Transcriptomic signatures of recurrent tuberculosis disease and treatment response in HIV-infected individuals

Fatoumatta Darboe

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Pathology

Faculty of Health Sciences

UNIVERSITY OF CAPE TOWN

On the 8th of March 2018

Supervisor: Associate Professor Thomas J. Scriba

Co-supervisor: Dr Adam Penn-Nicholson

Declaration

I, Fatoumatta Darboe, 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.

Signature: Signed by candidate

Date: 8th March 2018

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 PhD thesis, my co-authors have agreed that I may include the publication:

Darboe, Mbandi *et al.,* "Diagnostic performance of an optimized transcriptomic signature of risk in tuberculosis in cryopreserved peripheral blood mononuclear cells", Tuberculosis, 2018, Volume 108, pages 124-126.

Signature: Signed by candidate Date: 8th March 2018

Student Name: Fatoumatta Darboe Student Number: 1535104

Acknowledgements

The past three years have been the most difficult, yet amazing journey I have experienced. It was filled with personal growth, astounding academic experience and self-fulfilment. A lot of people contributed immmensely to this journey and I cannot thank them enough for all the support, encouragement and advice through this PhD. You all made the journey easier.

I would first like to thank my mum. You have been my inspiration, made me believe in myself, encouraged me all the times I wanted to give up. To my siblings, Mustapha and Isatou, thanks so much for all the encouragement and laughter meant to cheer me up. To two other brothers, Bubacarr and Ebrima thanks for always checking on me and making sure I was ok.

To my husband, Dawda Njie, thank you so much for everything. You provided a listening ear to my endless complaints and comfort from afar.

To my supervisor Thomas Scriba, thanks for agreeing to talk to me at that Keystone conference and inviting me to apply for a PhD at SATVI. There is no better lab I could have done this PhD. Thank you for the mentorship and guidance and being enthusiastic about results especially when I was not. Thank you to Adam Penn-Nicholson for your supervision. I am prepared for life as a scientist because of you two.

To the MRC-SHIP team, thank you for your guidance and mentorship during this period. Special shout-outs to, Mbandi Kimbung, Michelle Fisher, Ethan Thompson,

Daniel Zak, Nonhlanhla Y. Zuma, Lara Lewis, Nesri Padayatchi and Kogie Naidoo. This PhD would not have been possible without your collaborative efforts.

To Sara Suliman, thank you for introducing me to Tom and for your guidance and mentorship. To Munyaradzi Musvosvi thanks for all the advice, correction of R scripts, advice and early morning squash games.

To the other students, thank you so much for the friendship and lunch time conversations. Special thanks to Simon Mendelson and Richard Baguma for reviewing and proof reading this thesis. To my sister from another mother, Jor, thanks so much for going through this thesis and your invaluable support during my PhD.

To Anna Coussens, Hanif Esmail and Robert Wilkinson, thank you for the valuable samples from your cohort.

I would like to thank all the study participants from all the cohorts used in this thesis. A special thanks to the study teams from SATVI, CAPRISA and CIDRI for participant recruitment and sample processing.

Lastly, thanks to the SAMRC for funding this study, and to MMEG for funding my last year of studies.

Summary

HIV-infected persons are at particularly high risk of tuberculosis (TB) disease, especially in TB endemic countries where *M. tuberculosis* transmission is common. Although antiretroviral therapy (ART) reduces risk of TB, it does not return to that of HIV-uninfected persons. In addition, a previous history of TB disease significantly increases the risk of recurrent TB disease. Identification of HIV-infected individuals at greatest risk of recurrent TB for highly targeted therapy before disease manifestation would be a major advance in the fight against TB. This would also allow the provision of TB treatment in persons with subclinical TB.

Diagnosis of TB in HIV-infected persons is markedly undermined by the paucibacillary nature of HIV-associated disease. A non-sputum based diagnostic test that is highly sensitive and specific, such as a blood-based RNA signature, would be an important new tool. Such a test may also facilitate monitoring of TB treatment, opening the possibility for customizing the duration of TB treatment to that necessary for cure.

We previously discovered and validated a 16-gene transcriptomic signature with promising prognostic and diagnostic utility for TB in HIV-uninfected persons. The transcriptomic signature could predict progression to active TB disease up to a year before diagnosis and was shown to be a useful tool for TB treatment response monitoring in a treatment cohort of HIV-uninfected persons. The signature was reduced to an 11-gene signature to improve throughput with equivalent prognostic and diagnostic performance. In addition, we developed a smaller, 6-gene signature in preparation for translation to point-of-care testing. In this thesis, we aimed to determine (1) the diagnostic performance of these two transcriptomic signatures in

HIV-infected persons, (2) whether they could predict recurrent TB disease in HIV-infected persons, and (3) their utility to monitor TB treatment response in HIV-infected persons.

To assess aim 1, we designed a cross-sectional study of HIV-infected (n=40) and uninfected persons (n=60), each comprising equal numbers of active TB cases and QuantiFERON-positive controls. To assess aims 2 and 3, we designed retrospective substudies among participants enrolled into two clinical studies previously completed by our collaborators at CAPRISA, namely the TRuTH and IMPRESS studies. In the TRuTH cohort participants who developed recurrent TB diagnosis, diagnosed by microbiological testing of induced sputum, were assigned as progressors (n=43), while those who remained asymptomatic were assigned as non-progressors (n=86). In the IMPRESS cohort, participants with a new diagnosis of recurrent TB who initiated TB treatment were stratified into early (n=44) and late (n=19) converters based on time to sputum culture conversion from diagnosis. RNA was isolated from cryopreserved PBMC or PAXgene whole blood and gene expression measured by microfluidic qRT-PCR. Signature scores were generated using in-house customised scripts in R and performance of the signatures was measured using receiver operating characteristic area under the curve (ROC AUC), calculated using the pROC and verification packages in R.

The 11-gene and 6-gene signatures could diagnose active TB disease in HIV-infected persons with good accuracy (AUC = 0.83 and 0.92, respectively), although performances were lower than those observed in HIV-uninfected persons (AUC = 0.97 and 0.96). Signature performance was decreased in HIV-infected persons due

to higher signature scores, reflecting high expression of IFN-stimulated genes, especially in HIV-infected controls. In the TRuTH cohort, these signatures could identify those with recurrent TB within 3 months of diagnosis (AUC = 0.77, p = 0.003), suggesting detection of subclinical disease. Scores of both signatures decreased during TB treatment in the IMPRESS cohort, in participants with early or late sputum conversion. Importantly, two months after initiating TB treatment, the ACS 11-gene signature could differentiate early from late converters. Detectable plasma viral load was associated with higher signature scores in both cohorts, leading to a decrease in signature specificities.

We show that the 11-gene and 6-gene signatures performed well as blood-based diagnostic tests for active TB disease in HIV-infected persons. The signatures could detect recurrent TB disease during the subclinical phase of disease progression and demonstrated promise as treatment response markers in HIV-infected persons. The signatures performed best in persons with effectively suppressed HIV load, highlighting the importance of ART adherence and integration of HIV and TB care for effective clinical management of TB.

List of abbreviations

ACS	Adolescent cohort study
AIDS	Acquired immune deficiency syndrome
ART	Antiretroviral therapy
BMI	Body mass index
CAPRISA	Centre for the AIDS Programme of Research in South Africa
cDNA	Complimentary deoxyribonucleic acid
CI	Confidence Intervals
CIDR	Centre for Infectious Disease Research
CoR	Correlates of risk
CORTIS	Correlates of Risk Targetted Intervention Studies
CORTIS-HR	Correlates of Risk Targetted Intervention Studies High Risk
csv	comma separated values
Ct	cycle threshold
CTRC	Catalysis Foundation for Health Treatment Cohort
DC	Dendritic cells
dCt	delta Cycle threshold
DNA	Deoxyribonucliec acid
FDG	fluorodeoxyglucose
GC6-74	Grand challenges 6-74
GE	Gene Expression
GOI	Gene of interest
HAART	Highly active antiretroviral therapy
HIV	Human immunodeficiencyvirus
HR	Isoniazid, rifampicin

HR	Isoniazid, rifampicin
HRM	Isoniazid, rifampicin, moxifloxacin
HRZE	Isoniazid, rifampicin, pyrazinamide, ethambuthol
HRZM	Isoniazid, rifampicin, pyrazinamide, moxifloxacin
IFN	Interferon
IFNAR	Interferon alpha receptor
IGRA	Interferon gamma release assays
IMPRESS	Improving retreatment success
INH	Isoniazid
IPC	Internal positive control
IPT	Isoniazid preventive therapy
IQR	Interquartile range
IRF	Interferon-regulatory factor
IRIS	Immune reconstitution inflammatory syndrome
ISG	Interferon stimulated genes
JAK	Janus Kinase
K-RITH	KwaZulu-Natal research institute for TB-HIV
MDR	Multi-drug resistant
MGIT	mycobacterial growth indicator tube
mRNA	Messenger ribonucleic acid
Mtb	Mycobacterium tuberculosis
NAAT	Nucleic acid amplification tests
NHP	Non human primates
PBMC	Pheripheral blood mononuclear cells
PCR	Polymerase chain reaction

pDCs	Plasmacytoid dendritic cells
PET-CT	Positron emission tomography-computed tomography
pVL	Plasma viral loads
Pys	Person years
QC	Quality control
QFT	QuantiFERON
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
RFLP	Restrction fragment length polymorphism
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
ROC AUC	Area under the receiver operating characteristic curve
rPSVM.1	re-parameterised pair-wise support vector machine
RT-PCR	Reverse transcriptase polymerase chain reaction
SAPiT	Strating AIDS treatment at three time Points in TB treatment
SATVI	South African Tuberculosis Vaccine Research Initiative
SD	Standard deviation
SHIP	Strategic health innovations partnership
START	Starting Tuberculosis and Anti-Retroviral therapy
STAT	Signal Transducer and Activator of Transcription
SVM	Support vector machines
TB	Tuberculosis
TLR	Toll like receptors
TPP	Target product profile
TRuTH	TB recurrence upon treatment with highly active ART
TST	Tuberculin skin test

TYK	Tyrosine Kinase
WHO	World Health Organization
XDR	Extensively drug-resistant

List of Figures

List of Tables

List of Equations

Table of Contents

Chapter 1: Introduction and literature review

Tuberculosis (TB) is a grave global emergency. In 2015, approximately three times more people died from TB every week (~30,000) (World Health Organization 2017) than the deaths attributable to the Ebola outbreak in West Africa between 2014 and 2016 (~11, 000). Hence, TB remains the main cause of mortality due to an infectious disease and the ninth leading cause of all deaths, worldwide (World Health Organization 2017). Yet the global response to TB is inadequate, under resourced and lacks the appropriate urgency to combat the pandemic. Two of the specific aims of the End TB Strategy are an 80% reduction in incidence (new cases) and a 90% decrease in deaths due to tuberculosis (TB) by 2030 relative to incidence and deaths in 2015. In 2016, an increase of 0.2 million incident cases from 2015 was reported by National TB programmes to the World Health Organization (WHO) (World Health Organization 2017). This suggests that current tools are inadequate to prevent new cases, therefore more tools and strategies are needed to decrease incidence and prevent mortality due to TB disease.

Mycobacterium tuberculosis (Mtb), the causative agent of TB, is transmitted via aerosol in 2 to 5μm microdroplet particles that contain the bacterium from lungs of individuals with active disease (Etna et al., 2014). The most common form of disease that can develop after inhalation of particles and infection with *Mtb* is pulmonary TB. In addition to pulmonary TB, disease can occur in other parts of the body and is then referred to as extra-pulmonary TB or disseminated TB. Infection with the bacterium leads to three possible outcomes: 1) about 5% of individuals rapidly progress to TB

disease with microbiological confirmation in the first two years after infection; 2) ~90% of individuals are able to mount an effective immune response limiting proliferation of the bacilli, usually remaining healthy during their lifetime (known as latent infection, referred to as "*Mtb* infection" in this thesis); 3) in a proportion of *Mtb*-infected persons who did not rapidly progress, infection can be reactivated leading to disease (~1% to 2% per year) (Young, Gideon, and Wilkinson 2009; Wiker et al. 2010). The risk of developing TB disease in contacts of active TB patients is higher in low and middle-income countries than in high-income countries (Fox et al., 2013). Close contacts of TB cases are at their greatest risk of developing TB disease in the first two years after diagnosis of the index case (Fox et al. 2013; Wiker et al. 2010). However, they remain at risk of developing active TB disease for a lifetime with risk decreasing farther away from exposure. Immune suppression, aging, co-infections and repeated TB exposures may increase this risk of developing active TB disease.

1.1 From *Mtb* infection to active TB disease in the absence of immune suppression

Clinically, TB is classified into binary categories, active and latent TB (Barry et al., 2009; Esmail et al., 2014; Pai et al., 2016). However, it has recently become clear that in reality TB exists in a highly heterogeneous spectrum of states, including sterilized prior infection, quiescent or controlled infection and a range of asymptomatic phases of disease wherein patients do not present with clinical disease but have evidence of replicating bacteria, commonly referred to as incipient, subclinical or percolating TB (Pai et al. 2016). The different stages in this spectrum are associated with broad ranges of bacterial load and progression through these stages are dependent on the ability of the immune system to contain bacterial

replication (Young, Gideon, and Wilkinson 2009). Persons who progress from such asymptomatic phases to ultimately develop active disease are thought to transition through a phase of percolating, subclinical disease with active bacterial replication prior to development of symptoms that typically trigger care-seeking behaviour and TB diagnosis (Pai et al., 2016; Salgame et al., 2015). New tools that allow establishment of the stage into which a *Mtb*-infected person falls could help to identify persons at risk of progression to active TB disease and those with subclinical TB, for appropriate treatment.

1.1.1 *Mtb* infection: Epidemiology

The exact proportion of *Mtb*-infected persons is difficult to identify because individuals with *Mtb* infection are generally asymptomatic and the bacterium cannot be directly detected. Consequently, the magnitude of infection is estimated via mathematical models based on country-specific annual risk of infection, estimated from tuberculin skin test (TST) surveys, and transmission potential of active TB cases (Matteelli et al. 2017). A recent re-estimation of the global burden of *Mtb* infection based on such mathematical modelling suggested that an estimated 1.7 billion (~23% of the global population) individuals are infected with *Mtb* (Houben and Dodd 2016), representing a huge reservoir of probable TB disease. Estimates of prevalent infection vary between regions. At least 20% of persons from Southeast Asia, Western Pacific and African regions are estimated to be *Mtb*-infected contributing to over 80% of global estimates (Houben and Dodd 2016). On the other hand, Eastern-Mediterranean, Europe, and the Americas had an estimated prevalence of less than 17% of the general population. Thus, the reservoir of probable TB disease is highest in Southeast Asian, Western Pacific and African

regions. The risk of reactivation or progression to active TB disease is highest within two to five years after primary infection (Dye et al. 1999; Wiker et al. 2010), affecting 5% to 15% of *Mtb*-infected individuals, with some reactivating decades later (Lillebaek et al. 2002). Persons who reactivate within five years are termed “early” disease progressors, whilst those who reactivate later are termed “late” disease progressors (Lillebaek et al. 2002). Targeted interventions in *Mtb*-infected persons are critically important for the prevention of active TB disease in order to achieve TB elimination (Barry et al. 2009).

1.1.2 Diagnosis of *Mtb* infection

There is currently no “gold standard” for diagnosing *Mtb* infection because detection of the pathogen in asymptomatic persons is not possible. As a consequence infection is diagnosed via the detection of memory T-cell responses using *in vivo* (TST) and *in vitro* (Interferon gamma release assays (IGRAs)) assays. TST measures a delayed-type hypersensitivity reaction to an intradermal administration of a mixture of purified proteins derived from *Mtb* (purified protein derivative). Induration induced as a result of the hypersensitivity reaction is measured 48 to 72 hours after administration. An induration of five to 15 mm or more is considered positive, dependent on the setting, with an induration greater than 10mm classified as positive in most high-burden settings. Specificity of TST is confounded by Bacillus Calmette-Guerin vaccination, and exposure to environmental non-tuberculous mycobacteria, both of which can induce immune responses that cross-react with purified protein derivative, whilst sensitivity of the test in immunocompromised persons is usually low (O’Garra et al., 2013; Thillai et al., 2014). *In vitro* assays such as the QuantiFERON-TB Gold In Tube (QFT) and the T-SPOT TB assays measure T cell responses to *Mtb*

antigens such as the early secretory antigenic target-6 and culture filtrate protein-10 (Thillai et al. 2014). These antigens are absent in the Bacillus Calmette-Guerin vaccine and most non-tuberculosis mycobacteria, thereby making IGRAs more specific to *Mtb* sensitisation than the TST (Andersen et al., 2000; Diel et al., 2011). The *in vitro* IGRAs include the mitogen phytohemagglutinin A as a positive control antigen, that allows identification of persons who might be immunocompromised because they cannot mount an appropriate immune response to the antigen (Andersen et al. 2000). TST and IGRAs can only detect host sensitisation to *Mtb* antigens, indicating possible *Mtb* infection, and do not accurately determine risk of progression to TB disease, nor do they indicate the stage of *Mtb* infection (Matteelli et al. 2017). Furthermore, these tests do not have the ability to differentiate between active TB and other respiratory diseases in those with *Mtb* infection and are unable to indicate the presence of replicating or viable bacteria (Barry et al. 2009). Time of *Mtb* infection is difficult to establish in the general population (Wallgren 1948), making it difficult to determine how long someone has been infected. Hence, determining when a person is at the highest risk of developing active TB disease is very challenging.

1.1.3 Treatment of *Mtb* infection and prevention of progression to active TB disease

Eradication of TB would require prevention of disease or elimination of the pathogen in the 23% of the global population who are estimated to be *Mtb*-infected. An obvious intervention would be antibiotic treatment in all who have evidence of *Mtb* infection. The benefits of such preventive therapy to stop TB disease progression in *Mtb*-infected persons have been known for more than 60 years (Schluger 2017).

WHO currently recommends isoniazid (INH) preventive therapy (IPT) for individuals at risk of developing TB in low-burden countries (World Health Organization 2015; Getahun et al. 2015). Soon after its efficacy in treatment of disease was established, INH was found to also be effective in preventing TB disease when provided to asymptomatic individuals (Lobue and Menzies 2010). Administration of IPT for six to 12 months can lead to a 90% risk reduction in developing active TB disease in *Mtb*-infected persons (Dye et al., 2013). The first efficacy trial of IPT was conducted by George Comstock in Alaskan natives from the USA, known as the Bethel trial (Comstock, Ferebee, & Hammes, 1967; Zwerling, Hanrahan, & Dowdy, 2016). In this doubled-blinded, randomized, placebo-controlled study, participants received either INH or placebo for a period of 12 months (Comstock et al., 1967). A lower TB incidence (1.90%) was observed in the INH arm than in the placebo arm (4.67%) after four years of follow-up (Comstock et al., 1967). This effect was greater in persons with a TST measuring at least 5mm and a reduction of incidence up to 90% was observed in these individuals (Comstock et al., 1967). This landmark trial suggested that IPT might be more effective in persons with prior sensitisation to *Mtb* or persons with memory responses to an *Mtb* antigen. The effect of the prophylaxis treatment persisted up to almost two decades after the initial evaluation, with a 68% reduction in TB cases in the study setting attributable to INH (Comstock et al., 1979).

Many studies have confirmed the efficacy of INH since the first efficacy trial of IPT. INH taken daily for a year could prevent both pulmonary and extra-pulmonary TB in approximately 25,000 household contacts from the United States of America (Ferebee and Mount 1962). A study of contacts of TB cases observed a five-year cumulative risk of TB of 1.4 (95% CI: 0.03 to 2.6) in those who had preventive

therapy, compared to 2.4 (95% CI: 1.2 to 4.7) in those who did not receive therapy (Sloot et al., 2014). In South African children, INH reduced the risk of active TB disease in *Mtb*-infected children from a 16-fold to a 2.5-fold risk relative to *Mtb*-infected children who did not receive IPT (Bunyasi et al. 2017). A systematic review of 11 randomised double-blinded trials consisting of over 73,000 HIV-uninfected participants randomised to receive either INH or placebo found that among 796 participants who developed active TB disease, only 239 were from the INH arm, resulting in a relative risk of 0.40 (95% CI: 0.31 to 0.52) (Smieja et al. 1999).

Due to long treatment duration and hepatotoxicity of INH, several studies have evaluated shorter treatment regimens and evaluated the safety and adherence of new regimens (Lobue and Menzies 2010). At present, none of these is less efficacious or results in fewer adverse events than nine months of INH therapy (Lobue and Menzies 2010). Consequently, five different treatment regimens are recommended for preventive therapy (Getahun et al. 2015; World health Organization 2015): (i) six or nine months of isoniazid, (ii) three months of rifampicin, (iii) three to four months of rifampicin alone, (iv) Rifampicin plus isoniazid, or (iv) rifapentine plus isoniazid, weekly for three months. A meta-analysis of randomised controlled trials observed comparable efficacy of these different treatment regimens in different study populations from various settings (Stagg et al. 2014). In high-burden countries, provision of preventive therapy is not recommended to HIV-uninfected adults because it is very challenging to treat such a large proportion of the population when only a small subset of *Mtb*-infected persons will progress to active TB during their lifetime. Furthermore, continuous re-exposure to *Mtb* results in reinfection, placing people at risk of disease soon after they have completed

preventive therapy. High prevalence of infection and exposure preclude determination of recent *Mtb* infection with TSTs and IGRAs. When these issues have to be balanced against inadequate adherence to long preventive therapy regimens that is often reported and a risk of liver damage from hepatotoxic effect of IPT. Perhaps, identification of persons most likely to progress to active TB disease in high-burden countries would enhance IPT uptake. In addition, the duration of the treatment could be determined by a person's relative risk of developing active TB disease. For example, persons at highest risk of developing active TB disease within a short time (six to 12 months) can be given longer preventive therapy, whereas persons at a lower risk of developing active TB can be given shorter treatment regimens and be followed. This could be used to customise IPT to individual needs, thereby increasing adherence, but the tools to determine such risk are currently lacking.

1.1.4 Diagnosis of active TB disease

Active TB disease patients typically present with a number of the following signs and symptoms: fever, fatigue, drenching night sweats, chest pain, weight loss, anaemia, appetite loss, wasting and, in patients with pulmonary TB, persistent coughing (lasting at least two weeks) or haemoptysis (Davies and Pai 2008; Pai et al. 2016). Definite TB is defined by laboratory confirmation of *Mtb* from a clinical specimen, usually sputum (Davies and Pai 2008). Current diagnostic methods mainly rely on detection of the pathogen in sputum (microbiological confirmation), evaluation of clinical symptoms and radiological assessments of individuals presenting with the symptoms of the disease in clinics. The most common method of diagnosis used globally is sputum smear microscopy. However, the sensitivity of smear microscopy

is poor, ranging from between 34% to 80% in comparison to cultures, with 20% to 66% of active TB cases being smear negative (Davies and Pai 2008). The current "gold standard" for TB diagnosis is detection of *Mtb* by culture in liquid or solid media, which yields a sensitivity ranging from 80% to 93% and a specificity of 98%. In addition, cultures allow speciation and drug susceptibility testing because the cultured organism can be characterised further (Davies and Pai 2008). Poor sample collection and paucibacillarity of sputum samples, however, reduces the sensitivity of this test (Ndzi et al. 2016). A major disadvantage of current culture methods is that it can take up to three or four weeks before diagnosis can be confirmed, due to the slow growth of *Mtb* (Davies and Pai 2008). Nucleic acid amplification tests (NAAT) are molecular diagnostic tests that can generate results within hours, thus theoretically facilitating diagnosis and prescription of appropriate treatment during the same patient visit. As a result, many countries have invested very significantly into implementation of the NAAT, Xpert MTB-RIF (Cepheid), as the first-line TB diagnostic in clinics and hospitals. Although Xpert MTB-RIF has led to more rapid diagnosis of TB in many settings, this has not resulted in very significant changes in TB mortality (World Health Organization 2017). NAATs fail to differentiate between live and dead bacteria, and are therefore not useful to monitor whether patients have cleared *Mtb* during TB treatment (Davies and Pai 2008). In addition to these diagnostics, clinical diagnosis can be done using imaging techniques such as chest X-rays and, more recently, positron emission tomography-computed tomography (PET-CT) scanning (Pai et al. 2016). These radiographic imaging techniques can detect pulmonary features that are consistent with TB disease, but they are inherently non-specific and therefore serve as screening tools, whilst microbiological

tests result in confirmation of the presence of viable bacilli in sputum. Thus, better tools are needed for TB to improve TB diagnosis.

1.1.5 Immune response to TB disease

"Immune response to TB requires an interaction between the cellular and humoral compartments including both the innate and adaptive responses. Knowledge of early immune response to *Mtb* in humans is limited and current knowledge is based on experimental animal models, such as mice, rabbits, guinea pigs and non-human primates. Alveolar macrophages are the first immune cells of contact with the pathogen and engulf the bacteria by receptor-mediated phagocytosis, requiring different receptors (van Crevel et al. 2002). Phagosome-lysosome fusion is blocked by the mycobacterium ensuring its survival. Thereafter, several cell types are recruited to the site of infection (i.e lungs) leading to granuloma formation, which is the hallmark of infection. Granulomas consist of both innate cells such as macrophages, foamy macrophages, neutrophils, and cells involved in the adaptive immune response such as T and B cells (O'Garra et al. 2013). Despite their presence in the granuloma, the role of neutrophils has not been fully characterised in the immune response against TB. Neutrophils are the predominant cell types in the sputum and bronchoalveolar lavage fluid of TB patients (Eum et al. 2010). Furthermore, neutrophils have been observed to be the predominant cell type to express the genes implicated in numerous published transcriptomic signatures of TB disease (Berry et al. 2010; Scriba et al. 2017), discussed in section 1.7. Although little is known about the role of B cells and antibodies in the protective immune response against TB, they play an integral role in the maintenance and formation of granulomas. In addition, they can present antigen to T cells in the granuloma and

could modulate the immune inflammation through IL-10 production (Rao et al. 2015). T cells are the hallmark of the adaptive immune response against *Mtb* and have been the most studied immune cells in TB research. *Mtb*-specific responses are dominated by CD4 T cells producing TNF-α and IFNγ and recently CD8 T cells have also been shown to play significant roles in patients with TB disease (reviewed by Scriba et al. 2017). Furthermore, susceptibility to TB disease is associated with an increase in the monocyte to lymphocyte ratio in both children (Naranbhai et al. 2014) and adults (Naranbhai et al. 2015) including HIV-infected persons (Naranbhai et al. 2013).

1.2 Treatment of active TB disease

Standard pulmonary TB diagnosis requires six months of treatment, in most countries consisting of two months of intensive TB treatment with isoniazid, rifampicin, pyrazinamide and ethambutol (2HRZE) followed by four months of isoniazid and rifampicin (4HR) for the continuation phase of treatment (Parida and Kaufmann 2010). Such treatment results in cure in most patients, however due to poor adherence or poor treatment response cure is not achieved in some patients. This leads to a high risk of treatment failure and relapse due to insufficient treatment. The emergence and increasing prevalence of multiple drug resistant (MDR-TB) and extensively drug resistant (XDR-TB) strains presents a further challenge to the successful treatment of TB and has exacerbated the global TB problem (Raviglione and Smith 2007). MDR-TB is defined as resistance within a strain of *Mtb* to at least one first line drug, usually isoniazid or rifampicin. XDR-TB is resistance to fluoroquinolones and any of the injectable second-line aminoglycosides in addition to resistance to the first line drugs (Raviglione and Smith 2007). In 2016, approximately

600,000 new cases with drug resistance to rifampicin were reported, 490,000 of whom had MDR-TB (World Health Organization 2017). MDR-TB and XDR-TB are treated using second line drugs such as aminosalicylic acid, thioamides, ethionamide, prothionamide, capreomycin and fluoroquinolones and often for very long times, up to two years (Fernandes et al., 2017). Poor adherence to such long-term regimens can lead to increase in transmission and insufficient treatment, which is assumed to be the main cause of drug resistance (Friedrich et al. 2013), in addition to an increase risk of treatment failure and relapse. Furthermore, drug penetration in the lungs (Dheda et al. 2018), host genetic factors, heterogeneous treatment responses in patients and simplified treatment due to poor public health control strategies (reviewed in Koch et al. 2018), significantly contribute to acquisition of drug resistance. It is clear that better tools for monitoring the response of TB patients to TB treatment would be extremely helpful such that the duration of treatment and type of treatment can be tailored to each patient's need.

1.2.1 New regimens to shorten TB treatment and prevent relapse.

Several studies have investigated new treatment regimens to shorten treatment in order to possibly increase adherence (Gillespie et al. 2014; Burman et al. 2006; Rustomjee et al. 2008; Tuberculosis Research Center 2002; Jindani et al. 2014; Merle et al. 2014). Gillespie and colleagues replaced INH or ethambutol with moxifloxacin (a fluoroquinolone) for four months to determine its non-inferiority to the standard six-month regimen (Gillespie et al. 2014). However, the moxifloxacin regimen was inferior to the standard of care after 18 months of follow-up of nearly 2,000 newly diagnosed active TB cases. By substituting Moxifloxacin for ethambutol, Burman *et al* showed that despite resulting in higher proportions of negative cultures

at early time points, moxifloxacin did not have an effect on month-two culture conversion rates (Burman et al. 2006). This suggests that moxifloxacin has sterilising activity but has deficient activity when added to other drugs to support treatment shortening based on two-month culture conversion results (Burman et al. 2006). Burman and colleagues did not follow up patients to determine relapse rates, which is a more definitive measure of sterilizing activity. In two other studies, substituting ethambutol with fluoroquinolones to shorten treatment decreased time to culture negativity but did not decrease relapse rates in South Africa (Rustomjee et al. 2008) and India (Tuberculosis Research Center 2002). Higher relapse rates were observed in a treatment regimen of once-weekly rifapentine in the continuation phase compared to a twice-weekly rifampicin regimen (Benator et al. 2002). Higher relapse rates were also observed when a four-month regimen, in which moxifloxacin and rifapentine were given once weekly after the intensive phase in comparison to the six-month standard of care (Jindani et al. 2014). Culture conversion (i.e. a negative sputum *Mtb* culture) within two months after TB treatment is associated with lower relapse rates and treatment failure (Mitchison 1993). In the studies mentioned above, despite high culture conversion rates at two months the relapse rates in the intervention arms were equivalent or higher than those in the standard of care arm. This suggests that better markers of relapse and treatment failure are needed to monitor treatment response and predict relapse after clinical and microbiological cure.

1.3 Recurrent TB disease and treatment

A second episode of active TB disease after an initial “cured” TB episode is referred to as a recurrence (Lambert et al. 2003). In a five-year follow-up of “cured” active TB

persons, 18% of study participants developed recurrent TB (Verver et al. 2005). Of these, 14% had completed treatment and were classified as “clinically cured” (Verver et al. 2005). In 2016, an estimated 300,000 recurrent TB episodes were diagnosed by National TB programmes around the globe, representing an enormous burden on TB control efforts (World Health Organization 2017).

If the recurrent TB episode is due to a different *Mtb* strain relative to the primary disease episode, this is referred to as reinfection. In cases where the same strain is isolated from the two TB episodes, this is referred to as relapse. Risk of TB recurrence due to reinfection is determined by the risk of reinfection after a first episode and the risk of the reinfection progressing to a second episode of active disease (Lambert et al. 2003). Relapse generally occurs because of bacterial persistence after apparent “clinical cure”, mainly due to insufficient treatment efficacy either because of inadequate duration or an inefficient regimen (Lambert et al. 2003). Recurrence due to either relapse or reinfection is clinically indistinguishable. However, with recent developments in molecular techniques, such as whole genome sequencing and IS6110-based restriction fragment-length polymorphism, identification of relapse and reinfection TB is possible by identifying TB strains from the first and the second TB episode diagnosis. In a study in the Karonga district of Malawi, 73.3% of recurrent TB cases were due to relapse (Guerra-Assunçąõ et al. 2015). Recurrent TB generally occurred within a year of TB treatment completion, at a rate of 5.4 recurrences/100 person-years (py) (Guerra-Assunçąõ et al. 2015). A study in South Africa observed a similar trend whereby most relapses occurred in the first year, whilst reinfection featured mainly in the subsequent years (Marx et al. 2014). In contrast to the Malawian study, at least half of all TB recurrences in the

South African cohort were due to reinfections. These studies, in addition to other studies with low sample sizes (Boer et al., 2003; Van Rie et al., 1999), suggest that previous TB disease is a strong risk factor for subsequent recurrent TB disease and that prior TB disease does not induce immune memory responses after cure that provide protection against subsequent episodes of TB.

Due to the high incidence and probability of MDR-TB in recurrent TB cases, there is currently no standard of care for treating recurrent TB. The WHO recommends placing recurrent TB cases on empirical treatments whilst awaiting drug susceptibility testing results where available (World Health Organization 2010). In places where MDR-TB levels are low, an 8-month treatment regimen of either two months of HRZES or one month of HRZE and then five months of HRE should be administered, and, if MDR results become available, treatment should be adjusted accordingly.

1.4 HIV and TB co-infection

1.4.1 Epidemiology of TB disease and HIV co-infection

The advent of the HIV/AIDS pandemic has markedly changed the global TB pandemic and still poses one of the greatest challenges to health systems around the world. An interaction between *Mtb* and HIV was detected circa 40 years ago, fuelling the TB disease burden and mortality due to TB in humans (Harries et al. 2010). HIV co-infection is a major factor in the annual increase in global TB incidence rates between 1990 and 2004, leading to a five-fold increase in TB notification rates (Lawn et al., 2009). An estimated 46% of incident TB cases, in 2016, were HIV co-infected with the highest proportions from the African region

(World Health Organization 2017). In addition to the 1.3 million TB deaths in HIV-uninfected persons, 374,000 deaths occurred in HIV-infected persons, making TB the leading cause of mortality in HIV-infected persons (World Health Organization 2017). Active TB disease has been associated with reduced survival in Ugandan HIV-infected persons despite being highly adherent to TB treatment (Whalen et al. 2000). In a five-year follow up of 609 HIV-infected patients before the wide availability of antiretroviral therapy (ART), 103 (17%) deaths were recorded (Badri et al., 2002). Of these, 50 occurred in patients diagnosed with TB (Badri, Wilson, and Wood 2002). Progression to AIDS and AIDS defining illnesses were higher in persons with TB than in those without TB in the same cohort.

ART has caused a decrease in HIV-associated morbidity and mortality due to the partial restoration of CD4 T cells and suppression of the virus, but health is not fully restored (Klatt et al., 2013). Of the HIV co-infected notified TB cases in 2016, 85% were on ART (World Health Organization 2017). ART given alone, can decrease the risk of TB disease in HIV-infected persons up to 65% (Suthar et al. 2012). Despite this, HIV-infected persons on ART still have a higher TB incidence rate in comparison to HIV-uninfected cohorts from the same communities (Golub et al., 2009; Gupta et al., 2012; Lawn et al., 2009). In a well-defined South African township community, TB prevalence decreased from 44% to 6.7% over a three-year period in HIV-infected persons on ART (Middelkoop et al. 2010). TB incidence rates were higher in the first year after ART start, decreasing by almost half in the second year, from 12.43 cases per 100 py to 6.71 cases per 100 py (Gupta et al. 2012). These rates further decreased to 4.92 cases per 100 py after five years. A similar trend was observed in another study of TB incidence in persons beginning ART, wherein TB

incidence decreased from 13.9 per 100 py in the first three months to 1.6 per 100 py in the third year (Van Rie et al., 2011). A multi-centre study observed a decrease in incidence risk from 13.0 per 100 py in the first three months after ART start to 1.5 per 100 py after a year in eight sub-Saharan African countries (Nicholas et al. 2011). In these studies, the higher TB incidence in the first year can be due to the incidence of immune reconstitution inflammatory syndrome (IRIS). Incidence of immune reconstituting inflammatory syndrome was highest in the first three months after ART start.

The degree of immunosuppression, quantified by CD4 T cell counts, is a key determinant of TB risk in HIV-infected persons. ART increases CD4 T cell counts, thereby decreasing the risk of TB in HIV-infected persons. A 10% reduction in relative risk of TB disease attributable to ART was observed as CD4 T cell counts increase from less than 100 cells/mm^3 to more than 500 cells/mm^3 (Wood and Lawn 2011). In patients starting ART, 42% of persons with prevalent TB had CD4 counts of less than 50 cells/mm^3 (Van Rie, Westreich, and Sanne 2011). Early incident TB (TB within six months of ART start) was associated with advanced immunosuppression (Van Rie, Westreich, and Sanne 2011). In addition, patients with CD4 T cell counts less than 50 cells/mm^3 were twice as likely to develop TB disease than those with CD4 T cell counts of at least 100 cells/mm^3 (Van Rie, Westreich, and Sanne 2011). Patients initiated on ART during the early phases of HIV are only two times more likely to get TB in comparison to HIV-uninfected persons (Wood and Lawn 2011). Thus, reconstitution of functional CD4 T cell immune responses is used as a marker of effective ART in addition to a decline in HIV viral loads. Reconstitution of CD4 T cells occur initially for memory cells and

naïve cells, however, persistent impairments of specific CD4 T cell populations can be observed despite long-term ART (Walker et al., 2013). A 1.7-fold increased risk of TB was observed early after ART initiation in persons with CD4 T cell counts less than 100 cells/μl (Lawn et al., 2009). The authors attribute the high levels of TB during the early phases of ART treatment to the unmasking of subclinical TB in these patients.

1.5 Preventing TB in HIV-infected people

The WHO policy on collaborative TB/HIV activities has an aim to "reduce the burden of TB in people living with HIV, by ensuring the delivery of the three "I"s (intensified case finding, infection control, and IPT), for HIV/TB" (World Health Organization 2012b). In addition to these, ART has been proposed to be included in TB prevention strategies in HIV-infected persons (Lawn et al., 2009). In order to achieve this, it is recommended that adults and adolescents living with HIV should be regularly screened for TB (World Health Organization 2012b). All HIV-infected persons not diagnosed with active TB should receive at least six months of IPT irrespective of degree of immune suppression (World Health Organization 2012b). The implementation of the three "I"s complements the effects of early ART and must be scaled up to prevent mortality and morbidity from TB in HIV-infected persons (Harries et al. 2010). However, only ART has been implemented on a large scale in most countries (Gupta et al. 2012).

Intensified case finding aims to identify persons with subclinical TB and in need of treatment, leading to treatment start in those with TB disease and IPT in those without TB disease (Lawn et al., 2009). This is intended to reduce, 1) individual

morbidity and mortality through early diagnosis and treatment, 2) TB transmission via shortening of infectious period and, 3) exclude TB to allow preventive treatment (Corbett and MacPherson 2013). In a South African study passive case finding identified only 33% of HIV-infected persons with smear positive TB (Wood et al. 2007), strengthening the need to implement active and intensified case finding approaches in communities.

ART reduces the risk of active TB disease in HIV-infected persons, albeit in a time dependent manner (Lawn et al., 2011). The incidence of TB disease can be more than halved if ART is started within five years of HIV seroconversion (Harries et al. 2010). Mathematical modelling of HIV-associated TB in sub-Saharan African countries estimates a reduction in incidence rates by 66%, 95% and 98% if ART is started within five years, two years or one year of HIV seroconversion, respectively, by 2050 (Williams et al. 2010). Scale up of ART can thus result in an estimated reduction in TB incidence of up to 25% over a five-year period and up to 50% over a 20-year period, according to a modelling approach (Chindelevitch et al. 2015). In agreement with this, a risk reduction of up to 70% (range: 54% to 92%) was observed in observational ART cohort studies in low burden and high burden countries (Lawn et al., 2010; Lawn, et al., 2009).

Preventive TB therapy has been shown to reduce the risk of active TB disease by 36% to 64% in a review of more than 5,000 HIV-infected persons from 13 trials (Volmink and Woldehanna 2004). A systematic review of 12 randomized controlled trials of HIV-infected individuals showed that a six to twelve month course of daily INH therapy can reduce TB incidence up to 32% (Akolo et al. 2010). Mass

administration of IPT protects individuals against TB disease but has no effect on TB incidence in the community in a high incidence setting (Churchyard et al. 2014; Vynnycky et al. 2015). In a Zambian study a two-fold decrease in TB incidence was observed in HIV-infected persons who received six months of IPT, relative to those who received placebo (Mwinga et al. 1998). However, this protective effect of IPT was lost during a seven-year follow-up of participants from the same study (Quigley et al. 2001), with a higher TB incidence rate of 4.0 per 100 py observed in the IPT group relative to an incident rate of 2.1 per 100 py in the placebo group. Durability of protection by six months of IPT is short-lived as a limited benefit has been observed as early as 12 months after commencement of therapy (Johnson et al. 2001). No significant effect of six months of IPT daily was observed on both TB incidence and prevalence after TB therapy in a massive trial of gold miners (Churchyard et al. 2014). These studies suggest that there is limited duration of benefit after six months IPT in HIV-infected persons.

This has led to other randomised trials examining the durability of protection against TB offered by IPT and the effect of increasing IPT duration in HIV-infected persons. Extending IPT to 36 months reduced TB incidence by 43% in a study in Botswana, but protection was sustained only in TST-positive individuals who completed therapy (Samandari et al. 2015, 2011). In India, six-month treatment of daily IPT and ethambuthol was equivalent to a 36-month regimen of daily IPT only in preventing active TB (Swaminathan et al. 2012). Extending therapy to six years did not provide any additional benefits in preventing active TB when compared to six months of daily therapy in a South African cohort (Martinson et al. 2011). A meta-analysis of the three studies mentioned above detected a 38% reduction in risk of developing active

TB with 36 months IPT compared to six months of IPT (Boon et al., 2016). This suggests that long-term IPT is required to provide adequate protection against TB disease in HIV-infected persons. The lack of durable protection after completion of six months of INH was attributed to repeated exposure to *Mtb*, or high force of re-infection, seen in TB endemic settings (Havlir et al., 2008). However, a recent clinical trial demonstrated that a shortened treatment regimen consisting of once weekly high-dose rifapentine and isoniazid for three months was as effective as and non-inferior to nine months of daily INH (Sterling et al. 2016). Treatment completion rates were higher in the shorter rifapentine arm compared to the isoniazid only arm, suggesting better adherence in the three-month study arm. Protection against active TB was observed to last up to three years after a three-month prophylaxis of rifampicin containing drugs (Johnson et al. 2001). In addition to preventing active TB disease, IPT can prevent TB-associated mortality in HIV-infected persons up to 37% (Badje et al. 2017; Durovni et al. 2013). These studies suggest that shortening preventive therapy can prevent TB in HIV-infected persons and increases adherence, thereby reducing the risk of active TB in HIV-infected persons.

Daily IPT given in combination with ART for six months offers durable protection against TB, which lasted up to seven years after IPT initiation in a Brazilian study (Golub et al., 2007; Golub et al., 2015). Another study in both urban and rural settings in South Africa demonstrated TB risk reduction of 89% TB due to the combined effect of IPT and ART (Golub et al., 2009). In a South African township (informal residence), 12 months of IPT reduced TB incidence independent of ART by 37% (Rangaka et al. 2014). These studies suggest that combining IPT and ART will be more effective in preventing TB in HIV-infected persons. Thus, combining TB and

HIV care in a clinical setting is likely to effectively reduce the burden of active TB in HIV-infected persons.

The above-mentioned studies suggest that IPT should be given for extended periods in HIV-infected persons to prevent progression to active TB disease, as is recommended (World Health Organization 2011). In addition to poor adherence to treatment, as mentioned above, long-term IPT increases the risk of hepatotoxicity. Adherence may be increased if only those persons at risk are identified, provided counselling, and provided long-term directly observed therapy to decrease risk of progression to disease. In addition, such individuals should be further investigated for subclinical TB disease and if found to be sputum positive for *Mtb*, provided with TB treatment. Thus, markers to identify HIV-infected persons at greatest risk of progression to TB disease would enhance early case detection and customising length of preventive therapy or full treatment for TB disease.

1.6 Progression to active TB, treatment and recurrence in HIV infection

HIV infection increases the risk of progression and reactivation of *Mtb* infection to active TB disease (Kasprowicz et al. 2011; Gupta et al. 2012). Risk of TB disease doubles in the first year after HIV-seroconversion relative to HIV-uninfected persons (Sonnenberg et al., 2005) and remains elevated after the first year of seroconversion at levels 20 to 30 times higher than in HIV-uninfected persons (Getahun et al., 2010; World Health Organization, 2012). Clinical features of TB presentation are also highly dependent on the stage of HIV infection and immunodeficiency (Maher et al., 2005). Pulmonary TB in persons with early HIV infection with mild/moderate immunodeficiency usually presents with the same signs and symptoms as HIV-

uninfected persons whereas advanced/late HIV infection is associated with an increase in disseminated disease and less cavitary disease (Maher, Harries, and Getahun 2005; Tornheim and Dooley 2017). This makes TB much easier to diagnose in early HIV infection, as patients usually present with productive coughs and are more often smear positive TB (Maher, Harries, and Getahun 2005; Tornheim and Dooley 2017).

1.6.1 Diagnosis and treatment of active TB disease in HIV-infected persons

Diagnostic tests for active TB disease are not different for both HIV-infected and uninfected persons (Tornheim and Dooley 2017), and sputum culture or smear microscopy are the most commonly used methods of diagnosing pulmonary TB. Bacterial culture is considered as the "gold-standard" for diagnosis. However, diagnosis of TB in HIV-infected persons is complicated by the paucibacillary nature of disease, leading to smear- and culture-negativity in a large proportion of diseased individuals, particularly in non-cavitary pulmonary TB with advanced immunosuppression (Tornheim and Dooley 2017). In addition, chest X-rays are less reliable in HIV-infected persons as the disease presents with atypical patterns; up to 14% of individuals with positive cultures present with normal chest x-rays (Getahun et al., 2007). Using the current WHO screening recommendations for culture positive active TB cases with HIV co-infection from three Southeast Asian countries, 179 individuals had no coughs, whilst 75 individuals with coughs had normal chest X-rays and negative sputum smears from two sputum specimens (Cain et al. 2010). In the same study, mycobacterial culture was the most effective method of diagnosing TB, especially in participants with negative sputum smears. In a South African study

cultures were also more sensitive than sputum smears (81% and 41%, respectively) in diagnosing active TB in an ART roll out program (Lawn et al., 2006).

The standard six-month treatment regimen for active TB in HIV-uninfected persons is recommended for drug susceptible pulmonary TB cases on ART (World Health Organization 2017). Early initiation of ART after starting TB treatment decreases the risk of mortality due to TB (Marcy et al. 2014; S. Abdool Karim et al. 2011). The WHO recommends that all newly diagnosed TB cases that are HIV co-infected should be started on both ART and TB treatment regardless of CD4 T cell counts (World Health Organization 2017). However, due to the risk of immune reconstitution inflammatory syndrome (IRIS), ART should be delayed to start within eight weeks of start of TB treatment, whereas individuals with CD4 T cell counts less than 50 cells/mm^3 should be started on ART within two weeks (World Health Organization 2017). A systematic review of seven studies of HIV co-infected persons showed that treatment with a rifamycin-containing regimen for at least six months with ART is associated with lower treatment failure and relapse rates relative to treatment without ART (Khan et al. 2010). Hence, TB treatment given concurrently with ART decreases mortality due to TB in HIV co-infected persons.

1.6.2 Recurrent TB disease in HIV-infected persons

HIV infection is a risk factor for the development of recurrent TB disease (Korenromp et al., 2003). Several studies have shown that HIV-infected persons are more likely to develop recurrent TB disease than HIV-uninfected persons (Charalambous et al., 2008; Fitzgerald et al., 2000; Mallory et al., 2000; Sonnenberg et al., 2001). In a study of TB notification rates in Cape Town, recurrent TB cases constituted 30% of

all HIV cases accounting for almost 4,000 TB cases (Wood et al. 2011). Recurrent TB rates were higher in HIV-infected persons on ART in comparison to those not on ART, suggesting better diagnosis of TB in individuals on ART, failed initial treatment, multiple re-infections, or reactivations (Wood et al. 2011). In a community cohort of HIV-infected persons on ARTs, recurrent TB diagnosis accounted for 75% of the TB disease burden (Gupta et al. 2012). In a systematic review of 47 studies of recurrent TB disease, an average recurrence rate of 4.5 per 100 py in HIV-infected persons was observed (Korenromp et al. 2003). In comparison, the recurrence rate was 1.9 per 100 py in HIV-uninfected persons. Despite the higher number of recurrences in HIV-infected persons, median time to recurrent TB disease was not different between HIV-infected and uninfected persons, at 9.7 months (95% CI, 4.9 to 14 months) and 7.1 months (95% CI, 4 to 10 months), respectively (Korenromp et al. 2003). Rate of TB recurrence was two-fold higher in HIV-infected persons in comparison to HIV-uninfected persons in a study of recurrent TB disease over a 10-year follow up period (Crampin et al. 2010). Therefore, recurrent TB contributes significantly to the TB burden in HIV co-infection, and identification of HIV-infected persons at risk of recurrent TB disease is important if we are to decrease the burden of TB in HIV-infected populations.

1.6.3 HIV immmunopathogenesis and it's relevance for TB

Immune activation and systemic inflammation are hallmarks of HIV immmunopathogenesis and high levels of both are associated with risk of progression to AIDS. Hyper-activation and exhaustion of the immune system and depletion of CD4 T cells due to untreated HIV infection results in chronic immune dysfunction (Catalfamo et al., 2008; Klatt et al., 2013). Depletion of CD4 T cells and

HIV viremia lead to increased proliferation of CD4 T cells, which is accelerated by the inflammatory milieu induced by the virus (Catalfamo et al. 2008). This immune activation can be measured by CD38 and HLA-DR activation on CD8 and CD4 T cells, and is significantly higher in HIV-infected persons relative to uninfected controls (Joshi et al. 2016; Deeks et al. 2004).

Damage to the mucosal barriers and lymphoid structures in the gut are major drivers of immune activation and dysfunction in HIV infection (Brenchley et al., 2006; Deeks et al., 2013; Klatt et al., 2013). In addition, gut leakage is a contributing factor to inflammation in HIV infection as it leads to an increase in inflammatory cells such plasmacytoid dendritic cells (pDCs), neutrophils and monocytes in peripheral blood (Klatt et al. 2013; Somsouk et al. 2015). This inflammation drives T cell migration to the damaged tissue, leading to a higher concentration of activated target cells, thereby making HIV replication more efficient (Klatt et al. 2013). HIV infection depletes CD4 T cells, an important cell subset in TB control, increasing susceptibility of HIV-infected people to progress from *Mtb* infection to TB disease (Caruso et al., 1999; Diedrich & Flynn, 2011; Lawn et al., 2009; Lin et al., 2012). Furthermore, infection with HIV significantly decreases memory CD4 T cell responses to not only the virus but also to other opportunistic infections as early as a few days after infection (Veazey and Lackner 2005). Additionally, HIV infection impairs the ability of CD4 T cells to respond to *Mtb*-specific antigens by altering the balance of their functional profile through the selective depletion of IFNγ-producing cells (Riou et al. 2016). Hence, HIV infection increases susceptibilty to TB disease and other opportunistic infections and is associated with high levels of immune activation and systemic inflammation.

Antiretroviral drugs that target HIV to arrest viral replication should be given as combination treatment, referred to as ART. ART regimens typically prescribed in recent years consists of two nucleoside/nucleotide reverse transcriptase inhibitors and a third drug from a different class, such as protease inhibitors (Cihlar and Fordyce 2016). ART suppresses the replication of HIV thereby decreasing circulating viral loads, resulting in an asymptomatic aviremic state in most HIV-infected persons. The associated immune reconstitution has been shown to decrease morbidity and mortality. However, immune activation and persistent inflammation remain elevated in HIV-infected persons who receive effective ART, in comparison to HIV-uninfected persons, even if viral loads become undetectable (Klatt et al. 2013). Consequently, the risk of opportunistic infections, including TB, remain elevated in persons on ART relative to HIV-uninfected persons. A biomarker or blood-based diagnostic test that could identify individuals at high risk of TB would allow targeted investigation and treatment of individuals either with TB treatment in those with disease, or with preventive therapy in those who are not yet sick.

1.7 The search for a biomarker of TB disease and/or protection against TB: Transcriptomic signatures as biomarkers for TB disease

Despite what has been learnt of the relationship between *Mtb* and the host, components of protective immunity remain largely unknown. A biomarker is “a characteristic that is objectively measured and evaluated as an indicator of physiological or pathological processes or pharmacological responses to a therapeutic intervention” (Parida & Kaufmann, 2010). Biomarkers of TB would ideally allow the diagnosis, prediction of risk and monitoring of treatment response of active TB disease using easily accessible bodily fluids such as blood. This would facilitate

early diagnosis of active TB, provision of targeted preventive therapy and allow customisation of the length of TB treatment. The advent of new genome-wide technology, particularly the “omics” technologies, offers new tools in the discovery of biomarkers in diagnosis of active TB using unbiased, hypothesis-generating approaches (Cliff et al., 2015). This has led to an explosion of data in the field of biomarker studies. A profile of combined biomarkers is known as a signature (Parida & Kaufmann, 2010).

Early published studies showed that differential expression levels of selected mRNAs in blood leukocytes, measured by microarray, can accurately classify persons with *Mtb* infection and persons with active TB, demonstrating that combinations of transcriptomic biomarkers hold promise as diagnostic signatures (Jacobsen et al. 2007; Mistry et al. 2007). Mistry and colleagues identified nine gene transcripts that were differentially expressed between active TB, recurrent TB, cured and *Mtb*-infected persons (Mistry et al. 2007). Jacobsen *et al*, suggested that expression levels of only three genes could be used to discriminate between TB disease and healthy *Mtb*-infected controls with an accuracy of 85% (Jacobsen et al. 2007). A more recent publication, the first study to describe in detail, a gene signature due to the presence of active TB disease, identified a 393-gene signature that distinguished active TB from healthy *Mtb* infection and other diseases with promising accuracy (Berry et al. 2010). The signature also correlated with the extent of radiographic disease, suggesting that this blood signature reflected pulmonary disease severity (Berry et al. 2010). A common theme in these studies was that many if the most differentially expressed genes mapped to the type I/II interferon (IFN) pathways and were interferon-stimulated genes (ISGs). Interestingly, the most

prominent differences in expression of these ISGs were observed in myeloid cells such as neutrophils and monocytes, although they were also differentially expressed to a lesser degree in lymphoid cells such as CD4 and CD8 T cells (Berry et al. 2010). In a Gambian cohort, Maertzdorf *et al* discovered a signature that could differentiate active TB from healthy *Mtb* infection (Maertzdorf et al., 2011). An increase in complement and inflammation-associated genes in active TB disease has been observed in addition to differences in type I IFN genes (Maertzdorf et al., 2011). However, the signature could not differentiate healthy non-*Mtb*-infected (TST negative) persons from *Mtb*-infected persons (Maertzdorf et al., 2011). Pathway analysis of the 1,000 most differentially expressed genes in this signature again revealed an enrichment of type I IFN genes, which have also been shown to be involved in systemic lupus erythematosus. Furthermore, genes involved in the JAK-STAT and complement pathways were highly expressed in active TB patients. More recently, studies have been published using different transcriptomic analysis platforms that could differentiate active TB from healthy *Mtb* infection in whole blood and peripheral blood mononuclear cells (PBMCs) (Lee et al. 2016; Lu et al. 2011; Walter et al. 2016). These studies suggest that whole blood transcripts involved in type I/II IFN pathways are involved in TB disease pathology and that transcripts can be used to distinguish between TB diseased and *Mtb*-infected or healthy individuals.

Since several diseases, such as sarcoidosis, melioidosis and pneumonia, share a similar pathology with TB, it is particularly challenging to definitively make a TB diagnosis in the absence of microbiological confirmation of *Mtb*. Diagnosis of TB using non-specific methods such as symptom screening and chest X-rays may therefore lead to a misdiagnosis and provision of incorrect care. To determine if

blood-based transcriptomic signatures can improve the diagnosis of TB, several studies have identified gene signatures that distinguish TB from other diseases, which mimic and/or present in the same manner as TB. Maertzdorf and colleagues observed a significant overlap in type I/II IFN signalling genes identified to be differentially expressed between active TB disease or sarcoidosis and healthy controls (Maertzdorf et al., 2012). Another transcript signature of active TB disease was also found to be expressed in sarcoidosis patients and increased type I/II IFN gene modules was observed in both diseases relative to healthy controls (Bloom et al. 2013). In contrast, these signatures were different from signatures expressed in pneumonia and lung cancer patients, in whom upregulation of myeloid inflammatory modules was observed (Bloom et al. 2013). Koth *et al* observed the same phenomenon, wherein differentially expressed transcripts overlapped between sarcoidosis and active TB (Koth et al. 2011). Another bacterial disease, melioidosis with pathology similar to TB, has been shown to also have a signature that largely overlaps with active TB signatures (Koh et al. 2013). An attempt to develop a signature that could distinguish melioidosis from active TB was not successful (Koh et al. 2013). These studies suggest that blood transcriptomic signature profiles reflect underlying systemic activation of proinflammatory pathways, such as type I/II IFN signalling genes, and might not have the appropriate specificity to stratify active TB disease from other relevant diseases.

While considerable overlap is seen between the gene signatures of active TB disease identified in different studies mentioned above, each signature consists of unique sets of differentially expressed genes. This could be due to different clinical study designs, diagnosis methods, underlying epidemiology and the diverse

platforms, methodologies and statistical approaches used in assessing differentially expressed genes, which are very variable (Maertzdorf et al., 2015). A recent study attempted to combine analysis of all the genes identified in multiple TB biomarker studies (Berry et al., 2010; Cliff et al., 2013; Jacobsen et al., 2007; Maertzdorf et al., 2012; Maertzdorf et al., 2011; Maertzdorf et al., 2011; Mistry et al., 2007; Ottenhoff et al., 2012) to assess common pathways involved in the pathogenesis of TB (Joosten et al., 2013). The study identified 409 genes associated with active TB, which were strongly associated with myeloid-derived inflammatory pathways, with the TREM1 signalling pathway representing one of the most significant pathways. Genes involved in the type I/II IFN signalling pathway were identified, but expression differences were not as large as those associated with myeloid and TREM1 signalling pathways (Joosten, Fletcher, and Ottenhoff 2013). A more recent study re-analysed 16 published datasets in an attempt to update the study done by Joosten and colleagues (Blankley et al. 2016). A 380-gene signature, consistently up or down regulated in at least nine datasets, was identified. Five genes (AIM2, BATF2, FCGR1B, Haptoglobin and TLR5) were consistently differentially expressed in all 16 datasets. As observed by Joosten *et al*, this signature consisted mainly of IFN signalling and TREM-1 signalling genes in addition to genes involved in complement activation, pattern recognition and IL-6 signalling pathways.

