Figure 6.1**), and some have been developed specifically for TB diagnosis (Hong et al., 2022), such as the VerePLEX Biosystem (Cabibbe et al., 2015) and the Lab-on-a-Film assembly (Kukhtin et al., 2019).

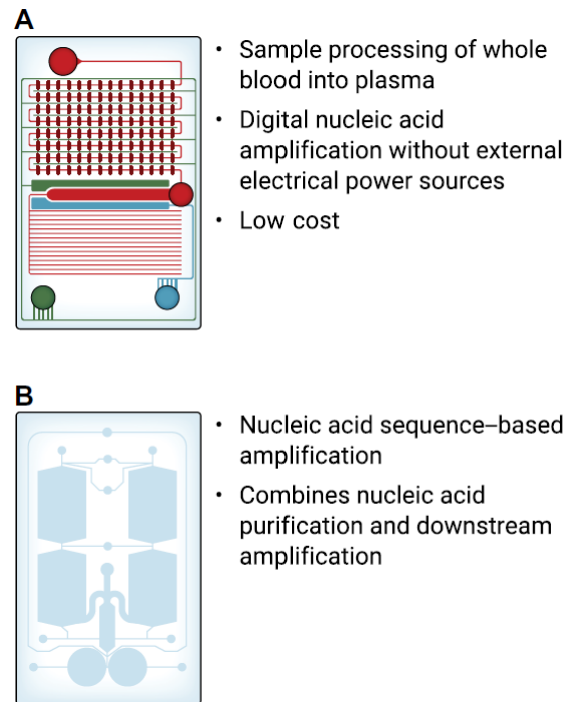


Figure 6.1: Innovations in microfluidic lab-on-a-chip systems for molecular diagnostics.

“(A) The self-powered integrated microfluidic POC low-cost enabling chip is a microfluidic device that processes whole blood samples into plasma and then performs digital nucleic acid amplification without external electrical power sources [adapted from (Yeh et al., 2017)]. (B) The nucleic acid sequence-based amplification device is another integrated microfluidic chip combining nucleic acid purification and downstream amplification [adapted from (Dimov et al., 2008)]”. (Hong et al., 2022)

Ultimately cost effectiveness, operating outside of a laboratory setting, user-friendliness and portability are all factors that will determine the extent to which these and any other tests sustainably become part of TB diagnostic algorithms in countries endemic for TB and LMIC settings.

Tests that can aid active case finding within communities are another goal for potential sputum-free TB diagnostics. Technologies that are in development, however, will have to demonstrate practicality within primary health care centres. This is because testing in communities may present more challenges due to a lower pre-test probability than that at health facilities (Nathavitharana et al., 2022). Our results showed that transcriptomic signatures performed better in symptomatic individuals with definite TB compared to those who had probable TB. These were all symptomatic individuals seeking care at primary healthcare facilities. From a study on biomarker-guided tuberculosis preventive therapy, RISK11 performed better as a screening test for TB in symptomatic individuals recruited from TB-endemic communities, compared to those who were asymptomatic. Altogether, these results demonstrate better performance of transcriptomic signatures in individuals who are not only symptomatic, but also the most ill. This would result from transcriptomic signatures representing higher inflammatory responses in these individuals.

6.4 Addressing the presence of ISGs as part of transcriptomic signatures

A majority of studies on transcriptomic signatures for active TB have provided evidence for the involvement of genes stimulated by type I or II interferons in TB pathogenesis (Berry et al., 2010; Qiu et al., 2022; Scriba et al., 2017; Singhania et al., 2018; Sweeney et al., 2016). Among genes that are differentially expressed during active TB, the magnitude of differential expression is the largest for ISGs. This has led to their inclusion in most transcriptomic signatures to date. This, however, comes with limitations in the performance of these signatures, since they are unable to distinguish active TB

from viral infections. Transcriptomic signatures dominated by ISGs have reduced specificity for active TB, in the presence of viral co-infections. This results from the upregulation of ISGs in response to both viral infections and active TB, leading to increased signature scores in individuals with either one of these conditions (Mendelsohn et al., 2022a; Mulenga et al., 2021; Singhania et al., 2018a). Viral infections among the control group in Chapter 4 would account for high signature scores among individuals with ORDs in some countries. This would be one factor that led to the reduced performance of some signatures in certain countries (**Chapter 4**). This would call for a change in the methodology used to identify host response biomarkers in whole blood, to be able to develop signatures that have a balance between ISGs and transcripts involved in other biological pathways. The application of multinomial modelling, to exclude viral infection as a confounder is another suggestion made to discover transcriptomic signatures that can ably distinguish TB from viral and other infections (Hamada et al., 2022; Mendelsohn et al., 2022b).

Singhania and colleagues note that most gene reduction methodologies concentrate on genes that are most differentially expressed, without accounting for the level of correlation between them (Singhania et al., 2018a). This leads to over-representation of specific pathways, due to a narrowed focus on particular aspects of the host immune response (Singhania et al., 2018a). Capturing a wider scope of the immune response may therefore be necessary. The approaches to accomplish that, without compromising the diagnostic accuracy required of biomarkers for TB diagnostic development, is a matter still yet to be addressed by the TB research community. The inclusion of genes that represent other pathways that are involved in TB pathogenesis may come with the need for a trade-off between a signature being either highly specific or sensitive for active TB (Singhania et al., 2018a).

6.5 Single biomarkers may not be enough

In a bid to support POC diagnostic development, researchers have focussed their efforts on identifying signatures across various omics platforms that are as parsimonious as possible. In the diagnostic study in Chapter 4, we were able to identify and include 18 transcriptomic signatures comprised of less than ten transcripts as part of our evaluation. The existence of more parsimonious signatures represents the strides that have been made to shift away from the initial identification of larger signatures with over 300 transcripts. We also evaluated a signature, Roe1, comprised of only one transcript, BATF2. Various studies, not limited to omics approaches, have assessed single and multiple biomarkers, with some researchers cautioning against the reliance on single biomarkers such as CRP, or on responses to IFN- γ only. Although single markers like BATF2 would be easier to translate into a POC test, more studies would have to generate evidence showing its superior performance as triage or screening tool for TB. In the field of proteomics, even though CRP tests exist, efforts have still been made to develop multi-biomarker strip tests which detect not only CRP but also SAA, IP-10, Ferritin, ApoA-I and IL-6 (Corstjens et al., 2019). In preparation for field testing within POC settings, this proteomic signature has undergone further refinement to include just three biomarkers: CRP, SAA and IP-10 (Richardson et al., 2023).

From studies that have evaluated both single and multiple transcripts, a signature with at least three transcripts may be more sufficient for POC test development. This would suggest that the reliance on biological pathways, represented by a single biomarker, may not be enough for any POC diagnostic test or for treatment monitoring purposes.

