Statistical Analysis Plan:

Evaluation of parsimonious host-blood tuberculosis transcriptomic signatures in HIV-infected and HIV-uninfected individuals

A sub-study of the CORTIS-01 and CORTIS-HR trials

Description: This document describes the statistical analysis plan to be performed on data obtained from the above-named sub-study of the CORTIS-01 and CORTIS-HR studies and is to be read in conjunction with the approved CORTIS-01 study protocol version 2.0 (01 August 2017) and SAP version 3.0 (24 October 2019), and CORTIS-HR study protocol version 1.0 (26th August 2016) and SAP version 1.0 (11 December 2019).

Authors: Simon C Mendelsohn[1], MBChB

Andrew Fiore-Gartland[2], Ph.D.

[1]South African Tuberculosis Vaccine Initiative, Department of Pathology and Institute of Infectious Diseases and Molecular Medicine, University of Cape Town

[2]Statistical Center for HIV/AIDS Research and Prevention, Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center

Date: 08 January 2020

Version: v 1.0

Document:	Statistical Analysis Plan: Evaluation of parsimonious host-blood tuberculosis transcriptomic signatures in HIV-infected and HIV-uninfected individuals. *A sub-study of the CORTIS-01 and CORTIS-HR trials*
SAP Version:	1.0
Version Date:	08 January 2020
Authorised By:	Prof Mark Hatherill, MBChB CORTIS-01 and CORTIS-HR Principal Investigator South African Tuberculosis Vaccine Initiative, University of Cape Town Sign: _____ Date: 09/JAN/2020
	Dr Andrew Fiore-Gartland, PhD CORTIS-01 and CORTIS-HR Lead Statistician Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center Sign: _____ Date: 13-Jan-2020

Table of Contents

1 LIST OF ABBREVIATIONS

AFB	Acid-fast bacilli
BMI	Body mass index
cDNA	Complementary DNA
CI	Confidence interval
CIR	Cumulative incidence ratio
COR	Correlate of risk
CRA	Clinical research associate
CRF	Case report form
eCRF	Electronic CRFs
GCP	Good clinical practice
H_0	Null hypothesis
HIV	Human immunodeficiency virus
IEC	Independent ethics committee
IGRA	Interferon gamma release assay
INH	Isoniazid
IPT	Isoniazid preventive therapy
LAM	Lipoarabinomannan
LTBI	Latent tuberculosis infection
MDR-TB	Multi-drug resistant tuberculosis
MGIT	Mycobacteria growth indicator tube
mRNA	Messenger RNA
Mtb	*Mycobacterium tuberculosis*
OR	Odds ratio
NNS	Number needed to screen (to detect one case)
NPV	Negative predictive value
PI	Principal investigator
PPV	Positive predictive value
QFT	QuantiFERON
RR	Relative risk
$RR_{COR}(15)$	Relative risk for TB disease over 15 months
SA	South Africa
SATVI	South African Tuberculosis Vaccine Initiative
TB	Tuberculosis
TCD	Triclinium Clinical Development
TST	Tuberculin skin test
WHO	World Health Organization

2 STATISTICAL ANALYSIS PLAN OVERVIEW

This statistical analysis plan (SAP) is for a sub-study of the CORTIS-01 and CORTIS-HR trails and is to be read in conjunction with the approved CORTIS-01 study protocol version 2.0 (01 August 2017) and SAP version 3.0 (24 October 2019), and CORTIS-HR study protocol version 1.0 (26th August 2016) and SAP version 1.0 (11 December 2019).

There is a need for earlier TB case identification, using novel non-sputum based diagnostics, linked to more effective preventive and curative strategies (World Health Organization, 2015). A blood-based triage test that allows targeted investigation for active and sub-clinical TB disease, including asymptomatic individuals at highest risk of progression from latency to disease, could shorten the time to TB treatment, or even prevent disease before symptoms emerge. The tuberculin skin test (TST) and interferon gamma release assay (IGRA) have poor specificity for incident TB disease in endemic populations, including HIV infected people (Auguste, *BMC Infect Dis*, 2017).

We have previously developed a highly specific predictive correlate of risk (COR) to identify healthy, HIV uninfected, South African adults at high risk of active TB disease (Zak, *Lancet*, 2016). This validated COR, based on mRNA expression signatures in blood, prospectively discriminates between TB cases and healthy controls among HIV uninfected persons. Based on published microarray case-control datasets, the COR has 87% diagnostic sensitivity and 97% specificity for prevalent TB disease in HIV uninfected South African adults (Zak, *Lancet*, 2016); and in two nested case-control studies, also among HIV uninfected persons, the COR has 70% predictive sensitivity and 84% specificity for incident TB disease occurring within one year of sampling (Penn-Nicholson, *S Afr Med J*, 2016). This PCR-based mRNA COR signature has been refined to 11-genes (RISK11) with equivalent diagnostic performance (Darboe, *Tuberculosis*, 2018).