There is a lack of concordance of signatures of active TB between studies, and significant overlap of transcript signatures with other diseases. Singhania and colleagues recently deposited a pre-peer reviewed article in the BioRxiv repository aimed at identifying a signature specific to active TB (Singhania et al. 2017). In this study, they re-identified the 393-transcript signature developed previously from

microarray data (Berry et al., 2010) by performing RNA sequencing (RNA-seq) on the same cohort. The newly developed 373-gene signature showed significant overlap with the 393-transcript microarray signature. They then applied their new 373-gene signature, the 27 and 44-gene signatures from Kaforou *et al*, (discussed below) and the 16-gene signature from Zak *et al*, (discussed below) to independent published datasets of influenza, bacterial pneumonia, sarcoidosis, lung cancer and other viral and bacterial infections. A 303-gene signature was derived from the 373-gene signature, which was specific to active TB, i.e. not detected in the other diseases. A smaller gene signature (n=20) was then refined, which could differentiate active TB from healthy *Mtb* infection with fewer false positives. The authors attributed this to the improved specificity since this signature did not detect viral and bacterial infections in the published datasets. This study suggests that despite overlap of TB signatures, a type I/II IFN signature highly specific for TB could be discovered if individuals with other diseases that mimic TB are included in the discovery and validation of TB signatures.

The biomarkers and signatures described above were developed using expensive tools such as microarray, which make implementation of tests based on such signatures difficult in clinical settings of TB endemic regions. These biomarkers need to be translated into “clinically useful algorithms, using relatively simple tests” that can be used at point of care (Walzl et al. 2014). Due to the number of genes in many of the TB signatures it would be difficult to translate these signatures into point of care tests. Several studies have assessed smaller signatures using microarrays, some of which have attempted to translate these to allow simpler measurement, such as polymerase chain reaction (PCR) assays. A four-gene signature discovered

in HIV-uninfected persons by microarray and translated into reverse transcriptase PCR (RT-PCR) could distinguish active TB disease from healthy *Mtb* infection with an area under the receiver operating characteristic curve (ROC AUC) of 0.98 (95% CI: 0.97-1.00) (Maertzdorf et al., et al., 2015). The signature, consisting of GBP1, IFITM3, P2RY14, and ID3 genes, was validated in Ugandan and Gambian cohorts, and applied to published microarray datasets and differentiated active TB from healthy *Mtb* infection and other diseases (sarcoidosis, pneumonia, and lung cancer) with excellent sensitivity and specificity (Maertzdorf et al., 2015).

Thirteen differentially expressed genes from the three published datasets of Berry *et al*, were applied to a Thai cohort to differentiate active TB from healthy *Mtb* infection (Satproedprai et al. 2015). A “TB sick score” was calculated from seven of 11 quantifiable genes using predictive logistic regression. The TB sick scores could differentiate active TB disease from *Mtb*-infected healthy controls, however the sample size for this study was small and no external validation cohort was available. These smaller gene signatures can be more easily integrated into clinical settings using battery-operated devices such as gene chip devices, which may become available for point of care applications in the near future (Maertzdorf et al., 2015). One such device is the GeneXpert® Omni, a small and portable battery-operated diagnostic device from Cepheid (Cepheid 2017), which can extract, amplify and detect *Mtb* DNA. Expression of the BATF2 gene alone could distinguish active TB disease from healthy controls and was identified by comparing changes in transcript signatures of active TB patients at diagnosis and two to four years after the end of TB treatment, (Roe et al. 2016). BATF2 expression could also diagnose active TB disease from healthy *Mtb* infection when applied to external datasets. BAFT2,

however, applied to a cohort of individuals with fever, could not differentiate between active TB and fever due to other infectious diseases (Roe et al. 2016).

1.8 Biomarkers of TB in HIV co-infection

TB mortality is higher in HIV-infected patients with smear negative TB compared to HIV-uninfected patients. This is likely due to many causes, but primarily the effect of HIV on immunosuppression, further underestimating disease intensity (Getahun et al. 2007). Low bacillary load in the sputum, as mentioned above, complicate the diagnosis of active TB disease in HIV infection. New tools are needed to diagnose and predict active TB disease in this population. Only very few studies have measured gene signatures for active TB disease in HIV-infected persons. Furthermore, to our knowledge only four published studies have developed transcript signatures aimed at differentiating active TB from other diseases in HIV-infected persons (Anderson et al., 2014; Dawany et al., 2014; Kaforou et al., 2013; Sweeney et al., 2016). Of these, three developed signatures in cohorts including HIV-infected persons (Kaforou et al. 2013; Anderson et al. 2014; Dawany et al. 2014), whilst one developed a signature from publicly available datasets which include both HIV-infected and uninfected persons (Sweeney et al. 2016). Kaforou and colleagues developed a 27-transcript and 44-transcript signature to distinguish active TB from *Mtb* infection and other diseases, respectively (Kaforou et al. 2013). The signatures were developed using samples from both HIV-infected and uninfected participants. Signatures developed by Berry *et al*, were applied to the Kaforou cohort and a lower performance of the Berry signatures was observed, particularly in the HIV-infected participants. Anderson and colleagues discovered a 51-transcript and a 42-transcript signature that could differentiate active TB from other diseases and *Mtb* infection,

respectively in children (Anderson et al. 2014). Currently, only one published study, by Dawany *et al*, has derived a transcriptomic signature from HIV-infected persons only. The study identified 251 differentially expressed genese that could distinguish active TB from controls in HIV-infected persons (Dawany et al. 2014). The signature was applied to published datasets of HIV-uninfected persons and could differentiate participants with active TB from participants with *Mtb* infection with an accuracy of 88.9% to 94.7%. The performance of the 393-gene signature from Berry and colleagues was also tested on this cohort, however it failed to accurately discriminate HIV-infected controls from HIV-infected TB cases (Dawany et al. 2014). Sweeney *et al* developed a three-gene signature from 14 publicly available datasets including HIV-infected persons, which was able to differentiate active TB from *Mtb* infection and other diseases (Sweeney et al. 2016). The signature was developed from the Berry (HIV-uninfected), Bloom (HIV-uninfected) and Kaforou (HIV-infected and uninfected) datasets (Berry et al., 2010; Bloom et al., 2013; Kaforou et al., 2013), and could differentiate active TB from *Mtb* infection and other diseases in both the discovery and validation sets. A lower performance of the Sweeney signature was observed in HIV-infected persons in differentiating active TB from *Mtb* infection ROC AUC, 0.89 (95 % CI: 0.87-0.91) in comparison to HIV-uninfected persons ROC AUC, 0.97 (95 % CI: 0.95-0.98) from the same setting.

Several recently published gene expression signatures have applied signatures developed from HIV-uninfected persons to the Kaforou and Dawany datasets. The four-gene signature of Maertzdorf and colleagues was applied to both datasets (Maertzdorf et al., 2015) and could differentiate active TB from *Mtb* infection in both cohorts and active TB from other diseases in the Kaforou dataset. The one-gene

signature from Roe and colleagues was also applied to the Kaforou HIV-infection dataset. A ROC AUC of 0.84 was observed in this cohort, possibly due to the higher expression of BATF2 in the HIV-infected *Mtb*-infected persons in comparison to their HIV-uninfected (AUC, 0.95) counterparts (Roe et al. 2016). Walter and coworkers applied their signatures to the Kaforou dataset and showed that the signatures could distinguish active TB from *Mtb* infection and other diseases in HIV-infected persons with a performance of 98.0% (95% CI: 97.8% to 98.3%) vs 82.8% (95% CI: 82.4% to 83.2%). A lower specificity of 76.3% vs 87.9% was observed in the HIV-infected cohort in comparison to their discovery cohort of HIV-uninfected persons (Walter et al. 2016). In the above-mentioned studies, due to high gene expression in HIV-infected *Mtb*-infected persons relative to HIV-uninfected controls, an increase in false positives was observed lowering the specificity of the signatures in HIV-infected persons. Hence, the signatures had a decreased specificity which led to consistently poorer performance in HIV-infected persons compared with HIV-uninfected cohorts. Consequently, these studies suggest that HIV-infection decreases the accuracy of transcriptomic signatures developed in HIV-uninfected persons. These signatures are promising for the development of point-of-care screening tools to identify individuals who would be tested further to confirm active TB disease. The Foundation for Innovative New Diagnostics (FIND) and the WHO developed a document describing the product specifications of a screening test and non-sputum based diagnostic tool (World Health Organization 2014). The performance of such tests (signatures) and operational characteristics (sensitivity and specificity) should meet minimal requirements set out in the target product profile (TPP) document to be considered as a triage tool. As a triage test, a diagnostic or triage signature would ideally be used during a patient's first encounter in a health care setting. Hence, it

has to perform with equivalent or greater sensitivity than the confirmatory test, otherwise patients would be missed by the test (Denkinger et al. 2015). The TPP for a screening tool states that it should have a diagnostic sensitivity of at least 95% and a specificity of 80% for culture-positive TB. Most publications of the signatures described above report ROC AUCs but not sensivities and specificities of the signatures, precluding assessment of performance against the TPP. It is important to note that in some of the studies mentioned in the sections above, the investigators aimed to identify gene expression profiles that could diagnose TB disease as well as describe the biological roles of the identified expression profiles in TB. Another limitation in some studies is the lack of validation cohorts making it near-impossible to determine the true performance of the suggested signatures and their true potential as triage tests. Thus, in designing studies for identification and development of triage tools, studies should include blinded validation in independent cohorts and report the performance as sensitivities and specificities as well as positive predictive values and negative predictive values of the signatures in such cohorts.

1.9 Biomarkers of risk of TB (prognostic biomarkers)

The post-2015 end-TB strategy aims to reduce the global TB incidence by 90% and mortality by 95% by the year 2035 (Uplekar et al. 2015). Another aim is a 50% and 75% reduction in incidence and mortality, respectively by 2025 (Uplekar et al. 2015). Elimination of TB by 2050 will require a reduction in the annual TB incidence rate at an average of 20% between 2015 and 2050, which translates to less than 9,000 cases amongst the estimated global population of nine billion people by 2050 (Dye et al. 2013). Attainment of the 2025 end TB strategy is feasible, but only if interventions

are carefully tailored to the epidemiology and existing TB control tools are used to their maximal capacity in individual countries (Houben et al. 2016). However, with current technologies, procedures and services TB elimination by 2050 is impossible (Dye et al. 2013). New technologies that improve diagnosis, treatment and pre-exposure vaccination have the ability to decrease TB incidence by 71% in 2050 (Abu-Raddad et al. 2009). However, such new technologies are still out of reach, emphasising the urgent need for research and development. Screening is still a public health priority in eliminating TB by the diagnosis and treatment of *Mtb* infection (Auguste et al., 2016). Without novel and more effective technologies, early case detection and effective treatment could be the best strategy to reduce TB incidence to achieve the stop-TB strategy by 2050 (Dye et al. 2013).

It is thought that a tool that can identify individuals at risk of developing active TB before they present with symptoms, allowing targeted provision of preventive therapy would be major advance (Dowdy et al., 2008). Mathematical modelling suggests that mass treatment of *Mtb*-infected persons with preventive therapy may help to reduce TB incidence by up to 64% (Abu-Raddad et al., 2009). However, it is not feasible to provide mass preventive therapy to the estimated one-quarter of the global population who are *Mtb*-infected, when only 5% to 10% of infected individuals are likely to progress to disease (Esmail et al., 2014). Although this may prevent many cases of incident TB, adherence to long-course antibiotics is a serious challenge and those who are not at risk of progression would be exposed to the harmful hepatotoxic effects of drugs such as INH.

Identification of those individuals at risk of developing TB disease has traditionally been very challenging (Lin et al. 2016) because individuals that comprise the

enormous *Mtb* infected reservoir are typically asymptomatic, detection of the infecting bacterium is not possible and the infection can last for decades (Rangaka et al. 2015). It is believed that adherence to preventive therapy can be increased if the subset of persons at real risk of progressing to active TB disease are identified (Dodd and Schlesinger 2017). Current *Mtb* infection diagnosis methods are very poor at predicting progression to active TB disease in *Mtb*-infected persons and close contacts of active TB persons. A weak association between positive IGRAs and development of active TB disease, incidence rate ratio, 2.10 (95% CI: 1.42 to 3.08) was observed in a meta-analysis of published data sets (Rangaka et al. 2012).

Biomarkers of risk of active TB disease will allow identification of persons at highest risk of developing TB at an early stage, before they are sick and ideally before transmission of *Mtb* has occurred. This would add to the arsenal of tools for *Mtb* elimination, because biomarkers of risk would allow targeted preventive antimicrobial therapy to prevent TB disease (Ottenhoff et al., 2012; Penn-Nicholson et al., 2016; Salgame et al., 2015). Many studies have discovered and validated blood transcriptomic signatures that differentiate individuals with active TB disease from healthy individuals with or without *Mtb* infection with promising accuracy (Berry et al. 2010; Sweeney et al. 2016; J. Maertzdorf et al. 2011). This body of work has opened the possibility that such diagnostic blood transcriptomic signatures could be developed as triage tests that could identify persons with TB for more intensive investigation. These advances also provide feasibility and impetus to determine if prognostic blood transcriptomic signatures can be identified that can predict progression to incident TB. Monocyte/lymphocyte ratios were associated with in vitro mycobacterial growth and transcriptional signatures of monocytes (Naranbhai et al.

2015), suggesting an association between disease progression and monocyte/lymphocyte ratios. Additionally, activated CD4 T cells have been associated with an increased risk of progression to active TB disease in HIV-uninfected South African adolescents (Fletcher et al. 2016). To date two studies have attempted to identify transcriptomic signatures that could predict active TB before the onset of disease (Sloot et al. 2015; Zak et al. 2016). Sloot and colleagues identified 50 genes that were differentially expressed in HIV-infected persons who developed active TB and controls who remained healthy, at least 180 days before diagnosis (Sloot et al. 2015). This study was cross-sectional, with a poorly characterised small cohort of 14 cases and 15 controls and the signature was not validated in an independent cohort. In a landmark study my colleagues at SATVI collaborated with investigators at the Center for Infectious Disease Research to identify a 16-gene, 57 transcript, signature by RNA-Seq that could predict active TB disease in a longitudinal cohort of South African adolescents more than a year before onset of disease (Zak et al. 2016). This signature is discussed in more detail in the section below.

1.9.1 The ACS 16-gene signature (Zak et al., 2016)

The adolescent cohort study was conducted in South African adolescents who were followed-up for two years to determine the prevalence and incidence of *Mtb* infection, TB disease and risk factors for developing TB (Mahomed et al. 2013). Healthy adolescents (n=6,363) were recruited at 11 high schools from the Worcester area outside Cape Town. Participants were allocated to two groups with different follow-up methods. Those in the active follow-up group were assessed every 3 months and had blood samples drawn every six months for QFT, PAXgene, plasma and PBMC

collection. Those in the passive follow-up group were assessed and had blood collected at baseline and two years after enrolment. The surveillance team accessed clinical records to identify TB cases diagnosed in the public health sector for both groups. Whole blood was collected in PAXgene tubes from each study participant at the follow-up time points in addition to clinical data. Microbiologically confirmed active TB disease (two consecutive sputum smear tests or one sputum culture positive test) was diagnosed during follow-up in 87 adolescents in total, and 46 qualified as progressors for our biomarker study (Zak et al. 2016). The progressors were matched to 107 healthy *Mtb*-infected controls by ethnicity, school, gender, age and prior episodes of TB disease at enrolment, and were referred to as non-progressors. The progressors and non-progressors were split into a training set for signature discovery and a test set for blind validation at a ratio of 3:1. All participants included in this analysis were *Mtb*-infected (QFT and/or TST positive) at baseline/enrolment.

RNA-sequencing was used to measure transcript expression in this cohort by comparing the progressors and non-progressors. Unlike microarrays, RNA-sequencing does not rely on pre-defined probe-based sequences, thereby allowing less biased identification of novel transcripts which were previously unknown (Blankley et al. 2014). Support vector machines (SVMs), a generalised linear classifier, was used to generate the signature in this population (Zak et al. 2016). This algorithm “maps data points onto a multidimensional space in such a way that it is possible to calculate a hyperplane (a multidimensional plane), which separates the classes of samples” (Maertzdorf et al., 2015). Multiple pair-wise ensemble models (representing two transcripts each) that could predict TB disease risk based on

splice junction counts measured by RNA-seq were developed. A transcript signature consisting of 16 differentially expressed interferon response genes representing 47 splice junctions plus 10 reference splice junctions could predict progression to TB disease in the progressors (Figure 1). The signature, henceforth referred to as the ACS 16-gene signature, was adapted to microfluidic quantitative reverse transcription polymerase chain reaction (qRT-PCR) to reproduce the findings using a different measurement platform, allow more high-throughput and cheaper measurement, and to prepare for translation to near-patient testing. Splice junctions from the RNA-seq data were matched to TaqMan primer-probe sets for this and the SVM models were re-parameterized to the PCR data.

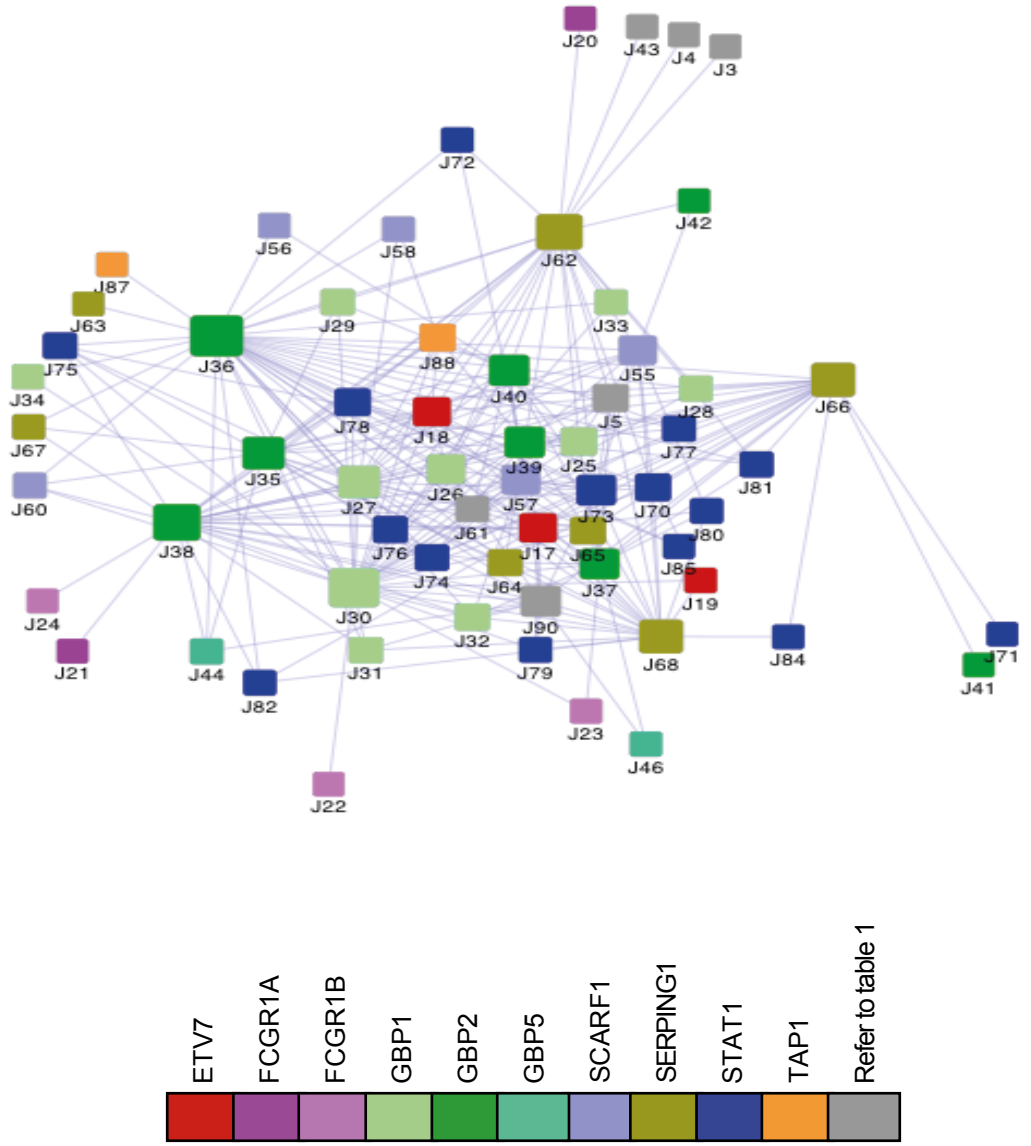


Figure 1: Network visualisation of the transcript pairs in the ACS 16-gene signature (adapted from Zak et al., 2016). The nodes represent splice junctions and the colours indicated genes represented by the splice junction.

The PCR-based signature model was fitted to the Ct value data in the training set. The qRT-PCR signature consisted of 47 genes of interest (GOI) primer-probes, representing 247 gene-pairs and 10 reference primer-probes for standardization (to calculate delta Ct values, Table 1). The signature is comprised of the following genes: FCGR1B, FCGR1A, STAT1, GBP1, GBP2, GBP4, GBP5, SERPING1, ETV7, BATF2, SCARF1, APOL1, TAP1, TRAFD1, ANKRD22 and SEPT4. Signature scores are calculated by summing the product of normalised abundances (delta Ct values) of two gene products (a gene-pair) with their coefficients and a coefficient of the gene-pair. Custom scripts, written in R, were applied on the results generated from the microfluidic qRT-PCR to calculate a single signature score per sample, expressed as a percentage. The signature scores are directly correlated with gene expression, thus an increasing signature score translates into a high gene expression. A higher score thus implies higher gene expression. The various transcripts could contribute differently to different sample types and signature scores would remain high in all these samples. The proportion of transcript-pairs voting a case relative to the total number of voting pairs is what determines the signature scores.

In the training set, the signature score in progressors increased closer to TB diagnosis. The signature could predict active TB disease up to a year before TB diagnosis with a sensitivity of 71.2% (ROC AUC, 0.79, 95% CI: 0.76 to 0.82) and 62.9% (ROC AUC, 0.77, 95% CI: 0.75 to 0.79) at 1-180 days and 181-360 days before TB diagnosis, respectively. RT-PCR data from the test set cohort were used to perform blind predictions, which demonstrated that the signature could predict

active TB in the test cohort up to a year before diagnosis with a sensitivity of 66% and specificity of 81%.

For external validation, the signature was applied to blinded samples from an independent cohort of household contacts of active TB disease from the Grand Challenges 6-74 (GC6-74) cohort. Samples from 73 progressors and 301 non-progressors from South Africa and Gambia were evaluated. The signature could predict active TB in this cohort up to a year before diagnosis with a sensitivity and specificity of 54% and 83%, respectively, with a ROC AUC of 0.72 (95% CI please). In addition, the signature was reparameterised to publicly available microarray datasets (Anderson et al., 2014; Berry et al., 2010; Bloom et al., 2012, 2013; Kaforou et al., 2013) to determine the diagnostic performance of the ACS 16-gene signature for discriminating active TB from uninfected or *Mtb*-infected controls or controls with other respiratory diseases. The signature performed well and showed excellent promise as a diagnostic and prognostic biomarker for TB. A more practical version of this signature (ACS 11-gene, described below) consisting of 48 transcripts was tested in this thesis.

1.9.2 The ACS 11-gene signature: a more practical PCR-based signature derived from the ACS 16-gene signature

Measurement of the 57 primer-probes (47 primer-probes and 10 reference primer-probes) that comprise the ACS 16-gene signature by microfluidic qRT-PCR is ideally done using the 96.96 gene expression platform with the BioMark HD instrument (Fluidigm). To improve throughput and practicality of measurement, we set out to remove nine primer-probes to allow 48 primer-probes to be measured in duplicate on

the 96 by 96 platform. This would allow simultaneous testing of 96 samples in a 96.96 gene expression chip. Primer-probes were removed based on three main criteria: 1) reproducibility in qRT-PCR assays, 2) redundant signal to overall signature and 3) number of pairs formed with other transcripts (Table 1) (Darboe et al., 2018, Appendix 1). Nine IFN response primer-probes were removed, representing ANKRD22, APOL1, FCGR1A, GBP4 and SEPT4 and the 10 reference genes were retained. The remaining 38 primer-probes represented 11 type I IFN genes and formed 231 gene-pairs. This new signature was referred to as the ACS 11-gene signature. The ACS 11-gene signature scores are calculated in the same way as the ACS 16-gene signature scores, as described above. New transcript pairs were not chosen during signature development, transcript pairs from the 16-gene signature were maintained. A pair was only removed if at least one of the transcripts in the pair was excluded from the signature.

To determine how reduction of the ACS 11-gene signature affected the signature we compared signature scores and prognostic performance to that of the ACS 16-gene signature in the ACS training and test cohort samples (progressors vs non-progressors). Signature scores were identical between the ACS 11 and 16-gene signatures in the ACS cohort (Figure 2A). In addition, the prognostic performance of the two signatures was equivalent in the ACS cohort (Figure 2B).

Table 1: Primer-probes included in the ACS 16-gene and 11-gene signatures

TaqMan primer probe in the ACS 16-gene signature	Number of gene pairs	Excluded from the ACS 11-gene signature
ACTR3.Hs01029159_g1	Reference	
ADRBK1.Hs01056345_g1	Reference	
CDC42.Hs03044122_g1	Reference	
CSDE1.Hs00918650_m1	Reference	
CYTIP.Hs00188734_m1	Reference	
TMBIM6.Hs01012081_m1	Reference	
TMBIM6.Hs00162661_m1	Reference	
TMBIM6.Hs01012082_g1	Reference	
TPM3.Hs01900726_g1	Reference	
USF2.Hs01100994_g1	Reference	
ANKRD22-j2	1	X
APOL1.Hs00358603_g1	1	X
FCGR1A.Hs02340030_m1	1	X
GBP4.Hs00925073_m1	1	X
STAT1.Hs01014001_m1	1	X
TAP1.Hs00388675_m1	1	
GBP5.Hs00369472_m1	2	
STAT1.Hs01014006_m1	2	X
STAT1.Hs01014008_m1	2	X
ETV7.Hs00903228_m1	3	
FCGR1C.Hs00417598_m1	3	
GBP2.Hs00894840_mH	3	
SCARF1.Hs01092482_g1	3	
STAT1.Hs01013990_m1	3	X
STAT1.Hs01014000_m1	3	
GBP5-j4	4	
SEPT4.Hs00910208_g1	4	X
STAT1.Hs01013989_m1	4	
STAT1.Hs01013991_m1	4	
STAT1.Hs01013992_g1	4	
STAT1.Hs01013994_m1	5	
STAT1.Hs01013996_m1	6	
STAT1.Hs01013997_m1	6	
SCARF1.Hs00186503_m1	7	
STAT1.Hs01013995_g1	7	
BATF2.Hs00912736_m1	8	
SCARF1.Hs01092483_m1	8	
SERPING1.Hs00163781_m1	9	
SERPING1.Hs00934328_g1	9	
STAT1.Hs01013993_m1	9	
STAT1.Hs01014002_m1	9	
TAP1.Hs00897093_g1	9	
ETV7-j2	10	
ETV7.Hs00903230_g1	10	
SCARF1.Hs01092485_g1	12	
TRAFD1.Hs00938765_m1	14	
STAT1.Hs01013998_m1	15	
GBP2-j1	16	
GBP2.Hs00894842_g1	16	
SERPING1.Hs00934329_m1	19	
SERPING1.Hs00934330_m1	22	
SERPING1.Hs00935959_m1	23	
GBP1.Hs00977005_m1	25	

GBP1-j1	32	
GBP2.Hs00894837_m1	33	
GBP1.Hs00266717_m1	40	
GBP2.Hs00894846_g1	64	

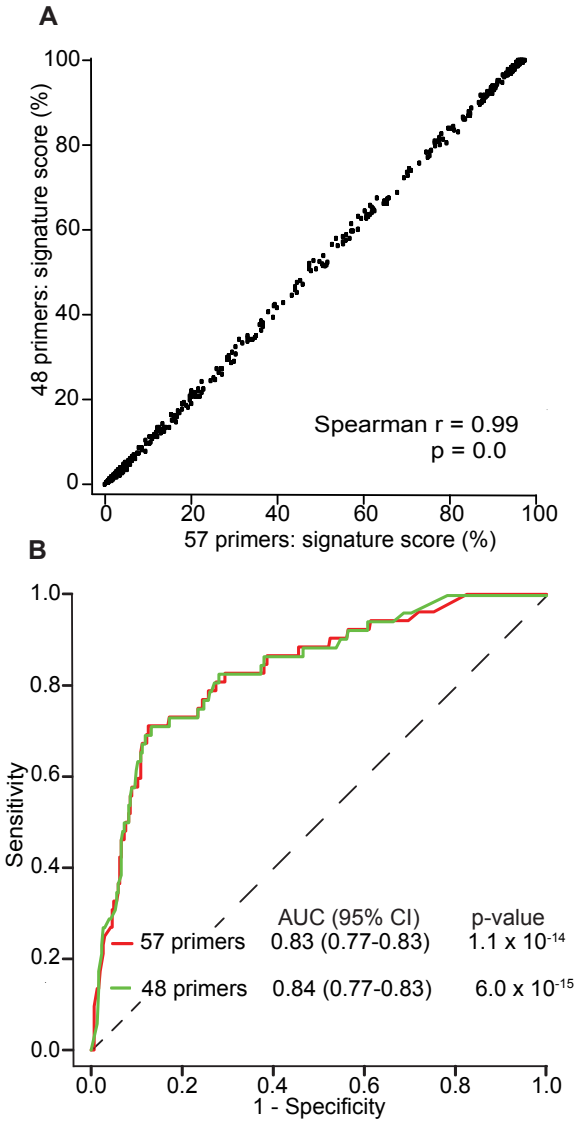


Figure 2: Comparison of signature scores and prognostic performance of the ACS 16 and 11-gene signatures in the ACS cohort. A) Correlation of signature scores between the ACS 16 (57 primer-probe) and 11-gene (48 primer-probe) signatures. The Spearman correlation coefficient is shown. B) Receiver Operating Characteristic curves depicting performance of the signatures in classifying progressors from non-progressors.

1.10 Biomarkers of treatment success in the presence and or absence of HIV infection

Adherence to a long-term TB treatment is a major problem in the control of TB, as poor adherence has been observed to be one of the main factors leading to drug resistance. Biomarkers of treatment success and or failure would allow better treatment monitoring, customisation of treatment duration to individuals and detection of poor treatment response much earlier. These advances may reduce non-adherence and accelerate the recognition of poor outcomes before patients complete treatment (Walzl et al. 2014). In addition, they would allow ineffective treatment regimens to be adjusted in individuals not responding to treatment. An added potential advantage of treatment response markers is that they may inform shortening of treatment regimens in persons who are “cured” well before the traditional 6-month regimen is completed. Currently, sputum conversion from a positive to a negative culture at two months is used as an early marker of treatment response. This takes very long because a negative culture result requires up to 42 days, and it a poor predictor of relapse. A meta-analysis of sputum conversion at two months observed a pooled sensitivity and specificity of 40% and 80% respectively for predicting relapse (Horne et al. 2010).

Biomarkers of TB treatment response would allow more rapid and potentially more accurate prediction of cure. Expression of active TB disease transcriptional signatures have been shown to decrease during TB treatment, suggesting that blood transcriptional signatures may be used to monitor TB treatment (Berry et al., 2010; Bloom et al., 2012; Ottenhoff et al., 2012). The 393-transcriptomic signature identified by Berry et al. diminished in active TB patients by two months after start of

TB treatment (Berry et al., 2010). In addition, the signature was shown to be strongly associated with extent of disease on chest radiography. A 664-transcript signature derived to distinguish active TB disease from *Mtb* infection by Bloom et al. decreased during TB treatment as early as two weeks after initiation of therapy (Bloom et al. 2012). Similarly, a PBMC-derived transcriptomic signature decreased during TB treatment, with a similar profile observed at 28 weeks between "cured" participants and healthy *Mtb*-infected controls (Ottenhoff et al., 2012). Another study aimed at deriving a transcriptomic signature of TB treatment reported a perturbation of 1,261 genes in the first week of TB treatment (Cliff et al. 2013). Of these, 780 genes remained differentially expressed after two and four weeks of therapy in comparison to at diagnosis. Genes involved in the IFN pathway were downregulated during TB treatment with the most significant changes occurring in the first week of therapy (Cliff et al. 2013). On the other hand, genes involved in the T cell, B cell and cytotoxic cells were upregulated at the end of 26 weeks of therapy, with the B cell modules being the most significant fold change from baseline. In the same study, a signature of 62 genes forming a neural network could determine the stage of TB therapy i.e. diagnosis from on-going treatment and end of treatment. Similarly, in a non-human primates (NHP) study of *Mtb*-infected cynomolgus macaques an association between transcriptomic signatures and bacterial load was observed that suggested blood IFN signatures may reflect pulmonary *Mtb* replication (Gideon et al., 2016).

We also showed that ACS-16 gene signature scores decreased during TB treatment in HIV-uninfected persons (Thompson et al. 2017). The signature could differentiate study participants based on their rate of bacteriological conversion measured by

sputum cultures. These data affirm that the ACS 16-gene signature could be a useful biomarker for predicting relapse and TB treatment failure.

1.10 Biomarkers of risk of recurrent TB disease in the presence and or absence of HIV infection

A history of previous TB is a risk factor for recurrent TB disease. Similarly, HIV infection increases risk of recurrent TB disease. Despite the risk of recurrent TB, only a few transcriptomic signatures have thus far been developed to identify individuals at risk of recurrent TB disease. Mistry *et al* derived a signature which could differentiate individuals who have completed antimycobacterial therapy from persons with recurrent TB in a cross-sectional study (Mistry et al. 2007). Another study identified 18 differentially expressed genes between relapse and “cured” HIV-uninfected participants in a longitudinal cohort (Cliff et al., 2015). Genes involved in cell-mediated cytotoxicity were highly expressed in samples from relapse patients in comparison to the “cured” group.

1.11 Type I IFN: Activation and role in active TB disease and HIV infection pathogenesis

Most of the gene signatures of TB disease described above, including the ACS 16 and 11-gene signatures comprise type I/II IFN transcripts. IFNs are cytokines that are classified into three classes (types I, II and III), with varying anti-viral and growth inhibitory functions, receptor usage, and structural features (Donnelly and Kotenko 2010). The type I IFN family encodes 16 homologous IFNα subtypes in humans, an IFNβ gene and several genes encoding for IFNε, IFNτ, IFNκ, IFNω and IFNδ (Chen et al., 2017; McNab et al., 2015). Of these, IFNα and IFNβ are the most widely

expressed and best-defined type I IFN sub-types, which have the ability to induce gene transcription programmes resulting in interference with multiple stages of viral replication cycles. Type II IFN, on the other hand, consist of only IFNγ, whilst type III IFNs consists of the IFNλ family. In this thesis, type I IFN generally refers to IFNα and IFNβ and type I/II IFN refers to both IFNα and IFNβ in combination with IFNγ.

Type I IFNs together with proinflammatory cytokines are the first effector molecules induced in the innate immune response against certain pathogens (Abel et al., 2002). They are expressed by multiple cell types in response to stimulation of innate immune cell receptors, such as toll like receptors (TLRs), or as a response to viral infections (McNab et al., 2015). *Mtb* can also induce type I IFN expression in macrophages and dendritic cells (DCs) via the TANK-binding kinase 1 (TBK1), a serine/threonine kinase (Donovan et al., 2017). TBK1 is activated by pathogen recognition receptors such as TLR4 and the Nod-like receptor, NOD2. Interferon-regulatory factors (IRFs) are required for the activation of interferon stimulated genes (ISGs); TBK1 facilitates the phosphorylation of IRFs in *Mtb*-infected cells inducing the production of IFNα and IFNβ (Donovan et al. 2017). Type I IFN signalling due to *Mtb* infection can hinder the host's immune response in controlling replicating bacteria. Type I IFN inhibits macrophage killing of *Mtb* in mice by impeding the production of host protective cytokines such as TNFα, IL-1β and IL-12, whilst promoting the production of IL-10 (McNab et al., 2014). In addition, type I IFNs have the ability to prevent macrophage killing by hindering IFNγ-mediated-growth inhibition of *Mtb*-infected macrophages (McNab et al., 2014). IFNβ prevented the induction of IFNγ in human monocytes and macrophages infected with *Mycobacterium leprae*, which was reversed by the blocking of IL-10 (Teles et al.

2013). Thus, type I IFN can exacerbate mycobacterial replication by inhibiting the protective action of the type II IFN, IFNγ, by modulating IL-10 induction. In contrast, Moreira-Toxeira and colleagues observed a protective role of type I IFN by *Mtb*-infected murine macrophages through the induction of nitric oxide synthase 2 and inhibition of arginase 1 gene expression thereby controlling bacterial replication (Moreira-Teixeira et al. 2016). Through this mechanism, induction of type I IFN seems to be able to provide protection to the murine host, particularly in the absence of IFNγ (Moreira-Teixeira et al. 2016). In humans, type I IFN signalling in the form of strongly upregulated expression of ISGs, has been observed in active TB cases relative to healthy controls, as discussed above. Blood leukocyte levels of ISG transcript expression have been associated with disease severity, progression to disease and treatment response, suggesting that type I/II IFN induction in humans is associated with bacterial replication and plays a role in exacerbating disease. It should be noted though that elevated type I IFN protein levels in peripheral blood have not been reported in persons with TB disease.

A well-known function of type I IFNs is their protective role against viral infections. In HIV-infection the main cellular producers of IFNα are plasmacytoid dendritic cells (pDCs), which constitute <1% of blood leukocytes, and can produce up to a 1,000-fold more IFNs than the other immune cell types (Herbeuval and Shearer 2007). IFNβ can be produced by all DCs (Chen, Liu, and Cao 2017). In response to HIV infection, some TLRs can induce IFN signalling via the TIR-domain-containing adaptor protein inducing IFNβ (TRIF) or the myeloid differentiation primary-response protein 88 (MyD88) dependent pathways (Gonzalez-Navajas et al. 2012). TLRs 3 and 4 on pDCs can induce IFN expression via the TRIF pathway in a wide variety of

cells, whilst TLRs 7, 8 and 9 induce IFNs in DCs through the MyD88-dependent pathway (Gonzalez-Navajas et al. 2012) (Figure 3). Induction of type I IFNs has been observed in acute HIV infection (Bosinger and Utay 2015) and type I IFN levels are positively associated with HIV progression (Lehmann et al. 2008). In Simian Immunodeficiency Virus (SIV) infection, type I IFN was observed to have a protective effect against disease progression (Sandler et al. 2014). In this study, blocking of the IFN alpha-receptor (IFNAR) signalling during acute infection resulted in the reduction of antiviral gene expression in two rhesus macaques. Despite a decrease in T cell activation, SIV reservoir was increased and CD4 T-cell depletion was accelerated and an increased rate of progression to AIDS was observed. In acute HIV infection, up to 159 days after presumed infection, *in vitro* type I IFN production in response to herpes simplex virus stimulation was impaired in HIV-infected persons relative to healthy controls (Kamga et al. 2005). Furthermore, viremia in ART-naïve participants with acute HIV infection was inversely associated with IFN production, suggesting that type I IFN is highly associated with viral control. This, however, was not observed in those who received ART. These studies suggest that type I IFN production protects not only against disease progression but can also protect against other opportunistic infections.

IFNα and IFNβ bind and signal through the heterodimeric IFN alpha-receptor subunits (IFNAR1 and 2), which are expressed on all cell types (Chen et al., 2017; McNab et al., 2015). Transcription of interferon stimulated genes (ISGs) primarily occurs upon the activation of the Janus Kinase 1 (JAK1) and tyrosine kinase 2 (TYK2), which are receptor-associated protein kinases (McNab et al., 2015). JAK1 and TYK2 are activated by the ligation of the IFNAR. Activation of the JAK1 and

TYK2 can result in type I IFN-mediated signalling through several mechanisms. In the canonical pathway, phosphorylation of STAT1 and STAT2 heterodimers leads to the formation of the ISG factor 3 (ISGF3) complex. This occurs via the binding of dimerized STAT1 and STAT2 heterodimers to IFN-regulatory factor (IRF) 9 (Figure 3). The newly formed ISGF3 translocates to the nucleus where it binds to the DNA sequence motif, IFN-stimulated response elements (ISRE, Figure 3) (Chen, Liu, and Cao 2017). This binding activates the transcription and expression of ISGs (Chen, Liu, and Cao 2017; Ivashkiv and Donlin 2015). In addition, type I IFNs can signal through STAT1 homodimers which are associated with other cytokine-mediated signalling pathways, as well as through the mitogen-activated protein kinase pathways (Ivashkiv and Donlin 2015) (Figure 3).

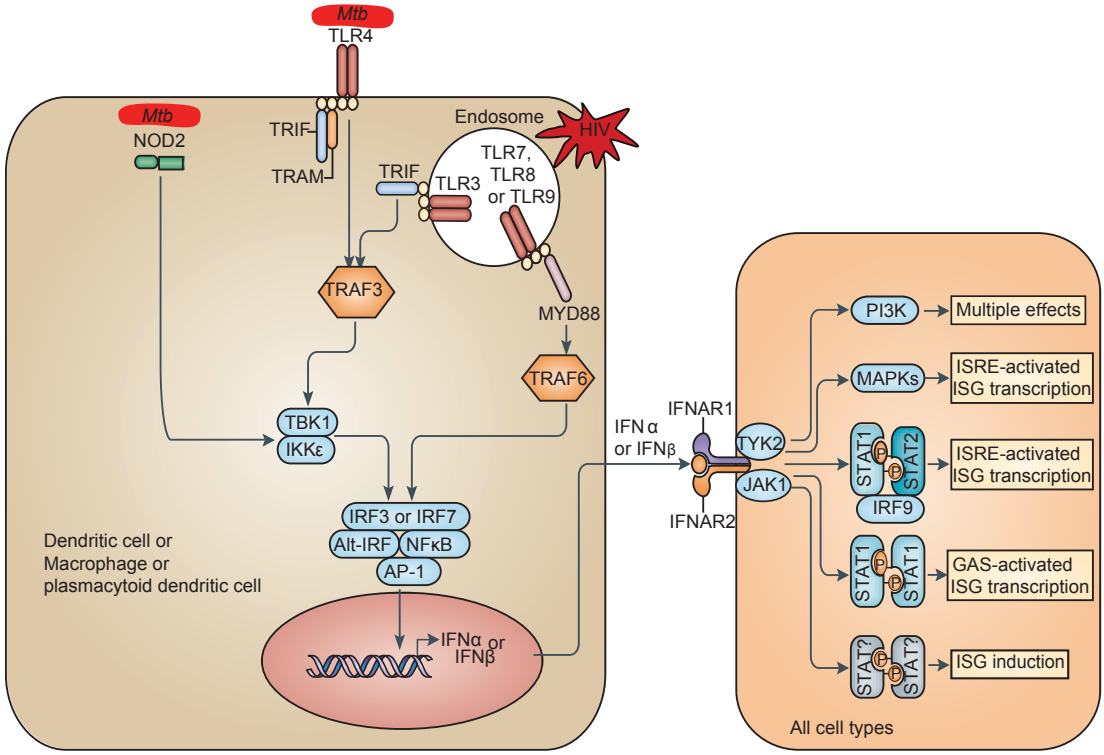


Figure 3: Type I IFN induction and signalling pathways (adapted from McNab et al., 2015). HIV and *Mtb* are recognised by a range of cell receptors on dendritic cells and macrophages (*Mtb*) and plasmacytoid dendritic cells (HIV). Upon induction, type I IFN bind to IFN-alpha receptors 1 and 2 activating type I IFN signalling pathways, leading to the transcription of interferon stimulated genes (ISGs).

In HIV-infected humans, ISG transcript levels were higher in participants with low CD4 T cell counts (<490 cells/μL) in comparison to HIV-infected persons with high CD4 T cell counts and HIV-uninfected healthy controls (Fernandez et al. 2011). The ISG levels in healthy controls were comparable to the levels in HIV-infected participants with high CD4 T cell counts. This study suggests that ISG expression normalisation may be associated with CD4 T cell reconstitution. This was probably virus driven, as CD4 T cell counts are biomarkers for viral replication-associated immunosuppression however, plasma viral load levels were not investigated in this study. Since HIV-infected persons with low CD4 T cell counts and high plasma viral loads are at a higher risk of developing active TB disease transcriptomic signatures of active TB disease mainly comprise ISGs, it is critical that the behaviour of IFN-based transcriptomic signatures of TB is studies in HIV-infected persons.

Therefore, in this thesis we explored the performance of ISG-based signatures, the ACS 11-gene signature, the rPSVM.1 signature and the ACS 6-gene signature at different stages in the TB timeline in HIV-infected persons.

1.12 Objectives of this thesis

The main aim of this thesis is to determine if the ACS 11-gene signature can predict recurrent TB disease and monitor treatment response in HIV-infected persons on ART. The signature was analysed in two prospective observational cohorts. In one of the cohorts, only PBMC samples were available for our analysis, thus we first determined if the ACS 11-gene and the ACS 6-gene (described in Chapter 7) signatures could be measured using cryopreserved PBMC.

Aims

I. To compare the diagnostic performance of the ACS 11-gene signature for TB disease from *Mtb* infection in HIV-infected and uninfected persons.

II. To determine if the ACS 11-gene signature can predict recurrent TB disease in HIV-infected persons on ART.

III. To determine if the ACS 11-gene signature can monitor retreatment response in HIV-infected persons on ART.

IV. To determine the diagnostic performance of a newly identified smaller ACS 6-gene signature, and if it can predict recurrent TB disease and monitor treatment response in HIV-infected persons on ART.

Hypotheses

I. Diagnostic performance of the ACS 11-gene signature for active TB disease is equivalent in HIV-infected and uninfected persons.

II. The ACS 11-gene signature has predictive utility for recurrent TB disease in HIV-infected persons on ART.

III. The ACS 11-gene signature will monitor treatment response in HIV-infected retreatment patients by differentiating early treatment responders from late treatment responders after two months of TB treatment.

IV. The newly identified small ACS 6-gene signature has predictive utility for recurrent TB disease and will differentiate early treatment responders from late treatment responders in HIV-infected persons on ART. The performance of the ACS 6-gene signature will be equivalent to the performance of the ACS 11-gene signature.

Chapter 2: Materials and methods

The aim of this chapter is to describe the methods used across the different aims in this thesis. Details of each human participant cohort and methods specific for each individual chapter are included in the results chapters.

2.1 Institutional ethics review

This PhD study was reviewed and approved by the Human Research Ethics Committee of the University of Cape Town Faculty of Health Sciences and assigned the ethics number HREC 761/2015. Details of the ethical considerations and review committees for human participation in each individual cohort (Table 2) are mentioned in the chapters where they are first described.

Table 2: HIV status and sample types of participant cohorts analysed in this thesis. TRuTH = TB recurrence upon treatment with HAART, IMPRESS = Improving Retreatment Success, PBMC = Peripheral blood mononuclear cells

Cohort	HIV status of study participants	Sample type analysed	Specific results chapter
Pilot study	Infected and uninfected	Whole blood and PBMC	Chapter 3
TRuTH	Infected	PBMC	Chapter 4
Esmail	Infected	Whole blood	Chapter 4
IMPRESS	Infected	Whole blood	Chapter 5

2.2: RNA extraction, processing and storage

2.2.1 Extraction from PAXgene tubes

Frozen PAXgene™ (Qiagen, Germany) tubes were left overnight at room temperature, in order to achieve complete lysis of samples. The tubes were centrifuged for ten minutes at 4,000 relative centrifugal force (rcf, g) to pellet the blood cells after which the supernatant was decanted. Pellets were resuspended in

4mL of RNAse free water and spun for 10 minutes at 4,000 rcf and supernatants decanted. RNA was either extracted manually with the PAXgene™ blood RNA kit (Qiagen, Germany) or the SimplyAmp extraction kit (Promega, USA) on an automated platform using the Freedom EVO 150 (Tecan, Switzerland).

RNA was extracted using the PAXgene™ blood RNA kit as per the manufacturer's instructions. Briefly, pellets were resuspended in the resuspension buffer (BR1) and pellets dissolved by vigorous vortexing. Binding buffer (BR2) and Proteinase K were added to samples upon transfer into 1.5mL Eppendorf tubes. Samples were incubated for 10 minutes at 55°C before transfer into QIAshredder columns and spun for three minutes at maximum speed (18,000 rcf) in a microfuge. To precipitate RNA, 96-100% ethanol (Sigma Aldrich, USA) was added to the flow-through after it was transferred into new Eppendorf tubes. Samples were transferred into PAXgene columns and spun for a minute at 8,000 rcf. BR3 wash buffer was added to the columns before spinning for a minute at 8,000 rcf. DNase I solution was added directly to the spin columns and incubated at room temperature for 15 minutes. The column was spun again using BR3 buffer and the flow-through was discarded. BR4 wash buffer was added to the columns before the columns were spun for a minute at 8,000 rcf. A second wash with BR4 was done and samples were spun for a further four minutes at maximum speed. To remove residual wash buffer from the columns, the empty columns were spun for a minute at maximum speed. The mRNA was eluted twice by adding 40µL of elution buffer BR5 and spinning for a minute at 8,000 rcf. The eluate was incubated at 65°C for five minutes. Upon completion of incubation the samples were immediately placed on ice and aliquoted into four vials before storage at -80°C.

2.2.2 RNA extraction from PBMC

Cryopreserved PBMC were placed on dry ice and thawed at 37°C with shaking in a water bath before placing on wet ice. In a biosafety hood, cells were added into 10 mL of Roswell Park Memorial Institute media (RPMI; Sigma Aldrich, USA) media already warmed at 37°C. PBMC vials were rinsed with 1mL of warmed RPMI and the tubes centrifuged at 400 rcf for 10 minutes. Supernatants were decanted and pellets placed on ice.

Thawed PBMC samples were extracted using the Qiagen RNeasy Plus Mini extraction kits (Qiagen, Germany) as per the manufacturer's instructions. Briefly, one percent 2-mercaptoethanol (Sigma-Aldrich, USA) in lysis buffer RLT was added to the samples and vortexed vigorously for proper disruption of cell walls. Lysates were added to QIAshredder columns (Qiagen, Germany) and centrifuged for two minutes at full speed (18,000 rcf). The homogenised lysate was added to a genomic DNA (gDNA) eliminator spin column placed in a 2mL collection tube and centrifuged for 30 seconds at 8,000 rcf. Seventy percent ethanol (Sigma Aldrich, USA) was added to the flow-through and samples were transferred into RNeasy MinElute spin columns before spinning at 8,000 rcf for 15 seconds. Wash buffer RW1 was added to the columns and centrifuged at 8,000 rcf for 15 seconds. Two more wash steps were carried out by adding another wash buffer, RPE, to the columns twice and centrifuging for 15 seconds at 8,000 rcf. Columns were placed into new collection tubes and spun for five minutes at full speed to remove any residual wash buffer. To elute the RNA, 40µL of RNase-Free water was added directly to the centre of the spin column membrane, twice, before centrifuging for a minute at full speed. Eluted RNA was aliquoted into four tubes for storage at -80°C.

2.2.3 RNA quantification

Extracted RNA was quantified using the Nanodrop ND2000 spectrophotometer (Thermo Fisher scientific, USA). Briefly, RNA background was measured using the elution buffer (BR5 or RNase-Free water). RNA concentration and purity were measured at a 260/230 wavelength and 230/280 wavelength, respectively. All RNA samples were processed regardless of concentration.

2.3 Measurement of gene expression

Biomark HD multiplex microfluidic real time quantitative polymerase chain reaction (qRT-PCR) technology on the Fluidigm (Fluidigm, US) platform was used to measure gene expression.

2.3.1 cDNA synthesis and pre-amplification PCR

Complementary deoxyribonucleic acid (cDNA) was synthesized from RNA in 96-well plates. Incorporated into the 96-well plates were two controls: a control with no sample added (non-template control,NTC) and a control with no reverse transcriptase added (no reverse transcriptase, NoRT). The NoRT control was to test that there was no genomic DNA contamination in the samples and that the gene expression results were not confounded by genomic DNA amplification. cDNA synthesis of the samples was done in singlets in 96-well plates (Applied Biosystems, USA). A master mix of 500µg/ml of oligonucleotide (Oligo (dT)) 12-18 primer (Thermo Fisher Scientific, USA), 10mM of deoxynucleotide triphosphates (Sigma Aldrich, USA), sterile endotoxin free water (Sigma Aldrich, US) was added to 3µL of RNA (±100ng), to initiate cDNA synthesis. Contents of the mix were briefly centrifuged and incubated at 65°C for five minutes in a T100 thermocycler (Biorad,

Belgium) and put on hold at 4°C. Another master mix containing 5x First-strand buffer (Invitrogen, USA), 0.1M of 1,4-Dithiothreitol (DTT) (Invitrogen, USA), 40 units/µL of RNAseOUT (Invitrogen, USA) and sterile endotoxin free water (Sigma Adrich, USA) was added to the samples after a brief centrifugation and incubated at 42°C for two minutes. Two hundred units of SuperScript II reverse transcriptase (Invitrogen, USA) was added to the samples and the samples further incubated at 42°C for 50 minutes, and then the reaction was inactivated at 70°C for 20 minutes. Samples were diluted to a final concentration of ±1ng/µL by adding 80µL of sterile endotoxin free water.

Each cDNA sample is split into 96 microchambers of a 96.96 Fluidigm gene Expression (GE) Array chip (Fluidigm, USA) to mix with each of the 96 primer-probe assays, consequently allowing 96 separate qRT-PCR reactions from each cDNA sample. The Fluidigm 96.96 GE chip technology requires sufficient cDNA template in each of the 96 microchambers to ensure optimal amplification of each primer-probe set (Dominguez et al. 2013). Thus, to ensure consistent distribution of cDNA template into each of the 96 microchambers, a preamplification of synthesized cDNA using a pool of the relevant primer-probe sets of interest is necessary before loading samples on the Fluidigm GE chip. Hence, transcripts of interest were amplified by conventional PCR using a mix of 2x PCR master mix (Thermo Fisher Scientific, USA) and 0.2x of the primer probe master mix of the TaqMan primer probe assays of interest (Thermo Fisher Scientific, USA). Final volume of 10µL of the mix was added in new 96-well plates and plates were sealed with strips of dome caps. The thermocycler was set to the following parameters for the PCR: 95°C for 10 minutes, 95°C for 15 seconds and 60°C for four minutes. Even distribution of mRNA in every

chamber requires at least 4,906 amplicons from 12 preamplification cycles (Dominguez et al. 2013). Therefore, steps two and three were run sixteen times (16 cycles) before the reaction was put on hold at 4°C. Samples were diluted 1:25 with endotoxin free water by adding 240µL per well. The samples generated from this step will henceforth be referred to as pre-amplified cDNA.

A confirmatory RT-PCR was performed using one of the reference primers to ensure that the samples were pre-amplified and that the reagents and the negative controls were not contaminated. The confirmatory RT-PCR consisted of two samples of interest, duplicate NTC and NoRT, and was run on the Rotor-Gene 6000 series RT-PCR instrument (Corbett life science). The RT-PCR was run using the following setting: 50°C for two minutes, 95°C for 10 minutes and for 40 cycles of 95°C for 15 seconds and 60°C for one minute. Preamplification of cDNA was repeated if amplification was detected in the NTC and/or NoRT samples, indicating a false positive result or contamination. If either the NTC or NoRT yielded amplified product a second time suggesting that the contamination occurred during cDNA synthesis, the assay ws repeated from the cDNA synthesis step. If the samples selected for the confirmatory RT-PCR were not amplified in addition to both controls, the confirmation RT-PCR was repeated to ensure that samples were amplified.

2.3.2 Fluidigm assay

The Fluidigm Biomark HD instrument was used to perform microfluidic measurement of gene expression in the amplified cDNA samples. Transcripts of interest were measured using 96.96 Dynamic GE Array chips (Figure 4A), allowing 9,216 simultaneous qRT-PCR reactions by quantifying abundance of 96 mRNA transcripts

in 96 samples. Samples and controls were run in duplicate, thus incorporating 45 amplified cDNA samples, an internal positive control (IPC), a NoRT and NTC. The IPC was pooled from different previously pre-amplified samples (using our primer-probes of interest) with varying levels of gene expression and was used to affirm that gene expression between the different experiments in each cohort (Chapters 4 and 5) was comparable. To prepare a chip for sample transfer, it was first injected with a control fluid into each accumulator and primed using the integrated fluidic circuit controllers. Assay loading reagent (Fluidigm, USA) was added on to the primer-probe assays at a 1:2 dilution, centrifuged and stored at 4°C until ready to load the chip. The samples were prepared by adding 2x TaqMan universal PCR master mix (Applied Biosystems, USA) and 20x GE sample loading reagent to 5.4μL of preAmped cDNA. Upon completion of priming, assays were loaded into the assay inlets in singlet and samples loaded into the sample inlets in duplicate (Figure 4A). The chips were returned to the integrated fluidic circuit controller for loading and mixing of samples and assays, after which the chips were loaded into the Fluidigm Biomark HD machine. Runs were performed using the Biomark data collection software (Fluidigm, USA), and pre-analysed using the Fluidigm real-time PCR analysis software v3.1.3 (Fluidigm, USA). The quality threshold, baseline correction and Ct threshold method were set to 0.65, linear derivative and auto global normalisation, respectively. Comma separated value (.csv) files of the data were exported for further analysis.

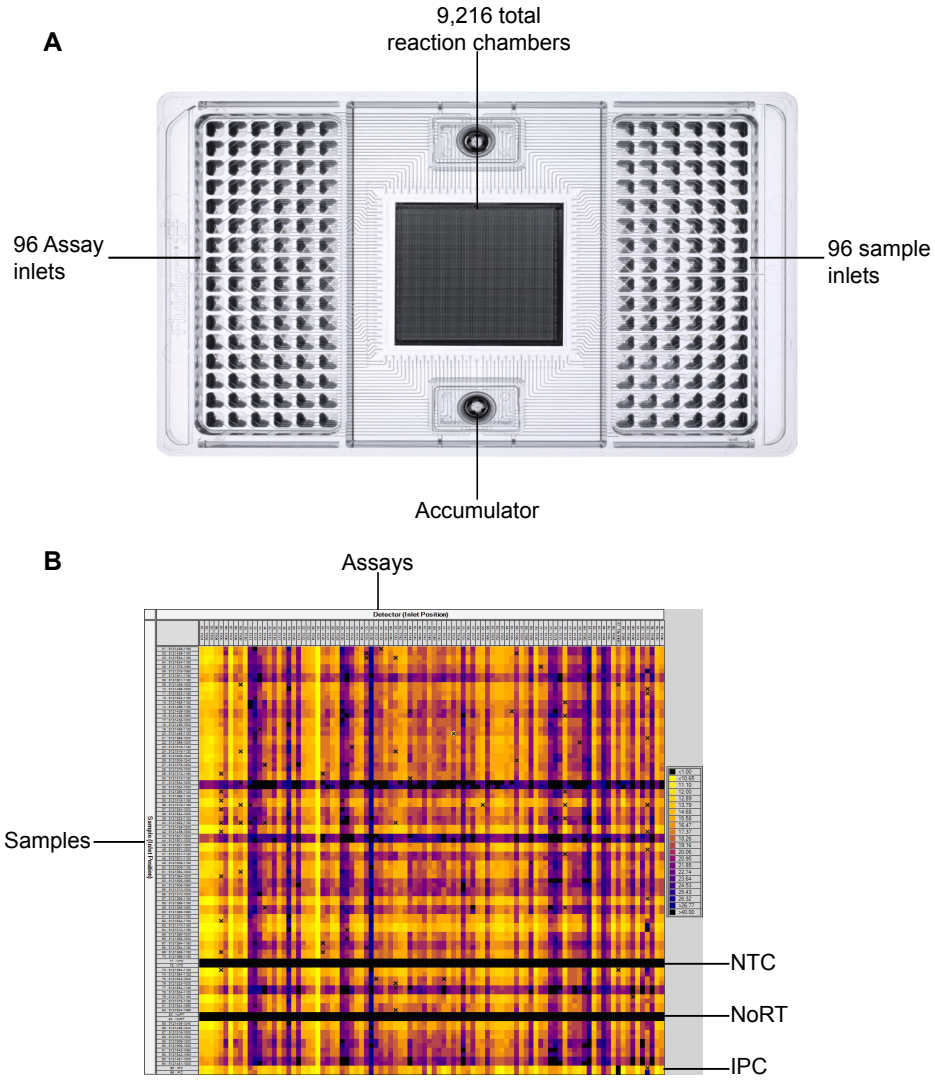


Figure 4: 96.96 dynamic array gene expression Fluidigm chip. (A) Layout and design of the 96.96 array chip for gene expression measurement. (B) Representative heatmap of gene expression measured from a 96.96 array chip, non-template control (NTC), no reverse transcriptase (NoRT) and internal positive control (IPC) shown.