6.6 Bottlenecks with studies on non-sputum based biomarkers for TB diagnosis and treatment monitoring in children

From the paediatric study in Chapter 5, there were differences in signature scores between participants who had different categories of TB. This arises from differences in the level of host immune inflammatory responses to the presence of *Mtb*. Another study evaluating the diagnostic performance of transcriptomic signatures in children reported higher sensitivity and specificity for distinguishing children who were culture-positive from those with other diseases, in comparison to children who were culture-negative. (Anderson et al., 2014). These results may warrant the identification of biomarkers that are specific for children within each of the three TB categories. Despite the differences present among children with different TB classifications, a larger part of paediatric studies, in literature, have not considered this when identifying biomarkers either for TB diagnosis or treatment monitoring.

A review of the role of transcriptomics in enhancing our understanding of paediatric TB highlighted the existence of transcriptomic signatures discovered and validated in paediatric cohorts for either TB diagnosis or tracking treatment responses (Kaforou et al., 2022). Although most studies included an appropriate control group to compare gene expression changes to, the TB cases included in the evaluations were not all-encompassing or did not clearly define the TB categories for children included as cases. Verhagen and colleagues (Verhagen et al., 2013) for example did not specify whether their cases had confirmed TB, unconfirmed TB, unlikely TB, culture-negative TB, culture-positive TB or definite and probable TB. Without a clear breakdown of TB cases or proper representation of children within different TB categories, limitations arise regarding the applicability of the signatures identified to other studies and cohorts.

In a systematic review of biomarkers for TB diagnosis in children, all biomarkers identified, apart from LAM, were host response biomarkers (Togun et al., 2018). Of these, cytokines and transcriptomic signatures made up majority of the biomarkers identified. When performance was benchmarked against WHO TPPs for diagnostic and triage tests, the following cytokines IP-10, IL-2 and IL-13 (Armand et al., 2014) and a micro-RNA signature (Zhou et al., 2016) met the minimal stipulated thresholds (Togun et al., 2018). In the study by Armand and colleagues (Armand et al., 2014), only children with confirmed or probable TB were included. The exclusion of children with possible and unlikely TB leaves out many children, most likely those who make up a larger part of paediatric TB cases. The study by Zhou and colleagues (Zhou et al., 2016) only included children with culture-confirmed TB. Although these two studies met the WHO TPPs for diagnostic and triage tests, the usefulness of the biomarkers in unconfirmed and unlikely TB cases cannot be ascertained. It is highly unlikely also, that these biomarkers would have met the WHO TPPs if these studies also included children with unconfirmed and unlikely TB. The application of different cut-offs for both diagnostic and treatment monitoring biomarkers in children, per TB category, may be a more feasible approach. Owing to the differences in biology and the host immune response to TB, there is a high likelihood that paediatric biomarkers for treatment response will fall short of the current joint WHO TPP benchmarks. Pediatric TB is typically mild, which is why most of the children in the SHINE trial were in the unconfirmed TB group. This group has lower inflammatory responses, hence there is less signal for transcriptomic signatures.

Aside from those identified in this systematic review (Togun et al., 2018), other potential tools and biomarkers for either triage or diagnosis of paediatric TB include: CAD (Tornheim et al., 2022; Vo et al., 2021), AI-enabled POC lung ultrasound (Bélard et al., 2018), circulating free DNA (Click et al., 2018), breath metabolites (Bobak et al., 2021; Mgone et al., 2018; Quesada et al., 2018), saliva (Khambati et al., 2021), and a T-cell activation assay (Hiza et al., 2021; Mpande et al., 2021; Musvosvi

et al., 2023, 2018; Portevin et al., 2014). These biomarkers have taken the step forward to address the reliance on sputum, however, to see broader utility for paediatric TB, evidence will have to be generated from studies that include children across all three TB categories.

From a systematic review of treatment monitoring biomarkers for active pulmonary TB in adults, host-response biomarkers accounted for more than half of those identified (Zimmer et al., 2022). These biomarkers fell into the following categories: cytokines, chemokines, blood-based transcriptomic signatures, and breath-based markers. The systematic review highlighted CRP, IL-6, IP-10, and TNF as biomarkers that could be explored further (Zimmer et al., 2022). Studies enrolling children were excluded from this systematic review due to difficulties in establishing a microbiological reference standard (Zimmer et al., 2022). The lack of a reference standard for paediatric TB was also highlighted by other researchers as another limitation towards the evaluation of non-sputum based biomarkers for paediatric TB (Anderson et al., 2014; Togun et al., 2018). The challenges of a microbiological standard arise from most children presenting with paucibacillary TB that is usually not detectable from smear and culture tests.

While CRP has been evaluated as a diagnostic tool for paediatric TB (Albuquerque et al., 2019; Jaganath et al., 2022; Kusumaningrum et al., 2021), to our knowledge, there are no publications on it as a treatment monitoring biomarker for children. While there are also no reports that exist on IL-6 and TNF as treatment response biomarkers for paediatric TB, one study has assessed and reported reduced levels of IP-10 in children treated for TB (Strzelak et al., 2024). Although the categorisation of children with confirmed, unconfirmed and unlikely TB was not explicitly stated in the paper (Strzelak et al., 2024), the authors included children suspected of having TB, with or without microbiological confirmation, as well as asymptomatic TB contacts.

A majority of studies within the literature that aimed to identify non-sputum based TB biomarkers have been carried out mainly in adult populations, with findings sometimes extrapolated to children (Togun et al., 2018). Most studies on non-sputum based biomarkers continue to discover and highlight host response biomarkers. Most, if not all host-response biomarkers essentially track changes in inflammation. Even though some may be non-specific for TB, these biomarkers continue to show promise as non-sputum based alternatives for treatment monitoring. Compared to tests such as microscopy, POC tests based on host immune response biomarkers could have a higher sensitivity for diagnosis and treatment monitoring in children, even for children with culture-negative TB. This is particularly essential for paediatric TB which is usually paucibacillary in nature, contributing to a majority of children being negative for microbiological tests. To ensure that these biomarkers are useful not just for confirmed TB cases, studies will need to include unconfirmed and unlikely TB cases and determine whether the same set of host response markers are useful across all three TB categories. Studies may also have to look at the possibility of identifying non-sputum based biomarkers that reflect more than just the inflammatory host immune response, or those that can track other biological pathways that are upregulated in response to TB disease.