Although relatively parsimonious, this signature is not ideal for a point-of-care (POC) device because of its size. Several concise mRNA signatures have recently been developed which are translatable into a POC test device (Gupta, *bioRxiv*, 2019; Warsinske, *PLoS Med*, 2019). A POC device could be employed for test-and-treat strategies in the community and clinic, both to detect early (and asymptomatic) TB disease for curative treatment and to identify *M.tb* infected individuals at high risk of progression to active disease for targeted short-course preventive therapy.

The aim of this sub-study is to evaluate and compare the performance of parsimonious host-blood TB transcriptomic signatures to identify prevalent TB disease and predict incident TB disease in HIV-infected and HIV-uninfected adults.

2.1 AIMS

*2.1.1 **PRIMARY AIMS***

2.1.1.1 **Primary Aim 1:** Estimate whether parsimonious host-blood TB transcriptomic signatures differentiate HIV-infected (CORTIS-HR) and HIV-uninfected (CORTIS-01) persons with prevalent TB disease from those without prevalent TB disease.

2.1.1.2 **Primary Aim 2:** Estimate whether parsimonious host-blood TB transcriptomic signatures differentiate HIV-infected (CORTIS-HR) and HIV-uninfected (CORTIS-01) persons at high risk for incident TB disease from those at low risk for incident TB disease.

*2.1.2 **SECONDARY AIMS***

2.1.2.1 **Secondary Aim 1:** Compare parsimonious host-blood TB transcriptomic signatures diagnostic performance in HIV-infected (CORTIS-HR) and HIV-uninfected (CORTIS-01) persons.

2.1.2.2 **Secondary Aim 2:** Compare parsimonious host-blood TB transcriptomic signatures predictive performance in HIV-infected (CORTIS-HR) and HIV-uninfected (CORTIS-01) persons.

2.1.2.3 **Secondary Aim 3:** Compare parsimonious host-blood TB transcriptomic signatures diagnostic performance, stratified by symptoms, in HIV-infected (CORTIS-HR) and HIV-uninfected (CORTIS-01) persons.

2.2 STUDY ENDPOINTS

2.2.1 TWO-SAMPLE ENDPOINT DEFINITION

TB disease confirmed by positive Xpert MTB/RIF and/or MGIT culture on two or more separate sputum samples, or samples from another site if extrapulmonary disease. All aims will be evaluated using this endpoint definition.

2.2.1 ONE AND TWO-SAMPLE ENDPOINT DEFINITION

TB disease confirmed by positive Xpert MTB/RIF and/or MGIT culture on at least one sputum sample, or sample from another site if extrapulmonary disease. As exploratory analyses, all aims may also evaluated using this endpoint definition.

2.3 TB DISEASE ENDPOINT ADJUDICATION ALGORITHM AND CENSORING

See CORTIS-01 and CORTIS-HR protocols and SAPs.

2.4 STUDY DESIGN

See CORTIS-01 and CORTIS-HR protocols and SAPs.

2.5 MEASUREMENT OF PARSIMONIOUS HOST-BLOOD TB TRANSCRIPTOMIC SIGNATURE SCORES

As part of the CORTIS-01 and CORTIS-HR studies, whole blood RNA was collected in PAXgene tubes at screening and shipped frozen to the SATVI Cape Town laboratory where RNA was extracted in a high-throughput, standardized, and reproducible fully automated procedure using a TECAN EVO Freedom robotic platform. Following cDNA-synthesis and pre-amplification steps, the parsimonious host-blood TB transcriptomic signatures listed below were run using the BioMark HD Fluidigm multiplex qRT-PCR machine and analysed using locked-down R quality control and analysis scripts:

Signature	**Model***	**Genes**	**TaqMan Assay**
Herberg2 / VIRAL2 (*JAMA*, 2016)	FAM89A – IFI44L	FAM89A IFI44L	Custom_ARZTE3U Hs00915292_m1
Maertzdorf4 / DIAG4 (*EMBO Mol Med*, 2016)	See Suliman (*Am J Respir Crit Care Med*, 2018)	GBP1 IFITM3 P2RY14 ID3	Hs00977005_m1 Hs03057129_s1 Hs01848195_s1 Hs00954037_g1
Penn-Nicholson6 / RISK6 (*medRxiv*, 2019)	See Penn-Nicholson (*medRxiv*, 2019)	GBP2 FCGR1B SERPING1 TUBGCP6 TRMT2A SDR39U1	Hs00894846_g1 Hs02341825_m1 Hs00934329_m1 Hs00363509_g1 Hs01000041_g1 Hs01016970_g1
Roe1 / DIAG1 (*JCI Insight*, 2016)	BATF2 – [(TMBIM6 + CDC42 + USF2 + ACTR3)/4]	BATF2 Reference probes: TMBIM6 CDC42 USF2 ACTR3	Hs00912736_m1 Hs00162661_m1 Hs03044122_g1 Hs01100994_g1 Hs01029159_g1
Roe3 / RISK3 (*Clin Infect Dis*, 2019)	[(SCARF1 + GBP5 + BATF2) / 3] – [(TMBIM6 + CDC42 + USF2 + ACTR3)/4]	SCARF1 GBP5 BATF2 Reference probes:	Hs01092483_m1 Hs00369472_m1 Hs00912736_m1 As for Roe1
Suliman4 / RISK4 (*Am J Respir Crit Care Med*, 2018)	See Suliman (*Am J Respir Crit Care Med*, 2018)	GAS6 SEPT4 CD1C BLK	Hs01090305_m1 Hs00910208_g1 Hs00957534_g1 Hs01017452_m1

Sweeney3 / DIAG3 (*Lancet Respir Med*, 2016)	[(GBP5 + DUSP3)/2] − KLF2 See Warsinske (*JAMA Netw Open*, 2018)	GBP5 DUSP3 KLF2	Hs00369472_m1 Hs01115776_m1 Hs00360439_g1
Thompson5 / RESPONSE5 (*Tuberculosis (Edinb)*, 2017)	See Suliman (*Am J Respir Crit Care Med*, 2018)	UCP2 MAP7D3 STT3A SMARCD3 RP11-295G20.2	Hs01075224_g1 Hs00226257_m1 Hs00967491_m1 Hs01088251_g1 Hs01373568_m1

* Raw Ct values from Fluidigm Biomark qRT-PCR will be used.

The COR (RISK11) signature was run independently during the CORTIS-01 and HR studies on samples from all participants (including non-enrolled CORTIS-01 participants). The Roe1 and Roe3 signatures scores are calculated from the same Fluidigm 96.96 integrated fluidic chip (IFC) qRT-PCR runs as the RISK11 signature. The other six signatures (Herberg2, Maertzdorf4, Penn- Nicholson6, Suliman4, Sweeney3, and Thompson5) were run side-by-side on Fluidigm 192.24 IFCs, from cDNA synthesised from stored RNA aliquots of RNA extracted from all enrolled participants of CORTIS-01 and HR. For this sub-study, the performance of these eight parsimonious host-blood TB transcriptomic signatures will be compared.

2.6 PROJECTED ENROLMENT, CASE ACCRUAL AND POWER

As per CORTIS-01 and CORTIS-HR protocols and SAPs.

2.7 ANALYSIS POPULATIONS AND WEIGHTING

The intention to treat (ITT) and modified intention to treat population (mITT) are as defined in the CORTIS-01 and CORTIS-HR SAPs. The ITT cohort will be used for addressing primary aim 1 and secondary aim 1 (diagnostic performance in differentiating prevalent cases from combined TB negative controls and incident TB cases). The mITT cohort, which excludes prevalent TB cases, will be used for assessing signature predictive performance for risk of progression to TB disease (primary aim 2 and secondary aim 2).

2.7.1 ***CORTIS-01***

Group B (COR+ observation arm) and Group C (COR- observation arm) will be included in analysis of both diagnostic and predictive performance aims, Group A (treatment efficacy arm) will be excluded from the predictive analysis, but will be included in estimating diagnostic performance at baseline (visit 2). Since the study is artificially enriched for COR+ (RISK11+) participants by design, participant-specific weights in the analysis are required to recover estimates that are applicable to the screened population, as opposed to the enrolled cohort. The COR- participants (Group C) will be up-weighted according to the empirical inverse probability of enrolling COR- participants that were

screened (i.e. inverse probability weighting, IPW). Weighting will be applied as per the CORTIS-01 SAP. The 95% confidence interval will be provided for each performance metric using a non-parametric bootstrap, with sampling stratified by COR status, since the number of COR+ and COR- enrolled is fixed by the study design.

2.7.1 ***CORTIS-HR***

All enrolled participants will be included in analysis. There was no treatment efficacy arm in CORTIS-HR. No weighting is required, as the cohort is not enriched for COR+ participants.

3 STATISTICAL CONSIDERATIONS

3.1 GENERAL PRINCIPLES

See CORTIS-01 and CORTIS-HR SAPs.

3.2 MISSING DATA

See CORTIS-01 and CORTIS-HR SAPs.