2.4 Quality control (QC) of Fluidigm data

Several QC steps were applied to the qRT-PCR data from each individual chip to verify that the data generated were reliable before gene expression analyses. A customised script in R programming software was used to analyse the data. The first QC step was to verify that the parameters and instrument settings listed in section 2.3.2 in the Fluidigm RT-PCR analysis software were applied during the PCR run to generate the data. The script also tested to ensure that all samples were in duplicate. The controls were then tested to ensure that none of the primer-probes yielded detectable amplification (i.e. a cycle threshold, Ct, value) in the two negative controls (NTC and NoRT, Figure 4B). If amplification was detected for these controls, the experiment was repeated. Thereafter, the script determined whether any samples yielded failed amplification of at least 80% in the reference primer-probes (Table 3) and if so, these samples were excluded from further analyses. More than 80% of transcripts of genes of interest (GOI) also had to demonstrate detectable amplification for a sample to be included in the analyses. Finally, each transcript (primer-probe) had to be detected in at least 20% of all samples to qualify for inclusion in the analyses. Delta Ct (dCt) values for each transcript of interest (non-reference transcripts) were calculated by subtracting the Ct of the mean of the ten reference transcripts (Table 3) from the Ct of the primer-probe of interest (Table 1). Data (Ct and dCt values) from the internal positive controls (IPC) were analysed to ensure that gene expression was consistent between the different chips and runs. The Ct values of all transcripts were analysed by creating a correlation matrix between the replicates using Spearman's correlation. Heat maps of the correlation matrices of the replicates are presented in Chapters 4 and 5. Distributions of the dCt of the transcripts representing the GOIs are shown as density plots.

Table 3: Ten reference primer-probes included in the gene expression analyses of the ACS 11-gene signature and calculation of dCt values.

Gene symbol	Assay ID
ACTR3	Hs01029159_g1
ADRBK1	Hs01056345_g1
CDC42	Hs03044122_g1
CSDE1	Hs00918650_m1
CYTIP	Hs00188734_m1
TMBIM6	Hs01012081_m1
TMBIM6	Hs00162661_m1
TMBIM6	Hs01012082_g1
TPM3	Hs01900726_g1
USF2	Hs01100994_g1

2.5 Analysis methods

Blind predictions of the signatures were applied on the gene expression data using R scripts that were prepared and tested by our collaborators at the Centre for Infectious Diseases Research (CIDR, Seattle) (manuscript in preparation). The sum of the relative abundance of two gene products (transcripts) forming a pair and coefficients representing the weights between the pairs is the first step in calculating signature scores, as previously described (Zak et al. 2016). A sample is assigned as a case or progressor by each transcript-pair if the product of the calculation is greater than zero and a non-progressor or control if the result is less than zero. An overall score was tallied using the number of progressor scores from each gene pair, expressed as a proportion of the total number of gene pairs (Equation 1).

Score = “progressor” if: a*N1 + b*N2 + c > 0
Score = “control” if: a*N1 + b*N2 + c ≤ 0,

Equation 1: Calculation of signature scores for the ACS 11-gene signature. N1 and N2 represent normalised abundances (dCt) of two transcripts in the pair and coefficients "a", "b" and "c" as identified and trained in the ACS progressor and control cohort (Zak et al. 2016)

Area under the receiver operating characteristic curves (ROC AUC) were generated using the calculated signature scores in GraphPad Prism version 7.0a (GraphPad Prism). AUC values and 95% confidence intervals (CI) were calculated using the pROC (Robin et al. 2011) and verification (Pocernich 2015) packages in R, respectively. Head to head comparisons of AUCs were done using the “roc.test” function in the pROC package. For analysis of participant characteristics, Fisher’s exact test and Mann-Whitney U test were respectively used to analyse categorical and continuous data. P-values <0.05 were considered statistically significant; p-values in this thesis have not been corrected for multiple comparisons.

2.6 Contributions

Ms F. Darboe wrote the study proposal, designed and conducted the experiments and wrote this chapter under the supervision of Dr A. Penn-Nicholson and A. Prof T.J. Scriba.

Chapter 3: The ACS11-gene signature can differentiate active TB from *Mtb* infection using PBMC from HIV-infected and uninfected persons

The main aim of this chapter is to determine if the ACS 11-gene signature can distinguish active TB from *Mtb* infection in PBMC samples.

3.1 Introduction

HIV infection increases the risk of progression to active TB in *Mtb*-infected persons (Parida and Kaufmann 2010). A five-fold increase in TB notification rates has been observed due to HIV infection (Lawn et al. 2009). In 2016, 46% of incident TB cases were HIV-infected (World Health Organization 2017). HIV-infected persons present with paucibacillary disease (low numbers of acid-fast bacilli in sputum) thereby making diagnosis of TB using current diagnostic tools difficult. Sputum culture is a more sensitive technique in diagnosing TB in HIV-infected persons than sputum smear microscopy, however, it takes days to week to confirm the diagnosis (Tornheim and Dooley 2017; Cain et al. 2010; Stephen D Lawn et al. 2006). Furthermore, a large proportion of HIV-infected persons do not have productive coughs (Cain et al. 2010) and are thus unable to provide sputum specimens. Hence, new tools to improve and shorten time to diagnosis of active TB in HIV-infected persons are essential in the fight against TB.

Transcriptomic signatures of TB are promising tools for the diagnosis of active TB disease. Many diagnostic TB signatures have been shown to yield good diagnostic performance in HIV-uninfected persons, either to differentiate between active TB

cases and healthy infected or uninfected controls, or symptomatic controls with other respiratory diseases (Berry et al. 2010; Blankley et al. 2016; Bloom et al. 2012, 2013; Kaforou et al. 2013; Maertzdorf et al. 2011; Maertzdorf et al. 2011; Walter et al. 2016; Zak et al. 2016). Only a few studies have developed transcriptomic TB diagnostic signatures in HIV-infected populations (Kaforou et al. 2013; Anderson et al. 2014; Dawany et al. 2014). Colleagues at SATVI and CIDR developed a transcriptomic 16-gene correlate of risk of TB signature capable of predicting TB disease in HIV-uninfected persons more than a year prior to TB diagnosis (Zak et al. 2016). This 16-gene risk signature also performed well in diagnosing active TB disease, when re-parameterised and applied to public, published microarray datasets from the Berry *et a*l 2010, Bloom *et al* 2012, Kaforou *et al* 2013 and Maertzdorf *et al* 2011 (Zak et al. 2016). The performance of this risk signature in diagnosing active TB disease in HIV-infected persons has not been compared to its performance in HIV-uninfected persons from the same population. Hence, we sought to determine the performance of the 11-gene TB risk signature derived from the 16-gene signature (described in Chapter 1) in diagnosing active TB in HIV-infected persons.

An important aim we sought to address in this project was to determine if our transcriptomic signatures could predict recurrent TB in HIV-infected persons. However, the only available clinical cohort to address this aim was the TRuTH study, in which PBMC samples and no whole blood RNA samples were collected and biobanked. Since the 16-gene signature from which the 11-gene signature was derived was developed on whole blood samples, and PBMC are distinct from whole blood in that they lack neutrophils and other granulocytes. We therefore reasoned

that a careful comparison of signature gene expression between whole blood and PBMC was necessary. This is especially relevant because the Type I IFN response genes found to be associated with active TB are most prominently expressed by neutrophils compared to other cell types (Berry et al. 2010; Singhania et al. 2017; Scriba et al. 2017).

Previous studies have shown that transcriptomic signatures developed from PBMC samples can diagnose active TB disease in HIV-uninfected persons (Lee et al. 2016; Lu et al. 2011). Lee and colleagues developed gene signatures to differentiate active TB from healthy *Mtb*-infected and *Mtb*-uninfected persons. The study identified 297 genes that were differentially expressed between active TB cases and *Mtb*-uninfected persons, and 127 genes that were differentially expressed between *Mtb*-infected and -uninfected persons. There were 169 genes differentially expressed genes active TB and *Mtb* infection, and 14 of which genes were found to be differentially expressed between *Mtb*-infected and -uninfected persons. These genes were enriched in immune activation and regulation, cell differentiation, and mRNA transcription and translation modules. Three of the differentially expressed genes in the three comparisons differentiated *Mtb*-infected from uninfected persons and active TB cases in an external validation cohort from the same setting (Taiwan). Lu and co-workers on the other hand, measured gene expression in active TB cases, *Mtb*-infected, and uninfected persons using PBMC stimulated with PPD for four hours or unstimulated PBMC as controls. A three-way pair-wise analysis revealed 506 differentially expressed genes amongst the three groups based on fold changes between PPD stimulated and unstimulated samples. Fifty-five genes mainly involved in T cell homeostasis were differentially expressed when active TB and *Mtb*-infected

individuals were compared to *Mtb*-uninfected persons. Furthermore, 229 genes related to chemotaxis and responses to external stimulus were differentially expressed for active TB (active TB relative to *Mtb*-infected and uninfected persons). Of the 506 genes from the three pair-wise analyses, 30 differentially expressed genes and 22 genes specific for active TB and *Mtb* infection were translated into qPCR. These genes were applied to a validation cohort of active TB cases, *Mtb*-infected and uninfected persons using qPCR. Of these, a three-gene signature (CXCL10, ATP10A, TLR6) was found to be the most discriminatory in identifying active TB disease from *Mtb* infection using decision trees. The three-gene signature had a sensitivity and specificity of 80% and 89% respectively with an accuracy of 85%. These studies suggest that gene signatures developed from PBMC samples can diagnose active TB with a high accuracy.

However, the diagnostic performance of whole blood based transcriptomic signatures in differentiating between active TB disease and *Mtb* infection using PBMC samples is unknown. We therefore set out to assess the mRNA signal strength of the ACS 11-gene signature in PBMC.

3.2 Aim and hypothesis

Aim:

To determine if the whole-blood based ACS 11-gene TB risk signature can diagnose active TB disease from *Mtb* infection in HIV-infected and uninfected persons using PBMC samples.

Hypothesis:

The ACS 11-gene signature measured on mRNA from PBMC can discriminate between *Mtb* infection from active TB disease with equivalent performance to mRNA from whole blood.

3.3 Materials and methods

3.3.1 Ethics clearance and study design

Donors had been recruited using the healthy donor protocol (HREC 126/2006) and the TB disease protocol (HREC 288/2008). These protocols were reviewed and approved by the Human Research Ethics committee of the Faculty of Health Sciences at the University of Cape Town. All study participants gave written informed consent to participate in this pilot study.

Donors were recruited at the SATVI Clinical Site in the Worcester municipality of the Western Cape province of South Africa. Written informed consent was obtained from all study participants. Healthy *Mtb*-infected individuals and newly diagnosed active pulmonary TB diseased adults with or without HIV infection were invited to participate in the study. *Mtb* infection was determined using the QuantiFERON-TB Gold In-Tube assay (Qiagen, IFNγ levels >0.35 IU/mL). Active TB disease cases were diagnosed by Xpert® MTB-RIF (Cepheid), performed in community clinics.

Participant enrolment

Adults who satisfied the inclusion criteria below were enrolled into the study.

Inclusion criteria:

Mtb-infected persons

- Healthy adults at least 18 years old
- QuantiFERON-TB Gold In-Tube (QFT) positive (>0.35 IU/mL)

Active TB persons

- Adults at least 18 years old
- Xpert® MTB-RIF-positive

Exclusion criteria:

We used the following clinical criteria to exclude adults from enrolment into this study:

- Anaemia with haemoglobin less than 8.0 g/dl
- BMI equal to or less than 17
- Signs of other chronic or systemic illnesses not consistent with TB

A basic questionnaire was administered to all participants for capturing demographic data including anthropometric data. HIV status was confirmed at screening using a rapid serological test by the SATVI clinical team. All HIV-infected persons were referred to ART clinics for ART and, if appropriate, INH prophylaxis.

To allow for the potential re-parameterisation of the gene expression signature to PBMC, in case our results would show poor signature performance in the cohorts, we predefined a random division of the cases and controls in the HIV-uninfected study arm into a training and a test set at a 1:1 ratio. Experiments and analysis were performed sequentially in a blinded manner, first for the training set, and then for the test set based on either satisfactory performance of the signature in the training set, or after re-parameterisation on the training set in the case of poor signature

performance. Finally, the HIV-infected samples were also analysed. In this analysis only the TB status of the participants was blinded during experiments and QC analyses.

3.3.2 Sample collection

Ex-vivo whole blood in PAXgene™ tubes and PBMC were collected in parallel from each person. PBMC were harvested from blood collected in Cell Preparation Tubes (BD) at the SATVI clinical immunology laboratory, in Worcester.

3.3.3 RNA extraction and quantification

RNA from PAXgene tubes was extracted using the PAXgene blood RNA kit, as per manufacturer's instructions (described in section 2.2.1). RNA from PBMC was extracted immediately after thawing and washing using the Qiagen RNeasy Plus Mini Extraction Kit as per manufacturer's instructions (described in section 2.2.2). RNA was extracted immediately to avoid deviations from in-vivo gene expression as much as possible. Thus we did not try stimulating or resting PBMC after thawing because cryopreservation and resting might result in upregulation or downregulation of some of the gene of interest transcripts.

3.3.4 Gene expression measurement

cDNA was synthesised from extracted RNA using SuperScript II reverse transcriptase (described in section 2.3.1). Gene expression levels of transcripts in the ACS 11-gene signature were measured by qRT-PCR using the Fluidigm Biomark HD (described in section 2.3.2).

3.3.5 Data analysis

Fisher's exact test and Mann Whitney U test were used to analyse baseline participant characteristics. Gene expression data was quality controlled using scripts generated in R (described in section 2.3.2) and signature scores were calculated. Signature scores were determined by mathematically associating a coefficient with the relative abundance of primer-probe pairs in the signature (as described in chapter 2.5) and ROC AUCs were generated using the pROC (Robin et al. 2011) and verification (Pocernich 2015) packages in R. Median and 95% CI of the relative difference of primer-probes of interest between *Mtb*-infected and active TB disease was calculated using the rank inversion method and bootstrapping 2000 times using the quantreg R package (Koenker 1994).

3.3.6 Re-parameterisation of the signature to PBMC and HIV infection

In an attempt to potentially improve the performance of the ACS 11-gene signature for PBMC samples in an HIV-infected population, we re-parameterised the signature by determining more optimal coefficients for each transcript-pair in the signature. To do this, transcript Ct values from the ACS 11-gene signature generated from PBMC of HIV-infected participants were standardised using the mean Ct values of all the reference primer-probes. This was done because the ACS 11-gene signature scores are calculated using the relative abundance of transcripts of interest to the reference primer-probes. All the transcripts were then ranked by standard deviation (SD) of their dCt values of PBMC relative to whole blood; SD was used because PCR data is already log transformed. As data was standardised to reference primer-probes, they were expected to have lower SDs than transcripts of interest. Consequently,

reference primer-probes with large SDs ranked lower than transcripts of interest. These reference primer-probes were removed to pave way for the generation of a new signature. Transcripts of interest were standardised to the new set of reference primer-probes and the ability of the transcripts to diagnose active TB disease was determined. Transcripts of interest that did not significantly discriminate between active TB disease and *Mtb* infection were removed from the signature ($p>0.05$). The remaining transcript-pairs were retrained using linear support vector machines, changing the weighting and pairing in the signature. To test the performance of the new signature termed "rPSVM.1", the model was applied to the HIV-uninfected participants in this cohort, and to progressor and control participants from the ACS and GC6-74 cohorts.

3.4 Results

3.4.1 Cohort characteristics

One hundred and fifty-three adults were approached to participate in this pilot study. Newly diagnosed Xpert® MTB-RIF-positive TB patients (n=50) and healthy *Mtb*-infected (QFT-positive) individuals (n=50) were enrolled (Figure 5). In each of these groups, 30 participants were HIV-uninfected and 20 were HIV-infected. In the HIV-uninfected cohort, more active TB cases were younger, males, and had lower body mass index (BMI) than the *Mtb*-infected control participants in the cohort (Table 4). Most (80%) of the healthy *Mtb*-infected participants were of Cape Mixed Ancestry (Coloured) whilst only about half (53.3%) of the active TB disease participants were of Cape Mixed Ancestry. HIV-uninfected participants were randomly split into equally sized training and test sets at a 1:1 ratio of active TB cases and *Mtb*-infected controls (Figure 5). There was no difference in baseline characteristics between the

training and test sets (Table 5). Similar demographics were collected in the HIV-infected persons and there were no differences between the active TB cases and *Mtb*-infected controls (Table 6). Plasma viral loads and CD4 T cell counts were not collected for the HIV-infected cohort.

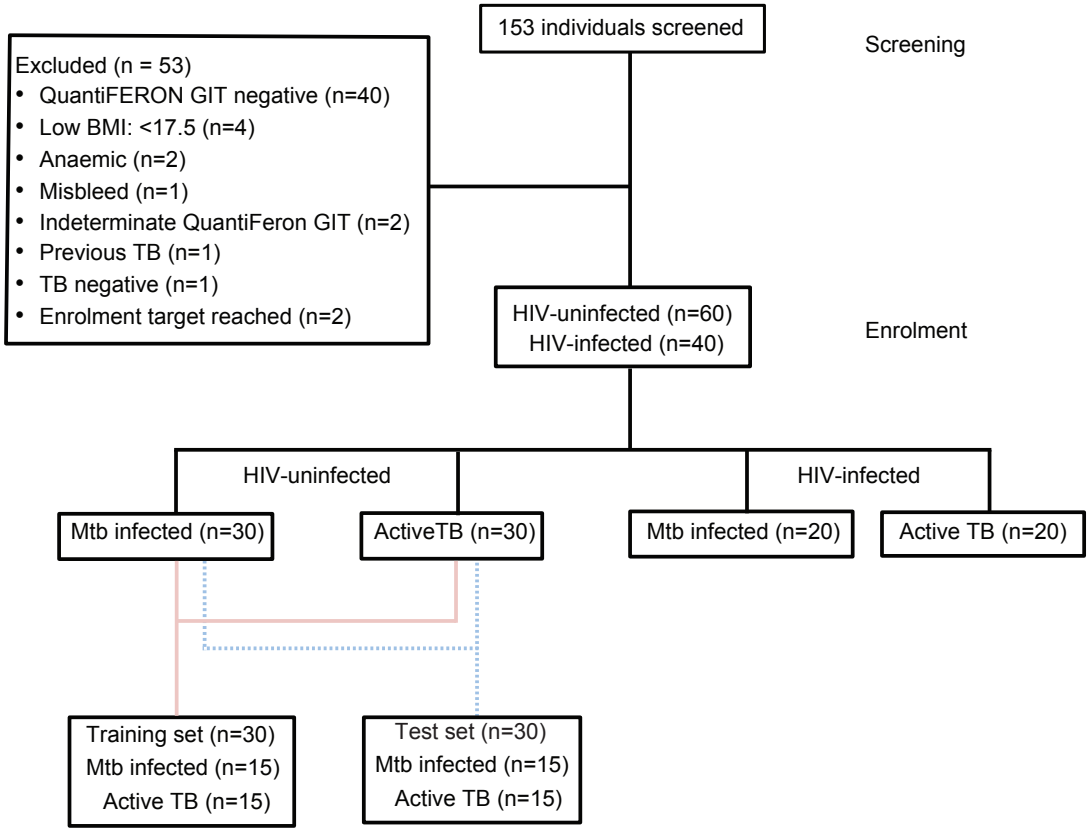


Figure 5: The pilot cohort. Participant screening and enrolment in the pilot cohort and assignment of participants in the HIV-uninfected cohort to training and test sets**.**

Table 4: Participant characteristics of the HIV-uninfected cohort. P <0.05 was considered significant. QFT- QuantiFeron, BMI- Body mass index

Participant variable/study group		**Active TB cases (n=30)**	***Mtb*-infected controls (n=30)**	**p-value**
Sex, n (%)	Male	22 (73)	8 (27)	0.0007
Median age, years (range)		34 (18-62)	40.5 (22-65)	0.01
Race, n (%)	Black African	14 (46.7)	5 (16.7)	0.03
	Caucasian	0 (0)	1 (3.3)	
	Cape Mixed Ancestry	16 (53.3)	24 (80)	
Median BMI, kg/m^2 (range)		20.5 (17.6-26.5)	30.4 (18.5-47.2)	<0.0001
Median QFT, IU/mL (range)		-	6.7 (0.5-13.8)	N/A

Table 5: Participant characteristics of the training and test set splits of the HIV-uninfected cohort. P <0.05 was considered significant. QFT- QuantiFeron, BMI- Body mass index

Participant variable/study group		**Training set (n=30)**	**Test set (n=30)**	**p-value**
Sex, n (%)	Male	19 (63.3)	11 (36.7)	0.07
Median age, years (range)		37.5 (18-65)	38 (19-54)	0.80
Race, n (%)	Black African	9 (30)	10 (33.3)	0.60
	Caucasian	1 (3.3)	0 (0)	
	Cape Mixed Ancestry	20 (66.7)	20 (66.7)	
Median BMI, kg/m^2 (range)		21.5 (17.6-44)	25.1 (18.1-47.2)	0.05
Median QFT, IU/mL (range)		5.7 (0.5-11)	8.4 (0.7-13.8)	0.40

Table 6: Participant characteristics in the HIV-infected cohort. P <0.05 was considered significant. QFT- QuantiFeron, BMI- Body mass index

Participant variable/study group		**Active TB cases (n=20)**	***Mtb*-infected controls (n=20)**	**p-value**
Sex, n (%)	Male	10 (50)	10 (50)	1.00
Median age, years (range)		36 (23-62)	31.5 (24-55)	0.10
Race, n (%)	Black African	14 (70)	16 (80)	0.70
	Cape Mixed Ancestry	6 (30)	4 (20)	
Median BMI, kg/m^2 (range)		22.5 (18.4-39.9)	22.1 (18.3-54.2)	0.8
Median QFT, IU/mL (range)		-	2.8 (0.35-13.1)	N/A

3.4.2: Indistinguishable diagnostic performance of the ACS 11-gene signature in whole blood and PBMC from HIV-uninfected participants

We first sought to determine if the signature could classify TB disease from healthy *Mtb* infection in PBMC in the training set. The 11-gene signature classified *Mtb* infection from active TB disease in whole blood and PBMC with comparable accuracy. The AUCs for whole blood and PBMC were equivalent at 0.93 (95% CI 0.80 to 1.00) and 0.96 (95% CI 0.89 to 1.00), respectively, p=0.71 (Figure 6A). Due to the excellent performance of the signature in PBMC, we proceeded to perform blind predictions of the unmodified 11-gene signature on the test set. An indistinguishable diagnostic performance was observed between whole blood and PBMC in the test set (Figure 6B). Performance in whole blood and PBMC was equivalent at 1.00 (95% CI 1.00 to 1.00) and 0.99 (0.98 to 1.00) respectively, confirming the results observed in the training set.

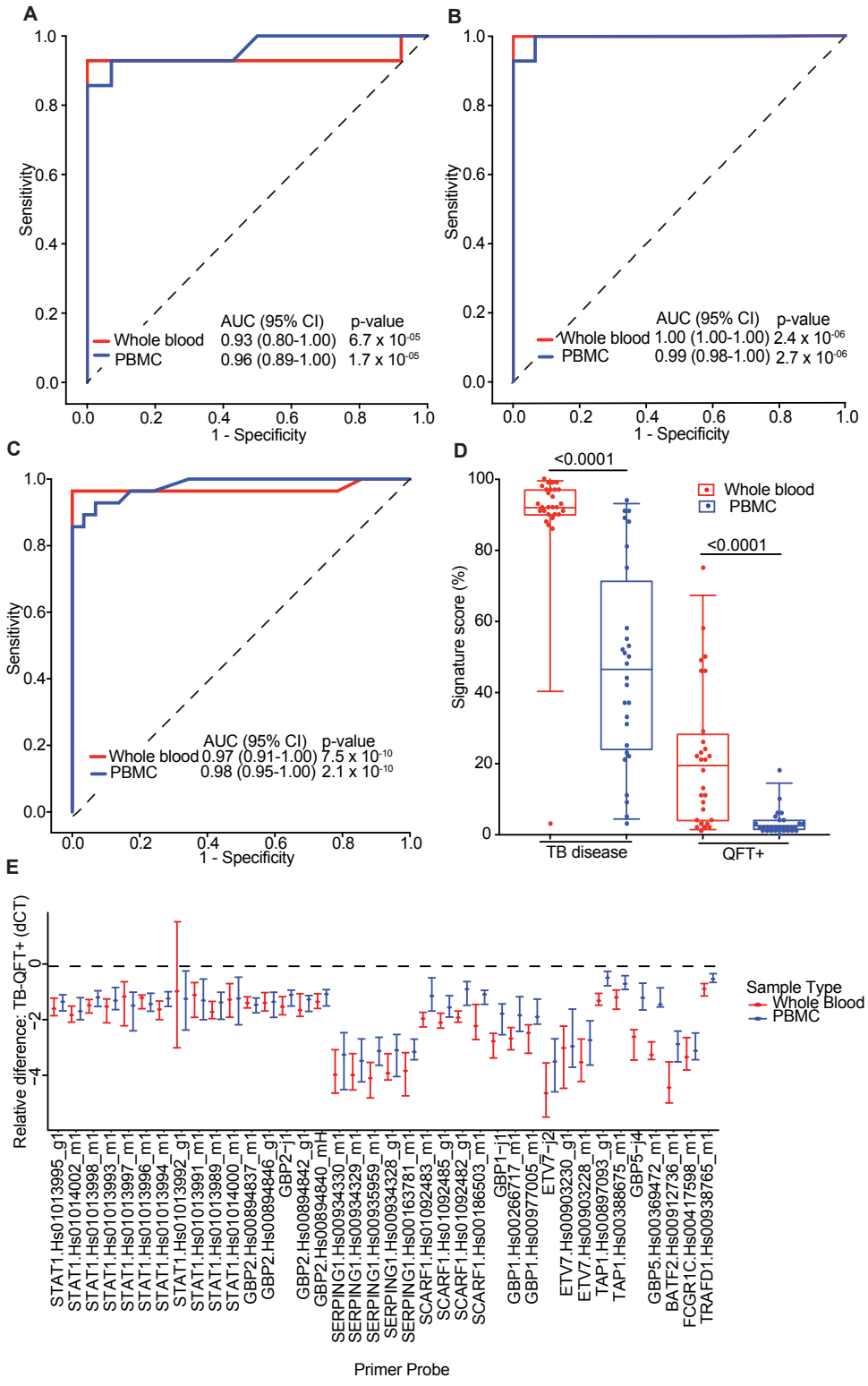


Figure 6: Performance of the ACS-11 gene signature in discriminating between active TB disease and *Mtb* infection in PBMC and whole blood from the HIV-uninfected cohort. A-C: Area under the Receiver Operating Characteristic curves illustrating the diagnostic performance of the 11-gene signature when measured in RNA from whole blood or PBMC samples from (A) the training set (n=30), (B) the test set (n=30), (C) the combined training and test sets (n=60). (D) Signature scores measured in whole blood or PBMC from active TB cases or healthy *Mtb*-infected persons. Horizontal lines represent medians, the box represents the interquartile range (IQR) and the whiskers represent the range. (E) Differences in transcript expression between TB cases and healthy controls when

measured in whole blood or PBMC samples. Negative differences indicate higher expression in TB than QFT+ controls. Dots represent medians and the bars 95% CI for each transcript identified by its TaqMan primer-probe set, computed with the rank inversion method and bootstrapping 2,000 times. Dashed lines represent the cut-off at which no difference was observed between cases and controls.

3.4.3: Expression differences of individual signature transcripts and signature scores in PBMC and whole blood

In light of the excellent performance results, we combined the training and test set data for subsequent analyses of the signature scores and to understand single transcript expression in PBMC and whole blood. The diagnostic performance of the signature in the combined datasets remained indistinguishable between whole blood, AUC 0.97 (95% CI 0.91 to 1.00) and PBMC, AUC 0.98 (95% CI 0.95 to 1.00) (Figure 6C). However, signature scores were lower in RNA from PBMC compared with RNA from whole blood (Figure 6D). As the lower signature scores from PBMC suggest lower gene expression in PBMC, we proceeded to determine the differences in single transcript expression between active TB cases and healthy *Mtb*-infected controls in whole blood and PBMC. Single transcript expression was higher in cases than in controls regardless of sample type (reflected by a negative difference in Figure 6E). Median relative differences in transcript expression between whole blood and PBMC were equivalent for most of the genes. However, for transcripts representing SCARF1, GBP1, GBP5, BATF2 and TRAFD1 the differences between active TB and healthy *Mtb* infection in whole blood were greater than in PBMC.

3.4.4: Performance of the signature in diagnosing active TB disease in the HIV-infected cohort

As observed in the HIV-uninfected cohort, the 11-gene signature could also successfully stratify *Mtb* infection from active TB disease in RNA samples from whole blood and PBMC from HIV-infected persons (Figure 7A). Again, AUCs were not significantly different (p=0.55) between whole blood, AUC 0.83 (95% CI 0.71 to 0.96) and PBMC AUC 0.88 (95% CI 0.76 to 0.99). Signature scores were lower in PBMC in comparison to whole blood (Figure 7B), although with a considerable overlap between *Mtb*-infected persons and TB cases. Single transcript expression was greater in TB cases in comparison to *Mtb*-infected controls in whole blood samples except for a few transcripts representing the STAT1 gene and TRAFD1 (Figure 7C). Single transcript expression was also greater in TB cases than controls in PBMC, except for a few transcripts representing the SCARF1 gene and again TRAFD1, for which no expression differences were observed.

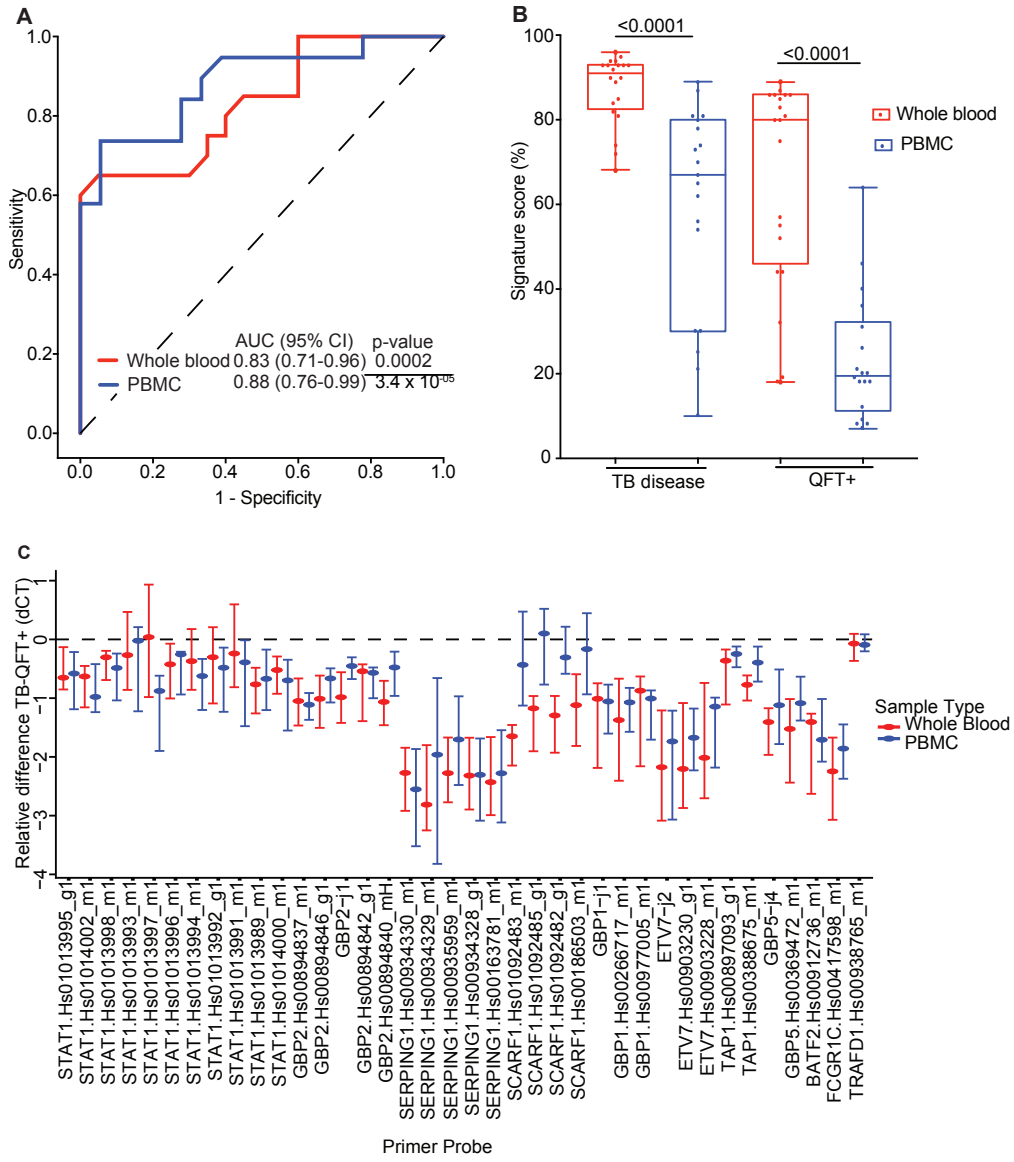


Figure 7: Performance of the ACS 11-gene signature in discriminating active TB disease and *Mtb* infection in PBMC and whole blood in HIV-infected participants (n=40). (A) ROC curves illustrating the diagnostic performance of the signature when measured in RNA from whole blood or PBMC. (B) Signature scores measured in whole blood or PBMC from HIV-infected active TB cases or *Mtb*-infected persons. Horizontal lines represent medians, the box represents the IQR, and the whiskers represent the range.(C) Differences in transcript expression between cases and controls when measured in whole blood or PBMC samples. Negative differences indicate higher expression in TB than QFT+ controls. Dots represent medians and the bars represent the 95% CI for each transcript, computed with the rank inversion method and bootstrapping 2,000 times. Dashed lines represent the cut-off at which no difference was observed between cases and controls.

3.4.5: HIV infection upregulates Type I IFN gene expression in whole blood and PBMC

The higher signature scores (representing higher IFN responses) in the QFT+ participants suggest lower specificity in HIV infection due to the higher number of false positives. The lower specificity resulted in a poorer diagnostic performance of the signature in both whole blood and PBMC samples in comparison to the HIV-uninfected cohort. We investigated if the difference significantly decreased the diagnostic utility of the signature by comparing the ROC AUCs using the pROC package. In addition, we investigated the difference in transcript expression between HIV-infected and uninfected active TB cases and *Mtb*-infected controls to determine the effect of HIV infection on gene expression. A decrease in ROC AUCs of at least 10% was observed in HIV-infected persons in comparison to HIV-uninfected persons in whole blood and PBMC (Figure 8A & B). The difference was not significant for PBMC (p=0.10) and approached significance in whole blood (p=0.06). Signature scores observed between HIV-infected and uninfected TB cases in both whole blood and PBMC were not significantly different (p= 0.05 and p = 0.15, respectively; Figure 8C & D). In both whole blood and PBMC samples, markedly increased signature scores were observed in *Mtb*-infected HIV-infected persons compared to their HIV-uninfected counterparts. We also observed generally higher single transcript expression in HIV-infected relative to HIV-uninfected participants (indicated by a negative difference in dCt between HIV-infected and uninfected persons in Figure 8E & F), irrespective of sample type (whole blood or PBMC). However, some STAT-1 transcripts in whole blood and PBMC as well as SCARF-1 transcripts in PBMC were expressed at lower levels in HIV-infected than in HIV-uninfected persons. These

patterns in signature scores and single transcript expression suggest that HIV infection drives type I IFN gene expression leading to lower specificity of the signature during HIV infection.

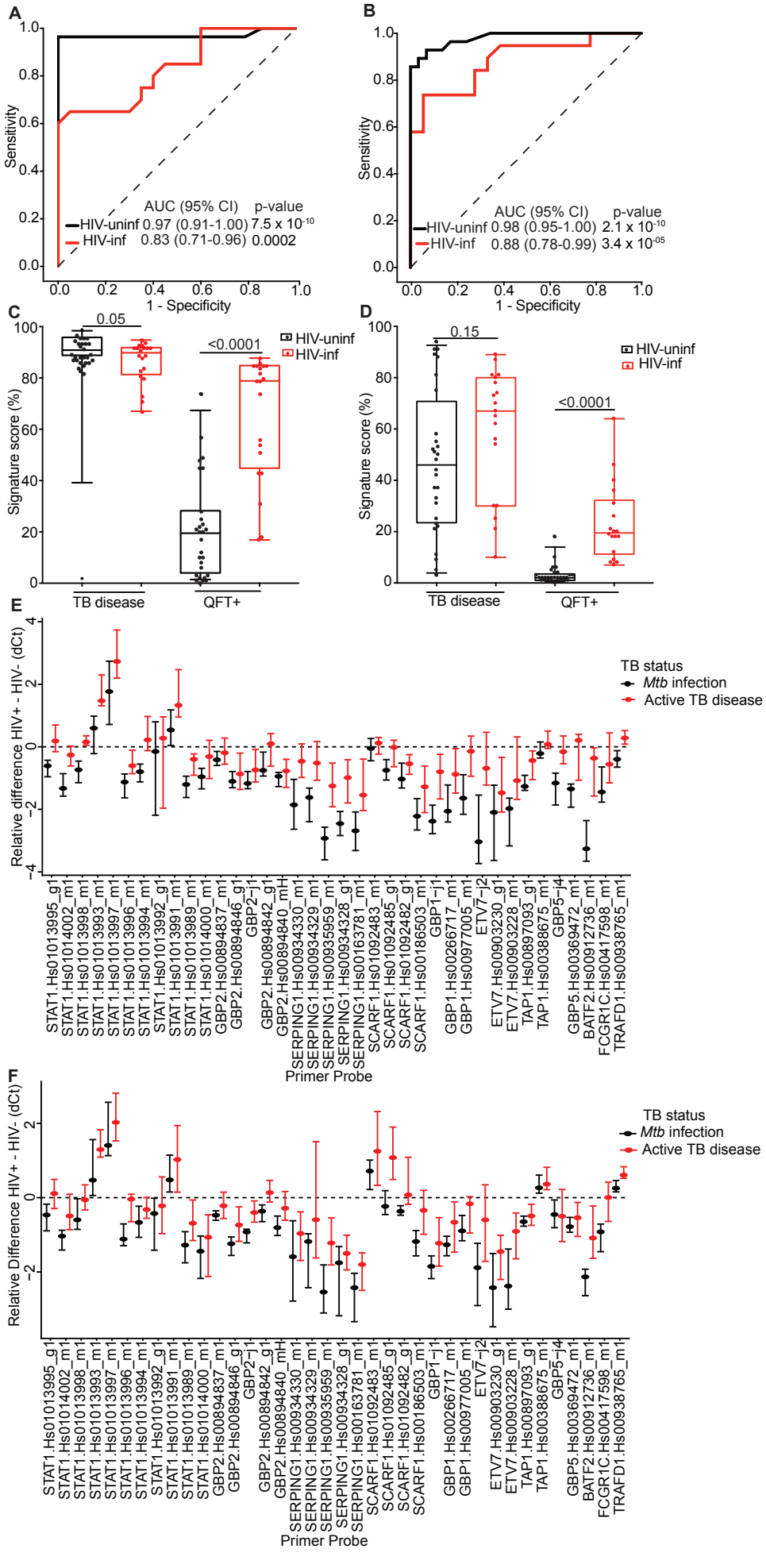

A
Sensitivity
1 - Specificity
AUC (95% CI) p-value
HIV-uninf 0.97 (0.91-1.00) 7.5 x 10-10
HIV-inf 0.83 (0.71-0.96) 0.0002
B
Sensitivity
1 - Specificity
AUC (95% CI) p-value
HIV-uninf 0.98 (0.95-1.00) 2.1 x 10-10
HIV-inf 0.88 (0.78-0.99) 3.4 x 10-05
C
Signature score (%)
0.05
<0.0001
HIV-uninf
HIV-inf
TB disease
QFT+
D
Signature score (%)
0.15
<0.0001
HIV-uninf
HIV-inf
TB disease
QFT+
E
Relative difference HIV+ - HIV- (dCt)
TB status
Mtb infection
Active TB disease
STAT1.Hs01013995_g1
STAT1.Hs01014002_m1
STAT1.Hs01013998_m1
STAT1.Hs01013993_m1
STAT1.Hs01013997_m1
STAT1.Hs01013996_m1
STAT1.Hs01013994_m1
STAT1.Hs01013992_g1
STAT1.Hs01013991_m1
STAT1.Hs01013989_m1
STAT1.Hs01014000_m1
GBP2.Hs00894837_m1
GBP2.Hs00894846_g1
GBP2-j1
GBP2.Hs00894842_g1
GBP2.Hs00894840_mH
SERPING1.Hs00934330_m1
SERPING1.Hs00934329_m1
SERPING1.Hs00935959_m1
SERPING1.Hs00934328_g1
SERPING1.Hs00163781_m1
SCARF1.Hs01092483_m1
SCARF1.Hs01092485_g1
SCARF1.Hs01092482_g1
SCARF1.Hs00186503_m1
GBP1-j1
GBP1.Hs00266717_m1
GBP1.Hs00977005_m1
ETV7-i2
ETV7.Hs00903230_g1
ETV7.Hs00903228_m1
TAP1.Hs00897093_g1
TAP1.Hs00388675_m1
GBP5-i4
GBP5.Hs00369472_m1
BATF2.Hs00912736_m1
FCGR1C.Hs00417598_m1
TRAFD1.Hs00938765_m1
Primer Probe
F
Relative Difference HIV+ - HIV- (dCt)
TB status
Mtb infection
Active TB disease
STAT1.Hs01013995_g1
STAT1.Hs01014002_m1
STAT1.Hs01013998_m1
STAT1.Hs01013993_m1
STAT1.Hs01013997_m1
STAT1.Hs01013996_m1
STAT1.Hs01013994_m1
STAT1.Hs01013992_g1
STAT1.Hs01013991_m1
STAT1.Hs01013989_m1
STAT1.Hs01014000_m1
GBP2.Hs00894837_m1
GBP2.Hs00894846_g1
GBP2-j1
GBP2.Hs00894842_g1
GBP2.Hs00894840_mH
SERPING1.Hs00934330_m1
SERPING1.Hs00934329_m1
SERPING1.Hs00935959_m1
SERPING1.Hs00934328_g1
SERPING1.Hs00163781_m1
SCARF1.Hs01092483_m1
SCARF1.Hs01092485_g1
SCARF1.Hs01092482_g1
SCARF1.Hs00186503_m1
GBP1-j1
GBP1.Hs00266717_m1
GBP1.Hs00977005_m1
ETV7-i2
ETV7.Hs00903230_g1
ETV7.Hs00903228_m1
TAP1.Hs00897093_g1
TAP1.Hs00388675_m1
GBP5-i4
GBP5.Hs00369472_m1
BATF2.Hs00912736_m1
FCGR1C.Hs00417598_m1
TRAFD1.Hs00938765_m1
Primer Probe

Figure 8: Comparison of the performance of the ACS 11-gene signature in HIV-infected and uninfected persons in RNA samples collected from whole blood and PBMC. A-B: ROC curves illustrating the diagnostic performance of the signature in classifying active TB from healthy *Mtb* infection when measured in HIV-infected and uninfected persons from (A) whole blood and (B) PBMC samples. C-D: Signature scores measured in HIV-infected or uninfected active TB cases or healthy *Mtb* infected persons in (C) whole blood and (D) PBMC. Horizontal lines represent medians, the box represents the IQR, and the whiskers represent the range. E-F: Differences in transcript expression between HIV-infected and uninfected persons when measured in active TB or *Mtb*-infected samples in (E) whole blood or (F) PBMC. Negative differences indicate higher expression in HIV-infected than HIV-uninfected persons. Dots represent medians and the bars 95% CI for each transcript, identified by its TaqMan primer-probe set, computed with the rank inversion method and bootstrapping 2,000 times.

3.4.6: Development and validation of a modified signature with enhanced diagnostic performance in HIV infection

It is well described that viral infections, and TB disease, increase type I IFN gene expression in blood leukocytes (Berry et al. 2010; Singhania et al. 2017; Scriba et al. 2017), as we also observed here. This elevation in type I IFN gene expression results in a reduction in specificity of the signature and is the most likely reason for the decrease in diagnostic performance in HIV-infected persons, relative to uninfected persons.

In an attempt to improve diagnostic performance of the 11-gene signature, we set out to adapt the signature for HIV infection to increase the accuracy of the signature in differentiating active TB disease from *Mtb* infection. We removed transcripts that resulted in diminished signature performance in HIV infection and re-parameterised the gene-pair weightings or coefficients that comprise the signature algorithm. First, the 10 reference (housekeeper) primers were ranked by their computed SD of the Ct values to determine variability in expression. The five reference primers with the largest SDs were removed (Figure 9). Ct values of GOI transcripts of PBMC samples from the HIV-infected cohort were then standardised to the mean of the five remaining reference primers, yielding dCt values. Ability of each individual GOI

transcript to discriminate active TB disease from *Mtb* infection in HIV-infected persons was assessed and eight primer-probes that could not classify active TB disease at a p-value cut-off of at least 0.05 were removed from the signature. Four SCARF1 primers, three STAT1 primers, and the TRAFD1 primer that yielded poor differentiation in PBMC from HIV-infected participants (Figure 8F) were removed. The remaining transcripts were re-trained (finding the best model by identifying transcript pairs and their optimal coefficients) to differentiate between active TB and *Mtb*-infected samples using linear SVM. The new signature, referred to as rPSVM.1 (re-parameterised pair-wise SVM), consisted of 30 transcripts or primer-probes (Table 7), representative of 115 transcript-pairs. However, when applied to both HIV-infected and uninfected groups of this cohort, rPSVM.1 did not yield significantly improved diagnostic performance than the original 11-gene signature (Table 8). We also tested if rPSVM.1 could differentiate between the HIV-uninfected TB progressors and controls from the ACS and GC6-74 cohorts to determine prognostic performance. Although rPSVM.1 yielded an AUC of 0.83 (95% CI, 0.77 to 0.88) in the ACS training set, which appeared higher than the 0.76 (95% CI, 0.71 to 0.82) obtained with the 11-gene signature, this difference was not significant (Table 8). Prognostic performance of rPSVM.1 in the ACS test set and the GC6-74 cohort were equivalent to the 11-gene signature.

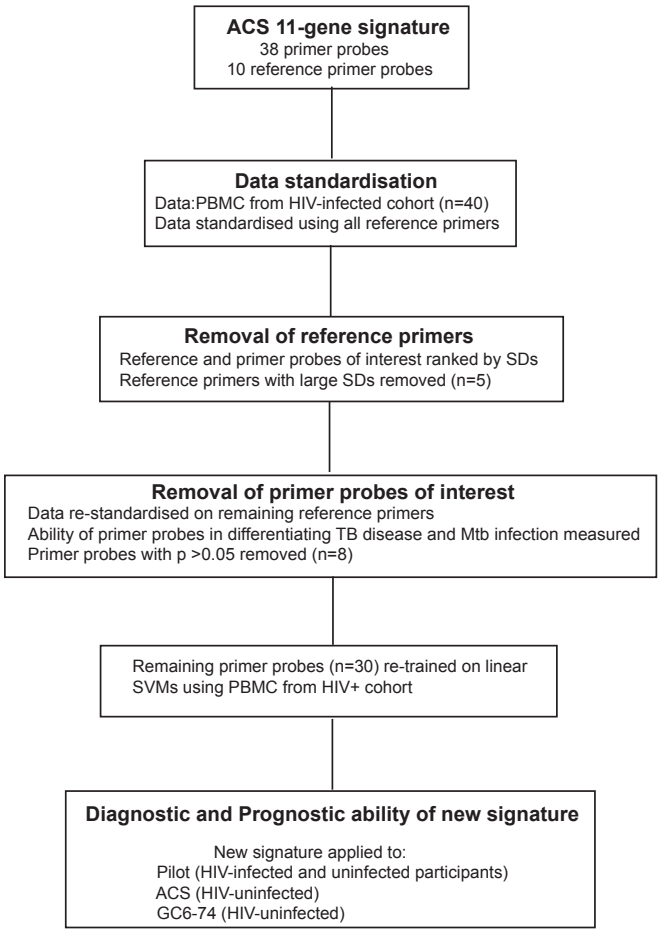


Figure 9: Development and validation of a re-parameterised signature (rPSVM.1) derived from the ACS 11-gene signature for HIV-infection. SD = standard deviation, SVM= support vector machine, ACS = Adolescent Cohort Study, GC6-74 = Grand Challenges 6-74 cohort.

Table 7: Genes and primer-probe IDs of transcripts in the rPSVM.1 signature.

Gene	Assay ID	Role in signature
ADRBRK1	Hs01056345_g1	Reference primer
CDC42	Hs03044122_g1	Reference primer
CSDE1	Hs00918650_m1	Reference primer
TMBIM6	Hs00162661_m1	Reference primer
USF2	Hs01100994_g1	Reference primer
GBP1	Hs00977005_m1	GOI primer
GBP2	Hs00894837_m1	GOI primer
GBP2	Hs00894846_g1	GOI primer
SERPING1	Hs00934330_m1	GOI primer
STAT1	Hs01013995_g1	GOI primer
SERPING1	Hs00934329_m1	GOI primer
GBP1	Hs00266717_m1	GOI primer
GBP1	j1	GOI primer
GBP2	Hs00894842_g1	GOI primer
GBP5	j4	GOI primer
STAT1	Hs01014002_m1	GOI primer
SERPING1	Hs00935959_m1	GOI primer
GBP2	j1	GOI primer
ETV7	j2	GOI primer
ETV7	Hs00903230_g1	GOI primer
TAP1	Hs00897093_g1	GOI primer
STAT1	Hs01013993_m1	GOI primer
SERPING1	Hs00934328_g1	GOI primer
SERPING1	Hs00163781_m1	GOI primer
BATF2	Hs00912736_m1	GOI primer
STAT1	Hs01013997_m1	GOI primer
STAT1	Hs01013996_m1	GOI primer
STAT1	Hs01013994_m1	GOI primer
STAT1	Hs01013991_m1	GOI primer
STAT1	Hs01013989_m1	GOI primer
FCGR1C	Hs00417598_m1	GOI primer
GBP2	Hs00894840_mH	GOI primer
ETV7	Hs00903228_m1	GOI primer
GBP5	Hs00369472_m1	GOI primer
TAP1	Hs00388675_m1	GOI primer

Table 8: Diagnostic and prognostic performance of the re-parameterised ACS 11-gene (rPSVM.1) signature in the pilot, ACS, and GC6-74 cohorts

Cohort	HIV status	Sample type	rPSVM.1 AUC (95% CI)	p-value	ACS 11-gene AUC (95% CI)	p-value of ROC comparison
Diagnostic						
Pilot	Uninfected	Whole blood	0.97 (0.90-1.00)	5.3×10^{-10}	0.97 (0.91-1.00)	0.93
Pilot	Uninfected	PBMC	0.99 (0.97-1.00)	5.2×10^{-11}	0.98 (0.95-1.00)	0.56
Pilot	Infected	Whole blood	0.84 (0.72-0.96)	1.9×10^{04}	0.83 (0.71-0.96)	0.93
Pilot	Infected	PBMC	0.93 (0.84-1.00)	3.5×10^{06}	0.88 (0.78-0.99)	0.11
Prognostic						
ACS training	Uninfected	Whole blood	0.83 (0.77-0.88)	3.2×10^{-14}	0.76 (0.71-0.82)	0.31
ACS test	Uninfected	Whole blood	0.75 (0.63-0.86)	3.8×10^{-05}	0.78 (0.68-0.89)	0.18
GC6-74 t	Uninfected	Whole blood	0.67 (0.60-0.74)	1.7×10^{-06}	0.69 (0.62-0.75)	0.36

3.5: Discussion

TB biomarkers are important tools that can help to identify individuals at the highest risk of developing active TB disease, or those who have subclinical or active TB disease, before and/or at TB diagnosis. Most published TB transcriptomic biomarker studies have focused on classifying active TB disease from healthy *Mtb*-uninfected or *Mtb*-infected controls, while some have demonstrated performance in differentiating between TB disease and other respiratory diseases such as pneumonia (Anderson et al., 2014; Berry et al., 2010; Bloom et al., 2013; Kaforou et al., 2013; Maertzdorf et al., 2012; Walter et al., 2016). These signatures typically return to levels observed in controls during TB therapy, suggesting that the selected transcripts reflect TB disease pathogenesis and the underlying immune responses (Berry et al. 2010; Bloom et al. 2012). Indeed, we have shown that the transcripts in our prognostic 16-gene and 11-gene signatures are highly upregulated at time points closest to TB diagnosis (Zak et al. 2016; Darboe et al. 2018). This chapter

demonstrates that these signatures performed well as diagnostic PCR-based biomarkers when differentiating between active TB disease and healthy, *Mtb*-infected controls. Similar diagnostic performance was also observed when the 16-gene signature was re-parameterised to published microarray datasets and applied to distinguish between active TB disease and healthy *Mtb*-infected persons (Zak et al. 2016). In addition, the 16-gene signature could differentiate active TB disease from diseases such as lung cancer, pneumonia, and sarcoidosis when applied to the published datasets of Bloom (2013), Kaforou (2014) and Anderson (2014) (Zak et al. 2016). A limitation in our analysis is the exclusion of other diseases that have similar clinical presentation as active TB disease, which would be encountered in a health-care setting. An advantage of PCR-based blood diagnostic tests in clinical settings is the ability to differentiate active TB disease from other diseases. The 11-gene signature has the potential to be used as such a diagnostic test, but this could not be confirmed in our cohorts.

The risk of progression to active TB disease is significantly increased by HIV infection (Parida and Kaufmann 2010). TB diagnosis in HIV-infected persons is complicated by the paucibacillary nature of sputum produced by HIV-infected and especially highly immunocompromised TB patients. This makes it difficult to diagnose active TB in HIV-infected persons, especially when using sputum smear microscopy, which has a poor sensitivity. Sputum cultures are more sensitive than sputum smears in diagnosing active TB (Cain et al. 2010; Stephen D Lawn et al. 2006), but diagnosis using sputum cultures can take up to 42 days. Blood-based RNA signatures could serve as important new tools for TB diagnosis in HIV-infected persons. We demonstrate here that the 11-gene signature can diagnose active TB in

HIV-infected persons using whole blood samples, could potentially be used as a triage test for TB disease in clinical settings.

Large clinical trials with long follow-up periods are needed to test the utility of prognostic signatures for identifying HIV-infected persons at risk of active TB disease. Biobanked PBMC samples from a longitudinal cohort of HIV-infected persons are available for such an evaluation, using the ACS 11-gene risk signature. In this chapter, we sought to prepare for these analyses by testing the performance of the whole-blood derived signature in PBMC samples. We reasoned that reliable performance of the signature using PBMC would allow testing of the signature in the biobanked samples from the TRuTH study. We show here that the signature yielded excellent diagnostic performance for active TB using PBMC from both HIV-infected and uninfected persons and that this performance was equivalent to that obtained with whole blood samples.

Despite the excellent performance of the signature in diagnosing active TB in PBMC, signature scores were lower in PBMCs than whole blood. These lower signature scores could be attributed to the lower mRNA expression in PBMCs for multiple transcripts, most likely due to the absence of certain cell types, such as neutrophils, in PBMCs, which are removed by gradient centrifugation. Neutrophils from active TB patients and healthy controls displayed very pronounced differences in expression of type I/II IFN response transcripts and appeared to contribute the strongest expression differences when compared to other blood cell subsets (Singhania et al. 2017; Berry et al. 2010; Scriba et al. 2017). We did not carry out any flow cytometry analyses on the samples to measure gene expression in different cell subsets. We

therefore do not know how the gene expression signatures performed in different cell subsets.

The signature performance in HIV-infected persons was at least 10% lower than in HIV-uninfected persons, as has been previously observed (Zak et al. 2016). However, the ROC AUCs were not statistically different, suggesting either an equivalent accuracy in the two populations or highlighting that the sample sizes in our study were not large enough. Some of the *Mtb*-infected HIV-infected persons could have been on INH prophylaxis, this could decrease gene expression in these participants. In addition, modulation of gene expression could occur due to its effects independent of *Mtb*. We, however, did not have enough statistical power to study the effect of INH on signature modulation.

We show that the reduction in signature performance in the HIV-infected cohort is likely attributable to the higher signature scores observed in *Mtb*-infected controls in this population. HIV infection was previously shown to increase immune activation and serum type I IFN response (Psomas et al. 2016; Bosinger and Utay 2015). Higher signature scores in the controls decreased the diagnostic performance and specificity of the signature. In addition, the degree of immune activation is associated with HIV viral load (Psomas et al. 2016; Klatt et al. 2013), thus the increase in signature scores in the *Mtb*-infected cohort are also likely to be affected by the viral load levels. This is investigated further in Chapter 6 of this thesis.

Despite the limited sample size of this cohort, the results generated are consistent with those of other studies of diagnostic TB biomarkers (Lee et al. 2016; Lu et al.

2011). Previous studies have also shown a reduction in the performance of TB diagnostic RNA signatures in HIV-infected persons (3 to12%) relative to HIV-uninfected persons (Sweeney et al. 2016; Walter et al. 2016; Dawany et al. 2014; Kaforou et al. 2013). Hence, a reduction in signature performance in HIV-infected persons is generally observed, further strengthening our observations.

We set out to generate an HIV-adapted signature that may perform better on PBMC samples by removing transcripts that decreased the specificity of the ACS 11-gene signature. We also adjusted the model parameters for PBMC samples by re-training transcript pairs on linear SVMs whilst maintaining the pair-wise ensemble in the signature. It was somewhat surprising that the diagnostic and prognostic performances of the rPSVM.1 signature were equivalent to the performances of the ACS 16 and 11-gene signatures. However, the result suggests that signatures of active TB of whole blood adjusted for PBMC samples and HIV-infected persons can accurately diagnose and predict active TB in whole blood samples and HIV-uninfected populations.