6.7 How to move biomarker studies for paediatric TB forward

Nicol and colleagues (Nicol et al., 2015) outline desired characteristics (**Table 6.3**) and priorities for TB biomarkers in children (**Table 6.4**). From our study on treatment monitoring in children, the rapid and significant decline in scores could be a result of some children being misclassified as having TB, particularly children in the unlikely TB group. Further evidence to support this hypothesis results from the unlikely TB group showing the least change in scores during TB treatment, in comparison to confirmed and unconfirmed TB cases. Some of these children could have had other infections, such as acute viral infections, that were self-resolving. Based on this, biomarkers that can distinguish TB in

children, from other childhood infections should be another priority for paediatric TB biomarker research. This is particularly important for biomarkers aimed at identifying children who have either unconfirmed or unlikely TB disease. This is because of two factors that contribute to the misclassification of children without TB as having unconfirmed or unlikely TB namely: the non-specific symptoms children present with and the current classification of children with unconfirmed and unlikely TB by Graham and colleagues (Graham et al., 2015).

Table 6.3: Desired characteristics of a biomarker for tuberculosis in children

Measurable in more readily available samples, apart from sputum. Examples include blood (e.g., by fingerstick), urine, stool, saliva, buccal mucosal transudates, or exhaled air.
Ability to identify <i>Mtb</i> with high sensitivity and specificity, in children with either pulmonary or extrapulmonary TB, irrespective of age, nutritional status, or HIV status.
Ability to distinguish between children with latent TB and those with active TB disease.
Suitable for incorporation into a rapid, accurate and easy-to-use point-of-care diagnostic platform

Adapted from (Nicol et al., 2015)

Table 6.4: Priorities for paediatric TB biomarker development

Biomarkers for the diagnosis of active TB disease among symptomatic children, with and without HIV
Biomarkers that target young children aged 0-5 years, since they are known to have the greatest disease burden.
Biomarkers focussed on pulmonary TB since it has the highest disease burden and diagnostic challenges.

Adapted from (Nicol et al., 2015)

In addition to the priorities highlighted in **Table 6.4**, by Nicol and colleagues, our results support the need for more focus on children with unconfirmed and unlikely TB. Children with confirmed TB have higher chances of being identified using already available microbiological and sputum-based diagnostic tests. Our results (**Chapter 5**) showed that scores were highest, at baseline, in children with confirmed TB. Transcriptomic signatures, and any other biomarkers that track the host immune response would therefore still be useful for children with confirmed TB.

We also identified significant differences in scores for children between all TB categories at baseline. This demonstrates that there are differences in the levels of inflammatory responses mounted against *Mtb*. If any biomarkers are to be useful for both unconfirmed and unlikely TB cases, researchers will have to take into account the differences in the levels of the inflammatory responses present and ensure that the biomarkers discovered are still sensitive enough to identify children in these TB categories.

Although Nicol and colleagues highlight children below 5 years as a priority focus of paediatric TB biomarkers, biomarkers should be able to identify children with TB at all ages. The major rationale for this is that most children that we identified with confirmed TB were children that were older.

Before we see further translation of promising biomarkers into tests for paediatric TB, several factors will still need to be addressed, regarding both the identification and validation of these biomarkers. Omics-based approaches in particular should include paediatric datasets in their discovery work, so that more paediatric-specific biomarkers can be identified (Basile et al., 2022). Preferably multi-site studies should be done that include children across all TB categories from geographically diverse settings. This would aid the identification of biomarkers that have wider applicability for children

with TB across different age groups and with different co-morbidities. Also, more studies should include paediatric cohorts to validate signatures that have been identified in adult populations (Basile et al., 2022).

The consistent use of recent, updated and universally accepted case definitions for paediatric TB, such as those by Graham and colleagues (Graham et al., 2015), would greatly aid the comparison and interpretation of results across biomarker studies (Togun et al., 2018). Finally, biomarker studies for paediatric TB will need to clearly state the intended use case of the biomarkers under evaluation (Togun et al., 2018). This will facilitate translation of biomarker findings into tests for implementation at the POC.

The current aims for any tests geared towards TB treatment monitoring and optimization, as outlined through proposed WHO TPPs (**Table 6.5**) would be to (i) determine patients at risk of poor treatment outcomes at the start of treatment, (ii) determine patients at risk of poor treatment outcomes during the course of treatment so that they can be switched to longer more effective treatment, and (iii) determine who has not achieved a favourable treatment outcome at the end of treatment. Since study participants who were part of the work in **Chapter 5** mostly had good treatment outcomes, we were not able to adequately assess the transcriptomic signatures evaluated for the use cases highlighted by the WHO TPPs. Based on upcoming results from the SATVI team, studies in adults have shown that transcriptomic signatures can predict treatment failure from the point of treatment initiation. That being said, practical use of signatures for treatment monitoring would require testing at the start, during and at the end of treatment to ensure that adequate treatment optimization occurs. The team of consultants involved in putting these TPPs together describe treatment optimization as starting an effective tuberculosis treatment regimen that results in a greater chance of a patient having a good treatment outcome. The matter of cost-effectiveness is essential as well and with this in mind, testing signatures at a single time point would be the best, however, testing at two selected time points would be better to ensure that comparative data is available for health care providers tasked with making decisions on treatment plans for patients. A greater task now for both researchers and product developers would be to determine which two time points would be the most adequate for testing and ensure that the same test/set of biomarkers can be used.

Table 6.5: WHO TPPs for tests aimed at TB treatment monitoring and optimization

	A test used at treatment initiation	
	Minimal target	Optimal target
Sensitivity	90%	95%
Specificity	70%	80%
	A test used during TB treatment	
Sensitivity	75%	90%
Specificity	80%	90%
	A test used at the end of treatment	
Sensitivity	80%	95%
Specificity	90%	95%

6.8 Future work in the field of TB transcriptomic signatures

Diagnostic efforts should concentrate more on identifying individuals who do not have a typical presentation of TB disease. Transcriptomic signatures that were evaluated in Chapter 4 all performed better at identifying individuals with microbiologically confirmed TB (definite TB). As there are already tests present for that, this is not the most pressing need to address. Probable TB cases and

individuals at earlier stages within the TB disease spectrum (incipient and sub-clinical TB disease) are the people who require more focus in non-sputum-based diagnostic biomarker studies.

Future studies should continue to assess the potential impact that co-morbidities could have on the expression and diagnostic accuracy of transcriptomic signatures. In the past five years, we have already seen an increase in co-morbidity studies for individuals with both TB and type 2 diabetes, driven majorly by results from projects such as the TANDEM consortium. The extent to which the work has been conducted has led to the discovery of both diagnostic and treatment monitoring biomarkers that are better suited for individuals who have both TB and type 2 diabetes. This, in turn, has also helped enhance our understanding of the host immune responses present among individuals with both diseases. We have also been able to learn more about *Mtb*.