3.3 RESPONSIBILITY

Management of the clinical and laboratory data as outlined in the protocol will be the responsibility of designated laboratory technologists and doctoral scientist under supervision of Deputy Director of Immunology, SATVI and the study PI and Director SATVI.

3.4 BLINDING

Participants, investigators, the data analysis team, and all members of the clinical trial team responsible for performing TB symptom and sputum screening for the purpose of endpoint determination, as well as the medical monitor, Sponsor, and data management personnel, will remain blind to participant parsimonious host-blood TB transcriptomic signature results until this SAP has been approved and signed by study sponsor. However, access to clinical and demographic data (include TB disease status) will be available to allow data cleaning and preparation of analysis script prior to database lock.

3.5 DATA STORAGE

All files containing the final data will be password-protected and backed-up to the SATVI server. Data will be saved as a .csv file and compiled into one master database using R. All analysis scripts and outputs will also be backed up onto the SATVI server upon completion of the study report.

4 STATISTICAL METHODS

4.1 SIGNATURE PERFORMANCE ANALYSIS

The performance of the signatures measured at screening will be evaluated on their ability to diagnose prevalent TB disease at baseline/enrolment and predict incident TB disease over 15-month follow up. In the CORTIS-01 and CORTIS-HR study, a pre-specified COR (RISK11) score threshold of 60% was used to differentiate correlate positive (COR+) from correlate negative (COR-) individuals. There are no pre-defined thresholds for the parsimonious host-blood TB transcriptomic signatures, thus we plan to perform receiver operating characteristic curve (ROC) analysis to select the best diagnostic and predictive signature score thresholds. Indeterminate signature score results will be excluded from primary analysis. The proportion of samples with successful signature results (i.e. those without PCR failures or indeterminate scores) will be reported. For all analyses, HIV-infected (CORTIS-HR) and HIV-uninfected (CORTIS-01) participants will be considered independently.

Score distributions for each signature in prevalent and incident TB cases, and incident-free controls will be visually represented using violin/box/scatter plots (**Figure 6.1**). Signature score distribution will be described using median and interquartile range. For CORTIS-01, we will present both weighted (section 2.7.1, above) and unweighted plots. For CORTIS-HR, plots will not be weighted (section 2.7.2, above).

Correlations between all signature scores (Spearman *rho*), as well as demographic and clinical data will be reported in a correlation matrix (**Figure 6.2**). CORTIS-01 correlations will be weighted as described in section 2.7.1 (above). For CORTIS-HR, correlations will not be weighted.

4.2 PRIMARY AIM 1: SIGNATURE DIAGNOSTIC PERFORMANCE

Binary diagnostic ROC analysis will be performed separately for each signature using both CORTIS-01 and CORTIS-HR ITT cohorts (independently). These analyses will include evaluation of the area under the curve (AUC), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and number needed to screen (NNS) to detect one case of TB at enrolment in HIV-infected and HIV-uninfected adults. Signature diagnostic accuracy versus the WHO target product profile (TPP) for a community-based triage or referral test to identify people suspected of having TB (World Health Organization, 2014) will be considered: For each signature, test accuracy with specificity benchmarked at 70% or sensitivity at 90% (minimum WHO triage test TPP criteria), and specificity at 80% or sensitivity 95% (optimum WHO triage test TPP criteria) will be reported.

Results will be reported in **Table 5.1** and graphically with ROC curves (**Figure 6.3**) and sensitivity/specificity versus signature score threshold plots (**Figure 6.4**).

4.3 PRIMARY AIM 2: SIGNATURE PREDICTIVE PERFORMANCE

We will evaluate the changes in biomarker performance over time, including AUC, sensitivity, specificity, PPV and NPV (**Table 5.2**). We will specifically consider the following time intervals to incident TB diagnosis: 0 to 3 months, 0 to 6 months, 0 to 9 months, 0 to 12 months, and 0 to 15 months windows. We will also look at 6 month sliding windows to provide better insight into performance at different times (e.g. 0 to 6 months, 1 to 7 months, 2 to 8 months… 9 to 15 month windows).

Plots of each of these measures as a function of follow-up time will aid interpretation of the primary results at month 15 (**Figure 6.5**). Signature accuracy in predicting progression from tuberculosis infection to active disease will be compared with the WHO TPP (World Health Organization, 2017): For each signature, test accuracy with specificity benchmarked at 75% or sensitivity at 75% (minimum WHO predictive test TPP criteria), and specificity at 90% or sensitivity 90% (optimum WHO predictive test TPP criteria) will be reported.

These analyses will make use of methods and R packages developed by Zheng (*J Am Stat Assoc*, 2008), Heagerty (*Biometrics*, 2000), Pepe (*Stata J*, 2009) and others (Bansal, *Diagn Progn Res*, 2019).