In conclusion, we show here that the ACS 11-gene signature can differentiate active TB disease from healthy *Mtb* infection in both whole blood and PBMC. The weaker IFN response signature in PBMC indicates the need for a different sample-type specific threshold for diagnosis of active TB disease due to the absence of signal from certain cell types, such as neutrophils. Regardless, the comparable performance between whole blood and PBMC demonstrates feasibility of applying the signature to biobanked cryopreserved PBMC for assessment of TB risk. This increased our confidence in testing the prognostic performance of the signature in

HIV-infected persons, and as a result we applied these signatures to the TRuTH (discussed in Chapter 4) and IMPRESS (discussed in Chapter 5) cohorts.

3.6: Contributions

Ms F. Darboe designed and conducted the experiments and wrote this chapter under the supervision of Dr A. Penn-Nicholson and A. Prof T.J. Scriba. The new signature was developed and tested by Dr E.G. Thompson. R scripts to calculate signature scores were written by Dr E.G Thompson.

Chapter 4: Predicting recurrent TB disease in HIV-infected persons on ART

The aim of this chapter is to determine if the TB risk signatures can predict recurrent TB disease in HIV-infected persons on ART.

4.1: Introduction

Two critical issues related to the management and care of active TB disease in HIV-infected populations is how to effectively reduce mortality and prevent TB recurrence after treatment. The World Health Organisation (WHO) policy on collaborative TB/HIV activities aims to "reduce the burden of TB in people living with HIV, by ensuring the delivery of the three "I"s (intensified case finding, infection control and IPT), for HIV/TB" (WHO, 2012). In order to achieve this, it is recommended that adults and adolescents living with HIV should be regularly screened for TB and should receive at least six months of IPT irrespective of degree of immune suppression (WHO, 2012). A systematic review of 12 randomized controlled trials of HIV-infected individuals showed that a six to12 month course of daily INH therapy reduced TB incidence by up to 60% (Akolo et al., 2010).

This approach is however undermined by poor adherence and incomplete *Mtb* sterilisation by IPT (Samandari et al. 2011). A study in more than 70,000 South African gold miners showed no significant effect of prophylaxis on both TB incidence and prevalence after the treatment period, likely in part due to poor adherence to the treatment regimen (Churchyard et al. 2014). Poor adherence may be attributable to the duration of the prophylaxis. In addition, mathematical modelling of the data from the study suggested that a small proportion of *Mtb* infection were cured, resulting in

the lack of detectable impact of the prophlaxis in the gold mining community (Vynnycky et al. 2015). More targeted provision of directly observed TB treatment, and novel regimens, such as provision of short-course, high-dose rifapentine and INH (Sterling et al. 2011), may be significantly more effective in improving adherence to prophylaxis, thus reducing progression to active TB disease. *Mtb*-infected persons appear to transition through a subclinical phase of disease when progressing to active TB disease (Esmail et al., 2014; Salgame et al., 2015; Young et al., 2009). In this phase patients are asymptomatic but may have evidence of replicating bacteria and may be identified by occasional sputum culture positivity (Pai et al. 2016; Salgame et al. 2015). In autopsy studies from the 1920s, replicating bacteria were identified in persons dying of other causes, suggesting on-going active infection or disease (Opie and Aronson 1927). Mass screening and surveys regularly identify asymptomatic persons with microbiologically positive subclinical TB (Corbett and MacPherson 2013; Onozaki et al. 2015) which would have otherwise been undetected, increasing the pool of infectious individuals and risk of TB transmission. Identification of HIV-infected persons at high risk of primary and recurrent TB disease, or with subclinical disease, would allow targeted provision of active TB treatment and could significantly contribute to decreasing transmission within communities.

To our knowledge, only one study has attempted to identify transcript signatures of recurrent TB in HIV-uninfected adults (Mistry et al. 2007). In this study, HIV-uninfected participants with recurrent TB disease were matched to active TB cases (with no previous history of TB), *Mtb*-infected persons, and participants who had been clinically cured of active TB disease after one TB episode. Three hundred and

thirty-seven genes were identified to be differentially expressed between the four groups from 27,566 cDNA probes analysed by microarray. Nine genes were derived from the 337 most differentially expressed genes by step-wise linear discriminant analysis, and were shown to accurately distinguish between the different groups. This study illustrates that TB treatment can be monitored using a blood-based transcriptomic signature, a tool that would circumvent the reliance on sputum-based diagnostic tests. More studies are necessary to determine if such signatures are applicable to other populations, epidemiological and environmental settings. The Mistry *et al* study was limited by the small sample size (n=10 for each group), while recurrent TB cases had two or three previous TB episodes. Importantly, the biomarker signatures they discovered were not validated in an external or independent cohort, a critical criterion for any promising biomarker.

We showed that active TB disease can be predicted and diagnosed by the ACS 11-gene signature (Chapter 3) and developed an adapted signature, termed rPSVM.1, with the aim of increasing its accuracy for diagnosing active TB disease in HIV-infected populations. We showed that these signatures could predict active TB disease a year before clinical symptoms manifested in HIV-uninfected persons (Zak et al. 2016), suggesting the detection of underlying incipient or subclinical TB disease. However, it is not known if these signatures can predict recurrent TB disease in HIV-infected persons. Thus, we set out to determine the prognostic potential of ACS 11-gene and rPSVM.1 signatures for TB recurrence in HIV-infected persons who are on ART.

We collaborated with investigators from the Centre for the AIDS Programme of Research in South Africa (CAPRISA), who completed a prospective observational cohort study, which aimed to determine the rate at which recurrent TB disease occurs in patients who completed TB therapy. The *TB Recurrence upon Treatment with Highly Active Antiretroviral therapy* (TRuTH) study was carried out at the CAPRISA eThekwini clinic, Durban. This study followed up newly diagnosed active TB cases who had completed the SAPiT (Starting AIDS treatment at three Points in TB treatment) (Abdool Karim et al., 2011; Abdool Karim et al., 2010) and START (Starting TB and ART) (Gengiah et al. 2012) trials, wherein they started ART. Participants who had completed participation in these trials were screened for participation in the TRuTH study after obtaining informed consent. The study was conducted between 2009 and 2013 and 92 episodes of recurrent TB was recorded from 82 patients during a median of three years of follow-up.

4.2: Aim and hypothesis

Aim

To determine the prognostic value of the ACS 11-gene and rPSVM.1 transcriptomic risk signatures for recurrent TB disease in HIV-infected persons on ART.

Hypothesis

The risk signatures will significantly predict recurrent TB disease in HIV-infected persons on ART.

4.3: Methods

4.3.1: Study design

Participant enrolment into TRuTH

Participants from the START and SAPiT trials who satisfied the following criteria were enrolled into the TRuTH study.

Inclusion criteria

- Adults at least 18 years old
- Previously enrolled in the SAPiT and START trials
- Previous history of active TB disease with HIV co-infection
- Consent to participate and contribute specimens to the KwaZulu-Natal Research Institute for TB and HIV (K-RITH) repository for future investigations

Exclusion criteria

- Refusal to consent
- Diagnosed with XDR TB

This study was reviewed and approved by the University of Kwazulu-Natal biomedical research ethics committee (BF051/09). All participants gave informed consent and consented to have samples stored for future studies.

Participants who met the above criteria were enrolled and longitudinally followed up for up to three years. Recurrent TB disease was detected by chest x-rays, sputum smear and culture using induced sputa, at study entry (baseline), at monthly intervals for the first three months and at three-monthly intervals thereafter. Additional cultures were performed when participants presented with TB symptoms and/or abnormal chest x-rays outside of the follow-up time points. Only cases with microbiologically

confirmed TB were included in our biomarker sub-study. In addition to TB investigations, serum and PBMC were collected at every visit.

Of the 82 microbiologically confirmed progressors, 43 had PBMC samples available and were all included in our sub-study to determine the prognostic value of the ACS 11-gene and rPSVM.1 signatures. Each progressor was assigned to two non-progressors, who did not develop recurrent TB during follow up. Non-progressors were matched to progressors based on time on ART and gender at TRuTH baseline. Due to the confounding effects of CD4 counts on risk of TB disease and viral loads on IFN responses, we compared these variables between the 43 progressors and 86 non-progressors at first TB episode diagnosis, end of TB treatment, and at TRuTH baseline and confirmed that they were not different. At least three PBMC samples were selected for analysis at three to six monthly intervals from baseline to recurrent TB diagnosis, where possible. Samples collected around the same study day (visit) were selected for progressors and matched non-progressor controls.

Because we could not predict how well the two signatures would perform, we prepared an analysis plan to allow the possible re-parameterisation of the signatures should they perform poorly. Thus, progressors and non-progressors were split into a training and test set. Ten different iterations, randomly selecting 29/14 progressors per split were done and a split with suitable distributions of two-thirds of samples in the training set and a third of samples in the test set was nominated.

We also sought to determine the ability of the signatures to differentiate subclinical TB from active TB disease or *Mtb* infection in HIV-infected persons in a cross-

sectional cohort of ART naïve HIV-infected individuals (Esmail cohort). We hypothesized that the signatures would differentiate subclinical TB from *Mtb* infection but not from active TB disease. HIV-infected persons with no previous history of active TB disease were screened from an outpatient clinic just outside Cape Town (Esmail et al., 2016, 2018). This study was reviewed and approved by the research ethics committees of the Universities of Cape Town (013/2011) and Stellenbosch (N12/11/079). Informed consent was obtained from all participants before screening and study entry. *Mtb*-infected (QFT-positive) persons, with no evidence of active TB on chest x-ray and CD4 counts greater than 350/mm^3 were recruited to undergo a 2-deoxy-2 [^{18}F] fluoro-D-glucose positron emission and computed tomography (PET-CT) scan. Included in this analysis were participants with no features of disease activity on PET-CT, classified as latently (*Mtb*)-infected persons. Those with features of disease activity on PET-CT, but who were asymptomatic with initial negative sputum culture, were classified as subclinical TB. In addition, culture or Xpert MTB-RIF positive symptomatic active TB disease cases were included (Esmail et al., 2016, 2018).

4.3.2: RNA extraction and quantification

RNA from the TRuTH cohort was extracted from cryopreserved PBMC as per manufacturer's instructions, as described in section 2.2.2. Extracted RNA was quantified using the Nanodrop ND 2000 spectrophotometer as described in section 2.2.3. Already extracted RNA samples from whole blood from participants enrolled into the Esmail cohort were received from our collaborators and these were not quantified.

4.3.3: Gene expression measurements

cDNA was synthesized, reverse transcribed and pre-amplified using TaqMan primer-probes (described in section 2.3.1). Gene expression levels of the signature genes were measured by microfluidic qRT-PCR on the Biomark HD instrument (described in section 2.3.2).

4.3.4: Data analysis

Data generated were quality controlled and analysed as described in sections 2.4 and 2.5. Blind predictions were applied to both the training and test sets simultaneously before unblinding of participants' progressor/non-progressor status in the training set. Analyses on the test set were only carried out after analyses of the training set were complete. Time to diagnosis of recurrent TB (days) was calculated for all progressors by subtracting the sample collection date from the date of recurrent TB diagnosis. For non-progressors, a "time to recurrent TB" was calculated by subtracting sample collection date from the recurrent TB diagnosis date of assigned progressors. Statistical analyses were done in GraphPad Prism (v7.0a) and R.

4.4: Results

4.4.1: Participant selection and characteristics

Five hundred and twenty participants who completed follow-up in the SAPiT (n=642) and START (n=58) clinical trials were enrolled into the TRuTH study (Figure 10). Ninety-two episodes of recurrent TB in 82 participants were recorded during the three-year follow-up period. Of these, 43 participants with pulmonary TB had PBMC

samples stored, which were available for this sub-study (Figure 10). TB recurrence occurred at a median time of 470 days (IQR: 250 to 672 days) after enrolment into the TRuTH cohort. Of the 438 participants who remained healthy during follow-up, 86 non-progressors were selected. Samples between enrolment and recurrent TB disease diagnosis were selected at approximately six monthly intervals. Where possible, at least three samples were selected prior to recurrent TB diagnosis. When realigned to time to recurrent TB diagnosis, most of the samples were collected up to a year before recurrent TB diagnosis (Figure 11A). A total of 103 and 196 samples were available from progressors and non-progressors, respectively. CD4 T cell counts, pVL and BMI at the start of first TB episode treatment were not different between the progressors and assigned non-progressors (Figure 11B-D). CD4 T cell counts and pVL were also not different at the end of primary TB treatment (Figure 11E-F). Time on ART was also not different between the two groups at TRuTH enrolment (Figure 11G). Participant characteristics such as sex, race, BMI and pVL were also similar between the two groups at TRuTH baseline with the exception of CD4 counts, which were higher in the non-progressors (Table 9). The training/test split with 66 progressors and 130 non-progressors samples in training set, and 37 progressors and 66 non-progressors samples in the test set, was nominated (Figure 12).

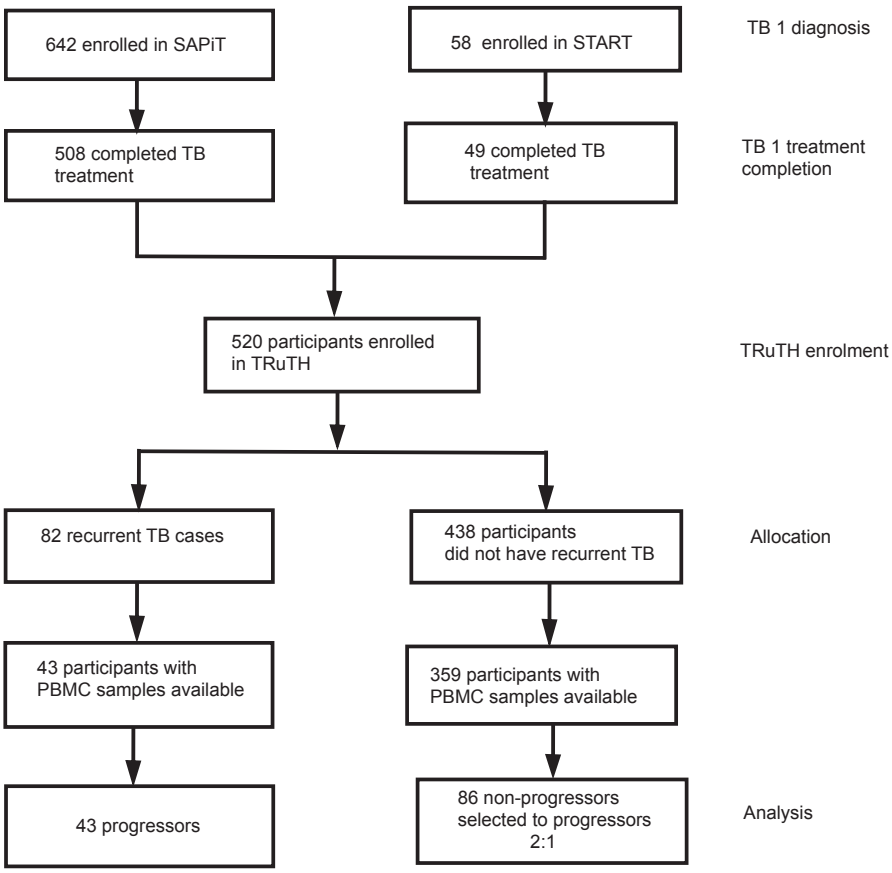


Figure 10: Selection of participants for the TRuTH cohort, in which prediction of recurrent TB disease in people living with HIV was tested. Progressors with PBMC samples stored before recurrent TB episodes were matched to eligible non-progressors (1:2).

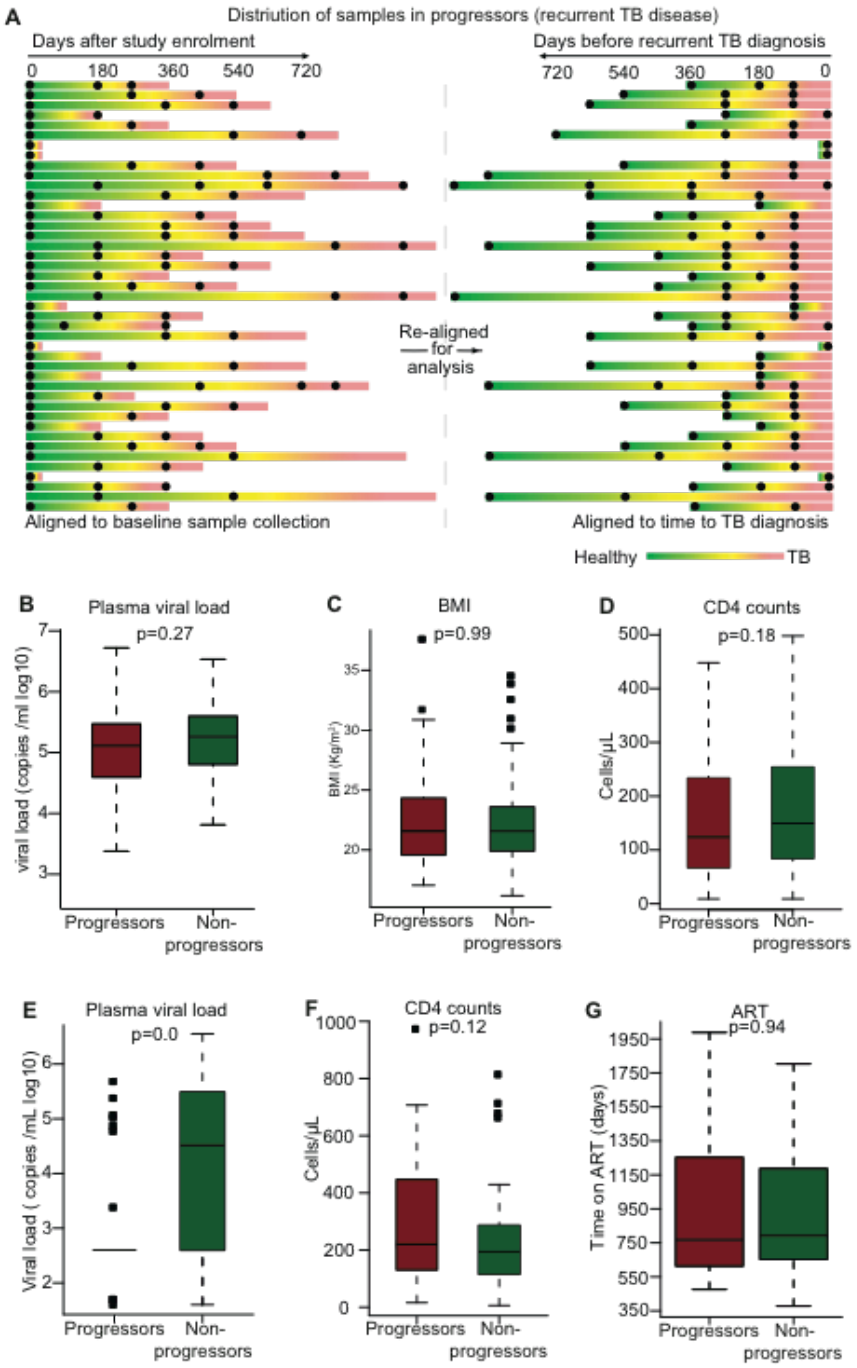


Figure 11: Selection of recurrent TB progressors and controls and their samples for transcriptomic signature testing. (A) Distribution of progressor samples depicted by time since TRuTH enrolment (left) and time to recurrent TB diagnosis (right). Each line depicts a progressor and the dots represent the time points of selected PBMC samples. (B-F) Differences in demographics and risk factors of TB between progressors and non-progressors at (B-D) the start of treatment of TB episode 1, (E&F) end of treatment of TB episode 1 and (G) at TRuTH enrolment.

Table 9: Basic demographics of participants who were included in this sub-study at TRuTH baseline. For each progressor, two non-progressors were selected from a set of available non-progressors with the same range of CD4 counts and plasma viral loads (pVL) as the progressors. pVL limit of detection = 400 copies/mL.

Participant variable		**Progressors (n=43)**	**Non-progressors (n=86)**	**p-value**
Sex, n (%)	Male	21 (49)	42 (49)	1.00
Median age, years (range)		35 (25-53)	37.5 (23-62)	0.38
Race, n (%)	Black African	43 (100)	85 (99)	1.00
	Other	0 (0)	1 (1)	
Median BMI, kg/m^2 (range)		24 (16.5-41.9)	24.3 (16.4-40.8)	0.58
Previous TB, n (%)	1	27 (62.8)	58 (67.4)	0.69
	2	16 (36.4)	28 (32.6)	
	Median time since previous TB, days (range)	876 (707-2009)	1079 (706-1866)	0.05
Plasma viral load, copies/mL	Undetectable, n (%)	35 (81.4)	79 (92.9)	0.07
	Median of log detectable pVL (range)	4.4 (2.8-5.7)	4.7 (3.8-6.0)	0.35
Median CD4 count (range), cells/μL		336 (14-884)	406 (66-1211)	0.03
Median time on ART, days (range)		758 (-22-1988)	775.5 (-28-1803)	0.95

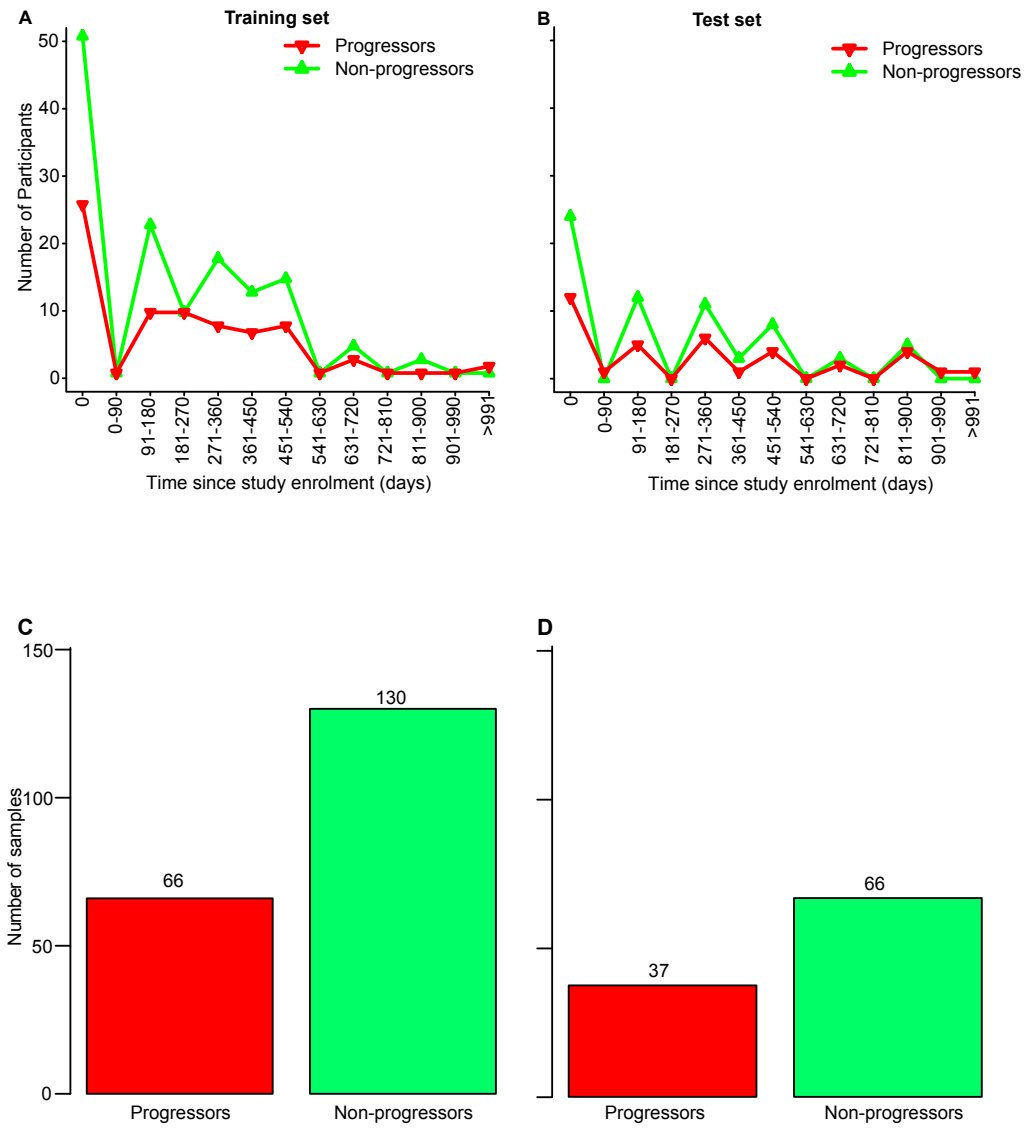


Figure 12: Distribution of samples in the training and test sets in the chosen iteration. (A&B) The number of participant samples at each 90-day time window in the training (A) and test (B) sets. (C&D) Total number of samples from progressors and non-progressors in the training (C) and test (D) sets.

4.4.2: QC of qRT-PCR data

Seven 96.96 GE chips were required to run qRT-PCR on all 299 samples in duplicate. Before blind predictions, Ct values of all assays for the duplicate IPC samples from the seven chips were analysed using Spearman's correlation test to determine consistency between the runs. The correlation coefficients between all the chips exceeded 0.95 (Figure 13A). We also compared the distributions of the average dCt of GOI primer-probes between the chip runs, to ensure that the distribution of primer of interest expression was similar between the chip runs. No marked differences in these distributions between the chip runs were observed as shown by overlapping histogram plots (Figure 13B). In addition, signature scores were compared between the seven chips and found to differ by less than 10% for both the ACS 11-gene (score range: 87 to 94) and the rPSVM.1 (score range: 90 to 94) signatures. To assess consistency of the QC, the signature score script, and its application to the Fluidigm qRT-PCR data, both SATVI and collaborators in CIDR performed blind predictions on the raw qRT-PCR data. The signature scores between the two institutes were compared before unblinding of progressor/non-progressor class, and were found to be identical (data not shown). The training set was unblinded first, and the performance of the signatures assessed.

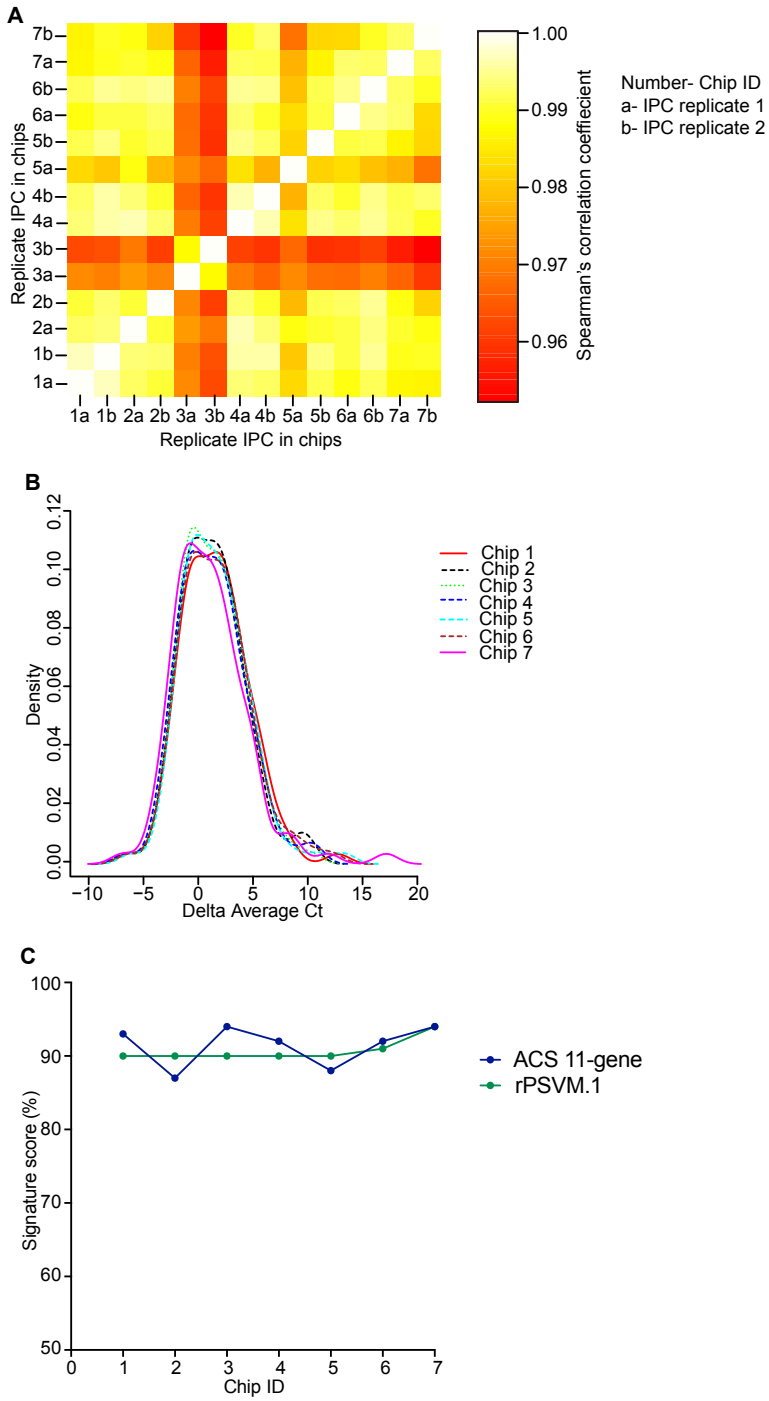


Figure 13: Correlation of duplicate internal positive controls (IPC) in the seven qRT-PCR chips. (A) Correlation of duplicate Ct values of all 96 assays of the IPC in the seven chips. Each square represents a correlation coefficient between two replicates from different chips. Numbers denote the chip IDs and the letters denote the replicates. Correlations were calculated by Spearman correlation. (B) Distribution plots depicting the distribution of delta average Ct values in the IPC of the seven chips. (C) Consistency in ACS 11-gene and rPSVM.1 signature scores obtained for the IPC sample run on all seven Fluidigm chips.

4.4.3: Signatures can only predict recurrent TB in the three months preceding diagnosis in the training set

To determine if the ACS 11-gene or rPSVM.1 can differentiate between recurrent TB progressors and non-progressors we stratified samples into six-monthly time windows relative to recurrent TB diagnosis. Samples within six months of recurrent TB diagnosis were further stratified into two three-monthly time windows. We first compared signature scores of all progressor and non-progressor samples, irrespective of time to recurrence, to determine if the signatures had any prognostic potential during the entire follow-up period. The signatures could significantly differentiate between progressors and non-progressors but only weakly (ACS 11-gene, AUC 0.65 (95% CI 0.57 to 0.73), p=0.0005; rPSVM.1, AUC 0.66 (95% CI 0.57 to 0.75), p=0.0002 (Figure 14A &B, Table 10). No difference in AUC was observed between the two signatures (p=0.55). When stratified into six-monthly time-to-TB diagnosis windows only the ACS 11-gene signature predicted recurrent TB disease in the six months preceding diagnosis (Figure 14A & B, Table 10). To understand the effect of time further we stratified samples within this six-month time window into two three-monthly time windows. Both signatures significantly differentiated between recurrent TB progressors and non-progressors only three months preceding diagnosis (Figure 14C & D, Table 10). Performance of the ACS 11-gene signature during this period was highly significant (p=0.003) whilst that of rPSVM.1 was barely significant (p=0.01) despite equivalent ROC AUCs of 0.77 (95% CI: 0.61 to 0.93) and 0.74 (0.57 to 0.91), respectively (Table 10). These AUCs were not different by pROC analysis, p=0.82. These results suggest that the differentiation observed for ACS 11-gene signature within six months of recurrence diagnosis was driven by signal detected in the last three months before recurrence.

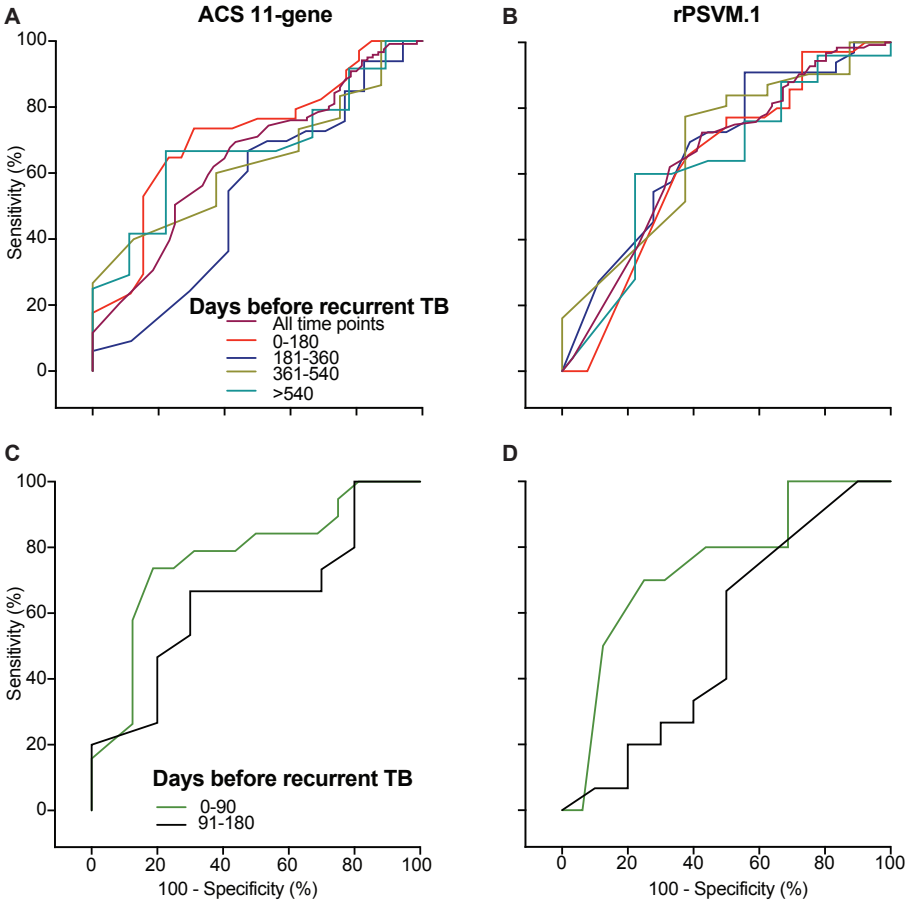


Figure 14: Performance of signatures in the TRuTH training set. ROC curves depict predictive potential of recurrent TB during HIV infection for discriminating progressors and non-progressors before diagnosis. (A&B): ROC curves correspond to 180-day intervals relative to TB recurrence and all time points combined for the ACS 11-gene (A) and rPSVM.1 (B) signatures. (C&D) Each ROC curve corresponds to a 90-day time interval relative to TB recurrence for the ACS 11-gene (C) and rPSVM.1 (D) signatures. Performance statistics are in Table 10 below.

Table 10: ROC AUC depicting the performance of the ACS 11-gene and rPSVM.1 signatures in the TRuTH training set

Days to recurrent TB	ACS 11-gene		rPSVM.1	
	ROC AUC (95% CI)	p-value	ROC AUC (95% CI)	p-value
All time points	**0.65 (0.57-0.73)**	**0.0005**	**0.66 (0.57-0.75)**	**0.0002**
By 3-month period				
0-90	0.77 (0.61-0.93)	0.003	0.74 (0.57-0.91)	0.01
91-180	0.64 (0.41-0.87)	0.13	0.47 (0.22-0.73)	0.60
By 6-month period				
0-180	0.72 (0.58-0.85)	0.0043	0.63 (0.49-0.78)	0.08
181-360	0.54 (0.36-0.71)	0.67	0.68 (0.52-0.84)	0.03
361-540	0.64 (0.45-0.83)	0.24	0.68 (0.46-0.89)	0.12
>540	0.68 (0.48-0.87)	0.12	0.63 (0.41-0.86)	0.24

4.4.4: Performance of transcriptomic signatures in the combined TRuTH training and test set

We expected the rPSVM.1 signature to perform better than the ACS 11-gene signature in predicting recurrent TB disease in the training set because the rPSVM.1 signature was specifically adapted and parameterised for PBMC and HIV infection. However, we observed equivalent performance between the rPSVM.1 and ACS 11-gene signatures showing that adaptation of such signatures does not necessarily improve performance. We also decided against generating yet another signature based on the premise that application of an existing signature to a new cohort provides a more valuable assessment of predictive performance than fitting a revised or adapted signature to a new dataset. We therefore decided that further re-parameterisation of signatures would not be warranted, unblinded the progressor/non-progressor status of the TRuTH test set, and combined this with the training set for analysis of the entire cohort providing a larger sample size and more statistical power. Stratified into six-monthly windows, the ACS 11-gene and rPSVM.1 signatures significantly predicted recurrent TB within six months of TB diagnosis

(Table 11). The signatures also predicted recurrent TB disease within three months of recurrent TB diagnosis; the ACS 11-gene signature with an AUC of 0.72 (95% CI: 0.58 to 0.85, p=0.004, Table 11) and the rPSVM.1 signature with an AUC of 0.73 (95% CI: 0.59 to 0.87, p=0.002; Table 11). These AUCs were equivalent by pROC analysis (p=0.49), confirming what was observed in the training set.

Table 11: Performance of the ACS 11-gene and rPSVM.1 signatures in the combined training and test sets of the TRuTH cohort.

Days to recurrent TB	ACS 11-gene		rPSVM.1	
	ROC AUC (95% CI)	p-value	ROC AUC (95% CI)	p-value
By 3-month period				
0-90	0.72 (0.58-0.85)	0.004	0.73 (0.59-0.87)	0.002
91-180	0.63 (0.43-0.82)	0.11	0.53 (0.33-0.74)	0.36
By 6-month period				
0-180	0.68 (0.57-0.79)	0.005	0.65 (0.53-0.76)	0.02
181-360	0.58 (0.45-0.71)	0.23	0.68 (0.56-0.80)	0.006
361-540	0.57 (0.39-0.75)	0.46	0.65 (0.46-0.83)	0.14
>540	0.59 (0.44-0.75)	0.26	0.63 (0.48-0.78)	0.13

4.4.5: Lower signature scores in non-progressors than in progressors in the three months preceding TB recurrence, despite heterogeneity of scores

Next, the longitudinal kinetics of the signature scores were investigated to determine any temporal changes in relation to time to recurrent TB diagnosis. Longitudinal median scores of either signature were not different over time between the progressors and non-progressors, except during the three months preceding recurrent TB diagnosis when scores were significantly higher in progressors than non-progressors (Figure 15A & B), mirroring what was observed in the ROC AUCs. Kinetics of longitudinal signature scores were highly variable in individual progressors and non-progressors (Figure 15C & D), highlighting the high degree of heterogeneity of IFN responses in HIV-infected persons with previous TB disease.

4.4.8: Signature scores were not associated with time to sputum culture positivity at recurrent TB diagnosis

In a recent study of TB disease progression in non-human primates, magnitudes of Type I IFN signatures in macaques was highly correlated with the extent of disease and FDG uptake as measured by PET-CT (Gideon et al., 2016). Similarly, in a human study of TB treatment, FDG uptake and Type I IFN were associated with the presence of mycobacteria in the sputum of HIV-uninfected persons during TB treatment (Malherbe et al. 2016; Thompson et al. 2017). These studies suggest that Type I IFN gene expression signatures and mycobacterial load are associated with disease severity at the site of infection. Hence, we set out to determine the relationship between signature scores of progressor samples collected within three months of diagnosis and mycobacterial load in sputum, using time to culture positivity as a measure of disease severity. There was no association between signature scores and time to culture positivity in a subset of cases (n=18) who had data available on time to culture positivity (Figure 15E &F).

4.4.9: Time since previous TB diagnosis, time on ART and number of previous TB episodes were not associated with signature scores

In the Malherbe *et al* study of TB treatment, inflammation detected in the form of PET-hot lesions measured by PET-CT was observed in HIV-uninfected persons up to a year after end of successful TB therapy (Malherbe et al., 2016). We investigated the effect of time since previous TB diagnosis and number of previous TB episodes on on-going inflammation measured by signature scores. Since effective ART results in lower risk of active TB, decreased immune activation and inflammation in HIV infection (Klatt et al., 2013), the effect of time on ART on signature scores was

measured to determine the effect of ART on on-going inflammation. Time since previous TB and time on ART were calculated by subtracting the dates of ART start and TB treatment start (within seven days of diagnosis) from the sampling date. Some participants had previous TB episodes at the start of primary TB study (i.e. START/SAPiT) but the number of previous TB episodes was not recorded. Thus, we stratified the number of previous TB episodes into one or more than one in our analyses of the TRuTH study. The median time from previous TB to sampling time was 1,325 days (IQR: 1,042 to 1,605 days). No association was observed between time since previous TB and ACS-11 gene or rPSVM.1 signature scores. Time on ART (median = 1,089 days, IQR: 791 to 1,436 days) was also not associated with signature scores at the corresponding sampling time points. Signature scores were not different between persons with one previous TB episode and those with more than one previous TB episode, p=0.49 and p=0.43 for the ACS 11-gene and rPSVM.1 signatures,respectively. This disproves our hypothesis that signature scores will be higher in persons with more than one episode of previous TB.

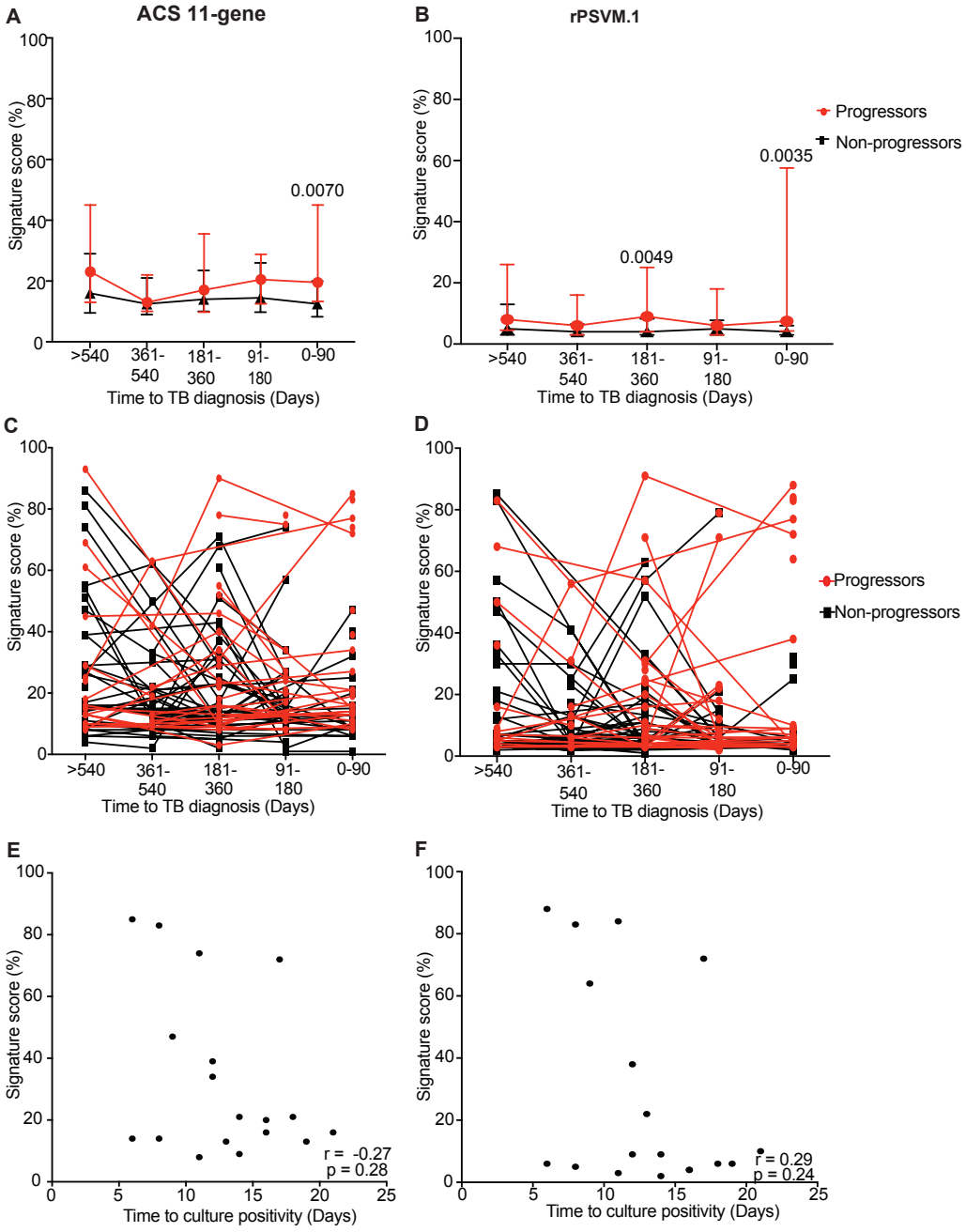


Figure 15: Longitudinal kinetics of signature scores in the TRuTH cohort. (A&B) Longitudinal kinetics of the ACS 11-gene (A) and rPSVM.1 (B) signature scores in progressors and non-progressors. Median and interquartile ranges are shown. (C&D) Longitudinal kinetics of the ACS 11-gene (C) and rPSVM.1 (D) signature scores in individual progressors and non-progressors. Dots represent time points and connecting lines represent individuals. (E&F) Association between the ACS 11-gene (E) and rPSVM.1 (F) signature scores and bacterial burden measured by days to culture positivity in sputum. Spearman r values are shown.

4.4.11: The ACS 11-gene and rPSVM.1 signatures could not differentiate *Mtb* infection from subclinical TB in ART naïve HIV-infected persons

The TRuTH study was designed to detect recurrent TB disease using very intensive case finding methods, including performing induced sputum collection at every study visit. It is important to note that most study participants who were diagnosed with recurrent TB were asymptomatic at the time of TB diagnosis (n=25), suggesting that many of the progressors had subclinical TB, rather than active TB disease. Our previous work with the ACS 16-gene signature showed that subclinical TB could be detected in asymptomatic HIV-uninfected progressors up to a year before active TB diagnosis (Zak et al. 2016). These data support our interpretation that the ACS 11-gene and rPSVM.1 signatures are detecting subclinical disease in the TRuTH cohort, within three months of recurrent TB disease diagnosis.

To address this further, we aimed to determine if the signatures can differentiate between subclinical TB and *Mtb*-infected controls and active TB cases in HIV-infected persons. We hypothesized that the signatures can differentiate subclinical TB from *Mtb* infection in HIV-infected persons.

In order to address this, the ACS 11-gene and rPSVM.1 signatures were applied to a small cross-sectional cohort of HIV-infected individuals established by Esmail and Wilkinson, wherein subclinical TB was diagnosed by PET-CT in 10 asymptomatic individuals with *Mtb* infection who were identified to have active pulmonary lesions (Esmail et al., 2016, 2018). A group of 21 asymptomatic HIV-infected individuals with *Mtb* infection and no active pulmonary lesions were included as controls while 15 patients with microbiologically-confirmed active TB were included as cases. As

observed in the pilot cohort in Chapter 3, the signatures could readily differentiate active TB from *Mtb*-infected controls (Figure 16, Table 12). The ACS 11-gene signature could also differentiate subclinical TB from active TB, whilst the rPSVM.1 signature did not significantly differentiate between these two groups (Figure 16A & B, Table 12). Neither signature could differentiate subclinical TB from *Mtb* infection. It was very notable that signature scores were generally very high in all three groups, including the *Mtb*-infected controls, perhaps suggesting very high HIV loads. Regardless, median ACS 11-gene signature scores were higher in active TB cases than in subclinical TB and *Mtb*-infected controls (Figure 16C). The rPSVM.1 signature scores from the active TB cohort were only higher than that of the *Mtb*-infected control group (Figure 16D).

Table 12: Performance of the signatures in distinguishing subclinical TB from *Mtb* infection and active TB disease in ART naïve HIV-infected persons from the Esmail cohort (Esmail et al., 2016, 2018)

	ACS 11-gene		**rPSVM.1**	
	ROC AUC (95% CI)	**p-value**	**ROC AUC (95% CI)**	**p-value**
Subclinical vs *Mtb* infection	0.55 (0.33-0.76)	0.69	0.64 (0.42-0.86)	0.20
Subclinical vs Active TB	0.77 (0.56-0.98)	0.03	0.71 (0.49-0.92)	0.09
Active TB vs *Mtb* infection	0.80 (0.65-0.95)	0.0027	0.85 (0.72-0.98)	0.0004

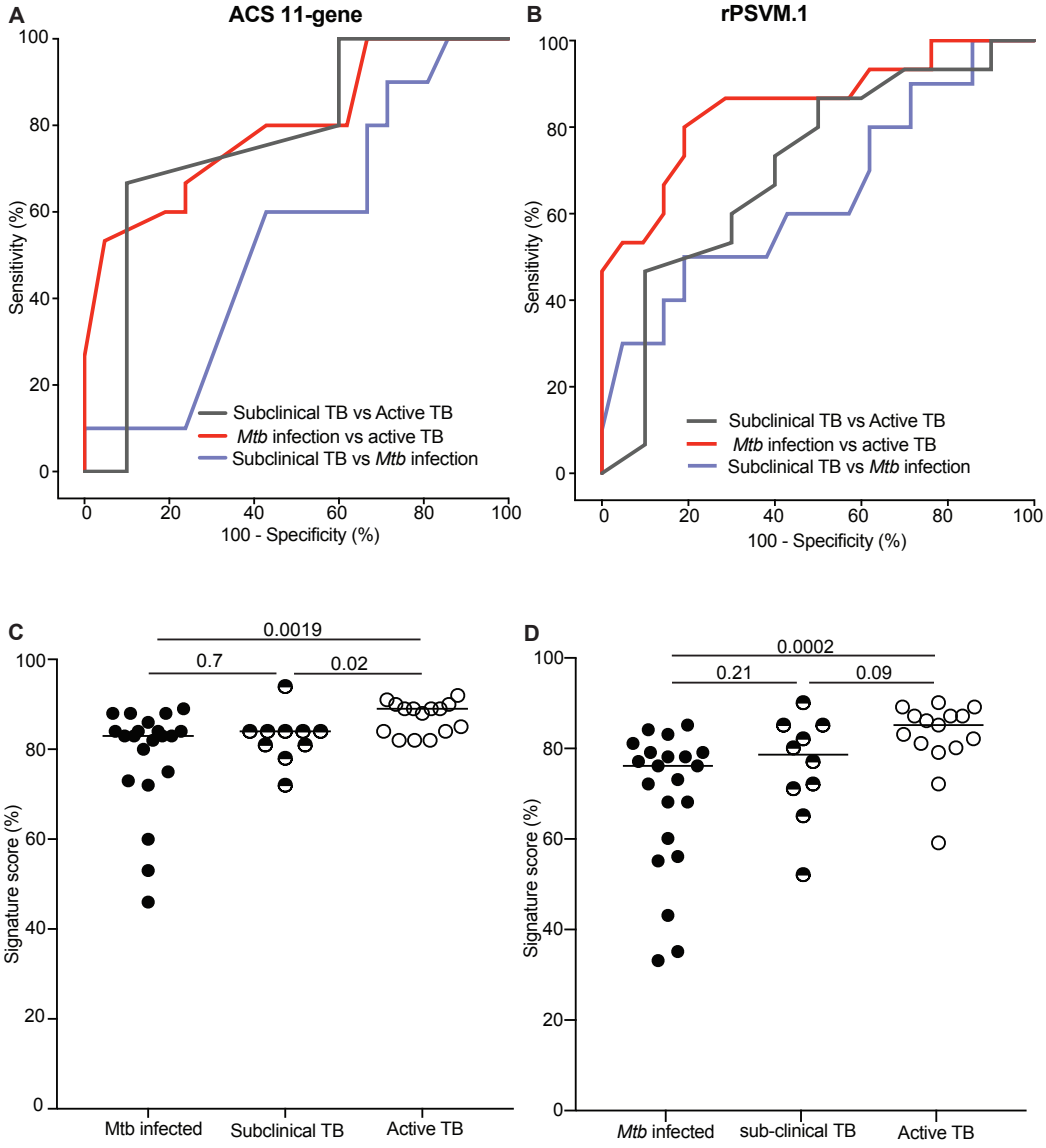


Figure 16: Ability of the signatures to distinguish subclinical TB from *Mtb* infection and active TB disease in the Esmail cohort. (A and B) ROC curves depict classification potential of the ACS 11-gene (A) and rPSVM.1 (B) signatures in distinguishing subclinical TB from *Mtb* infection and active TB disease in HIV-infected persons. (C and D) Distribution and differences in ACS 11- gene (C) and rPSVM.1 (D) signature scores between the three classes of TB.

4.5: Discussion

HIV infection, as well as a history of previous TB disease, increases risk of progression to recurrent TB disease. According to current policy, *Mtb*-infected persons who are HIV-infected should receive INH prophylaxis (World Health Organization 2011). However, poor adherence and incomplete sterilisation of the bacilli undermines the effectiveness of the prophylaxis and in practice many HIV-infected persons do not receive or adhere to IPT. Recurrent TB disease adds to the infectious pool of TB cases. It is therefore imperative to identify those most at risk of developing recurrent TB disease irrespective of HIV status. In this project, we aimed to identify HIV-infected persons already on ART who are at risk of progressing to recurrent TB disease.

Both the ACS 11-gene and rPSVM.1 signatures could differentiate between recurrent TB cases and controls, but only within three months of diagnosis in this cohort. Intensive active case finding was performed in the TRuTH study by induced sputum at every visit and most participants (25 out of 43, 58%) were asymptomatic at time of diagnosis. This suggests that the signatures detected subclinical TB disease. This is consistent with our finding in the ACS study (Zak et al. 2016) where the ACS 16-gene signature significantly predicted incident TB at least a year before diagnosis suggesting detection of subclinical TB many months before symptoms manifested in HIV-uninfected persons. A number of the asymptomatic TRuTH participants with positive sputum tests who refused to start TB therapy developed clinical symptoms after diagnosis (Dr Kogie Naidoo, personal communication). This further strengthens the hypothesis that many participants had subclinical TB at

diagnosis and may explain the short prognostic window of the transcriptomic signatures relative to those reported in the ACS study. Epidemiological field studies have also shown that subclinical culture positive TB disease can be diagnosed when intensified case finding for active TB is carried out in household contacts or at HIV care clinics in high TB transmission settings (Corbett and MacPherson 2013). Furthermore, a review of surveys conducted in Asia observed that up to 40% to 60% of bacteriologically positive TB cases were asymptomatic (Onozaki et al. 2015). This study confirms that active TB case-finding in asymptomatic persons can result in detection of subclinical TB that might have remained undetected, and are likely sources of on-going transmissions in their communities. HIV infection is known to rapidly increase progression to active TB disease and IPT should be given to HIV-infected persons to decrease the risk of progression to active TB disease. However, several studies have observed an increase in TB incidence shortly after the end of IPT (J. E. Golub et al. 2015; Samandari et al. 2015) suggesting rapid progression to active TB immediately after stopping IPT in HIV-infected persons. This may be due to incomplete cure or immediate reinfection. In the TRuTH cohort, study participants eligible for IPT (n=212) were initiated on six months of daily IPT (Maharaj et al. 2017) and most developed recurrent TB within three months of stopping IPT as observed previously. It is tempting to speculate that cessation of IPT in these individuals could also have increased the risk of progression to recurrent TB disease in these individuals, which may have contributed to the short prognostic window of the signature for recurrent TB.

To further determine if the transcriptomic signatures can distinguish between subclinical TB and active TB, and between subclinical TB and asymptomatic *Mtb*-

infected persons, we applied the signatures to the the well-characterised Esmail cohort of subclinical TB (Esmail et al., 2016, 2018). The finding that the signatures could not differentiate subclinical TB from *Mtb* infection in HIV-infected persons is to some degree consistent with the finding that recurrent TB cases in the TRuTH study could not be differentiated from controls until three months before diagnosis. This is very likely also due to the small range in signature scores, which were also universally high in the different groups in the Esmail cohort. The elevated scores were likely due to high HIV loads (further discussions in Chapter 6).

In the ACS cohort, signature scores in progressors increased closer to TB diagnosis indicating increasing IFN responses, with low and more stable signature scores in non-progressors (Thompson et al. 2017; Zak et al. 2016). However, this trend was not clear in progressors and controls in the TRuTH cohort. Kinetics of signature scores in individuals showed very high degrees of heterogeneity over time in both progressors and non-progressors. Samples analysed in this study were collected as early as 706 days and up to 2,295 days after diagnosis of the first TB episode. We propose that ongoing inflammation from this previous disease process could have been detected in our analyses. This is supported by the study by Malherbe *et al*, which shows that pulmonary inflammation, measured by PET-CT, was present in many HIV-uninfected persons up to one year after end of TB treatment, even in patients with microbiological cure (Malherbe et al. 2016). To delve deeper into this, we analysed if time since previous TB episode was associated with the magnitude of Type I IFN gene expression at the early study time points in the TRuTH cohort. No association between time since previous TB and signature scores was observed. Although not significant, signature scores appeared to be higher at time points closer

to previous TB disease, again suggesting an effect of the previous TB episode on signature scores. ART decreases viral replication and associated immune activation and inflammation in HIV-infected persons. Duration of time on ART is therefore another important consideration. However, we did not observe an association between signature scores and time on ART. This is consistent with another study that showed ART duration was also not associated with risk of TB in a South African community cohort with a recurrent TB burden of 75% (Gupta et al., 2012). Since ART decreases plasma viral load levels and undetectable plasma viral load levels are associated with a decrease in the risk of TB, the effect of plasma viral loads on the signature scores is also a critical consideration. This is discussed in detail in Chapter 6.

A previous study in NHPs suggest that type I/II IFN expression signatures are associated with the amount of *Mtb* bacilli in the lungs (Gideon et al., 2016). The degree of pulmonary inflammation as detected by FDG uptake on PET-CT scans was associated with type I/II IFN signatures in NHPs and humans (Gideon et al. 2016; Thompson et al. 2017). We observed no association between signature scores and days to culture positivity measured by MGIT culture positivity as a proxy for bacterial burden in the lungs. A major limitation of this analysis was the small number of available samples with time to culture positivity data (n=18).

The poor prognostic performance of the signatures in predicting recurrent TB at time points earlier than three months before diagnosis in this population could be due to the rapid progression to recurrent TB disease in HIV-infected persons. Epidemiological studies in high-risk populations have observed a higher recurrent TB

disease progression rate in HIV-infected persons in comparison to uninfected persons (Charalambous et al., 2008; Fitzgerald et al., 2000; Mallory et al., 2000). Recurrent TB disease occurred within two years after end of TB treatment in HIV-infected persons from these cohorts. ART was not readily available when these studies were carried out, thus ART status of participants is unknown. In our cohort, recurrent TB disease occurred at a median of 3.5 years after end of previous TB episode. This longer time to recurrent TB diagnosis was very likely due to provision of ART while IPT was also given to a number of participants.