Although most children with TB end up falling under the unconfirmed TB category, work in the literature on transcriptomic signatures and other host response biomarkers for paediatric TB has still focussed only on confirmed TB cases. With regards to any inflammatory-driven biomarkers, those that better represent the host immune responses in children with unconfirmed and unlikely TB may be required. This arises from clear differences in the levels of inflammation, mirrored by differences in signature scores for children across all three TB categories (**Chapter 5**).

Although we were not in position to address this from our sub-study in children, the next step with work on treatment response monitoring should also look at the ability for transcriptomic signatures, and any other potential non-sputum based biomarkers, to predict treatment outcomes. The ability to predict and stratify children, from the onset of TB treatment, based on differences in signature scores would greatly aid efforts to shorten and even individualize treatment plans. This would also facilitate children receiving the right care and medication sooner. Since extensive work has already been carried out regarding diagnostic biomarkers for TB in adults, a shift in focus to treatment monitoring and prediction of relapse, TB recurrence, as well as death, should be another priority for transcriptomic studies. This could lead to identification of a new sets of biomarkers, expand on the application of transcriptomic signatures that have already been identified and allow us to learn more about the biology that drives the host immune responses to *Mtb*. These insights will potentially aid biomarker work, and potentially address other gaps in the TB field, such as the evaluation of TB vaccine candidates.

More multistakeholder collaborations, involving academic and industry partners, would increase the likelihood of transcriptomic signatures reaching the point of field testing. Conducting pilot studies on prototype POC devices may give researchers and test developers more insight into the necessary refinements that are required regarding both the technology behind the tests, but also the biology. An example of one successful partnership would be between the inventors of the Sweeney3 signature and Cepheid. The development and deployment of the Xpert HR cartridge led to one transcript within the signature being replaced by another that was more effective for the fingerpick test. Joint efforts from multiple research groups and the Foundation for Innovative New Diagnostics (FIND) have led to initiatives such as the TriageTB study that will support field testing of both the Multiple Biomarker Test (MBT) and the Cepheid Xpert HR cartridge. Partnerships such as these are necessary for generating more evidence on the feasibility of host biomarker-based tests at the POC.

Since omics biomarkers have been identified across multiple platforms, merging data from different techniques would be a great next step in advancing our understanding of host immune responses to TB. The combination of proteomic, metabolomic and transcriptomic biomarkers could also increase the sensitivity and specificity for TB tests aimed at screening, triage or diagnosis. Studies with larger sample sizes are needed that compare different biomarker approaches to allow comparison of

different technologies and types of biomarkers. This should be coupled with novel methods for identifying or quantifying *Mtb* load, to better understand the performance within the context of more definitive classification within the TB spectrum.

References

- Abdulgader, S.M., Okunola, A.O., Ndlangalavu, G., Reeve, B.W.P., Allwood, B.W., Koegelenberg, C.F.N., Warren, R.M., Theron, G., 2022. Diagnosing Tuberculosis: What Do New Technologies Allow Us to (Not) Do? *Respiration* 101, 797–813. <https://doi.org/10.1159/000525142>
- Albuquerque, V.V.S., Kumar, N.P., Fukutani, K.F., Vasconcelos, B., Arriaga, M.B., Silveira-Mattos, P.S., Babu, S., Andrade, B.B., 2019. Plasma levels of C-reactive protein, matrix metalloproteinase-7 and lipopolysaccharide-binding protein distinguish active pulmonary or extrapulmonary tuberculosis from uninfected controls in children. *Cytokine* 123, 154773. <https://doi.org/10.1016/j.cyto.2019.154773>
- Anderson, S.T., Kaforou, M., Brent, A.J., Wright, V.J., Banwell, C.M., Chagaluka, G., Crampin, A.C., Dockrell, H.M., French, N., Hamilton, M.S., Hibberd, M.L., Kern, F., Langford, P.R., Ling, L., Mlotha, R., Ottenhoff, T.H.M., Pienaar, S., Pillay, V., Scott, J.A.G., Twahir, H., Wilkinson, R.J., Coin, L.J., Heyderman, R.S., Levin, M., Eley, B., 2014. Diagnosis of Childhood Tuberculosis and Host RNA Expression in Africa. *N. Engl. J. Med.*
- Armand, M., Chhor, V., Lauzanne, A. de, Khourouj, V.G.-E., Pédrón, B., Jeljeli, M., Gressens, P., Faye, A., Sterkers, G., 2014. Cytokine responses to quantiferon peptides in pediatric tuberculosis: A pilot study. *J. Infect.* 68, 62–70. <https://doi.org/10.1016/j.jinf.2013.08.005>
- Basile, F.W., Nabeta, P., Ruhwald, M., Song, R., 2022. Pediatric Tuberculosis Diagnostics: Present and Future. *J. Pediatr. Infect. Dis. Soc.* 11, S85–S93. <https://doi.org/10.1093/jpids/piac082>
- Bélard, S., Heuvelings, C.C., Banderker, E., Bateman, L., Heller, T., Andronikou, S., Workman, L., Grobusch, M.P., Zar, H.J., 2018. Utility of Point-of-care Ultrasound in Children With Pulmonary Tuberculosis. *Pediatr. Infect. Dis. J.* 37, 637. <https://doi.org/10.1097/INF.0000000000001872>
- Berry, M.P.R., Graham, C.M., McNab, F.W., Xu, Z., Bloch, S.A.A., Oni, T., Wilkinson, K.A., Banchereau, R., Skinner, J., Wilkinson, R.J., Quinn, C., Blankenship, D., Dhawan, R., Cush, J.J., Mejias, A., Ramilo, O., Kon, O.M., Pascual, V., Banchereau, J., Chaussabel, D., O'Garra, A., 2010. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature* 466, 973–977. <https://doi.org/10.1038/nature09247>
- Bobak, C.A., Kang, L., Workman, L., Bateman, L., Khan, M.S., Prins, M., May, L., Franchina, F.A., Baard, C., Nicol, M.P., Zar, H.J., Hill, J.E., 2021. Breath can discriminate tuberculosis from other lower respiratory illness in children. *Sci. Rep.* 11, 2704. <https://doi.org/10.1038/s41598-021-80970-w>
- Cabibbe, A.M., Miotto, P., Moure, R., Alcaide, F., Feuerriegel, S., Pozzi, G., Nikolayevskyy, V., Drobniowski, F., Niemann, S., Reither, K., Cirillo, D.M., TM-REST Consortium, TB-CHILD Consortium, 2015. Lab-on-Chip-Based Platform for Fast Molecular Diagnosis of Multidrug-Resistant Tuberculosis. *J. Clin. Microbiol.* 53, 3876–3880. <https://doi.org/10.1128/jcm.01824-15>
- Click, E.S., Murithi, W., Ouma, G.S., McCarthy, K., Willby, M., Musau, S., Alexander, H., Pevzner, E., Posey, J., Cain, K.P., 2018. Detection of Apparent Cell-free M. tuberculosis DNA from Plasma. *Sci. Rep.* 8, 645. <https://doi.org/10.1038/s41598-017-17683-6>
- Corstjens, P., Hooij, A.V., Fat, E.T.K., Herdigein, S., Namuganga, A.R., Diergaardt, A., Mutavhatsindi, H., Gindeh, A., Mihret, A., Spuy, G.V.D., Gunther, G., Howe, R., Mayanja-Kizza, H., Sutherland, J., Chegou, N.N., Ottenhoff, T., Walzl, G., Geluk, A., 2019. Oc 8435 Multi-Biomarker Test Strip for Point-of-Care Screening for Active Tuberculosis: A Five-Country Multi-Centre Test Evaluation. *BMJ Glob. Health* 4. <https://doi.org/10.1136/bmjgh-2019-EDC.14>
- da Costa, L.L., Delcroix, M., Dalla Costa, E.R., Prestes, I.V., Milano, M., Francis, S.S., Unis, G., Silva, D.R., Riley, L.W., Rossetti, M.L.R., 2015. A real-time PCR signature to discriminate between tuberculosis and other pulmonary diseases. *Tuberc. Edinb. Scotl.* 95, 421–425. <https://doi.org/10.1016/j.tube.2015.04.008>