4.4 CONFIDENCE INTERVALS

The 95% confidence interval (95%CI) will be provided for each performance metric using a non-parametric bootstrap with 10,000 iterations. For CORTIS-01 (HIV-uninfected), bootstrap sampling will be stratified by COR status, since the number of COR+ and COR- enrolled is fixed by the study design.

4.5 SECONDARY AIMS: COMPARISON OF SIGNATURE PERFORMANCE

4.5.1.1 **HIV-infected versus HIV-uninfected population signature performance:** Difference in signature performance (AUC) between HIV-infected (CORTIS-HR) and HIV-uninfected (CORTIS-01). The difference in a signature's AUC will be considered significant if the 95%CIs do not overlap. Score distributions will be visually represented with violin/box-and-scatter plots (**Figure 6.1**).

4.5.1.2 **Symptomatic versus asymptomatic population signature performance (sub-group analysis):** Difference in signature diagnostic performance (AUC) between symptomatic and asymptomatic persons will also be explored in both HIV-infected (CORTIS-HR) and HIV-uninfected (CORTIS-01) populations. The difference in a signature's AUC will be considered significant if the 95%CIs do not overlap. Score distributions will be visually represented with violin/box-and-scatter plots (**Figure 6.1**).

4.6 EXPLORATORY SUB-GROUP ANALYSES

Exploratory diagnostic and predictive performance analyses may be performed by the below sub-groups. These analyses may have limited power due to small sample size and limited numbers of active TB cases. Antiretroviral (ART) status and isoniazid preventive therapy (IPT) status is determined by concomitant medication recorded during follow up and is coded as per the CORTIS-HR SAP. The difference in a performance between groups will be considered significant if the 95%CIs do not overlap. Score distributions will be visually represented with violin/box-and-scatter plots (**Figure 6.1**).

4.6.1 ***CORTIS-01 AND CORTIS-HR***

4.6.1.1 Age at enrolment

4.6.1.2 Sex (male/female)

4.6.1.3 Ethnicity (Black African, Cape Mixed Ancestry)

4.6.1.4 Study site

4.6.1.5 Smoking history at enrolment (yes/no)

4.6.1.6 BMI at enrolment

4.6.1.7 Prior TB episode (yes/no)

4.6.1.8 QFT status at enrolment (<0.35, negative; ≥0.35, positive)

4.6.2 ***CORTIS-HR ONLY***

4.6.2.1 ART-naïve versus ART experienced at enrolment

4.6.2.2 CD4 cell count at enrolment

4.6.2.3 Viral load at enrolment (lower than the detectable limit (<100), versus ≥ 100)

4.6.2.4 Enrolment IPT status

4.6.2.5 IPT duration during study (No IPT, IPT <6 months, IPT > 6 months)

5 TABLES

5.1 PRIMARY AIM 1: SIGNATURE DIAGNOSTIC PERFORMANCE AT ENROLMENT (BINARY ANALYSIS) IN HIV-INFECTED (CORTIS-HR) / HIV-UNINFECTED (CORTIS-01) COHORT

Signature	Optimal signature threshold‡	PCR failure or indeterminate result†	Sample pass rate (%)	TP	FN	TN	FP	AUC, % (95% CI)	Sensitivity‡ % (95% CI)	Specificity‡ % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)	NNS, n (95% CI)
RISK11 (COR)	x	x	x	x	x	x	x	x% (x–x)	x% (x–x)	70%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		x% (x–x)	80%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		95%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
Herberg2	x	x	x	x	x	x	x	x% (x–x)	x% (x–x)	70%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		x% (x–x)	80%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		95%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
Maertzdorf4	x	x	x	x	x	x	x	x% (x–x)	x% (x–x)	70%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		x% (x–x)	80%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		95%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
Penn-Nicholson6	x	x	x	x	x	x	x	x% (x–x)	x% (x–x)	70%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		x% (x–x)	80%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		95%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
Roe1	x	x	x	x	x	x	x	x% (x–x)	x% (x–x)	70%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		x% (x–x)	80%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		95%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
Roe3	x	x	x	x	x	x	x	x% (x–x)	x% (x–x)	70%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		x% (x–x)	80%	x% (x–x)	x% (x–x)	x (x–x)

	x	x	x	x	x	x	x		95%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
Suliman4	x	x	x	x	x	x	x	x% (x–x)	x% (x–x)	70%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		x% (x–x)	80%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		95%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
Sweeney3	x	x	x	x	x	x	x	x% (x–x)	x% (x–x)	70%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		x% (x–x)	80%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		95%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
Thompson5	x	x	x	x	x	x	x	x% (x–x)	x% (x–x)	70%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		x% (x–x)	80%	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		95%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)
	x	x	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)	x (x–x)

TP = True positive; FN = False negative; TN = True negative; FP = False positive; AUC = Area under the receiver operating characteristics (ROC) curve; PPV = Positive predictive value; NPV = Negative predictive value; NNS = Number needed to screen to detect 1 case of incident TB; COR = Correlate of risk.