We show here that the prognostic value of the ACS 11-gene and rPSVM.1 signatures for predicting recurrent TB disease in HIV-infected persons was low. Prognostic ability of these signatures may be better in HIV-infected persons who have not previously had TB. We are currently testing the diagnostic and prognostic value of the ACS 11-gene signature in a study of 860 South African HIV-infected adults, namely the *Correlates of Risk Targetted Intervention Study-High Risk* (CORTIS-HR) study. Results from this study will address important questions about the prognostic ability of the signature in asymptomatic HIV-infected persons and the effect of ART and viral load in a much larger sample size.

4.6: Contributions

Sample selection and matching were done by Mrs N. Yendeh-Zuma and Dr S.M. Kimbung. R scripts to calculate signature scores were written by Drs S.M Kimbung and E.G Thompson. Ms F. Darboe designed,conducted the experiments, performed data analyses, and wrote this chapter under the supervision of Dr A. Penn-Nicholson and A. Prof T.J. Scriba.

Chapter 5: Treatment response monitoring in HIV-infected persons on ART: results from the IMPRESS cohort

The aim of this chapter is to determine if the transcriptomic signatures can differentiate late culture conversion from early conversion in HIV-infected persons on ART.

5.1 Introduction

The current standard chemotherapy for pulmonary TB entails at least six months of drugs heavily reliant on isoniazid and rifampicin. For the majority of TB patients, treatment results in disease cure, however some do not achieve sterilisation or clearance of *Mtb* and are at risk of relapse (Wallis et al. 2000). Poor treatment adherence is one of the factors contributing to insufficient treatment, leading to a high risk of treatment failure and relapse. Shortening the six-month regimen is one of the goals of the global strategy to combat the TB pandemic and is an active area of research. Mathematical modelling predicts that shortening of treatment regimens would increase adherence and successful treatment rates (Abu-Raddad et al. 2009). Several studies have evaluated new treatment regimens aimed at shortening TB treatment, for example using fluoroquinolones (Gillespie et al. 2014; Burman et al. 2006; Rustomjee et al. 2008; Merle et al. 2014; Jindani et al. 2014). Sterilising activity of the new TB treatment regimens is typically measured by their ability to prevent treatment failure and relapse after treatment completion (Burman et al. 2006). However, large clinical trials with long follow-up times are needed to study relapse and treatment failure rates. Human studies to test treatment regimen shortening with an endpoint to measure relapse rates require at least two years

follow-up after the end of chemotherapy of at least 1,000 individuals (Mitchison 1993; Burman et al. 2006).

Currently, sputum conversion from positive to negative (either by sputum smear microscopy or culture) is used as the marker of cure and potential relapse. The rate of relapse is strongly correlated with sputum conversion after two months of therapy (Mitchison, 1993). Despite high proportions of culture conversion in the trials evaluating new shortened treatment regimens, relapse rates were higher than the standard of care regimen, two months of HRZE and four months of HR (Gillespie et al. 2014; Burman et al. 2006; Rustomjee et al. 2008). A systematic review of 28 published studies observed a low sensitivity (40%) and modest specificity (85%) of sputum culture conversion in predicting treatment failure and relapse (Horne et al. 2010). These studies suggest that culture conversion at two months has a poor positive predictive value (18%) in predicting relapse. Recently, the Xpert MTB/RIF assay has been tested as a treatment response marker in South African and Tanzanian adults with smear positive TB (Friedrich et al. 2013). *Mtb* DNA was associated with both sputum smear positivity (or grade) and the time to culture positivity. Despite the high sensitivity of Xpert, it had a low specificity (48.6%), possibly due to detection of dead bacillli, impeding its' usage as a biomarker for treatment response monitoring. Thus, better markers are therefore needed to determine treatment response.

As mentioned in previous chapters, uptake of ^{18}F-labelled fluorodeoxyglucose ([^{18}F] FDG) viewed on PET-CT scans is indicative of glucose metabolism. [^{18}F] FDG uptake is thus a useful marker of inflammation and has been successfully used to

track pulmonary inflammation during TB treatment (Malherbe et al. 2016). Observed inflammation patterns in lungs were associated with clinical outcomes in this study, suggesting PET-CT scans can be used to determine treatment outcomes in HIV-uninfected persons. However, measurement of [^{18}F] FDG uptake by PET-CT requires expensive CT and PET imaging scanners, which are not readily available in resource-poor settings with the highest disease burden. In addition, PET-CTs are not feasible due to large number of patients would require scans requiring trained radiologists, who are scarce in resource-poor settings, to interprete scans.

As discussed in previous chapters, blood transcriptomic gene signatures based on IFN response genes are able to distinguish TB disease *Mtb* infection and other diseases. These signatures could also be useful to monitor inflammatory resolution during treatment. Several studies have shown that gene expression of these signatures decrease upon TB treatment as early as two weeks post-treatment start (Berry et al., 2010; Bloom et al., 2013; Cliff et al., 2013; Sweeney et al., 2016), suggesting that the signatures may be used to monitor treatment response. The recently published ACS-16 gene signature was applied to the Catalysis Foundation for Heath Treatment Response Cohort (CTRC) described by Malherbe and colleagues (Malherbe et al. 2016) to determine if the signature can predict treatment outcome (Thompson et al. 2017). One hundred and thirty-one HIV-uninfected adults with newly diagnosed pulmonary TB were enrolled in the CTRC study and followed-up during TB treatment. Ninety-nine of these participants completed study follow-up and TB treatment was stratified based on sputum culture results at the end of six months of treatment. Eighty-four were defined as cured, eight were treatment failures (sputum culture positive at the end of the treatment) and the remaining seven with

unevaluable culture results were not included in the analyses. Eighteen of the 84 cured participants attained mycobacterial conversion within four weeks of start of treatment and were referred to as rapid converters. Treatment failure and cure could be differentiated as early as four weeks after commencement of TB therapy with the ACS-16 gene signature (AUC = 0.72, p = 0.02). At the end of 24 weeks of treatment the signature could strongly discriminate between treatment cures and failures (AUC=0.92, p=0.00005), confirming that the ACS 16-gene signature can predict treatment response in HIV-uninfected active TB cases undergoing therapy. To the best of our knowledge, no study has investigated blood transcriptomic signatures as treatment response monitoring tools in HIV-infected persons. Furthermore, it is not clear whether blood transcriptomic signatures can differentiate between HIV-infected persons who have achieved microbiological clearance after undergoing TB therapy for two months and those who have not. A blood-based biomarker that can make this differentiation at two months, when the intensive phase of TB treatment is typically completed, would be useful to guide the length of the continuation phase.

Diagnosis of active TB disease in HIV-infected persons is dependent on sputum smears or cultures, similar to HIV-uninfected individuals. However, sputum collection from unwell and immunocompromised individuals is often difficult and sometimes not possible. Furthermore, HIV-infected persons often have paucibacillary TB disease and as a result may have smear and culture negative sputum test results. Because TB treatment response is currently monitored by sputum smear or culture conversion, HIV co-infected persons could be incorrectly classifies as “cured” could be a source of ongoing *Mtb* transmission in their communities. Blood-based transcriptomic signatures thus have the potential to be more accurate and more

universal TB treatment monitoring tools in HIV-infected persons and could allow customisation of the length of TB treatment.

The *Improving Retreatment Success* (IMPRESS) study was an open label, randomized controlled clinical trial in HIV-infected TB retreatment patients conducted by CAPRISA at the CAPRISA eThekwini clinical research site in Durban, South Africa. The aim of the study was to determine if a moxifloxacin-containing regimen, wherein moxifloxacin was substituted for ethambutol, was superior to the standard of care in improving treatment outcomes in patients with recurrent TB. Study participants were randomized to receive either the moxifloxacin-containing regimen or RIFAFOUR (HRZE), at TB diagnosis. The moxifloxacin-containing regimen consisted of a two-month intensive phase of daily isoniazid, rifampicin, pyrazinamide and moxifloxacin (2HRZM) and four months of the continuation phase of isoniazid, rifampicin and moxifloxacin (4HRM). Participants randomized into the control arm (standard of care) had a two-month intensive phase of isoniazid, rifampicin, pyrazinamide and ethambuthol (2HRZE) followed by a four-month continuation phase of daily isoniazid and rifampicin (4HR). All study participants also received a daily dose of the vitamin B6 supplement, pyridoxine, regardless of randomization arm.The two arms will henceforth be referred to as HRZM and HRZE arms, respectively. Participants were followed up for a total of 24 months, including six months of treatment and 18-months post treatment. HIV testing was mandatory to confirm HIV status. HIV-infected participants already on ART were only allowed into the study if their treatment regimen was not contraindicated with the study drugs. Participants were enrolled irrespective of CD4 T cell counts and whether they were on ART or not. Those not yet initiated on ART were provided ART. The study

protocol was amended after the study started enrolment to allow PAXgene sample collection so that the TB risk signatures could be measured. As a result, whole blood RNA was not available for all enrolled participants.

5.2: Aims and hypotheses

5.2.1: Aims

1. Determine if the ACS 11-gene and rPSVM.1 signatures could differentiate HIV-infected early sputum culture converters from late sputum culture converters at the start of therapy, two months into therapy, at the end of TB treatment, and six months after treatment in TB retreatment patients.
2. Determine if signature scores resolve in HIV-infected TB retreatment patients after successful TB treatment.
3. Determine if there is an association between signature scores and time to sputum culture conversion at baseline, two months after the start of therapy, at the end of treatment, and six months after treatment completion in HIV-infected patients.

Hypotheses:

1. The ACS 11-gene and rPSVM.1 signatures will differentiate early from late converters at the start of TB treatment, two months after treatment start, and at the end of TB treatment.
2. Signature scores will significantly decrease in HIV-infected TB retreatment patients from TB diagnosis to the end of TB treatment.

3. Sputum culture conversion at TB diagnosis will be associated with signature scores from baseline, two months, end of treatment and post treatment samples.

5.3: Methods

5.3.1: Study design

Participant enrolment into IMPRESS

Participants who satisfied the following criteria were enrolled into the study

Inclusion criteria:

- At least 18 years of age
- History of previous TB treatment
- Smear positive and rifampicin susceptible TB determined by Xpert MTB-RIF
- Karnofsky score (measurement of functional impairment) of at least 70
- A negative pregnancy test
- Female participants were required to agree to use one form of contraception

Exclusion criteria:

- Patients on a Nevirapine containing ART regimen at screening
- Pregnant or lactating women
- Received an antibiotic active against TB in the 14 days prior to screening
- *Mtb* resistant to any of the study drugs
- Allergies or intolerance to study drugs

Sputum samples for cultures were collected every two weeks in the first two months of treatment and monthly thereafter until completion of TB chemotherapy. PAXgene samples collected at baseline, two, six, eight and fourteen months after start of

therapy were available for this project. End of treatment time points were calculated according to when TB therapy was stopped. In addition to these time points, PAXgene samples were collected at further follow-up visits. These additional samples were not investigated in this chapter. Signature performance was evaluated in HIV-infected participants with a baseline sample and at least two other visit time points mentioned above.

Patient clinical outcomes in our analyses were defined based on time from study baseline (i.e. start of TB treatment) for sputum cultures to convert to negative. Time to culture conversion was calculated by subtracting the date of the first positive culture at diagnosis from the date of the first persistent negative culture result. Participants were stratified into “early converters” where sputum culture conversion occurred within two months (56 days) and “late converters” where culture converted after two months of treatment start. This data was unavailable until after all experiments and analyses were completed to ensure that wet-lab experiments and analyses were done in a blinded manner. Some participants stopped TB treatment at six months, whilst others completed treatment later based on clinical reasons. As such, the time at which they stopped TB treatment was used as the end of treatment time point.

5.3.2: RNA extraction

RNA from the PAXgene samples was extracted using the Tecan Freedom EVO 150 automated system as described in Chapter 2. RNA was not quantified after extraction because this extraction protocol is highly optimised and a very consistent

RNA yield is typically obtained, based on extensive experience from the *Correlate of Risk Targeted Intervention Study* (CORTIS) trial.

5.3.3: Gene expression measurement

cDNA was synthesized by reverse transcription using Taqman primer-probes as described in section 2.3.1. Gene expression levels of the signature transcripts were measured by Biomark HD multiplex microfluidic qRT-PCR as described in section 2.3.2.

5.3.4: Data analysis

Quality control of the data was performed as described in sections 2.4 and 2.5. Blind predictions were applied to the entire cohort before calculation of time to culture conversions. Culture conversion dates were made available by the clinical team only at the time of unblinding, once all signature scores had been generated for all samples and other blinded analyses had been completed. Differences in basic demographics were calculated using Mann-Whitney U test for continuous data and Fishers' exact test for categorical data in GraphPad Prism (v7.0a). ROC analyses was done in Prism, AUCs were compared using pROC and verification packages in R. Non-parametric tests, Mann-Whitney U and Kruskal-Wallis tests were used to analyse differences between two and three groups, respectively.

5.4: Results

5.4.1: Participant characteristics and sample stratification

Newly diagnosed HIV-infected and uninfected persons with a history of previous TB disease were screened and enrolled into the IMPRESS study (n=197), the majority

of whom were HIV-infected (n=139, 71%) (Figure 17). Of the 197 participants enrolled, only 154 had PAXgene samples collected at any time point and 115 of these participants were HIV co-infected. Due to the few HIV-uninfected participants with PAXgene samples available, these were excluded from analysis and we only focused on treatment response in HIV-infected patients. Furthermore, only participants with samples collected at three time points or more, including study baseline, were included in analysis in order that TB treatment response could be monitored. Of the 70 participants that were selected, seven were excluded from analyses. The reasons for these exclusions were prolonged TB treatment (n=2), treatment interruption (n=1), treatment failure (n=1), and no culture results at baseline/diagnosis (n=3) (Figure 17). Hence, 63 participants were included in the analysis, 33 of these were from the treatment arm (HRZM) and 30 were from the control arm (HRZE). Forty-four participants culture converted (became culture negative) within two months of diagnosis and 19 participants culture converted after 56 days (Figure 17).

Most of the participants in this substudy were male (n=44) (Table 13). When comparing early and late converters the proportion of males was not different between the two groups. In addition, race and age were also not different between the groups (Table 13). Median BMI of 20.0 Kg/m^2 was observed in the substudy; late converters had lower median BMI than early converters (p=0.05). Most of the early converters were from the HRZM arm, whilst late converters were mainly in the HRZE arm. No differences were observed in median detectable plasma viral load nor in the proportion of persons with undetectable plasma viral loads early and late

converters. Median CD4 T cell counts were lower in late converters in comparison to the early converters (p=0.05).

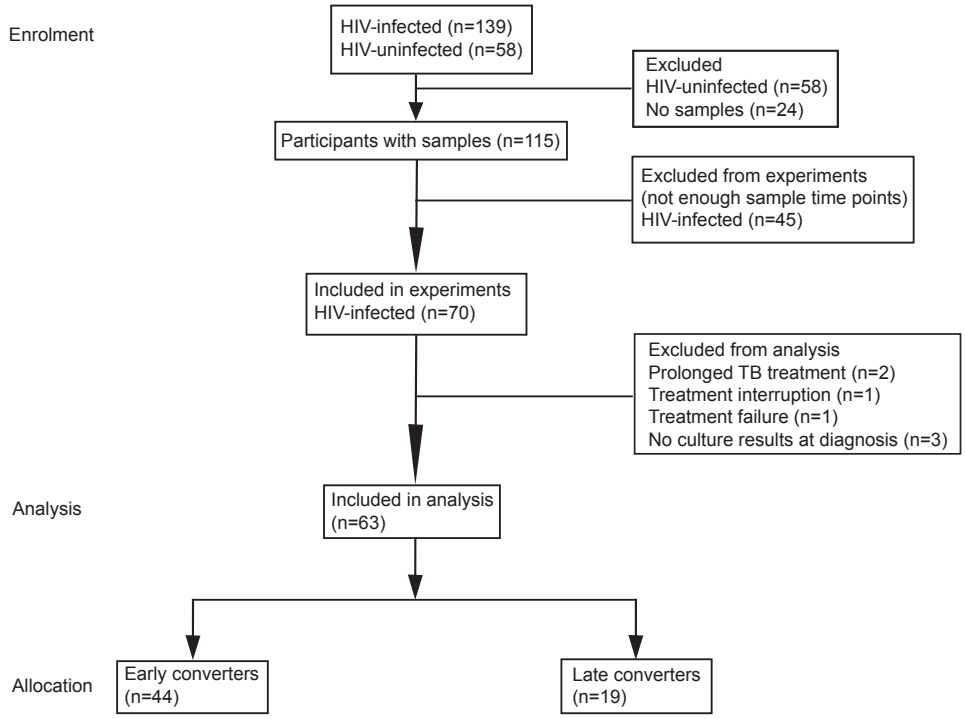


Figure 17: The IMPRESS cohort for the prediction of recurrent TB disease in people living with HIV. HIV-infected participants starting retreatment for TB disease were included in our biomarker study. Participants were allocated to early converters (culture conversion within two months) and late converters (culture conversion after two months).

Table 13: Characteristics of participants included in this biomarker sub-study at IMPRESS baseline.

Participant characteristic		Entire cohort (n=63)	Early converters (n=44)	Late converters (n-19)	p-value
Sex, n (%)	Male	44 (69.8)	28 (63.6)	16 (84.2)	0.1
Median age, years (range)		37 (19-51)	37 (21-51)	35 (19-51)	0.3
Race, n (%)	Black African	62 (98.4)	43 (97.7)	19 (100)	>0.9
	Cape Mixed Ancestry	1 (1.6)	1 (2.3)	0 (0)	
Median BMI, kg/m^2 (range)		20.0 (15.6-37.8)	21.1 (16.6-37.8)	19.4 (15.6-26.1)	0.05
Study arm, n (%)	HRZM	33 (52.4)	28 (63.6)	5 (26.3)	0.01
On ARV*, n (%)		28 (45.2)	19 (35.2)	9 (50)	0.8
Median CD4 counts**, (range)		268 (46-849)	285 (46-849)	236.5 (69-477)	0.05
Plasma viral load ***	Undetectable, n (%)	27 (45.0)	19 (45.2)	8 (44.4)	>0.9
	Median detectable pVL, log copies/mL (range)	4.7 (2.8-6.1)	4.7 (3.3-6.1)	4.3 (2.8-5.6)	0.6

* Only 62 participants have data available, early converters (n=44), late converters (n=18).
** Only 61 participants had CD4 T cell count information, early converters (n=43), late converters (n=18).
*** Only 60 participants had plasma viral load data, early converters (n=42), late converters (n=18). pVL limit of detection (LOD) was 400 copies/mL.

5.4.2: QC of qRT-PCR data

Samples were run in duplicate across a total of seven Fluidigm 96.96 GE chips. Ct values of the data from duplicate IPC reactions were analysed between the different chip runs using Spearman correlation test to determine consistency between the runs. The correlation coefficients between the between the chip runs exceeded 0.90 (Figure 18A). Distributions of normalised Ct (delta Ct) values of genes of interest were compared between the chip runs using histograms and were found not to be markedly different between the chip runs (Figure 18B). Finally, signature scores of the IPC were very consistent between the chips for both the ACS 11-gene (range: 92% to 96%) and rPSVM.1 (range: 86% to 93%) signatures (Figure 18C). These results confirm that the qRT-PCR data generated from different GE chips is consistent, inter-assay variability is low between the different chips and no obvious batch effects were observed.

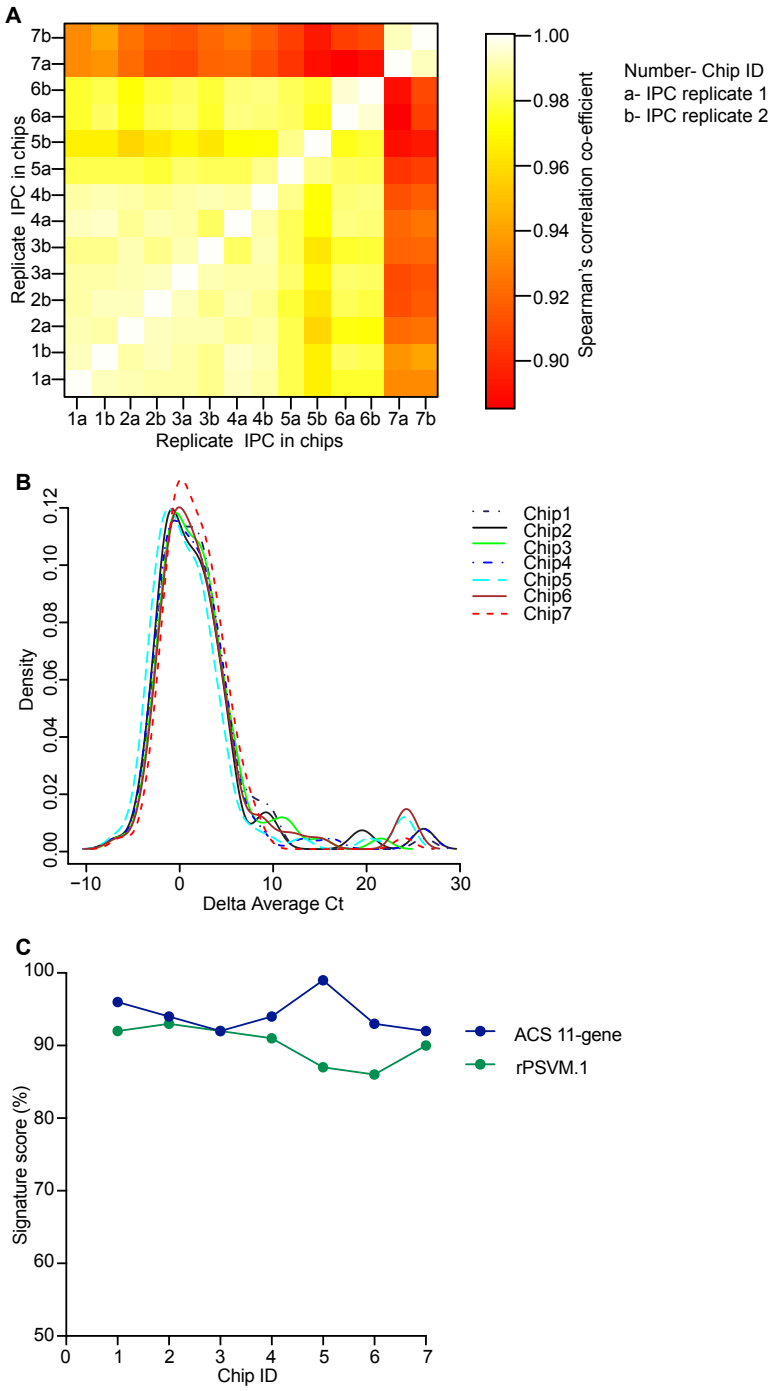


Figure 18: Variability of duplicate internal positive controls (IPC) in the seven qRT-PCR chips. (A) Correlation of duplicate Ct values of all 96 assays of the IPC in the seven chips. Each square represents the Spearman correlation coefficient computed between different chips. Numbers denote the chip IDs and the letters denote the replicates. (B) Distribution plots depicting the average delta Ct values in the IPC of the seven chips. (C) Signature scores of IPC in ACS 11-gene and rPSVM.1 signatures.

5.4.3: Signature scores decrease over treatment duration but do not resolve at the end of TB treatment

We set out to determine if retreatment of HIV-infected TB patients leads to resolution of inflammation, which we proposed would be reflected by decreases in ACS 11-gene and rPSVM.1 signature scores. Median signature scores decreased gradually from TB diagnosis to the end of treatment, but a considerable proportion of patients still had very high signature scores at the end of treatment, indicating ongoing type I/II IFN inflammation (Figure 19A & B). To determine if this inflammation would resolve further after the end of treatment, we also measured ACS 11-gene and rPSVM.1 signature scores six to eight months after treatment cessation. Surprisingly, ACS 11-gene signature scores six to eight months after TB treatment completion were not significantly different to scores observed at the end of TB treatment (Figure 19A). The rPSVM.1 scores increased significantly after TB therapy was stopped (Figure 19B). ROC analyses showed that both signatures could differentiate between the pre-treatment time points (i.e. at TB diagnosis) and the month-two treatment time point, the end of treatment time point and samples collected six to eight months after TB treatment (Figure 19C & D).

5.4.4: Signature scores were not different between early and late converters during TB treatment.

Next we sought to determine if patients could be classified into early and late converters based on their culture conversion status at two months, typically the end of the intensive phase of TB therapy. No striking differences in the kinetics of signature scores were observed between early and late converters (Figure 20A & B).

Median signature scores were not different at baseline, two months, and at the end of treatment in early and late converters for both the ACS 11-gene and rPSVM.1 signatures (statistical analyses are shown in Figure 22A & B). A very high degree of inter-individual heterogeneity in signature scores was observed in early and late converters (Figure 20C & D). To further unpack the kinetic changes longitudinal signature scores of participants with at least three time points were classified into; 1) no change (<10% difference between two treatment time points), 2) decrease (>10% change), 3) increase (>10% difference between consecutive time points), or 4) mixed. The 10% cut-off was derived from analysis of variability observed in signature scores of the internal positive control sample run across different chips, which was below 10%.

Forty-three participants in the early conversion group had at least two time points with ACS 11-gene signature scores available. Amongst these, signature scores decreased in 24 (56%), increased in one, did not change in five, and were mixed in 13 (30%) participants. Amongst 17 late converters with at least two available ACS 11-gene signature scores, scores decreased in 10 (59%), did not change in one, and were mixed in six participants. A similar pattern was observed for the rPSVM.1 signature. All 44 early converters had at least two time points with available signature scores for rPSVM.1. Signature scores decreased in 22 (50%), increased in one, did not change in two, and were mixed in 19 (43%) participants. Only 18 late converters had rPSVM.1 scores for at least two time points and the score decreased in seven (39%), increased in two, did not change in one participant, and were mixed in eight (44%) participants.

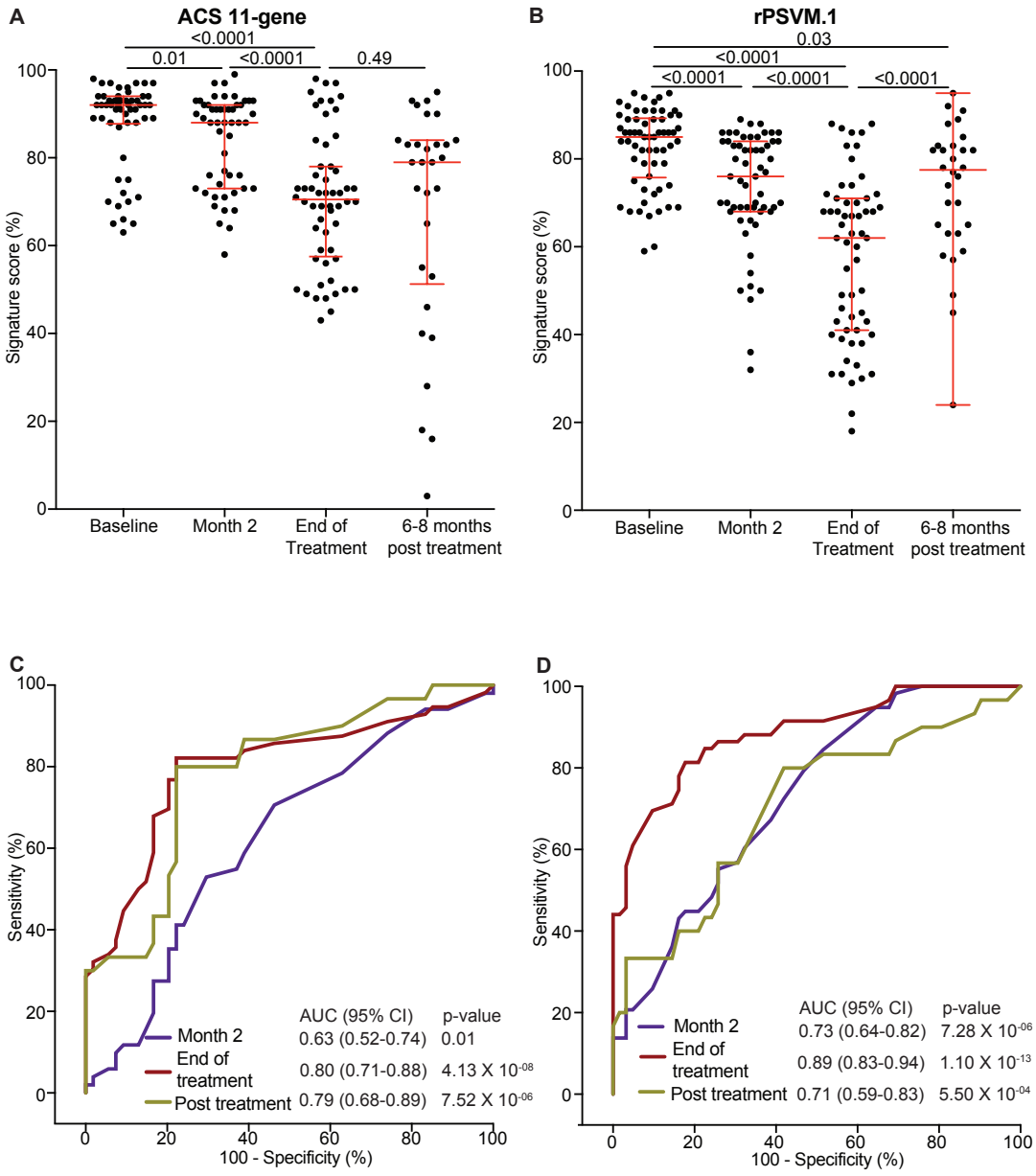


Figure 19: ACS 11-gene and rPSVM.1 signature score kinetics in the IMPRESS cohort over treatment duration. A and B) Signature scores decrease during TB therapy in the ACS 11-gene (A) and rPSVM.1 (B) signatures. C and D) ROC curves depict differentiation of baseline samples (collected before start of TB therapy) and samples collected during TB treatment (Month two), at the end of treatment or six to eight months post treatment stop.

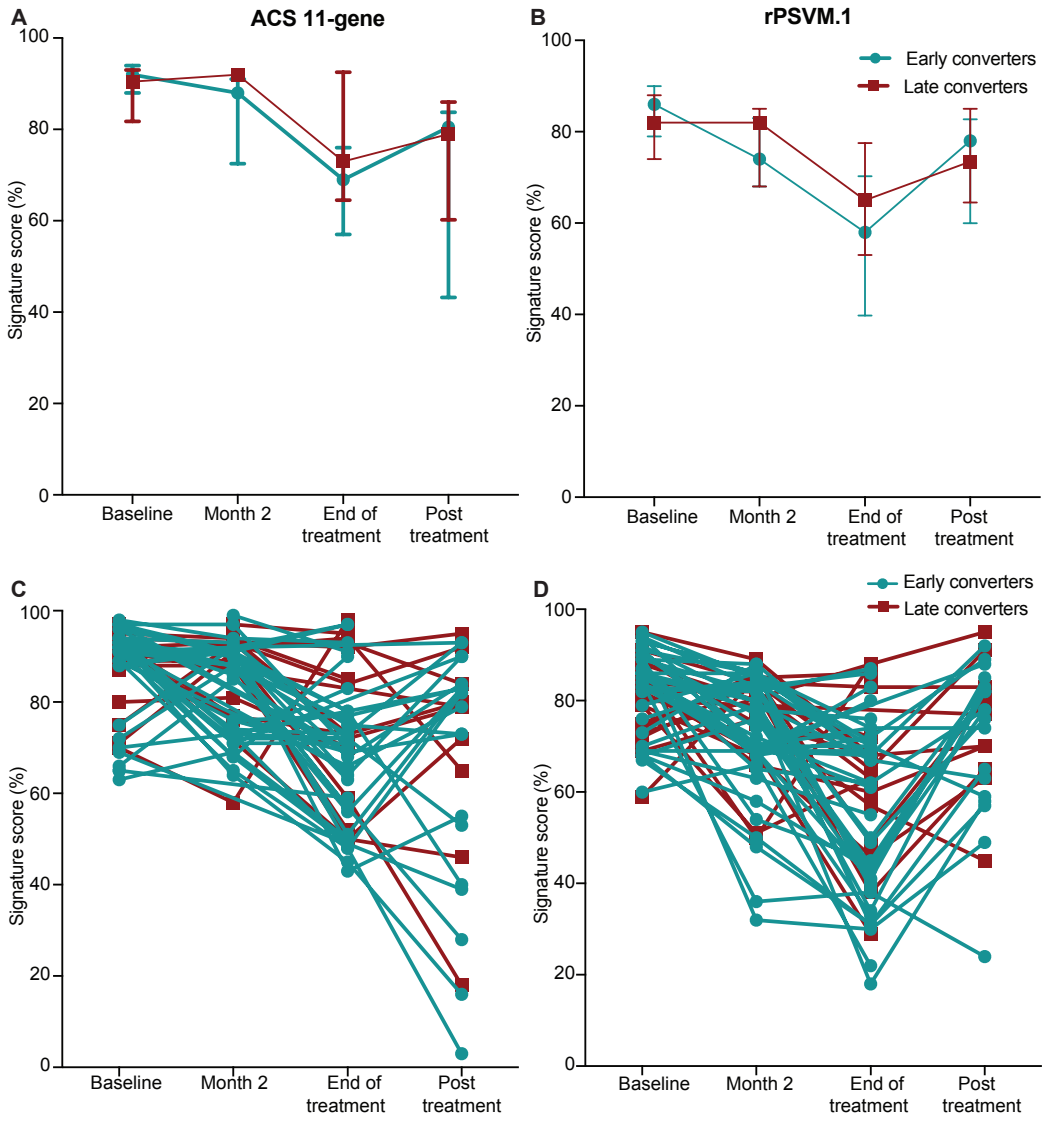


Figure 20: Kinetics of ACS 11-gene and rPSVM.1 signature scores in early and late TB treatment converters. A and B) Longitudinal kinetics of signature scores during and post TB treatment in early and late converters in the ACS 11-gene (A) and rPSVM.1 (B) signatures. Medians and interquartile ranges (IQR) are shown. C and D) Longitudinal kinetics of individual signature scores over time in early and late converters for the ACS 11-gene (C) and rPSVM.1 (D) signatures.

5.4.5: The ACS 11-gene signature can monitor treatment response in HIV-infected persons

At TB diagnosis (TB treatment initiation), neither the ACS 11-gene nor rPSVM.1 signature could predict early sputum conversion (Figure 21A & B). However, at two months the ACS 11-gene signature could differentiate early from late converters, AUC, 0.73 (95% CI 0.58 to 0.89), p=0.006. Ability of the signature to differentiate these groups was lost by the end of TB treatment and six to eight months after end of treatment (Figure 21A). The rPSVM.1 signature on the other hand could not classify late from early converters at any of the time points measured, whether before, during or after TB treatment (Figure 21B).

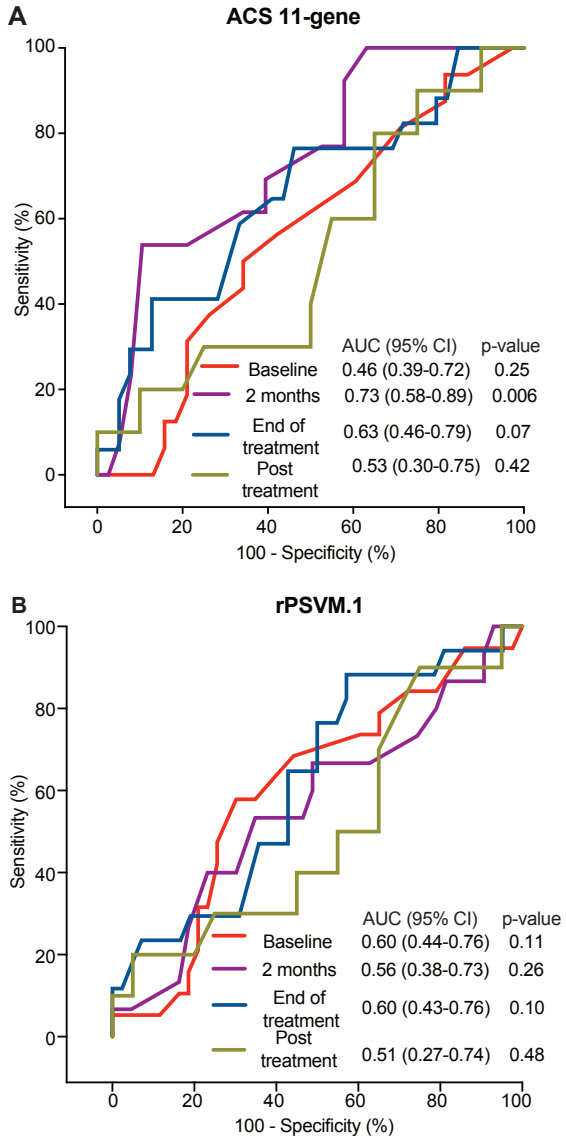


Figure 21: Differentiation between early and late converters by signatures in the IMPRESS cohort. Each ROC curve depicts signature differentiation between early and later converters. Each ROC curve corresponds to a time point during TB treatment for the ACS 11-gene (A) and rPSVM.1 (B) signatures.

5.4.6: Early sputum culture conversion is not associated with signature scores and time to culture positivity at diagnosis

ACS 11-gene signature scores were only significantly different between early and late conversion at two months after treatment start (Figure 22A), while rPSVM.1 signature scores were not different at any of the time points (Figure 22B). In HIV-uninfected TB patients, Xpert MTB-RIF Ct values measured as early as seven days post diagnosis have been shown to be predictive of time to culture conversion and treatment failure (Shenai et al. 2016). This suggests that sputum bacterial load measured early during treatment by Xpert MTB-RIF may be an important variable in understanding the behaviour of transcriptomic signatures. Unfortunately, Xpert MTB-RIF Ct values were not recorded in the IMPRESS study. However, in a paediatric study Xpert MTB-RIF detected 75% of mycobacterial growth indicator tube (MGIT) culture positive cases in HIV-infected and uninfected children with a high specificity (Nicol et al. 2011), suggesting that MGIT time to culture positivity data may reveal similar clues about sputum bacterial load and host inflammation. We therefore analysed time to culture positivity of MGIT cultures at IMPRESS baseline as a measure of bacterial load. However, time to culture positivity on MGIT cultures was not different between early and late converters (Figure 22C). We also sought to determine whether time to culture positivity would be associated with signature scores at TB diagnosis and if the combination of these two variables may be used to predict early and late converters.

However, signature scores were not associated with mycobacterial load in sputum and combining these two variables did not improve the prediction of early and late converters (Figure 22D & E).

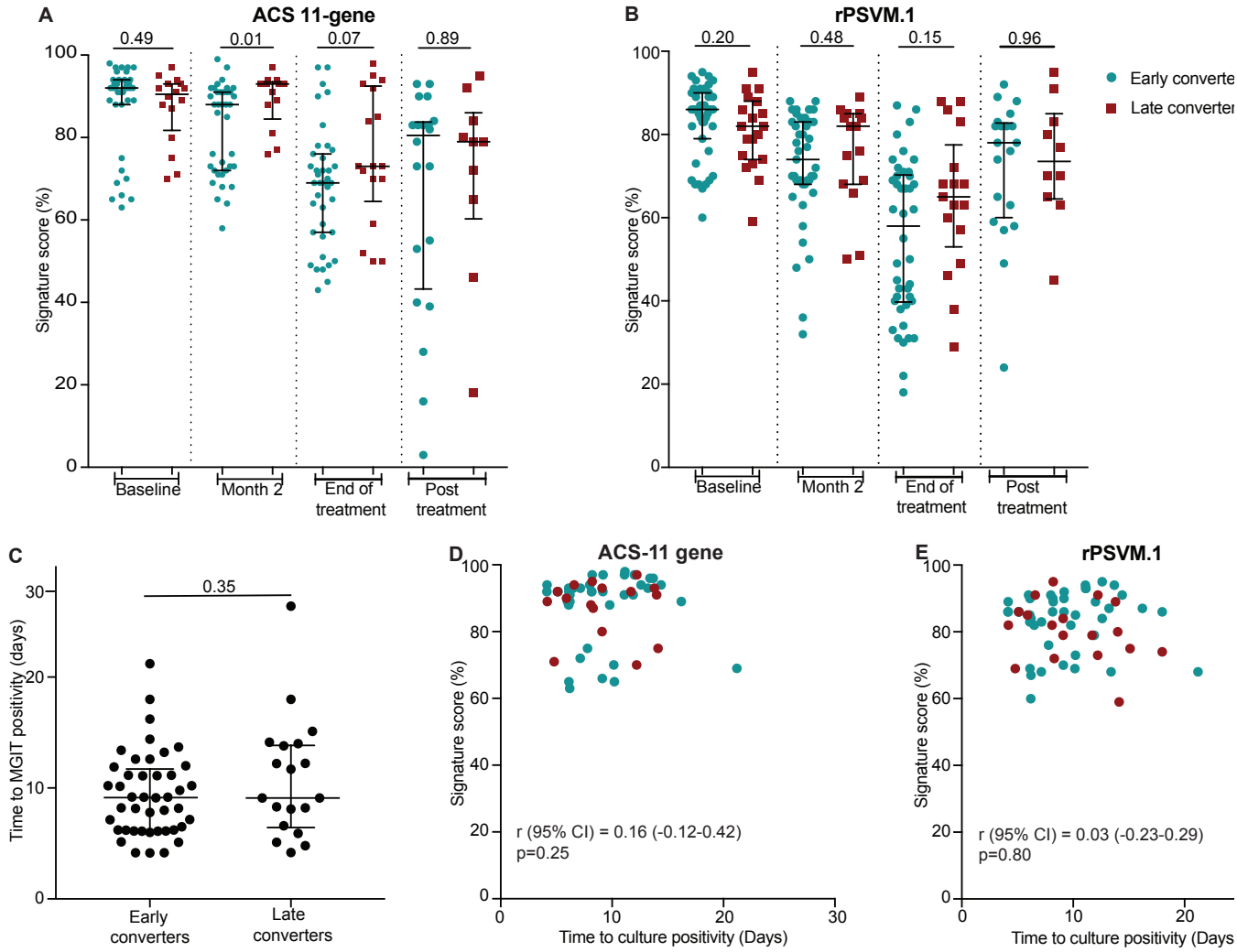


Figure 22: Association between signature scores and culture conversion in early and late treatment converters. A and B) ACS 11-gene (A) and rPSVM.1 (B) signature scores in early and late converters. C) Time to MGIT culture positivity at TB diagnosis in early and late converters. D) & E) Association between baseline ACS 11-gene (D) and rPSVM.1 (E) signature scores and time to culture positivity at TB diagnosis. Spearman's r correlation is shown.

5.5: Discussion

Poor adherence to TB therapy is a risk factor for treatment failure and relapse. Shortening of treatment regimens could increase adherence and consequently successful treatment rates (Abu-Raddad et al. 2009). High treatment success rates would result in a decrease in relapse cases, which are mostly due to treatment failure. Thus, predicting treatment response could play an important role in personalising therapy and testing the efficacy of drug regimens for active TB (Burman et al. 2006). Culture conversion after two months of intensive TB treatment is typically used as a proxy for treatment cure and microbiological sterilisation (Wallis et al. 2009). Furthermore, culture conversion after two months of therapy can predict relapse after completion of TB therapy (Mitchison 1993). Despite this, high relapse rates have been observed in clinical trials in participants who achieved mycobacterial clearance in their sputa after two months of therapy (Gillespie et al. 2014; Rustomjee et al. 2008; Merle et al. 2014). Data from these trials suggest that culture conversion after two months of TB treatment is insufficient for predicting relapse and monitoring treatment response, hence better markers are needed. Gene expression signatures of active TB disease could be useful markers in monitoring treatment response.

Gene expression signatures of active TB have been used to monitor treatment response in HIV-uninfected persons (Bloom et al. 2013; Sweeney et al. 2016; Berry et al. 2010), and the ACS 16-gene signature is no exception (Thompson et al. 2017). We show here that gene expression of both the ACS 11-gene and rPSVM.1 signatures can monitor treatment response in HIV-infected TB retreatment patients. Expression of the signatures in the participants was highly heterogeneous in the

IMPRESS participants, with some participants having persistently high signature scores at the end of treatment, as was observed in HIV-uninfected persons (Sweeney et al. 2016; Thompson et al. 2017). This suggests underlying inflammation and high type I/II IFN responses in these participants despite clinical and microbiological cure. Indeed, a heterogeneous pattern of pulmonary inflammation, measured by FDG-uptake, was also observed in the CTRC study, with some participants developing new lesions whilst a mixed response was observed in some participants (Malherbe et al. 2016). FDG uptake improved in 60% and resolved in 14% of the CTRC participants. Signature scores decreased in approximately 56% of the IMPRESS cohort at the end of treatment. However, because we did not have a matched control group to define signature scores in HIV-infected persons without TB with a history of previous TB disease, we could not define what a resolved response would be. This precluded distinguishing between resolved and improved response in this cohort. Underlying pulmonary inflammation was observed up to a year after successful cure in the CTRC study, suggestive of residual mycobacterial disease or bacterial activity despite control of disease after successful cure. This was also suggested by the detection of mycobacterial RNA in bronchoalveolar lavage samples (Malherbe et al. 2016). Consistent with this observation in HIV-uninfected patients, we observed underlying type I/II IFN response up to eight months after end of TB treatment in the IMPRESS cohort. The rPSVM.1 signature suggested that type I/II IFN activity increased up to eight months after successful cure, however the reason for this in not clear to us. Residual bacterial products may continue to trigger inflammation after recent TB disease, while circulating HIV loads may also trigger gene expression patterns in the periphery (the effect of HIV plasma viral loads on gene expression is discussed in the next chapter). Finally, this study was performed

in a high *Mtb*-transmission setting and reinfection may also contribute to elevation of signature scores. We conclude that our TB risk signatures can monitor TB retreatment response in HIV-infected persons, but with a reduced accuracy to that observed in HIV-uninfected persons.

When we applied the ACS 11-gene signature to identify rapid converters at baseline the signature could not correctly identify rapid converters at this time point. Our cohort consisted of HIV co-infected persons, most of who were not on ART (n=34) and had unsuppressed pVL (n=36) (see Chapter 6) at diagnosis. The ACS 16-gene signature could distinguish rapid from late converters at baseline in the CTRC and most importantly, treatment failures were identified at 24 weeks after the start of therapy (Thompson et al. 2017). This suggests that the ACS 16-gene signature could be a useful tool for customising duration of therapy, especially in individuals with treatment failure. The ACS 11-gene and rPSVM.1 signatures measure expression of ISGs. Expression of the transcripts in our signatures were higher around the TB diagnosis time points. However, both HIV infection and active TB disease induce expression of ISGs. Thus, the effect of active TB and HIV co-infection on type I/II IFN signature gene expression resulted in the inaccurate differentiation between early and late converters at TB diagnosis. After two months of treatment, signature expression remained elevated in some persons likely due to the presence of actively replicating bacteria. Thus late converters could be distinguished from early converters by the ACS 11-gene signature at month two. The gene signatures does not improve upon the diagnostic performance of culture conversion, but can be performed in individuals who cannot produce sputum and can be done in a much shorter time than culture, offereing some advantages.

Xpert MTB-RIF Ct values have been used as a measure of bacterial load and have been shown to predict treatment response and treatment failure in HIV-uninfected persons at diagnosis (Shenai et al. 2016). Time to culture positivity at diagnosis has been associated with sputum culture conversion at two months of TB treatment (Hesseling et al. 2010). In addition, sputum smear grading at diagnosis in the same study could predict treatment response. Thus, mycobacterial loads measured at baseline can be used to monitor treatment response in HIV-uninfected persons. In the IMPRESS study, culture conversion at month two was not associated with time to positivity at diagnosis. This raises the question whether mycobacterial load may be poorly predictive of treatment response in HIV-infected persons. In addition to HIV infection, the methodologies used to determine the time to positivity is different between the two studies. Hesseling and colleagues used the BACTEC 12B liquid radiometric culture method to determine time to positivity whilst we used the more sensitive MGIT automated system, thus we might be identifying participants as early converters who could have been identified as late converters using the radiometric system. Time to culture positivity used in conjunction with a biomarker may increase its' accuracy in monitoring treatment response (Hesseling et al. 2010). The ACS 11-gene and rPSVM.1 signature scores were not associated with time to culture positivity in our cohort, contrary to what was observed in the CTRC study (Thompson et al. 2017). In addition, Thompson and colleagues observed an increase treatment response monitoring when mycobacterial load was combined with signature score into a new model. In our cohort, assessing signature scores and time to culture positivity at baseline did not improve clustering accuracy at baseline, month two, or at the end of treatment in monitoring treatment response. Based on this result, we

believed that it was not warranted to develop a model that combined time to culture positivity with signature scores into a classifier. We observed that HIV infection increased ACS 11-gene signature scores (Chapter 3) leading to a decreased specificity and diagnostic accuracy for active TB disease. We propose that high signature scores in all participants in the IMPRESS cohort resulting from HIV load is the most likely explanation for the poor classification of early and late converters (more discussion in Chapter 6).

The small sample size, particularly in the late converters study arm, was a limitation of our study. In addition, we could not determine the utility of signature scores in identifying treatment failures because no participants were identified with clinically defined treatment failure.

To conclude, the ACS 11-gene and rPSVM.1 signatures could not predict early conversion at TB diagnosis but showed promise as treatment monitoring tools in HIV-infected TB retreatment patients.

5.6: Contributions

Ms F. Darboe designed, conducted the experiments, performed the analyses, and wrote this chapter under the supervision of Dr A. Penn-Nicholson and A. Prof T.J. Scriba.

Chapter 6: The effect of HIV viraemia on transcriptomic signature performance in predicting TB recurrence and monitoring treatment response.

The aim of this chapter is to determine the effect of HIV plasma viral load (pVL) on transcriptomic signature scores and performance of the signatures.

6.1: Introduction

HIV pVL is a key factor in the determination of HIV-related morbidity and mortality (Suthar et al. 2015). The pVL remains relatively stable during untreated asymptomatic chronic HIV infection (Fraser et al. 2014). Immune activation and inflammation are directly associated with pVL levels (Deeks et al., 2013; Joshi et al., 2016). HIV replication is thought to be responsible for increased immune activation and inflammation (Deeks, Tracy, and Douek 2013). Viraemic individuals express higher levels of soluble inflammation and activation markers, such as soluble CD14 and Ki-67, than aviraemic individuals (pVL < 40 copies/mL, or lower than the detectable limit) and HIV-uninfected controls (Somsouk et al. 2015).

ART suppresses the replication of HIV and reduces pVL resulting in an asymptomatic aviraemic state in most HIV-infected persons. Immune activation and persistent inflammation due to HIV infection remain elevated in HIV-infected persons who have suppressed viral replication due to ART, when compared to uninfected persons (Klatt et al. 2013). Despite ART, up to a quarter of HIV-infected persons remain immunosuppressed in the asymptomatic phase, as CD4 cells are not restored to their full capacities (Corbeau and Reynes 2011). The relative risk of TB disease in HIV-infected individuals with pVL >1000 copies /mL is 1.74 (95% CI: 1.39-

2.17) higher than the risk in persons with pVL<1000 copies/mL (Gupta et al., 2012). HIV-infected persons with suppressed pVL of <50 copies per mL had higher activation of the immune, endothelial and coagulation systems in comparison to HIV-uninfected persons (Psomas et al. 2016). Despite a partially restored immune system, an inflammatory milieu is still observed in HIV-infected persons on ART.

In addition to contributing to inflammation and gut leakage, activated pDCs also produce Type I IFNs, such as IFN-α and β, which interfere with stages of viral replication thereby mediating antiviral immunity (F. McNab et al. 2015). The pro-inflammatory antiviral activities of type I IFN may paradoxically hasten HIV disease progression (Stifter and Feng 2015). Progression of HIV disease has been associated with an increase in IFNα production from pDCs (Lehmann et al. 2008). Early studies of Type I IFN in untreated acute HIV infection observed a widespread induction of ISGs in peripheral blood cells regardless of viral load (Bosinger and Utay 2015).

Type I IFN signalling is associated with TB disease severity as measured by bacterial load in mice (Manca et al., 2001; Manca et al., 2005; Stanley et al., 2007). Expression of ISGs by blood immune cells is highly elevated at TB diagnosis relative to controls without disease (Berry et al., 2010; Bloom et al., 2013; Cliff et al., 2013; Maertzdorf et al., 2011, 2015). Elevated ISG expression was also the basis for detecting *Mtb*-infected individuals who progresed to TB disease which was used to develop the 16-gene transcriptomic signature of risk of TB (Zak et al. 2016). These studies confirm the observation in mice that the presence of type I IFN is indicative of TB disease and is a useful measure of disease severity. In addition, type I IFN

production is correlated with mycobacterial virulence and host susceptibility (Stifter and Feng 2015).

Collectively these results suggest that high expression of ISGs due to HIV infection can lead to a decrease in accuracy of transcriptomic signatures of TB due to an increase in false positives. Most published transcriptomic signatures of TB disease were developed in HIV-uninfected cohorts and some have been applied to HIV-infected cohorts to determine diagnostic signature performance. Few studies have developed RNA transcriptomic signatures to specifically distinguish TB disease from other infectious diseases in HIV-infected persons (Dawany et al., 2014; Kaforou et al., 2013; Sweeney et al., 2016). Kaforou and colleagues developed a 27-transcript signature that could distinguish TB disease from *Mtb* infection in HIV-infected (AUC 97, 95% CI 95-100) and uninfected cohorts, (AUC, 100 (95% CI, 100-100)) (Kaforou et al. 2013). Another study of HIV-infected persons developed a 251-transcript signature, which could differentiate HIV-TB co-infection from HIV mono-infection with an accuracy of 81.4% (Dawany et al. 2014). Sweeney and colleagues observed an 8% reduction in performance of a 3-gene signature in discriminating of TB from *Mtb* infection in HIV-infected persons, relative to HIV-uninfected persons (Sweeney et al., 2016). When the ACS 16-gene signature, re-parameterised to microarray data, was applied to the published datasets from Kaforou and colleagues, a similar 10% reduction in signature performance in HIV-infected relative to HIV-uninfected cohorts was also observed (Zak et al. 2016). These findings suggest that IFN response signatures yield lower accuracy in HIV-infected cohorts, likely due to the effects of viral replication. When the diagnostic performance of the ACS 11-gene signature was assessed (described in Chapter 3, Figure 7), a 10% decrease in diagnostic

performance was observed in HIV-infected persons compared with uninfected persons. The rPSVM.1 signature was developed with the explicit aim of increasing the accuracy of the ACS 11-gene signature for HIV-infected cohorts. However, we did not observe a significantly improved performance in the TRuTH and IMPRESS cohorts (Chapters 4 and 5). To address the question of HIV load in the context of these IFN response signatures, we set out to determine the effect of HIV viraemia (plasma viral loads) on the performance of the signatures in detecting recurrent TB and in monitoring treatment response.

Aim

To determine the effect of HIV pVL on the performance of the signatures in predicting recurrent TB disease and monitoring of treatment response.

Hypothesis

IFN response signature scores will be higher when pVL is detectable than when pVL is undetectable and signature scores will correlate with pVL.

6.2: Methods

6.2.1: Study design

Samples collected in the TRuTH and IMPRESS cohorts were stratified into two groups based on whether pVL was detectable or undetectable. Different pVL assays were used during the conduct of the TRuTH and IMPRESS studies and we sought to avoid performing analyses with different lower limits of viral detection. Hence, the detection limit for pVL was 400 copies/mL. The performance of the ACS 11-gene and rPSVM.1 signatures was measured by area under the ROC curve. Viral load data was available for all sampling time points in the TRuTH study. However, viral

load data in the IMPRESS cohort was only available at baseline and end of treatment.

The Mann Whitney U test was used to compare signature scores between undetectable and detectable pVL samples.

6.3: Results

6.3.1: Stratification of samples in the TRuTH and IMPRESS cohorts for pVL analysis.

Out of 299 samples in the TRuTH study, 272 samples were collected when pVL was undetectable and 27 when pvL was detectable. For some samples (n=37), pVL data was not available at the sample time point. In these circumstances we assigned the pVL values collected within two weeks of sampling. Fifteen of the 27 (56%) samples with detectable pVL were from progressors, as compared to only 88 of the 272 (32%) samples from those with undetectable pVL (Figure 23A).

In the IMPRESS cohort only 86 samples had associated pVL data. Fifty of these had undetectable pVL whilst 36 had detectable plasma viral loads (Figure 23B). Of the samples with pVL data, 60 were baseline samples, 25 were end of treatment samples and one was at the 2-month time point. Of the baseline samples 27 had undetectable pVL, and 22 of the end of treatment samples had undetectable pVL (Figure 23B). The single two-month sample with undetectable pVL data was not included in the analysis.

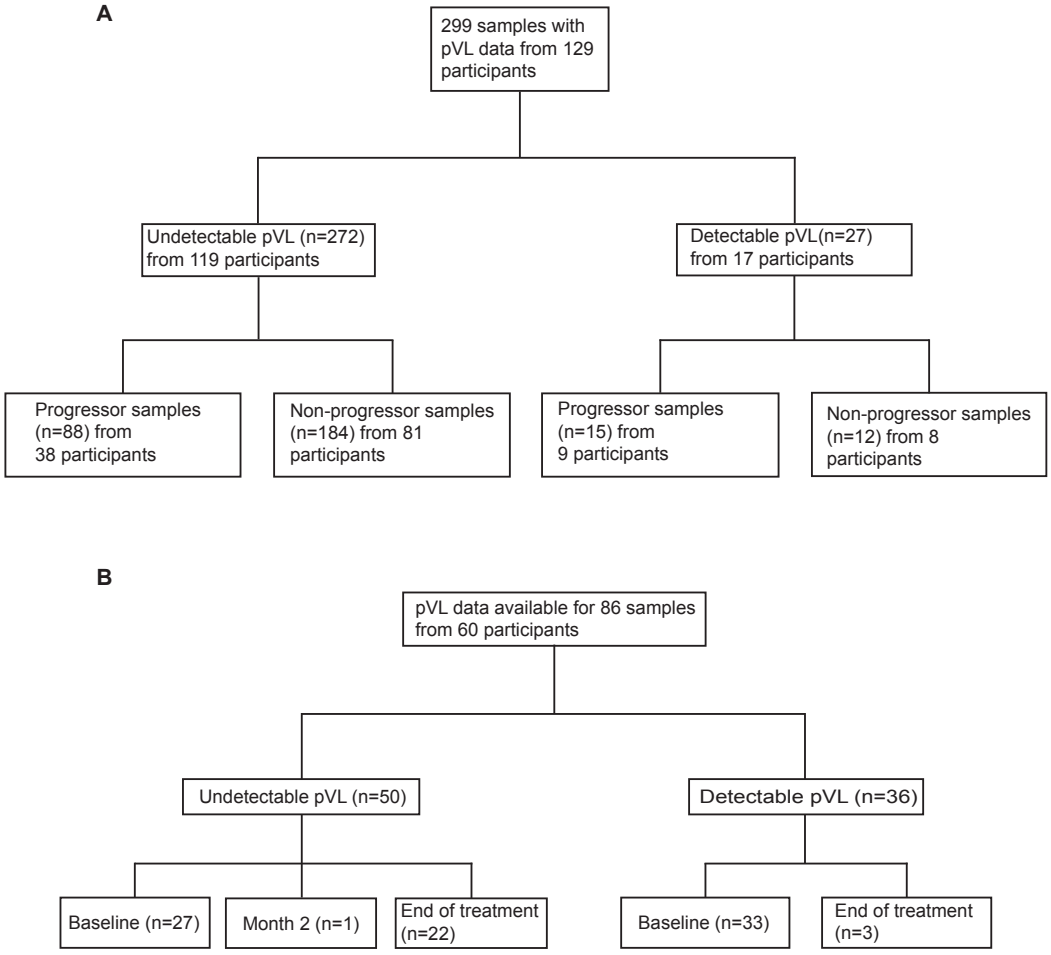


Figure 23: Sample selection for analysis of effect of plasma viral loads (pVL) on signature performance. (A) Sample stratification for analyses in the TRuTH cohort. (B) Sample allocation in the IMPRESS cohort**.**

6.3.2: HIV viraemia levels are associated with signature scores

In the TRuTH cohort, ACS 11-gene and rPSVM.1 signature scores were higher when HIV pVL was detectable (median score 51%, IQR, 25% to 78%) than when pVL was undetectable (median score 14%, IQR, 10% to 22.3%), irrespective of whether the samples were from progressors or controls (Figure 24A). A weak positive association was observed between pVL and signature scores when only data from detectable pVL samples were analysed (Figure 24C and D). pVL and CD4 T cell counts were not predictive of recurrent TB disease at any time point before diagnosis (data not shown).