- Den Boon, S., Lienhardt, C., Zignol, M., Schwartzman, K., Arinaminpathy, N., Campbell, J.R., Nahid, P., Penazzato, M., Menzies, D., Vesga, J.F., Oxlade, O., Churchyard, G., Merle, C.S., Kasaeva, T., Falzon, D., 2022. WHO target product profiles for TB preventive treatment. *Int. J. Tuberc. Lung Dis.* 26, 302–309. <https://doi.org/10.5588/ijtld.21.0667>
- Dimov, I.K., Garcia-Cordero, J.L., O’Grady, J., Poulsen, C.R., Viguier, C., Kent, L., Daly, P., Lincoln, B., Maher, M., O’Kennedy, R., Smith, T.J., Ricco, A.J., Lee, L.P., 2008. Integrated microfluidic tmRNA purification and real-time NASBA device for molecular diagnostics. *Lab. Chip* 8, 2071–2078. <https://doi.org/10.1039/B812515E>
- (FIND), F. for I.N.D., 2023. Target product profiles [WWW Document]. FIND. URL <https://www.finddx.org/tools-and-resources/rd-and-innovation/target-product-profiles/> (accessed 12.29.23).
- Flanagan, M., 2024. The Benefits of Rapid Diagnostics at the Point of Care [WWW Document]. QuantuMDx. URL <https://www.quantumdx.com/blog/the-benefits-of-rapid-diagnostics-at-the-point-of-care/> (accessed 2.19.24).
- Francisco, N.M., Fang, Y.-M., Ding, L., Feng, S., Yang, Y., Wu, M., Jacobs, M., Ryffel, B., Huang, X., 2017. Diagnostic accuracy of a selected signature gene set that discriminates active pulmonary tuberculosis and other pulmonary diseases. *J. Infect.* 75, 499–510. <https://doi.org/10.1016/j.jinf.2017.09.012>
- Global POC Molecular Diagnostics Market Driven by Increasing Demand for Rapid and Accurate Testing [WWW Document], 2023. . Labmedica.com. URL <https://www.labmedica.com/industry-news/articles/294797621/global-poc-molecular-diagnostics-market-driven-by-increasing-demand-for-rapid-and-accurate-testing.html> (accessed 2.19.24).
- Graham, S.M., Cuevas, L.E., Jean-Philippe, P., Browning, R., Casenghi, M., Detjen, A.K., Gnanashanmugam, D., Hesselning, A.C., Kampmann, B., Mandalakas, A., Marais, B.J., Schito, M., Spiegel, H.M.L., Starke, J.R., Worrell, C., Zar, H.J., 2015. Clinical Case Definitions for Classification of Intrathoracic Tuberculosis in Children: An Update. *Clin. Infect. Dis.* 61, S179–S187. <https://doi.org/10.1093/cid/civ581>
- Gupta-Wright, A., den Boon, S., MacLean, E.L., Cirillo, D., Cobelens, F., Gillespie, S.H., Kohli, M., Ruhwald, M., Savic, R., Brigden, G., Gidado, M., Goletti, D., Hanna, D., Hasan, R., Hewison, C., Koura, K.G., Lienhardt, C., Lungu, P., McHugh, T.D., McKenna, L., Scott, C., Scriba, T., Sekaggya-Wiltshire, C., Kasaeva, T., Zignol, M., Denkinger, C.M., Falzon, D., 2023. Target product profiles: tests for tuberculosis treatment monitoring and optimization. *Bull. World Health Organ.* 101, 730–737. <https://doi.org/10.2471/BLT.23.290901>
- Gutiérrez-González, L.H., Juárez, E., Carranza, C., Carreto-Binaghi, L.E., Alejandre, A., Cabello-Gutiérrez, C., Gonzalez, Y., 2021. Immunological Aspects of Diagnosis and Management of Childhood Tuberculosis. *Infect. Drug Resist.* 14, 929–946. <https://doi.org/10.2147/IDR.S295798>
- Hamada, Y., Penn-Nicholson, A., Krishnan, S., Cirillo, D.M., Matteelli, A., Wyss, R., Denkinger, C.M., Rangaka, M.X., Ruhwald, M., Schumacher, S.G., 2022. Are mRNA based transcriptomic signatures ready for diagnosing tuberculosis in the clinic? - A review of evidence and the technological landscape. *eBioMedicine* 82, 104174. <https://doi.org/10.1016/j.ebiom.2022.104174>
- Hiza, H., Hella, J., Arbués, A., Magani, B., Sasamalo, M., Gagneux, S., Reither, K., Portevin, D., 2021. Case-control diagnostic accuracy study of a non-sputum CD38-based TAM-TB test from a single milliliter of blood. *Sci. Rep.* 11, 13190. <https://doi.org/10.1038/s41598-021-92596-z>
- Hong, J.M., Lee, H., Menon, N.V., Lim, C.T., Lee, L.P., Ong, C.W.M., 2022. Point-of-care diagnostic tests for tuberculosis disease. *Sci. Transl. Med.* 14, eabj4124. <https://doi.org/10.1126/scitranslmed.abj4124>
- Jaganath, D., Reza, T.F., Wambi, P., Nakafeero, J., Kiconco, E., Nanyonga, G., Oumo, E.A., Nsereko, M.C., Sekadde, M.P., Nabukenya-Mudiope, M.G., Kato-Maeda, M., Andama, A., Yoon, C.,