Note: Participants who were unable to produce satisfactory sputum samples at enrolment or at end-of-study visit were assumed to be sputum negative at those time-point.

† Indeterminate results excluded. ‡ Threshold with specificity benchmarked at 70% or sensitivity at 90% (minimum WHO triage test TPP), and specificity at 80% or sensitivity 95% (optimum WHO triage test TPP).

5.2 PRIMARY AIM 2: SIGNATURE PREDICTIVE PERFORMANCE FOR IDENTIFICATION OF TB DISEASE OVER A 15-MONTH PERIOD, STRATIFIED BY THE TIME INTERVAL TO DISEASE (TIME-DEPENDENT ANALYSIS) IN HIV-INFECTED (CORTIS-HR) / HIV-UNINFECTED (CORTIS-01) COHORT

Time intervals: 0 to 3 months, 0 to 6 months, 0 to 9 months, 0 to 12 months, and 0 to 15 months

1 to 6 month, 2 to 7 month, 3 to 8 month… 10 to 15 month sliding windows

Test[†] and time interval to TB disease	Optimal signature threshold[‡]	PCR failure or indeterminate result[†]	Sample pass rate (%)	Incident TB Cases	Non-incident controls	AUC, % (95% CI)	Sensitivity[‡] % (95% CI)	Specificity[‡] % (95% CI)	PPV[‡], % (95% CI)	NPV[‡], % (95% CI)
Window: X to Y months										
RISK11 (COR)	x	x	x	x	x	x% (x–x)	x% (x–x)	75%	x% (x–x)	x% (x–x)
	x	x	x	x	x		75%	x% (x–x)	x% (x–x)	x% (x–x)
	x	x	x	x	x		x% (x–x)	90%	x% (x–x)	x% (x–x)
	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)
Herberg2	x	x	x	x	x	x% (x–x)	x% (x–x)	75%	x% (x–x)	x% (x–x)
	x	x	x	x	x		75%	x% (x–x)	x% (x–x)	x% (x–x)
	x	x	x	x	x		x% (x–x)	90%	x% (x–x)	x% (x–x)
	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)
Maertzdorf4	x	x	x	x	x	x% (x–x)	x% (x–x)	75%	x% (x–x)	x% (x–x)
	x	x	x	x	x		75%	x% (x–x)	x% (x–x)	x% (x–x)
	x	x	x	x	x		x% (x–x)	90%	x% (x–x)	x% (x–x)
	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)
Penn-Nicholson6	x	x	x	x	x	x% (x–x)	x% (x–x)	75%	x% (x–x)	x% (x–x)
	x	x	x	x	x		75%	x% (x–x)	x% (x–x)	x% (x–x)
	x	x	x	x	x		x% (x–x)	90%	x% (x–x)	x% (x–x)
	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)
Roe1	x	x	x	x	x	x% (x–x)	x% (x–x)	75%	x% (x–x)	x% (x–x)
	x	x	x	x	x		75%	x% (x–x)	x% (x–x)	x% (x–x)
	x	x	x	x	x		x% (x–x)	90%	x% (x–x)	x% (x–x)

	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)
Roe3	x	x	x	x	x	x% (x–x)	x% (x–x)	75%	x% (x–x)	x% (x–x)
	x	x	x	x	x		75%	x% (x–x)	x% (x–x)	x% (x–x)
	x	x	x	x	x		x% (x–x)	90%	x% (x–x)	x% (x–x)
	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)
Suliman4	x	x	x	x	x	x% (x–x)	x% (x–x)	75%	x% (x–x)	x% (x–x)
	x	x	x	x	x		75%	x% (x–x)	x% (x–x)	x% (x–x)
	x	x	x	x	x		x% (x–x)	90%	x% (x–x)	x% (x–x)
	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)
Sweeney3	x	x	x	x	x	x% (x–x)	x% (x–x)	75%	x% (x–x)	x% (x–x)
	x	x	x	x	x		75%	x% (x–x)	x% (x–x)	x% (x–x)
	x	x	x	x	x		x% (x–x)	90%	x% (x–x)	x% (x–x)
	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)
Thompson5	x	x	x	x	x	x% (x–x)	x% (x–x)	75%	x% (x–x)	x% (x–x)
	x	x	x	x	x		75%	x% (x–x)	x% (x–x)	x% (x–x)
	x	x	x	x	x		x% (x–x)	90%	x% (x–x)	x% (x–x)
	x	x	x	x	x		90%	x% (x–x)	x% (x–x)	x% (x–x)

AUC = Area under the receiver operating characteristics (ROC) curve; PPV = Positive predictive value; NPV = Negative predictive value

† Indeterminate results excluded. ‡ Threshold with specificity benchmarked at 75% or sensitivity at 75% (minimum WHO TPP for test predicting progression to incident TB), and specificity at 90% or sensitivity 90% (optimum WHO TPP for test predicting progression to incident TB).