Similarly, signature scores in the IMPRESS cohort were higher in samples with detectable pVL than those with undetectable pVL (Figure 24B). However, contrary to what was observed in the TRuTH cohort, signature scores were not correlated with detectable pVL for either the ACS 11-gene or rPSVM.1 signatures in the small IMPRESS cohort. Most of the samples in the TRuTH and IMPRESS studies were collected when study participants had undetectable pVL. Sample numbers were too low to perform ROC analyses in those with detectable pVL (Figure 23).

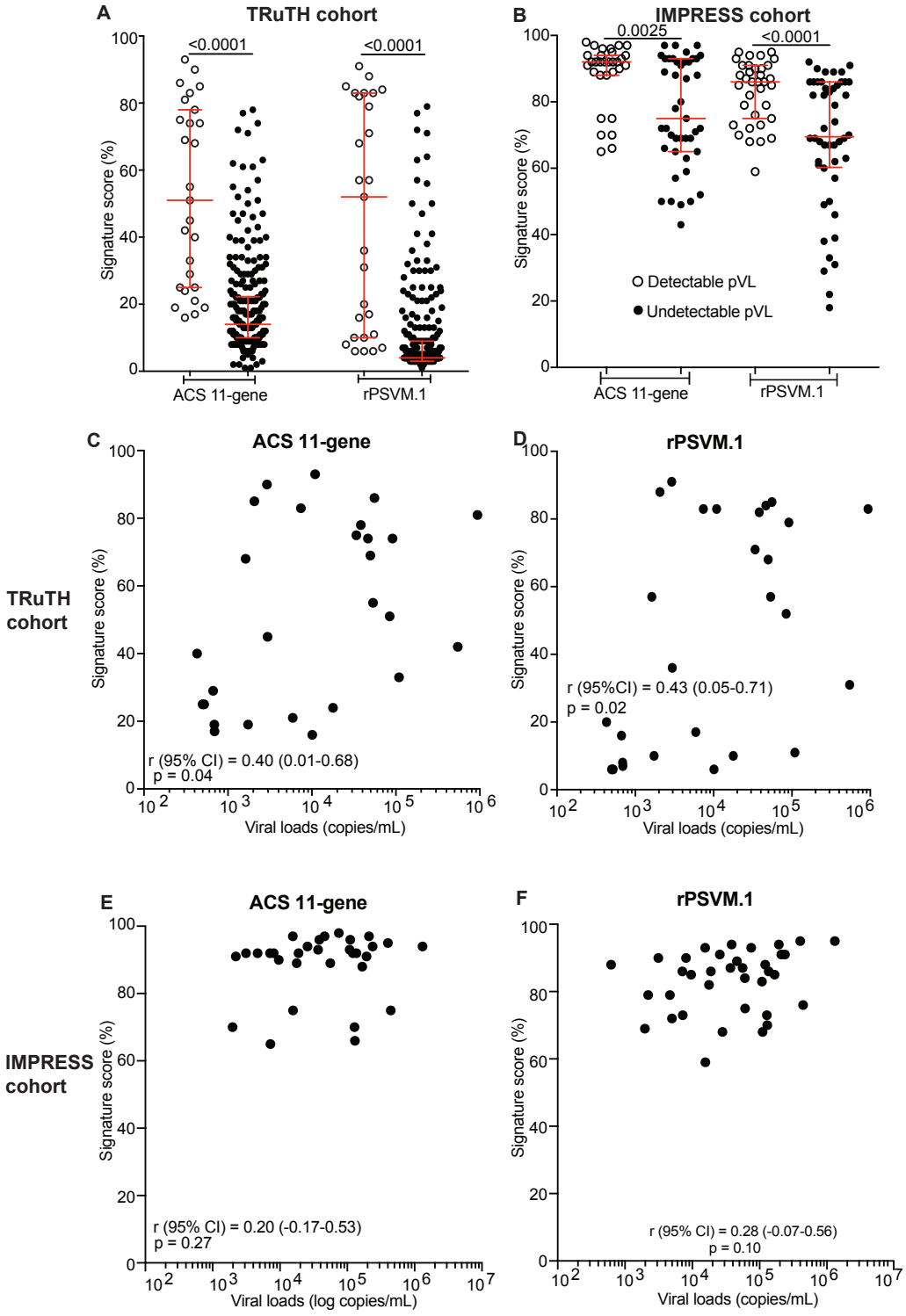


Figure 24: Association between ACS 11-gene and rPSVM.1 signature scores and pVL. (A) Effect of pVL on signature scores in the TRuTH cohort. Progressors and controls were combined. (B) Effect of pVL on signature scores in the IMPRESS cohort with all time points combined. (C-F) Association between signature scores and detectable pVL from the TRuTH (C and D) and the IMPRESS (E and F) cohorts. Spearman correlation coefficients are shown. P-values are not adjusted for multiple comparisons. TRuTH cohort samples were PBMC and IMPRESS cohort samples were whole blood.

6.3.3: Differentiating between recurrent TB progressors and controls and monitoring treatment response when HIV pVL is undetectable

Signature scores from TRuTH cohort samples with undetectable HIV pVL were stratified into six-month (180-day) time to TB windows. The period within six months of recurrent TB diagnosis was further subdivided into two three-month (90-day) windows as done in Chapter 4. In the three months preceding TB recurrence signature performance was equivalent when all samples (samples with both detectable and undetectable pVL) or only those with undetectable pVL were included in the same ROC (Figure 25).

In the IMPRESS cohort performance of the signatures in differentiating early from late converters was also not different when the entire cohort was analysed or when only undetectable pVL samples were analysed (Figure 26A & B). In addition, the performance of the signatures in differentiating samples taken at diagnosis from samples taken at treatment completion was also equivalent when the entire cohort or only undetectable pVL samples were analysed (Figure 26C & D).

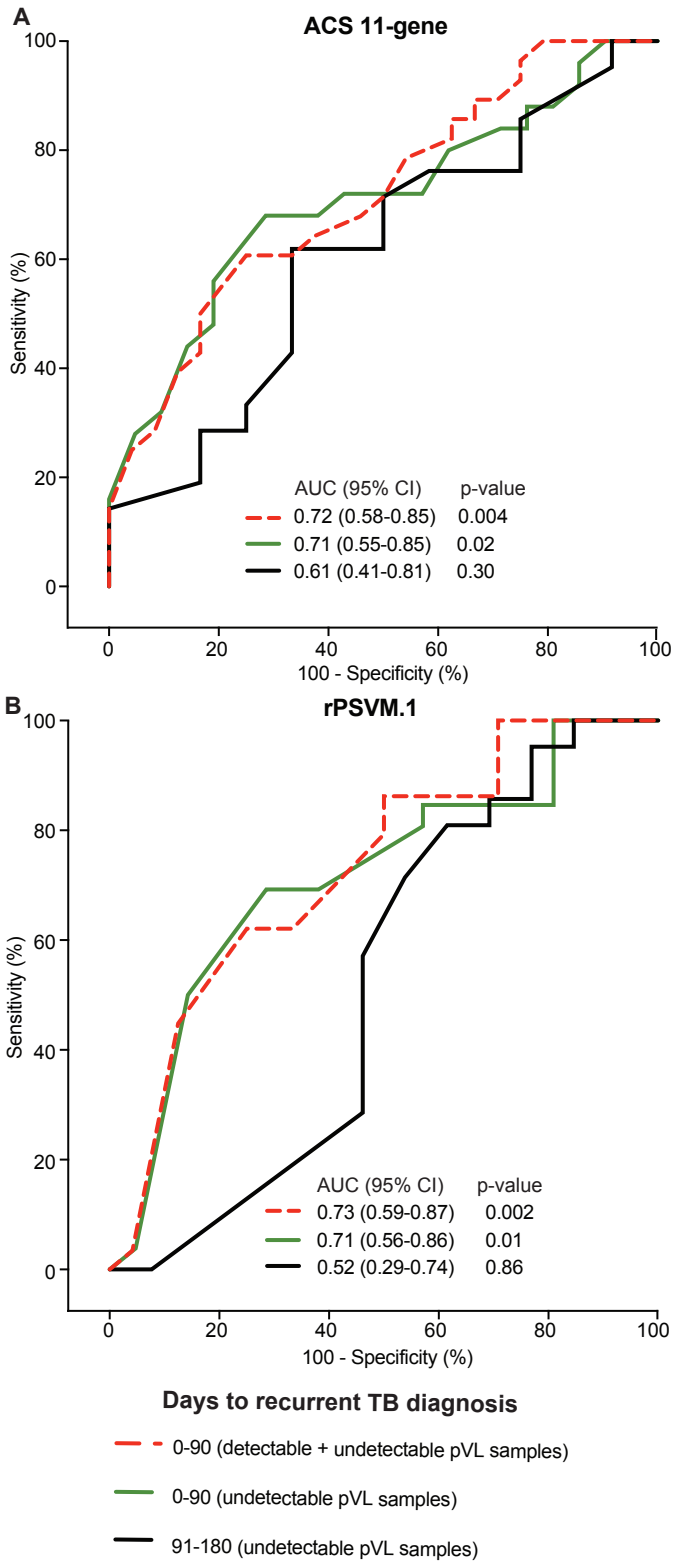


Figure 25: Classification of progressor and control samples from only those with undetectable pVL in the TRuTH cohort. Each ROC curve depicts the AUC for progressor vs. control at the indicated time point by the ACS 11-gene (A) and rPSVM.1 (B) signatures.

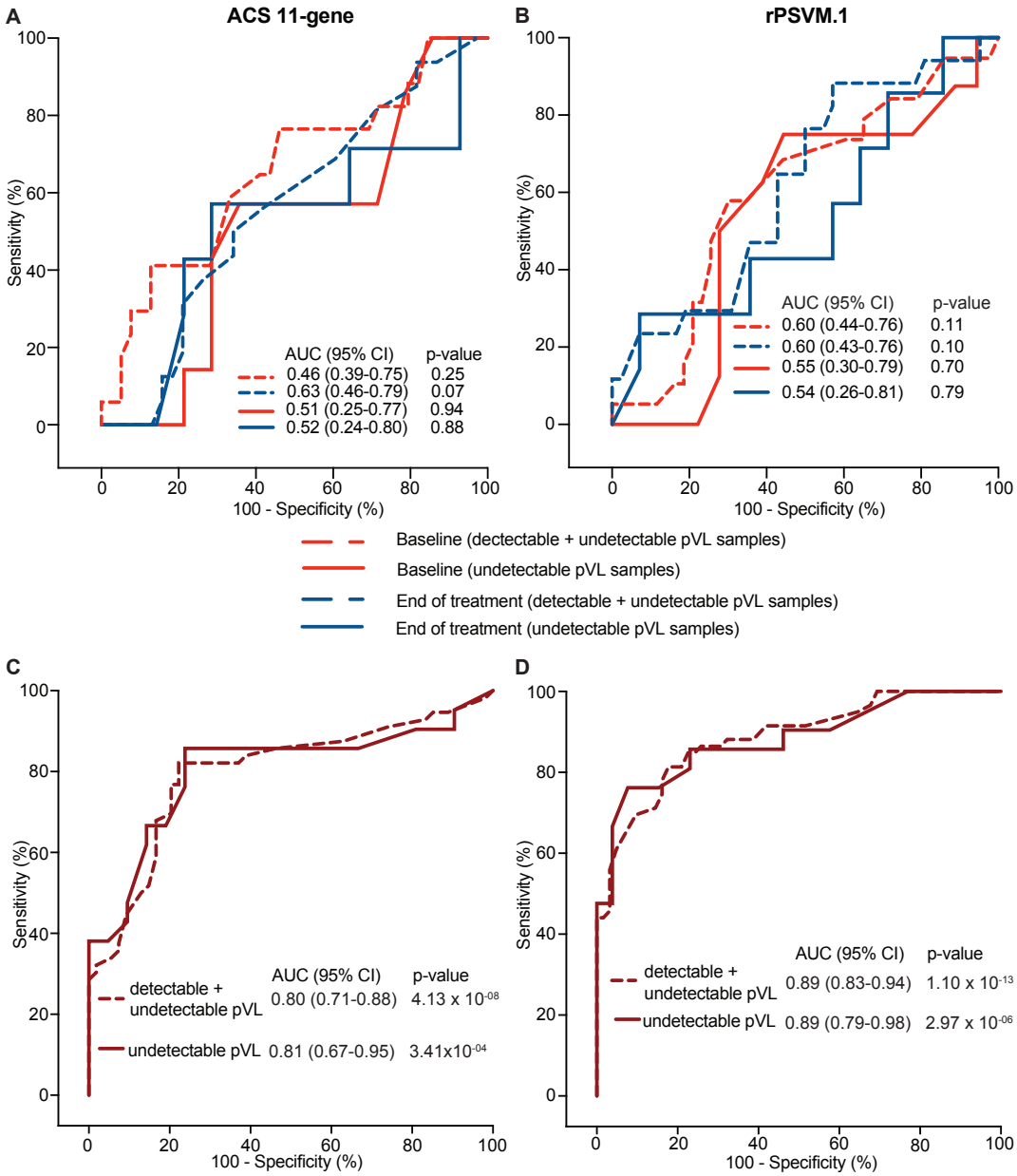


Figure 26: Effect of pVL on performance of signatures in treatment response monitoring in the IMPRESS cohort. (A &B) Each ROC curve depicts the ability of the ACS 11-gene (A) and rPSVM.1 (B) signatures in differentiating early from late converters at study time points. (C&D) Each ROC curve depicted the ability of the ACS 11-gene (C) and rPSVM.1 (B) signatures in differentiating samples collected at diagnosis from samples collected at the end of TB treatment.

6.4: Discussion

In this chapter, we investigated the effect of HIV viraemia on the performance of the ACS 11-gene and rPSVM.1 signatures in the TRuTH and IMPRESS studies. We show that signature scores were significantly higher when participants had detectable HIV virus in the plasma and a weak association was observed between pVL and signature scores for both ACS 11-gene and rPSVM.1 signatures in the TRuTH cohort. These data corroborate our findings in the previous chapters that underlying HIV infection increases IFN response signature scores and that the diagnostic performance in HIV-infection, and especially in viraemic individuals, is reduced by virus-induced inflammation.

A limitation of the work in this chapter is that the clinical studies were not sufficiently powered to investigate the performance of the signatures in samples with detectable or high pVLs. This will be possible in a study currently underway at SATVI, the CORTIS-HR study, in which we are assessing the performance of the ACS 11-gene signature in 860 HIV-infected persons.

Despite this limitation and small sample sizes, to our knowledge this is the first time that the effect of HIV pVL on transcriptomic gene signatures of TB has been described. Our work suggests that underlying HIV infection has a marked effect on performance of IFN response signatures, which requires further examination. Of note, a recent transcriptomic study by Esmail, Wilkinson and colleagues demonstrated that a transcriptomic TB signature based on complement pathway genes may have greater utility in ART naïve HIV-infected persons (Esmail et al.,

2018). In this study, pVL did not affect circulating immune complexes, which were associated with transcripts involved in the complement pathway.

We conclude that the decrease in the performance of type I IFN signatures of TB in HIV-infected persons is due to the stimulatory effect of circulating virus on type I IFNs production. We hypothesise that TB signatures developed in HIV-infected cohorts may perform better in HIV-uninfected persons.

6.6: Contributions

Ms F. Darboe designed, conducted the experiments, performed the analyses, and wrote this chapter under the supervision of Dr A. Penn-Nicholson and A. Prof T.J. Scriba.

Chapter 7: Signature reduction to pave the way for a point-of-care diagnostic test: the ACS 6-gene signature

7.1 Introduction

Most recently published gene signatures for diagnosing active TB have been developed on microarray and RNA sequencing platforms, which are expensive and measure expression of the entire transcriptome. Although these platforms are ideal for the discovery of transcriptomic signatures, the size of most signatures that have been discovered, including the 11-gene, 48-transcript ACS signature, makes them impractical for clinical use at a point-of-care level. There is an urgent need to develop these signatures into cheap and easy-to-use diagnostic tools especially for use in resource-poor settings. Cheaper and highly targeted gene expression quantification methods such as qRT-PCR assays can make the signatures more accessible for the development of point-of-care diagnostics. This underlies the motivation for translating the ACS correlate of risk signatures from RNA-seq into qRT-PCR, before they were validated on independent clinical cohorts (Zak et al. 2016).

For translational purposes it is believed that fewer transcripts in a signature will have a higher probability of success when implementing such an assay into a hand-held diagnostic device, due to lower costs, and simplicity. Several recent studies have discovered small signatures in a bid to develop cheaper and simpler diagnostic tests for TB disease (Maertzdorf et al., 2016; Roe et al., 2016; Sweeney et al., 2016). Sweeney and colleagues discovered and validated a three-gene signature that distinguished active TB disease from *Mtb* infection in published datasets of both HIV-

infected and uninfected individuals with excellent accuracy (Sweeney et al., 2016). This 3-gene signature also distinguished active TB disease from other diseases, including sarcoidosis and pneumonia. Maertzdorf and colleagues developed and validated a 4-gene signature that can classify TB disease from healthy persons (Maertzdorf et al., 2016). This signature was compared to a 15-gene signature developed in the same cohort; both signatures were applied to publicly available datasets and could distinguish active TB disease from *Mtb* infection, also with excellent accuracy. In another study, transcript levels of a single gene, BATF2, allows differentiation between active TB disease and healthy HIV-uninfected persons in individuals with diverse ethnic backgrounds (Roe et al. 2016). The performance of these signatures in diagnosing TB disease suggest that signatures based on a few transcripts may provide equivalent diagnostic performance to signatures with large number of transcripts.

We developed a qRT-PCR signature consisting of six transcripts representing six genes that could predict active TB in adolescents before onset of symptoms and diagnosis in the ACS cohort. The signature was independently validated on the GC6-74 cohort. Referred to as the “ACS 6-gene signature”, it can be measured using six Taqman primer-probe assays (Penn-Nicholson, unpublished data). The 6-gene signature consists of three transcripts, GBP2, FCGR1B, and SERPING1, that are upregulated in progressors or TB patients (Taqman primer-probe assays: GBP2.Hs00894846_g1, FCGR1B.Hs0234185_m1, and SERPING1.Hs00934329_m1, respectively) and three transcripts, TUBGCP6, TRMT2A, and SDR39U1, that are downregulated in progressors or TB patients (Taqman primer-probe assays: TUBGCP6.Hs00363509_g1,

TRMT2A.Hs01000041_g1 and SDR39U1.Hs01016970_g1, respectively), resulting in nine-transcript pairs (Figure 27). The score for the ACS 6-gene signature is derived from computing the ratio between the nine transcript pairs, where each pair contains one transcript that is upregulated in active TB with one that is downregulated in active TB, relative to healthy controls. This pair-ratio format presents two advantages over other signatures. Firstly, the up-down pairing provides a "self-standardisation" function that eliminates the need for housekeeper transcript-based standardisation to account for variability in input RNA. Secondly, the pair-ratio format allows a signature score to be calculated even if a Ct value for one primer probe is not derived, due to a failed PCR reaction for example. As observed for the pair-wise ACS 16-gene signature, this provides important robustness to the signature (Zak et al. 2016). The sum of the ratio of the transcript pairs and a coefficient "d" (weighting between the pairs, which was originally derived from model parameterisation on the ACS study) results in a binary score (Table 14). Samples are classified as "progressors or TB cases" or "non-progressors or controls" by each pair based on whether this calculation is greater (progressor/case) or less (non-progressor/control) than zero. A final score is given as the summation of the progressor voting pairs divided by the total number of voting pairs. For example, if six out of the nine pairs vote "progressor or TB case" and all nine pairs voted (a Ct value was obtained for all six transcripts), then the score will be calculated as: (6/9)*100 = 66.7%.

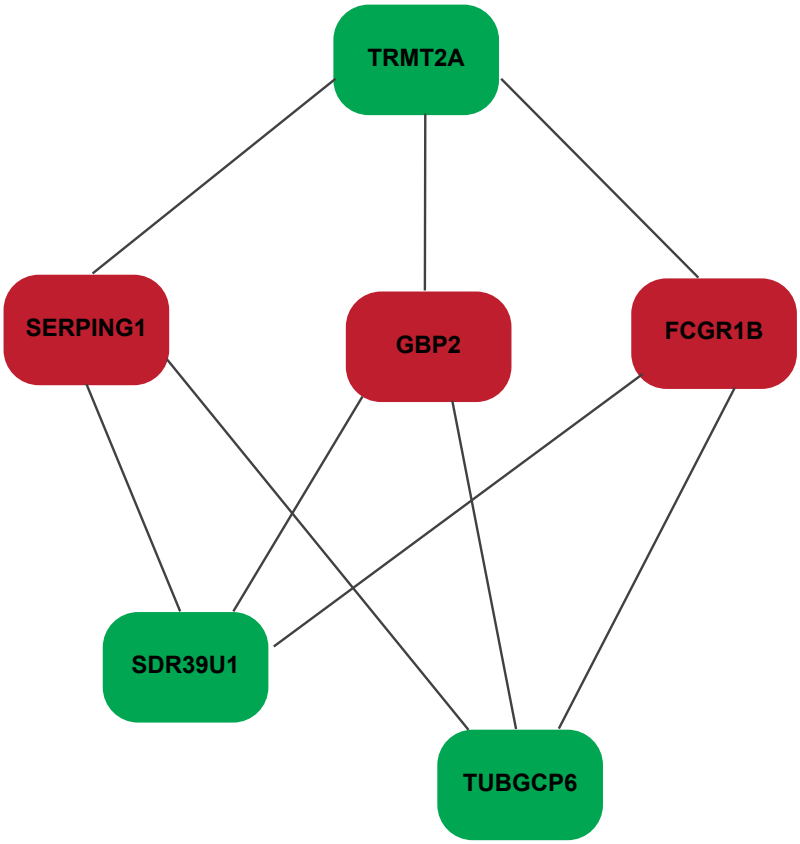


Figure 27 Pair structure of the ACS 6-gene signature. The signature comprises nine transcript pairs that each links a transcript that is upregulated in active TB with one that is downregulated in active TB relative to healthy controls. Lines indicate the pairing of the transcripts. Transcripts that are upregulated in progressors are in red nodes and those that are downregulated in progressors are in green.

Table 14: Taqman primer-probe pairs in the ACS 6-gene signature and coefficients for calculating signature scores.

Primer #1	Primer #2	Coefficient d
GBP2.Hs00894846_g1	TUBGCP6.Hs00363509_g1	-2,3
GBP2.Hs00894846_g1	TRMT2A.Hs01000041_g1	-5,7
GBP2.Hs00894846_g1	SDR39U1.Hs01016970_g1	-4,7
FCGR1B.Hs02341825_m1	TUBGCP6.Hs00363509_g1	2,4
FCGR1B.Hs02341825_m1	TRMT2A.Hs01000041_g1	-1,2
FCGR1B.Hs02341825_m1	SDR39U1.Hs01016970_g1	-0,2
SERPING1.Hs00934329_m1	TUBGCP6.Hs00363509_g1	0,7
SERPING1.Hs00934329_m1	TRMT2A.Hs01000041_g1	-2,5
SERPING1.Hs00934329_m1	SDR39U1.Hs01016970_g1	-1,5

In this chapter, we sought to assess the diagnostic, prognostic and treatment response monitoring performance of the ACS 6-gene signature in HIV-infected persons.

7.2: Aims and Hypotheses

7.2.1 Aims

7.2.1.1 To determine if the ACS 6-gene signature can differentiate active TB from *Mtb*-infected persons in PBMC from HIV-uninfected persons.

7.2.1.2 To determine the diagnostic performance of the ACS 6-gene signature in HIV-infected participants.

7.2.1.3 To determine if the ACS 6-gene signature can predict TB recurrence in HIV-infected persons on ART.

7.2.1.4 To determine the performance of the ACS 6-gene signature as a treatment monitoring tool in HIV-infected persons on TB retreatment.

7.2.2 Hypotheses

7.2.2.1 The ACS 6-gene signature will diagnose active TB disease in PBMC from HIV-uninfected persons and its diagnostic performance will be equivalent to the diagnostic performance of the ACS 11-gene signature.

7.2.2.2 The ACS 6-gene signature will diagnose active TB in whole blood and PBMC from HIV-infected persons with an equivalent performance to the ACS 11-gene signature.

7.2.2.3 The ability of the ACS 6-gene signature in predicting recurrent TB disease in HIV-infected persons on ART will be equivalent to the prognostic ability of the ACS 11-gene signature in the same cohort.

7.2.2.4 The ACS 6-gene sigature will be able to differentiate rapid treatment responders from late treatment responders in HIV-infected TB retreatment patients.

7.3 Methods

7.3.1 Study design

Samples from the cross-sectional study of TB cases and *Mtb*-infected controls (described in Chapter 3) were used to analyse the diagnostic performance of the ACS 6-gene signature in classifying active TB disease from *Mtb* infection. In the TRuTH (described in Chapter 4) and IMPRESS (described in Chapter 5) cohorts, we aimed to determine if the ACS 6-gene signature could predict recurrent TB disease and predict treatment response respectively, as was described for the 11-gene and rPSVM.1 signatures in Chapters 4 and 5. The study designs described in the respective chapters were maintained in this chapter.

7.3.2 Experimental set up and data analyses

Primer-probes for the ACS 6-gene signature were included in each 96.96-gene expression chip along with the other signatures that were run. The chip quality control results shown in Figures 13 and 18 in Chapters 4 and 5 therefore apply to the assays performed here. However, one of the primer-probe assays, TRMT2A.Hs01000041_g1, was not included in the gene expression chips that were used for comparing PBMC and whole blood gene expression from TB cases and controls (Chapter 3). This TRMT2A primer-probe assay was not included in the pilot study because the ACS 6-gene signature was developed after completion of the experiments for this cohort. However, since five of the six primer-probes for this signature were co-incidentally already assayed in the gene expression chips, we proceeded to calculate scores for the ACS 6-gene signature from this cohort. This therefore provided a test of the robustness of the pairwise formulation, because it allowed the assessment of the effect of the missing transcript on signature performance. Signature scores were calculated using in-house scripts in R developed with our collaborators at CIDR (Seattle, USA). ROC AUCs were calculated in R using the pROC and verification packages. Statistical analyses were done using Mann Whitney U for differences between two groups, Wilcoxon ranked sum and Kruskal-Wallis tests for differences between three groups in GraphPad Prism.

7.3 Results

7.3.1 Diagnostic performance of the ACS 6-gene signature in HIV-infected and uninfected persons

To measure the diagnostic performance of the ACS 6-gene signature in PBMC, it was applied to whole blood and PBMC samples collected in parallel from participants enrolled in the pilot cohort (Chapter 3). Results of this experiment would also allow us to make go/no-go decisions about the suitability of the ACS 6-gene signature for assessing the prognostic potential of the ACS 6-gene signature in the TRuTH cohort.

The diagnostic performance of the ACS 6-gene signature in HIV-uninfected persons was excellent, whether RNA from whole blood or PBMC was tested. The AUCs for the ACS 6-gene signature of 0.96 (95% CI 0.89 to1.00) and 0.99 (95% CI 0.99-1.00) in whole blood and PBMC respectively (Figure 28A) were equivalent to the ACS 11-gene signature with AUCs of 0.97 (95% CI 0.91-1.00) and 0.98 (95% CI 0.95-1.00) in whole blood and PBMC respectively (Figure 6C). As also observed for the ACS 11-gene signature, differentiation between active TB disease and *Mtb* infection using whole blood or PBMC samples was equivalent (Figure 28A), p=0.25. Again, ACS 6-gene signature scores were lower in PBMC samples than RNA from whole blood samples (Figure 28B).

We proceeded to determine the difference in single transcript expression between TB cases and controls in whole blood and PBMC, to assess the contribution of each transcript to the signature scores. Expression of two gene transcripts (FCGR1B and SERPING1) was significantly higher in TB cases than in *Mtb*-infected controls in whole blood (Figure 28C). Expression was not different between the two groups for

GBP2, SDRR39U1, and TUBGCP6 in whole blood, as indicated by 95% CIs that crossed zero. In PBMC, expression of SDR39U1 and TUBGCP6 transcripts was significantly lower in TB cases than *Mtb*-infected controls whilst there was no difference in expression between the two groups for the remaining three genes measured (Figure 28C). These data suggest an interesting picture where different transcripts within the ACS 6-gene signature appear to drive the signal that underlies diagnostic potential, when PBMC and whole blood are tested.

We next determined the ability of the signature to diagnose active TB in HIV-infected persons. The ACS 6-gene signature (whole blood AUC 0.92, 95% CI 0.85-1.00 and PBMC AUC 0.89, 95% CI 0.77-1.00, Figure 29A) yielded similar diagnostic performance to the ACS 11-gene signature (whole blood AUC 0.83, 95% CI 0.71-0.96 and PBMC AUC 0.88, 95% CI 0.76-0.99, Figure 7A) in HIV-infected persons. Although there was an indication of superior performance of the smaller signature in whole blood, this was not significantly different, p=0.24. As observed with the HIV-uninfected cohort, AUCs of the ACS 6-gene signature were not different between whole blood and PBMC in HIV-infected individuals, p = 0.60 (Figure 29A). Similarly, signature scores were lower in PBMC than in whole blood samples as observed in the HIV-infected cohort (Figure 29B).

As observed in the HIV-uninfected cohort, significant differences in transcript expression between TB cases and *Mtb*-infected controls were observed for FCGR1B and SERPING1 transcripts in whole blood (Figure 29C). In PBMC samples, significant differences in gene expression between active TB cases and *Mtb*-infected controls were observed only in the TUBGCP6 gene. No differences in GBP2

expression were observed between active TB cases and *Mtb*-infected controls regardless of sample type (Figure 29C).

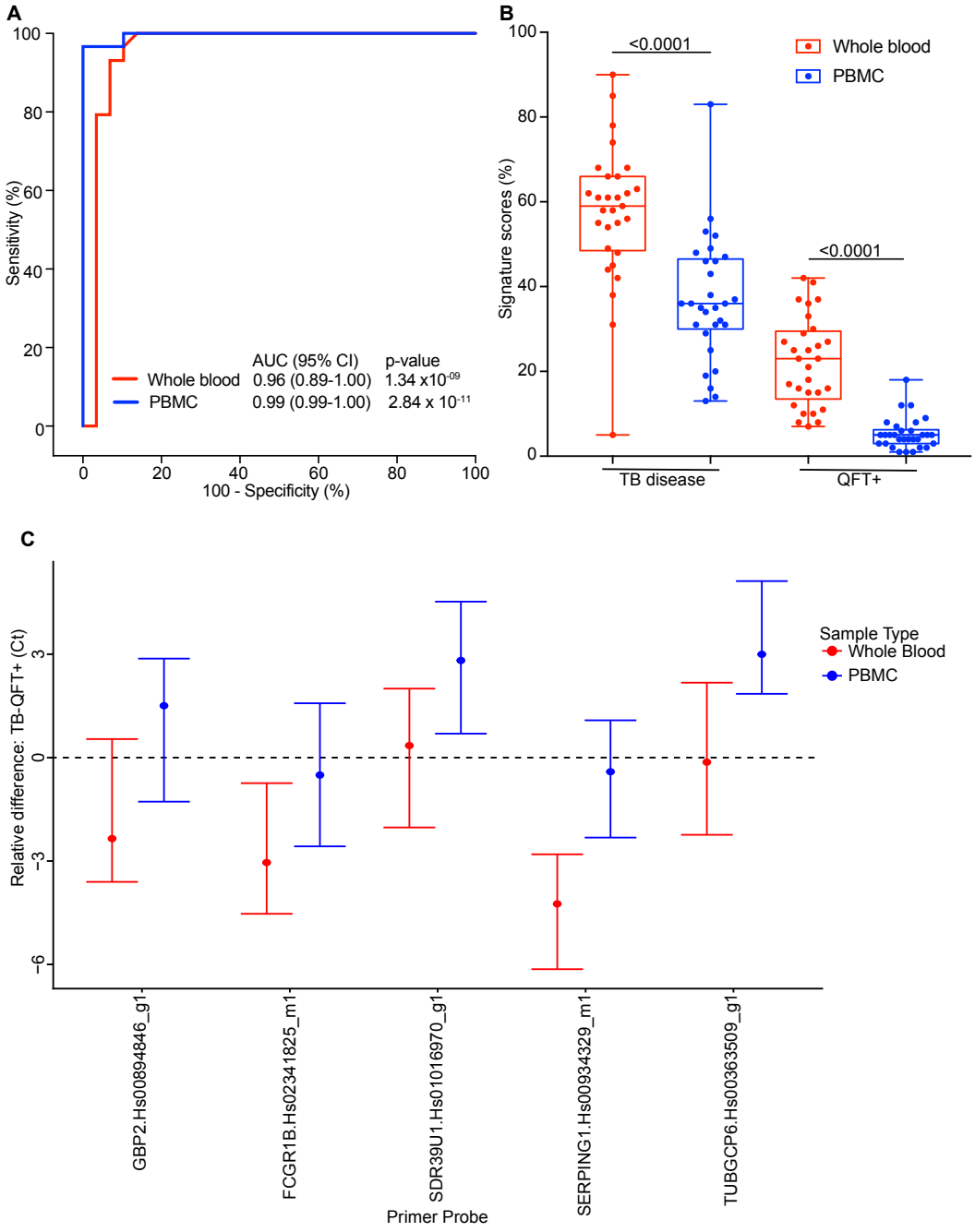


Figure 28: Diagnostic performance of the ACS 6-gene signature in whole blood and PBMC from HIV-uninfected persons from the pilot cohort. A) ROC curves for classification of active TB disease (n=30) from *Mtb* infection (n=30) in whole blood and PBMC. B) Signature scores measured in whole blood and PBMC from HIV-uninfected persons. Horizontal lines represent medians, the box represents the IQR and the whiskers the range. Dots represent individuals. C) Differences in transcript expression between TB cases and healthy *Mtb*-infected persons when measured in whole blood or PBMC samples. Negative differences indicate higher expression in active TB cases than *Mtb-infected* controls. Dots represent the medians and the bars the 95% CI for each transcript identified by its Taqman primer-probe set, computed with the rank inversion method and bootstrapping 2,000 times. Dashed lined represent the cut-off of which no difference was observed between cases and controls.

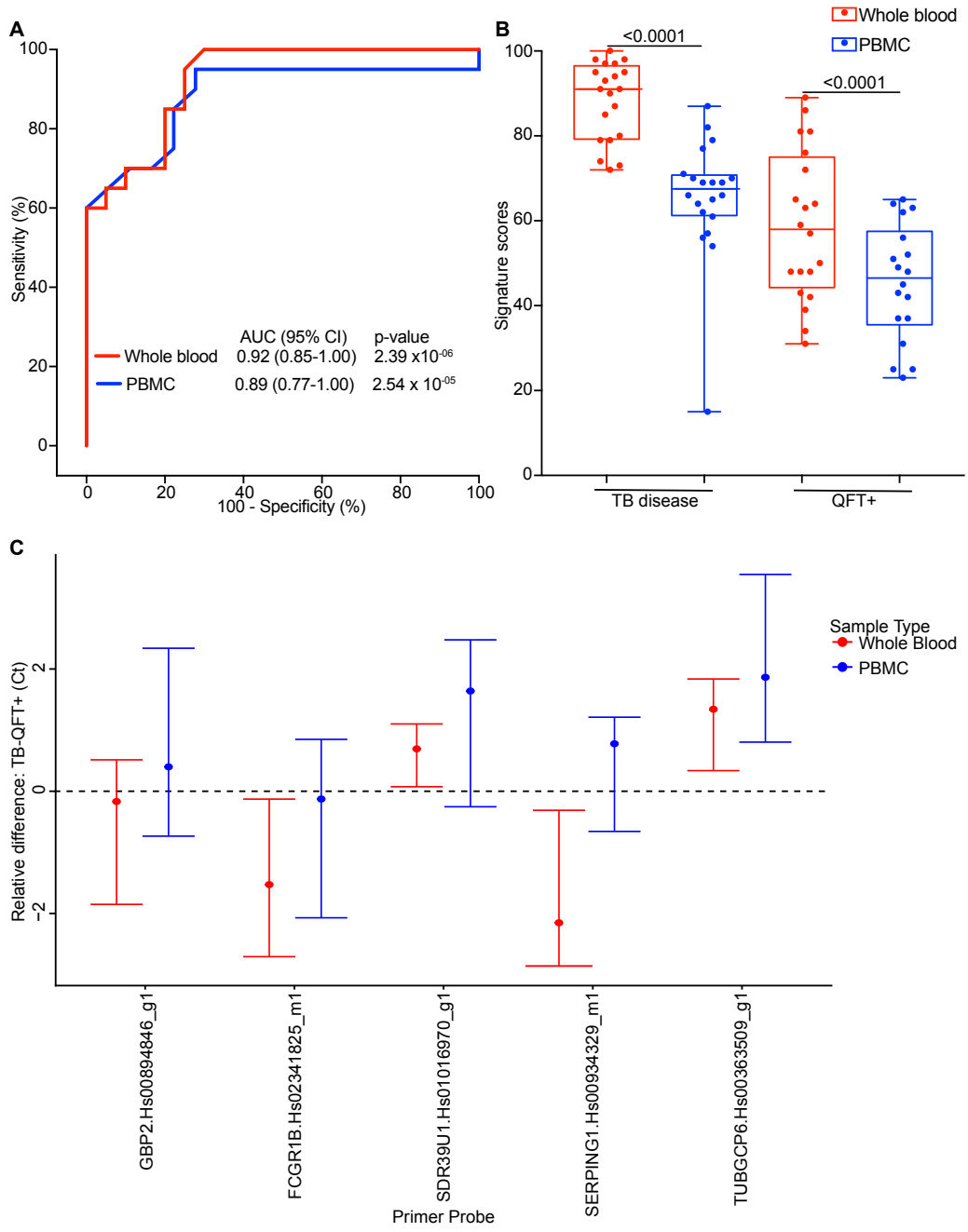


Figure 29: Diagnostic performance of the ACS 6-gene signature in whole blood and PBMC from the HIV-infected group from the pilot cohort. A) ROC curves for classification of active TB cases (n=20) from *Mtb* infection (n=20) in whole blood and PBMC. B) Signature scores measured in whole blood from PBMC from HIV-infected active TB cases or *Mtb*-infected persons. Horizontal lines represent medians, the box represents the IQR, and the whiskers the range. Dots represent individuals C) Differences in transcript expression between TB cases and *Mtb*-infected persons when measured in whole blood or PBMC samples. Negative differences indicate higher expression in active TB cases than *Mtb*-infectd controls. Dots represent medians and the bars 95% CI for each transcript identifies by its TaqMan primer-probe set, computed with the rank inversion method and bootsrapping 2,000 times. Dashed lines represent the cut-off at which no difference was observed between cases and controls.

The ACS 6-gene signature performance in HIV-infected persons was 4% and 10% lower than in HIV-uninfected persons in whole blood and PBMC, respectively (Figure 30A & B). However, the lower ROC AUCs observed in HIV-infected persons were not statistically different from the ROC AUCs observed in HIV-uninfected persons, either for whole blood ($p=0.55$) or PBMC ($p=0.07$) samples. Signature scores in TB cases and *Mtb*-infected controls were higher in HIV-infected individuals than in their HIV-uninfected counterparts ($p<0.0001$) for each of the comparisons (Figure 30C & D), as observed for the ACS 11-gene and rPSVM.1 signatures. To unpack this further, we compared the differences in individual transcript expression between HIV-infected and uninfected persons. Expression of GBP2, FCGR1B, SDR39U1 and SERPING1 in whole blood was significantly higher in HIV-infected than uninfected individuals from the *Mtb*-infected control group (Figure 31A). In TB cases, only GBP2 was expressed at significantly higher levels in HIV-infected than uninfected persons (Figure 31A). In PBMC, transcript expression of FCGR1B and SERPING1 were significantly higher in HIV-infected than HIV-uninfected *Mtb*-infected and TB cases (Figure 31B). The relative difference in gene expression by HIV-status was not different between *Mtb*-infected controls and active TB cases as indicated by overlapping 95% CI, except for SERPING1 in whole blood (Figure 31A & B).

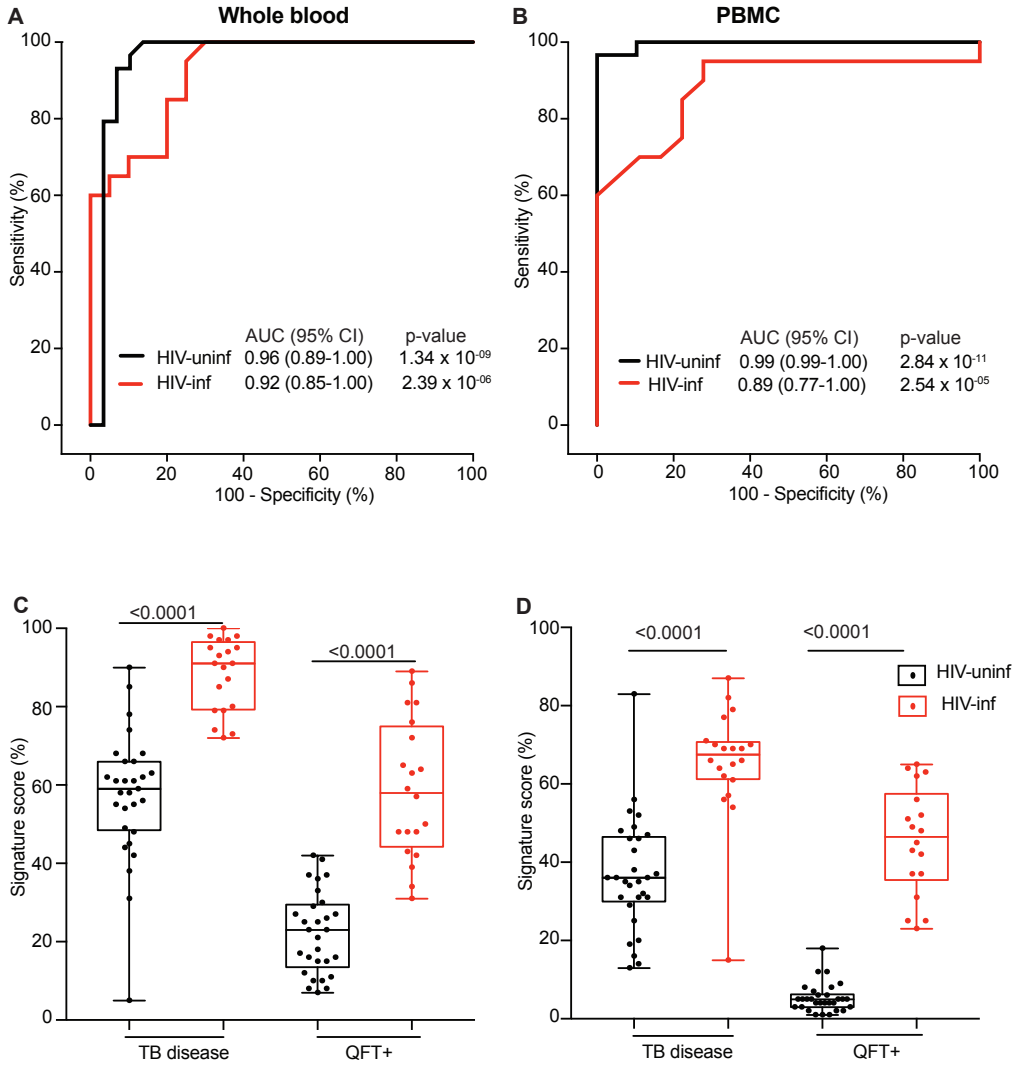


Figure 30: Comparative diagnostic performance of the ACS 6-gene signature between HIV-infected and HIV-uninfected individuals. ROC curves for classification between active TB cases and *Mtb*-infected controls in whole blood (A) and PBMC (B). Signature scores measured in whole blood (C) and PBMC (D) from HIV-infected and uninfected active TB cases or healthy *Mtb*-infected persons. Horizontal lines represent medians; the box and whiskers represent the IQR and range, respectively. Dots represent individuals.

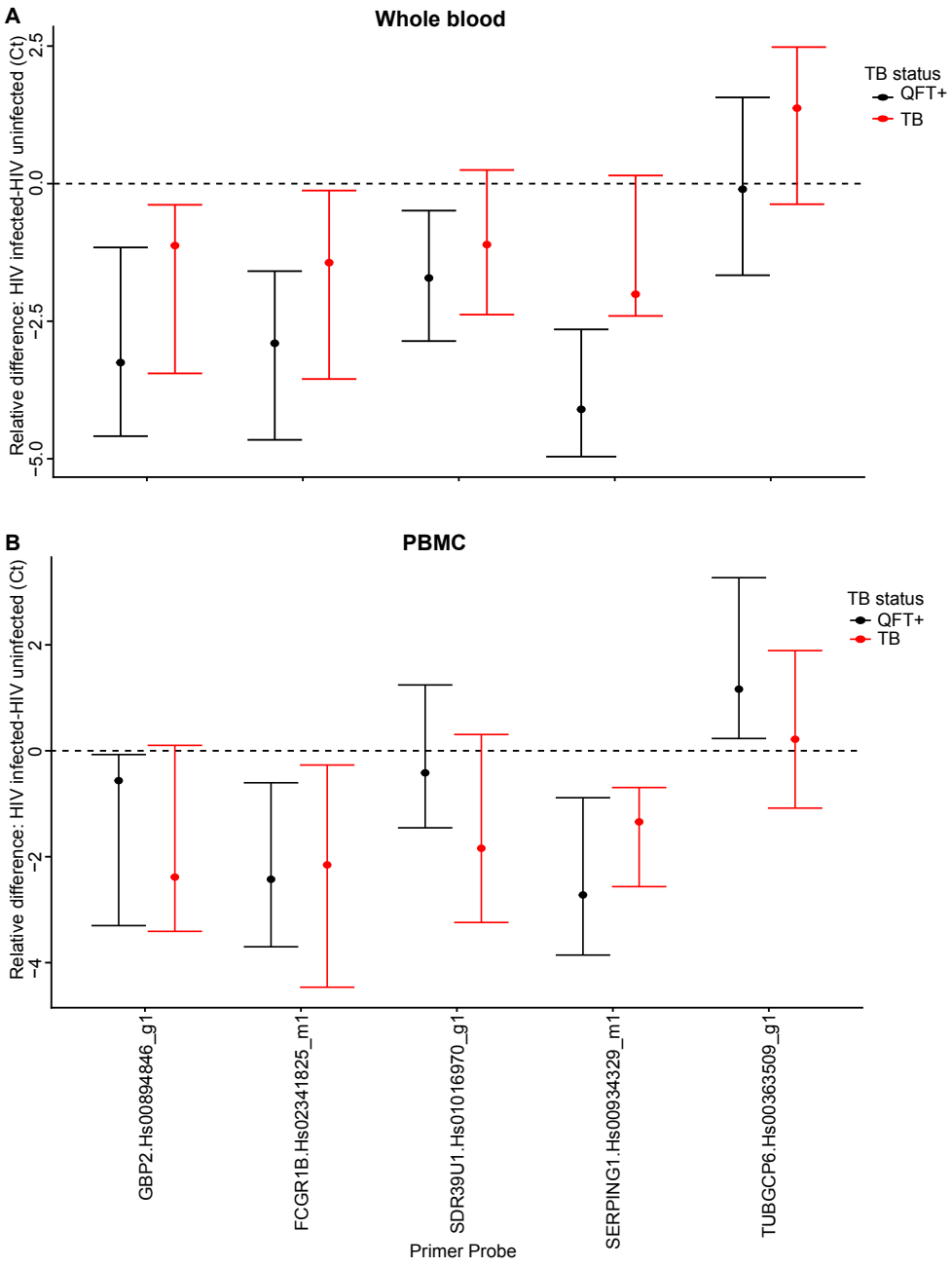


Figure 31: Differences in transcript expression between HIV-infected and uninfected persons when measured in whole blood (A) or PBMC (B) samples from TB cases or *Mtb*-infected controls. Negative differences indicate higher expression in TB than QFT+ controls. Dots represent medians and the bars 95% CI for each transcript identified by its Taqman primer-probe set, computed with the rank inversion method and bootstrapping 2,000 times. Dashed lines represent the cut-off at which no difference was onserved between cases and controls.

7.3.2 Predicting recurrent TB disease in HIV infected persons on highly active antiretroviral therapy from the TRuTH cohort

To determine the prognostic performance of the ACS 6-gene signature for recurrent TB disease in HIV-infected persons, samples from participants in the TRuTH cohort were stratified as previously described (Chapter 4). Briefly, participants who developed recurrent TB during the three-year follow-up (n=43) were assigned to two controls each by gender and time on ART. Samples from study participants were split into six-month time windows relative to diagnosis of recurrent TB disease. The six-month window preceding TB diagnosis was further divided into two three-month time periods. The ACS 6-gene signature could not differentiate those who developed recurrent TB from those who remained disease free in any of the six-month time windows (Figure 32A). However, within the three-month window preceding TB diagnosis the signature differentiated those who developed recurrent TB from those who did not with an AUC of 0.68 (95% CI 0.53 to 0.83, Figure 32B). Signature scores were highly heterogeneous in both progressors and non-progressors over the follow-up period, as was observed for the ACS 11-gene signature (Figure 32C). ACS 6-gene signature scores of samples collected from progressors in the three months preceding recurrent TB diagnosis correlated weakly (Spearman r = -0.46, p = 0.06) with MGIT time to sputum culture positivity at diagnosis suggesting a possible association between blood IFN response and sputum bacterial load (Figure 32D).

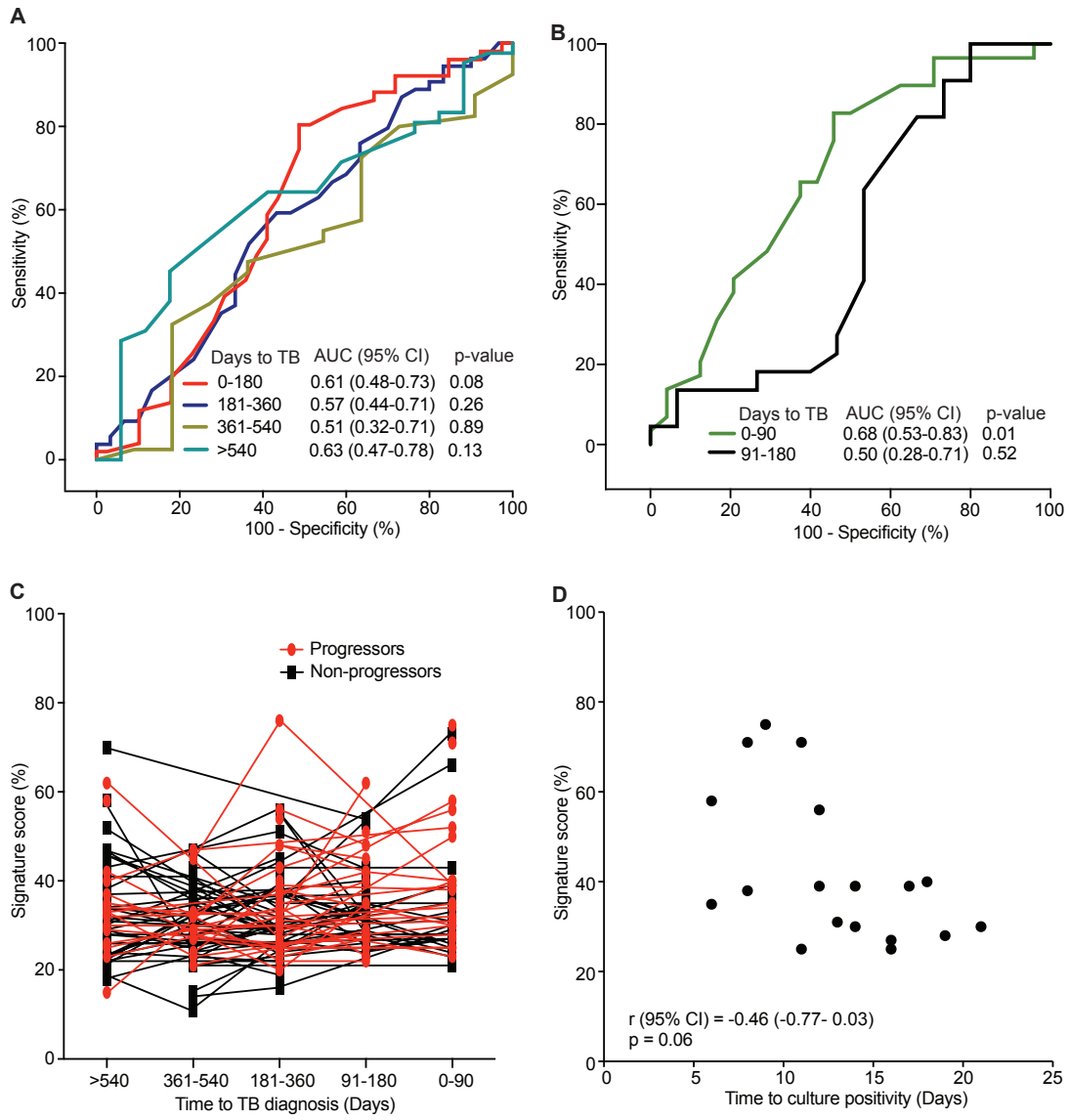


Figure 32: The ACS 6-gene signature can differentiate between progressors and non-progressors within three months of diagnosis in the TRuTH cohort. A and B) ROC curve illustrating the prognostic performance of the signature in diagnosing recurrent TB disease in HIV-infected persons. Each ROC curve depicts time to recurrent TB diagnosis: times to TB diagnosis are stratified into six-monthly time windows (A) and three-monthly time windows (B). p-values have not been corrected for multiple comparisons. C) Longitudinal signature scores of samples stratified by time to recurrent TB diagnosis and recurrent TB status at the end of follow-up. D) Association between signature scores within three months of diagnosis and time to culture positivity at recurrent TB diagnosis as an indicator of bacterial load in the TB cases.

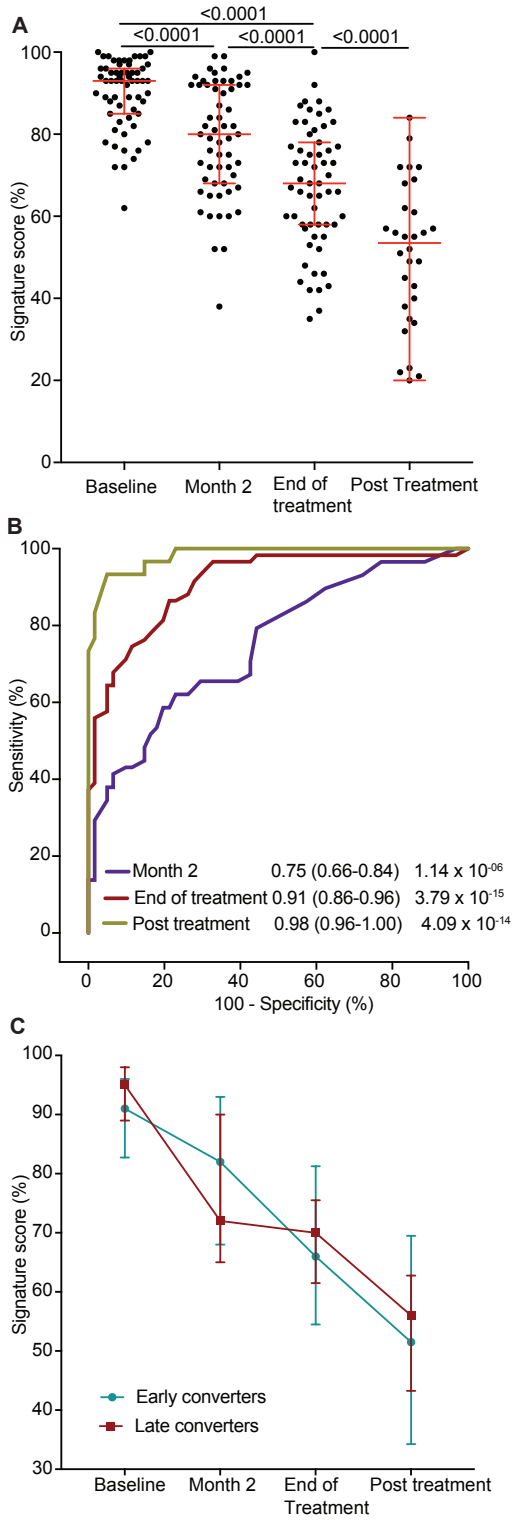


Figure 33: Treatment monitoring using the ACS 6-gene signature in the IMPRESS cohort. A) Signature scores during TB treatment. Horizontal lines depict medians and error bars depict the IQR. Dots represent individual scores. B) ROC curves depicting signature performance in differentiating samples collected at diagnosis and samples collected during TB treatment (two months, at the end of treatment, and post treatment). C) Kinetics of signature scores in early and late converters, median and IQR shown.

7.3.3 Treatment response monitoring using the ACS 6-gene signature

We applied the ACS 6-gene signature to the IMPRESS cohort (Chapter 5) to determine its ability to monitor treatment response. Participants were stratified according to their time on TB treatment and further stratified into early and late converters based on time to culture conversion by month two. During the first two months of TB treatment, a significant decrease of approximately 10% was observed in median signature scores (Figure 33A). A further decrease was observed during the continuation phase of TB treatment from month three to the end of treatment (Figure 33A). As opposed to the results obtained with the ACS 11-gene and rPSVM.1 signatures (Figures 19A & B), the ACS 6-gene signature scores continued to decrease up to eight months after the end of TB treatment. The ACS 6-gene signature differentiated between pre-treatment baseline (TB disease diagnosis), two-month treatment (end of intensive phase), end of treatment, and eight-month post-treatment samples with good accuracy (Figure 33B), confirming that the signature can monitor treatment response. The decrease in signature scores was observed in both early and late converters, but scores were not different between the two arms at any of the time points (Figure 33C).