- Mohanty, S., Wobudeya, E., Cattamanchi, A., 2022. The Role of C-Reactive Protein as a Triage Tool for Pulmonary Tuberculosis in Children. *J. Pediatr. Infect. Dis. Soc.* 11, 316–321. <https://doi.org/10.1093/jpids/piac015>
- Kaforou, M., Broderick, C., Vito, O., Levin, M., Scriba, T.J., Seddon, J.A., 2022. Transcriptomics for child and adolescent tuberculosis*. *Immunol. Rev.* 309, 97–122. <https://doi.org/10.1111/imr.13116>
- Khambati, N., Olbrich, L., Ellner, J., Salgame, P., Song, R., Bijker, E.M., 2021. Host-Based Biomarkers in Saliva for the Diagnosis of Pulmonary Tuberculosis in Children: A Mini-Review. *Front. Pediatr.* 9.
- Kik, S.V., Denking, C.M., Casenghi, M., Vadnais, C., Pai, M., 2014. Tuberculosis diagnostics: which target product profiles should be prioritised? *Eur. Respir. J.* 44, 537–540. <https://doi.org/10.1183/09031936.00027714>
- Kukhtin, A.C., Sebastian, T., Golova, J., Perov, A., Knickerbocker, C., Linger, Y., Bueno, A., Qu, P., Villanueva, M., Holmberg, R.C., Chandler, D.P., Cooney, C.G., 2019. Lab-on-a-Film disposable for genotyping multidrug-resistant *Mycobacterium tuberculosis* from sputum extracts. *Lab. Chip* 19, 1217–1225. <https://doi.org/10.1039/C8LC01404C>
- Kusumaningrum, R., Anam, M.S., Dadiyanto, D.W., Mexitalia, M., Sidhartani, M., 2021. High sensitivity C-reactive protein level in various manifestations of tuberculosis in children. *Paediatr. Indones.* 61, 253–60. <https://doi.org/10.14238/pi61.5.2021.253-60>
- Lawn, S.D., Mwaba, P., Bates, M., Piatek, A., Alexander, H., Marais, B.J., Cuevas, L.E., McHugh, T.D., Zijenah, L., Kapata, N., Abubakar, I., McNerney, R., Hoelscher, M., Memish, Z.A., Migliori, G.B., Kim, P., Maeurer, M., Schito, M., Zumla, A., 2013. Advances in tuberculosis diagnostics: the Xpert MTB/RIF assay and future prospects for a point-of-care test. *Lancet Infect. Dis.* 13, 349–361. [https://doi.org/10.1016/S1473-3099\(13\)70008-2](https://doi.org/10.1016/S1473-3099(13)70008-2)
- MacLean, E.L.-H., Miotto, P., Angulo, L.G., Chiacchiaretta, M., Walker, T.M., Casenghi, M., Rodrigues, C., Rodwell, T.C., Supply, P., André, E., Kohli, M., Ruhwald, M., Cirillo, D.M., Ismail, N., Zignol, M., 2023. Updating the WHO target product profile for next-generation *Mycobacterium tuberculosis* drug susceptibility testing at peripheral centres. *PLOS Glob. Public Health* 3, e0001754. <https://doi.org/10.1371/journal.pgph.0001754>
- Mendelsohn, S.C., Mbandi, S.K., Fiore-Gartland, A., Penn-Nicholson, A., Musvosvi, M., Mulenga, H., Fisher, M., Hadley, K., Erasmus, M., Nombida, O., Tameris, M., Walzl, G., Naidoo, K., Churchyard, G., Hatherill, M., Scriba, T.J., 2022a. Prospective multicentre head-to-head validation of host blood transcriptomic biomarkers for pulmonary tuberculosis by real-time PCR. *Commun. Med.* 2, 26. <https://doi.org/10.1038/s43856-022-00086-8>
- Mendelsohn, S.C., Mbandi, S.K., Fiore-Gartland, A., Penn-Nicholson, A., Musvosvi, M., Mulenga, H., Fisher, M., Hadley, K., Erasmus, M., Nombida, O., Tameris, M., Walzl, G., Naidoo, K., Churchyard, G., Hatherill, M., Scriba, T.J., 2022b. Prospective multicentre head-to-head validation of host blood transcriptomic biomarkers for pulmonary tuberculosis by real-time PCR. *Commun. Med.* 2, 1–13. <https://doi.org/10.1038/s43856-022-00086-8>
- Mgode, G.F., Cox, C.L., Mwimanzu, S., Mulder, C., 2018. Pediatric tuberculosis detection using trained African giant pouched rats. *Pediatr. Res.* 84, 99–103. <https://doi.org/10.1038/pr.2018.40>
- Mpande, C.A.M., Musvosvi, M., Rozot, V., Mosito, B., Reid, T.D., Schreuder, C., Lloyd, T., Bilek, N., Huang, H., Obermoser, G., Davis, M.M., Ruhwald, M., Hatherill, M., Scriba, T.J., Nemes, E., Mahomed, H., Hanekom, W.A., Kafaar, F., Workman, L., Mulenga, H., Ehrlich, R., Erasmus, M., Abrahams, D., Hawkridge, A., Hughes, E.J., Moyo, S., Gelderbloem, S., Tameris, M., Geldenhuys, H., Hussey, G., 2021. Antigen-Specific T-Cell Activation Distinguishes between Recent and Remote Tuberculosis Infection. *Am. J. Respir. Crit. Care Med.* 203, 1556–1565. <https://doi.org/10.1164/rccm.202007-2686OC>
- Mulenga, H., Musvosvi, M., Mendelsohn, S.C., Penn-Nicholson, A., Kimbung Mbandi, S., Fiore-Gartland, A., Tameris, M., Mabwe, S., Africa, H., Bilek, N., Kafaar, F., Khader, S.A., Carstens, B., Hadley, K., Hikuam, C., Erasmus, M., Jaxa, L., Raphela, R., Nombida, O., Kaskar, M., Nicol,