6 FIGURES

6.1 Test score distribution: violin/box-and-scatter plots. Plots will be weighted for CORTIS-01 (section 2.7.1) and unweighted for CORTIS-HR (section 2.7.2).

6.2 Test score correlations: correlation matrix of signature scores versus signature scores and demographic variables (spearman *rho* correlation coefficient). Plots will be weighted for CORTIS-01 (section 2.7.1) and unweighted for CORTIS-HR (section 2.7.2).

6.3 Test performance: ROC curves (sensitivity versus 100-specificity) with relevant WHO TPP criteria indicated

6.4 Test accuracy: sensitivity/specificity versus test score threshold plots (x-axis: score threshold, y-axis: sensitivity or specificity)

6.5 Time dependent analysis: AUC versus time. Sensitivity, specificity, NPV and PPV (at different thresholds) versus time, over 15-month follow-up

7 APPENDIX

7.1 WHO DIAGNOSTIC, TRIAGE, AND PREDICTIVE TB TEST PERFORMANCE TARGET PRODUCT PROFILE

1. Rapid biomarker-based non-sputum-based test for detecting (diagnosing) PTB (World Health Organization, 2014)

	Minimum diagnostic	Optimal diagnostic
Sensitivity in adult PTB (overall pooled sensitivity in culture-positive cases)	≥65% (among both smear-positive & -negative cases)	≥68% (among smear-negative cases only)
Sensitivity in adult PTB (among smear-positive culture-positive cases only)	>98%	≥98%
Specificity	≥98%	Not specified

2. Community-based triage or referral test to identify people suspected of having TB (World Health Organization, 2014)

	Minimum screening	Optimal screening
Sensitivity in adult PTB (compared with confirmatory testing)	>90%	>95%
Specificity in adult PTB (compared with confirmatory testing)	>70%	>80%

3. Test predicting progression from tuberculosis infection to active disease (WHO, 2017) (World Health Organization, 2017)

	Minimum predictive	Optimal predictive
Predictive sensitivity	≥75%	≥90%
Predictive specificity	≥75%	≥90%

8 REFERENCES

Auguste P, Tsertsvadze A, Pink J, Court R, McCarthy N, Sutcliffe P, and Clarke A. 2017. Comparing interferon-gamma release assays with tuberculin skin test for identifying latent tuberculosis infection that progresses to active tuberculosis: systematic review and meta-analysis. *BMC Infect Dis.* 17 (1):200. doi: 10.1186/s12879-017-2301-4.

Bansal A, and Heagerty PJ. 2019. A comparison of landmark methods and time-dependent ROC methods to evaluate the time-varying performance of prognostic markers for survival outcomes. *Diagn Progn Res.* 3:14. doi: 10.1186/s41512-019-0057-6.

Darboe F, Mbandi SK, Thompson EG, Fisher M, Rodo M, van Rooyen M, Filander E, Bilek N, Mabwe S, Hatherill M*, et al.* 2018. Diagnostic performance of an optimized transcriptomic signature of risk of tuberculosis in cryopreserved peripheral blood mononuclear cells. *Tuberculosis.* 108:124-126. doi: 10.1016/j.tube.2017.11.001.

Gupta RK, Turner CT, Venturini C, Esmail H, Rangaka MX, Copas A, Lipman M, Abubakar I, and Noursadeghi M. 2019. Concise whole blood transcriptional signatures for incipient tuberculosis: A systematic review and patientlevel pooled meta-analysis. *bioRxiv.* doi: 10.1101/668137.

Heagerty PJ, Lumley T, and Pepe MS. 2000. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics.* 56 (2):337-44.

Herberg JA, Kaforou M, Wright VJ, Shailes H, Eleftherohorinou H, Hoggart CJ, Cebey-Lopez M, Carter MJ, Janes VA, Gormley S*, et al.* 2016. Diagnostic Test Accuracy of a 2-Transcript Host RNA Signature for Discriminating Bacterial vs Viral Infection in Febrile Children. *JAMA.* 316 (8):835-45. doi: 10.1001/jama.2016.11236.