ROC analysis of differentiation between early and late converters using the ACS 6-gene signature showed that treatment response could not be predicted at baseline or the other time points (Figure 34A). Signature scores were also not associated with time to culture conversion at the measured study time points (Figure 34B) and time to MGIT culture positivity at baseline was not associated with signature scores measured at treatment baseline (Figure 34C). A combination of the two variables to

determine clustering of the late and early converters did not correctly stratify early and late converters (Figure 34C).

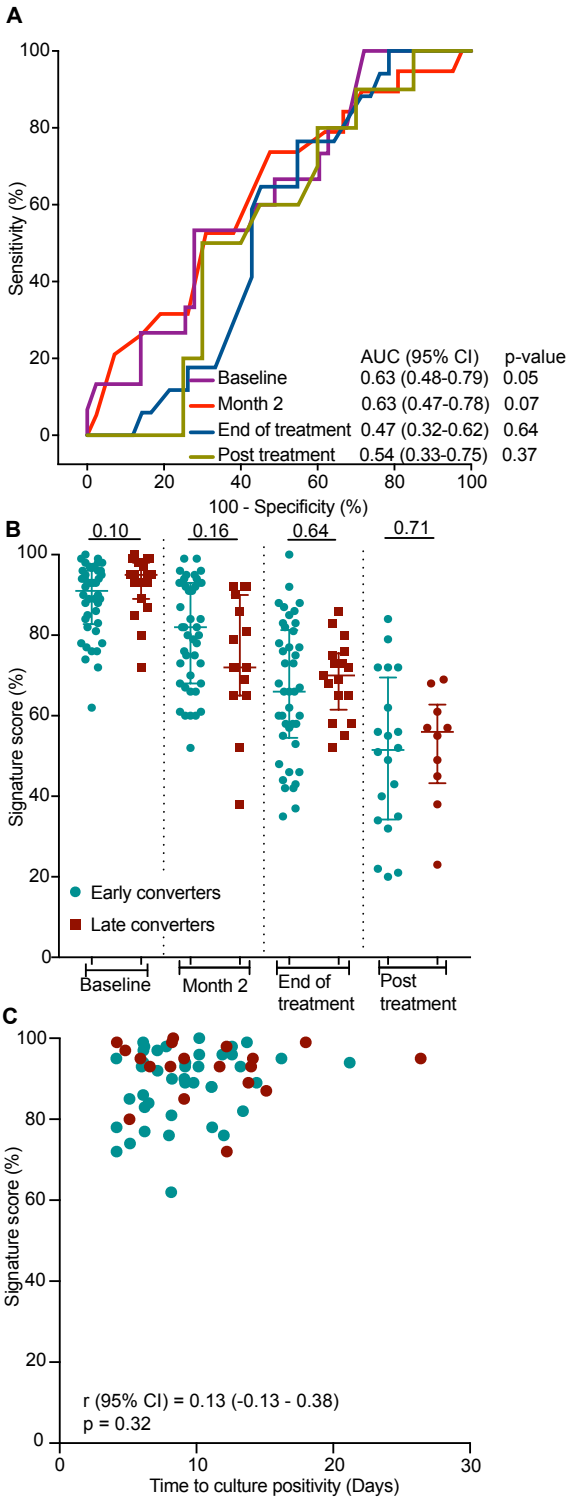


Figure 34: Application of the ACS 6-gene signature to predict treatment response in the IMPRESS cohort. A) ROC curves show differentiation between early and late converters at the indicated time points during TB treatment. B) Signature scores in early and late converters. Dots and squares represent individuals. Horizontal lines and error bars depict median and IQR, respectively. C) Association between signature scores at treatment baseline and time to culture positivity at diagnosis. Red dots represent late converters and teal dots represent early converters.

7.3.4 Effect of plasma viral load levels on ACS 6-gene signature scores in the TRuTH and IMPRESS cohorts

Samples from the TRuTH and IMPRESS cohorts were stratified into detectable and undetectable pVL samples, as described in Chapter 6. ACS 6-gene signature scores from samples collected when participants had undetectable pVL were significantly lower than when pVL was detectable in both the IMPRESS and TRuTH cohorts (Figure 35A). In the TRuTH cohort, equivalent ROC AUCs from the three-month time window prior to recurrent TB diagnosis were observed in samples with undetectable pVL and all samples (i.e. irrespective of pVL (Figure 35B). In the IMPRESS cohort similar ROC AUCs were obtained from samples with undetectable pVL to those from all samples, in differentiating early converters from late converters at baseline and end of treatment (Figure 35C). In addition, an equivalent performance was observed between samples with undetectable pVL and all samples in differentiating baseline samples from end of TB treatment samples (Figure 35D).

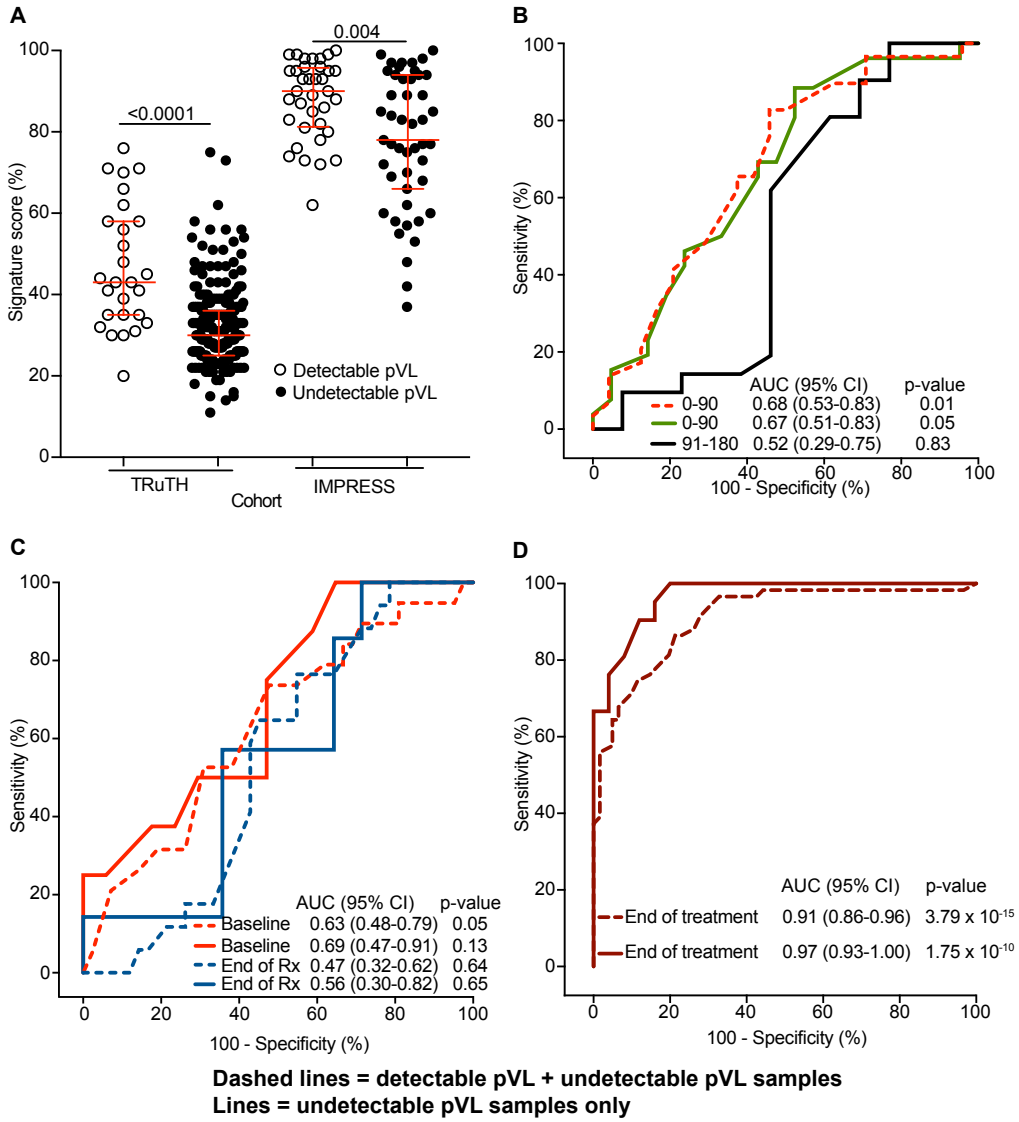


Figure 35: Effect of HIV viraemia on ACS 6-gene signature scores and performance. A) Signature scores stratified by samples with detectable or undetectable pVL in the TRuTH (PBMC samples) and IMPRESS (whole blood samples) cohorts. Dots represent individual samples collected when viral loads were detectable (open dots) or undetectable (closed dots). Median and IQR are shown. (B-D) ROC curves depicting ability of the ACS 6-gene signature to detect recurrent TB disease (B) or monitor treatment response (C and D). Solid ROC curves represent samples collected when viral loads were undetectable and dashed curves depict signature performance in the whole cohort (including both detectable and undetectable pVL samples). C) ROC curves depict ability of the signature in differentiating early converters from late converters. D) ROC curves ability of the signature to differentiate samples collected at the end of treatment from samples collected at diagnosis.

7.4 Discussion

In this chapter we sought to develop a reduced transcriptomic signature of TB that is more parsimonious than the ACS 11-gene signature for possible future application as a point-of-care test. We reduced signature to six transcripts by training a new pair-wise signature from qRT-PCR data from ACS progressor and non-progressor samples. We demonstrated that this ACS 6-gene signature could diagnose active TB disease in HIV-infected and uninfected persons, predict recurrent TB disease, and monitor treatment response in HIV-infected persons.

Diagnostic performance of the 6-gene signature was equivalent between whole blood and PBMC in both HIV-infected and uninfected persons. Importantly, the 6-gene signature performed at least as well as the ACS 11-gene signature, which comprised 48 primer-probe assays, even though measurements for one primer-probe were missing. These data demonstrate that the pair-wise formulation of the signature makes it robust to missing primer-probes, a very important feature for such a biomarker.

Three of the six genes are ISGs, which are predominantly expressed by neutrophils during TB disease (Berry et al. 2010; Singhania et al. 2017; Scriba et al. 2017). It is thus not surprising that signature scores were lower in PBMC, which is deficient in neutrophils. Our results that HIV infection increases signature scores and thereby decreases the diagnostic accuracy of the signature in HIV-infected persons is consistent with the reduced performance of other published signatures (Sweeney et al. 2016; Zak et al. 2016).

The ACS 6-gene signature shows promise as a point-of-care triage test that could be deployed in resource-poor settings in both HIV-infected and uninfected persons to identify individuals who require further investigation for TB disease. Such a triage test ideally needs to be as sensitive as the confirmatory (likely microbiological) test in diagnosing TB, or else participants would be missed by the test (Denkinger et al. 2015). The diagnostic ROC AUC of the ACS 6-gene signature appeared equivalent to previous studies of small gene signatures (Laux Da Costa et al., 2015; Maertzdorf et al., 2016; Roe et al., 2016; Sweeney et al., 2016) (Table 15).

Table 15: Performance of small gene signatures in diagnosing active TB disease using whole blood samples.

Number of transcripts in signature	HIV status	Model fit or validation	Sample size	Method of signature measurement	ROC AUC (95% CI)	Reference
Three	Negative	Model fit	123	RT-PCR	0.96 (0.89-1.00)	(Laux Da Costa et al. 2015)
Four	Negative	Model fit and validation	189	RT-PCR derived from two microarray datasets	0.98 (0.97-1.00)	(Maertzdorf et al., 2015)
	Negative	Validation	75		0.89 (0.81-0.96)	
	Negative	Validation	62		0.82 (0.71-0.93)	
Fifteen	Negative	Model fit and validation	189		0.98 (0.96-1.00)	
Three	Negative	Model fit	532	Different methods (publicly available datasets)	0.93 (0.91-0.95)	(Sweeney et al. 2016)
	Negative	Validation	296		0.93 (0.84-0.97)	
	Positive	Validation	N/A		0.89 (0.87-0.91)	
One	Negative	Model fit	46	Microarray	0.95	(Roe et al. 2016)
38 plus 10 (ACS 11-gene)	Negative	Validation	60	RT-PCR	0.97 (0.91-1.00)	Chapter 3
	Positive	Validation	40	RT-PCR	0.83 (0.71-0.96)	
30 plus 5 rPSVM.1 (10 genes)	Negative	Validation	60	RT-PCR	0.97 (0.90-1.00)	Chapter 3
	Positive	Validation	40	RT-PCR	0.84 (0.72-0.96)	
Six (ACS 6-gene)	Negative	Validation	60	RT-PCR	0.96 (0.89-1.00)	Chapter 7
	Positive	Validation	40	RT-PCR	0.92 (0.85-1.00)	

As discussed previously (Chapter 4), most participants in the TRuTH cohort had subclinical TB with more than half of the recurrent TB patients (25 out of 43) being asymptomatic identified through intensified case finding using induced sputa. The finding that the ACS 6-gene signature could predict recurrent TB disease within three months of microbiological confirmation suggests that this signature can detect the inflammatory signal of subclinical disease, as has been reported in HIV-negative persons in the ACS study (Zak et al. 2016; Scriba et al. 2017). This is further

supported by the observation that most asymptomatic participants who elected not to receive TB treatment manifested with TB symptoms some months later. These results suggest that the ACS 6-gene signature performs in a similar manner to the larger ACS 11-gene signature. We propose that the ACS 6-gene signature should perform with equivalent accuracy to the ACS 11-gene signature in other studies with larger sample sizes, such as the CORTIS and CORTIS HR trials.

We also show that the ACS 6-gene signature score tracked the progress of TB treatment in the IMPRESS cohort and could differentiate with good accuracy between pre-treatment samples and samples collected at the end of treatment. Despite the promising diagnostic performance of the signature for active TB disease in HIV-infected persons, the signature could not differentiate early from late converters. It is unclear to what degree the signature performance was poor because participants in this cohort were TB retreatment patients, most of whom had a TB episode within five years before the IMPRESS study.

As discussed in previous chapters, there is evidence that patients with recent TB disease have ongoing inflammation that can be detected by IFN response signatures. This inflammation was also detected by FDG uptake on PET-CT in HIV-uninfected persons up to a year after mycobacterial cure (Malherbe et al. 2016). Furthermore, we showed that HIV-infection and pVL increased signature scores, while Type I IFNs are known to be upregulated in HIV-infected persons with high levels of immune activation (Herbeuval and Shearer 2007). We hypothesise that the poor distinction by the ACS 6-gene signature between patients who have achieved

and not achieved mycobacterial clearance as measured by sputum culture at two months is due to underlying HIV and the ongoing effects of the previous TB episode.

The 6-transcript pairwise configuration of the ACS 6-gene signature has several advantages. The most important feature of this signature is its robustness to missing primer-probes despite the small numbers of genes. This suggests that the signature is as robust to missing primer-probes as the large ACS 11-gene signature, which comprises of 48 primer-probes. The lack of reference (or housekeeper) genes means that only six primer-probes are needed to measure the gene signature and simplifies the data processing steps since no standardisation is necessary. Surprisingly, the performance of the signature was equivalent to that of the ACS 11-gene signature including diagnosis of active TB disease, prediction of recurrent TB (or detection of subclinical TB), and monitoring treatment response in HIV-infected persons. In light of its small size, we propose that the ACS 6-gene signature is an ideal TB biomarker in HIV-uninfected and HIV-infected persons. The test could be used a triage test in clinical stteines, however we currently do not have an estimate of the required sensitivity and specificity in patient clicare that is missng. The signature could be a useful tool for host-direct therapy and work is already underway to adapt the signature assay such that it can be measured in a simple, point-of-care battery operated device to be used as a triage test in primary healthcare facilities, ART clinics, and in community settings.

7.6: Contributions

Ms F. Darboe designed, conducted the experiments, performed the analyses, and wrote this chapter under the supervision of Dr A. Penn-Nicholson and A. Prof T.J. Scriba. The R scripts to calculate signatures was written by Dr E.G Thompson.

Chapter 8: General Discussion and Conclusions

The targets set by the STOP-TB Partnership to eradicate TB by 2050 will not be reached unless highly innovative and new approaches are promoted. This includes high-quality, non-sputum based diagnostic or triage tests that can be done either in community settings or in primary clinics where individuals first present for healthcare. Such tests, especially if found to be sensitive in persons with subclinical disease, will increase early TB case detection, reduce missed cases at the initial stage of the care cascade, and perhaps even reduce transmission of *Mtb*.

Many novel blood transcriptional signatures, consisting mainly of Type I/II IFN stimulated genes (ISGs), have recently been developed as diagnostic (Berry et al., 2010; Maertzdorf et al., 2011, 2012; Ottenhoff et al., 2012; Sweeney et al., 2016) or prognostic (Zak et al. 2016; Darboe et al. 2018) biomarkers of active TB disease. The readouts of such signatures also correlate with the extent of disease, suggesting that ISG signatures can be used for TB treatment response monitoring (Berry et al. 2010; Cliff et al. 2013; Thompson et al. 2017). Despite the massive disease burden of TB in HIV-infected persons and the high risk of progression to active TB in this population, only a few studies have investigated diagnostic gene expression biomarkers in HIV-infected persons (Kaforou et al. 2013; Anderson et al. 2014; Sweeney et al. 2016). Increasing evidence suggests these signatures may have universal utility, with diagnostic, prognostic, and treatment response applications (Singhania et al. 2017; Zak et al. 2016; Thompson et al. 2017). However, to the best of our knowledge this universal application has not yet been tested in a single study, either in HIV-infected or in HIV-uninfected persons. The aim of this thesis was to determine the diagnostic, prognostic and treatment response performance of

previously discovered ISG-based prognostic gene signatures of TB, applied in blinded analyses, in HIV-infected individuals. Our findings represent the first steps in the endeavour to develop blood-based gene expression signatures for TB, highlight some important challenges with the approach, and present interesting conclusions.

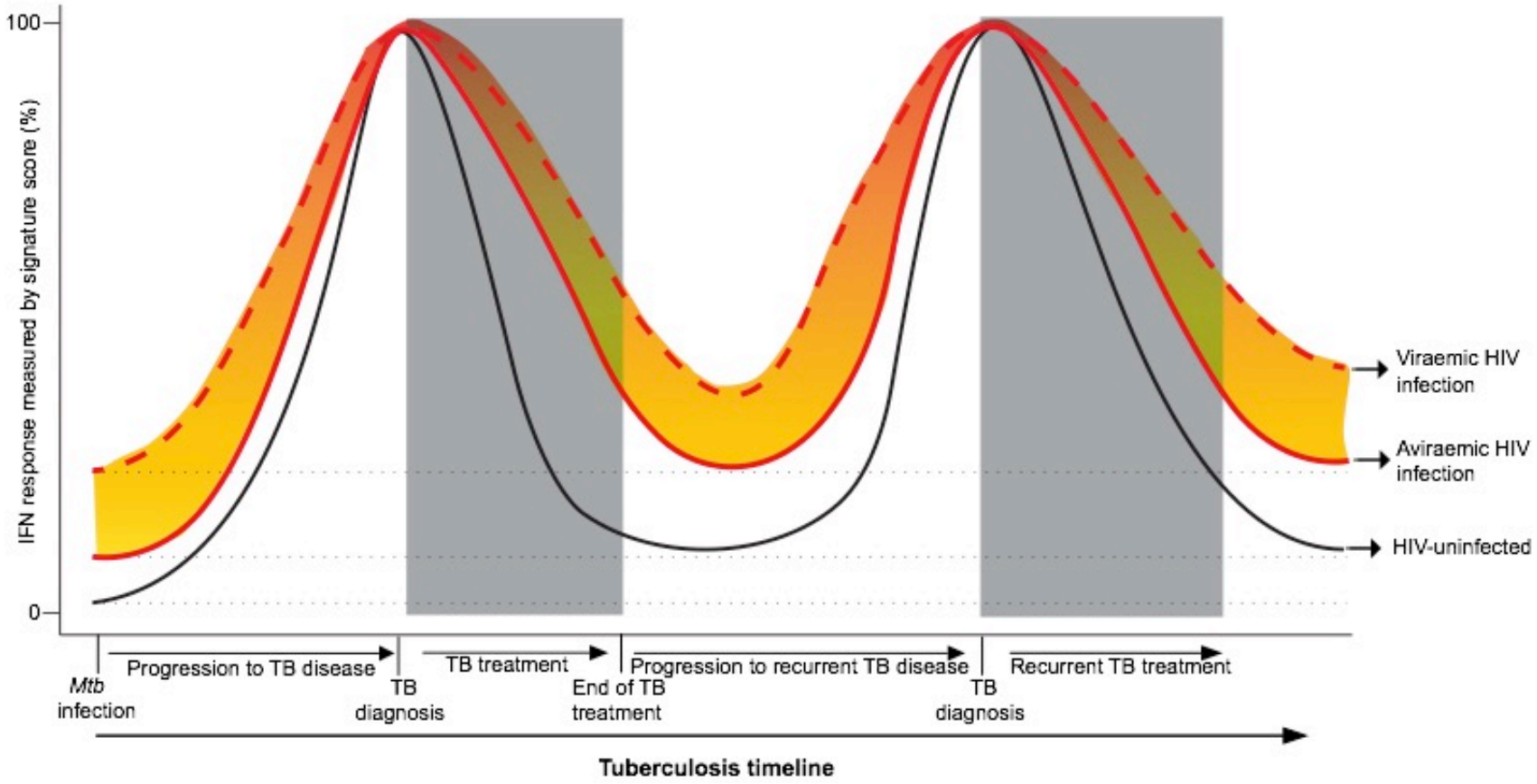
Our first major finding was that signatures developed in HIV-uninfected populations for prognostic application, (i) identify HIV-infected persons with active TB disease, (ii) predict incident recurrent TB three months before microbiological diagnosis and, (iii) monitor treatment response of recurrent TB disease. However, the diagnostic performance observed in HIV-infected persons was approximately 10% lower than that observed in HIV-uninfected persons. We show that this is most likely due to higher false positive rates of signature scores in HIV-infected populations. Our data also corroborate previous findings that underlying HIV infection drives Type I IFN responses (Fernandez et al. 2011; Kamga et al. 2005), thus undermining the specificity of the TB signatures in HIV-infected populations. The data further suggest that whole transcriptome analysis (or other omics platforms) would be ideal to discover novel biological pathways that are not confounded by HIV. We acknowledge that the lack of viral load data in the pilot cohort of TB cases and QFT-positive controls (Chapter 3) was a major limitation. Future studies should collect detailed clinical, immunological and virological data to allow careful and comprehensive analyses of these factors in biomarker discovery and performance.

Our analyses of the ACS 11-gene, the rPSVM.1, and the ACS 6-gene signatures in the TRuTH cohort showed that prediction of recurrent TB disease was only possible very close to (within three months) of microbiological TB diagnosis. We showed that

the poor performance was not due to utilizing PBMC samples instead of whole blood, since we observed equivalent diagnostic performance in PBMC and whole blood in our pilot cohort. The poor predictive performance of the signatures in the TRuTH cohort is likely due to a combination of the following five factors: 1) ongoing inflammation from the previous episode of TB disease may have affected signature sensitivity and specificity; 2) underlying HIV-infection likely affected signature performance (as observed for diagnostic performance); 3) most patients had subclinical TB disease at diagnosis, which is more likely to present with lower IFN signatures than active clinical disease; 4) intensive active case finding by induced sputum was frequently performed, allowing cases to be detected early during disease progression; and 5) many individuals received IPT, possibly affecting disease progression and affecting the risk of recurrent TB disease. Due to intensified case finding, most recurrent cases (25 of 43) detected were asymptomatic at the time of diagnosis and therefore had subclinical disease. Subclinical TB disease has been identified in other studies of active TB screening and case finding (Corbett and MacPherson 2013; Onozaki et al. 2015). More than half (n=212) of participants eligible for IPT in this cohort were initiated on six months of daily IPT (Maharaj et al. 2017). Interestingly, recurrent TB disease was diagnosed shortly (median, 279 days) after end of preventive therapy in many of these individuals (Maharaj et al. 2017), as has been previously observed in other studies of HIV-infected persons who underwent six months of IPT (Samandari et al. 2011, 2015; J. E. Golub et al. 2015). This highlights that an important novel application of blood-based RNA signatures, which could be used to determine if HIV-infected persons are at risk of developing TB or already have subclinical TB when they complete their IPT course. This would allow prolongation of IPT treatment in individuals with high signature scores, thus

allowing their risk of active TB disease to be reduced. Alternatively, provision of full TB treatment may be necessary in cases were IPT does not reduce signature scores. In addition, gene signatures could be used to monitor preventive treatment response in HIV-infected retreatment cases, as was demonstrated for full TB treatment in HIV-uninfected persons (Thompson et al. 2017).

Another major finding of this study is the effect of HIV infection and plasma viral load on signature scores, which shows that HIV infection elevates ISG expression as previously shown (Bosinger and Utay 2015; Fernandez et al. 2011; Kamga et al. 2005). Within the HIV-infected population, aviraemic persons had lower expression of ISGs in comparison to their viraemic counterparts. The association of high signature scores in individuals with detectable viral loads suggests an underlying state of on-going immune activation and inflammation, decreasing the specificity of ISG-based gene expression signatures. These data highlight the importance of accounting for HIV co-infection and plasma viral load levels when developing or validating signatures of active TB disease. This is clearly illustrated by the finding that the 393-gene signature developed in HIV-uninfected persons by Berry *et al* (2010), had 0% specificity in an HIV-infected cohort due to the high signature expression in HIV-infected controls (Dawany et al. 2014). A major limitation of our sub-analyses was the small sample size such that the performance of the signatures could not be determined in highly viraemic samples.


100
0
IFN response measured by signature score (%)
Viraemic HIV infection
Aviraemic HIV infection
HIV-uninfected
Progression to TB disease
TB treatment
Progression to recurrent TB disease
Recurrent TB treatment
Mtb infection
TB diagnosis
End of TB treatment
TB diagnosis
Tuberculosis timeline

Perhaps the most insightful outcome from our study is that it allowed us to summarise the kinetics of type I/II IFN responses during progression to TB disease, during treatment, and recurrence of TB disease (Figure 36). Expression of ISGs increased as individuals transitioned to subclinical and then active clinical TB disease, likely due to the effect of bacterial replication and/or host cell mitochondrial damage during cell death (Wiens and Ernst 2016). A rapid decrease in ISG expression was observed during TB treatment, more so in HIV-uninfected (Thompson et al. 2017) than infected persons. This likely reflects lower levels of bacteria that are sensed through pathogen recognition receptors and perhaps reduced host cell death. However, in both HIV-infected and uninfected persons, gene expression of ISGs did not return to the original levels typically seen in healthy persons despite clinical and microbiological cure (Malherbe et al. 2016; Thompson et al. 2017). As observed in primary TB disease progression, recurrent TB disease progression was associated with an elevation in signature scores. The rapid increase of signature scores in HIV-infected persons before recurrent TB disease appears to reflect the increased risk and more rapid progression to recurrent TB disease relative to HIV-uninfected persons (Korenromp et al., 2003; Maher et al., 2005). As observed during treatment of a first TB episode, ISG expression rapidly decreased in HIV-uninfected persons during TB treatment in comparison to HIV-infected persons (Figure 36). These upward and downward waves of type I/II IFN response kinetics appear to neatly tie together the transitions between quiescent, subclinical, and active disease as well as antibiotic treatment of TB, and even recurrence. However, more research on independent cohorts is required to determine if this proposed model can be reproduced.

The blood transcriptomic signatures could be useful as triage, diagnostic, or prognostic tests for active TB disease in resource-poor settings if measurement of gene expression can be translated into a cost-effective and near-patient device. A triage test ideally should be more sensitive than the confirmatory test in diagnosing TB, otherwise participants would be missed by the test (Denkinger et al. 2015). The ACS 11-gene, rPSVM.1, and ACS 6-gene signatures could diagnose Xpert MTB-RIF-positive TB with a high sensitivity and specificity in both HIV-infected and uninfected persons. However, we show that the utility of the signatures as triage tests may be limited in HIV-infected persons due to viraemia, especially in persons not on ART and not virologically suppressed (Table 16). Despite this, due to effective roll out of ART in most TB endemic countries (World Health Organization 2017), most HIV-infected persons are expected to be aviraemic, which may make IFN response signatures particularly useful in XpertMTB-RIF or sputum smear negative individuals, before culture results become available. Furthermore, these signatures would be particularly useful as a diagnostic tool in HIV-infected persons with paucibacillary disease (Cain et al. 2010), and could serve as an important diagnostic tool in children, who also cannot provide sputum samples. To detect one case of TB using the signatures as triage tests (using the WHO target product profile (TPP) specificities), using the 6-gene signature would require lower numbers to screen and yield less false-positive tests regardless of HIV-status than the 11-gene signature (Table 16). This suggests that the 6-gene signature may perform better than the 11-gene signature as a triage test in HIV-infected persons. However, a major limitation in HIV-infected persons is the interpretation of results, as underlying immune activation and inflammation can still lead to false positive results, especially in persons with previous TB disease. Regardless, the sensitivity and specificity of the

transcriptomic signatures (Table 16) still suggest better diagnostic performance than symptom screening alone or symptom screening in parallel with chest radiography in HIV-infected persons (Table 16) (Getahun et al. 2011).

Table 16: Performance of the 11-gene and 6-gene signatures as triage tests in HIV-infected and uninfected persons based on results from the pilot cohort.

Signature	HIV status	Sensitivity	Specificity	Number needed to screen	False positives	False Negatives
Triage test (WHO TPP)		>95%	>80%	<55	<11	1
11-gene	Uninfected	100%	80%	50	10	0
11-gene	Infected	65%	80%	77	16	1
6-gene	Uninfected	100%	80%	50	10	0
6-gene	Infected	85%	80%	59	12	1
Performance of current screening tools in HIV-infected persons						
Symptom screening only		79%	50%	64	32	1
Symptom screening and chest radiography		91%	39%	55	33	1

Number needed to screen to detect one case of TB is shown. For these calculations prevalent TB was set at 2%. False positives and negatives were calculated per number screened to detect one cases of TB. Calculations for symptom screening and chest radiography are based on results from Getahun *et al* 2011. TPP-Target product profile (WHO, 2012).

HIV infection represents a highly heterogeneous spectrum of immunosuppression, viraemia, lymphoid tissue destruction and co-infections. This study highlighted the different effects of viraemic and aviraemic HIV-infection on the sensitivity and specificity of ISG signatures in predicting and diagnosing active TB disease. In addition, we identified plasma viral load as the main confounder in the application of TB signatures based on ISGs in HIV-infected persons. Successful provision of ART and good adherence leading to effective suppression of viral replication would decrease ISG expression in HIV-infected persons and improve the performance of whole blood-based ISG transcriptomic signatures. Despite ART, it is apparent that a certain level of elevated IFN expression persists due to HIV infection. Integrating HIV and TB clinical care would allow for the effective monitoring of virological

suppression, provision of targeted preventive therapy, or further investigations such as sputum induction to determine if the person has TB disease. The utility of gene signatures in the control of TB disease would be increased in such an integrated setting, thereby reducing transmission and increasing early TB case detection rates.

References

Abdool Karim, Salim, Kogieleum Naidoo, Anneke Groble, Nesri Padayatchi, Cheryl Baxter, Andrew L Gray, Tanuja Gengiah, et al. 2011. “Integration of Antiretroviral Therapy with Tuberculosis Treatment.” *The New England Journal of Medicine* 365:1492–1501.

Abdool Karim, Salim S, Kogieleum Naidoo, Anneke Grobler, Nesri Padayatchi, Cheryl Baxter, Andrew Gray, Tanuja Gengiah, et al. 2010. “Timing of Initiation of Antiretroviral Drugs during Tuberculosis Therapy.” *The New England Journal of Medicine* 362 (8):697–706.

Abel, Kristina, Michelle J Alegria-Hartman, Kristina Rothaeusler, Marta Marthas, and Christopher J Miller. 2002. “The Relationship between Simian Immunodeficiency Virus RNA Levels and the mRNA Levels of Alpha/beta Interferons (IFN-Alpha/beta) and IFN-Alpha/beta-Inducible Mx in Lymphoid Tissues of Rhesus Macaques during Acute and Chronic Infection.” *Journal of Virology* 76 (16):8433–45.

Abu-Raddad, Laith J., Lorenzo Sabatelli, Jerusha T. Achterberg, Jonathan D. Sugimoto, Ira M. Longini, Christopher Dye, and M. Elizabeth Halloran. 2009. “Epidemiological Benefits of More-Effective Tuberculosis Vaccines, Drugs, and Diagnostics.” *PNAS* 106 (33):13980–85.

Akolo, Christopher, Ifedayo Adetifa, Sasha Shepperd, and Jimmy Volmink. 2010. “Treatment of Latent Tuberculosis Infection in HIV Infected Persons.” *Cochrane Database Syst Rev*, no. 1.

Andersen, P, M E Munk, J M Pollock, and T M Doherty. 2000. “Specific Immune-Based Diagnosis of Tuberculosis.” *The Lancet* 356 (9235):1099–1104.

Anderson, Suzanne T., Myrsini Kaforou, Andrew J. Brent, Victoria J. Wright, Claire

M. Banwell, George Chagaluka, Amelia C. Crampin, et al. 2014. “Diagnosis of Childhood Tuberculosis and Host RNA Expression in Africa.” *New England Journal of Medicine* 370 (18):1712–23.

Auguste, Peter, Alexander Tsertsvadze, Rachel Court, and Joshua Pink. 2016. “A Systematic Review of Economic Models Used to Assess the Cost-Effectiveness of Strategies for Identifying Latent Tuberculosis in High-Risk Groups.” *Tuberculosis* 99. Elsevier Ltd:81–91.

Badje, Anani, Raoul Moh, Delphine Gabillard, Calixte Guéhi, Mathieu Kabran, Jean-Baptiste Ntakpé, Jérôme Le Carrou, et al. 2017. “Effect of Isoniazid Preventive Therapy on Risk of Death in West African, HIV-Infected Adults with High CD4 Cell Counts: Long-Term Follow-up of the Temprano ANRS 12136 Trial.” *The Lancet Global Health* 5 (11):e1080–89.

Badri, Motasim, Douglas Wilson, and Robin Wood. 2002. “Effect of Highly Active Antiretroviral Therapy on Incidence of Tuberculosis in South Africa: A Cohort Study.” *The Lancet* 359 (9323):2059–64.

Barry, Clifton E., Helena I. Boshoff, Véronique Dartois, Thomas Dick, Sabine Ehrt, JoAnne Flynn, Dirk Schnappinger, Robert J. Wilkinson, and Douglas Young. 2009. “The Spectrum of Latent Tuberculosis: Rethinking the Biology and Intervention Strategies.” *Nature Reviews Microbiology* 7 (12). Nature Publishing Group:845–55.

Benator, Debra, Mondira Bhattacharya, Lorna Bozeman, William Burman, Antonino Catanzaro, Richard Chaisson, Fred Gordin, et al. 2002. “Rifapentine and Isoniazid Once a Week versus Rifampicin and Isoniazid Twice a Week for Treatment of Drug-Susceptible Pulmonary Tuberculosis in HIV-Negative Patients: A Randomised Clinical Trial.” *Lancet* 360 (9332):528–34.

Berry, Matthew P.R. R, Christine M. Graham, Finlay W. McNab, Zhaohui Xu, Susannah A.A. a Bloch, Tolu Oni, Katalin A. Wilkinson, et al. 2010. “An Interferon-Inducible Neutrophil-Driven Blood Transcriptional Signature in Human Tuberculosis.” *Nature* 466 (7309):973–77.

Blankley, Simon, Matthew Paul Reddoch Berry, Christine M Graham, Chloe I Bloom, Marc Lipman, and Anne O’Garra. 2014. “The Application of Transcriptional Blood Signatures to Enhance Our Understanding of the Host Response to Infection: The Example of Tuberculosis.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 369 (1645):20130427.

Blankley, Simon, Christine M Graham, Joe Levin, Jacob Turner, Matthew P R Berry, Chloe I Bloom, Zhaohui Xu, Virgina Pascual, and Jacques Banchereau. 2016. “A 380 Gene Meta-Signature of Active TB Compared to Healthy Controls.” *Eurpoean Respiratoy Journal* 47 (6):1873–76.

Bloom, Chloe I., Christine M. Graham, Matthew P.R. R Berry, Fotini Rozakeas, Paul S. Redford, Yuanyuan Wang, Zhaohui Xu, et al. 2013. “Transcriptional Blood Signatures Distinguish Pulmonary Tuberculosis, Pulmonary Sarcoidosis, Pneumonias and Lung Cancers.” *PLoS ONE* 8 (8).

Bloom, Chloe I., Christine M. Graham, Matthew P.R. R Berry, Katalin A. Wilkinson, Tolu Oni, Fotini Rozakeas, Zhaohui Xu, et al. 2012. “Detectable Changes in The Blood Transcriptome Are Present after Two Weeks of Antituberculosis Therapy.” *PLoS ONE* 7 (10).

Bock, Naomi N., Timothy R. Sterling, Carol D. Hamilton, Connie Pachucki, Yong Cheng Wang, Donna S. Conwell, Ann Mosher, Mary Samuels, and Andrew Vernon. 2002. “A Prospective, Randomized, Double-Blind Study of the Tolerability of Rifapentine 600, 900, and 1,200 Mg plus Isoniazid in the

Continuation Phase of Tuberculosis Treatment." *American Journal of Respiratory and Critical Care Medicine* 165 (11):1526–30.

Boer, A S De, M W Borgdorff, E Vynnycky, M M Sebek, and D Van Soolingen. 2003. "Exogenous Re-Infection as a Cause of Recurrent Tuberculosis in a Low-Incidence Area." *European Respiratory Journal* 7 (May 2002):145–52.

Boon, Saskia Den, Alberto Matteelli, Nathan Ford, and Haileyesus Getahun. 2016. "Continuous Isoniazid for the Treatment of Latent Tuberculosis Infection in People Living with HIV." *Aids* 30 (5):797–801.

Bosinger, Steven E., and Netanya S. Utay. 2015. "Type I Interferon: Understanding Its Role in HIV Pathogenesis and Therapy." *Current HIV/AIDS Reports* 12:41–53.

Brenchley, Jason M., David A. Price, Timothy W. Schacker, Tedi E. Asher, Guido Silvestri, Srinivas Rao, Zachary Kazzaz, et al. 2006. "Microbial Translocation Is a Cause of Systemic Immune Activation in Chronic HIV Infection." *Nature Medicine* 12 (12):1365–71.

Bunyasi, E W, A K K Luabeya, M Tameris, H Geldenhuys, H Mulenga, B S Landry, T J Scriba, et al. 2017. "Impact of Isoniazid Preventive Therapy on the Evaluation of Long-Term Effectiveness of Infant MVA85A Vaccination." *The International Journal of Tuberculosis and Lung Disease* 21 (7).

Burman, William J., Stefan Goldberg, John L. Johnson, Grace Muzanye, Melissa Engle, Ann W. Mosher, Shurjeel Choudhri, et al. 2006. "Moxifloxacin versus Ethambutol in the First 2 Months of Treatment for Pulmonary Tuberculosis." *American Journal of Respiratory and Critical Care Medicine* 174 (3):331–38.

Cain, K P, K D McCarthy, C M Heilig, P Monkongdee, T Tasaneeyapan, N Kanara, M E Kimerling, et al. 2010. "An Algorithm for Tuberculosis Screening and

Diagnosis in People with HIV." *The New England Journal of Medicine* 362 (8):707–16.

Caruso, A M, N Serbina, E Klein, K Triebold, B R Bloom, and J L Flynn. 1999. "Mice Deficient in CD4 T Cells Have Only Transiently Diminished Levels of IFN-Gamma, yet Succumb to Tuberculosis." *Journal of Immunology* 162 (9):5407–16.

Catalfamo, M., M. Di Mascio, Z. Hu, S. Srinivasula, V. Thaker, J. Adelsberger, A. Rupert, et al. 2008. "HIV Infection-Associated Immune Activation Occurs by Two Distinct Pathways That Differentially Affect CD4 and CD8 T Cells." *Proceedings of the National Academy of Sciences* 105 (50):19851–56.

Cepheid. 2017. "GeneXpert Omni." 2017. http://www.cepheid.com/us/genexpert-omni.

Charalambous, S, A D Grant, V Moloi, R Warren, J H Day, P van Helden, R J Hayes, et al. 2008. "Contribution of Reinfection to Recurrent Tuberculosis in South African Gold Miners." *International Journal of Tuberculosis and Lung Disease* 12 (8):942–48.

Chen, Kun, Juan Liu, and Xuetao Cao. 2017. "Regulation of Type I Interferon Signaling in Immunity and Inflammation: A Comprehensive Review." *Journal of Autoimmunity* 83. Elsevier Ltd:1–11.

Chindelevitch, L., N. A. Menzies, C. Pretorius, J. Stover, J. A. Salomon, and T. Cohen. 2015. "Evaluating the Potential Impact of Enhancing HIV Treatment and Tuberculosis Control Programmes on the Burden of Tuberculosis." *Journal of The Royal Society Interface* 12 (106):20150146–20150146.

Churchyard, G J, K L Fielding, J J Lewis, L Coetzee, E L Corbett, P Godfrey-Faussett, R J Hayes, R E Chaisson, A D Grant, and T B Study Team Thibela.

2014. “A Trial of Mass Isoniazid Preventive Therapy for Tuberculosis Control.” *New England Journal of Medicine* 370 (4):301–10.

Churchyard, Gavin J., Katherine L. Fielding, James J. Lewis, Leonie Coetzee, Elizabeth L. Corbett, Peter Godfrey-Faussett, Richard J. Hayes, Richard E. Chaisson, and Alison D. Grant. 2014. “A Trial of Mass Isoniazid Preventive Therapy for Tuberculosis Control.” *New England Journal of Medicine* 370 (4):301–10.

Cihlar, Tomas, and Marshall Fordyce. 2016. “Current Status and Prospects of HIV Treatment.” *Current Opinion in Virology* 18. Elsevier B.V.:50–56.

Cliff, Jacqueline M., Stefan H.E. E Kaufmann, Helen Mcshane, Paul van Helden, Anne O’Garra, Paul Van Helden, Anne O Garra, et al. 2015. “The Human Immune Response to Tuberculosis and Its Treatment: A View from the Blood.” *Immunological Reviews* 264 (1):88–102.

Cliff, Jacqueline M., Ji Sook Lee, Nicholas Constantinou, Jang Eun Cho, Taane G. Clark, Katharina Ronacher, Elizabeth C. King, et al. 2013. “Distinct Phases of Blood Gene Expression Pattern through Tuberculosis Treatment Reflect Modulation of the Humoral Immune Response.” *Journal of Infectious Diseases* 207 (1):18–29.

Comstock, G W, C Baum, and D E Snider. 1979. “Isoniazid Prophylaxis among Alaskan Eskimos: A Final Report of the Bethel Isoniazid Studies.” *American Review of Respiratory Disease* 119 (5):827–30.

Comstock, George, Shirley H. Ferebee, and Laurel M Hammes. 1967. “A Controlled Trial of Community-Wide Isoniazid Prohylaxis in Alaska.” *American Review of Respiratory Disease* 95 (6):935–43.

Corbeau, Pierre, and Jacques Reynes. 2011. “Immune Reconstitution under

Antiretroviral Therapy : The New Challenge in HIV-1 Infection." *Blood* 117 (21):5582–90.

Corbett, E.L, and P. MacPherson. 2013. "Tuberculosis Screening in High Human Immunodeficiency Virus Prevalence Settings : Turning Promise into Reality." *International Journal of Tuberculosis and Lung Disease.* 17 (9):1125–38.

Crampin, Amelia C., J. Nimrod Mwaungulu, Frank D. Mwaungulu, D. Totah Mwafulirwa, Kondwani Munthali, Sian Floyd, Paul E M Fine, and Judith R. Glynn. 2010. "Recurrent TB: Relapse or Reinfection? The Effect of HIV in a General Population Cohort in Malawi." *Aids* 24 (3):417–26.

Crevel, Reinout Van, Tom H M Ottenhoff, and Jos W M Van Der Meer. 2002. "Innate Immunity to Mycobacterium Tuberculosis." *Clinical Microbiology Reviews* 15 (2):294–309. https://doi.org/10.1128/CMR.15.2.294.

Darboe, Fatoumatta, Stanley Kimbung S.K. Mbandi, Ethan G. E.G. Thompson, Michelle Fisher, Miguel Rodo, Michele van Rooyen, Elizabeth Filander, et al. 2018. "Diagnostic Performance of an Optimized Transcriptomic Signature of Risk of Tuberculosis in Cryopreserved Peripheral Blood Mononuclear Cells." *Tuberculosis* 108:124–26.

Davies, P D O, and M Pai. 2008. "The Diagnosis and Misdiagnosis of Tuberculosis." *International Journal of Tuberculosis and Lung Disease* 12 (11):1226–34.

Dawany, Noor, Louise C. Showe, Andrew V. Kossenkov, Celia Chang, Prudence Ive, Francesca Conradie, Wendy Stevens, Ian Sanne, Livio Azzoni, and Luis J. Montaner. 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 (2):1–8.

Deeks, Steven G, Christina M R Kitchen, Lea Liu, Hua Guo, Ron Gascon, Amy B

Narváez, Peter Hunt, et al. 2004. “Immune Activation Set Point during Early HIV Infection Predicts Subsequent CD4 + T-Cell Changes Independent of Viral Load.” *Blood* 104 (4):942–47.

Deeks, Steven G, Russel Tracy, and Daniel C Douek. 2013. “Systemic Effects of Inflammation on Health during Chronic HIV Infection.” *Immunity* 39 (4):633–45.

Denkinger, Claudia M., Sandra V. Kik, Daniela Maria Cirillo, Martina Casenghi, Thomas Shinnick, Karin Weyer, Chris Gilpin, et al. 2015. “Defining the Needs for Next Generation Assays for Tuberculosis.” *Journal of Infectious Diseases* 211 (suppl 2):S29–38.

Dheda, Keertan, Laura Lenders, Gesham Magombedze, Shasshikant Srivastava, Prithvi Raj, Erland Arning, Paula Azhcraft, et al. 2018. “Drug Penetration Gradients Associated with Acquired Drug Resistance in Tuberculosis Patients.” *American Journal of Respiratory and Critical Care Medicine*.

Diedrich, Collin R., and JoAnne L. Anne L. Flynn. 2011. “HIV-1/Mycobacterium Tuberculosis Coinfection Immunology: How Does HIV-1 Exacerbate Tuberculosis?” *Infection and Immunity* 79 (4):1407–17.

Diel, R., D. Goletti, G. Ferrara, G. Bothamley, D. Cirillo, B. Kampmann, C. Lange, et al. 2011. “Interferon-γ Release Assays for the Diagnosis of Latent Mycobacterium Tuberculosis Infection: A Systematic Review and Meta-Analysis.” *European Respiratory Journal* 37 (1):88–99.

Dodd, Claire E., and Larry S. Schlesinger. 2017. “New Concepts in Understanding Latent Tuberculosis.” *Current Opinion in Infectious Diseases* 30 (3):316–21.

Dominguez, Maria H., Pratip K. Chattopadhyay, Steven Ma, Laurie Lamoreaux, Andrew McDavid, Greg Finak, Raphael Gottardo, Richard A. Koup, and Mario Roederer. 2013. “Highly Multiplexed Quantitation of Gene Expression on Single

Cells.” *Journal of Immunological Methods* 391 (1–2). Elsevier B.V.:133–45.

Donnelly, Raymond P, and Sergei V Kotenko. 2010. “Interferon-Lambda: A New Addition to an Old Family.” *Journal of Interferon & Cytokine Research* 30 (8):555–64.

Donovan, Meg L., Thomas E. Schultz, Taylor J. Duke, and Antje Blumenthal. 2017. “Type I Interferons in the Pathogenesis of Tuberculosis: Molecular Drivers and Immunological Consequences.” *Frontiers in Immunology* 8 (NOV).

Dowdy, D. W., R. E. Chaisson, G. Maartens, E. L. Corbett, and S. E. Dorman. 2008. “Impact of Enhanced Tuberculosis Diagnosis in South Africa: A Mathematical Model of Expanded Culture and Drug Susceptibility Testing.” *Proceedings of the National Academy of Sciences* 105 (32):11293–98.

Durovni, Betina, Valeria Saraceni, Lawrence H Moulton, Antonio G Pacheco, Richard E Chaisson, and Jonathan E Golub. 2013. “Impact of Tuberculosis Screening and Isoniazid Preventive Therapy on Incidence of Tuberculosis and Death in Patients with HIV Infection Receiving Care in Public Clinics in Rio de Janeiro, Brazil: The Tuberculosis/HIV in Rio de Janeiro (THRio) Study: A Ste.” *Lancet Infectious Diseases* 13 (10):852–58.

Dye, Christopher, Philippe Glaziou, Katherine Floyd, and Mario Raviglione. 2013. “Prospects for Tuberculosis Elimination.” *Annual Review of Public Health* 34 (1):271–86.

Dye, Christopher, Suzanne Scheele, Paul Dolin, Vikram Pathania, and Mario C Raviglione. 1999. “Global Burden of Tuberculosis Estimated Incidence , Prevalence , and Mortality by Country.” *JAMA* 282 (7):677–86.

Esmail, H., C. E. Barry, D. B. Young, and R. J. Wilkinson. 2014. “The Ongoing Challenge of Latent Tuberculosis.” *Philosophical Transactions of the Royal*

Society B 369 (1645).

Esmail, H, C E Barry, D B Young, R J Wilkinson, and Phil Trans R Soc B. 2014. “The Ongoing Challenge of Latent Tuberculosis.” *Philosophical Transactions of Biological Scinces*, no. May:1–14.

Esmail, Hanif, Rachel P. Lai, Maia Lesosky, Katalin A. Wilkinson, Christine M. Graham, Stuart Horswell, Anna K. Coussens, Clifton E. Barry, Anne O’Garra, and Robert J. Wilkinson. 2018. “Complement Pathway Gene Activation and Rising Circulating Immune Complexes Characterize Early Disease in HIV-Associated Tuberculosis.” *Proceedings of the National Academy of Sciences* 115 (5):E964–73.

Esmail, Hanif, Rachel P Lai, Maia Lesosky, Katalin A Wilkinson, Chrisitine M Graham, Anna K Coussens, Tolu Oni, et al. 2016. “Characterization of Progressive HIV-Associated Tuberculosis Using 2-Deoxy-2-[18F]fluoro-D-Glucose Positron Emission and Computed Tomography.” *Nature Medicine* 22 (10). Nature Publishing Group:1090–93.

Etna, Marilena Paola, Elena Giacomini, Martina Severa, and Eliana Marina Coccia. 2014. “Pro-and Anti-Inflammatory Cytokines in Tuberculosis: A Two-Edged Sword in TB Pathogenesis.” *Seminars in Immunology* 26 (6). Elsevier Ltd:543–51.

Eum, Seok Yong, Ji Hye Kong, Min Sun Hong, Ye Jin Lee, Jin Hee Kim, Soo Hee Hwang, Sang Nae Cho, Laura E. Via, and Clifton E. Barry. 2010. “Neutrophils Are the Predominant Infected Phagocytic Cells in the Airways of Patients with Active Pulmonary TB.” *Chest* 137 (1):122–28.

Ferebee, Shirley H, and Frank W Mount. 1962. “Tuberculosis Morbidity in a Controlled Trial of the Prophylactic Use of Isoniazid among Household

Contacts.” *American Review of Respiratory Disease* 85:490–510.

Fernandes, Guilherme Felipe dos Santos, Hérida Regina Nunes Salgado, and Jean Leandro dos Santos. 2017. “Isoniazid: A Review of Characteristics, Properties and Analytical Methods.” *Critical Reviews in Analytical Chemistry* 0 (0). Taylor & Francis:1–11.

Fernandez, Sonia, Sara Tanaskovic, Karla Helbig, Reena Rajasuriar, Marit Kramski, John M. Murray, Michael Beard, et al. 2011. “CD4+ T-Cell Deficiency in HIV Patients Responding to Antiretroviral Therapy Is Associated With Increased Expression of Interferon-Stimulated Genes in CD4+ T Cells.” *The Journal of Infectious Diseases* 204 (12):1927–35.

Fitzgerald, Daniel W., Moïse Desvarieux, Patrice Severe, Patrice Joseph, Warren D. Johnson, and Jean William Pape. 2000. “Effect of Post-Treatment Isoniazid on Prevention of Recurrent Tuberculosis in HIV-1-Infected Individuals: A Randomised Trial.” *Lancet* 356 (9240):1470–74.

Fletcher, Helen A, Margaret A Snowden, Bernard Landry, Wasima Rida, Iman Satti, Stephanie A Harris, Magali Matsumiya, et al. 2016. “T-Cell Activation Is an Immune Correlate of Risk in BCG Vaccinated Infants.” *Nature Communications* 7 (April). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.

Fox, Gregory J., Simone E. Barry, Warwick J. Britton, and Guy B. Marks. 2013. “Contact Investigation for Tuberculosis: A Systematic Review and Meta-Analysis.” *European Respiratory Journal* 41 (1):140–56.

Fraser, Christophe, Katrina Lythgoe, Gabriel E. Leventhal, George Shirreff, T. Déirdre Hollingsworth, Samuel Alizon, and Sebastian Bonhoeffer. 2014. “Virulence and Pathogenesis of HIV-1 Infection: An Evolutionary Perspective.”

Science 343 (6177).

Friedrich, Sven O., Andrea Rachow, Elmar Saathoff, Kasha Singh, Chacha D. Mangu, Rodney Dawson, Patrick P J Phillips, et al. 2013. “Assessment of the Sensitivity and Specificity of Xpert MTB/RIF Assay as an Early Sputum Biomarker of Response to Tuberculosis Treatment.” *The Lancet Respiratory Medicine* 1 (6). Elsevier Ltd:462–70.

Gengiah, Tanuja N., Nicholas H G Holford, Julia H. Botha, Andrew L. Gray, Kogieleum Naidoo, and Salim S Abdool Karim. 2012. “The Influence of Tuberculosis Treatment on Efavirenz Clearance in Patients Co-Infected with HIV and Tuberculosis.” *European Journal of Clinical Pharmacology* 68 (5):689–95.

Getahun, Haileyesus, Christian Gunneberg, Reuben Granich, and Paul Nunn. 2010. “HIV Infection–Associated Tuberculosis: The Epidemiology and the Response.” *Clinical Infectious Diseases* 50 (s3):S201–7.

Getahun, Haileyesus, Mark Harrington, Rick O’Brien, and Paul Nunn. 2007. “Diagnosis of Smear-Negative Pulmonary Tuberculosis in People with HIV Infection or AIDS in Resource-Constrained Settings: Informing Urgent Policy Changes.” *Lancet* 369 (9578):2042–49.

Getahun, Haileyesus, Wanitchaya Kittikraisak, Charles M. Heilig, Elizabeth L. Corbett, Helen Ayles, Kevin P. Cain, Alison D. Grant, et al. 2011. “Development of a Standardized Screening Rule for Tuberculosis in People Living with HIV in Resource-Constrained Settings: Individual Participant Data Meta-Analysis of Observational Studies.” *PLoS Medicine* 8 (1).

Getahun, Haileyesus, Alberto Matteelli, Ibrahim Abubakar, Mohamed Abdel Aziz, Annabel Baddeley, Draurio Barreira, Saskia Den Boon, et al. 2015. “Management of Latent Mycobacterium Tuberculosis Infection: WHO Guidelines

for Low Tuberculosis Burden Countries." *European Respiratory Journal* 46 (6):1563–76.

Gideon, Hannah P., Jason A. Skinner, Nicole Baldwin, JoAnne L. Flynn, and Philana Ling Lin. 2016. "Early Whole Blood Transcriptional Signatures Are Associated with Severity of Lung Inflammation in Cynomolgus Macaques with Mycobacterium Tuberculosis Infection." *The Journal of Immunology* 197 (12):4817–28.

Gillespie, Stephen H., Angela M. Crook, Timothy D. McHugh, Carl M. Mendel, Sarah K. Meredith, Stephen R. Murray, Frances Pappas, Patrick P.J. Phillips, and Andrew J. Nunn. 2014. "Four-Month Moxifloxacin-Based Regimens for Drug-Sensitive Tuberculosis." *New England Journal of Medicine* 371 (17):1577–87.

Golub, JE, V Saraceni, Cavalcante Solange, Pacheco Antonio, Lawrence Moulton, Bonnie King, Anne Efron, Richard Moore, Richard Chaisson, and Betina Durovni. 2007. "The Impact of Antiretroviral Therapy and Isoniazid Preventive Therapy on Tuberculosis Incidence in HIV-Infected Patients in Rio de Janeiro, Brazil." *AIDS* 21 (11):1441–48.

Golub, Jonathan E., Silvia Cohn, Valeria Saraceni, Solange C. Cavalcante, Antonio G. Pacheco, Lawrence H. Moulton, Betina Durovni, and Richard E. Chaisson. 2015. "Long-Term Protection from Isoniazid Preventive Therapy for Tuberculosis in HIV-Infected Patients in a Medium-Burden Tuberculosis Setting: The TB/HIV in Rio (THRio) Study." *Clinical Infectious Diseases* 60 (4):639–45.

Golub, Jonathan E, Paul Pronyk, Lerato Mohapi, Nkeko Thsabangu, Mosa Moshabela, Helen Struthers, Glenda E Gray, James A McIntyre, Richard E Chaisson, and Neil A Martinson. 2009. "Isoniazid Preventive Therapy, HAART and Tuberculosis Risk in HIV-Infected Adults in South Africa: A Prospective

Cohort." *Aids* 23 (5):631–36.

Gonzalez-Navajas, Jose M., Jongdae Lee, Michael David, and Eyal Raz. 2012. "Immunomodulatory Functions of Type I Interferons." *Nature Reviews Immunology* 12 (2):125–35.

Guerra-Assunçąõ, José Afonso, Rein M.G.J. Houben, Amelia C. Crampin, Themba Mzembe, Kim Mallard, Francesc Coll, Palwasha Khan, et al. 2015. "Recurrence due to Relapse or Reinfection with Mycobacterium Tuberculosis: A Whole-Genome Sequencing Approach in a Large, Population-Based Cohort with a High HIV Infection Prevalence and Active Follow-Up." *Journal of Infectious Diseases* 211 (7):1154–63.

Gupta, Ankur, Robin Wood, Richard Kaplan, Linda Gail Bekker, and Stephen D. Lawn. 2012. "Tuberculosis Incidence Rates during 8 Years of Follow-up of an Antiretroviral Treatment Cohort in South Africa: Comparison with Rates in the Community." *PLoS ONE* 7 (3):1–10.

Harries, Anthony D., Rony Zachariah, Elizabeth L. Corbett, Stephen D. Lawn, Ezio T. Santos-Filho, Rhehab Chimzizi, Mark Harrington, Dermot Maher, Brian G. Williams, and Kevin M. De Cock. 2010. "The HIV-Associated Tuberculosis Epidemic-When Will We Act?" *The Lancet* 375 (9729). Elsevier Ltd:1906–19.

Havlir, Diane V, Ian Sanne, F C P Sa, and Paul Nunn. 2008. "Opportunities and Challenges for HIV Care in Overlapping HIV and TB Epidemics." *Jama* 300 (4):423.