- M.P., Mbhele, S., Van Heerden, J., Innes, C., Brumskine, W., Hiemstra, A., Malherbe, S.T., Hassan-Moosa, R., Walzl, G., Naidoo, K., Churchyard, G., Hatherill, M., Scriba, T.J., 2021. Longitudinal Dynamics of a Blood Transcriptomic Signature of Tuberculosis. *Am. J. Respir. Crit. Care Med.* 204, 1463–1472. <https://doi.org/10.1164/rccm.202103-0548OC>
- Musvosvi, M., Duffy, D., Filander, E., Africa, H., Mabwe, S., Jaxa, L., Bilek, N., Llibre, A., Rouilly, V., Hatherill, M., Albert, M., Scriba, T.J., Nemes, E., 2018. T-cell biomarkers for diagnosis of tuberculosis: candidate evaluation by a simple whole blood assay for clinical translation. *Eur. Respir. J.* 51, 1800153. <https://doi.org/10.1183/13993003.00153-2018>
- Musvosvi, M., Huang, H., Wang, C., Xia, Q., Rozot, V., Krishnan, A., Acs, P., Cheruku, A., Obermoser, G., Leslie, A., Behar, S.M., Hanekom, W.A., Bilek, N., Fisher, M., Kaufmann, S.H.E., Walzl, G., Hatherill, M., Davis, M.M., Scriba, T.J., 2023. T cell receptor repertoires associated with control and disease progression following *Mycobacterium tuberculosis* infection. *Nat. Med.* 29, 258–269. <https://doi.org/10.1038/s41591-022-02110-9>
- Nalugwa, T., Shete, P.B., Nantale, M., Farr, K., Ojok, C., Ochom, E., Mugabe, F., Joloba, M., Dowdy, D.W., Moore, D.A.J., Davis, J.L., Cattamanchi, A., Katamba, A., 2020. Challenges with scale-up of GeneXpert MTB/RIF® in Uganda: a health systems perspective. *BMC Health Serv. Res.* 20, 162. <https://doi.org/10.1186/s12913-020-4997-x>
- Nicol, M.P., Gnanashanmugam, D., Browning, R., Click, E.S., Cuevas, L.E., Detjen, A., Graham, S.M., Levin, M., Makhene, M., Nahid, P., Perez-Velez, C.M., Reither, K., Song, R., Spiegel, H.M.L., Worrell, C., Zar, H.J., Walzl, G., 2015. A Blueprint to Address Research Gaps in the Development of Biomarkers for Pediatric Tuberculosis. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* 61, S164–S172. <https://doi.org/10.1093/cid/civ613>
- Penn-Nicholson, A., Mbandi, S.K., Thompson, E., Mendelsohn, S.C., Suliman, S., Chegou, N.N., Malherbe, S.T., Darboe, F., Erasmus, M., Hanekom, W.A., Bilek, N., Fisher, M., Kaufmann, S.H.E., Winter, J., Murphy, M., Wood, R., Morrow, C., Van Rhijn, I., Moody, B., Murray, M., Andrade, B.B., Sterling, T.R., Sutherland, J., Naidoo, K., Padayatchi, N., Walzl, G., Hatherill, M., Zak, D., Scriba, T.J., 2020. RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response. *Sci. Rep.* 10, 8629. <https://doi.org/10.1038/s41598-020-65043-8>
- Portevin, D., Moukambi, F., Clowes, P., Bauer, A., Chachage, M., Ntinginya, N.E., Mfinanga, E., Said, K., Haraka, F., Rachow, A., Saathoff, E., Mpina, M., Jugheli, L., Lwilla, F., Marais, B.J., Hoelscher, M., Daubenberger, C., Reither, K., Geldmacher, C., 2014. Assessment of the novel T-cell activation marker–tuberculosis assay for diagnosis of active tuberculosis in children: a prospective proof-of-concept study. *Lancet Infect. Dis.* 14, 931–938. [https://doi.org/10.1016/S1473-3099\(14\)70884-9](https://doi.org/10.1016/S1473-3099(14)70884-9)
- Qiu, Q., Peng, A., Zhao, Y., Liu, D., Liu, C., Qiu, S., Xu, J., Cheng, H., Xiong, W., Chen, Y., 2022. Diagnosis of pulmonary tuberculosis via identification of core genes and pathways utilizing blood transcriptional signatures: a multicohort analysis. *Respir. Res.* 23, 125. <https://doi.org/10.1186/s12931-022-02035-4>
- Quesada, L., Frías, S.P.D., Toro, G.D., Salas, E., Gutierrez, A., Bello, J., Tovar, I., España, M., Perez, H., Villalba, J., Villalon, M., Waard, J.D., 2018. Late Breaking Abstract - An Electronic Nose as a screening tool for childhood Tuberculosis. *Eur. Respir. J.* 52. <https://doi.org/10.1183/13993003.congress-2018.PA4757>
- Richardson, T.R., Smith, B., Malherbe, S.T., Shaw, J.A., Noor, F., MacDonald, C., van der Spuy, G.D., Stanley, K., Carstens, A., Fisher, T.-L., van Rensburg, I., Flinn, M., Snyders, C., Johnson, I., Fransman, B., Dockrell, H., Thwaites, G., Thuong, N.T.T., Schacht, C., Mayanja-Kizza, H., Nsereko, M., Tjon Kon Fat, E.M., Corstjens, P.L.A.M., Geluk, A., Ruhwald, Morton, Penn-Nicholson, A., Chegou, N.N., Sutherland, J., Walzl, G., Hiemstra, A., Tonsing, S., Tromp, G., Owolabi, M., Mendy, J., Gindeh, A., Barry, A., Mbayo, G., Buech, J., Streitz, M., Nalukwago, S., Namuganga, A.R., Lamunu, D., Odie, M., Pierneef, L., van Hooij, A., Ruhwald, Morten, Belisle, J., Dobos, K., Hatherill, M., Scriba, T., Winter, J., TrENDx consortium, 2023. Field