Maertzdorf J, McEwen G, Weiner J, 3rd, Tian S, Lader E, Schriek U, Mayanja-Kizza H, Ota M, Kenneth J, and Kaufmann SH. 2016. Concise gene signature for point-of-care classification of tuberculosis. *EMBO Mol Med.* 8 (2):86-95. doi: 10.15252/emmm.201505790.

Penn-Nicholson A, Scriba TJ, Hatherill M, White RG, and Sumner T. 2016. A novel blood test for tuberculosis prevention and treatment. *S Afr Med J.* 107 (1):4-5. doi: 10.7196/SAMJ.2016.v107.i1.12230.

Penn-Nicholson A, Mbandi SK, Thompson E, Mendelsohn S, Suliman S, Chegou NN, Malherbe ST, Darboe F, Erasmus M, Hanekom WA*, et al.* 2019. RISK6, a universal 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response. *medRxiv.* doi: 10.1101/19006197.

Pepe M, Longton G, and Janes H. 2009. Estimation and Comparison of Receiver Operating Characteristic Curves. *Stata J.* 9 (1):1.

Roe J, Venturini C, Gupta RK, Gurry C, Chain BM, Sun Y, Southern J, Jackson C, Lipman MC, Miller RF*, et al.* 2019. Blood transcriptomic stratification of short-term risk in contacts of tuberculosis. *Clin Infect Dis.* ciz252. doi: 10.1093/cid/ciz252.

Roe JK, Thomas N, Gil E, Best K, Tsaliki E, Morris-Jones S, Stafford S, Simpson N, Witt KD, Chain B*, et al.* 2016. Blood transcriptomic diagnosis of pulmonary and extrapulmonary tuberculosis. *JCI Insight.* 1 (16):e87238. doi: 10.1172/jci.insight.87238.

Suliman S, Thompson E, Sutherland J, Weiner Rd J, Ota MOC, Shankar S, Penn-Nicholson A, Thiel B, Erasmus M, Maertzdorf J*, et al.* 2018. Four-gene Pan-African Blood Signature Predicts Progression to Tuberculosis. *Am J Respir Crit Care Med.* 197 (9):1198-1208. doi: 10.1164/rccm.201711-2340OC.

Sweeney TE, Braviak L, Tato CM, and Khatri P. 2016. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. *Lancet Respir Med.* 4 (3):213-24. doi: 10.1016/S2213-2600(16)00048-5.

Thompson EG, Du Y, Malherbe ST, Shankar S, Braun J, Valvo J, Ronacher K, Tromp G, Tabb DL, Alland D*, et al.* 2017. Host blood RNA signatures predict the outcome of tuberculosis treatment. *Tuberculosis (Edinb).* 107:48-58. doi: 10.1016/j.tube.2017.08.004.

Warsinske H, Vashisht R, and Khatri P. 2019. Host-response-based gene signatures for tuberculosis diagnosis: A systematic comparison of 16 signatures. *PLoS Med.* 16 (4):e1002786. doi: 10.1371/journal.pmed.1002786.

Warsinske HC, Rao AM, Moreira FMF, Santos PCP, Liu AB, Scott M, Malherbe ST, Ronacher K, Walzl G, Winter J*, et al.* 2018. Assessment of Validity of a Blood-Based 3-Gene Signature Score for Progression and Diagnosis of Tuberculosis, Disease Severity, and Treatment Response. *JAMA Netw Open.* 1 (6):e183779. doi: 10.1001/jamanetworkopen.2018.3779.

World Health Organization. 2014. High-priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting. Geneva: WHO. www.who.int/tb/publications/tpp_report/en/ (Accessed 12 August 2017).

World Health Organization. 2015. The End TB Strategy: Global strategy and targets for tuberculosis prevention, care and control after 2015. Geneva: WHO. https://www.who.int/tb/strategy/End_TB_Strategy.pdf?ua=1 (Accessed September 6, 2019).

World Health Organization. 2017. Consensus Meeting Report: Development of a Target Product Profile (TPP) and a framework for evaluation for a test for predicting progression from tuberculosis infection to active disease. Geneva: WHO. http://apps.who.int/iris/handle/10665/259176 (Accessed 1 October 2017).

Zak DE, Penn-Nicholson A, Scriba TJ, Thompson E, Suliman S, Amon LM, Mahomed H, Erasmus M, Whatney W, Hussey GD*, et al.* 2016. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. *Lancet.* 387 (10035):2312-2322. doi: 10.1016/S0140-6736(15)01316-1.

Zheng Y, Cai T, Pepe MS, and Levy WC. 2008. Time-dependent Predictive Values of Prognostic Biomarkers with Failure Time Outcome. *J Am Stat Assoc.* 103 (481):362-368. doi: 10.1198/016214507000001481.