Herbeuval, Jean Philippe, and Gene M. Shearer. 2007. "HIV-1 Immunopathogenesis: How Good Interferon Turns Bad." *Clinical Immunology* 123 (2):121–28.

Hesseling, A C, G Walzl, D A Enarson, N M Carroll, K Duncan, P T Lukey, C

Lombard, et al. 2010. “Baseline Sputum Time to Detection Predicts Month Two Culture Conversion and Relapse in Non-HIV-Infected Patients.” *International Journal of Tuberculosis and Lung Disease* 14 (5):560–70.

Horne, David J., Sarah E. Royce, Lisa Gooze, Masahiro Narita, Philip C. Hopewell, Payam Nahid, and Karen R. Steingart. 2010. “Sputum Monitoring during Tuberculosis Treatment for Predicting Outcome: Systematic Review and Meta-Analysis.” *The Lancet Infectious Diseases* 10 (6). Elsevier Ltd:387–94.

Houben, Rein M. G. J., and Peter J. Dodd. 2016. “The Global Burden of Latent Tuberculosis Infection: A Re-Estimation Using Mathematical Modelling.” *PLoS Medicine* 13 (10):e1002152.

Houben, Rein M.G.J., Nicolas A. Menzies, Tom Sumner, Grace H. Huynh, Nimalan Arinaminpathy, Jeremy D. Goldhaber-Fiebert, Hsien Ho Lin, et al. 2016. “Feasibility of Achieving the 2025 WHO Global Tuberculosis Targets in South Africa, China, and India: A Combined Analysis of 11 Mathematical Models.” *The Lancet Global Health* 4 (11):e806–15.

Ivashkiv, Lionel B., and Laura T. Donlin. 2015. “Regulation of Type I Interferon Responses.” *Nature Reviews Immunology* 14 (1):36–49.

Jacobsen, Marc, Dirk Repsilber, Andrea Gutschmidt, Albert Neher, Knut Feldmann, Hans J. Mollenkopf, Andreas Ziegler, and Stefan H E Kaufmann. 2007. “Candidate Biomarkers for Discrimination between Infection and Disease Caused by Mycobacterium Tuberculosis.” *Journal of Molecular Medicine* 85 (6):613–21.

Jindani, Amina, Thomas S Harrison, Andrew J Nunn, Patrick PJ Phillips, Gavin J. Churchyard, Salome Charalambous, Mark Hatherill, et al. 2014. “High-Dose Rifapentine with Moxifloxacin for Pulmonary Tuberculosis.” *New England*

Journal of Medicine 371 (17):1599–1608.

Johnson, J L, A Okwera, D L Hom, H Mayanja, C Mutuluuza Kityo, P Nsubuga, J G Nakibali, et al. 2001. “Duration of Efficacy of Treatment of Latent Tuberculosis Infection in HIV-Infected Adults.” *AIDS* 15 (16):2137–47.

Joosten, Simone A., Helen a. Fletcher, and Tom H M Ottenhoff. 2013. “A Helicopter Perspective on TB Biomarkers: Pathway and Process Based Analysis of Gene Expression Data Provides New Insight into TB Pathogenesis.” *PLoS ONE* 8 (9).

Joshi, Anjali, Melina Sedano, Bethany Beauchamp, Erin B Punke, Zuber D Mulla, Armando Meza, Ogechika K Alozie, Debabrata Mukherjee, and Himanshu Garg. 2016. “HIV-1 Env Glycoprotein Phenotype along with Immune Activation Determines CD4 T Cell Loss in HIV Patients.” *The Journal of Immunology* 196 (4):1768–79.

Kaforou, Myrsini, Victoria J. Wright, Tolu Oni, Neil French, Suzanne T. Anderson, Nonzwakazi Bangani, Claire M. Banwell, et al. 2013. “Detection of Tuberculosis in HIV-Infected and -Uninfected African Adults Using Whole Blood RNA Expression Signatures: A Case-Control Study.” *PLoS Medicine* 10 (10).

Kamga, Isabelle, Sandrine Kahi, Leyla Develioglu, Miriam Lichtner, Concepción Marañón, Christiane Deveau, Laurence Meyer, et al. 2005. “Type I Interferon Production Is Profoundly and Transiently Impaired in Primary HIV-1 Infection.” *The Journal of Infectious Diseases* 192 (2):303–10.

Kasprowicz, Victoria O., Gavin Churchyard, Stephen D. Lawn, S. Bertel Squire, and Ajit Lalvani. 2011. “Diagnosing Latent Tuberculosis in High-Risk Individuals: Rising to the Challenge in High-Burden Areas.” *Journal of Infectious Diseases* 204 (SUPPL. 4).

Khan, Faiz A., Jessica Minion, Madhukar Pai, Sarah Royce, William Burman,

Anthony D. Harries, and Dick Menzies. 2010. “Treatment of Active Tuberculosis in HIV-Coinfected Patients: A Systematic Review and Meta-Analysis.” *Clinical Infectious Diseases* 50 (9):1288–99.

Klatt, Nichole R, Nicolas Chomont, Daniel C Douek, and Steven G Deeks. 2013. “Immune Activation and HIV Persistence: Implications for Curative Approaches to HIV Infection.” *Immunological Investigations* 254 (1):326–42.

Koch, Anastasia, Helen Cox, and Valerie Mizrahi. 2018. “Drug-Resistant Tuberculosis: Challenges and Opportunities for Diagnosis and Treatment.” *Current Opinion in Pharmacology* 42. Elsevier Ltd:7–15. https://doi.org/10.1016/j.coph.2018.05.013.

Koenker, Roger. 1994. “Confidence Intervals for Quantile Regression.” In *Asymptomatic Statistics: Proceedings of the 5th Prague Symposium*. http://www.econ.uiuc.edu/~roger/research/rqci/rqci.

Koh, Gavin C.K.W. K W, M. Fernanda Schreiber, Ruben Bautista, Rapeephan R. Maude, Susanna Dunachie, Direk Limmathurotsakul, Nicholas P.J. J Day, Gordon Dougan, and Sharon J. Peacock. 2013. “Host Responses to Melioidosis and Tuberculosis Are Both Dominated by Interferon-Mediated Signaling.” *PLoS ONE* 8 (1):1–13.

Korenromp, E L, F Scano, B G Williams, C Dye, and P Nunn. 2003. “Effects of Human Immunodeficiency Virus Infection on Recurrence of Tuberculosis after Receipt of Rifampin-Based Treatment: An Analytical Review.” *Clinical Infectious Diseases* 37 (1):101–12.

Koth, Laura L., Owen D. Solberg, Jeffrey C. Peng, Nirav R. Bhakta, Christine P. Nguyen, and Prescott G. Woodruff. 2011. “Sarcoidosis Blood Transcriptome Reflects Lung Inflammation and Overlaps with Tuberculosis.” *American Journal*

of Respiratory and Critical Care Medicine 184 (10):1153–63.

Lambert, Marie Laurence, Epco Hasker, Armand Van Deun, Dominique Roberfroid, Marleen Boelaert, and Patrick Van der Stuyft. 2003. “Recurrence in Tuberculosis: Relapse or Reinfection?” *Lancet Infectious Diseases* 3 (5):282–87.

Laux Da Costa, Lucas, Melaine Delcroix, Elis R. Dalla Costa, Isaías V. Prestes, Mariana Milano, Steve S. Francis, Gisela Unis, Denise R. Silva, Lee W. Riley, and Maria L.R. Rossetti. 2015. “A Real-Time PCR Signature to Discriminate between Tuberculosis and Other Pulmonary Diseases.” *Tuberculosis* 95 (4). Elsevier Ltd:421–25.

Lawn, S.D, A.D Harries, B.G Williams, R.E Chaisson, E. Losina, K.M De Cock, and R. Wood. 2011. “Antiretroviral Therapy and the Control of HIV-Associated Tuberculosis. Will ART Do It?” *International Journal of Tuberculosis and Lung Disease* 2011 (15):571–81.

Lawn, Stephen D., Katharina Kranzer, and Robin Wood. 2009. “Antiretroviral Therapy for Control of the HIV-Associated Tuberculosis Epidemic in Resource-Limited Settings.” *Clinics in Chest Medicine* 30 (4):685–99.

Lawn, Stephen D., Robin Wood, Kevin M. De Cock, Katharina Kranzer, James J. Lewis, and Gavin J. Churchyard. 2010. “Antiretrovirals and Isoniazid Preventive Therapy in the Prevention of HIV-Associated Tuberculosis in Settings with Limited Health-Care Resources.” *The Lancet Infectious Diseases* 10. Elsevier Ltd:489–98.

Lawn, Stephen D, Landon Myer, Linda-Gail Bekker, and Robin Wood. 2006. “Burden of Tuberculosis in an Antiretroviral Treatment Programme in Sub-Saharan Africa: Impact on Treatment Outcomes and Implications for Tuberculosis

Control.” *AIDS* 20 (12):1605–12.

Lawn, Stephen D, Landon Myer, David Edwards, Linda-Gail Bekker, and Robin Wood. 2009. “Short-Term and Long-Term Risk of Tuberculosis Associated with CD4 Cell Recovery during Antiretroviral Therapy in South Africa.” *AIDS* 23 (13):1717–25.

Lee, Shih-Wei, Lawrence Shih-Hsin Wu, Guan-Mau Huang, Kai-Yao Huang, Tzong-Yi Lee, and Julia Tzu-Ya Weng. 2016. “Gene Expression Profiling Identifies Candidate Biomarkers for Active and Latent Tuberculosis.” *BMC Bioinformatics* 17 (S1):S3.

Lehmann, C, J M Harper, D Taubert, P Hartmann, G Fatkenheuer, N Jung, J van Lunzen, H J Stellbrink, R C Gallo, and F Romerio. 2008. “Increased Interferon Alpha Expression in Circulating Plasmacytoid Dendritic Cells of HIV-1-Infected Patients.” *J Acquir Immune Defic Syndr* 48 (5):522–30.

Lillebaek, Troels, Asger Dirksen, Inga Baess, Benedicte Strunge, Vibeke Ø Thomsen, and Ase B Andersen. 2002. “Molecular Evidence of Endogenous Reactivation of Mycobacterium Tuberculosis after 33 Years of Latent Infection.” *The Journal of Infectious Diseases* 185 (3):401–4.

Lin, Philana Ling, Pauline Maiello, Hannah P. Gideon, M. Teresa Coleman, Anthony M. Cadena, Mark A. Rodgers, Robert Gregg, et al. 2016. “PET CT Identifies Reactivation Risk in Cynomolgus Macaques with Latent M. Tuberculosis.” *PLoS Pathogens* 12 (7):1–17.

Lin, Philana Ling, Tara Rutledge, Angela M. Green, Matthew Bigbee, Carl Fuhrman, Edwin Klein, and JoAnne L. Flynn. 2012. “CD4 T Cell Depletion Exacerbates Acute Mycobacterium Tuberculosis While Reactivation of Latent Infection Is Dependent on Severity of Tissue Depletion in Cynomolgus Macaques.” *AIDS*

Research and Human Retroviruses 28 (12):1693–1702.

Lobue, Philip, and Dick Menzies. 2010. “Treatment of Latent Tuberculosis Infection: An Update.” *Respirology* 15 (4):603–22.

Lu, Chanyi, Jing Wu, Honghai Wang, Sen Wang, Ni Diao, Feifei Wang, Yan Gao, et al. 2011. “Novel Biomarkers Distinguishing Active Tuberculosis from Latent Infection Identified by Gene Expression Profile of Peripheral Blood Mononuclear Cells.” *PLoS ONE* 6 (8):1–10.

Maertzdorf, J., D. Repsilber, S. K. Parida, K. Stanley, T. Roberts, G. Black, G. Walzl, and S. H E Kaufmann. 2011. “Human Gene Expression Profiles of Susceptibility and Resistance in Tuberculosis.” *Genes and Immunity* 12 (1). Nature Publishing Group:15–22.

Maertzdorf, J., J. Weiner, H.-J. Mollenkopf, T. Network, T. Bauer, A. Prasse, J. Muller-Quernheim, et al. 2012. “Common Patterns and Disease-Related Signatures in Tuberculosis and Sarcoidosis.” *Proceedings of the National Academy of Sciences* 109 (20):7853–58.

Maertzdorf, Jeroen, Stefan H E Kaufmann, and January Weiner. 2015. “Toward a Unified Biosignature for Tuberculosis.” *Cold Spring Harbor Perspective in Medicine* 4 (a018531).

Maertzdorf, Jeroen, Gayle McEwen, January Weiner, Song Tian, Eric Lader, Ulrich Schriek, Harriet Mayanja-kizza, et al. 2015. “Concise Gene Signature for Point-of-Care Classification of Tuberculosis.” *EMBO Molecular Medicine* 8 (2):86–95.

Maertzdorf, Jeroen, Martin Ota, Dirk Repsilber, Hans J. Mollenkopf, January Weiner, Philip C. Hill, and Stefan H.E. Kaufmann. 2011. “Functional Correlations of Pathogenesis-Driven Gene Expression Signatures in Tuberculosis.” *PLoS ONE* 6 (10):1–8.

Maharaj, B., T N Gengiah, N. Yende-Zuma, S. Gengiah, A Naidoo, and K. Naidoo. 2017. “Implementing Isoniazid Preventive Therapy in a Tuberculosis Treatment-Experienced Cohort on ART.” *The International Journal of Tuberculosis and Lung Disease* 21 (5):537–43.

Maher, Dermot, Anthony Harries, and Haileyesus Getahun. 2005. “Tuberculosis and HIV Interaction in Sub-Saharan Africa: Impact on Patients and Programmes; Implications for Policies.” *Tropical Medicine and International Health* 10 (8):734–42.

Mahomed, Hassan, Rodney Ehrlich, Tony Hawkridge, Mark Hatherill, Lawrence Geiter, Fazlin Kafaar, Deborah Ann Abrahams, et al. 2013. “TB Incidence in an Adolescent Cohort in South Africa.” *PLoS ONE* 8 (3).

Malherbe, Stephanus T, Shubhada Shenai, Katharina Ronacher, Andre G Loxton, Gregory Dolganov, Magdalena Kriel, Tran Van, et al. 2016. “Persisting Positron Emission Tomography Lesion Activity and Mycobacterium Tuberculosis mRNA after Tuberculosis Cure.” *Nature Medicine* 22 (10):1094–1100.

Mallory, K F, G J Churchyard, I Kleinschmidt, K M De Cock, and E L Corbett. 2000. “The Impact of HIV Infection on Recurrence of Tuberculosis in South African Gold Miners.” *International Journal of Tuberculosis and Lung Disease* 4 (5):455–62.

Manca, C., L. Tsenova, A. Bergtold, S. Freeman, M. Tovey, J. M. Musser, C. E. Barry, V. H. Freedman, and G. Kaplan. 2001. “Virulence of a Mycobacterium Tuberculosis Clinical Isolate in Mice Is Determined by Failure to Induce Th1 Type Immunity and Is Associated with Induction of IFN- A/b.” *Proceedings of the National Academy of Sciences* 98 (10):5752–57.

Manca, Claudia, Liana Tsenova, Sherry Freeman, Amy K. Barczak, Michael Tovey,

Peter J. Murray, Clifton Barry, and Gilla Kaplan. 2005. “Hypervirulent M. Tuberculosis W/Beijing Strains Upregulate Type I IFNs and Increase Expression of Negative Regulators of the Jak-Stat Pathway.” *Journal of Interferon & Cytokine Research* 25 (11):694–701.

Marcy, Olivier, Didier Laureillard, Yoann Madec, Sarin Chan, Charles Mayaud, Laurence Borand, Narom Prak, et al. 2014. “Causes and Determinants of Mortality in HIV-Infected Adults with Tuberculosis: An Analysis from the CAMELIA ANRS 1295-CIPRA KH001 Randomized Trial.” *Clinical Infectious Diseases* 59 (3):435–45.

Martinson, Neil A., Grace L. Barnes, Lawrence H. Moulton, Reginah Msandiwa, Harry Hausler, Malathi Ram, James A McIntyre, Glenda E. Gray, and Richard E. Chaisson. 2011. “New Regimens to Prevent Tuberculosis in Adults with HIV Infection.” *New England Journal of Medicine* 365 (1):11–10.

Marx, Florian M., Rory Dunbar, Donald A. Enarson, Brian G. Williams, Robin M. Warren, Gian D. Van Der Spuy, Paul D. Van Helden, and Nulda Beyers. 2014. “The Temporal Dynamics of Relapse and Reinfection Tuberculosis after Successful Treatment: A Retrospective Cohort Study.” *Clinical Infectious Diseases* 58 (12):1676–83.

Matteelli, Alberto, Giorgia Sulis, Sussana Capone, Lia D’Ambrosio, Giovanni Battista Migliori, and Haileyesus Getahun. 2017. “Tuberculosis Elimination and the Challenge of Latent Tuberculosis.” *La Presse Medicale* 46 (2). Elsevier Masson SAS:e13–21.

McNab, Finlay, Katrin Mayer-Barber, Alan Sher, Andreas Wack, Anne O’Garra, and Anne O Garra. 2015. “Type I Interferons in Infectious Disease.” *Nature Reviews Immunology* 15 (2). Nature Publishing Group:87–103.

McNab, Finlay W., John Ewbank, Ashleigh Howes, Lucia Moreira-Teixeira, Anna Martirosyan, Nico Ghilardi, Margarida Saraiva, and Anne O'Garra. 2014. "Type I IFN Induces IL-10 Production in an IL-27–Independent Manner and Blocks Responsiveness to IFN-γ for Production of IL-12 and Bacterial Killing in Mycobacterium Tuberculosis –Infected Macrophages." *The Journal of Immunology* 193 (7):3600–3612.

Merle, Corinne S., Katherine Fielding, Omou Bah Sow, Martin Gninafon, Mame B. Lo, Thuli Mthiyane, Joseph Odhiambo, et al. 2014. "A Four-Month Gatifloxacin-Containing Regimen for Treating Tuberculosis." *New England Journal of Medicine* 371 (17):1588–98.

Middelkoop, Keren, Linda Gail Bekker, Landon Myer, Andrew Whitelaw, Alison Grant, Gilla Kaplan, James McIntyre, and Robin Wood. 2010. "Antiretroviral Program Associated with Reduction in Untreated Prevalent Tuberculosis in a South African Township." *American Journal of Respiratory and Critical Care Medicine* 182 (8):1080–85.

Mistry, Rohit, Jacqueline M Cliff, Christopher L Clayton, Nulda Beyers, Yasmin S Mohamed, Paul a Wilson, Hazel M Dockrell, et al. 2007. "Gene-Expression Patterns in Whole Blood Identify Subjects at Risk for Recurrent Tuberculosis." *The Journal of Infectious Diseases* 195 (3):357–65.

Mitchison, D A. 1993. "Assessment of New Sterilizing Drugs for Treating Pulmonary Tuberculosis by Culture at 2 Months." *American Journal of Respiratory and Critical Care Medicine* 147 (4):1062–63.

Moreira-Teixeira, Lúcia, Jeremy Sousa, Finlay W. McNab, Egídio Torrado, Filipa Cardoso, Henrique Machado, Flávia Castro, et al. 2016. "Type I IFN Inhibits Alternative Macrophage Activation during Mycobacterium Tuberculosis Infection

and Leads to Enhanced Protection in the Absence of IFN-γ Signaling." *The Journal of Immunology* 197 (12):4714–26.

Mwinga, A, M Hosp, P Godfrey-Faussett, M Quigley, P Mwaba, B N Mugala, O Nyirenda, et al. 1998. "Twice Weekly Tuberculosis Preventive Therapy in HIV Infection in Zambia." *AIDS* 12:2447–57.

Naranbhai, Vivek, Helen A. Fletcher, Rachel Tanner, Matthew K. O'Shea, Helen McShane, Benjamin P. Fairfax, Julian C. Knight, and Adrian V.S. Hill. 2015. "Distinct Transcriptional and Anti-Mycobacterial Profiles of Peripheral Blood Monocytes Dependent on the Ratio of Monocytes: Lymphocytes." *EBioMedicine* 2 (11). The Authors:1619–26.

Naranbhai, Vivek, Adrian V S Hill, Salim S Abdool Karim, Kogieleum Naidoo, Quarraisha Abdool Karim, George M Warimwe, Helen McShane, and Helen Fletcher. 2013. "Ratio of Monocytes to Lymphocytes in Peripheral Blood Identifies Adults at Risk of Incident Tuberculosis Among HIV-Infected Adults Initiating Antiretroviral Therapy." *The Journal of Infectious Diseases* 209 (4):1–10. https://doi.org/10.1093/infdis/jit494.

Naranbhai, Vivek, Soyeon Kim, Helen Fletcher, Mark F Cotton, Avy Violari, Charles Mitchell, Sharon Nachman, et al. 2014. "The Association between the Ratio of Monocytes:lymphocytes at Age 3 Months and Risk of Tuberculosis (TB) in the First Two Years of Life." *BMC Medicine* 12 (120).

Ndzi, Elvis Ndukong, Celine Nguefeu Nkenfou, Luc Christian Gwom, Nadine Fainguem, Joseph Fokam, and Yone Pefura. 2016. "The Pros and Cons of the QuantiFERON Test for the Diagnosis of Tuberculosis, Prediction of Disease Progression, and Treatment Monitoring." *International Journal of Mycobacteriology* 5. Asian African Society for Mycobacteriology:0–7.

Nicholas, Sarala, Kalpana Sabapathy, Cecilia Ferreyra, Francis Varaine, Mar Pujades-Rodríguez, and Mar Pujades-rodr. 2011. “Incidence of Tuberculosis in HIV-Infected Patients before and after Starting Combined Antiretroviral Therapy in 8 Sub-Saharan African HIV Programs.” *Journal of Acquired Immune Deficiency Syndromes* 57 (4):311–18.

Nicol, Mark P., Lesley Workman, Washiefa Isaacs, Jacinta Munro, Faye Black, Brian Eley, Catharina C. Boehme, Widaad Zemanay, and Heather J. Zar. 2011. “Accuracy of the Xpert MTB/RIF Test for the Diagnosis of Pulmonary Tuberculosis in Children Admitted to Hospital in Cape Town, South Africa: A Descriptive Study.” *The Lancet Infectious Diseases* 11 (11):819–24.

O’Garra, Anne, Paul S Redford, Finlay W Mcnab, Chloe I Bloom, Robert J Wilkinson, and Matthew P R Berry. 2013. “The Immune Response in Tuberculosis.” *Annual Review of Immunology* 31:475–527.

Onozaki, Ikushi, Irwin Law, Charalambos Sismanidis, Matteo Zignol, Philippe Glaziou, and Katherine Floyd. 2015. “National Tuberculosis Prevalence Surveys in Asia, 1990-2012: An Overview of Results and Lessons Learned.” *Tropical Medicine and International Health* 20 (9):1128–45.

Opie, E, and J Aronson. 1927. “Tubercle Bacilli in Latent Tuberculosis Lesions and Lung Tissue without Tuberculosis Lesions.” *Archives of Pathology and Laboratoy Medicine*.

Ottenhoff, Tom H.M. M, Jerrold J. Ellner, and Stefan H.E. E Kaufmann. 2012. “Ten Challenges for TB Biomarkers.” *Tuberculosis* 92. Elsevier Ltd:S17–20.

Ottenhoff, Tom H M, Ranjeeta Hari Dass, Ninghan Yang, Mingzi M. Zhang, Hazel E E Wong, Edhyana Sahiratmadja, Chiea Chuen Khor, et al. 2012. “Genome-Wide Expression Profiling Identifies Type 1 Interferon Response Pathways in Active

Tuberculosis." *PLoS ONE* 7 (9).

Pai, Madhukar, Marcel A. Behr, David Dowdy, Keertan Dheda, Maziar Divangahi, Catharina C. Boehme, Ann Ginsberg, et al. 2016. "Tuberculosis." *Nature Reviews Disease Primers* 2:16077.

Parida, Shreemanta K., and Stefan H.E. Kaufmann. 2010. "The Quest for Biomarkers in Tuberculosis." *Drug Discovery Today* 15 (3–4):148–57.

Penn-Nicholson, A, T J Scriba, M Hatherill, R G White, and T Sumner. 2016. "A Novel Blood Test for Tuberculosis Prevention and Treatment." *South African Medical Journal* 107 (1):4.

Pocernich, Matt. 2015. "Package 'Verification.'" CRAN.R-Project.org. 2015. https://cran.r-project.org/web/packages/verification.

Psomas, Christina, Mehwish Younas, Christelle Reynes, Renaud Cezar, Pierre Portalès, Edouard Tuaillon, Adeline Guigues, et al. 2016. "One of the Immune Activation Profiles Observed in HIV-1-Infected Adults with Suppressed Viremia Is Linked to Metabolic Syndrome: The ACTIVIH Study." *EBioMedicine* 10. Elsevier B.V.:318–22.

Quigley, Maria A., Alwyn Mwinga, Maria Hosp, Ida Lisse, Dietmar Fuchs, John D. H. Porter, and Peter Godfrey-Faussett. 2001. "Long-Term Effect of Preventive Therapy for Tuberculosis in a Cohort of HIV-Infected Zambian Adults." *Aids* 15 (2):215–22.

Rangaka, Molebogeng X., Solange C. Cavalcante, Ben J. Marais, Sok Thim, Neil A. Martinson, Soumya Swaminathan, and Richard E. Chaisson. 2015. "Controlling the Seedbeds of Tuberculosis: Diagnosis and Treatment of Tuberculosis Infection." *The Lancet* 386 (10010). Elsevier Ltd:2344–53.

Rangaka, Molebogeng X., Robert J. Wilkinson, Judith R. Glynn, Andrew Boulle,

Gilles Van Cutsem, Rene Goliath, Shaheed Mathee, and Gary Maartens. 2012. "Effect of Antiretroviral Therapy on the Diagnostic Accuracy of Symptom Screening for Intensified Tuberculosis Case Finding in a South African HIV Clinic." *Clinical Infectious Diseases* 55 (12):1698–1706.

Rangaka, Molebogeng X, Robert J Wilkinson, Andrew Boulle, Judith R Glynn, Katherine Fielding, Gilles Van Cutsem, Katalin A Wilkinson, et al. 2014. "Isoniazid plus Antiretroviral Therapy to Prevent Tuberculosis : A Randomised Double-Blind Placebo-Controlled Trial." *Lancet* 384 (9944):682–90.

Rao, Martin, Davide Valentini, Thomas Poiret, Ernest Dodoo, Shreemanta Parida, Alimuddin Zumla, Sussana Brighenti, and Markus Maeurer. 2015. "B in TB: B Cells Are Mediators of Clinically Relevant Immune Responses in Tuberculosis." *Clinical Infectious Diseases* 61 (Suppl 3):S225–34.

Raviglione, Mario C, and Ian M Smith. 2007. "XDR Tuberculosis--Implications for Global Public Health." *The New England Journal of Medicine* 356 (7):656–59.

Rie, Annelies Van, Robin Warren, Madeleine Richardson, Thomas C. Victor, Robert P. Gie, Donald A. Enarson, Nulda Beyers, and Paul D. van Helden. 1999. "Exogenous Reinfection as a Cause of Recurrent Tuberculosis after Curative Treatment." *The New England Journal of Medicine* 341 (16).

Rie, Annelies Van, Daniel Westreich, and Ian M Sanne. 2011. "Tuberculosis in Patients Receiving Antiretroviral Treatment: Incidence, Risk Factors, and Prevention Strategies." *J Acquir Immune Defic Syndr* 56 (4):349–55.

Riou, Catherine, Rubina Bunjun, Tracey L. Müller, Agano Kiravu, Zekarias Ginbot, Tolu Oni, Rene Goliath, Robert J. Wilkinson, and Wendy A. Burgers. 2016. "Selective Reduction of IFN-γ Single Positive Mycobacteria-Specific CD4+ T Cells in HIV-1 Infected Individuals with Latent Tuberculosis Infection."

Tuberculosis 101. Elsevier Ltd:25–30.

Robin, Xavier, Natacha Turck, Alexandre Hainard, Natalia Tiberti, Frédérique Lisacek, Jean-Charles Sanchez, and Markus Müller. 2011. “pROC: An Open-Source Package for R and S+ to Analyze and Compare ROC Curves.” *BMC Bioinformatics* 12 (1):77.

Roe, Jennifer K, Niclas Thomas, Eliza Gil, Katharine Best, Evdokia Tsaliki, Stephen Morris-Jones, Sian Stafford, et al. 2016. “Blood Transcriptomic Diagnosis of Pulmonary and Extrapulmonary Tuberculosis.” *JCI Insight* 1 (16):1–14.

Rustomjee, R, C Lienhardt, T Kanyok, G R Davies, J Levin, T Mthiyane, C Reddy, et al. 2008. “A Phase II Study of the Sterilizing Activities of Ofloxacin, Gatifloxacin and Moxifloxacin in Pulmonary Tuberculosis.” *Int J Tuberc Lung Dis* 12 (October 2007):128–138(111).

Salgame, Padmini, Carolina Geadas, Lauren Collins, Edward Jones-López, and Jerrold J. Ellner. 2015. “Latent Tuberculosis Infection - Revisiting and Revising Concepts.” *Tuberculosis* 95 (4):373–84.

Samandari, Taraz, Tefera B. Agizew, Samba Nyirenda, Zegabriel Tedla, Thabisa Sibanda, Barudi Mosimaneotsile, Oaitse I. Motsamai, Nong Shang, Charles E. Rose, and James Shepherd. 2015. “Tuberculosis Incidence after 36 Months’ Isoniazid Prophylaxis in HIV-Infected Adults in Botswana: A Posttrial Observational Analysis.” *AIDS* 29 (3):351–59.

Samandari, Taraz, Tefera B. Agizew, Samba Nyirenda, Zegabriel Tedla, Thabisa Sibanda, Nong Shang, Barudi Mosimaneotsile, et al. 2011. “6-Month versus 36-Month Isoniazid Preventive Treatment for Tuberculosis in Adults with HIV Infection in Botswana: A Randomised, Double-Blind, Placebo-Controlled Trial.” *The Lancet* 377 (9777):1588–98.

Sandler, Netanya G, Steven E Bosinger, Jacob D Estes, Richard T R Zhu, Gregory K Tharp, Eli Boritz, Doron Levin, et al. 2014. “Type I Interferon Responses in Rhesus Macaques Prevent SIV Infection and Slow Disease Progression.” *Nature* 511 (7511):601–5.

Satproedprai, N., N. Wichukchinda, S. Suphankong, W. Inunchot, T. Kuntima, S. Kumpeerasart, S. Wattanapokayakit, et al. 2015. “Diagnostic Value of Blood Gene Expression Signatures in Active Tuberculosis in Thais: A Pilot Study.” *Genes and Immunity* 16 (4). Nature Publishing Group:253–60.

Schluger, Neil W. 2017. “Improving the Diagnosis of Latent TB Infection: Tools for TB Elimination?” *Chest* 151 (6):1207–8.

Scriba, Thomas J T.J., Adam Penn-Nicholson, Smitha Shankar, Tom Hraha, Ethan G E.G. Thompson, David Sterling, Elisa Nemes, et al. 2017. “Sequential Inflammatory Processes Define Human Progression from M. Tuberculosis Infection to Tuberculosis Disease.” *PLoS Pathogens* 13 (11):1–24.

Shenai, Shubhada, Katharina Ronacher, Stefanus Malherbe, Kim Stanley, Magdalena Kriel, Jill Winter, Thomas Peppard, et al. 2016. “Bacterial Loads Measured by the Xpert MTB/ RIF Assay as Markers of Culture Conversion and Bacteriological Cure in Pulmonary TB.” *PLoS ONE* 11 (8):1–13.

Singhania, Akul, Raman Verma, Christine M Graham, Jo Lee, Tran Trang, Matthew Richarsdon, Patrick Lecine, et al. 2017. “A Modular Transcriptional Signature Identifies Phenotypic Heterogeneity of Human Tuberculosis Infection.” *bioRxiv*, 1–32.

Sloot, Rosa, Maarten F.Schim Van Der Loeff, Peter M. Kouw, and Martien W. Borgdorff. 2014. “Risk of Tuberculosis after Recent Exposure: A 10-Year Follow-up Study of Contacts in Amsterdam.” *American Journal of Respiratory and*

Critical Care Medicine 190 (9):1044–52.

Sloot, Rosa, Maarten F. Schim van der Loeff, Erik W. van Zwet, Mariëlle C. Haks, Sytze T. Keizer, Maarten Scholing, Tom H.M. Ottenhoff, Martien W. Borgdorff, and Simone a. Joosten. 2015. "Biomarkers Can Identify Pulmonary Tuberculosis in HIV-Infected Drug Users Months Prior to Clinical Diagnosis." *EBioMedicine* 2 (2). Elsevier B.V.:172–79.

Smieja, M, C Marchetti, D Cook, and FM Smaill. 1999. "Isoniazid for Preventing Tuberculosis in Non-HIV Infected Persons (Review) Smieja M , Marchetti C , Cook D , Smaill FM." *Cochrane Database of Systematic Reviews*, no. 1.

Somsouk, Ma, Jacob D Estes, Claire Deleage, Richard M Dunham, Rebecca Albright, John M Inadomi, Jeffrey N Martin, Steven G Deeks, Joseph M Mccune, and Peter W Hunt. 2015. "Gut Epithelial Barrier and Systemic Inflammation during Chronic HIV Infection." *Aids* 29 (1):43–51.

Sonnenberg, Pam, Judith R Glynn, Katherine Fielding, Jill Murray, Peter Godfrey-Faussett, and Stuart Shearer. 2005. "How Soon after Infection with HIV Does the Risk of Tuberculosis Start to Increase? A Retrospective Cohort Study in South African Gold Miners." *Journal of Infectious Diseases* 191 (2):150–58.

Sonnenberg, Pamela, Jill Murray, Judith R. Glynn, Stuart Shearer, Bupe Kambashi, and Peter Godfrey-Faussett. 2001. "HIV-1 and Recurrence, Relapse, and Reinfection of Tuberculosis after Cure: A Cohort Study in South African Mineworkers." *Lancet* 358 (9294):1687–93.

Stagg, Helen R., Dominik Zenner, Ross J. Harris, Laura Muñoz, Marc C. Lipman, and Ibrahim Abubakar. 2014. "Treatment of Latent Tuberculosis Infection a Network Meta-Analysis." *Annals of Internal Medicine* 161 (6):419–28.

Stanley, Sarah A, James E Johndrow, Paolo Manzanillo, and Jeffery S Cox. 2007.

"The Type I IFN Response to Infection with Mycobacterium Tuberculosis Requires ESX-1-Mediated Secretion and Contributes to Pathogenesis." *Journal of Immunology* 178 (5):3143–52.

Sterling, Timothy R., Nigel A. Scott, Jose M. Miro, Guilherme Calvet, Alberto La Rosa, Rosa Infante, Michael P. Chen, et al. 2016. "Three Months of Weekly Rifapentine and Isoniazid for Treatment of Mycobacterium Tuberculosis Infection in HIV-Coinfected Persons." *Aids* 30 (10):1607–15.

Sterling, Timothy R., M.Elsa Villarino, Andrey S. Borisov, Nong Shang, Fred Gordin, Erin Bliven-Sizemore, Judith HAckman, et al. 2011. "Three Months of Rifapentine and Isoniazid for Latent Tuberculosis Infection." *The New England Journal of Medicine* 365 (23):215–24.

Stifter, Sebastian A., and Carl G. Feng. 2015. "Interfering with Immunity: Detrimental Role of Type I IFNs during Infection." *The Journal of Immunology* 194:2455–65.

Suthar, Amitabh B., Reuben M. Granich, Masaya Kato, Sabin Nsanzimana, Julio S.G. Montaner, and Brian G. Williams. 2015. "Programmatic Implications of Acute and Early HIV Infection." *Journal of Infectious Diseases* 212 (9):1351–60.

Suthar, Amitabh B., Stephen D. Lawn, Julia del Amo, Haileyesus Getahun, Christopher Dye, Delphine Sculier, Timothy R. Sterling, et al. 2012. "Antiretroviral Therapy for Prevention of Tuberculosis in Adults with HIV: A Systematic Review and Meta-Analysis." *PLoS Medicine* 9 (7).

Swaminathan, Soumya, Pradeep Aravindan Menon, Narendran Gopalan, Venkatesan Perumal, Ramesh Kumar Santhanakrishnan, Ranjani Ramachandran, Ponnuraja Chinnaiyan, et al. 2012. "Efficacy of a Six-Month versus a 36-Month Regimen for Prevention of Tuberculosis in HIV-Infected Persons in India: A Randomized Clinical Trial." *PLoS ONE* 7 (12):1–9.

Sweeney, Timothy E., Lindsay Braviak, Cristina M. Tato, and Purvesh Khatri. 2016. "Genome-Wide Expression for Diagnosis of Pulmonary Tuberculosis: A Multicohort Analysis." *The Lancet Respiratory Medicine* 4 (3). Elsevier Ltd:213–24.

Teles, Rosane M B, Thomas G Graeber, Stephan R Krutzik, Dennis Montoya, Mirjam Schenk, Delphine J Lee, Evangelia Komisopoulou, et al. 2013. "Type I Interferon Suppresses Type II Interferon-Trigerred Human Anti-Mycobacterial Responses." *Science* 339 (6126):1448–53.

Thillai, M, K Pollock, M Pareek, and A Lalvani. 2014. "Interferon-Gamma Release Assays for Tuberculosis: Current and Future Applications." *Expert Review of Respiratory Medicine* 8 (1). Informa UK, Ltd.:67–78.

Thompson, Ethan G., Ying Du, Stephanus T. Malherbe, Smitha Shankar, Jackie Braun, Joe Valvo, Katharina Ronacher, et al. 2017. "Host Blood RNA Signatures Predict the Outcome of Tuberculosis Treatment." *Tuberculosis* 107:48–58.

Tornheim, J.A., and K.E. Dooley. 2017. "Tuberculosis Associated with HIV Infection." *Microbiology Spectrum* 5 (1).

Tuberculosis Research Center. 2002. "Shortening Short Course Chemotherapy: A Randomized Clinical Trial for Treatment of Smear Positive Pulmonary Tuberculosis with Regimens Using Ofloxacin in the Intensive Phase." *Indian Journal of Tuberculosis* 49:27–38.

Uplekar, Mukund, Diana Weil, Knut Lonnroth, Ernesto Jaramillo, Christian Lienhardt, Hannah Monica Dias, Dennis Falzon, et al. 2015. "WHO's New End TB Strategy." *The Lancet* 385 (9979):1799–1801.

Veazey, Ronald S, and Andrew A Lackner. 2005. "HIV Swiftly Guts the Immune System." *Nature Medicine* 11 (5):469–70.

Verver, Suzanne, Robin M. Warren, Nulda Beyers, Madalene Richardson, Gian D. Van Der Spuy, Martien W. Borgdorff, Donald A. Enarson, Marcel A. Behr, and Paul D. Van Helden. 2005. “Rate of Reinfection Tuberculosis after Successful Treatment Is Higher than Rate of New Tuberculosis.” *American Journal of Respiratory and Critical Care Medicine* 171 (12):1430–35.

Volmink, Jimmy, and Sara Woldehanna. 2004. “Treatment of Latent Tuberculosis Infection in HIV Infected Persons (Review).” *The Cochrane Collaboration*, no. 1:1–3.

Vynnycky, Emilia, Tom Sumner, Katherine L. Fielding, James J. Lewis, Andrew P. Cox, Richard J. Hayes, Elizabeth L. Corbett, Gavin J. Churchyard, Alison D. Grant, and Richard G. White. 2015. “Tuberculosis Control in South African Gold Mines: Mathematical Modeling of a Trial of Community-Wide Isoniazid Preventive Therapy.” *American Journal of Epidemiology* 181 (8):619–32.

Walker, N F, G Meintjes, and R J Wilkinson. 2013. “HIV-1 and the Immune Response to TB.” *Future Virology* 8 (1):57–80.

Wallgren, Arvid. 1948. “The Time-Table of Tuberculosis.” *Tubercle*, 245–51.

Wallis, Robert S., T. Mark Doherty, Phillip Onyebujoh, Mahnaz Vahedi, Hannu Laang, Ole Olesen, Shreemanta Parida, and Alimuddin Zumla. 2009. “Biomarkers for Tuberculosis Disease Activity, Cure, and Relapse.” *The Lancet Infectious Diseases* 9 (3):162–72.

Wallis, Robert S., Mark D. Perkins, Manijeh Phillips, Moses Joloba, Alice Namale, John L. Johnson, Christopher Christoper C. Whalen, et al. 2000. “Predicting the Outcome of Therapy for Pulmonary Tuberculosis.” *American Journal of Respiratory and Critical Care Medicine* 161 (401):1076–80.

Walter, Nicholas D, Mikaela A. Miller, Joshua Vasquez, Marc Weiner, Adam

Chapman, Melissa Engle, Michael Higgins, et al. 2016. “Blood Transcriptional Biomarkers for Active Tuberculosis among Patients in the United States : A Case-Control Study with Systematic Evaluation Cross-Classifier.” *Journal of Clinical Microbiology* 54 (2):274–82.

Walzl, Gerhard, C Haks, Simone A. Joosten, Katharina Ronacher, Tom H.M. M Ottenhoff, Marielle C. Haks, Simone A. Joosten, Leanie Kleynhans, Katharina Ronacher, and Tom H.M. M Ottenhoff. 2014. “Clinical Immunology and Multiplex Biomarkers of Human Tuberculosis.” *Cold Spring Harbor Perspective in Medicine*, 1–14.

Whalen, Christopher C., Peter Nsubuga, Alphonse Okwera, John L. Johnson, David L. Hom, Nelson L. Michael, Roy D. Mugerwa, and Jerrold J. Ellner. 2000. “Impact of Pulmonary Tuberculosis on Survival of HIV-Infected Adults: A Prospective Epidemiologic Study in Uganda.” *Aids* 14 (9):1219–28.

Wiens, Kirsten E., and Joel D. Ernst. 2016. “The Mechanism for Type I Interferon Induction by Mycobacterium Tuberculosis Is Bacterial Strain-Dependent.” *PLoS Pathogens* 12 (8):1–20.

Wiker, Harald G., Tehmina Mustafa, Gunnar A. Bjune, and Morten Harboe. 2010. “Evidence for Waning of Latency in a Cohort Study of Tuberculosis.” *BMC Infectious Diseases* 10 (37).

Williams, B. G., R. Granich, K. M. De Cock, P. Glaziou, A. Sharma, and C. Dye. 2010. “Antiretroviral Therapy for Tuberculosis Control in Nine African Countries.” *Proceedings of the National Academy of Sciences* 107 (45).

Wood, Robin, and Stephen D Lawn. 2011. “Antiretroviral Treatment as Prevention: Impact of the ‘Test and Treat’ Strategy on the Tuberculosis Epidemic.” *Current HIV Research* 9 (6):383–92.

Wood, Robin, Stephen D. Lawn, Judy Caldwell, Richard Kaplan, Keren Middelkoop, and Linda Gail Bekker. 2011. “Burden of New and Recurrent Tuberculosis in a Major South African City Stratified by Age and HIV-Status.” *PLoS ONE* 6 (10):1–9.

Wood, Robin, Keren Middelkoop, Landon Myer, Alison D. Grant, Andrew Whitelaw, Stephen D. Lawn, Gilla Kaplan, Robin Huebner, James McIntyre, and Linda Gail Bekker. 2007. “Undiagnosed Tuberculosis in a Community with High HIV Prevalence: Implications for Tuberculosis Control.” *American Journal of Respiratory and Critical Care Medicine* 175 (1):87–93.

World health Organization. 2015. “Global Strategy and Targets for Tuberculosis Prevention, Care and Control after 2015.” *World Health Organisation Document*, 1–2.

World Health Organization. 2010. “Treatment of Tuberculosis.” *World Health Organization Document*.

———. 2011. “Guidelines for Intensified Tuberculosis Case-Finding and Isoniazid Preventive Therapy for People Living with HIV in Resource- Constrained Settings.” *World Health Organization Document* 53 (9):4–5.

———. 2012a. “Antiretroviral Treatment as Prevention (TASP) of HIV and TB.” *WHO Programmatic Update*, no. june:1.

———. 2012b. “WHO Policy on Collaborative TB/HIV Activities. Guidelines for National Programmes and Other Stakeholders.” *World Health Organization Document*.

———. 2014. “High-Priority Target Product Profiles for New Tuberculosis Diagnostics: Report of a Consensus Meeting.” *World Health Organization Meeting Report*.

———. 2015. “Guidelines on the Management of Latent Tuberculosis Infection.” *WHO Guideline Document*.

———. 2017. “Global Tuberculosis Report 2017.” *World Health Organization Document*, 1–262.

Young, Douglas B., Hannah P. Gideon, and Robert J. Wilkinson. 2009. “Eliminating Latent Tuberculosis.” *Trends in Microbiology* 17 (5):183–88.

Zak, Daniel E., Adam Penn-Nicholson, Thomas J. Scriba, Ethan Thompson, Sara Suliman, Lynn M. Amon, Hassan Mahomed, et al. 2016. “A Blood RNA Signature for Tuberculosis Disease Risk: A Prospective Cohort Study.” *The Lancet* 387. Elsevier Ltd:2312–22.

Zwerling, Alice, Colleen Hanrahan, and David W. Dowdy. 2016. “Ancient Disease, Modern Epidemiology: A Century of Progress in Understanding and Fighting Tuberculosis.” *American Journal of Epidemiology* 183 (5):407–14.

Tuberculosis 108 (2018) 124–126

Contents lists available at ScienceDirect

Tuberculosis

journal homepage: www.elsevier.com/locate/tube

Diagnostic performance of an optimized transcriptomic signature of risk of tuberculosis in cryopreserved peripheral blood mononuclear cells

ARTICLE INFO

Keywords:
Biomarker
Correlate of risk
Diagnostic
Whole blood
PBMC
mRNA

A staggering 23% of the global population is estimated to be infected with *Mycobacterium tuberculosis (M.tb)* [1], representing an enormous reservoir of individuals at risk of progressing to active tuberculosis (TB) disease. Identification of those at risk of TB disease would allow targeted preventive therapy, potentially curtailing transmission [2]. We recently identified and validated a 16-gene whole blood transcriptomic correlate of risk (CoR) signature that can predict active TB disease in *M.tb*-infected individuals more than a year before disease manifestations and diagnosis [3]. The transcriptomic signature is measured by microfluidic quantitative reverse transcription polymerase chain reaction (qRT-PCR), comprising a set of 57 primer-probes (47 primer-probes detecting 47 transcripts representing 16 interferon (IFN) response genes, and 10 primer-probes used as a reference for standardization). To improve high-throughput processing we sought to trim 9 primer-probes from the signature such that the 48 primer-probes can be conveniently accommodated in duplicate in a typical 96.96 Fluidigm gene expression platform.

We selected the transcripts for removal based on 3 criteria: (1) those that showed the least reproducibility in qRT-PCR assays, (2) those that yielded redundant signal to the overall signature, and (3) when transcripts were ranked by the number of pairings they form with other transcripts in the signature only those that form very few pairs were removed (thus contributing little to the signature; data not shown). Based on these criteria, we removed nine transcripts detected by 9 primer-probes, which represented five IFN response genes.

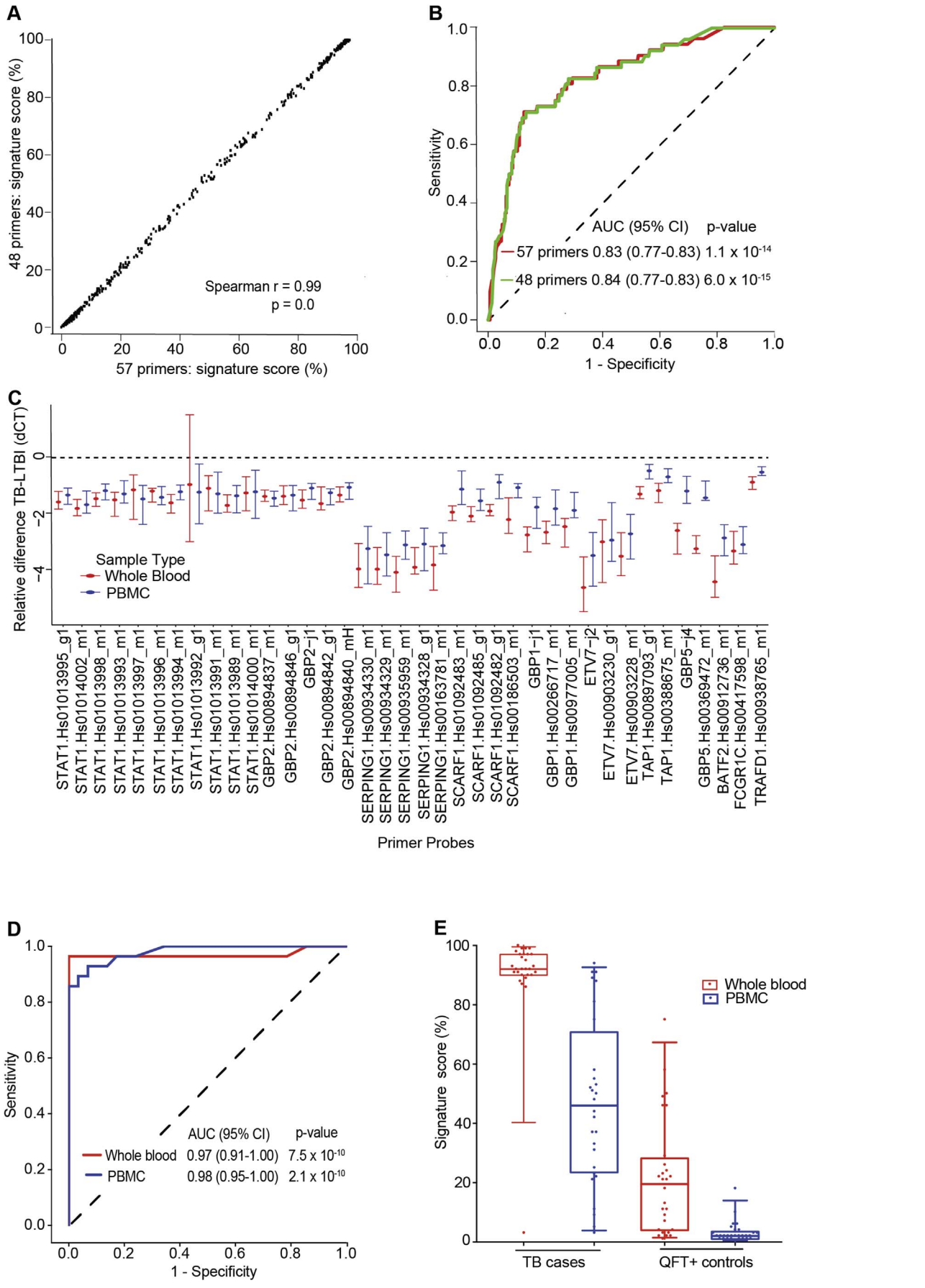
The reduced 48-primer-probe signature comprised 38 primer-probes representing 11 IFN response genes, and the 10 reference primer-probes. The removed primer-probes represented the genes, SEPT4, ANKRD22, APOL1, FCGR1A and GBP4. We compared the prognostic performance of the original, 57 primer-probe signature to the reduced, 48 primer-probe signature (Adolescent Cohort Study (ACS) 11-gene signature), measured by microfluidic qRT-PCR on samples collected from adolescent progressors and controls that comprised the original training and test sets from the ACS, in which the signature was discovered [3]. Ribonucleic acid (RNA) was isolated with the PAXgene™ blood RNA kit (QIAGEN); complimentary DNA (cDNA) was synthesised by reverse transcription and pre-amplified using PCR, as described [3]. Signature transcripts were quantified using TaqMan primer-probes on 96.96 Gene Expression Chips (Fluidigm) on a Biomark HD multiplex microfluidic qRT-PCR instrument. Signature scores were calculated as previously described [3]. Risk scores of the original signature and reduced ACS 11-gene signature were essentially identical (Spearman $r = 0.99$, $p = 0.0$ (Fig. 1A)). Predictive performance of TB risk, expressed as receiver operating characteristic (ROC) area under the curve (AUC) also did not differ (Fig. 1B). Thus, reduction of the CoR signature to 48 primer-probes resulted in equivalent performance.

Next, we sought to determine the performance of the 48 primer-probe whole blood ACS 11-gene signature when measured in peripheral blood mononuclear cell (PBMC) samples. Reliable measurement of the signature using RNA from PBMC would allow testing and/or further validation using biobanked samples from historical clinical studies. Because the signature score increases most proximal to the onset of TB disease, we evaluated diagnostic performance using RNA from PAXgene whole blood and cryopreserved PBMC samples (RNEasy mini kit, QIAGEN) collected in parallel from 30 HIV-uninfected adults with newly diagnosed active TB disease (sputum Xpert MTB/RIF-positive) and 30 healthy, *M.tb*-infected (QuantiFERON Gold In-tube-positive, QIAGEN) adult controls.

We first compared messenger RNA (mRNA) expression differences for each IFN response transcript in the reduced signature between TB cases and controls in whole blood and PBMC. The difference in median mRNA expression between TB cases and controls (delta cycle threshold (Ct)) was estimated and 95% confidence intervals (CI) of the median were obtained using the rank inversion method [4], with code from the R package QUANTREG. All 38 IFN response transcripts were expressed at higher levels in TB cases than in controls, whether measured in whole blood or PBMC (Fig. 1C). However, for transcripts including SCARF1, GBP2, GBP5, BATF2 and TRAFD1, the relative differences in transcript expression between TB cases and controls were greater in whole blood than in PBMC. This suggests a substantial contribution of cell subsets present in whole blood but absent in PBMC, such as granulocytes, to the disease-associated IFN response, as previously described [5].

Nevertheless, the ACS 11-gene signature classified samples from TB cases and healthy controls with equivalent accuracy (Fig. 1D). ROC-AUCs for whole blood and PBMC were indistinguishable at 0.97 (95% CI 0.91–1.00) and 0.98 (0.95–1.00), respectively. Risk signature scores were markedly lower in PBMC than in whole blood regardless of disease status, again reflecting the comparatively weaker IFN response transcriptomic signal in

https://doi.org/10.1016/j.tube.2017.11.001
Received 28 July 2017; Received in revised form 30 October 2017; Accepted 2 November 2017



(caption on next page)

Fig. 1. Performance of the reduced ACS 11-gene signature of TB risk as a classifier of TB disease from *M.tb*-infection in PBMC. (**A**) Correlation of the signature scores generated from the original 57 primer-probe (16 genes) and the reduced, 48 primer-probe (11-genes) qRT-PCR transcriptomic signatures of risk of TB disease in progressors and controls from the Adolescent Cohort Study (ACS). The Spearman correlation coefficient is shown. (**B**) Performance of the original and reduced CoR signatures in classifying progressor and control samples from the Adolescent Cohort Study. (**C**) Differences in transcript expression between TB cases (Xpert MTB/RIF-positive) and healthy controls (QuantiFERON-positive) when measured in whole blood or PBMC samples. Dots represent medians and the bars 95% CI for each transcript, identified by its TaqMan primer-probe set. (**D**) Receiver operating characteristic curves illustrating diagnostic performance of the 48-primer-probe ACS 11-gene signature in distinguishing individuals with active TB disease from healthy QuantiFERON-positive controls when measured in RNA from whole blood or PBMC samples. (**E**) CoR signature scores measured in whole blood or PBMC from active TB cases or healthy QuantiFERON-positive persons. Horizontal lines represent medians, the box represents the IQR and whiskers the range.

PBMC (Fig. 1E). This necessitates the use of independent, sample type-specific thresholds for classification of signature-positive or negative samples.

The comparable performance of the CoR signature in discriminating active TB cases from healthy controls in PBMC and whole blood demonstrates feasibility for applying this signature to ascertain risk of TB in historical studies that have prospectively collected cryopreserved PBMC.

In conclusion, the ACS 11-gene, 48-primer-probe transcriptomic signature has diagnostic utility for TB with identical accuracy as the original 16-gene, 57- primer-probe signature, and can be measured in cryopreserved PBMC.

Funding

This work was supported by the Strategic Health Innovation Partnerships (SHIP) Unit of the South African Medical Research Council with funds received from the South African Department of Science and Technology.

F.D. is supported by the Margaret McNamara educational grant for women in developing countries.

Conflicts of interest

EF, MH, NB, MR, MF, SM, FD and MvR declare no conflicts of interest.

APN, DEZ, EGT and TJS report pending patent of the gene signature.

DEZ and TJS report receiving grants from the South African Medical Research Council, National Institutes of Health and/or Bill and Melinda Gates Foundation related to the gene signature during course of study.

References

[1] Houben RMGJ, Dodd PJ. The global burden of latent tuberculosis infection: a Re-estimation using mathematical modelling. PLoS Med 2016;13:1–13. http://dx.doi.org/10.1371/journal.pmed.1002152.

[2] Penn-Nicholson A, Scriba TJ, Hatherill M, White RG, Sumner T. A novel blood test for tuberculosis prevention and treatment. South Afr Med J 2016;107:4. http://dx.doi.org/10.7196/SAMJ.2017.v107i1.12230.

[3] Zak DE, Penn-Nicholson A, Scriba TJ, Thompson E, Suliman S, Amon LM, et al. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. Lancet 2016;6736:1–11. http://dx.doi.org/10.1016/S0140-6736(15)01316-1.

[4] Koenker R. Confidence intervals for quantile regression. Asymptomatic stat. Proc. 5th prague symp. 1994.

[5] Berry MPR, Graham CM, McNab FW, Xu Z, Bloch SA, Oni T, et al. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. Nature 2010;466:973–7. http://dx.doi.org/10.1038/nature09247.

Fatoumatta Darboe[1], Stanley Kimbung Mbandi[1]
South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular Medicine and Division of Immunology, Department of Pathology, University of Cape Town, 7925, Cape Town, South Africa

Ethan G. Thompson
The Center for Infectious Disease Research, Seattle, 98109, WA, USA

Michelle Fisher
South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular Medicine and Division of Immunology, Department of Pathology, University of Cape Town, 7925, Cape Town, South Africa

Miguel Rodo
South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular Medicine and Division of Immunology, Department of Pathology, University of Cape Town, 7925, Cape Town, South Africa
Department of Statistical Sciences, University of Cape Town, Cape Town, 7925, South Africa

Michele van Rooyen, Elizabeth Filander, Nicole Bilek, Simbarashe Mabwe, Mark Hatherill
South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular Medicine and Division of Immunology, Department of Pathology, University of Cape Town, 7925, Cape Town, South Africa

Daniel E. Zak
The Center for Infectious Disease Research, Seattle, 98109, WA, USA

Adam Penn-Nicholson[2], Thomas J. Scriba[*,2]
South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular Medicine and Division of Immunology, Department of Pathology, University of Cape Town, 7925, Cape Town, South Africa
E-mail address: thomas.scriba@uct.ac.za

The SATVI Clinical Immunology Team

[*] Corresponding author.
[1] These authors contributed equally.
[2] Shared senior authors.
Sindile Matiwane, Lungisa Jaxanee NkantsuNoncedo Xoyana, Constance Schreuder, Janelle Botes, Hadn Africa, Lebohang Makhethe, Marcia Steyn.