- evaluation of a point-of-care triage test for active tuberculosis (TriageTB). *BMC Infect. Dis.* 23, 447. <https://doi.org/10.1186/s12879-023-08342-5>
- Sambarey, A., Devaprasad, A., Mohan, A., Ahmed, A., Nayak, S., Swaminathan, S., D'Souza, G., Jesuraj, A., Dhar, C., Babu, S., Vyakarnam, A., Chandra, N., 2017. Unbiased Identification of Blood-based Biomarkers for Pulmonary Tuberculosis by Modeling and Mining Molecular Interaction Networks. *eBioMedicine* 15, 112–126. <https://doi.org/10.1016/j.ebiom.2016.12.009>
- Scriba, T.J., Penn-Nicholson, A., Shankar, S., Hraha, T., Thompson, E.G., Sterling, D., Nemes, E., Darboe, F., Suliman, S., Amon, L.M., Mahomed, H., Erasmus, M., Whatney, W., Johnson, J.L., Boom, W.H., Hatherill, M., Valvo, J., Groote, M.A.D., Ochsner, U.A., Aderem, A., Hanekom, W.A., Zak, D.E., Team, other members of the A. cohort study, 2017. Sequential inflammatory processes define human progression from *M. tuberculosis* infection to tuberculosis disease. *PLOS Pathog.* 13, e1006687. <https://doi.org/10.1371/journal.ppat.1006687>
- Singhania, A., Wilkinson, R.J., Rodrigue, M., Haldar, P., O'Garra, A., 2018. Transcriptomics in TB: the immune response and diagnosis. *Nat. Immunol.* 19, 1159–1168. <https://doi.org/10.1038/s41590-018-0225-9>
- Strzelak, A., Komorowska-Piotrowska, A., Borowa, A., Krasińska, M., Feleszko, W., Kulus, M., 2024. IP-10 for the Diagnosis and Treatment Monitoring of Tuberculosis in Children. *Diagnostics* 14, 177. <https://doi.org/10.3390/diagnostics14020177>
- Su, W., Gao, X., Jiang, L., Qin, J., 2015. Microfluidic platform towards point-of-care diagnostics in infectious diseases. *J. Chromatogr. A* 1377, 13–26. <https://doi.org/10.1016/j.chroma.2014.12.041>
- Sutherland, J.S., van der Spuy, G., Gindeh, A., Thuong, N.T.T., Namuganga, A., Owolabi, O., Mayanja-Kizza, H., Nsereko, M., Thwaites, G., Winter, J., Dockrell, H.M., Scriba, T.J., Geluk, A., Corstjens, P., Stanley, K., Richardson, T., Shaw, J.A., Smith, B., Malherbe, S.T., Walzl, G., TrENDx-TB Consortium, Sutherland, J., Owolabi, O., Secka, A., Daffeh, G., Gindeh, A., Mendy, J., Sarr, B., Riley, A.-J., Jobe, A., Davies, M., Kanyi, K., Jallow, M., Barry, S., Cham, O., Nkereuwem, E., Walzl, G., Malherbe, S., Smith, B., van der Spuy, G., Stanley, K., Shaw, J., Chetram, A., Richardson, T., Finn, M., Hiemstra, A., Chegou, N., Kuivaniemi, H., Tromp, G., Tonsing, S., Smit, E., Carstens, B., Mayanja-Kizza, H., Nsereko, M., Namuganga, A., Nalukwago, S., Akol, J., Lamunu, D., Ordie, M., Thwaites, G., Nguyen, T., Le, V., Vo Thanh, S., Thi, H.N., Thi Ngoc, H.V., Le Hong, N., Belisle, J., Dobos, K., Dockrell, H., Scriba, T., Hatherill, M., Hadley, K., Shenje, J., Kimbung, S., Mulenga, H., Oelofse, R., Bilek, N., van Rooyen, E., Mabwe, S., Corstjens, P., Geluk, A., Kon Fat, E.T., Pierneef, L., van Hooij, A., Winter, J., Ruhwald, M., Moreau, E., Penn-Nicholson, A., Schacht, C., Büech, J., Streitz, M., 2022. Diagnostic Accuracy of the Cepheid 3-gene Host Response Fingerstick Blood Test in a Prospective, Multi-site Study: Interim Results. *Clin. Infect. Dis.* 74, 2136–2141. <https://doi.org/10.1093/cid/ciab839>
- Sweeney, T.E., Braviak, L., Tato, C.M., Khatri, P., 2016. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. *Lancet Respir. Med.* 4, 213–224. [https://doi.org/10.1016/S2213-2600\(16\)00048-5](https://doi.org/10.1016/S2213-2600(16)00048-5)
- Tay, A., Pavesi, A., Yazdi, S.R., Lim, C.T., Warkiani, M.E., 2016. Advances in microfluidics in combating infectious diseases. *Biotechnol. Adv.* 34, 404–421. <https://doi.org/10.1016/j.biotechadv.2016.02.002>
- Togun, T.O., MacLean, E., Kampmann, B., Pai, M., 2018. Biomarkers for diagnosis of childhood tuberculosis: A systematic review. *PLoS ONE* 13, e0204029. <https://doi.org/10.1371/journal.pone.0204029>
- Tornheim, J.A., Paradkar, M., Zhao, H., Kulkarni, V., Pradhan, N., Kinikar, A., Kagal, A., Gupte, N., Mave, V., Gupta, A., Karakousis, P.C., 2022. The Kynurenine/Tryptophan Ratio Is a Sensitive

- Biomarker for the Diagnosis of Pediatric Tuberculosis Among Indian Children. *Front. Immunol.* 12.
- Verhagen, L.M., Zomer, A., Maes, M., Villalba, J.A., del Nogal, B., Eleveld, M., van Hijum, S.A., de Waard, J.H., Hermans, P.W., 2013. A predictive signature gene set for discriminating active from latent tuberculosis in Warao Amerindian children. *BMC Genomics* 14, 74. <https://doi.org/10.1186/1471-2164-14-74>
- Vo, L.N.Q., Codlin, A., Ngo, T.D., Dao, T.P., Dong, T.T.T., Mo, H.T.L., Forse, R., Nguyen, T.T., Cung, C.V., Nguyen, H.B., Nguyen, N.V., Nguyen, V.V., Tran, N.T., Nguyen, G.H., Qin, Z.Z., Creswell, J., 2021. Early Evaluation of an Ultra-Portable X-ray System for Tuberculosis Active Case Finding. *Trop. Med. Infect. Dis.* 6, 163. <https://doi.org/10.3390/tropicalmed6030163>
- World Health Organization, 2014. High priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting, 28-29 April 2014, Geneva, Switzerland. World Health Organization, Geneva.
- Yeh, E.-C., Fu, C.-C., Hu, L., Thakur, R., Feng, J., Lee, L.P., 2017. Self-powered integrated microfluidic point-of-care low-cost enabling (SIMPLE) chip. *Sci. Adv.* 3, e1501645. <https://doi.org/10.1126/sciadv.1501645>
- Zhou, M., Yu, G., Yang, X., Zhu, C., Zhang, Z., Zhan, X., 2016. Circulating microRNAs as biomarkers for the early diagnosis of childhood tuberculosis infection. *Mol. Med. Rep.* 13, 4620–4626. <https://doi.org/10.3892/mmr.2016.5097>
- Zimmer, A.J., Lainati, F., Aguilera Vasquez, N., Chedid, C., McGrath, S., Benedetti, A., MacLean, E., Ruhwald, M., Denking, C.M., Kohli, M., 2022. Biomarkers That Correlate with Active Pulmonary Tuberculosis Treatment Response: a Systematic Review and Meta-analysis. *J. Clin. Microbiol.* 60, e01859-21. <https://doi.org/10.1128/jcm.01859-21>
- Zimmer, A.J., Schumacher, S.G., Södersten, E., Mantsoki, A., Wyss, R., Persing, D.H., Banderby, S., Strömqvist Meuzelaar, L., Prieto, J., Gnanashanmugam, D., Khatri, P., Ongarello, S., Ruhwald, M., Denking, C.M., 2021. A novel blood-based assay for treatment monitoring of tuberculosis. *BMC Res. Notes* 14, 247. <https://doi.org/10.1186/s13104-021-05663-z>